OPERATIONAL RESEARCH SOCIETY OF NEW ZEALAND

40th ANNUAL CONFERENCE

2-3 December 2005

Victoria University of Wellington
Wellington
New Zealand

Sponsored by
Concept Consulting
Hoare Research
The Electricity Commission
Victoria Management School
Victoria University of Wellington

ISSN 0114-6602

The Operational Research Society of New Zealand
Acknowledges the support of the following sponsors:

Victoria Management School
Te Kura Whakahaere

School of Mathematics, Statistics and Computer Science
Te Kura Tatau

Operational Research Society of New Zealand (Inc.)
P.O.Box 6544, Wellesley Street, Auckland, New Zealand
www.orsnz.org.nz

Kia Ora and Haere Mai to…

ORSNZ's 40th Annual Conference
Reflecting Back – Looking Forward
Celebrating 40 Years of OR in New Zealand

Welcome to the 40th Annual Conference of the Operational Research Society of New Zealand, hosted by the Victoria Management School, at Victoria University of Wellington's recently developed Pipitea Campus. The Victoria Management School is involved in a number of projects, courses, degrees and programmes with an Operational Research and Operations Management flavour, as can be found from our website www.vms.ac.nz

This conference is ORSNZ's 40th Annual Conference and is held in Wellington, as was the first in 1964, and many conferences in intervening years, including the 21st in 1985. So this 40th provides an opportunity to reflect on where we have been, as a Society and as a discipline of OR, and where we are heading.

Wellington was the early home of OR in NZ, with sizable OR groups in major government departments such as DSIR, Electricity, Energy and Transport being located here. Such groups provided many of the key office holders of the Society and Wellington provided the bulk of Council members and was the geographic base of Council until the early 90s. It is interesting to note that the majority of members back then were practitioners, whereas now the bulk of members are in academia. The discipline of OR has become much more diverse, as witnessed by the range of papers on offer at this conference, summarised in these Conference Proceedings.

Organising this conference has been a wonderful team drawn from practitioners and academics – at least we have managed to keep the practitioner spirit alive, even if we have been less successful in persuading those practitioners to speak. The team has worked hard to ensure a well-organised event, both in terms of the papers and in terms of some special added features, for our 40th.

We extend our thanks to all those who have made this conference possible, including our sponsors, whose generosity has covered many of the expenses, our speakers, and our attendees. Officeholders of the ORSNZ Council have provided support and we especially thank John Scott and the panel of judges for taking care of the Young Practitioners Prize awards this year.

We hope you enjoy your time in Wellington, both at the Conference and in taking advantage of all our fair Capital City has to offer.

Vicky Mabin
(Chair, Conference Committee)

Conference Committee:	Vicky Mabin
	Stefanka Chukova
	John Davies
	John Hayes
	Mark Johnston
	Mark Pickup
Proceedings Editor:	Tom Halliburton
Administrative Support:	Sophia Lum

Conference Organising Committee:

Mark Johnston, Tom Halliburton, Vicky Mabin, Mark Pickup, John Davies, Stefanka Chukova, John Hayes

Table of Contents

Author Session Index

Author	Session	Time
Arnold, Richard	5B	Saturday 10:50-11:30
Arthanari, Tiru	4A	Saturday 10:00-10:20
Basnet, Chuda	5A	Saturday 11:10-11:30
Belton, Val	6A	Saturday 1:30-2:00
Beltrán, Fernando	6B	Saturday 2:10-2:30
Bleach-Wood, Simon	3A	Friday 3:50-4:10
Bourke, Natalie	2B	Friday 1:40-2:00
Bramley, Richard	2B	Friday 2:40-3:00
Brkovic, Vladimir	1B	Friday 12:00-12:20
Bruce, Gregor	4A	Saturday 9:20-9:40
Cavana, Bob	1A	Friday 11:00-11:40
Chattopadhyay, Deb	2A	Friday 2:40-3:00
Chen, Mei-hua	3A	Friday 4:10-4:30
Chukova, Stefanka	5B	Saturday 12:10-12:30
Craigie, David	3B	Friday 4:30-4:50
Daniell, Emma	2B	Friday 1:40-2:00
Davies, John	4B	Saturday 9:40-10:00
	6A	Saturday 2:20-2:40
Donovan, Stuart	3B	Friday 4:50-5:10
Downward, Anthony	2B	Friday 2:00-2:20
Dye, Shane	4A	Saturday 9:00-9:20
Ehrgott, Matthias	5A	Saturday 12:10-12:30
Everett, Graeme	2A	Friday 2:00-2:20
Ford, Andrew	1A	Friday 11:00-11:40
Foulds, L.R.	5A	Saturday 11:10-11:30
Goddard, Kathy	3A	Friday 3:50-4:10
Gordon, M	4A	Saturday 9:40-10:00
Gould, Helyn	6A	Saturday 1:30-2:00
Hannah, Luke	1B	Friday 11:00-11:20
Harding, Dean	3A	Friday 3:50-4:10
Haywood, John	5B	Saturday 11:50-12:10
Hinze, Tonje	2B	Friday 1:40-2:00
James, Ross	5A	Saturday 10:50-11:10
Jarden, Kathryn	2B	Friday 1:20-1:40
Johnston, Mark	5B	Saturday 12:10-12:30
Joshan, Amir	3B	Friday 3:30-3:50
Kildare, Carolyn	3A	Friday 3:50-4:10
Kim, Kunhong	6B	Saturday 1:30-2:10
Koh, Joshua	3B	Friday 3:50-4:10
Krawczyk, Jacek	6B	Saturday 1:30-2:10
Lermit, Jonathan	1A	Friday 11:40-12:00
	2A	Friday 1:20-2:00
Liu, Ivy	5B	Saturday 11:30-11:50
Lusby, Richard	1B	Friday 11:40-12:00
Mabin, Vicky	4B	Saturday 9:40-10:00
	6A	Saturday 2:20-2:40
Mason, Andrew	5A	Saturday 11:50-12:10
McNaughton, Alastair	3A	Friday 4:10-4:30
McNickle, Don	2A	Friday 2:40-3:00
Messenger, Steve	3A	Friday 3:50-4:10
Mitchell, S	5A	Saturday 11:30-11:50

Mukherjee, Bhramar	5B	Saturday 11:30-11:50
O'Sullivan, M	4A	Saturday 9:40-10:00
Paynter, John	3A	Friday 4:30-4:50
	4B	Saturday 10:00-10:20
Peko, Gabrielle	3A	Friday 4:30-4:50
Pickup, Mark	1A	Friday 12:00-12:20
Poot, Tineke	3B	Friday 4:10-4:30
Randal, John	5B	Saturday 11:50-12:10
Read, Grant	2A	Friday 2:20-2:40
	2A	Friday 2:40-3:00
Richards, David	1B	Friday 11:20-11:40
Rodrigo, Allen	4B	Saturday 10:00-10:20
Ryan, David	5A	Saturday 12:10-12:30
Santamaria, Natalia	6B	Saturday 2:10-2:30
Schöbel, Anita	5A	Saturday 12:10-12:30
Scott, John	6A	Saturday 1:30-2:00
Shao, Lizhen	4B	Saturday 9:20-9:40
Steer, Tim	2B	Friday 2:40-3:00
Thomas, Mike	2A	Friday 2:20-2:40
Tipping, James	2A	Friday 2:40-3:00
Tompkins, Oliver	2B	Friday 2:20-2:40
Too, Megan	4B	Saturday 10:00-10:20
Vélasquez, Rafael	5A	Saturday 12:10-12:30
Vignaux, Tony	6A	Saturday 2:40-3:10
Walker, Cameron	4A	Saturday 9:40-10:00
Walker, Colin	6A	Saturday 2:00-2:20
Wang, Dong	3A	Friday 4:50-5:10
Waterer, Hamish	5A	Saturday 11:30-11:50
Weide, Oliver	4B	Saturday 9:00-9:20
Wilson, J.M.	5A	Saturday 11:10-11:30
Wilson, Scott	1B	Friday 11:00-11:20
Winch, Graham	Plenary	Friday 9:45-10:30

Systems Thinking and Modelling Interventions in Complex Problem Areas: *Park the Dogma at the Door!*

Graham W. Winch
Professor Emeritus in Business Analysis, Faculty of Social Sciences and Business,
University of Plymouth, UK PL4 8AA
President, System Dynamics Society, c/o Rockefeller School of Public Administration,
SUNY – The University at Albany, NY12222
graham.winch@plymouth.ac.uk

Abstract

There can be little dispute that increasing complexity is behind the major challenges in managing modern situations. This complexity derives from the overlap between contexts that might historically have been separate – widespread public-private ventures, business impacts on the environmental and *vice versa,* industrialised-nation perspectives and subsistence/ sustainability concerns in developing economies. However, not only are the core systems complex, but any study is within a further layer of complexity - the highly complex environments that surround them: interacting and possibly conflicting business, political and social agenda, widely extended stakeholder groups, and the baggage that all these people bring. The question has been raised as to whether OR – at least in its classical form - is capable anymore of supporting critical decision-making amidst this complexity. Using one complex intervention case as an example, this paper suggests that in trying to find solutions to intractable problems in these spheres, analysts can beneficially use a range of methodologies deriving from both the hard and social sciences, and OR processes have a clear role within this.

1. Introduction

In the late 1970s, the eminent scholar and consultant Russell Ackoff was invited to write an article presenting the interesting views he had expressed in an address to the British OR Society. The paper was entitled “The future of OR is past” (Ackoff 1979). In it he asserted his view that, essentially, the pressing managerial problems were moving from the management of physical resources – production and logistics – into areas of marketing, strategy, mergers/acquisitions, and industry re-organisation. As a result OR approaches were finding themselves no longer capable of addressing the contemporary problems that mattered. His criticism was scathing. In 2001 he wrote a further article (Ackoff 2001) in which he repeated his original points, reported that in his view no resurrection had taken place, and worse still catalogued the new strategic issues that have emerged – rationalisation of organisational structures, knowledge management, leadership, e-commerce. He believes that despite whatever advances there have been in the mathematical methods, OR is no better suited to these problems areas than it has ever been. In this latter article he dissociates himself from Operations Research, because “it has increasingly devoted itself too doing the wrong things righter”. He argues that systems thinking and practice are what OR could and should have become.

This is a pretty extreme view. If this is exactly the case and the complete story, then OR departments around the world are redundant, students and the organisations that eventually employ them are deluded, and we should find something better to do for the next two days. As far as I am aware there may be a dip in undergraduate recruits common with many technical subjects, but there is no glut of unemployed OR graduates, graduates are not demanding their fees back, and there has been no wholesale closing of university departments where funding-squeezed institutions have been forced to rationalise their offerings. Is Ackoff complete wrong therefore? I think that I can agree with much of what he says, but not that the only logical conclusion is a decent burial. In my view that would represent just as dogmatic a view as those who persist in refining techniques beyond their usefulness. I think it also ignores that progressive OR departments happily include "soft OR", systems thinking, soft system methodologies, and other similar advances, and people like Checkland and Eden are as widely cited as the fathers of the traditional scientific methods of OR.

There is an almost exactly parallel (or perhaps better a sub-set) debate in the area of system dynamics. There are those who argue that it cannot be system dynamics if a quantitative dynamic model is not constructed to simulate system behaviour, and the use of causal loop diagrams alone to infer dynamic behaviour is false science. The counter argument is that attempts to model quantitatively vague, contentious and invented constructs to capture social or "soft" processes do not always lead to better analyses, but rather confuse or mislead. (See Coyle 2000, Homer and Oliva 2001, Coyle 2001)

I believe that dogmatic extremism is most unhelpful in modern analysis. The complexity inherent these days in all problems that matter demands flexibility in applying our tools, and maybe a portfolio approach. I would like to describe briefly one particular study in which I am presently involved that typifies my recent experiences. The client here, a public healthcare provider, is facing a major challenge within a highly complex environment. This high level of complexity is quite typical for modern times, where organisations face situations which derive from the overlap between contexts that might historically have been separate. Here the complexity comes from overlapping EU and national agenda for work place practice, changing objectives and criteria for doctor training, and a dual funding system where the stakeholders have concurrent but different performance criteria. In such cases, not only are the core systems themselves complex, but studies must be conducted within the highly complex environments.

2. A Knotty Problem – *'trying to get a quart into a pint-pot'*

One of my on-going projects is with a NHS Trust in the UK, looking at what we have described elsewhere in a full description of the work (Winch and Derrick 2005) as the "knotty" problem of redefining and redesigning training programmes for junior hospital doctors at the large teaching hospital they operate. The word "knotty" is defined by Merriam-Webster Online as: "so full of difficulties and complications as to be likely to defy solution", and offers 'complex' as a synonym. An initial review of the circumstances certainly suggests that the problem posed here could well defy solution.

The original stimulus for the project was a directive from the European Union – the EU Working Time Directive (EUWTD) – which requires the application of stringent restraints on working hours, break time, and holiday arrangements for all workers in the EU. These apply equally to all professionals, including doctors. The restrictions of the EUWTD are compounded by other changes in working practice and significant changes in the training regime demanded at national government level. The term 'junior doctors' refers to all doctors who have graduated from medical school and are passing through the seven years or so of practical and specialist training to enable them to qualify for a consultant post. Historically, excessive hours have been an integral part of the life of junior doctors worldwide, with long hours on-call. Increased recognition of the safety implications for patients, and its adverse effects on learning and the health of doctors has lead to the introduction of the *New Deal* working hours regulations in the UK and a revised junior doctor contract. Alongside the EUWTD, this has limited junior doctor hours to 56 hours a week (from infamous figures of 80-100 + hours) with additional restrictions on duty lengths and rest breaks. Further reductions to 48 hours a week are currently scheduled to be in force by 2009.

Further changes to junior doctor working will be forced by the implementation of *Modernising Medical Careers*, a new initiative which aims to restructure and formalise junior doctor training, and since August 2005 a new training regime has been operating. This requires doctors to undergo a two-year general training programme – called the foundation programme – before moving into specialist training called ‘run through’ which is expected to last five years. The project described here focuses on the training programme for the first two foundation years. This is a radical change as in most cases it will shorten the total period doctors typically spent in training in the past, in addition to the reduced weekly hours. However, a critical element is still that whilst doctors are trying to complete all the requirements of their ‘training’ – knowledge acquisition and practical experience – they must also treat patients – provide ‘service’. Nominally their time should be devoted 50% to training and 50% to service, and the dual source of funding reflects this split. The challenge then is to design working regimes that allow the same, essential levels of training within a much shorter period of time, while at the same time enabling the doctors to provide similar levels of care to their patients. Any proposed new ways of working must be evaluated in light of all the potential consequences, beyond simple compliance and costs. This seems a problem where there can be no solution, but *no solution* is not an option!

The project is funded jointly from central government and by the local NHS Trust operating the hospital. The hospital provides acute care services to around ½ million people in the local region as well as covering most specialist treatments for a population of almost 2 million people. It has 1300 beds and employs approximately 450 junior doctors and 230 consultants. From the outset, it was realised that system dynamics was an appropriate approach for this situation (as did Ratnarajah & Morecroft, 2004, in a related study) and the original proposal specified this as the core approach in the analysis. However, the complexities inherent in both the system under study and its environment, the continuous implementation of new initiatives, and the need to satisfy a disparate group of stakeholders has meant the project has had to progress in a very flexible manner to keep on track and involve all key stakeholder groups.

3. The choice of an adaptive system-dynamics study process

The classic approach in this kind of case – the scientific method – would most likely be something like the four-phase consulting approach described by Lyneis (1999), as shown in Figure 1.

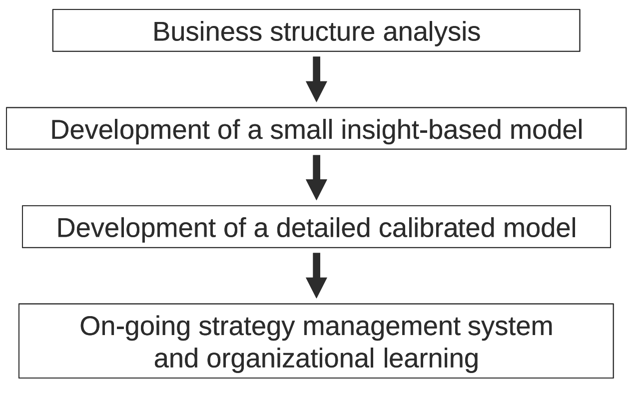


Figure 1 Typical study approach: *Jim Lyneis’ 4-phase process*

The project set out with the intention of following this kind of study process. However, it became quite apparent that simply working through the phases sequentially was simply not going to deliver. We therefore had to be flexible, to back-track, to include soft analysis using causal loop diagrams, and re-consider the level of detail in the final quantitative model.

3.1 Step 1 – Process analysis and project scoping

Early in the process, a scoping exercise was undertaken using cognitive mapping techniques to identify the perceived impacts of moving to the new regime. From this a cognitive map was developed with around twenty-five nodes representing different aspects of doctors' working practice, patient and other needs, and performance measures. The nodes were linked by arrows reflecting the direction of causal influences. This scoping diagram was then used in presentations and discussions with a wide audience within the Trust, including patient groups, consultants and management. Particular focus concentrated on the key factor that training and service are both requirements of the junior doctors, and they have to occur concurrently throughout. It was therefore hypothesised that all activities doctors undertake lay on a continuum where some are all, or nearly all, training with minimal service content, while others could be largely categorised as service with some small training content. A substantial database was set up on how the junior doctors actually spend their time, and this is reported in full in Derrick (2004), Derrick et al. (2005), Derrick et al. (2006).

This was a promising start. The scoping exercise produced a consensus view on the driving forces and critical interdependences, which along with the database, should then provide a firm foundation for quantitative model development (Eden 1994; McFadzean & O'Loughlin 2000). The next phase was to move to a relatively simple insight model.

3.2 Step 2 – First steps in model development - envisioning the system in terms of operations on the ground

This problem appears, *prima facie*, to be simply a matter of balancing the number of junior doctors needed as they work through the training process, senior doctors (consultants) and the hours they work, within the constraints of the EUWTD and other changes on training and service. Therefore, the main structure was a classic progression model with an 'ageing chain' of junior doctors progressing through training until they are qualified for consultant status. In addition, in order to examine the impacts on the critical soft variables like training effectiveness, the model must include further sub-sectors and formulation interacting with the ageing chain.

The modelling process was a group effort involving significant interaction with, and inputs from, the medical members of the team who were all active hospital consultants alongside their posts relating to the hospital's training programmes. The SD stock-flow diagram was developed relatively quickly with much of the detail relating to the career progression chain; this is the staffing structure that the hospital must manage in order to maintain its complement of doctors at all grades. While this was a reasonable representation of the system, it quickly became apparent that it had serious drawbacks. The level of complexity in the diagram was growing much faster than necessary for showing the overall feedback relationships and impacts, and for a first working model. While separating out each and every grade of junior doctor enabled the decision-makers to identify readily with the model as a representation of the real-life "stocks", it meant that they were focusing at a very detailed level on the operational structure, rather than concentrating on a simple representation so that an early simulation model could be developed.

This is typical of situations where the 'hard' elements of the system are readily assimilated – a hard model tends to stimulate hard thinking. Day (2000) has identified a similar phenomenon in software development which he blames for IT tending to over-focus on the technical issues. He also asserts that there is then a critical need to "moderate the dominance of the engineering metaphor" to ensure that the complexity in “human populated systems” is fully incorporated. It was also becoming so complicated that the medical team members (and the analysts!) were finding it very difficult to isolate the feedback loops visually, and these are the key structures in determining the true impacts of the new training regime. While this may have been genuine progress towards a quantitative model, it was not helping the team understand the structural changes envisaged nor the dynamic implications of the changes. Continuing with further development would therefore have been counterproductive, simply leading to a more refined and detailed portrayal of the ‘aging chain’ structure, whilst further diverting attention away from the key elements surrounding quality and effectiveness.

3.3 Step 3 – Back-tracking to re-establish objectives and boundaries.

At this point it was decided to shelve the emerging model, to refocus on critical feedbacks surrounding the training/service balance, and move from operational to more strategic thinking. The team reverted to developing a causal loop diagram which summarised the major feedback loops and relationships. This still reflected some of the more important day-to-day operational issues, but crucially incorporated qualitative dimensions that had been raised in earlier focus groups. The CLD has been valued as capturing well the broad system at the aggregate, overview level (See Figure 2).

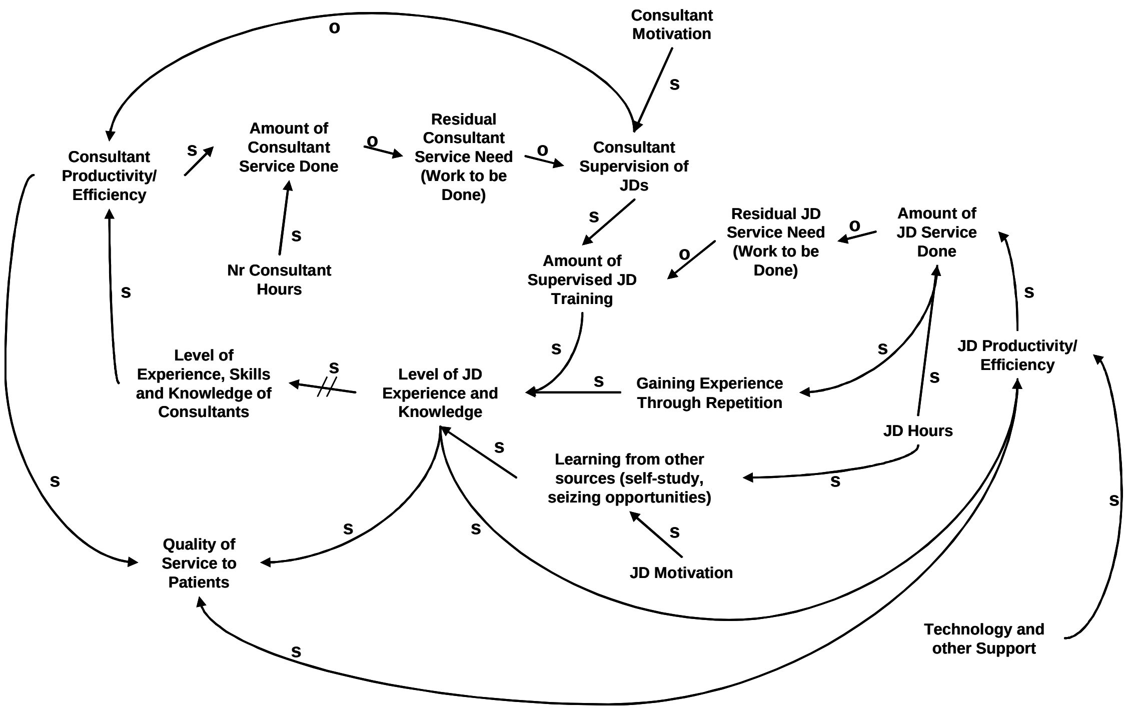


Figure 2 – CLD reverting to a broader reflection of the training system

One specific issue that had emerged was that it was perceived that the level of supervised training junior doctors receive from senior colleagues is a strong determinant of their acquisition of experience and knowledge. Consequently, focus has turned onto how supervision is being affected by the EUWTD and revised working patterns. The diagram clarifies how the reduction in hours of consultants alongside the continuing, and even increasing, demands on service impacts on the availability of consultants for junior doctor supervision. Other key issues that were being overlooked were aspects like the willingness of junior doctors to be more proactive in their learning – to look independently for learning opportunities rather than relying on the consultants to pass knowledge to them, the willingness of consultants to provide attentive supervision and the impact on their own effectiveness in patient service, and the impacts of different hospital working patterns like ‘hospital at night’.

The development of this CLD was most helpful in refocusing attention on the broader system, encompassing all the dimensions that drive the system and must be considered in any solution. It also proved a valuable aid to the team and other stakeholders in understanding the potential impacts of the changes.

3.4 Step 4 – Renewed work to produce the simulation model.

Even with the very positive benefits from the qualitative analysis, it was still considered essential that a simulation model be developed, both to cement the inferred results from Step 3 and to provide further quantitative impact analyses. The model is now in its final stages of development, and is a much simpler model. The buy-in of the medical people was essential, and the detailed representation of doctor career path supported this strongly, but working with the CLD confirmed to everyone that the level of detail emerging was not needed for the kinds of strategy and policy support required. As another example of the need for flexibility, the career progression structure had also to change to reflect the new post-August 2005 training structure with 'Foundation phase' and 'Run-through-grade' (RTG), replacing the traditional House Officer – Registrar path. This also includes 'peer training' – supervision and training undertaken not by consultants but by advanced junior doctors in RTG posts. It was also decided that to include within the model all the mechanisms that would capture processes in sufficient detail to allow scenarios that could reflect new practices like hospital-at-night, more self-directed learning, greater or less supervision, etc. would be very time consuming and dependent on many soft variable interactions. However, the database is available containing patterns of work typically undertaken at the junior levels, the activities they perform, whether or not they are supervised, and research into how each of these activities might be considered in terms of the training/service balance.

The final use of the model is therefore now a two stage process. In order to set up a scenario, the medical team members must first reset a spreadsheet that captures the work patterns, supervision, and the training/service split for each activity. This is a discursive process and itself a valuable learning and insight opportunity for the managers. Experience in other modelling situations requiring the active parameterisation of a model by a group has also proved this to be a strongly valued phase (see, for example, Winch & Arthur 2002). The spreadsheet crunches these data into a single overall split of junior doctors' work into training and service, with a figure for the anticipated consultant training overhead, which are then input to the model.

3.5 The emerging and adaptive approach

The project has gone through four phases, but is quite different from the classical consulting/scientific approach epitomised by the Lyneis model. During the project the environment was rapidly changing causing some rethinking in terms of model structure, but the main reason was project management needs, and the need to maintain focus on the critical drivers of the problem. Table 1 summarises the outputs from each of the phases and gives some idea of the nature of the activity, level of detail and focus.

The most glaring statistic is the balance of variables in the two simulation model variants; the first model was larger but included over three-quarters of the variables in the ageing/career progression chain structure. (Not only that, but with the attention of the medical team members on this section, there was every probability that this proportion would rise further.) In the final model the balance is significantly more on the softer service/training aspects – these reflect the critical performance measures for the hospital in terms of training received by doctors, service and care provided to patients, and the management of consultant and junior doctor hours. The final model was also much smaller than the first emerging model. This was largely because of the decision to pre-process much of the scenario related data externally in the spreadsheet, and then input the scenario particulars as three critical parameters.

Project Phase	Number of model variables	Propn. in Ageing chain	Propn. in service/ Training	Primary Study Focus
Step 1 – Project Scoping	--	--	--	Project scope & key system drivers
Step 2 – First SD Model	118	77%	23%	Operational detail
Step 3 – Back-tracking	--	--	--	Critical feedbacks in service/training
Step 4 – Final SD Model	55	35%	65%	Strategy/policy formulation & scenario building

Table 1 Summary of activities and model characteristics at each step

While final model experiments are continuing and final policy implementation has not yet occurred, the study has already yielded many valuable conclusions to support decision-making and has lead to successful funding applications for trials of other efficiency measures. The process itself has been judged valuable even before the final model is fully operational.

4. Concluding Comments

This presentation has attempted to show how a flexible, adaptive and recursive process has lead to the production of an effective simulation model. Even though not yet finally complete, the project has been a great success and is already supporting much greater understanding of the impacts of enforced changes and the development of robust and effective new strategies. But success has come at a price, and the price has been in terms of backtracking and changes of course. On the face of things this might have meant much wasted effort - for example, the detailed emerging model was shelved, and now has been discarded in favour of the simpler final version. However, the flexibility has been essential to keep the medical team members fully on board. This is also consistent with an observation by Winch (1990) that in large consulting projects, especially with a disparate client group, a sub-optimal modelling approach is often needed, and models may have to be 'over-engineered' to gain buy-in from all. Allowing them to develop a highly detailed representation of the ageing chain for the doctors' career path was a necessary step. It enabled them to see that the approach was able to accurately reflect operations on the ground, but at the same time it allowed them to decide that this detail was not necessary in the model to provide the kind of support they needed in strategy and policy formation.

Deborah Campbell (2001) in an amusing article titled "The long and winding (and frequently bumpy) road to successful client engagement" listed a number of important lessons for her and her team. One lesson was "Let the complexity come". She observed that some may not agree, but for her giving team members frequent and early opportunities get their contributions onto the table and into the model is a winning formula. Our experience of quickly going to a detailed operational representation to reassure team members that the system that they know so well will be modelled explicitly and accurately, and then letting a higher level view develop through a further causal loop diagram, while also unorthodox, has been very effective.

The case reviewed here has involved a number of different approaches including both hard and soft. That said, the distinction between the two is not that clear-cut, though Maani and Cavana (2000) have attempted to clarify the distinction. Even so, the use of cognitive mapping and causal loop diagrams here might still be regarded as 'hard' by Checkland (2000) and other Soft System Methodology purists, in that throughout, the modelling process is based on a 'world as a system' view. However, we would argue that discussions around the diagrams were not limited to extrapolating likely system behaviour purely as a deterministic process but it did draw in considerations of political processes, emotions and 'mystery'. We would further argue that it is easier and more effective to interpret the outcomes from the quantitative models in terms of world reality because it has been developed through this adaptive and recursive process. Like Mayo *et al*. (2001), the team have observed the value of the parallel use of a model as a qualitative framework and a quantitative analytic tool. It is also similar to the reasoning of Liddell & Powell (2004) who used their developed CLD – the QPID – to deal with "hybrid" systems comprising both hard elements and intractable soft elements.

While admissions of failure in the literature and practice are inevitably and understandably relatively scarce, there is a growing list of examples wherein a rigid insistence on pursuing a narrow, mono-tonic (dogmatic, can we say?) methodological pathway has lead to disappointing interventions and dissatisfied clients. On the other hand flexible approaches that combine and juxtaposition different approaches, or that are prepared to backtrack, diverge and re-converge, have been able to engage effectively the disparate stakeholders groups, take account of different agenda, and adapt to changing objectives and performance measures as externalities impinge.

This paper has reviewed modern thinking in terms of OR and related disciplines and used one particular case in an attempt to generate some insights for effective system thinking and modelling interventions in complex problem areas. Among the emerging conclusions, a couple of factors seem major determinants of success. Firstly, it is essential to manage clients' expectations, especially where early successes seem to point the way to increasing value as the study proceeds. Secondly, many analysts come to problems with a limited range of favoured or 'pet' approaches - those they are comfortable with and have seen work in previous studies, though maybe in less complex situations. I do not totally believe Ackoff's assertion that Operations Research is no longer capable of supporting decisions that matter -it is possible to combine soft and hard OR approaches, and to use a portfolio of tools to produce successful and valued outcomes. Slavish application of single tools is less likely to succeed than flexible, diverse, and integrated approaches – when entering the client's office *'park the dogma at the door'!*

5. Acknowledgement

The modelling study referred to in this paper is part of a larger project progressing under the Knowledge Transfer Partnership (KTP) scheme, and is jointly funded from government sources and by the Plymouth NHS Trust. I would like to acknowledge the contribution of the other team members on the project – Beryl Badger and Joan Chandler of the University of Plymouth and Doctors Tim Nokes, Adrian Dashfield and Jenny Lovett of Derriford Hospital, and particularly the KTP Associate, Sonja Derrick. The discussion of factors relating to the role of OR and OR project management processes presented in this paper represents the views of the author alone.

6. References

Ackoff R.L. 1979, The future of operational research is past, Journal of Operational Research Society 30 93-104

Ackoff R.L. 2001, OR: after the post mortem, *System Dynamics Review*, 17(4) 341-346

Campbell D. 2001, The long and winding (and frequently bumpy) road to successful client engagement: one team's journey, *System Dynamics Review: Special Issue on Consulting and Practice*, 17(3) 195-215.

Checkland P. 1999, System Think, System Practice (Including a 30-year Retrospective), Chichester: John Wiley and Sons

Coyle 2000, Qualitative and Quantitative modelling in system dynamics: some research problems, *System Dynamics Review*, 16(3) 225 - 244

Coyle 2001, Rejoinder to Homer and Oliva, *System Dynamics Review*, 17(4) 357 - 364

Day J. 2000, Software development as organizational conversation: analogy as a systems intervention, *Systems Research and Behavioural Science* 17(4) 349 - 358

Derrick S. 2004, Beyond reducing hours: Reconciling the junior doctor systems, Proceedings of the 46th. Operational Research Society Conference, York, September

Derrick S., G. Winch, B. Badger, J. Chandler & T. Nokes 2006, Exploring the Training/ Service Balance of Senior House Officer Activities" *Medical Education* (forth.)

Derrick S., G. Winch, B. Badger, J. Chandler, J. Lovett & T. Nokes 2005, Evaluating the impacts of time-reduction legislation on junior doctor training and service, Proceedings of the International SD Conference Boston, SD Society.

Eden C. 1994, Cognitive mapping and problem structuring for system dynamics model building, *System Dynamics Review Special Issue: System Thinkers, Systems Thinking* (eds. Richardson G.P., E.F., Wolstenholme and J.D.W. Morecroft) 11(2-3) 257 - 276

Homer J and R. Oliva 2001, Maps and models in system dynamics: a response to Coyle, *System Dynamics Review*, 17(4) 347 - 356.

Liddell W.G. & Powell J.H. 2004, Agreeing access policy in a general medical practice: a case study using QPID, *System Dynamics Review* 20(1) 49 - 73

Lyneis J.M. 1999. System dynamics for business strategy: a phased approach. *System Dynamics Review* 15(1): 37-70.

Maani K.E. & R.Y. Cavana 2000, Systems Thinking and Modelling, Auckland: Pearson Education

Mayo D.D., Callaghan M.J., & Dalton W.J. 2001, Aiming for restructuring success at London Underground, *System Dynamics Review: Special Issue on Consulting and Practice*, 17(3) 261-289.

McFadzean E. & O'Loughlin A. 2000, Five strategies for improving group effectiveness, *Strategic Change*, 9(2) 103 - 114

Ratnarajah M & Morecroft J. 2004, How might the European Union Working Time Directive, designed to limit doctors' hours, contribute to Junior Doctor Attrition from the British National Health Service and can desirable outcomes be achieved within these constraints? *JEMBA2004*.

Winch G.W. 1990, The Role of System Dynamics in Strategy Evaluation: A Consultant's View, In Applied Simulation and System Dynamics, (eds. Moscardini, A.O. and E.J. Fletcher), Northallerton: Emjoc Press.

Winch G.W., and D. Arthur 2002, "User-parameterised Generic Models: a Solution to the Conundrum of Modelling Access for SMEs? "*Special issue on Small Medium Enterprises - System Dynamics Review*, 18:3, 315 – 338.

Winch G. W. and S. Derrick 2005, Flexible Study Processes in 'Knotty' System Dynamics Projects, *2nd. European SD Workshop*, Radboud University, Nijmegen, Netherlands, March, (in review with J. of System Research and Behavioral Science)

Environmental and resource systems: Editors' introduction*

Robert Y. Cavana
Victoria Management School
Victoria University of Wellington
PO Box 600, Wellington, New Zealand
bob.cavana@vuw.ac.nz

Andrew Ford
Environmental Science and Regional Planning
Washington State University
Pullman, WA 99164-4430 USA
FordA@mail.wsu.edu

Abstract

This paper provides an introduction to the special issue of the System Dynamics Review on environmental and resource systems. The quantity and quality of previous system dynamics publications related to environmental and resource systems are briefly outlined. The background to the special issue is provided, together with a summary and comparison of the five papers and models in the issue. The papers relate to forestry in Indonesia, irrigated lands in Spain, renewable resource management in Norway, wildlife management in USA and blue-green algae bloom in the coastal waters of Australia.

Previous work

System dynamics has been applied to a wide variety of environmental and resource systems. One of the earliest applications occurred shortly after Dynamo became available in the 1960s. The problem was declining harvest of salmon in the Pacific Northwest. Researchers at the University of Washington used a Dynamo model to explain the over-investment in "gear" that contributed to the declining harvest (Paulik and Greenough 1966; Watt 1968). The most widely known application is probably the collection of models associated with the study of population growth in a world with finite resources (Forrester 1973; Meadows *et al.* 1972, 1973, 1974, 1992). Some of the recent applications are described in a special issue of *System Dynamics Review* devoted to celebrating the life of Dana Meadows (Sterman 2002).

Many of the applications are listed in the bibliography of the System Dynamics Society (2003). Figure 1 reveals the distribution of citations with key words "environmental" or "resource" over the time period from 1960 to 2002. There are 635 citations, approximately 10 percent of the total publications in the 2003 bibliography. The number of publications increased in the early 1970s, probably as a result of the

* The editors' introduction to the Special Issue of System Dynamics Review on Environmental and Resource Systems. Reproduced by kind permission from *System Dynamics Review* Vol. 20, No. 2, (Summer 2004): pp. 89–98. Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/sdr.293.

tremendous interest in the collection of models associated with *Limits to Growth* (Meadows *et al.* 1972). Figure 1 shows around ten citations per year during the 1980s and early 1990s followed by a dramatic increase at the turn of the century. The appearance of 70–80 annual citations in the past few years is certainly a dramatic development that might be attributed to growth in the System Dynamics Society, a growing awareness of environmental and resource problems, and the interest generated by the theme "Sustainability in the Third Millennium" (Davidsen *et al.* 2000) of the 18th International Conference of the System Dynamics Society held in Bergen, Norway. The surge of interest is also supported by publication of special issues of *System Dynamics Review* devoted to sustainable development (Saeed and Radzicki 1998) and to "The Global Citizen" (Sterman 2002), and other initiatives like the establishment of the Environmental Dynamics Special Interest Group of the System Dynamics Society by Tasso Perdicoulis.[1]

Figure 2 provides further analysis of the number of environmental and resource publications during the past four decades. This chart shows that energy and resource applications dominate the frequency of citations. Resource applications cover a lot of ground, so it is reasonable for this keyword to extract a large number of citations. Energy issues have received major attention of system dynamics practitioners, as explained by Ford (1999) and by Bunn and Larsen (1997). Figure 2 shows a significant number of citations for more narrowly defined keywords such as fisheries and earth systems.

The citation analysis shows a tremendous quantity of work, and these citations are mostly limited to those who elect to publish their models in the system dynamics literature. The number of citations would be even greater if we were to count the many modelling applications using system dynamics software such as Stella. These citations would include, for example, work by Costanza and Ruth (1997), Costanza (1998), Saysel and Barlas (2001) and the text by Grant *et al.* (1997). Counting such applications in the analysis would reinforce the trend evident in Figure 1. There is a major surge in interest in the application of system dynamics methods to environmental and resource systems.

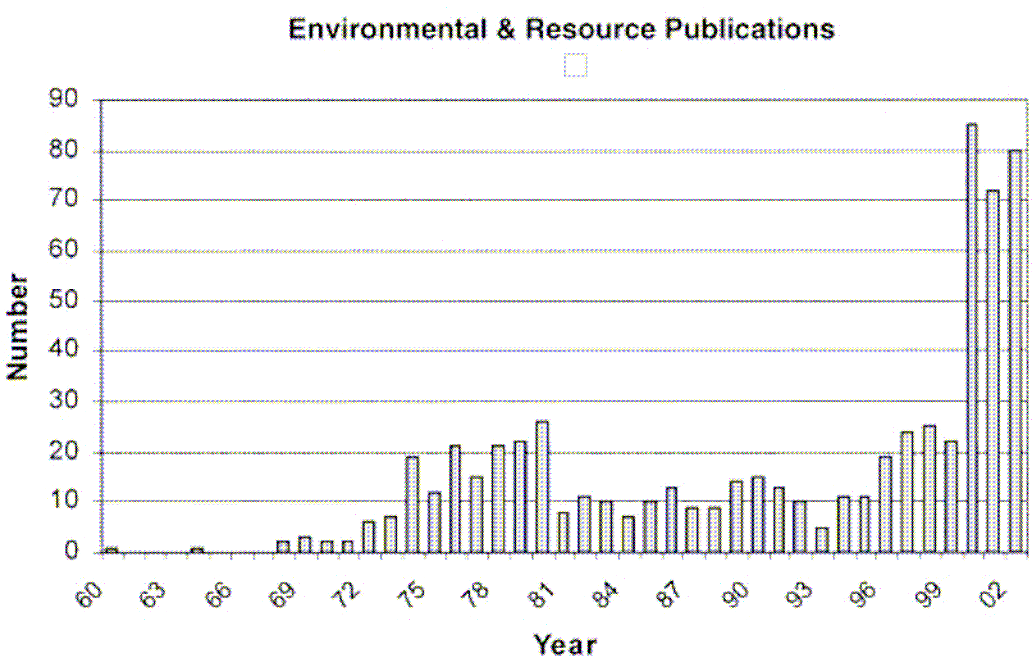


Figure 1. Number of publications listed in the System Dynamics Society bibliography extracted using the keywords "environmental" or "resource"

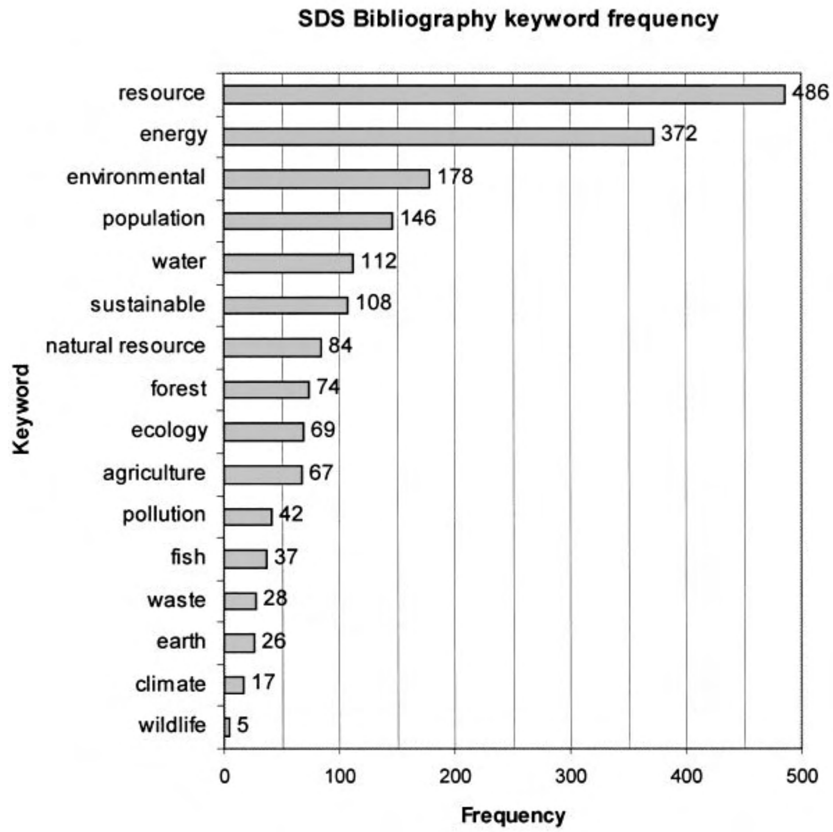


Figure. 2. Analysis of System Dynamics Society 2003 bibliography by keyword frequency

We now consider the quality of applications. Since we are interested in pragmatic applications, the ultimate measure of quality is success in model implementation and in shaping decisions and attitudes about the environment. From an academic perspective, an important indicator of quality is the Jay Wright Forrester Award. This award is granted up to once a year for the best contribution to the field of System Dynamics. The award was first granted in 1983, and there have been 18 awards through the year 2003. Four of the awards have honoured publications on environmental and resource systems:

- The 1991 award went to Dennis Meadows for the development of model based games such as "Stratagem." (Stratagem is a game to help national leaders plan development with limited resources, available from the University of New Hampshire, http://www.unh.edu/ipssr/Lab/Stratagem/html)
- The 1995 award honoured Khalid Saeed's (1991) work on modelling of sustainable development and for his book on *Towards Sustainable Development: Essays on System Analysis of National Policy*.
- The 1996 award honoured the work on electricity and conservation modeling by Andrew Ford (1990), as reported in an *Operations Research* article on uncertainty in the northwest electric system.
- The most recent award-winning work on environmental problems went to Erling Moxnes for his experiments on a cod fishery with Norwegian subjects. The

award-winning article was published in *Management Science*, "Not Only the Tragedy of the Commons: Misperceptions of Bioeconomics" (Moxnes 1998).

The Moxnes fishery experiment is the most relevant to this special issue, and it merits additional description. The experiment builds from the widely used "Fish Banks, Ltd" game by Meadows, Fiddaman and Shannon (1989). Participants in the Fish Banks exercise compete for harvest in an ocean fishery with open access. The participants typically over-invest in boats, deplete the fish populations and grow to appreciate how their own decisions lead to the "tragedy of the commons." Moxnes builds upon these findings in an experiment with Norwegian subjects operating a cod fishery with monopoly ownership. (Moxnes granted the subjects ownership to avoid the open access problem.) Consequently, his participants were able to avoid depleting the fish population, but they exhibited a strong tendency to over-invest in the number of boats. Moxnes documented this pattern and compared it to a similar pattern of over-investment in Norwegian fleets.

The tendency for decision makers to over-invest in capacity to exploit renewable resources is a chronic, serious problem that limits our ability to manage renewable resources in a sustainable, efficient manner. The problem is well known in fisheries, where Clark reports that investment in fishing capacity is often "much larger than twice the optimum level" (Clark 1985, p. 7). Over-investment also contributes to problems of sustainable management of forests, of watersheds and in grazing systems, as will be explained in articles in this special issue.

History and purpose

The special issue has been under discussion and development over the past three years. We volunteered as guest editors because of our own interests in resources and the environment (Cavana *et al.* 1984, 1996, 2003; Ford 1990, 1996, 1999). Our call for proposals was answered with over 30 proposals from researchers around the world. We received a truly impressive collection of proposals dealing with water resources, fisheries, forest management, energy consumption, land degradation, air pollution and conflict resolution. The quantity and quality of the proposals is further evidence to confirm the surge of interest shown in Figure 1.

The purpose of the special issue is to share examples of pragmatic applications of system dynamics modelling of important environmental and resource systems in a single issue. We have asked each of the authors to make their models available so that other researchers can build from their work. The models described here have been carefully documented, both in the papers and in the model files, which can be downloaded from a website devoted to the special issue (http://www.wsu.edu/~forda).

Comparison of papers

Table 1 provides a summary of the papers. We report the authors, models, location of the case studies and the issues under consideration. Table 1 then summarizes the models' purposes and the clients for whom the models were constructed.

Table 2 continues the tabular comparison with our summary of the main insights from each paper. We then list the value of each model and the authors' suggestions for future work.

Table 1. Summary and comparison of papers

Author(s)	*Dudley*	*Martinez Fernandez & Esteve Selma*	*Moxnes*	*Faust, Jackson, Ford, Earnhardt & Thompson*	*Arquitt & Johnstone*
Model	Log export ban	Irrigated landscapes	Misperception of. feedback	Wildlife population management (spectacled & grizzly bears)	Blue-green algae blooms
Geographic location	Forests in Indonesia	Irrigated lands in SE Spain	Reindeer rangelands of Norway	AZA zoos, & Greater Yellowstone Ecosystem, USA	Coastal waters of Queensland, Australia
Issue(s)	Substantial increases in illegal logging & deforestation	New irrigated lands leading to overexploitation of available water resources	Over use of renewable resources due to misperceptions of the dynamics	Improved management and conservation of wildlife populations in captivity and natural habitats	Algae blooms threatens water quality, coastal ecosystems & harmful to humans
Model purpose	To understand the effects of a log export ban on the forestry sector	To analyse the key socio-economic and environmental factors leading to overexploitation.	Simulators to test decision making of reindeer management with lichen as limiting resource	To assess the impact of management actions, help guide data collection and allocation of conservation resources	Scoping and consensus model to understand the dynamics of algae blooms and for developing research directions
Client/ sponsor/ decision maker	Centre for International Forestry Research, public & private forestry managers	Public policy makers (Spanish National Water Plan), local decision makers, general public, and technical people	Public and private managers of renewable resource systems	Population biologists, zoo managers, national park administrators and advisors	Committee established by Environmental Protection Agency, Queensland from Govt depts, academics and technical people, and community organisations.

Table 2. Further comparison of the papers

Author(s)	*Dudley*	*Martinez Fernandez & Esteve Selma*	*Moxnes*	*Faust, Jackson, Ford, Earnhardt & Thompson*	*Arquitt & Johnstone*
Model	Log export ban	Irrigated landscapes	Misperception of. feedback	Wildlife population management	Blue-green algae blooms
Main insights	For a log export ban to be effective tool in combating over-harvesting and illegal logging, limits must be placed on potential domestic milling and logging capacity	Current policies based on expanding irrigated agricultural land following increased planned water supplies lead to further expansion and unsustainable water deficits	The experiments show that even in simple systems information feedback is not sufficient to make up for misperceptions.	Monte Carlo simulations help population biologists understand and combat the frequency and risks associated with rare outcomes (eg population extinction or explosion).	Simulations show that occurrence of Lyngbya bloom is sensitive to a number of uncertain parameters and model structures, such as the size of the iron pulse entering the coastal waters.
Model value	Provides a framework for thinking about the system wide implications of a log export ban	Contributes to the systemic analysis of intensive irrigated lands and water management in Spain	The experiments and simulators provide both a motivation for and an introduction to studies of system dynamics	Population models can be useful tools to establish scientific based management decisions and to facilitate experimentation and communication.	Contributes to the understanding of the dynamics of algae blooms of interest to the wider international community facing similar problems.
Future work	Further development of model and framework to assist strategic planning and policy making for sustainable forestry in Indonesia	Group model building workshops with stakeholders to change culture to a more participative management approach. Further links with GIS for spatial distribution analysis	Further studies of dynamic systems, and signals a need for improved education in dynamic systems in general	Further development of the models in conjunction with other forms of population modelling for managing other species also in captivity or in the wild.	Iterative development of second-stage research model and third-stage management model that are well grounded scientifically and accepted by stakeholders.

Comparison of models

Table 3 provides a tabular summary of the models described in the papers. The table compares the models in terms of attributes that are easy to count and tabulate. The comparison shows that Dudley's model of log exports operates over the longest time horizon. (The forest model runs for 60 years, sufficient time to see the impact on a slowly changing forest in Indonesia.) The model of hydrology and irrigation runs for nearly 40 years to keep track of the gradual depletion of aquifers in south-east Spain. The model of grizzly bears in the GYE (the Greater Yellowstone Ecosystem area) also requires a long simulation length of 50 years.

Two of the models in Table 3 require unusually short time steps to accurately simulate fast acting dynamics within a model with a much longer simulation time horizon. Dudley's model of log exports simulates for 60 years, but requires a DT of 0.0078125 years to simulate the fast-acting dynamics. The toxic bloom model requires a DT of 0.03125 half days, an extremely short time step for a model that runs for 180 half-days. The toxic bloom model uses higher order integration methods (fourth-order Runge Kutta) to aid in the numerical accuracy. All of the other models use the standard, Euler method of integration.

The models range considerably in size. The smallest is Moxnes's model of grazing intensity, a model with one stock and five equations. The largest model, as measured by equations, is the model of irrigation in south-east Spain. It uses 12 stocks and 265 equations. The model of the spectacled bears is the only model in the group to make use of "arrays" (or subscripts).

Table 3. Comparison of Models

Author(s)	*Dudley*	*Martinez-Fernandez & Esteve-Selma*	*Moxnes*	*Moxnes*	*Faust, Earnhardt & Thompson*	*Jackson & Ford*	*Arquitt & Johnstone*
Model	Log export ban	Irrigated landscape s in Spain	Misperceive d feedback T1	Misperceive d feedback T2	Spectacled bears in AZA	Grizzly bears in GYE	Toxic algae blooms
Software	Vensim	Vensim	Excel	Powersim	Stella	Stella	Stella
Unit of time	Years	Months	Years	Years	Months	Months	Half-days
Time step (DT)	0.0078 years	0.25 months	1 year	1 year	1 month	1 month	0.03125 half-days
Simulation length	60 years	468 months	15 years	15 years	362 months	600 months	180 half-days
Equations	133	265	5	14	85	171	183
Stocks	8	12	1	2	16	15	15
Sectors	6	5	1	1	4	7	4

Concluding remarks

The special issue begins with models with the highest degree of human involvement in shaping the system. We begin with Dudley's model of log exports from Indonesia. He shows the difficulty in managing logging in a sustainable manner when there is over-investment in saw-mill capacity. He argues for limits on domestic milling and logging capacity if bans on log exports are to be effective. The second paper describes the over-investment in irrigated lands in south-east Spain. The authors present a model to explain the over-investment despite the existence of a national plan for water resources. They use the model to show the depletion and degradation of the aquifers and the decline in yields from poor irrigation water. The third paper describes the tendency for reindeer managers to over-stock the grazing lands even though the subjects are provided with sufficient information to construct perfect mental models. These three papers expand upon important contributions by previous system dynamics practitioners who have used models to show the problems of managing renewable resources in a sustainable and efficient manner. The papers help us understand the adverse impacts of over-investment and the reasoning that leads decision makers into the trap of over-investment.

The fourth paper deals with wildlife management with examples from animals in captive zoo populations and animals in the wild. The zoo example is the spectacled bear population, one that is heavily managed by humans who control the bears' diets, their location and veterinary care. The wild example is the grizzly bear population of Yellowstone, a population that is influenced by humans who control the location of human infrastructure in Yellowstone and who deal with the bears who stray into the human zones in search for food.

The final paper differs from the others in that human involvement is limited only to the release of iron into the coastal waters, the event which is hypothesized to trigger an outbreak of a toxic bloom. The model presents a novel application of system dynamics in an area with considerable, scientific uncertainty. The model is presented as a device to explore alternative hypotheses on the cause of a toxic bloom. (The authors describe the model as a first-stage, scoping model for setting priorities for empirical investigations.) Although the authors take care to limit the intent of the model, the model has already delivered pragmatic benefits to one referee, who reports an improvement in a costly environmental monitoring plan based on the insights gained from the paper.

Acknowledgements

First, and most importantly, we wish to thank the authors who have worked to write and rewrite their papers on a tight schedule. We also acknowledge the authors' extra effort to provide a clearly documented version of their models, which will be available for downloading on the website. We thank the many anonymous referees who provided reviews on an accelerated schedule, and we are also very grateful to Brian Dangerfield for his very helpful editorial support throughout the preparation of this special issue. We would also like to thank our universities for their support for our time spent on the issue. Robert Cavana thanks the University of Wellington, and the Centre for Systemic Development and the Department of Agriculture and Environment at the University of Western Sydney for their support during his 2003 sabbatical leave. Andrew Ford thanks the Program in Environmental Science and Regional Planning at Washington State University.

Note

1. The Environmental Dynamics Special Interest Group (ED SIG) of the System Dynamics Society was established in 2002, and focuses on systems involving human activities and their natural environment, at any scale, with the intent to understand and administer them better—an objective also referred to as Sustainable Development. The ED SIG meets once per year, at the SDS conference, but marks its presence with a website (http://home.utad.pt/ed) that features, *inter alia*, a Resource Database, a discussion list (SDsustain), and an activity calendar. The ED SIG is preparing its own publication, designated as the Repository, to collect and share academic communications in its field of interest. (Anastassios Perdicoulis, personal communication, 2004).

References

Bunn, D.W., E.R. Larsen. 1997. *Systems Modelling for Energy Policy*. Wiley, Chichester.

Cavana, R.Y. 2003. "Book review: *Modeling the Environment: An Introduction to System Dynamics Models of Environmental Systems* by Andrew Ford." *System Dynamics Review* **19**(2): 171–173.

Cavana, R.Y., R.G. Coyle. 1984. "A policy analysis model for New Zealand's Plantation Forestry System." *New Zealand Journal of Forestry* **29**(1): 24–50.

Cavana, R.Y., M.W. Lee, J. Bennett, and R.J. Taylor. 1996. "Possum and gorse control on a farm woodlot: a system dynamics analysis." *Asia-Pacific Journal of Operational Research* **13**(2): 181–207.

Clark, C. 1985. *Bioeconomic Modeling and Fisheries Management*. Wiley, New York.

Costanza, R. (ed.). 1998. *Ecological Modeling* (special issue on modeling in Stella). Volume 112.

Costanza, R., M. Ruth. 1997. "Dynamic systems modeling for scoping and consensus building", in *Sustainability and Global Environmental policy: New Perspectives*, Dragun, A.K., K.M. Jakonson (eds). Edward Elgar, Cheltenham.

Costanza, R., M. Ruth. 1998. "Using dynamic modeling to scope environmental problems and build consensus." *Environmental Management* **22**: 183–196.

Davidsen, P.I., D.N. Ford, and A.N. Mashayekhi (eds). 2000. *Sustainability in the Third Millennium: Proceedings of the 18th International Conference of the System Dynamics Society*. Bergen, Norway. System Dynamics Society: Albany, NY.

Ford, A. 1990. "Estimating the impact of efficiency standards on the uncertainty of the northwest electric system." *Operations Research* **38**(4): 580–597.

Ford, A. 1996. "Testing the Snake River Explorer." *System Dynamics Review* **12**(4): 305–329.

Ford, A. 1999. *Modeling the Environment: An Introduction to System Dynamics Modeling of Environmental Systems*. Island Press: Washington, DC.

Forrester, J.W. 1973. *World Dynamics* (2nd edn). Productivity Press, Cambridge, MA. (Now available from Pegasus Communications, Waltham, MA.)

Grant, W., E. Pedersen, and S. Marin. 1997. *Ecology & Natural Resource Management: Systems Analysis and Simulation*. Wiley, New York.

Meadows, D.H., D.L. Meadows, J. Randers. 1992. *Beyond the Limits: Confronting Global Collapse, Envisioning a Sustainable Future*. Chelsea Green, Post Mills, VT.

Meadows, D.H., D.L. Meadows, J. Randers, and W.W. III Behrens. 1972. *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*. Universe Books, New York.

Meadows, D.L., W.W. III Behrens, D.H. Meadows, R.F. Naill, J. Randers, E.K.O. Zahn. 1974. *Dynamics of Growth in a Finite World*. Productivity Press, Cambridge, MA. (Now available from Pegasus Communications, Waltham, MA.)

Meadows, D.L., T. Fiddaman, D. Shannon. 1989. *Fish Banks, Ltd*. Institute for Policy and Social Science Research, University of New Hampshire, Durham, NH 03824.

Meadows, D.L., D.H. Meadows (eds). 1973. *Toward Global Equilibrium: Collected Papers*. Productivity Press, Cambridge, MA. (Now available from Pegasus Communications, Waltham, MA.)

Moxnes, E. 1998. Not only the tragedy of the commons: misperceptions of bioeconomics. *Management Science* **44**(9): 1234–1248.

Paulik, G, J. Greenough. 1966. Management analysis for a salmon resource system, in *Systems Analysis in Ecology*, Kenneth Watt (ed.). Academic Press, New York.

Saeed, K. 1991. *Towards Sustainable Development: Essays on System Analysis of National Policy*. Progressive Publishers, Lahore.

Saeed, K., M. Radzicki. 1998. "Sustainable Development." A special issue of *System Dynamics Review* **14**(2–3).

Saysel, A.K., Y. Barlas. 2001. "A dynamic model of salinization on irrigated lands." *Ecological Modelling* **139**(2–3): 177–199.

Sterman, J.D. (ed.). 2002. "Special Issue. The Global Citizen: Celebrating the Life of Dana Meadows." *System Dynamics Review* **18**(2).

System Dynamics Society, 2003. *System Dynamics Bibliography* (User Notes Version 03 B, July 2003). System Dynamics Society: Albany, USA. Available at http://www.albany.edu/cpr/sds/biblio/sdbib.html [20 March 2004].

Watt, K. 1968. *Ecology and Resource Management: A Quantitative Approach*. McGraw-Hill, New York.

A Historical Review of Electricity Demand Forecasting

Jonathan Lermit

Jonathan.Lermit@paradise.net.nz

Abstract

Changes in forecasting methods from the 1970s onwards are reviewed. Initially bottom up approaches were used where forecasting was done by adding up a lot of individual estimates from Supply Authorities each forecasting in their particular geographical areas. While this technique had worked successfully in the 1950s and 1960s, by the 1970s this was no longer effective because of the interaction between geographical areas and the saturation of consumption. This led to the development by a number of people of alternative regression and econometric techniques which proved to be more accurate for national load. Subsequent developments of the electricity market in the 1980s and 90s has again seen a shift to other techniques which have proved satisfactory for individual companies on the much shorter planning horizons on which they operate. We are now seeing a return to some of the older techniques operating at a national level being undertaken by organisations such as the Electricity Commission. The likely evolution of forecasting methods over the next decade remains an open question.

Keywords: Forecasting, electricity

Through the looking glass: A personal history of OR in the electricity sector over the last 20 years.

Mark Pickup
Electricity Commission

1. Pre History

Today I would like to touch very briefly on some of the history of OR in the New Zealand electricity sector, as seen though my own personal experiences.

I joined the NZED in 1986 and was employed as a scientist or power planner, which was a very impressive title for a recent OR graduate. I was doubly lucky in that the group I joined was heavily involved in the use and development of a large OR simulation model for the purposes of long term investment planning.

This model, known as PRISM which if I remember correctly stood for Plant and Reservoir Integrated Simulation Model. PRISM was a two reservoir stochastic dynamic program and was used to study investment options over a 20 to 30 year horizon. The model prior to this was known as MONTE, a Monte Carlo style simulation with extensive heuristics. The break through with PRISM was the use of a stochastic DP to evaluate water values.

PRISM I recall took just over two minutes to solve on the then Ministry of Works IBM 3090 mainframe. If memory serves me correctly, this was charged between government departments at about $100/minute. While this may not seem much money, a typical investment plan would require many hundreds of optimisations. Hence it was not an idle boast that our top power planner of that time was knick named the million-dollar man.

1987 brought the corporatisation of many government departments, of which one of the largest was the new Electricorp, created from the NZED. However, from an OR/modelling perspective, the most significant change I remember at the time was a shift in discount rate from 5% to 7%, and a sudden emphasis on learning about WACC and tax shields.

1.1 Enter SPECTRA

In the early 90's SPECTRA was developed as the replacement for PRISM. This model incorporated improvements over PRISM, but remained in essence a two reservoir stochastic DP.

The characteristics of the investment problem of the time were:

- Minimum cost objective function
- Single investor for transmission and generation
- Low, Medium and High demand scenarios
- Complexity in representing physical characteristics of the problem - hydrology and stochastic reservoir management

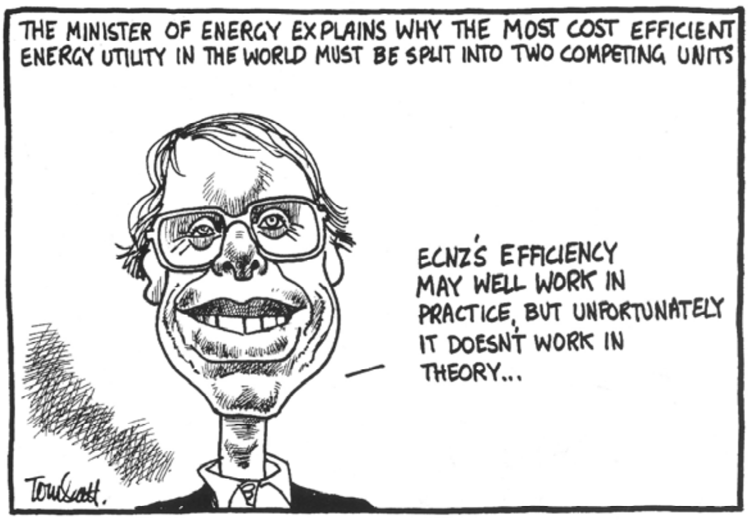
SPECTRA remained Electricorp's and later ECNZ's main investment analysis tool for the following decade until the ECNZ's demise in 1999. During the 90's however, several significant events occurred that affected the use of these OR tools for investment planning in the corporation.

The first event was the winter crisis of 91/92 and the resultant drought. Among other things this led to a reassessment of many of the key concepts in the SPECTRA model, and the planning margins that were being adhered to.

Criticism was very public at times, as the following cartoon alludes, and the shortage review committee was very critical of some of the assumptions made by ECNZ's models; but I would like to think that the end result was an improvement in how ECNZ used SPECTRA.

1.2 Market Crisis

The early to mid 90's bought several further changes. The first was the split of Transpower, in 1994. The second was the formation of Contact Energy in early 1996, followed shortly after with the wholesale electricity market in October 1996. This measure was also controversial with some of the public.



Together these events bought about a fundamental change in the nature of the investment problem. - That is, there was no longer a single central investor. Now the invisible hand of the "market" would guide investment decisions.

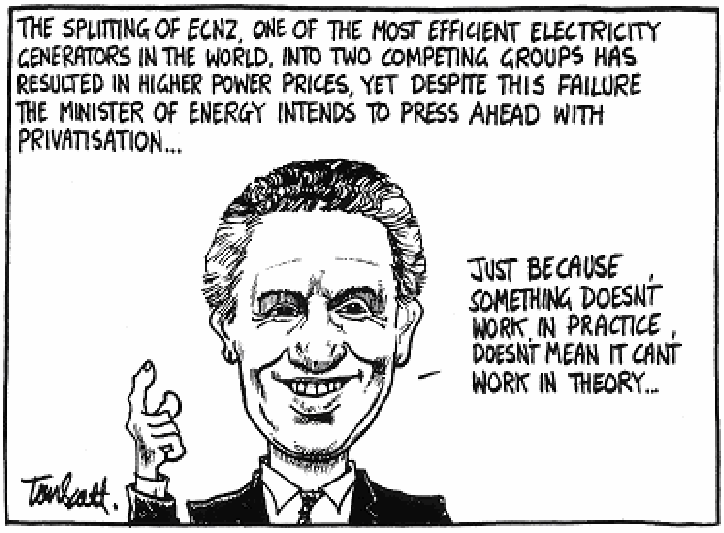
SPECTRA had been designed for a simpler world; a world where the main unknown was hydrology, not potential competitor behaviour. Was SPECTRA sufficient for this new paradigm? In the absence of a better replacement, SPECTRA continued in use.

1.3 Gaming models

Starting in the mid 90's the concepts of gaming theory were adapted to these planning models. Here the assumption could be made that participant behaviour would follow a profit maximising or Cournot strategy, and the least cost optimised strategy could be perturbed in some fashion in order to maximise participant profit at each stage of the simulation.

Several Cournot models have been developed in NZ. DUBLIN, as successor to SPECTRA was under development when ECNZ came to an end. Subsequently it has been used by the Commerce Commission for merger and acquisition studies in the electricity market.

Gaming models have been used for risk studies, hedging decisions and competition analysis as mentioned. But for longer term market supply and demand balance issues, the gaming aspect is of less utility, as it is very difficult to forecast generator fuel and hedge positions. Both of which are a key determinant of gaming behaviour.



The mid 90's saw further change in the market with a final split in ECNZ. Again as the cartoon by Tom Scott illustrates this also was a controversial move.

2. The Present Day

2.1 The Commissions Role

This brings us to the present day, where the most significant recent change has been the abandonment of the market's self-governance model, for a regulated market, under the oversight of the Electricity Commission.

It is here, from an OR perspective, that things start to get even more interesting.

While the problems of forecasting and modelling a competitive market remain as complex as before, the regulator has a different modelling perspective to individual participants. Paradoxically, those very same tools – SPECTRA, that were perhaps less suited to the competitive market, come back into their own again.

Hence my illusion in the title to "through the looking glass", those of you familiar with Lewis Carol may recall the scene in the Red Queens garden where Alice has to run as fast as she can just to stay in to the one spot. Well our "OR" Alice has been running in the one spot carrying out her SPECTRA optimisations for many years now, and now twenty years later she's still running in the same spot.

As a regulator, the Electricity Commission has a neutral perspective and must consider the market and security of supply from a national perspective. Unlike the central planners in the past however, the Commission lacks the authority or ability to direct market investment; this function remains with the market.

In particular, the Commission is required to:

1. forecast supply and demand,
2. investigate and rule on Transpower transmission investments, but
3. it does not regulate generation investment decisions.

To carry out these functions the Commission developed its own forecasting models and is in the process of developing the tools it needs for generation and transmission investment assessment.

2.2 The Commissions modelling/OR approach

Initially the Commission adopted SPECTRA for expediency. But while there is still life in the old model, some of its decades old limitations have now come back to bite us, such that the Commission early on decided to adopt the Brazilian model SDDP.

SDDP adds the capability to model the transmission system, and can be extended over multiple reservoirs. But its significant shortcoming is speed. Currently a 20-year monthly optimisation over the full grid takes over 20 hours on my PC. Of course minus the grid and with fewer reservoirs a shorter time can be achieved. For example a ten year monthly SDDP run, without the grid, is on the order of two hours. Faster yes, but still no where near as friendly as the 20 seconds it takes SPECTRA for a 30 year weekly run on my low spec Linux PC.

For that reason the Commission continues to explore other technologies to enable us to evaluate our supply and demand scenarios, including continuing to use SPECTRA for an initial problem exploration.

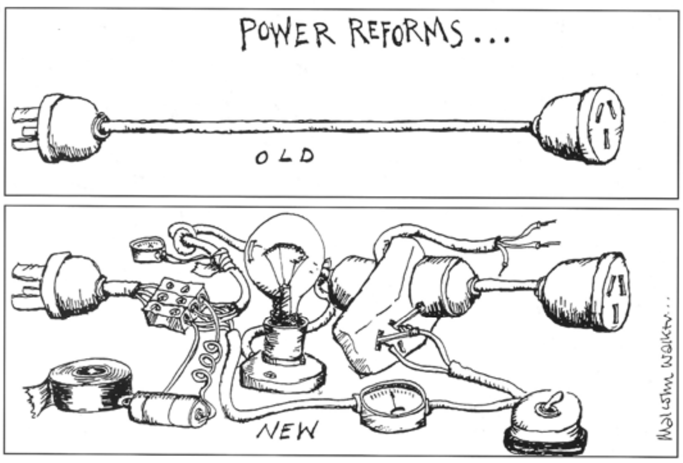
The Commission's approach to supply and demand modelling at present uses a scenario approach. A scenario approach postulates a credible set of future circumstances, and builds a limited set of optimised plans around these scenarios to map the extreme points of the solution space. Within these scenarios, Monte Carlo methods are used to map out distributions of key decision variables for sensitivity testing.

The scenario approach is appealing for its simplicity. But reality is not a convex combination of the extreme scenarios, hence the optimal result, with perfect foresight, can not be replicated by simply taking a weighted combination of the scenarios. To date, the Commission in its scenarios has not picked a preferred or 'average" strategy. However, we are aware that in some circumstances a "base" scenario may be useful.

Proper scenarios construction is more of an art than a science. International consultants even conduct courses on how to construct scenarios. The Commission's energy scenarios however, must be considered in the context for which they are intended. They are required to assist people to assess transmission and alternatives, that is, generation and demand response. Thus these scenarios must present a "reasonable" or credible and plausible set of possibilities in that context. Last year the Commission used fuel availability and fuel type as the primary determinant of its scenarios. This year the Commission intends taking a wider economic perspective (not just fuel) as its basis for scenario development. This is work still in progress.

In terms of model usage the Commission remains committed to the use of SDDP as a tool. However, due to the long solution times with full grid optimisation, the Commission continues to use SPECTRA for broad scenario development, changing to SDDP for finer tuning against the grid.

The Commission is also working on how best to incorporate options pricing and related risk methods to investment uncertainty. This work is preliminary and still in progress. It is the Commissions hope that through these efforts, the confusion alluded to in the following cartoon, can eventually be reversed.



The Allocation of Specialised Equipment to Stations for the New Zealand Fire Service

Luke R. Hannah and Scott D. Willson
Department of Management
University of Canterbury
New Zealand
luke.hannah@gmail.com

Abstract

The New Zealand Fire Service (NZFS) has specialised equipment which is used during fire incidents and emergencies throughout the country. This equipment is expensive and its use is becoming more important at incidents for the rescue and safety of personnel as well as in the salvage stage of an incident. A model for allocating both passive and aggressively-used equipment is developed and Tabu Search is used to solve this model with a multi-criteria objective function. The primary function of the model is to evaluate different allocations of equipment to fire stations. The model could also be used for analysing where further expenditure on equipment could improve the level of service.

1 Introduction

The aims and goals of the New Zealand Fire Service are summarised by their mission statement – "To reduce the incidence and consequence of fire and to provide a professional response to other emergencies." This paper and the analysis undertaken aims distinctly at supporting this professional response to emergencies by ensuring the allocations of equipment to stations, and their appliances, are robust. The software package - PYRO (Planning Your Resource Organisation) is developed to model the aforementioned problem situation to assist in achieving a satisfactory level of service, which is of fundamental importance to community residents. Of equal importance is the secondary analysis which aims at providing information regarding future allocations when more or less equipment is available.

1.1 Problem Situation

In this paper, the Transalpine region of New Zealand is holistically viewed as a network of points. The stations act as supply nodes, from which a service can be provided. The zone centroids act as demand nodes; the demand attributed to each of these nodes is classified as the number of callouts that are estimated to occur within each of the zones. Response time estimates are used for every station and zone combination to provide quantitative data, in order to differentiate between allocations.

To create the model, three types of valuable equipment have been used to base our research and development on. The first is a Positive Pressure Ventilation Fan (PPV Fan); which is used in the salvage stage of an incident to ventilate smoke-filled dwellings. This equipment is used passively, as its use has no bearing on when the fire is extinguished or the time necessary needed to remove the threat to life. Typically this equipment will be used after the threat of further damage to personnel or property has been removed and when the salvage/clean-up process begins.

The second and third pieces of equipment are both used aggressively. An aggressively-used piece of equipment helps to extinguish the fire and provide emergency response to incidents, thereby reducing the threat of casualty. The second test item, the Thermal Imaging Camera or TIC, is used to detect areas of heat, such as body heat or any remaining hot spots within smoke filled dwellings. They are valued at $30,000, and can be used at a wide range of fire incident classifications. The third test item is the Gas Detection Unit (GDU), valued at $10,000, which has only recently been made available to the NZFS for use at a restricted number of fires. The amount of data available on its usage is therefore limited.

The primary goal of our allocation of these pieces of equipment is to enable them to reliably arrive at an appropriate fire incident as quickly as possible. For passively-used items, the onus is on placing the equipment around the network of stations in a fashion that reflects incident demographics with respect to the travel time. For aggressively-used items the focus shifts toward getting the equipment to as many incidents as possible within a time specified by the NZFS.

1.2 System Boundaries

Figure 1:

Aside from the obvious geographical boundary that defines the Transalpine region shown in Figure 1, boundaries placed on the analysis and scope of this project necessarily include a number of limitations and assumptions regarding operational aspects.

The model will only deal with equipment that is of either an aggressive or passive nature. Equipment that could possibly be classified as both has been excluded from the model. Equipment is allocated on a station basis, and the decision as to which appliance at the station that the equipment is assigned is left to the fire service to decide.

2 Data

Given the nature of the problem situation outlined above, large amounts of data had to be sourced and maintained appropriately so that inconsistencies could be detected in the overall quality of each set, and as a consequence, its validity reviewed. This section will provide a summary of the data that has been acquired, for what purpose, and the intentions of its use in the overall model.

2.1 Zone/Station Information

As the wider network is separated into a network of demand and supply nodes, accurate estimates pertaining to these exact locations are useful in generating the required response/arrival time estimates. The data we were provided with included unique identifiers, X-Y coordinates of predefined NZFS geographical centres for zones, exact X-Y coordinates for stations and a name of the location for every station and zone. These centres were a valuable reference point for where the demand for incidents within each respective zone occurred if no data on callouts within the zone were available.

2.2 Callout Information

Arguably the most important aspect of modelling emergency response over any region is the supporting callout information. We acquired from our client a set of data which included a unique identifier or CAD Number, call creation/stop times/departure times, X-Y geographical co-ordinates, zone number and classification grouping for 49,330 callouts over a five-and-a-half year period. Various aspects of this data were used for different aspects of the modelling process.

2.2.1 Classification/Zone Grouping

The callout data was first grouped by the classification of each callout. In total there are 18 different fire classifications, such as Structure Fires, Mobile Property Fires, or Flammable Liquid, Gas Incidents, to name a few, of which subsets are created that reflect incident types potentially requiring the respective pieces of equipment. Secondly, grouping sets of callout data by zone can not only be used to visually display the distribution of incidents within a given zone, but is also an effective means of storage.

2.2.2 Average Centre Calculations

As the wider network exists as a set of demand points, coordinates representing any location within each zone are reflective of where incidents occur within the zone and are important to base arrival time estimates on. As mentioned above in section 2.1, the zone-indexed callout data identifies predefined NZFS approximate geographical centre coordinates for each zone, most of which do not lie on a road network. This has proved an important piece of information when using ArcGIS 9.0, the Geographical Information Systems software used to provide the arrival time estimates. Average centres, rounded to the closest call (based on the callouts within the zone) were calculated for each zone, as a representation of the zone's 'true' centre. In the case of zones for which no data was available the geographical centre was used.

2.2.3 Assigning Weightings

The most important aspect in assigning weightings within the given network of zones is that they reflect the incidence of events within each zone. Using the callout data grouped by zone, the expected number of callouts within each zone becomes a useful representation of the weighting for each given demand node (i.e. a zone).

2.2.4 Calculating Overlaps

When PYRO is run, a large emphasis is placed on the time a servicing piece of equipment is spent in overlap, that is, when it is needed at more than one incident at one time. To calculate the overlap the number of pieces of equipment servicing each particular zone from within a contiguous set of zones is used to estimate the expected fraction of time spent in overlap. Overlaps are calculated by contrasting the difference in time between the creation time of a specific call with the stop (or departure) time (i.e. the time at which an event is deemed 'under control') of any previous call. In the instance of a particular zone being serviced by $X > 1$ pieces of equipment, it determines the fraction of time $Y > X$ calls are occurring at one time within this zone. This ensures that heavily weighted zones, which may potentially require service from more than one piece of equipment, are accounted for when each equipment items' overlap is calculated.

2.3 Time Estimates

The allocation of one or more pieces of each type of equipment to a station needs to reflect the zones an allocated station can realistically service within an acceptable time, if called upon to do so. Therefore time estimates between stations and zones need to be sufficiently accurate that the zones an allocated station needs to service is reflective of reality.

Initially, estimating these times was focussed on using the aforementioned callout data grouped by zones and finding the mean time, with an accompanying standard deviation, for each station/zone combination for which data existed. This proved highly unsuccessful when only 2.3% of these combinations could be calculated, of which only half were reliable; aspects of the callout data such as non-emergency responses and appliances arriving at a zone from somewhere other than their domicile station offset the distribution, and further data surrounding these issues was not available.

Fortunately, with the help of ArcGIS9.0, and the aid of an expert, these issues were alleviated with the use of Geographical Information Systems. Firstly every zone centre X-Y coordinate estimate that had previously been calculated was displaced to its nearest node on New Zealand's road network, and the distance of displacement in metres given in each case. ArcGis9.0 then calculated the travel time from the station to the average zone centre's closest node for every station/zone combination. ArcGis9.0 assumes a travel speed equal to that of the speed limit, slowing to allow for corners, and while this is most probably not the case in most urban areas (travelling as low as the speed limit for emergency response), it is consistent over every time estimate. By assuming a travel speed of 40 km/h for off-road travel, an accurate arrival time estimate can also be calculated for the zones that do not have callouts occurring within them. Finally, for all time estimates from a rural station, a penalty of five minutes has been imposed to represent the average time between when a callout is created and the first unit departs from the station, reflecting the volunteer nature of their operation.

The increased help to accuracy of time estimates should improve the quality of the allocation of equipment around the Transalpine station network, and the zones each allocated station is required to service.

2.4 Equipment Current Information/Allocation

Information on the current allocation of each type of equipment, its usage type and the fire classifications it is used for provides the model with 'test' items that are both passively-and aggressively-used. These equipment types have driven the development of our generic allocation model, and will form the basis for the analysis in this report.

3 Objective function motivation

Due to the different usages of a passive and aggressive piece of equipment, there are different objectives for the two equipment types.

3.1 Passive equipment

For the passively-used piece of equipment there is no time restriction on when it should arrive at an incident. However a piece of equipment being used at one incident is then unable to be sent to at any other incident. Therefore in order to reduce the amount of time during which a piece of equipment is unable to provide service to other incidents, it would be beneficial for the equipment to be placed as closely as possible to as many potential incidents as possible.

This is the motivation behind the first objective used to determine a solution for a passively used piece of equipment. For each of the zones within the Transalpine region, the difference in time between the closest station to this zone, and the closest station which has a piece of equipment allocated to it, is calculated and multiplied by the weighting for that respective zone. If the closest allocated station has more than one piece of equipment allocated to it then the result is divided by the number of pieces of equipment allocated. The objective aims to minimise the maximum of these values for each the zones, such that the maximal weighted difference is of the greatest concern.

The motivation behind the second passive objective is the same as the motivation behind the first. For every allocated station there is a collection of callouts that are considered "served" by that station. Using the start and close times for every one of these callouts we can calculate the amount of time each piece of equipment spent in "overlap" over the last five and a half year time horizon. This objective works to minimise the average of these values for all of the allocated stations as this is a better and more realistic proportion of time in overlap than the station with the maximum overlap, as in most instances heavily loaded zones would receive service from more than one station.

The third objective attempts to encapsulate the information lost within the first objective, and is represented as the sum product of all the response times and the weighting for every zone. This objective works to minimise the average case, while the primary objective works toward minimising the worst case.

3.2 Aggressive equipment

A time based objective for the NZFS is that 90% of the time two appliances arrive at the scene of an urban incident within ten minutes (twenty minutes for a rural incident). For an aggressively-used piece of equipment it would be beneficial for it to arrive at the scene of an incident as quickly as possible in order to maximise its ability to provide the emergency response it is designed for. The difference between the arrival time estimate for the closest allocated station and the time limit for each zone is multiplied by the weighting for that zone to give the value to be used in the objective. The primary objective is to minimise the maximum of these values for an allocated station.

The second objective for an aggressively-used piece of equipment is to minimise the sum of these values calculated in the above objective. This objective will be penalised if a zone is reached after the specified time limit, by adding to it a fixed value for every zone that is reached after the time limit. The third objective used to determine the performance of an aggressively-used piece of equipment is the same as the second objective used to assess a passively-used piece of equipment, i.e. to minimise the average proportion of time spent in overlap by a particular piece of equipment.

4 Heuristics and Search Procedure

4.1 PYRO

The software package PYRO (Planning Your Resource Organisation) is the result of this analysis. It provides a user interface allowing the NZFS the ability to access information surrounding their current and subsequent allocations of equipment. It also adapts to the addition of new items of equipment and presents the option of fixing already allocated equipment within the network of stations.

4.2 Heuristic

PYRO makes use of the Tabu Search meta-heuristic to seek a "good" allocation. Tabu Search was the preferred choice for a search heuristic due to its ability to handle a variety of stochastic solution spaces.

The following framework is used when deciding which of the currently unallocated stations is chosen to have a piece of equipment allocated to it.

Table 1:
Tabu Search Framework

Construction heuristic	When finding an improved current allocation, use the original allocation as an initial solution from which to make changes. When adding a new piece of equipment, use the current allocation with the additional piece allocated to an arbitrarily chosen currently unallocated station. When removing a piece of equipment, take the current allocation, and remove the piece of equipment from the station that services the lowest overall weighting.
Move	Swaps one piece of equipment between one station and another.
Tabu restrictions	An allocation that has been removed cannot be restored within a number of iterations determined by the prevailing Tabu list size
Dynamic Tabulist	Adjusts its length accordingly by increasing in size as it steers away from local optima, and decreasing in size as it finds new parts of the solution space of interest.
Tabu Criteria	For each possible equipment item swap, evaluates the primary and tertiary objectives and remembers swap indices if the resulting combined objective is less than the best found so far.
Aspiration Criteria	The tabu list can be overridden if the move that is inadmissible will generate a solution which is currently better than any previous one.
First Improvement	For moves that have met tabu criteria, the secondary objective value is first calculated. Accepts a swap and sets to current solution if the swap results in an overall objective value less than that of the current solution.
Stopping Criteria	Stop once the search has been run for a set number of iterations

4.2 Search Procedure

The following procedure is used to search through different allocations, and is the same logical argument called within the PYRO software implementation.

Step 1. Make the next swap, given by the relevant indices within each of the allocated and non-allocated station lists.

Step 2. Adjust time differences, i.e. for passive equipment this is the difference in time between the best service time and the response time for which it is serviced in. Re-allocate zones to stations for those which were serviced by the removed station (if a station that had equipment is left with none), for those zones which are serviced faster by the newly allocated station, adjust primary objective accordingly.

Step 3. Recalculate the tertiary objective value.

Step 4. If the combined scaled weighting of the new primary and tertiary objectives is less than the Best Value So Far (which is set arbitrarily high before the first swap), then check the tabu list for the station added to the allocation. If the station is not on the tabu list then check if this scaled weighting is less than the equivalent weighting of the Current Solution, if it is calculate the secondary objective and go to *Step 5a*. If the station is on the tabu list, then check to see if the station's scaled weighting meets the initial aspiration criteria, i.e. its primary and secondary scaled weighted objective value is less than the best solutions equivalent value. If so then calculate the secondary objective and go to *Step 5b*. Else, reverse the swap remembering the indices as the best swap so far and return to *Step 1*.

Step 5a. Now that all objectives values are known for the solution an overall scaled weighted objective value is known. If this value is less than the overall objective value for the current solution, then this a first improvement go to *Step 6*. Else reverse the swap remembering the indices as the best swap so far and return to *Step 1*.

Step 5b. With the now known overall scaled weighted objective value; check if the new overall objective value is less than the Best Solutions overall objective value, and if so the new Best Solution is found. Set this as Current Solution and go to *Step 7*. Else, reverse the swap remembering the indices as the best swap so far and return to *Step 1*.

Step 6. Change this solution to the Current Solution, resize list size if necessary and check if overall scaled objective is less than best known, if so go to *Step 7*, else reset test solution and go to next move

Step 7. Set Current Solution to the Best Solution and go to next move

5 Average Case Scenario Setting

A fundamental assumption of the proposed model is that it is based on callout information, that is, by using average case scenarios. While there may be a case where in the future, a major scale bushfire will occur within an uninhabited zone where no incident has occurred over the five and half year time frame; the probability this could happen is taken to be negligible. Furthermore, as with the allocation of any scarce resource, the primary focus is to maximise its utilisation, and in the case of aggressively used equipment maximising this utilisation potentially acts as a life-saving objective.

6 Recommendations and Analysis

6.1 Allocation of currently owned equipment

The three pieces of equipment used to develop the model are the PPV Fans, TICs, and GDUs. To calculate the level of service, the weightings placed on each of the three objectives needed to be determined. To do this the model was run for two hundred iterations using a variety of weighting combinations. These results are summarised in Table 2.

The results for the two types of aggressively-used equipment indicate that the weightings on the objectives which give the most consistent results in terms of the number of iterations to find the best solution and the improvement found in the objective value are 0.2 on the primary, 0.7 on the secondary and 0.1 on the tertiary. For each of the three objectives used to determine the level of service for a passively used piece of equipment the weightings were 0.5, 0.4 and 0.1 respectively.

The results show that by reallocating the PPV fans, the normalised weighted objective value decreases from 1 to 0.5709. This solution was found after 21 iterations

Table 2.

	PPV Fans			TICs			GDUs		
	PPV – scaled overall objective value	iteration number, solution was found	number of equipment placements that are different from the status quo allocation	TIC – scaled overall objective value	iteration number, solution was found	number of equipment placements that are different from the status quo allocation	GDU – scaled overall objective value	iteration number, solution was found	number of equipment placements that are different from the status quo allocation
status quo	1		0	1		0	1		0
Prim = 0.2 second = 0.7	0.743	174	13	0.2226	25	5	0.1922	9	2
Prim = 0.3 second = 0.6	0.6749	40	10	0.5139	149	5	0.1643	124	4
Prim = 0.4 second = 0.5	0.6559	20	9	0.313	182	5	0.6559	20	2
Prim = 0.5 second = 0.4	0.5709	21	8	0.3103	130	5	0.1564	7	3
Prim = 0.6 second = 0.3	0.4559	60	8	0.4386	177	6	0.1208	115	3
Prim = 0.7 second = 0.2	0.3424	74	10	0.3266	5	2	0.1094	61	3

Table 3. PPV Fans

ADDITIONAL AMOUNT	ALL FIXED	ITERATION	ONE VARYING	ITERATION	TWO VARYING	ITERATION	THREE VARYING	ITERATION
1	0.87228121	6	0.623095125	37	0.611054783	15	0.603666461	61
2	0.591569007	10	0.580405405	43	0.545261668	17	0.553042349	13
3	0.535793434	70	0.511302862	73	0.519872438	17	0.513578074	19
4	0.50062781	42	0.496019512	34	0.488772019	20	0.489878834	26
5	0.468518347	32	0.468942625	25	0.466571	22	0.473964366	15

Table 4. TICs

ADDITIONAL AMOUNT	ALL FIXED	ITERATION	ONE VARYING	ITERATION	TWO VARYING	ITERATION	THREE VARYING	ITERATION
1	0.29067139	6	0.26523473	69	0.279178283	7	0.236287745	12
2	0.218515223	8	0.259440105	7	0.198046553	83	0.210917848	73
3	0.193167188	15	0.19263304	94	0.145468296	17	0.186691633	74
4	0.146663742	22	0.137364278	17	0.167055293	83	0.134595418	23
5	0.129271841	19	0.129775673	17	0.121113878	76	0.130037865	78

Table 5. GDUs

ADDITIONAL AMOUNT	ALL FIXED	ITERATION	ONE VARYING	ITERATION	TWO VARYING	ITERATION	THREE VARYING	ITERATION
1	0.164546614	3	0.164850105	91	0.157949816	58	0.157322553	5
2	0.137065016	84	0.122180821	96	0.12190355	56	0.128595639	20
3	0.110584338	82	0.108902364	37	0.105866696	90	0.118584697	71
4	0.099588182	43	0.097374279	49	0.101003186	16	0.101003186	16
5	0.088223742	43	0.091400686	16	0.091126968	19	0.092160888	21

and the change from the original allocation, results in the moving of 8 of the 18 items away from their domicile station.

By reallocating the TICs, the weighted objective value changes by 67.34%, from 1 to 0.3266. The TIC is an aggressively-used type of equipment, and the objective weightings used for this type correspond to ones identified in Table 2. These are a weighting of 0.7 for the primary objective, 0.2 for the secondary objective, and 0.1 for the tertiary objective. The solution was found after only five iterations, and the stations allocated equipment in the new solution differ from the original station by only two.

The GDUs are also an aggressively used type of equipment. The weighted improvement of the objective function is 89.16%, form 1 to 0.1094. However this solution took a lot longer to find, occurring after 61 iterations. Again the change from the original allocation is small, with three pieces of equipment moving from their current allocation.

6.2 Allocation of further equipment purchases

Due to the nature of the New Zealand Fire Service operation, the current allocation may not change, even though it has been shown above that a reallocation will result in a much improved objective value. This is because it can often crate friction moving a piece of equipment from one station to another, due to the way personnel feel "attached" to the equipment on their appliances. Therefore, the following section, which allocates additional pieces of equipment, is based on the original allocation given to us by the NZFS, not the improved allocation calculated above. The process that is more likely to happen within the NZFS is that additional pieces of equipment are requested for, and the NZFS wish to know where on top of the current allocation they can be placed in order to maximise their level of service.

Tables 3, 4 and 5 on the previous page, outline the results found from the running of the model with the specified objective weightings found in Section 6.1, with a varying number of additional pieces of equipment. For each level of new equipment allowed, the heuristic is run allowing varying numbers of equipment to move from its position in the current allocation. This process is done to find out what will happen if the current allocation is allowed to change. These Tables all indicate an improvement in the objective value should an additional piece of equipment become available, which is to be expected.

The second column in each of the Tables 3, 4 and 5 indicates the solution if the current allocation is not allowed to change. Reading across the columns it shows that by increasing the number of equipment that is allowed to change, the objective function generally improves. For each of these runs, the heuristic was run for a fixed length of 100 iterations; this was shown to be insufficient when the number of equipment allowed to move increased from two to three. When the complexity of the model increased, with an increasing number of pieces of equipment allowed to move within the solution, the length of time required in order to find the best solution was found to increase. Due to this, the heuristic did not always find an improved solution as an additional piece of equipment was allowed to move. In order to ensure an improvement, the heuristic should be run longer for every additional piece of equipment allowed to move within the solution.

Reading the results in Tables 3, 4 and 5, the increase for every additional piece of equipment allowed to move is minimal. This finding backs up the stance held by the

NZFS, in which they are reluctant to remove equipment from stations that have become skilled in using that type of equipment.

7 Limitations of PYRO and further research

There are a number of assumptions made throughout the model that place limitations on what the PYRO software is capable of and on the results produced. It is assumed that the likelihood of an incident occurring is consistent throughout the day, and that the times calculated between the stations and each nodes are reflective of a response time any time of the day. Although this is not true, the viability of collecting data on the distribution of fires throughout the day, and as such the differing response times to these fires, was deemed to be a level of detail too intense to relate to the recommendations given by the results.

PYRO also assumes that every piece of equipment can be identified by the fire classifications that it is used to serve. It also assumes that each piece of equipment is either passively used or aggressively used, and never both. Should a new piece of equipment become available to the NZFS which can be used either passively or aggressively dependent on the fire classification it serves, then an adjustment to the model would be needed to provide an accurate re-allocation. There is also no allowance within our model for the addition of new fire stations to the region. Our model is able to cope within additional appliances relatively easily, although not on a user based level, however an additional station would require a new set of station/zone response time estimates.

The same problem will occur should new fire service zones be created. Due to the application within the project of ArcGIS9.0 to calculate the station zone arrival times, any zone change of this nature would require a technical update to the model. There remains a possibility of having the program extended to allow automatic updating of the callout data for the NZFS database. We see this extension as an internal IT project. The model works as a stand alone package now, but having it connected to the NZFS database would allow the data set used to reflect changes in the demographics of fire incidents throughout the region.

8 Acknowledgements

The authors of this paper would like to thank:

- Dr. Clive Sabel of the University of Canterbury for the his help and expertise in operating ArcGIS9.0
- Paul Burns, Steve Kennedy, and Anton Daish of the NZFS
- Simon Quirke, Team Leader at NZFS National Headquarters
- Dr. Ross James and Dr. John Giffin of the University of Canterbury

Planning a New Model for Dynamically Reallocating Ambulances

David P. Richards
Department of Engineering Science
University of Auckland
New Zealand
dric038@ec.auckland.ac.nz

Abstract

The Melbourne Ambulance Service is currently using an ambulance simulation system called Siren, to assist in their planning. One concern raised, was that in the event that a hole appeared in their coverage plan, the Melbourne Ambulance Service had limited tools to assist in the effective relocation of ambulances in real time. A dynamic model was needed to be developed that was able to relocate idle ambulances, leaving no areas unprotected.

The aim of this paper was to develop a model that would calculate good locations for a fixed number of ambulances, by using the details of the ambulance system operating procedure to determine where ambulances should be based, in order to maximise the percentage of calls responded to within their target time. The optimisation model was designed to work within a dynamic model.

To evaluate the configurations of bases that have an ambulance, this paper used the number of calls each base was closest to, and the associated response times of these calls. Queuing theory, which treats the ambulances as servers and the calls as customers, was also used to account for the fact that the closest base may not always be able to respond due to already attending to another call. A model was developed using these various methods and the different evaluations were compared using Siren simulations. It was found that the queuing theory model produced the best response times for the call sets analysed.

1 Introduction

Ambulance organisations are always under pressure to respond faster to emergency 111 calls. To try to improve ambulance response times, the locations of ambulance bases are optimised to provide the best coverage of calls. These coverage plans can start to perform poorly when ambulances become busy. When ambulances are dispatched to calls, we often find that holes occur; where the ambulances that had been planned to serve calls in an area are all busy serving other calls. This means that if a call occurs in this area, the first choice ambulance would not be able to respond to the call and therefore, the call may not responded within the target time. To remedy this, a 'move-up' operation is needed to be performed in which, the available ambulances are redistributed to cover the hole.

The Melbourne Ambulance Service is currently using an ambulance simulation system called Siren, to assist in their planning. One concern raised, was that in the event that a hole appeared in their coverage plan, the Melbourne Ambulance Service had limited tools to assist in the effective relocation of ambulances in real time.

For this reason we aimed to develop an optimisation model that will work inside the dynamic model, to calculate good locations for a given number of ambulances. This optimisation model will use the details of how an ambulance system operates to determine where ambulances should be based to satisfy the largest percentage of call response targets. Therefore, in order to develop this model we first need an understanding of how ambulance systems operate.

The main performance measure of an ambulance system is the percentage of calls that are responded to within a certain target time. The response time for a call, is the elapsed time from when the call is received, until the time an ambulance first arrives at the scene. The Melbourne Ambulance Service response process is shown in Figure 1.1, with the response time illustrated by bold red arrows.

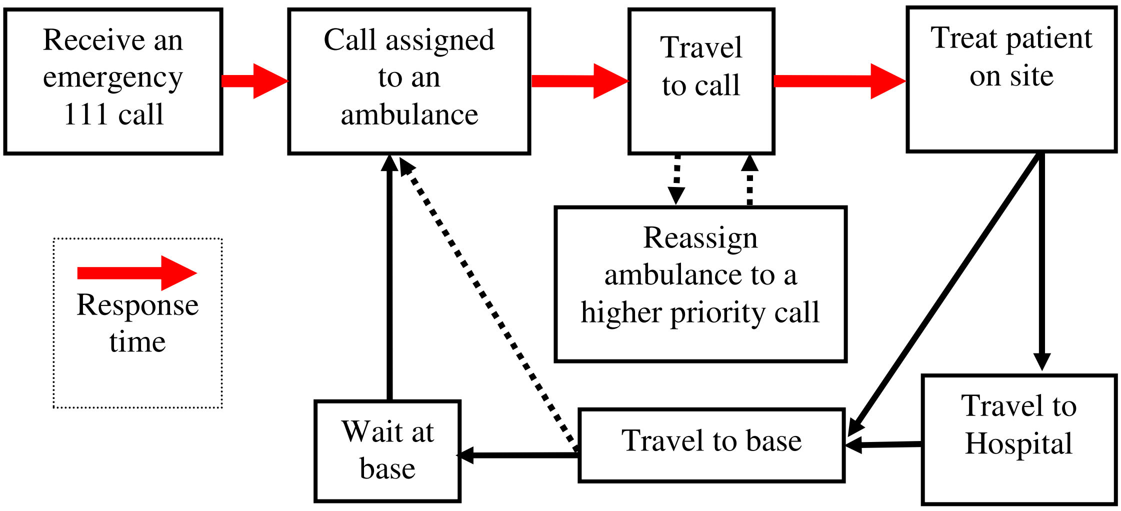


Figure 1.1 Flow chart of the operational process for ambulances, adapted from Kirkpatrick [1]

The problem of relocating ambulances to a new base has several factors that increase the complexity of the problem. The location of an ambulance when it responds to a call is difficult to predict in advance and the exact location of a call and its type is not known prior to being received. When they are sent to respond to a call, ambulances are not necessarily at the base they are assigned to. Ambulances can be redirected to a higher priority call or could be sent to respond to a call on their way back to base. These two situations are shown in Figure 1.1 by the dotted arrows.

3 The Optimisation Model

The optimisation model's role in the dynamic model is to calculate good base locations for a given fixed number of ambulances. To calculate good locations of a set of ambulances, we first need to be able to evaluate the effectiveness of an ambulance at each base. When an emergency call is received the closest available suitable ambulance is sent to respond to the call. The closest ambulance is the ambulance that can travel from its base to the location of the call in a shorter time then any other available ambulance. An ambulance's base location can then be evaluated by its response times to all the calls it is the closest base to.

This model uses many assumptions to make it easier and more viable to construct. The model uses historical calls to determine where ambulances should be relocated; this is due to the fact that it is impossible to calculate in advance where an emergency will occur, at what time and how serious the emergency will be. The other assumptions used in this model are:

- Each ambulance travels from its assigned base to the call. The fact that ambulances can be reassigned to higher priority calls or can be assigned to a call while travelling back to base is not taken into account;
- Each call is served by a vehicle that responds from the closest active base (defined below);
- Each base has 1 ambulance vehicle;
- Each ambulance is of the same type, with the same equipment, with staff at the same skill level, and that these skills and equipment are able to handle any type of emergency.

The first step of this model is to determine which calls an ambulance is closest to.

3.1 Determining closest calls

The closest calls to a base are dependent on the status of the surrounding bases. The meaning of a base state is whether the base has an idle ambulance or not. A base state is then said to be either active or inactive, respectively. If the base states of all the surrounding bases are known then the closest calls to the base can be calculated.

The local neighbours of a base can easily be visualised by the partitioning of the area into sets of all points that are closest to the same base. Any point in the base's area is closer to this base then any other dot. This is an approximation of the area of closest calls, where the set of closest points to a base is actually dependent on both the road network and time.

To calculate the calls closest to base 1 in Figure 3.1, we therefore need to know the states of at least, all bases in the local neighbourhood which include the bases 2, 3, 4, 5, 6, 7 and 8. We can see from the Figure 3.1 that these bases will determine the shape and size of the area for base 1, and thus will determine the calls that are closest to base 1.

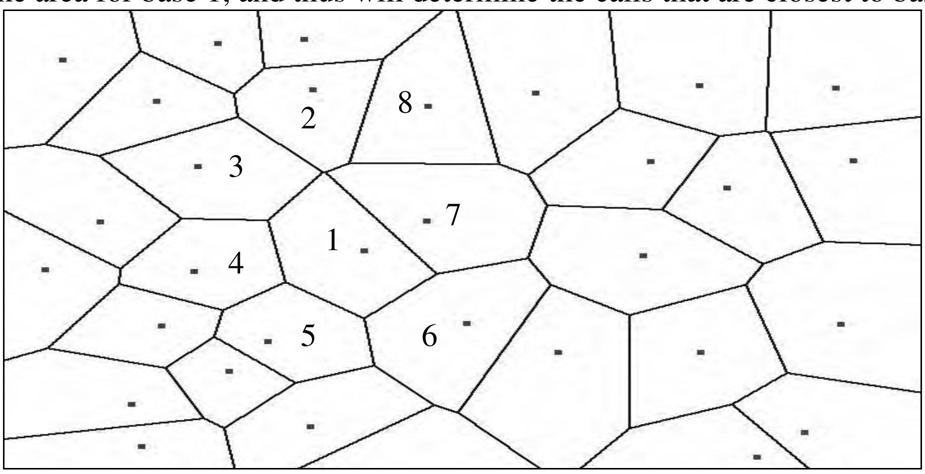


Figure 3.1 Diagram of bases assuming all bases are active

3.2 Call based partitioning

Our model works with actual historic calls, which are allocated to a base rather then allocating a base an area. Depending on the time of the day, different bases might be better suited to respond to an area. This difference is created mainly by the variation in traffic throughout the day. Some calls can be reached faster from different bases by the route it would take to get to the call. Therefore call partitioning is more accurate then area partitioning. To calculate all the calls that a base is closest to, we need to know the states of the neighbouring bases for the base in question.

3.3 Call response scores

The call response scores can be calculated for a base by summing the response scores for each of the calls the base is closest too, given that we know what neighbours have an available ambulance. To evaluate the response score of a call, the call details and the estimated response time is needed. The call will have a specific code (1, 2 or 3) depending on what type of emergency it is. Each code has a target response time. The response score is then calculated by taking the target response time and the estimated response time. The response score, as a function of the code and the estimated response time, is shown in Figure 3.2. The severity of codes are ranked from 1 to 3; this is seen in Figure 3.2 as the response times for code 1 calls having higher scores then code 2 and 3 calls, for fast response times.

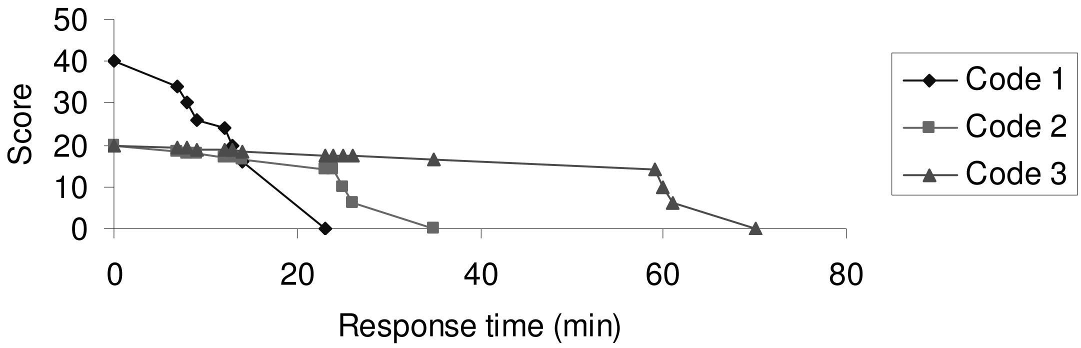


Figure 3.2 Call response target score

4 Adjacent Bases

4.1 Target base

To calculate the calls that a base is closest to we need to know the status (active or inactive) of each of the adjacent bases. The ‘target base’ is the term we use for the base we calculate the closest calls for.

4.2 Adjacent Bases

Given a target base, that we assume is active, and assuming for now that all other bases are also active, we can define an adjacent base of the target to be a base whose status, as active (having an available ambulance) or inactive, affects the number of calls that the target base is closest to. Geographically, adjacent bases tend to be the surrounding bases of the target base in 2-d space.

In Figure 4.1, it is also observed that by making an adjacent base inactive a previously non-adjacent base becomes adjacent.

Making different combinations of adjacent bases inactive can create different adjacent base configurations, for the target base. By looking at different sets of adjacent bases we can see how the effect of different combinations of active bases will affect the number of calls the target base will be assigned to respond to. Configurations of adjacent bases can be calculated using a base usage tree.

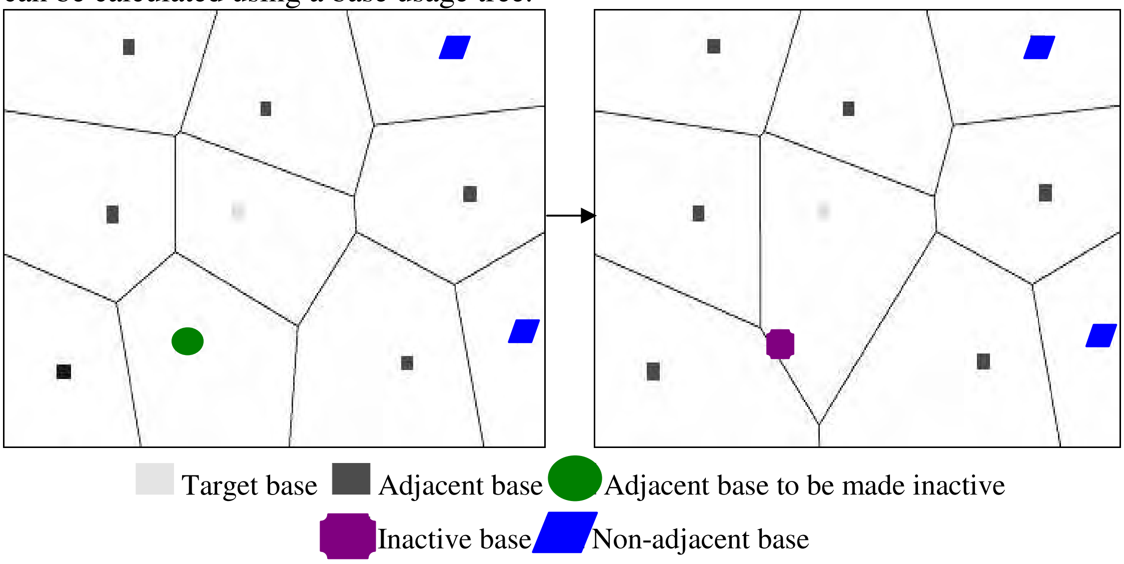


Figure 4.1 The effect of making an adjacent base inactive

5 Base Usage Tree

A Base Usage Tree, is a tree containing all historical calls in a given time period and represents the closest bases to each of the individual calls. The Base Usage Tree is constructed by first calculating the ith closest bases to each call, where i = 1 to m, and m is the maximum depth of the tree. The calls are then grouped together into nodes so that all the calls with the same 1st closest base start at the same node at the beginning of the tree, at a tree depth of 1. The calls with the same 1st and 2nd closest base are then grouped in the next nodes at a tree depth of two (or the 2nd node in a branch). This is carried on until the maximum depth of the tree is reached. At the start of the tree there is an artificial node called the ‘root node’, which we say is closest to all calls. Each node has associated with it a base, the number of calls counted by this base for this node, and the total response score these calls generate when responded to, from this base. For example in Figure 5.1, node 1 has an associated base 1, the number of calls counted by base 1 for this node is 10 and the associated total response scores for these 10 calls if responded to from base 1, is 285.

The Base Usage Tree in Figure 5.1, can be interpreted, so that the information that can be obtained from the tree is as follows; of the 25 historical calls used to construct the tree, 10 calls are closest to base 1, 7 to base 2 and 8 to base 3. Of the 10 calls that are closest to base 1, 7 calls are 2nd closest to base 2 and 3 are 2nd closest to base 3, and so on. The Base Usage Tree can then be easily used to calculate the calls that are closest to each base given that we know the states of each base.

In this paper we use certain terms to refer to parts of the Base Usage Tree. These terms are children, parents, and peers. For the tree in Figure 5.1, we can say that node 2 is a child of node 1, therefore node 1 is a parent of node 2, and node 2 and 4 are peers. Peer nodes always occur at the same tree depth in the tree, node 2 and node 4 are both at the same tree depth of 2, and have the same parent, node 1.

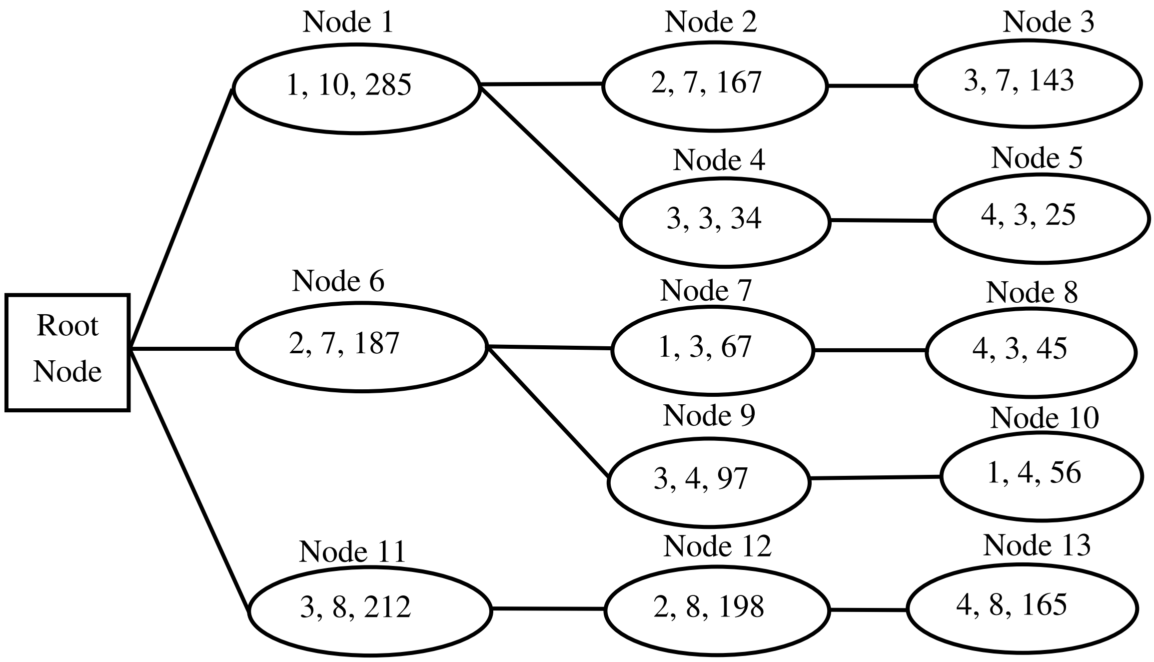


Figure 5.1 A Simple Base Usage Tree

5.1 Using the Base Usage Tree to Calculate the Adjacent Base

The reason for including the Base Usage Tree in the optimisation model is so that we have a method of easily calculating adjacent bases. In the tree, if a base is made inactive then any node that contains this base is ignored. For example in Figure 5.1, if base 2 were inactive, then we would ignore nodes 2, 6 and 12. The result of this would be that now, base 1 is 1st closest to 3 more calls and base 3 is 1st closest to 4 more calls. Also, nodes 3 and 13 are now treated as a child of nodes 1 and 11 respectively.

We can therefore use the Base Usage Tree to calculate which bases are adjacent to each target base for a given input of base states. This can be done, because an adjacent base in the tree is an active parent of the target base. For example, assuming that all bases are active, in Figure 5.1, if the target base was base 2 then base 1 and base 3 would be adjacent bases.

We can use this fact in an algorithm to calculate all the adjacent bases for each target base, as seen in Figure 5.2. Any base, which has a child base that is the target base, is adjacent to this target base. An adjacent base is the first active parent of the target base; if there are bases that are all inactive in between the two bases then as described above, these nodes are ignored. We can see this in Figure 5.1, if base 2 is inactive and the target base is base 3, then base 1 becomes an adjacent base to base 3. The algorithm that uses the Base Usage Tree to calculate the configurations of adjacent bases is shown in Figure 5.2. This algorithm takes in a set of initial states of the bases in the problem, and sets as 'active' those currently 'undecided' bases that are adjacent to our target base, where our test for adjacent bases now assumes that inactive bases have been effectively removed from the problem. The base states are stored in an array called base_states. The value for the ith value in base_states represents the state of base i. States are represented as follows: 1 active base;

0 Non-adjacent base, therefore its state is left undecided;

-1 inactive base.

In the tree we have to treat non-adjacent base sates as both active and inactive as the base state is not being constrained in any way.

Some other terms that are used in this algorithm are the last_active_node, First_child and Next_peer. The last_active_node is the last node that was stepped over that contained a base that was active. The First_child of a node is the first child node that appears at the top of the other child nodes, for example in Figure 5.1 the First_child of node 1 is node 2. The Next_peer is the next peer node in the tree, and again looking at Figure 5.1, it ca be seen that the Next_peer of node 1 is node 6.

The algorithm starts at the root node, with the target base set to the first base, and walks down the Base Usage Tree from the first child to the next first child. When a node is reached that does not have an inactive base (i.e. is active or undecided), the node is stored as the last active node. The algorithm then walks to the next first child and checks to see if this node's base is the 'target base'. If it is, then the last active node's base state is changed to active, as this base is an adjacent base. If it is not the 'target base' but the base state of the current node is 'undecided' then the algorithm sets the last active node to be the current node and walks to the next first child. If the node's base state was active then we walk to the next peer of the current node, and carry on the search.

```
Sub Search(target_base, base_states, search_node=0, last_actvie_node = NULL) {
  node = search_node
  repeat
            If the nodes base is not active then
                      If node has a child
                         Search(target_base, base_states, node->First_child, last_actvie_node)
            Else If last_active_node = NULL
                     If base_states[node] =0  //undecided and node has a child
                         Search(target_base, base_states, First_child of  node, node)
           Else
                    If node = target_base
                        base_states[node] = 1  //active
                    Else If base_states[node] = 0  //undecided and node has a child
                        Search(target_base, base_states, First_child of node, node) //recursive
            node = next_peer of  node
  while node != NULL or last_active_node != NULL and base_states[last_active_node] >0
}
```

Figure 5.2 Algorithm for Finding Adjacent Bases

6 Building the Optimisation Model

This model takes in a set of configurations for the active status of the neighbours of each base. It then constructs columns for each of these configurations so that the model can be formulated as an Integer program (IP). The columns of the base sates are represented with the values ‘on’ if the base is active, ‘off’ if the base is inactive and ‘?’ if the base is non-adjacent. These columns can be seen in Figure 6.1, displayed vertically. Each column configuration will also have a corresponding score, as discussed later.

This IP has two sets of binary variables, the chosen configuration and the base states. The chosen configuration variable has the value 1 if the configuration is in the solution, otherwise it has the value zero. The base state variable has the value 1 if the base is active, and 0 if the base is inactive. The configurations constrain the base states so that if base i is set to be ‘on’ in the configuration then a lower bound of 1 is placed on the sate of base i. If base i is set to be ‘off’ in the configuration then an upper bound of 0 is placed on the sate of base i. Bases whose state is undecided (?) have a lower bound of 0 and a upper bound of 1. These constraints can be seen in the bottom part of the IP in Figure 6.1. These constraints ensure that base states across configurations are constant. The IP also has a generalised upper bound constraint, which can be seen in the middle of the IP in Figure 6.1. This constraint means the model can only select one configuration, or column, of adjacent bases for each base so that some objective (discussed in Section 7) is maximised. By selecting a configuration, active statuses of certain bases are defined.

Base 1 Options					Base 2 Options					Base 3 Options									
x10	x11	x12	x13	x14	x20	x21	x22	x23	x24	x30	x31	x32	x33	x34	x1	x2	x3		
OFF	ON	ON	ON	ON	?	ON	ON	OFF	OFF	?	ON	ON	OFF	OFF				base 1	
?	ON	ON	OFF	OFF	OFF	ON	ON	ON	ON	?	ON	OFF	ON	OFF				base 2	
?	ON	OFF	ON	OFF	?	ON	OFF	ON	OFF	OFF	ON	ON	ON	ON				base 3	
1	1	1	1	1														=	1
					1	1	1	1	1									=	1
										1	1	1	1	1				=	1
0	1	1	1	1														<=	x1
0	1	1	0	0														<=	x2
0	1	0	1	0														<=	x3
0	1	1	1	1														>=	x1
1	1	1	0	0														>=	x2
1	1	0	1	0														>=	x3
					0	1	1	0	0									<=	x1
					0	1	1	1	1									<=	x2
					0	1	0	1	0									<=	x3
					1	1	1	0	0									>=	x1
					0	1	1	1	1									>=	x2
					1	1	0	1	0									>=	x3
										0	1	1	0	0				<=	x1
										0	1	1	1	1				<=	x2
										0	1	0	1	0				<=	x3
										1	1	1	0	0				>=	x1
										0	1	1	1	1				>=	x2
										1	1	0	1	0				>=	x3

Figure 6.1 Formulation of the model for three bases

7 Evaluating a Base State Scenario

Once we have all the possible configurations for some target base, we need to rank each scenario configuration in order to calculate the optimal solution of active bases.

7.1 Penalising Overloaded Bases

The likelihood of an ambulance at a base responding to a call within the required time is directly related to the number of calls the ambulance is expected to respond to. Bases that are assigned a large number of calls are unlikely to be able to respond to all of them. If this is the case then an ambulance at another base will have to respond to the call instead. Having the responding ambulance further away from the call than the closest ambulance is obviously going to make it less likely to respond to the call within the required time. Therefore one way of trying to improve the response times of ambulances is to punish an adjacent base configuration for a base if the configuration assigns too many calls (through the closest base partitioning) to the base. The objective function of this method works out the number of calls that should be assigned to a base if each base has an equal number of calls to respond to. Then any configuration that assigns more calls than this average value is given a negative score to punish it.

7.2 Time Taken to Respond to the Calls Allocated

Part of the benefit of using the Base Usage Tree is that not only do we know how many calls are allocated to each base; we also know exactly which calls are allocated to the base. From knowing which calls each base responds to we can use the details of the call to calculate how long it would likely take to respond to this call. The details of the call describe the time of day of the call and what priority the call is. The priority code of the call tells us how fast the ambulance will travel i.e. with or without lights and sirens. The code also gives us information on what the desired time to respond to the call is. With the time of the call and information stored in Siren, we can calculate how busy the roads from the base to the call are and thus how long it will take for the ambulance to get to the call. Knowing how long the ambulance will take to respond to the call and the calls priority code we can give the ambulance a score. Summing the scores of all calls allocated to a base for a given scenario configuration will provides us with a score for that configuration. We can then optimise the allocation of ambulances to bases by maximising the sum of the scores of all active bases.

7.3 Busy Probability Factor

The previous methods, only look at the response times from the base closest to the call. It does not take into account the response time if this closest base is busy. To calculate the probability a base is busy we use queuing theory. The score for each configuration can thus be calculated by taking the scores for each of the calls assigned to the target base and multiplying this value by the probability the base is not busy. The scores from the 2nd closest base are also taken into account. For each of the calls assigned to the target base we also add to the response score given by the 2nd closest base responding to the call multiplied by the fact the 1st closest base is busy, multiplied by the mean system-wide probability the 2nd closest base is not busy. The 2nd closest bases to each of the calls are those active bases defined as adjacent bases in the configuration under consideration.

8 Results

The results generated for this paper have been from randomly generated call sets and thus do not reflect the actual performance of the Melbourne ambulance service. However, the response performance differences reported in the results are likely to be similar to what would be seen using real call sets.All the results in this section were produced using the simulation inside Siren. The Base Usage Tree used in these results was constructed using a randomly generated call set based on historical calls. The size of the call set is 64633 calls, and these calls are for the elapsed time of 3 months or 92 days.

8.1 Comparison of objective functions

From the simulations it was seen that evaluating base locations using queuing techniques is more effective then just distributing the calls as evenly as possible across all active bases. The queuing evaluation not only responds to calls on average faster, it also responds to more calls. The queuing evaluation model responded to 3054 calls while the overloading evaluation model only responds to 2975. Evaluating base locations using queuing techniques was also more effective than just maximising the response scores. The queuing evaluation model's average response time is 31.11 minutes, 1.18 minutes faster then the overloading evaluation model's response time of 32.28 minutes. For these scenarios simulated, evaluating local neighbourhoods of active bases using queuing techniques is the best method. This is likely due to the fact that this method recognises that bases are occasionally busy when an emergency call that they would ideally respond to is received.

9 Conclusions

The aim of this paper was to develop an optimisation model that could be used in a dynamic model to calculate the optimal location for a fixed number of ambulances. An optimisation model was created in Siren, by creating object-orientated functions that work inside the already existing Base Usage Tree class. The model developed provides a good foundation and starting point for calculating the optimal location for a fixed number of ambulances. There are a lot of assumptions in this model and it would be desirable to remove these assumptions so that the results produced by this model were more credible.

From the results it was observed that for the scenarios simulated, the best method to evaluate active base configurations was the queuing theory objective function. This method was most probably seen as the best method because not only does it take into account the call loads at a base and the response scores, it also estimates the probability that each base is busy and if a base is busy, what response scores are generated from the next closest ambulances. The current model appears to provide good results, but comparisons to other preexisting models would need to be done, before the value of the models results could truly be evaluated.

10 Acknowledgements

I would like to thank Dr Andrew Mason for his advice and help in providing Siren code and adding new features to Siren for my work.

11 References

1. Kirkpatrick, S. Better Base Locations for the Melbourne Ambulance Services, University of Auckland School of Engineering, Auckland, New Zealand, 2004

2. Budge, S. Approximating Vehicle Dispatch Probabilities for Emergency Service Systems, University of Alberta School of Business, Edmonton, Alberta, Canada, Jan 2005

Routing Trains Through Railway Junctions: Dual Formulation of A Set Packing Model

Richard Lusby

Department of Engineering Science

The University of Auckland

rlus005@ec.auckland.ac.nz

Abstract.

The problem of routing trains through a railway junction is an integral part of railway operations. Given the track configuration of the junction as well as a proposed timetable for the trains, the problem involves the task of assigning each train a particular route through the junction ensuring that at most one train occupies a track segment in any time period and also ensuring that further restrictions based on safety considerations are satisfied. Previous models in the literature have been based on node-packing formulations. We present a novel formulation as a set packing problem. The primal set packing problem has a very large number of constraints and relatively few variables which can be generated through a column generation procedure. However, the corresponding dual problem has a small number of constraints and many variables, most of which we expect will be zero at optimality. We implement the dual formulation with primal column generation, which corresponds to the addition of dual constraints Implementing it this way allows us to solve the more computationally attractive dual problem. We believe that significant improvements can be made to solution time and quality using this approach.

1. Introduction

The problem of routing trains through a railway junction is an integral part of railway operations. At large, busy junctions track capacity is a sparse resource. As a result of this, one can almost guarantee that at certain times of the day many trains will compete for the same track segment. Hence finding an assignment of trains to routes that ensures maximum utilization of the junction without compromising the safety requirements is no easy task.

This scheduling problem arises at a number of levels in the planning process of a railway company. Firstly, it occurs at the strategic planning level of determining the future capacity of a particular junction as a result of infrastructural changes to the track configuration. Secondly, it occurs at the tactical level of timetable generation to validate proposed timetables. Finally, it occurs on a day-to-day operational basis when predetermined train routes need to be adjusted due to unforeseen disruptions such as late train arrival, track maintenance or even accidents, which make the published timetable no longer feasible. This paper focuses on the second instance of the problem outlined above- timetable feasibility.

Due to the interconnected nature of a railway junction train routes are highly interdependent. Changes to one part of the timetable will inevitably cause conflicts in other parts. Effective and efficient solution approaches are vital. However, despite its noticeable importance this particular problem has received somewhat limited attention in the literature.

Earlier approaches by Zwaneveld et al. (1996) and Zwaneveld et al. (2001) have been based on node packing formulations. The key idea behind each of these is that every node represents a possible route for a particular train through the junction. Any two nodes that are incompatible, in other words, any two train routes that cannot be simultaneously assigned without conflict are connected by an edge. A maximal node packing is then one that represents the maximum number of trains that can be successfully and safely routed through the junction. This type of approach was first proposed by Zwaneveld et al. (1996) and combines extensive pre-processing, valid inequalities, and a branch-and-cut approach in finding the optimal solution. It was later revised to a weighted node packing formulation by Zwaneveld et al. (2001) as it was unable to solve routing problems for two of the larger stations in the Netherlands.

While the previous approach is effective in the sense that it can accurately represent any constraint regarding train route compatibility, the resulting conflict graph usually has many nodes. The ensuing many constraints and the weakness of the linear programming relaxation problem are contributing factors as to why only relatively small instances of the problem can be solved to optimality quickly.

The purpose of this paper is to present the novel approach of formulating this problem as a set packing problem, and in particular to focus on the corresponding dual of this formulation as it has several nice properties that one can take advantage of when solving the problem. We will also discuss our solution approach in which we implement primal column generation as dual constraint addition.

1.1 Paper Outline

This paper is organized as follows. In section 2 we describe the problem in more detail. Section 3 introduces some notation as well as both the primal and dual formulations. In section 4 we discuss our primal/dual solution approach and in particular the innovative use of primal column generation as dual constraint addition. A detailed description of the entire solution approach is also given.

2. Problem Definition

Throughout this paper the word junction refers to the network of track segments defining a railway station as well as an outer neighbourhood. It thus consists of a number of platform track segments (those that are adjacent to platforms) as well as normal track segments. Trains may enter the junction via a number of *entering points* and depart the junction via several *leaving points*. It is not uncommon for a leaving point to act as an entering point and vice versa. Figure 1 below illustrates the characteristics of a junction.

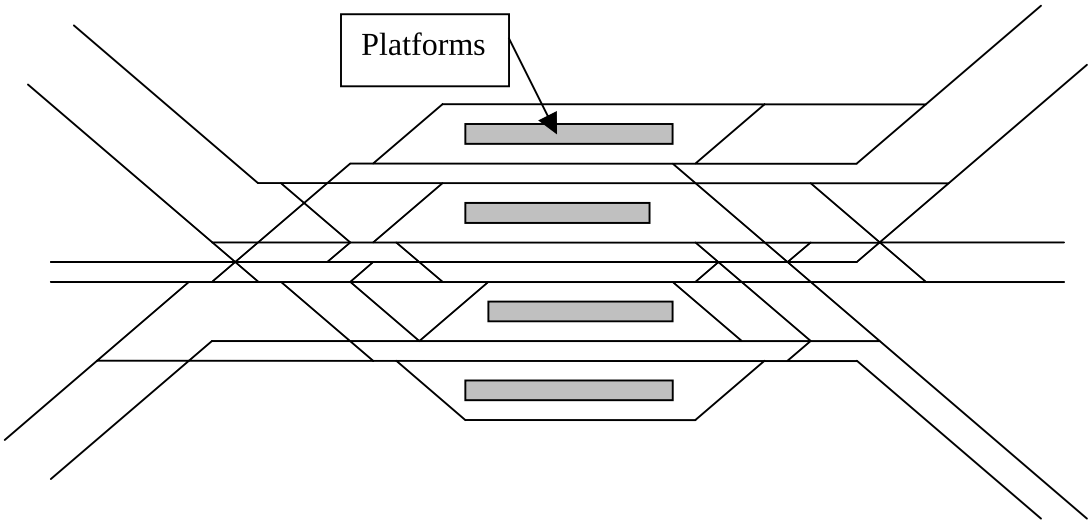


Fig 1. A possible junction.

The safety requirements of most European railway systems stipulate a *route locking and sectional release system*. Essentially this means that when a train arrives at an entering point to the junction, it claims all the track segments it will use in reaching its designated platform (*inbound route*). This prevents any other train from using any track segment in the locked inbound route as no track segment can be claimed by more than one train at the same time. On traversing its inbound route, the train will successively release each of the track segments comprising the route (after a short buffer time has elapsed), allowing them to be claimed by other trains. An identical procedure holds when a train departs the junction on an *outbound route*.

The problem of routing trains through a junction can therefore be formally defined as follows. Given the track configuration of the junction as well as a proposed timetable (i.e. the respective arrival and departure times for each of the trains) what is the maximum number of trains that can be assigned a route through the junction ensuring that no track segment is occupied by more than one train at any given time as well as enforcing the safety requirement outlined above. Other objectives are of course possible; however we have chosen this one for explanatory purposes.

3. The Set Packing Model & its Dual

In this section we give our set packing mathematical formulation of the above train routing problem as well as its dual. Set packing problems typically appear in scheduling applications where we need to assign elements to sets under the strict requirement that no element is allowed to be contained in more than one set. The problem at hand is certainly such a case, as we shall see now:

As was mentioned earlier, a railway junction is divided into a number of track segments on which there can be at most one train at any given time period. In order to enforce this, we discretize the timetable period into equal intervals of time and at each time interval observe all the track segments making up the junction. The elements for our set packing model can then be considered as a time interval track segment pair, while the possible train routes would correspond to the sets. A possible route through the junction is then simply a vector of zeros and ones. A one in a particular row would correspond to the train occupying that particular track segment at that particular time, while a zero would indicate otherwise. As is evident from Figure 1, it is quite possible for a train to have a number of possible routes through the junction, including of course the *null route*, which refers to the train not being scheduled (such routes are heavily penalised in the objective function). To ensure only one route is chosen for a particular train, we need to include a generalized upper bound constraint for each train. A detailed description of the integer programming model is given below.

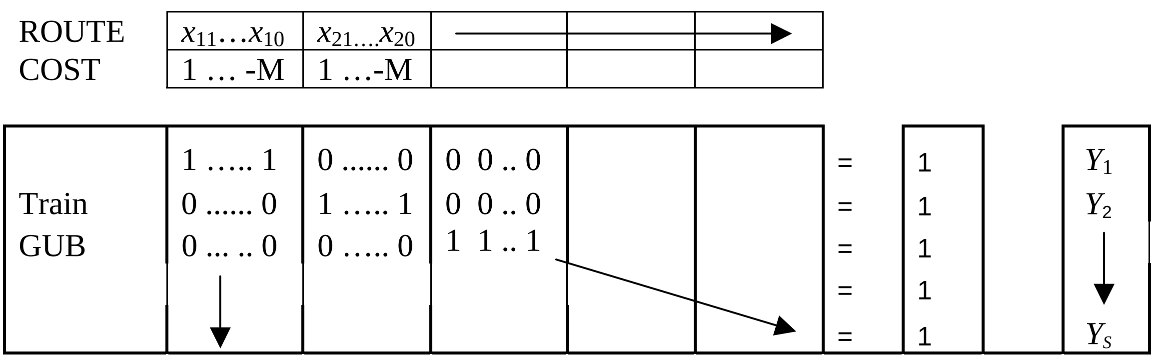
Suppose the proposed timetable asks us to assign routes to S trains through a junction consisting of W track segments indexed by w. Further assume that the timetable period is V and that this has been discretized into equal intervals of time and is indexed by v. The set packing model can be logically partitioned into a set of S train constraints, one for each of the trains in the proposed timetable, and a set of VW time period track segment constraints, one for every track segment in every time period which must be satisfied. The decision variables of the problem x_{tr} (equal to one if train t is assigned to route r where r is one of the allowable routes for train t, and zero otherwise) may also be partitioned to correspond to the possible routes for a particular train. The A matrix of this set packing model is a zero one matrix partitioned as

$$A = \begin{bmatrix} T_1 \; T_2 \; T_3 \; \ldots\ldots\ldots T_S \\ R_1 \; R_2 \; R_3 \ldots\ldots\ldots R_S \end{bmatrix}$$

where $T_i = \boldsymbol{e}_i \boldsymbol{e}^{\mathrm{T}}$ is a $(S \times n_i)$ matrix with $\boldsymbol{e}_i$ the ith unit vector and $\boldsymbol{e}^{\mathrm{T}}$ = (1,1,1,…,1). The n_i possible routes for train i form the columns of the $(VW \times n_i)$ matrix R_i with elements r_{jk} defined as r_{jk} =1 if the kth route for train i uses the jth time track segment constraint and r_{jk} =0 otherwise. The total dimensions of the A matrix is $(S + VW) \times \sum_i^S n_i$. The right-hand-side vector $\boldsymbol{b}$ is given by b_i =1, i =1,2,..,$(S+VW)$. In this particular instance as we are maximizing the number of trains assigned non-null routes, all non-null routes contribute one to the objective, while null routes contribute –M (where M is a large positive number).

Each possible route for a particular train must be a permissible route through the junction. That is, columns must satisfy the route locking and sectional release system discussed earlier as well as incorporate buffer times for when track segments are released. We can account for such requirements in the variable generation stage, and thus do not need to explicitly define them in the set packing model.

Figure 2 gives a diagrammatic representation of the above primal model. It has the following characteristics. Firstly, there are relatively few variables (there are a limited number of possible train routes), all of which can be generated through a column generation procedure. In determining the exact claim and release times for each of the track segments comprising a particular route the column generation procedure takes into account the train speed on arrival at each track segment, as well as its maximum acceleration and deceleration along the track segment. Secondly, there are a significant number of constraints; one for each track segment in each time interval plus S train GUB constraints. This is an obvious limitation.



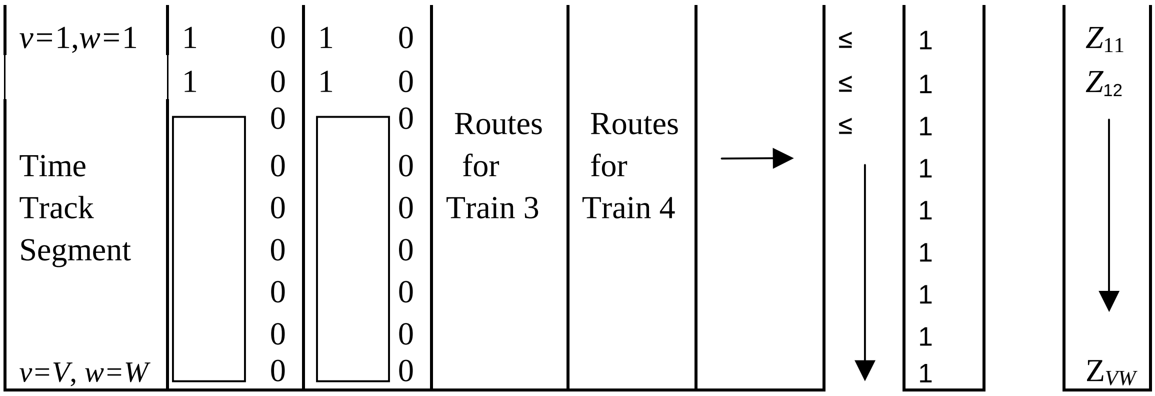


Fig 2. A diagrammatic representation of the primal formulation.

Whereas the above primal formulation is computationally unattractive due to the significant number of constraints it has, the corresponding dual possesses some nice properties that we can exploit when solving this problem. In the dual formulation, possible train routes are now represented as constraints. Thus the maximum number of constraints we could expect would be the number of primal variables. This is considerably less than the number of constraints in the primal formulation.

On the other hand, the dual does have many variables. The A-matrix of the dual is a zero one matrix portioned as

$$A = \begin{bmatrix} T_1^{\mathrm{T}} & R_1^{\mathrm{T}} \\ T_2^{\mathrm{T}} & R_2^{\mathrm{T}} \\ T_3^{\mathrm{T}} & R_3^{\mathrm{T}} \\ \ldots\ldots\ldots\ldots & \\ T_S^{\mathrm{T}} & R_S^{\mathrm{T}} \end{bmatrix}$$

Where the matrices T_i and R_i are the same as that defined earlier. However, due to the nature of this particular problem we would expect most of these variables to be zero at optimality. The vast majority of the variables in the dual formulation are the dual variables on the time track segment constraints, and it would be highly unlikely that every track segment would be claimed in every time period.

When solving the dual model we are obviously minimizing the sum of the dual variables; the primal right-hand-side vector is a vector of ones. The decision variables can be categorized into two groups; those belonging to the primal train generalized upper bound constraints, and those belonging to the time track segment constraints. The former are the dual variables on equality constraints and hence can be either positive or negative, while the latter are restricted to be non-negative.

The right-hand-side vector for the dual formulation is the coefficient vector for the primal objective. Hence in this particular case it too can be categorized into two groups. It is equal to one for all constraints corresponding to non-null train routes, while it is equal to –M otherwise. The constraints of the dual are simply enforcing the requirement that the accumulation of the dual variables along any route be greater than or equal to the benefit we would receive in scheduling that particular route.

Figure 3 below schematically represents the dual formulation. This formulation possesses nice structure also, and is clearly the more preferable model to use from a computational point of view as it has a much smaller basis. In the next section we introduce our solution procedure.

Dual V	Y_1	Y_2	...	Y_S	Z_{11}	Z_{12}	…	Z_{1W}	Z_{21}	…	Z_{VW}
Cost	1	1	…	1	1	1	…	1	1	…	1

	1	0	0	0	Train 1 non null routes							≥	1	x_{11}
	1	0	0	0								≥	1	
	1	0	0	0	0	0	0	0	0	0	0	≥	-M	x_{10}
Train	0	1	0	0	Train 2 non null routes							≥	1	x_{21}
Route	0	1	0	0								≥	1	
	0	1	0	0	0	0	0	0	0	0	0	≥	-M	x_{20}
	0	0	..	0	Other train routes							≥		
	0	0	..	0								≥		
	0	0	0	1	Train S non null routes							≥	1	X_{S1}
	0	0	0	1								≥	1	
	0	0	0	1	0	0	0	0	0	0	0	≥	-M	x_{T0}

Fig 3.A diagrammatic representation of the dual formulation.

4. Solution Approach

The solution approach that we propose utilizes the inherent advantages of the dual formulation outlined above. Essentially it involves solving the dual problem via the implementation of primal column generation as dual constraint addition. By employing it in this manner, we have the ability to dynamically update the dual problem, and in essence work with a small basis. In other words, we need not have all the possible routes in our initial dual problem, but rather just enough to have a primal basic feasible solution. We can then add constraints to this problem on finding primal entering variables.

As was mentioned earlier, each dual constraint represents a valid route through the junction, thus the respective dual variable on each of the constraints represents the value of the corresponding primal variable. In solving the dual problem we are finding the dual variable for each of the primal constraints. This solution vector can be easily used to find the reduced cost of any non-basic primal variable. If there does exist a non-basic primal variable that prices out favourably, it is simply added to the dual problem as a constraint. The problem is then resolved and a new set of dual variables is determined.

If on solving a particular instance of the dual problem, a constraint becomes inactive it is removed. Inactive constraints in the dual indicate the corresponding primal variable is at value zero. This dynamic approach of adding and removing constraints of the dual allows us to maintain a small basis throughout.

When solving a particular instance of the dual problem we consider only a subset of the primal non-basic variables. If it so happens that none of these have a favourable reduced cost, a column generation procedure is invoked and it returns a number of potential primal entering variables. The column generation procedure knows with what speed and at what time each train arrives at the junction. Using the dual variables and taking into account the maximum acceleration and deceleration along each track segment for each train, it will generate possible primal entering variables (if they exist). It even permits the train to come to a standstill on a given track segment if such an action is possible. On reaching the point where we cannot find a primal entering variable, even after running the column generator, we conclude that we are optimal. The optimal primal solution is then just the dual variables for the optimal dual solution.

In initializing the dual problem, we assign each of the trains the null route. This is obviously a primal basic feasible solution, albeit the worst. Our initial dual problem has *S* constraints, and is trivial to solve. We then proceed in the manner described above.

4.1 Example of the solution procedure.

This section reinforces the ideas of the previous section with use of some examples. Figure 4 illustrates what the initial dual problem is. Obviously, there is only one solution to the problem. That is to set all the GUB dual variables to –M. This vector is then used to price out the non-basic primal variables. At this initial stage it is apparent that any valid route for any train would have a favourable reduced cost. Without loss of generality assume that we decide to add the first route for the first train.

Dual V	Y_1	Y_2	Y_3	….	Y_S	Z_{11}	…	Z_{1W}	Z_{21}	…	Z_{VW}			
Cost	1	1	1	1	1	1	…	1	1	…	1			
	1	0	0	0		0	0	0	0	0	0	≥	-M	x_{10}
	0	1	0	0		0	0	0	0	0	0	≥	-M	x_{20}
Train	0	0	1	0		0	0	0	0	0	0	≥	-M	x_{30}
Route	0	0	0			0	0	0	0	0	0	≥	-M	x_{40}
						0	0	0	0	0	0	≥		
					1	0	0	0	0	0	0	≥	-M	x_{S0}

Fig. 4 Diagrammatic representation of the initial dual basis.

Hence this particular route is added as a constraint to the dual problem. Figure 5 below shows our new dual problem. The current dual solution is no longer feasible with this additional constraint. Including an artificial variable *A* whose value is equal to the reduced cost of the primal entering variable allows us to easily identify a starting basis for the new dual problem

Dual V	Y_1	Y_2	Y_3	….	Y_S	Z_{12}	…	Z_{1W}	..	Z_{VW}	A			
Cost	1	1	1	1	1	1	…	1	…	1	M			
	1	0	0	0		0	0	0	0	0	0	≥	-M	x_{10}
	0	1	0	0		0	0	0	0	0	0	≥	-M	x_{20}
Train	0	0	1	0		0	0	0	0	0	0	≥	-M	x_{30}
Route	0	0	0			0	0	0	0	0	0	≥	-M	x_{40}
						0	0	0	0	0	0	≥		
					1	0	0	0	0	0	0	≥	-M	x_{S0}
	1	0	0	…	0	TRAIN 1 ROUTE					1	≥	1	x_{11}

Fig 5. Constraint addition.

Starting from this initial basis it takes relatively few iterations to re-optimize the dual problem and obtain a new vector of dual variables for the primal problem. Clearly for the example above, the addition of the new constraint results in the first constraint becoming inactive. Thus it can be deleted and we can return to a problem with just S constraints. We continue in this way, employing column generation to generate subsets of potential primal entering variables, until we have reached optimality.

5. Conclusion

In this paper we have introduced the problem of routing trains through railway junctions. We briefly described how this problem has been previously treated, introduced both the primal and dual formulations of our set packing model emphasizing the attractiveness of the dual. We have also discussed our solution approach in which we solve the dual problem by implementing primal column generation as dual constraint addition. Although we have no computational results to substantiate our claims that we can significantly reduce solution time and quality using our novel approach (it is very much work in progress), we are confident of doing so given the computational simplicity of our procedure.

References

Zwaneveld, P.J., S Dauzère-Peres, S.P.M van Hoesel, L.G. Kroon, H.E. Romeijn, M.Salomon, H.W. Ambergen. 1996. "Routing Trains Through Railway Stations: Model Formulation and Algorithms." *Transportation Science* **30** (3): 181-194.

Zwaneveld, P.J., L.G. Kroon, S.P.M van Hoesel. 2001 "Routing Trains Through a Railway Station Based on a Node Packing Model." *European Journal of Operational Research* **128** (2001): 14-33

Supply Boat Routing for Statoil

Vladimir Brkovic
Department of Engineering Science
Faculty of Engineering
University of Auckland
New Zealand
vbrk001@ec.auckland.ac.nz

Abstract

Statoil is an integrated oil and gas company that operates 60% of all Norwegian oil and gas production. They are the world's third largest seller of crude oil and the largest retailer of oil products in Scandinavia. Each year, Statoil ships over one million tonnes of various supplies from their bases to their offshore installations, utilising some fifty ships. Due to the enormous costs associated with such grand scale operations, they seek to utilise these supply vessels as cost effectively as possible.

This paper investigates the routing of a single supply boat from a single supply base by developing a model and evaluating its performance under different complexities of the problem. This model must take into account aspects of the problem such as the boat capacity, installation capacity, necessary working space, and commodity delivery and pick-up demands. A mixed integer programming formulation is used to model this problem by extending the formulation of the Merchant Subtour Problem to include the installation capacity constraints and the necessary working space constraints. This extended formulation appears to be valid for the simple case of routing a single supply boat from a single supply base.

1 Introduction

1.1 Problem Description

A model for the routing of a single supply boat from a single supply base is required. This model must take into account aspects of the problem such as the boat capacity, installation capacity, necessary working space, and commodity delivery and pick-up demands. Here, commodities are goods (food, equipment etc) that are each defined as unique source-destination ordered pairs.

For the boat to be able to deliver to or pick up commodities from an installation, there must be either spare space on the boat, on the installation, or on both, for the exchange to take place, since these exchanges take place out on the open sea, with no spare space, such as the ground. The only places where a commodity can be situated are on the boat, on an installation or on the crane. Consequently, our model must ensure that there is either free space on the boat, on the installation, or on both. This free space will be referred to as the 'necessary working space'.

1.2 The Merchant Subtour Problem (MSP)

Verweij and Aardal describe a problem that they termed the Merchant Subtour Problem (MSP) where a merchant travels in a vehicle of fixed capacity, and makes money by buying commodities where they are cheap and selling them where he can make a profit. A merchant subtour is a directed closed walk, starting and ending at a given distribution centre, with a description of the load of the vehicle between each centre.

Given the prices of all of the commodities in all of the distribution centres, and the cost of driving from one centre to another, the merchant must select a subset of centres that he can visit within a prescribed time limit, and to determine the order in which the centres are visited, with the maximised total profit, while never exceeding the capacity of the vehicle.

2 Model Formulation

In Case 1, we consider the situation of the boat being allowed at most one visit to each of the offshore installations and to the depot (supply base). In Case 2, we consider the more complicated problem of allowing the boat to visit each installation at most twice, while still only permitting it to visit the depot once, i.e. a single subtour.

In both cases, we have modified the Merchant Subtour Problem formulation. The in- and out-degree constraints are necessary to our problem, representing either the boat visiting an installation and then leaving it, or the boat not visiting an installation, as represented in Figure 1 by self-loop arcs at installations 'A' and 'B'. The vehicle (boat) capacity constraints and the demand constraints are also needed. In general, we need the subtour elimination constraints, but due to the problem sizes of the above two cases, they have been replaced in our formulation by a simple arc constraint that states that the arc leaving the 'Depot' node must always be 1. This means that any subtour will always contain the 'Depot' node.

Unlike Verweij and Aardal assumption of the infinite installations' (distribution centres) capacity, in our problem all offshore installations have limited storage capacity; therefore one of the major modifications we have made to the MSP formulation is the addition of the installation capacity constraints, to ensure installations' capacity is never to be exceeded.

Another underlying difference is that in Verweij and Aardal's problem, there was always extra space available, like the 'ground', to facilitate the exchange of commodities. They could sit there until the necessary space was freed up for them. In our problem the commodities would have to be placed in the sea; therefore our models has to include the necessary working space constraints to ensure that there is either free space on the boat, on the installation, or on both.

2.1 Model Assumptions

Before formulating a model for the problem, two assumptions were made:

- An installation will not order more commodities than it has capacity to accept.
- A commodity cannot be picked up from one installation and delivered to another - it can be transported from the depot to an installation and vice-versa. The commodities, though, can be transported to an installation or to the depot via other installations, but they do not leave the boat.

2.2 Definitions of Variables and Problem Parameters

Our problem is formulated on a directed graph $G = (V, A)$, in the same way as discussed in section 1.2. The set of vertices (nodes) V corresponds to the set of installations, the depot and the boat, while A is the set of directed arcs (including self-loops). Arcs are represented by binary variables $x_{(i,j)}$, where $(i, j)\in A$. The variable $x_{(i,j)}=1$ if the arc between nodes i and j is used in the subtour, and $x_{(i,j)}=0$ otherwise.

We define P as the set of all paths along which commodities can be transported by the boat. A path from node i to j is denoted by $i...j$. The amount of commodity transported on the boat travelling along a path $i...j$ is represented by the decision variable $f_{i...j}$.

Commodities are defined as source destination ordered pairs $(u, v) \in V \times V$, $u \neq v$, with u representing the source and v the destination of the commodity. Because we are assuming that a commodity cannot be picked up from one installation and delivered to another, the commodities considered in this project are depot-installation and installation-depot ordered pairs. The demands are represented by the vector $\boldsymbol{d}$, where the maximum amount of commodity (u, v) that can be shipped is $d_{(u, v)}$.

The capacity of the boat is D, and the capacity of installation i is D_i. The necessary working space at installation i is S_i. The free space available on installation i immediately after visit v is represented by the variable $s_{(i, v)}$. These variables are only used in the model formulation for Case 2, as we need to include the amount of free space on the installation before the second visit.

2.3 Construction of the Graphs

For Case 1, the nodes of the graph are the depot, the boat and each of the installations. In Case 2, where the boat is allowed to visit each installation more than once (at most twice in this case), we use a network transformation called vertex duplication to define nodes that represent an installation being visited by the boat for the first or second time. The 'Depot' and 'Boat' nodes remain the same.

2.4 Model Formulation

2.4.1 In- and Out-degree Constraints

Figure 1 below is the graph for Case 1. From this diagram we formulated the in- and out-degree constraints for all nodes. As mentioned previously, the arc leaving the depot node is always 1, thus

$$x_{(Depot, Boat)} = 1.$$

This is also the in-degree constraint for the node 'Boat'.

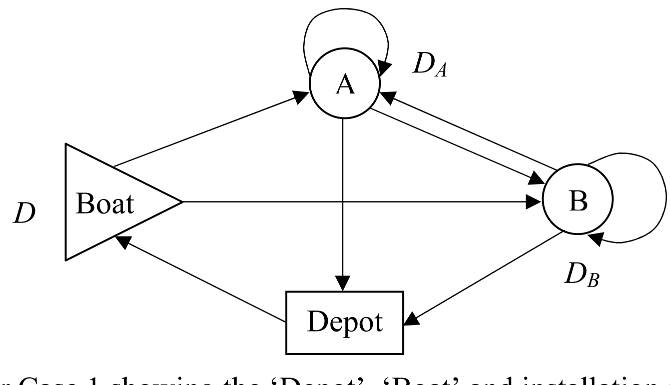


Figure 1 Graph for Case 1 showing the 'Depot', 'Boat' and installation nodes 'A' and 'B', and all of the relevant arcs along which the boat can travel while transporting commodities.

The out-degree constraints for the 'Boat' node and nodes 'A' and 'B' (installations) are written by summing all x variables corresponding to arcs that are leaving a particular node (including the node's self-loop), and making the sum equal to 1. That is, the total number of arcs that are leaving a particular node, including the node's self-loop, must be 1. In the case of nodes 'Boat' and 'Depot', there is no self-loop, and it means that the boat will either visit installation ‘A’ or ‘B’ to deliver or pick up commodities before returning to the depot. The out-degree constraints for nodes ‘Boat’ and ‘A’ are shown below.

$$x_{(Boat,A)} + x_{(Boat,B)} = 1 \qquad x_{(A,A)} + x_{(A,B)} + x_{(A,Depot)} = 1$$

The in-degree constraints state that the total number of arcs entering a node (including self-loops) must be 1. These constraints for nodes ‘B’ and ‘Depot’ are as written below.

$$x_{(Boat,B)} + x_{(A,B)} + x_{(B,B)} = 1 \qquad x_{(A,Depot)} + x_{(B,Depot)} = 1$$

Let A be the set of all arcs. We can write the in- and out-degree constraints for node v as

$$\text{In-degree}: \sum_{\substack{(u,v)\in A \\ u\neq v}} x_{(u,v)} + x_{(v,v)} = 1 \qquad \text{Out-degree}: \sum_{\substack{(v,u)\in A \\ v\neq u}} x_{(v,u)} + x_{(v,v)} = 1 \qquad (2.1)$$

For Case 2, we make use of a network transformation called vertex duplication (also used by Verweij and Aardal) to allow for the situation in which each installation can be visited twice. Nodes ‘Depot’ and ‘Boat’ remain the same, while the installation nodes now become installation visit pairs, stating the installation that the boat is visiting and whether it is its first or second visit to that installation. Therefore nodes ‘A’ and ‘B’ become ‘A1’ and ‘A2’, and ‘B1’ and ‘B2’, respectively. The graph for Case 2 is shown below in Figure 2.

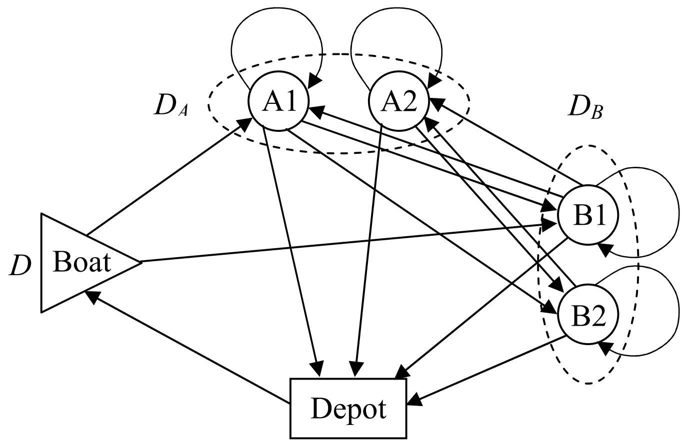


Figure 2 Graph for Case 2 showing the ‘Depot’ and ‘Boat’ nodes, the installation visit pair nodes ‘A1’, ‘A2’, ‘B1’ and ‘B2’, and all of the relevant arcs along which the boat can travel while transporting commodities.

There is no real change to these constraints from the corresponding ones in Case 1. The only difference is that instead of having the in- and out-degree constraints for each of installations ‘A’ and ‘B’ like in Case 1, we have to define these constraints for each of the installation visit pairs.

2.4.2 Boat Capacity Constraints

These constraints ensure that the capacity of the boat as it is traversing any arc (excluding self-loops) is never exceeded. Therefore, if a particular arc is being traversed, the sum of the flows of commodities along all paths that contain that arc must not exceed the capacity of the boat. For example, let the arc being traversed be the one that connects nodes 'Boat' and 'A', i.e. arc (Boat, A). The paths that contain this arc are (see Figure 3):

$$\text{Depot} \rightarrow \text{Boat} \rightarrow \text{A} \quad \text{and} \quad \text{Depot} \rightarrow \text{Boat} \rightarrow \text{A} \rightarrow \text{B}$$

The total flow of commodities along these paths is the sum of the *f*-values corresponding to these paths. The capacity of the boat is *D*. Therefore, the boat capacity constraint for arc (Boat, A) is

$$f_{Depot \rightarrow Boat \rightarrow A} + f_{Depot \rightarrow Boat \rightarrow A \rightarrow B} \leq Dx_{(Boat, A)}.$$

When a particular arc is part of the subtour that the boat uses to deliver or pick up commodities, the *x*-value corresponding to that arc is equal to 1. This will make the right-hand side of these constraints equal to the boat capacity *D*. The left-hand side is equal to the sum of the flows of commodities along all paths that contain this arc. If the boat does not traverse a particular arc, the corresponding *x*-value will be 0, and there will be no paths that contain this arc. Therefore, the boat capacity constraint for this arc will reduce to

$$f_{\ldots} + f_{\ldots} + \cdots \leq 0.$$

Let *P* be the set of all depot-installation and installation-depot paths in *G*. If *p* is a path in *P*, then the boat capacity constraint for any arc (i, j) can be expressed as

$$\sum_{\substack{p \in P \\ (i,j) \in p}} f_p \leq Dx_{(i,j)} \qquad (2.2)$$

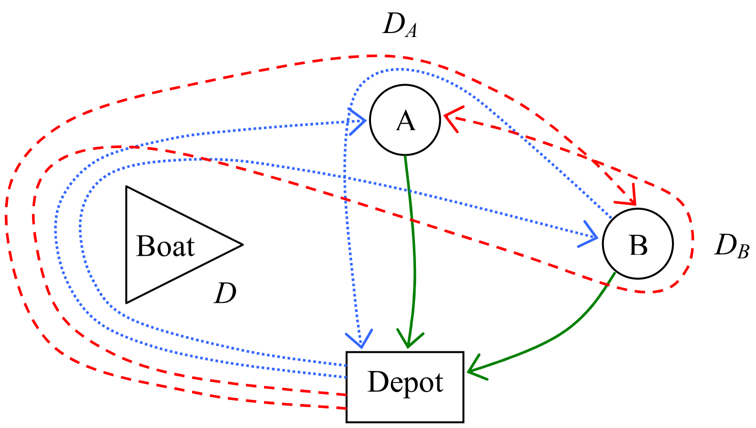


Figure 3 Path diagram for Case 1 showing all possible paths the boat may take to transport the different commodities.

For Case 2, these constraints also do not change from their formulations for Case 1 because they do not take into consideration the number of visits that the boat makes to an installation. There are now more of these constraints due to the greater number of arcs needed for Case 2 (excluding self-loops).

2.4.3 Installation Capacity Constraints

These constraints ensure that the capacity of any installation is never exceeded, i.e. the difference between the sum of commodities delivered to an installation and the sum of commodities picked up from that installation must not exceed the free space at the installation.

Commodities are source-destination pairs. The boat traverses a subtour picking up and delivering commodities along a path. This path lies "within" the subtour traversed by the boat. Since the subtour of the boat was chosen arbitrarily, we need to enumerate all possible paths along which a commodity may be transported, i.e. all paths between each source-destination pair.

Sum of commodities delivered to the installations
To deliver commodities to installation 'A', the boat must traverse one of the arcs (Boat, A) and (B, A). To deliver commodities to installation 'B', the boat must traverse one of the arcs (Boat, B) and (A, B).

Sum of commodities picked up from the installations
To pick up commodities from installation 'A', the boat must traverse one of the arcs (A, B) and (A, Depot). To pick up commodities from installation 'B', the boat must traverse one of the arcs (B, A) and (B, Depot).

The initial amount of free space at an installation is expressed as:
Capacity of the installation – Commodities to be picked up from the installation ≥ 0, or
$D_A - d_{(A,\ Depot)}$, for installation 'A'.

Therefore, the installation capacity constraints for installations 'A' and 'B' respectively are:

$$f_{Depot \to Boat \to A} + f_{Depot \to Boat \to B \to A} - f_{A \to Depot} - f_{A \to B \to Depot} \le D_A - d_{(A,Depot)}$$
$$f_{Depot \to Boat \to B} + f_{Depot \to Boat \to A \to B} - f_{B \to Depot} - f_{B \to A \to Depot} \le D_B - d_{(B,Depot)}$$

Let $p \in P$, and *start*(p) and *end*(p) represent the start and end nodes of path p respectively. The installation capacity constraint for installation i can be written as:

$$\sum_{\substack{p \in P: start(p) = Depot, \\ end(p) = i}} f_p - \sum_{\substack{p \in P: start(p) = i, \\ end(p) = Depot}} f_p \le D_i - d_{(i,Depot)} \tag{2.3}$$

The installation capacity constraints for Case 2 differ from those in Case 1 as they must ensure that upon every boat visit to an installation, the storage capacity of that installation is never exceeded. Consequently, we introduce a new set of variables, $s_{(i,\ v)}$, that represent the free space at installation i immediately after visit v to that installation. These s-values are equal to:

Free space at the installation before the visit
- Total amount of commodities delivered to the installation (taking up space)
+ Total amount of commodities picked up from the installation (freeing up space), or

Let $s_{(i,0)} = D_i - d_{(i,Depot)} \ge 0$. Then

$$s_{(i,v)} = s_{(i,v-1)} - \left(\sum_{\substack{p \in P: start(p) = Depot, \\ end(p) = i}} f_p - \sum_{\substack{p \in P: start(p) = i, \\ end(p) = Depot}} f_p \right) = s_{(i,v-1)} - \sum_{\substack{p \in P: start(p) = Depot, \\ end(p) = i}} f_p + \sum_{\substack{p \in P: start(p) = i, \\ end(p) = Depot}} f_p \tag{2.4}$$

If the boat is visiting installation i for the first time ($v = 1$), the free space on the installation is equal to the initial free space:

$$D_i - d_{(i,Depot)}.$$

However, if the boat is making its second visit to installation i, the free space at the installation before this second visit is equal to $s_{(i,\ v\text{-}1)}$, the free space at the installation immediately after the previous (in this case first) visit.

2.4.4 Necessary Working Space Constraints

These constraints ensure that there is enough necessary working space either on the boat, on an installation, or on both, for the exchange of commodities to take place. The constraints state that before the boat can visit an installation, the space available on the boat plus the free space on that installation must be at least equal to the necessary working space at that installation. For installation i this constraint can be written as

$$D\sum_{(j,i)\in A} x_{(j,i)} - \sum_{\substack{p\in P \\ (j,i)\in p}} f_p + \left(D_i - d_{(i,Depot)}\right) \geq S_i\left(1 - x_{(i,i)}\right). \qquad (2.5)$$

If the boat visits installation i, this means that the x-value of the arc representing the self-loop at that installation will be zero ($x_{(i,\ i)} = 0$), making the right-hand side of the inequality equal to the necessary working space S_i at installation i.

If $x_{(i,\ i)} = 1$, $x_{(j,\ i)} = 0\ \ \forall(j,i)\in A$ and $f_p = 0\ \forall p\in P$ where $(j,i)\in p$, meaning that the boat does not visit installation i. The right-hand side of the inequality will be zero, whereas the left-hand side will be reduced to the expression for the initial free space at installation i ($D_i - d_{(i,\ Depot)}$). This value for free space will always be greater than or equal to zero, thus satisfying the inequality.

In Case 2 we have to incorporate the situation when the boat comes back to visit for the second time. The free space on the installation before the second visit is equal to the free space at the installation immediately after the first visit. Therefore we again make use of the $s(i,\ v)$ variables introduced in the previous section. If we let $(i,\ v)$ represent a node (installation-visit pair), then, for example, (A, 1) will represent the node 'A1'. Using this notation the necessary working space constraints for Case 2 are

$$D\sum_{(j,i)\in A} x_{(j,i)} - \sum_{\substack{p\in P \\ (j,i)\in p}} f_p + s_{(i,v-1)} \geq S_i\left(1 - x_{((i,v),(i,v))}\right) \qquad v\in\{1,2\}. \qquad (2.6)$$

2.4.5 Demand Constraints

In Case 1, there are four separate types of demands because there are four commodities that need to be considered. They are the delivery and pick-up demands of each of installations 'A' and 'B'. These demands are expressed as:

$$d_{(Depot,A)}, d_{(Depot,B)}, d_{(A,Depot)}, d_{(B,Depot)}$$

The amount of commodities delivered to or picked up from an installation must not exceed the actual demand for those deliveries and pick-ups. In other words, the total flow of commodities transported by the boat along paths that start at node i and end at node j must not exceed $d_{(i,j)}$. If we define *start*(p) and *end*(p) as the starting and ending nodes of path p respectively, then the demand constraint for commodity (i, j) can be written as

$$\sum_{\substack{p \in P: start(p)=i, \\ end(p)=j}} f_p \leq d_{(i,j)} \tag{2.7}$$

where i is the depot and j is an installation, or vice versa.

The demand constraints for Case 2 must ensure that the amount of commodity that the boat delivers or picks up both on the first and second visits to each installation does not exceed the actual demand for those deliveries and pick-ups. They are shown below.

$$\sum_{\substack{p \in P, v \in \{1,2\} \\ start(p)=i \\ end(p)=(j,v)}} f_p \leq d_{(i,j)} \qquad v \in \{1,2\}, \tag{2.8}$$

where i is the depot and j is an installation.

For both cases, the objective is to maximise the total flow of commodities being transported along the subtour to try to satisfy as much of the delivery and pick-up demands of every installation as possible:

$$\max \sum_{p \in P} f_p \,.$$

3 Selected Results

3.1 Case 1 Results

Problem Parameters 3.1:

		Storage Capacity	Necessary Working Space	Boat Capacity
Installation	A	2	1	2
	B	1	1	

Demands	
(Depot, A)	1
(Depot, B)	1
(A, Depot)	1
(B, Depot)	1

Result 3.1

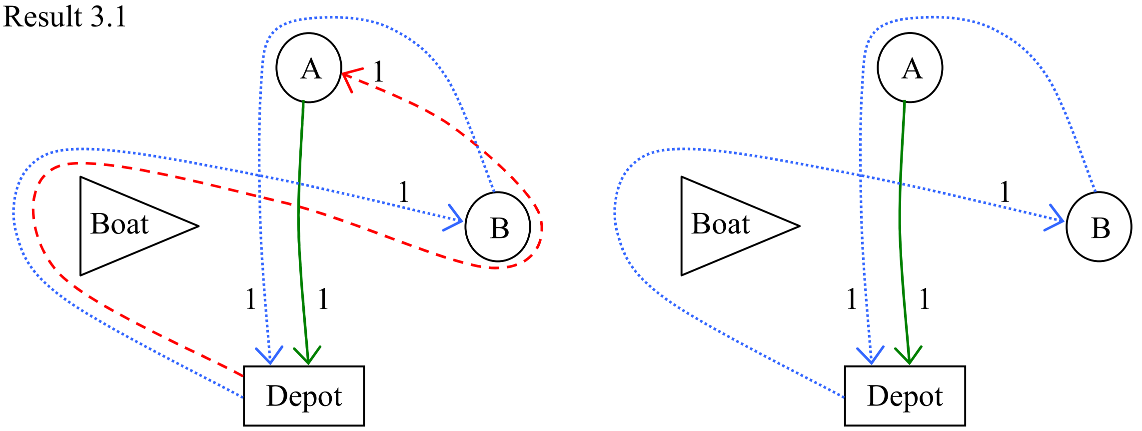


Figure 4 Graphs for Result 3.1: result obtained without (left graph) and with (right graph) the necessary working space constraints included in the model.

The total flow of commodities on the left graph in Figure 4 is 4. The boat firstly visits installation ‘B’, where it delivers one commodity and picks up one commodity, thus satisfying both the delivery and pick-up demands of that installation. Upon leaving installation ‘B’, the boat is fully laden. It then visits installation ‘A’ where it also delivers one and picks up one commodity, thus also satisfying the demands of installation ‘A’. The boat is full as it returns to the depot.

The objective function value becomes 3 with the inclusion of the necessary working space constraints to the model (right graph in Figure 4). This time, the boat again visits installation ‘B’ first. Since installation ‘B’ is at full capacity, and the necessary working space at ‘B’ is 1, the boat must therefore be only half-full (there is one free space available on the boat). The boat uses that free space to make the exchange by picking up the one commodity on installation ‘B’ and delivering one. It then visits installation ‘A’ to pick up one commodity and goes back to the depot without delivering any commodity to ‘A’.

Other possible solutions for this problem are for the boat to:

- Deliver one commodity to ‘A’, and deliver one and pick up one commodity from ‘B’
- Pick up one commodity from ‘B’, and deliver one and pick up one commodity from ‘A’

One of the deliveries or pick-ups will never be made because, in this case, there is not enough total free space on the boat and on either installation for the required exchange to take place.

3.2 Case 2 Results

Problem Parameters 3.2:

		Storage Capacity	Necessary Working Space	Boat Capacity
Installation	A	5	3	5
	B	4	2	

Demands	
(Depot, A)	2
(Depot, B)	3
(A, Depot)	3
(B, Depot)	2

Result 3.2

The total flow of commodities shown on the left graph in Figure 6 is 10 (all demands are satisfied). Initially, before the boat visits either installation, the amount of free space at each installation is 2. These values are calculated by subtracting the amount that needs to be picked up from an installation from that installation’s capacity.

The boat firstly arrives at installation ‘B’ where the necessary working space is 2. Since the free space at ‘B’ is 2, the boat arrives at ‘B’ fully laden. It delivers three commodities and picks up one (it firstly delivers two, picks up one and finally delivers the one remaining commodity) at ‘B’. As a result, the boat carries three commodities as it leaves ‘B’, while ‘B’ is at full capacity. This is confirmed by $s_{(B,\ 1)} = 0$ in the solution, which states that the free space on installation ‘B’ after the boat makes its first visit there is 0.

The boat then visits installation ‘A’ where the necessary working space is 3. Since installation ‘A’ has free space of 2, the boat has one available free space as it arrives at ‘A’. It delivers two commodities to ‘A’, loading it to capacity, therefore $s_{(A,\ 1)} = 0$ in the solution. The boat then visits installation ‘B’ for the second time, picking up one commodity before making a second visit to installation ‘A’. Picking up three commodities from ‘A’, the boat returns to the depot fully laden. The amount of free space at ‘A’ and ‘B’ after the boat makes two visits to both is 3 and 1 respectively. The inclusion of the necessary working space constraints does not affect the solution.

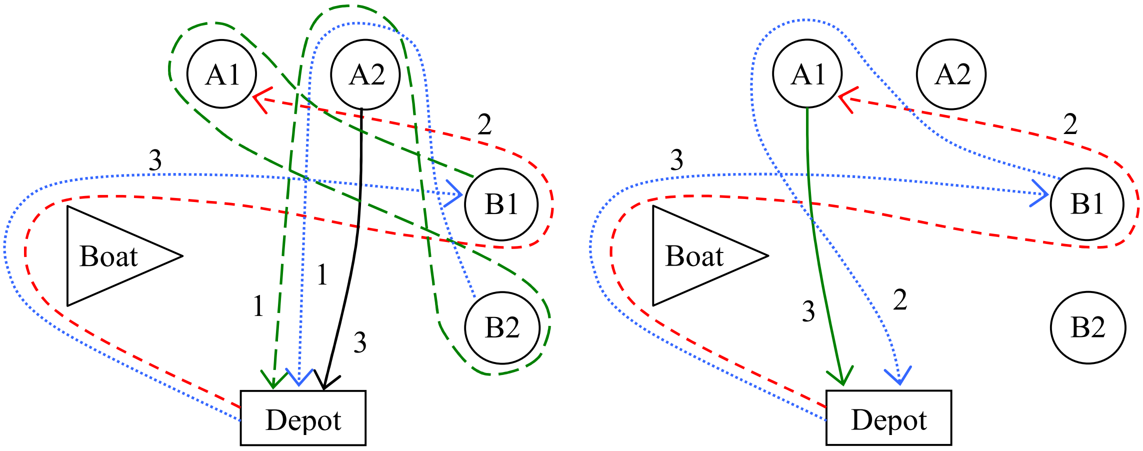


Figure 5 Graph for Result 3.2: Left – result obtained is the same with or without the necessary working space constraints included in the model. Right – alternative solution.

The above solution (left graph) is not the best one in terms of the distance travelled by the boat. An alternative solution that also satisfies all delivery and pick-up demands is shown on the right graph in Figure 5. It satisfies all of the delivery and pick-up demands by visiting each installation only once, unlike the solution on the left graph, where the boat visits each installation twice. Therefore, the boat travels a much shorter distance.

4 Conclusions

- Our extended formulation of the Merchant Subtour Problem appears to be valid for the simple case of routing a single supply boat from a single supply base.
- Certain results produced by AMPL could be simplified to provide more effective and logical routing solutions.
- This project represents only the ground work for further applications to be developed in order to replicate the very complex, real-life, day-to-day operations vital for optimal functioning of the installation supply system.

5 References

Verweij, B., Aardal, K. (2003). The Merchant Subtour Problem. *Mathematical Programming*, Ser B 94, 295-322.

Tutorial: Electricity Demand Forecasting Methods

Jonathan Lermit

Jonathan.Lermit@paradise.net.nz

Abstract

Both short term, up to a year ahead, and long term, a decade or more ahead, are discussed. The short term problem is essentially one of pattern recognition in which the effect of season, holidays and weather have to be separated out and each of these projected forward independently. Electricity generation requires forecasting at least at the half hourly level, using appropriate regression, pattern recognition and Kalman filter techniques. Long term forecasting on the other hand requires knowledge of the main drivers of consumption, such as population and economic growth and the demands of large industry. It is usually necessary to fit econometric models to past annual data and forecast forward on the basis of expected population and economic growth. The appropriateness of these techniques and their success and failure over the last few decades will be discussed.

Keywords: Forecasting, electricity, econometrics

Modelling of Frequency Keeping in the New Zealand Electricity Industry

G. R. Everett
Norske Skög Tasman Limited
Private Bag, Kawerau
graeme.everett@norskeskog.com

Abstract

The New Zealand national grid operator operates a market to procure frequency regulation services from generators. We discuss the frequency market and its relationship with the electricity market.

1 Summary

The grid operator includes a number of restrictions in the frequency keeping market design. We show that if these restrictions can be relaxed the costs of procuring frequency keeping services may be reduced. We show that minimum and maximum output constraints for frequency keeping agents affect the nodal energy prices produced by SPD. Furthermore we show that the constrained on compensation paid to frequency keeping agents can potentially be manipulated and briefly review two proposals that have been suggested to solve this problem. Finally we consider an approach that enables the frequency keeping auction to be included in the SPD dispatch software.

2 Frequency Keeping Auctions

In New Zealand the national grid operator has a number of principal performance obligations to meet. One of these is to maintain frequency at or near 50 Hz under normal operating conditions. The grid operator procures frequency keeping services from generators whose speed governors are capable of balancing frequency variations caused by unanticipated changes to demand and differences between modelled and actual parameters such as generation ramping. Simple auctions are used by the grid operator to select generators to provide frequency keeping services (often referred to as frequency keepers). Suppose there are G generation stations indexed by g. Each generation station supplies up to $t \in T$ tranches of frequency volume offers which we define as f_{gt},with associated costs c_{gt}. We define non-negative variables F_{gt} that represent the volume of frequency rights granted to generator g in tranche

t. The grid operator wishes to award q units of frequency rights, and achieves this by solving the following simple auction problem:

$$\begin{array}{lrcl} min & \sum_{g=1}^{G}\sum_{t=1}^{T} F_{gt}c_{gt} & & \\ s.t. & F_{gt} & \leq & f_{gt}, \quad g \in G, \quad t \in T. \\ & \sum_{g=1}^{G}\sum_{t=1}^{T} F_{gt} & = & q \end{array} \tag{1}$$

The grid operator currently adds a further constraint that frequency keeping rights can only be awarded to multiple generation units or machines if they are owned and operated by the same company. Suppose there are J generation companies indexed by j. Let the set $\mathcal{G}[j]$ represent the generation units or machines that are owned by generation company j, and let us define binary decision variables,

$$\sigma_j = \begin{cases} 1, & \text{if generation company } j \text{ is awarded frequency keeping rights} \\ 0, & \text{otherwise.} \end{cases}$$

The frequency auction problem now becomes:

$$\begin{array}{lrcl} min & \sum_{g=1}^{G}\sum_{t=1}^{T} F_{gt}c_{gt} & & \\ s.t. & F_{gt} & \leq & \sigma_j f_{gt}, \quad g \in \mathcal{G}[j], \quad t \in T. \\ & \sum_{j=1}^{J} \sigma_j & = & 1 \\ & \sum_{g=1}^{G}\sum_{t=1}^{T} F_{gt} & = & q \end{array} \tag{2}$$

From various discussions in the industry we understand that there is some doubt regarding the speed of response to frequency variations by thermal units. There is a view that if thermal units and hydro machines are allocated frequency keeping rights in the same trading period then close coordination is required and this can realistically only be achieved by a single trading desk. Further concerns exist regarding the differences in response times of hydro machines that could result in one machine responding more quickly than others. We do not consider any discussions regarding the relative merits of different generation types but simply make the observation that if technical solutions can be found to obviate these problems then the restrictions applied to the auction can be relaxed and the cost of procuring frequency keeping services is likely to be reduced.

2.1 Relaxation 1

We note that (2) restricts the choice of frequency keeper to a single company regardless of whether thermal units are involved or not. We propose a relaxation of (2) to allow more than one company to be eligible for the provision of frequency keeping services, as long as thermal units are not involved.

We define the set $\mathcal{T} \subseteq G$ indexed by g as the set of thermal generation units and let the set $\mathcal{J}[g]$, $g \in G$ represent the company that owns generation unit g. Note that only one company can own each generation unit so the set $\mathcal{J}[g]$ contains single elements.

We define further binary decision variables:

$$\beta_g = \begin{cases} 1, & \text{if thermal generation unit } g \text{ is awarded frequency rights} \\ 0, & \text{otherwise.} \end{cases}$$

We introduce new constraints that prevent thermal generation unit g from providing frequency keeping services unless $\beta_g = 1$

$$F_{gt} \leq \beta_g f_{gt}, \quad g \in \mathcal{T}, \quad t \in T. \tag{3}$$

We introduce further constraints that ensure that if thermal generation unit g is awarded frequency keeping rights then no rights are awarded to other generation units unless they share the same owner as unit g.

$$F_{kt} \leq (1 - \beta_g) f_{kt}, \quad g \in \mathcal{T}, \quad k \notin \mathcal{G}[\mathcal{J}[g]], \quad t \in T. \tag{4}$$

$$F_{gt} \leq f_{gt}, \quad g \in G, \quad t \in T. \tag{5}$$

Suppose a thermal generator g_1, owned by company j_1 is awarded frequency keeping rights. The set $\mathcal{J}[g_1] = \{j_1\}$ and $\beta_{g_1} = 1$. Then by (4)

$$F_{gt} = 0, \quad g \notin \mathcal{G}[j_1], \quad t \in T.$$

i.e. all generation units not owned by company j_1 are unable to provide frequency keeping services, and by (3) and (5)

$$F_{gt} \leq f_{gt}, \quad g \in \mathcal{G}[j_1], \quad t \in T.$$

i.e. all units owned by company j_1 are able to provide frequency keeping services. (3) is consistent with (5)

$$F_{g_1 t} \leq f_{g_1 t}, \quad t \in T.$$

allowing thermal generator g_1 to provide frequency keeping services.

It can be readily observed that (3), (4), and (5) are equivalent to the first two constraints of (2) with $\sigma_{j_1} = 1$.

Alternatively suppose that generator g_1, owned by company j_1 is not awarded frequency rights. The set $\mathcal{J}[g_1] = \{j_1\}$ and $\beta_{g_1} = 0$. Then by (4)

$$F_{gt} \leq f_{gt}, \quad g \notin \mathcal{G}[j_1], \quad t \in T.$$

i.e. all units not owned by company j_1 are free to provide frequency keeping services.

By (3)

$$F_{g_1 t} = 0, \quad t \in T.$$

i.e. unit g_1 is prevented from providing frequency keeping services.

By (3) and (5)

$$F_{gt} \leq f_{gt}, \quad g \in G\backslash\{g_1\}, \quad t \in T.$$

i.e. all units other than generator g_1 are free to provide frequency keeping services. This result is true for any $g \in \mathcal{T}$ meaning that hydro machines from multiple owners can provide frequency keeping servies simultaneously.

Consequently we can remove the first two constraints of (2) and replace them with (3), (4), and (5) to give:

$$\begin{array}{lllll} min & \sum_{g=1}^{G}\sum_{t=1}^{T} F_{gt}c_{gt} & & & \\ s.t. & \sum_{g=1}^{G}\sum_{t=1}^{T} F_{gt} & = & q & \\ & F_{kt} & \leq & (1-\beta_g)f_{kt}, & g \in \mathcal{T}, \quad k \notin \mathcal{G}[\mathcal{J}[g]], \quad t \in T. \\ & F_{gt} & \leq & \beta_g f_{gt}, & g \in \mathcal{T}, \quad t \in T. \\ & F_{gt} & \leq & f_{gt}, & g \in G, \quad t \in T. \end{array} \tag{6}$$

2.2 Relaxation 2

If the requirement for thermal units to only be allocated frequency rights with units operated by the same owner (6) can be relaxed further by removing (3) and (4) to give the same formulation as (1):

$$\begin{array}{llll} min & \sum_{g=1}^{G}\sum_{t=1}^{T} F_{gt}c_{gt} & & \\ s.t. & \sum_{g=1}^{G}\sum_{t=1}^{T} F_{gt} & = & q \\ & F_{gt} & \leq & f_{gt}, \quad g \in G, \quad t \in T. \end{array}$$

2.3 Example

Consider an auction for 50 MW of frequency keeping services with three generators all offering two 25 MW tranches, as in Figure 1. Note that Hydro 1 and Thermal 1 have the same owner. The solution to the auction is shown in Figure 2. Not surprisingly the cost of procuring frequency keeping services decreases as the auction restrictions are relaxed.

Company	Unit	Tranche 1 Offer Price	Tranche 2 Offer Price
A	Hydro 1	100	500
A	Thermal 1	80	200
B	Hydro 2	75	500

Figure 1: Frequency Keeping Offers

	Hydro 1 Rights	Thermal 1 Rights	Hydro 2 Rights	Cost
Formulation 1	25	25	0	180
Relaxation 1	25	0	25	175
Relaxation 2	0	25	25	155

Figure 2: Frequency Keeping Auction Solutions

3 Effect on SPD model of Frequency Keeping Rights

The NZ electricity industry is centrally dispatched according to a network flow optimisation problem (see [1] for a description). The particular model used in New Zealand is known as SPD, standing for Scheduling, Pricing and Dispatch. It is convenient for our purposes to simplify this model to involve only one node thus allowing us to remove all transmission, loop and security constraints. We also remove reserve constraints to give the very simple formulation:

$$\begin{array}{llllll} \min & \sum_{i=1}^{N} c_i x_i & & & & \\ s.t. & 0 & \leq & x_i & \leq & g_i, \quad i = 1, ...N \end{array}$$

Here there are N offers of generation indexed by i with a quantity g_i at a cost c_i. The variables x_i are the volume of energy accepted from offer i.

We simplify the problem further by including just two generation offers. The network flow problem becomes:

$$\begin{array}{lllll} P1: & \min & c_1x_1 + c_2x_2 & & \\ & s.t. & x_1 + x_2 & = & d \\ & & x_1 & \leq & g_1 \\ & & x_2 & \leq & g_2 \\ & & x_1, x_2 & \geq & 0 \end{array}$$

The dual problem to *P1* is as follows:

$$\begin{array}{lllll} D1: & \max & d\pi + g_1\lambda + g_2\mu & & \\ & s.t. & \pi + \lambda & \leq & c_1 \\ & & \pi + \mu & \leq & c_2 \\ & & \pi & \mathit{unrestricted} & \\ & & \lambda, \mu & \leq & 0 \end{array}$$

Of particular interest are the values of π, the dual variables associated with the energy balance constraint, since these define the prices paid for electricity (commonly known as nodal prices).

Now if $c_1 > c_2$ and $d \geq g_1$ and $d \geq g_2$ then the optimal solution to *D1* is:

$$
\begin{aligned}
D1^* &= c_1 d + g_2(c_2 - c_1)\\
\pi &= c_1\\
\mu &= c_2 - c_1\\
\lambda &= 0
\end{aligned}
$$

3.0.1 Constraints from the Frequency Keeping Auction

Each generation unit awarded frequency keeping rights must be able to deviate output from its dispatch setpoint in response to variations in the system frequency. Each frequency keeper nominates minimum and maximum plant output that are used as the right hand sides of constraints that ensure the unit is generating within a band that enables it to perform frequency keeping services. In this section we examine the effect of the minimum output constraint.

Suppose *generator 1* has been awarded frequency keeping rights. We now add a minimum output constraint for *generator 1*, $x_1 \geq m$

$$
\begin{array}{lllcl}
P2: & \min & c_1x_1 + c_2x_2 & & \\
& s.t. & x_1 + x_2 & = & d\\
& & x_1 & \leq & g_1\\
& & x_2 & \leq & g_2\\
& & x_1 & \geq & m\\
& & x_1, x_2 & \geq & 0
\end{array}
$$

Assume that the capacity of *generator 1* and the demand d is greater than the frequency keeping instruction, and the remaining demand can be met by *generator 2*. This means that $g_1 > m$, $m < d$ and $g_2 \geq d - m$. The optimal solution to *P2* is therefore:

$$
\begin{aligned}
P2^* &= c_1 m + c_2(d - m)\\
x_1 &= m\\
x_2 &= d - m
\end{aligned}
$$

The dual problem to *P2* is as follows:

$$
\begin{array}{lllcl}
D2: & \max & d\pi + g_1\lambda + g_2\mu + m\varphi & & \\
& s.t. & \pi + \lambda + \varphi & \leq & c_1\\
& & \pi + \mu & \leq & c_2\\
& & \pi & unrestricted & \\
& & \lambda, \mu & \leq & 0\\
& & \varphi & \geq & 0
\end{array}
$$

The primal constraints $x_1 \leq g_1$ and $x_2 \leq g_2$ are not binding and consequently the dual variables $\lambda, \mu = 0$. The dual problem simplifies as follows:

$$
\begin{array}{lllcl}
D2: & \max & d\pi + m\varphi & & \\
& s.t. & \pi + \varphi & \leq & c_1\\
& & \pi & \leq & c_2\\
& & \pi & unrestricted & \\
& & \varphi & \geq & 0
\end{array}
$$

Recall that $c_1 > c_2$. Therefore the optimal solution to *D2* is:

$$\begin{aligned} D2^* &= c_2 d + m(c_1 - c_2) \\ \pi &= c_2 \\ \varphi &= c_1 - c_2 \end{aligned}$$

Here demand (d) can be met from (the cheaper) *generator 2* and the constrained on volume from the Frequency Keeper ($x_1 = m$). Consequently the more expensive generation tranche (g_1) is no longer required and the nodal price is reduced as a consequence of the minimum output constraint on *generator 1*. The value of π from *D2** is smaller than the value of π from *D1**, since $c_2 < c_1$.

This result shows that the application of a minimum output constraint on frequency keepers can suppress the price for electricity. This result is illustrated in Figures 3 and 4. For actual examples that have been observed in the NZ electricity market, see [4].

Note that we do not perform a similar analysis of the maximum output constraints here but we note that if binding these constraints can increase nodal prices. We observe that this is unlikely to be of material concern, as thermal plant operating at maximum output are unlikely to seek frequency keeping rights and hydro plant generally only operate at their maximum output at times of high inflows (and low spot prices).

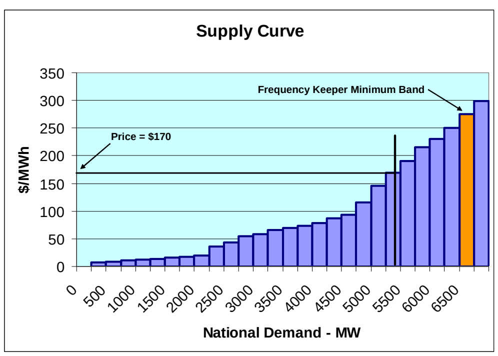


Figure 3: Supply curve without a minimum output constraint for the frequency keeper

3.1 Constrained On Compensation

Another outcome of the minimum output constraint for the frequency keeper is that *generator 1* receives c_2 for each unit of power produced, a lower price than it offered. The grid operator compensates *generator 1* by making a constrained on payment of $(c_1 - \pi)x_1 = (c_1 - c_2)x_1$, see [5]. Note that if $\pi \geq c_1$ there is no constrained on payment accruing. Note also that in a larger problem with more participants in the energy market the nodal clearing price may be higher than the generation offered by any of the frequency market participants. In this case there would be no requirement for constrained on compensation. However it is clearly apparent

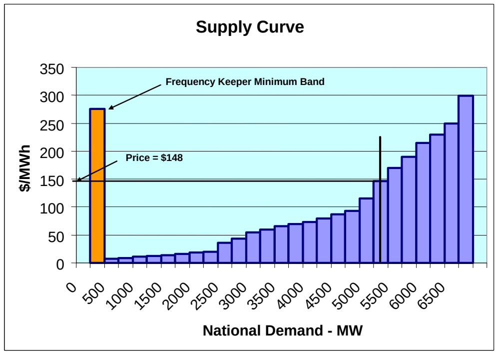


Figure 4: Supply curve with a minimum output constraint applied to the frequency keeper

that *generator 1* is able to manipulate a constrained on payment by setting a high value for its minimum output and offering energy at a high price and offering into the frequency keeping auction at a very low price. In this way *generator 1* will be awarded frequency keeping rights and consequently receives whatever price it chooses for the energy that is constrained on by the minimum output constraint. This possibility has been noted and well publicised in the electricity industry, see [2]. At least two solutions have been proposed.

3.1.1 Include Constrained On Costs in Frequency Keeping Auction

The first solution involves including the constrained on costs in the frequency keeping auction [4], though no details are provided. We note that constrained on payments depend on the values of π from solutions to SPD. We propose the following algorithm that involves running SPD twice:

1. Solve SPD without and frequency keeping related constraints to obtain values for for π and dispatch instructions for frequency market participants, $x_g, g \in G$.

2. If $x_g \geq m_g, \forall g \in G$, then solve FK auction without considering constrained on costs. Stop.

3. If $\exists$ at least one $g \in G : x_g \leq m_g$ then find

$$\widetilde{c}_g = \begin{cases} \text{the offer price for the tranche corresponding to } m_g, & \forall g \in G : x_g < m_g \\ -\pi, & \forall g \in G : x_g \geq m_g \end{cases}$$

4. Solve FK auction (1) with constrained on costs added to the objective function:

$$
\begin{array}{llcl}
min & \sum_{g=1}^{G}\sum_{t=1}^{T} F_{gt}c_{gt} + \sum_{g=1}^{G}(\widetilde{c}_g - \pi)(m_g - \sum_{t=1}^{T} F_{gt}) & & \\
s.t. & F_{gt} & \leq & f_{gt}, \quad g \in G, \quad t \in T. \\
 & \sum_{g=1}^{G}\sum_{t=1}^{T} F_{gt} & = & q
\end{array}
\tag{7}
$$

5. Solve SPD with constraints for frequency keeper minimum output included

The second term in the objective function of (7) will equal the constrained on payment for frequency market participants that have a minimum output value with an associated energy offer price that is lower than the market clearing price.

3.1.2 Deviations from dispatch instructions by Frequency Keepers

Variations in system frequency in real time cause frequency keepers to alter output, resulting in variations from dispatch setpoints and ensuing constrained on and off payments. A proposal to utilise historical probability distributions to predict constrained on and off payments has been provided in [3]. This approach would supercede that given in (3.1.1) but whether or not it would result in lower cost outcomes depends on the accuracy of the probability distributions.

3.1.3 Remove Constrained On and Off Costs to Frequency Keepers

If frequency keeping participants were not awarded constrained on (or off) costs then there is no possibility for frequency keeping agents to take advantage of the situation we discussed in (3.1). Instead agents would be expected to adjust their offers for frequency to reflect the possibility of having to deviate from their dispatch setpoint, probably by applying a similar approach to (3.1.2). This may have deleterious outcomes if the competition in the frequency keeping auction is low, but at least it does not expose the industry to potentially unlimited constrained on payments.

4 Including Frequency Keeping Auction in SPD

If either (3.1.1) or (3.1.2) are adopted there may be advantages for the grid operator if the frequency keeping auction was included in the SPD model. We propose a general formulation of SPD that includes (1), as follows in (8). Here d represents demand for energy and q represents the volume of frequency rights the grid operator wishes to award. For each generator $i \in N : y_i$ represents dispatch for frequency keeping, f_i represents frequency offer costs, h_i represents frequency offer volumes, m_i represents minimum output and n_i represents maximum output. We also define binary variables:

$$
a_i = \begin{cases} 1, & \text{if generator } i \text{ is awarded frequency rights} \\ 0, & \text{otherwise.} \end{cases}
$$

$$
\begin{array}{llccll}
P3: & \min & \sum_{i=1}^{N} c_i(x_i+y_i) + \sum_{i=1}^{N} f_i y_i & & & \\
& s.t. & \sum_{i=1}^{N}(x_i+y_i) & = & d. & \\
& & x_i & \leq & g_i & i \in N. \\
& & y_i & \leq & h_i & i \in N. \\
& & \sum_{i=1}^{N} y_i & = & q & \\
& & y_i & \geq & m_i\alpha_i & i \in N. \\
& & y_i & \leq & n_i\alpha_i & i \in N. \\
& & x_i+y_i & \leq & n_i\alpha_i+(1-\alpha_i)g_i & i \in N. \\
& & x_i, y_i & \geq & 0 & i \in N. \\
& & \alpha_i & & binary & i \in N.
\end{array} \tag{8}
$$

We remove constraints that for simplicity we will assume are not binding to give:

$$
\begin{array}{llccll}
P3: & \min & \sum_{i=1}^{N} c_i(x_i+y_i) + \sum_{i=1}^{N} f_i y_i & & & \\
& s.t. & \sum_{i=1}^{N}(x_i+y_i) & = & d. & \\
& & \sum_{i=1}^{N} y_i & = & q & \\
& & y_i & \geq & m_i\alpha_i & i \in N. \\
& & x_i, y_i & \geq & 0 & i \in N. \\
& & \alpha_i & & binary & i \in N.
\end{array}
$$

Suppose we fix the binary variables α_i. Then the dual problem to *P3* is:

$$
\begin{array}{llccll}
D3: & \max & d\pi + \omega q + \sum_{i=1}^{N} \theta_i m_i \alpha_i & & & \\
& s.t. & \pi + \theta_i & \leq & c_i, & i \in N. \\
& & \pi + \omega + \theta_i & \leq & c_i + f_i, & i \in N. \\
& & \pi, \omega & unrestricted & & i \in N. \\
& & \theta_i & \geq & 0, & i \in N.
\end{array}
$$

which can be re-written as:

$$
\begin{array}{llccll}
D3: & \max & d\pi + \omega q + \sum_{i=1}^{N} \theta_i m_i \alpha_i & & & \\
& s.t. & \pi + \theta_i & \leq & c_i, & i \in N. \\
& & \omega & \leq & f_i, & i \in N. \\
& & \pi, \omega & unrestricted & & \\
& & \theta_i & \geq & 0, & i \in N.
\end{array}
$$

Here ω represents the (marginal) price paid to frequency keepers. We note that as in (3.0.1) the minimum output constraints for frequency keepers may decrease prices for energy. However the frequency price is only influenced by frequency offer prices. Note also that the market-clearing prices for frequency (ω) and energy (π) resulting from *D3* are valid if the values for α_i from *P3* are fixed, see [6].

References

[1] G. Pritchard, *The must-run dispatch auction in an electricity market*, Energy Economics 24 (2002) 199–216.

[2] http://www.electricitycommission.govt.nz/pdfs/rulesandregs/uts/utspdfs/final-decision-uts-12July05.pdf

[3] http://www.electricitycommission.govt.nz/pdfs/advisorygroups/comqual/13Oct05/4-Ltr-DReeve-to-SO.pdf

[4] Transpower NZ, *Frequency Keeping Selection (FK) - August 2005*, http://www.transpower.co.nz/?id=3540

[5] Electricity Governance Regulations, *Part G, Section V, Rule 5.3.1*, http://www.electricitycommission.govt.nz

[6] R.P. O'Neill, P.M. Sotkiewicz, B.F. Hobbs, M.H. Rothkopf, W.R. Stewart, *Efficient market-clearing prices in markets with nonconvexities*, EJOR 164 (2005) 269-285.

Risk-Adjusted Discount Rates and Optimal Plant Mix: A Conceptual Analysis for Electricity Markets

E. Grant Read
Energy Modelling Research Group
Department of Management
University of Canterbury
New Zealand

Mike Thomas
CRA International
Asia-Pacific Office
Wellington
New Zealand

grant.read@canterbury.ac.nz

Abstract

The appropriate discount rate to be used for electricity sector investments has been debated over many decades. Governments have considered the relative merits of "commercial" vs "social" discount rates, but generally tended to use something like "risk-free" commercial rates. Being lower, these favour investment in projects which only pay off over the longer term. But, as the history of public sector electricity sector investment in New Zealand makes clear, such investment involves substantial risks, eg of cost over-run and over capacity. And, in a market environment, all participants are expected to account for risk when determining their optimal trade-off between investing in the electricity sector, and any other sector. We will show how adjusting the discount rate for risk can impact on traditional perceptions with respect to the optimal level of investment in electricity supply capacity, and mix of plant types.

A Mixed Integer Programming model can be formulated, but these effects can be demonstrated conceptually using a simple spreadsheet, which modifies a traditional "screening curve" analysis for determining optimal plant mix. Two effects may be distinguished. The overall increase in discount rates discourages investment in traditional capital-intensive base-load plant, while the recognition of risk associated with "dry year support" discourages investment in traditional peaking plant. The outcome depends on the balance of these effects, but it seems quite possible that some technologies which have hitherto been seen as desirable contributors to generation portfolios may henceforth be excluded entirely from the market plant mix. And it also seems possible that the overall reliability of supply will be reduced, implying greater reliance on "demand side management".

Despite widespread public concern, we are not aware of any convincing evidence that electricity markets, ether here or overseas, actually are systematically under-investing in generation. And it is an open question whether a reduction in electricity sector investment, and supply reliability, should be welcomed as a triumph for the environment, abhorred as a disaster for civilised society, or both. We will not attempt to resolve these philosophical questions here, but note several issues, including the likely discrepancy between private and public risk perceptions and attitudes, and weaknesses in the "contracting chain:" linking suppliers to ultimate consumers, which

might tend to imply the likelihood of systematic “market failure” with respect to provision of “backup” capacity in the long run. We also note the potential role which Governments may play in increasing investment risk by intervention, or decreasing it by fascilitating mechanisms such as “capacity tickets”.

Can the shoe be made to fit? – Cournot modelling of Australian electricity prices

James Tipping[1], Grant Read, Deb Chattopadhyay[2] and Don McNickle
Energy Modelling Research Group
Department of Management
University of Canterbury
New Zealand
james.tipping@canterbury.ac.nz

Abstract

Tools for simulating market behaviour, such as Cournot models, have for many years been used to simulate both the operation of hypothetical electricity markets and the effects of hypothetical changes to existing electricity markets. However, very little work has been undertaken in discovering how successful these models were at estimating future price levels *after* either the markets have begun operation or the changes have been implemented. The research presented in this paper, which calibrates a Cournot model with real market data from Australia's National Electricity Market, increases the credibility of these models for future use.

Aside from input information on the supply mix, transmission and load, two of the crucial inputs required by a Cournot model of an electricity market are the elasticity of demand and the extent to which generating companies are contracted. However, while the aforementioned inputs are generally publicly available, the elasticity of demand can only be estimated from observed market data, while contract levels are confidential, and often only implicit in the need to meet retail load requirements. In fact it is not even clear how either parameter should be defined in this context. Since the market is not actually Cournot, the "demand elasticity" somehow also serves as a proxy for the response of other generators. Similarly "contract levels" may also serve as a proxy for other price-restraining factors such as regulation, or the threat thereof. Nor is it clear over what time horizon participants will want to assess either parameter. Thus it seems most reasonable to deduce these parameters from observed market behaviour. The major focus of this work was to determine the effective contract rates and elasticities of demand that best enabled the Cournot model to model spot prices in peak, shoulder and offpeak periods throughout the day.

As a forecasting tool, such an analytical model cannot model all the price volatility that is realised in practice due to such unforeseen events as generating plant outages and unexpected transmission constraints. As supply-side inputs were assumed fixed

[1] Supported by the Tertiary Education Commission (TEC) Te Amorangi Matauranga Matua.
[2] CRA International (Asia Pacific) Ltd, Melbourne.

We would like to acknowledge CRA International (Asia Pacific) Ltd. for the financial support of this study, and for the many hours of assistance provided.

throughout the sample period in this study, the aim of the calibration was therefore to capture accurately the underlying or deterministic movements in the price series from each of the three load-defined periods. The second part of this study involved fitting an existing spot price time series model to the residual price movements that were not captured by the Cournot model. The combination with a stochastic process model greatly improves the overall ability of the Cournot model to estimate and forecast both the frequency of extreme prices and the persistence in the unexpected volatility. Combining these two types of models has not previously been reported in the academic literature, and such a combination is a valuable tool for future spot price forecasting projects.

Applying the theory of constraints with LP to Fonterra

Kathryn M. Jarden
Victoria Management School
Victoria University of Wellington
New Zealand
kjarden@gemtechnz.com

Abstract

This paper explores the implications of applying the Theory of Constraints (TOC) in combination with Linear Programming to Fonterra Kapuni. The objective of the research was to apply the TOC to a New Zealand business in order to develop solutions to improve its performance. Since the application of TOC was a first for the researcher there was an additional objective, which was to gain an understanding of how to apply TOC successfully and to test the relevance of TOC methods in the "real world". Case study methodology and the TOC combined with Linear Programming (LP) were used and applied to Fonterra Kapuni.

As a result of applying TOC with LP, the research highlighted the importance of a systemic perspective, which is a key philosophy of TOC. The lesson learnt was that a LP model should be used in combination with TOC to get the best results. Although LP would have provided solutions for improvement, TOC was required to identify the core problem, which the LP model alone would not have identified, to achieve significant improvement to Fonterra Kapuni's performance. While the results of the application were important, this paper will specifically focus on the implications of applying TOC in combination with LP.

1 Introduction

This paper describes a case study relating to a New Zealand manufacturing organisation, Fonterra Kapuni, which produces pharmaceutical lactose. Goldratt's Theory of Constraints (TOC) and spreadsheet linear programming (LP) were the methods used in combination to provide performance improvement solutions. The purpose of this research was to provide solutions to improve the performance of an organisation through the application of the Theory of Constraints (TOC). The LP was used to assist as a decision aid in the identification of the constraint and to provide further solutions.

The case study described in this paper was the result of research carried out at Fonterra Kapuni in 2004 for the completion of my thesis in a Masters of Management Studies. The thesis topic was chosen as a result of Fonterra Kapuni describing a problem in its manufacturing process. Fonterra Kapuni suspected its primary evaporators were a bottleneck as the available capacity was unable to keep up with the supply of raw materials. TOC endeavours to overcome the constraint or bottleneck and create a process of continuous improvement, which I thought would provide an appropriate methodology for the problem.

The Fonterra Kapuni problem also provided an opportunity to contribute to the further application of TOC in New Zealand manufacturing. Mabin and Balderstone (2003) conducted an extensive literature search to answer the question of whether TOC provides manufacturing organisations with a source of competitive advantage. The conclusion from the review was that TOC can provide a substantial source of competitive advantage to organisations. Although this has been proven using overseas examples, the application of the TOC is less well established in NZ (Baldertsone 1999, Mabin and Balderstone, 2000, Ho, 2001.) Since the application of TOC was a first for me, there was an additional objective, which was to gain an understanding of how to apply TOC successfully and to test the relevance of TOC methods in the "real world". In this research, LP was used within the TOC framework to assist with identifying the constraint, and making resource and product mix decisions. Not being a specialist in LP or mathematical optimisation tools, I found LP with TOC able to simplify a complex problem successfully that also generated useful results. It is my experience of applying TOC and LP to a NZ manufacturing organisation that I will discuss in this paper and the implications of applying such a combination.

In the remainder of this paper, I will first describe the case study, its analysis using the combined TOC and LP approach, and then discuss the implications of applying TOC in combination with LP.

2 The case study

2.1 Background

Fonterra Kapuni manufactures pharmaceutical grade lactose for Fonterra Cooperative Group Ltd (Fonterra). Fonterra is New Zealand's largest company, making its money through exporting over 90 per cent of its produce and its pharmaceutical lactose business is one of the largest global suppliers of this product. Pharmaceutical lactose is used in the manufacture of pills and in other drug delivery systems as the 'carrier' for the active drugs. The lactose at Fonterra Kapuni is produced from whey that is a by-product from the production of cheese, milk and whey protein concentrate, and casein from other Fonterra manufacturing sites.

Fonterra Kapuni currently wants to improve its performance (like any organisation). To improve its performance it wants to overcome bottlenecks in its process to be able to optimise the production of lactose. The TOC addresses these measures, operations and the cost focus to increase performance. Tools are provided for managers to give them the ability to better focus improvement activities so that their organisation can continue to improve its performance.

2.2 Analysis

The TOC consists of a theoretical framework, a measurement system and a philosophy. For this research, the methodology was chosen based on the presented problem. Mabin and Balderstone (2003) observed in their review that significant gains were achieved even when only one or two TOC tools were used. Even though not all the TOC methods were used in this research, it seemed likely that the methods below would provide major improvement:

- **Three Questions for Change –** framework to achieve successful change.
- **Five Focusing Steps –** were used to answer the three questions to find a solution to Fonterra Kapuni's problem. Other TOC tools were used within each step such

as the Evaporating Clouds and the Current Reality Tree in the first step to help identify the constraint.

- **Linear Programming –** was used in combination with TOC to create a product mix that maximises throughput through the constraint and to provide a decision aid for change.

Overall the research endeavoured to answer *what to change*, *what to change to*, and *how to cause the change* by using the five focusing steps and LP to answer the questions. These questions are the first four steps of the Five Focusing Steps but consolidated into three. By answering *what to change* the core problem can be identified by completing an analysis. *What to change to* is the strategic part of the process that involves constructing the solution to the problem. Once the solution is known, designing an implementation plan, the tactical part, satisfies *how to cause the change* (Goldratt, 1994).

2.3 The theory of constraints

The theory of constraints is a system-based methodology that has been developed to assist people and organisations to think through problems and their solutions logically and systematically (Mabin & Balderstone, 2000). The developed TOC framework gives managers the ability to better focus on improvement activities so that their organisation can continue to improve its performance (Goldratt & Cox, 1992). The literature on theory of constraints and the results from Mabin and Balderstone's review has demonstrated that it can be successfully applied to any type of organisation. Since unlimited profit generation has yet to be achieved, it can be said there is always at least one constraint that is limiting an organisation's performance. As long as there is a constraint limiting an organisation's performance, there is always an opportunity for improvement. Therefore, an assumption can be made that Fonterra Kapuni has an opportunity for improvement and the theory of constraints is a theoretical framework that can be used to realise this opportunity.

The five focusing steps are five sequential steps developed by Goldratt to identify exactly where and what improvement efforts should be concentrated to achieve the maximum global impact on the system (Goldratt & Fox, 1986). The following is an overview of the steps; further information can be found elsewhere (Goldratt, 1990, 1992; and Dettmer, 1997).

The five focusing steps:

Step 1: Identify the System Constraint – What is the weakest link? What limits the system's performance? Is the constraint inside or outside the system (the market, material, capacity or policy)? The first step is to locate the constraint that is limiting the performance of the organisation or the system's throughput (Schragenheim & Dettmer, 2001).

Step 2: Decide How to Exploit the Constraint – As the constraint is what limits the system's throughput, the constraint's performance must be improved. We first investigate how this might be achieved without any additional resources (Lepore & Cohen, 1999; and Schragenheim & Dettmer, 2001). In the second step, we seek to exploit the constraint by squeezing every bit of capability out of the constraining component as any time gained on the constraint will result in gains in the overall system (Goldratt, 1992).

Step 3: Subordinate Everything Else to the Decision made in Step 2 – Since the constraint has been identified and exploited, it is now necessary for all the non-constraints to be managed in such away that the performance of the constraint is never threatened (Schragenheim, 1999).

Step 4: Elevate the Constraint – If steps 2 and 3 were not sufficient to eliminate the constraint then next we consider whether the constraint needs to be elevated by increasing its capacity (Dettmer, 1997).

Step 5: Go Back to Step 1, Beware of Inertia – Once the constraint is broken it is necessary to go back to step 1, as a new constraint will exist that needs to be managed or eliminated. That is why the five focusing steps are known as a process of ongoing improvement. The caution to be aware of inertia refers to when a constraint is broken, not to get too comfortable because the cycle never ends.

The remainder of the paper will deal with the application. Figure 1 presents exactly *what* was trying to be improved in the form of a business system diagram that describes the system's business transformation process, and the inputs that were transformed into outputs. TOC uses the term *throughput* instead of *output* to emphasise the focus of the business on meeting customer needs. Therefore in this system, throughput is the sale of lactose to Fonterra Kapuni's customers. The importance of doing this is the focus is on producing products that will sell rather than the products that are the fastest to produce or the products that have the best efficiency rates.

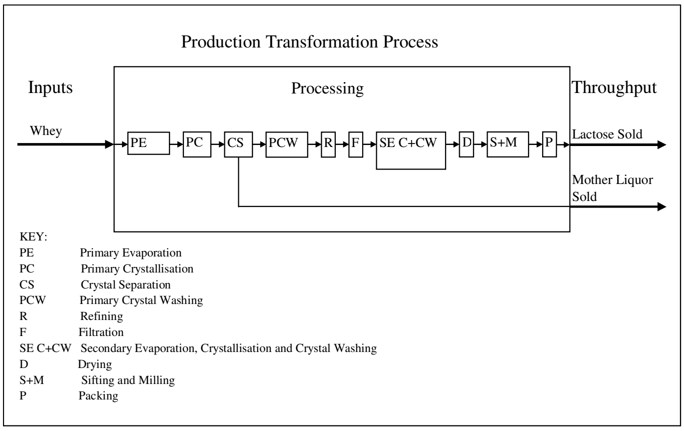


Figure 1: Fonterra Kapuni's Business System

2.4 Application

The application is presented in the five focusing steps framework:

1) *Identify the constraint*. The first step was to identify Fonterra Kapuni's constraint in order to answer what to change. Before starting, the project staff at Kapuni thought the primary evaporation stage of the process was a problem area. They also thought there were more constraints and were aware that they all had a different idea of what the constraint was, due to their position in the business process and what was relevant to them. For example, possible constraints were:
   - Primary evaporation – as they struggle to process incoming raw materials.
   - Bunkers – bunkers keep filling up with condensed whey, more storage space needed.
   - Dry process – driers or the sifters and mills are capacity constrained.

   Identifying the primary evaporation and bunkers as problem areas or possible constraints is quite natural as they are the resources that are having problems coping with the extra inventory. The only places the inventory can pile up in Fonterra

Kapuni's production system is at the start of the process and at the bunkers. Therefore it is possible to assume at this point there is a constraint that is limiting the performance of Fonterra Kapuni's production system and as a result there is a build up of inventory at the bunkers and at the start of the process before the prefinishers.

In order to gain an understanding of the cause and effect relationships in Fonterra Kapuni and its problems, Thinking Process Tools were used to create a current reality tree (CRT). The logic diagram helped to identify *what to change* and to link their undesirable effects (UDEs) to identify the core problem. The *core problem* was identified as the uncertainty behind the organisational focus due to the existence of two opposing aims. One aim was to maximise tonnes per hour by having a focus on processing RMs. The other aim was to maximise sales by meeting customer demand. Based on the TOC philosophy, in order to increase throughput, the constraint needs to be identified and the following steps of the five focusing steps carried out. In completing the five focusing steps the core problem can be resolved as there is no longer a dual focus, as the focus will be on maximising throughput through the constraint. However this will resolve the core problem but the physical constraint still has not been identified and resolved. Therefore the five focusing steps has helped to resolve the core problem, ie the split focus, but at this point the constraint still needs to be identified.

The CRT suggested that the evaporators were not the physical constraint. The constraint was identified to most likely be in dry processing that included the drying, sifting and milling of the product. The assumption was that maybe the resource was being inefficiently used due to product mix decisions. Hence the decision to try LP, along the lines of Mabin (1995) and Mabin and Gibson (1998). LP was used to identify the physical constraint in the manufacturing process that was limiting Fonterra Kapuni from achieving its goal. In this case the goal was to maximise profits (see Figure 2). Therefore if we try to maximise profits by trying to meet market demand the *constraint* is BM3. The operators have flexibility in that more than one mill can be used to make the same product. For the purpose of this paper the spreadsheet presents the most common example. It is important to note (as Mabin and Gibson (1998) did) that "product mix" problems formulated as an LP would identify more than one constraint. However, in TOC as Goldratt (1990a) argues is that one of these constraints is more critical than the others and is what we should focus actions on. Also note that the data in the spreadsheet has been modified and some omitted for commercial reasons, but the underlying structure of the problem has not been altered, so it provides a fair representation of the actual situation.

BM3 as the Constraint	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
THROUGHPUT	1,650	342	342	342	1,230	342	342	342	342	342	510	480	480	318	330	300	300	300	300	330	306	312	528	528	528	34
Quantity supplied to Market (tonnes/6 weeks)	38	51	87	14	195	148	129	1,438	14	0	73	0	77	0	14	230	46	835	965	15	272	18	0	1	6	6,897
DEMAND (tonnes/6 weeks)	38	51	87	14	195	148	129	1,438	14	0	73	53	77	30	14	230	46	835	965	28	272	18	0	1	6	6,897

PRODUCTION PLAN (mins/tonne)	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	Y	X	Y	Z	Totals	Available	Slack
Quantity to make (FG)	38	51	87	14	195	148	129	1,438	14	0	73	0	77	0	14	230	46	835	965	15	272	18	0	1	6	6,897			
RM processed (RM)	113	149	255	41	573	436	379	4,230	40	1	215	0	228	0	41	675	134	2,457	2,839	45	801	52	1	4	17	3,651	**17,377**		
Evaporation 1	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	15,675	35,127	19,452
Evaporation 2	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	13,188	35,127	21,939
Primary Crystallisation	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	15,826	42,153	26,327
Crystal Separation	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	29,013	40,396	11,383
PCW, R and F	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	0.0	23,051	40,396	17,345
Evaporation	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	2.8	0.0	18,534	40,396	21,862
Crystallisation and CW	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	3.8	0.0	18,520	40,396	21,876
Driers	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	0.0	35,048	40,396	5,349
BM1	0.0	0.0	0.0	0.0	0.0	31.3	31.3	31.3	31.3	31.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	31.3	31.3	31.3	31.3	0.0	0.0	13,494	40,396	26,902
BM2	0.0	0.0	31.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	31.3	31.3	31.3	31.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	23,688	36,884	13,196
BM3	12.5	0.0	11.4	0.0	12.5	12.5	12.5	12.5	0.0	0.0	0.0	20.9	15.7	12.5	12.5	11.4	11.4	11.4	11.4	12.5	12.5	0.0	0.0	0.0	0.0	0.0	40,396	40,396	0
GWS	0.0	55.0	0.0	460.0	0.0	0.0	0.0	0.0	0.0	0.0	69.0	0.0	69.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	69.0	0.0	19,919	40,396	20,477
Super Tab	345.0	0.0	0.0	0.0	69.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	26,656	40,396	13,740
GROSS PROFIT																											**1,981,611**		

NOTE: Some data has been modified and some omitted

Figure 2: Fonterra Kapuni LP model

2) *Exploit the constraint.* This step exploits the constraint by squeezing every bit of capability out of the constraint. As the constraint is what limits the system's throughput, it has to be worked to the maximum. At the time of the analysis, the mills were scheduled based on product demand and the RM push. The objective was to meet demand but there was a desire to produce products that were the fastest to make to keep up with the incoming RM. There are guidelines of how the mills should be operated but actual decisions predominately based on the operators' judgement. The result was that everyone had a different understanding of what is the best method to operate the mills. With all attention focused on the constraint and throughput maximised through it, operators would not have a dual focus and decisions would be made based on the effect those decisions had on throughput. Therefore, in order to maximise throughput through the constraint it is necessary to find suggestions for improvement to the reasons for lost time on the constraint and to create a product mix that generates the most throughput per constraint minute.

   Every minute the constraint is not running for whatever reason, this equates to lost throughput. Therefore any reason for this lost time should be identified and suggestions to prevent this lost time on the constraint. Examples of reasons for lost time at Fonterra Kapuni were breakdowns, slow changeovers, ineffective buffers, "push" production system and work place stress. Suggestions for improvement of these reasons for lost time are to:

   - Have a focus on throughput generation
   - Reduce operational variability
   - Create a pull system from the bunkers
   - Buffer strategically

   In order to achieve these improvements a product mix needs to be created based on the amount of throughput the constraint generates per minute. From the product mix a production schedule can be created, buffers can be placed strategically, a pull system can be created from the bunkers and since the focus is on throughput generation with attention on the constraint, operational variability can be reduced. Since it is not possible to make every product it is necessary to decide which products should be given higher priority. Therefore, when calculating the product mix, the products should be prioritised based on the products that provide the highest contribution per constraint minute (Schragenheim & Dettmer, 2001). The products are prioritised as it is not possible to make everything demanded. To calculate the product ratio, compare each product by the throughput they generate per unit based on the constraint. The calculation for this is (Cox et al., 2003):

   Throughput per constraint unit = (revenue per unit – raw material expense)/ constraint time per unit

   The calculation shows the product with the highest throughput per constraint unit. The LP solution is based on the premise that all products should be made until the constraint becomes binding. When the constraint becomes binding the quantity to produce is reduced on the products that generate the lowest throughput per constraint minute. In Figure 2 L, N and O (T and O are interchangeable as they have the same rank) are the products that the LP has chosen to reduce the quantity to make in order to maximise profit.

3) *Subordinate other activities to the decision made in step 2.* The third step subordinates the non-constraints to the pace of the constraint. The purpose of subordinating is to ensure the constraint is never idle or starved because a non-

constraint is working on another task. Scheduling and buffers were also used in this step to ensure the constraint performance is never threatened. Buffers placed at the control points pull material through the plant. Goldratt's *drum-buffer-rope* can be used to help plan the transformation process by buffering strategically in order to generate throughput. The rate of throughput is dictated by the "drum" (BM3); where the "buffer" protects the throughput of the system against any disruptions so the constraint is never starved or sitting idle; and the rope is a communication mechanism from the constraint that limits material released into the system at the rate that best suits the constraint. Therefore at Fonterra Kapuni, any spare capacity should to be used to build up modest buffers to ensure the constraint is never sitting idle or starved. For example, the crystalliser has quite a bit of slack (21876 minutes) providing an opportunity to place a constraint buffer before the constraint to ensure it is never starved. I also recommended a space buffer after the mills to protect against downtime downstream to ensure the constraint is never blocked and a shipping buffer at the end of the process to protect due dates. To protect against the seasonal nature of the raw material supply, Fonterra Kapuni have large bunkers at the end of primary evaporation and crystallisation processes. These bunkers are where material would be released at the pace of the constraint therefore creating a pull system from the bunkers. The key here is to maximise throughput through the constraint and all other activities subordinated to this so that the constraint will perform at its best. By not focusing on the constraint this part of the process's performance will not increase as the constraint is the "weakest link", production rate will decrease and the bunkers will no longer be effective as they will fill up.

4) *Elevate the Constraint.* This is when capacity is added but only if necessary and desired. It may not be necessary once the previous recommendations have been implemented. However, if the constraint has not been broken, steps 2 and 3 were not sufficient enough to eliminate it, and then the constraint needs to be elevated if higher performance is desired. At this point it is not possible for the constraint to perform any better. There are alternative ways to increase the capacity of the constraint. The investment may be in time, energy, money, or other resources (Dettmer, 1997). Therefore step 4 is last step to break the constraint, so that throughput can be increased and the system is able to perform at a higher level. The 'Sensitivity Analysis' is useful here to find out what it is worth to add extra capacity and how much extra resource is needed. The focus here is on the constraint as additional capacity is worthless since other resources have spare capacity.
5) *Go Back to Step 1 but Beware of Inertia.* The caution to be aware of inertia refers to when a constraint is broken, not to get too comfortable because the cycle never ends. Also, when change is made to an organisation there are subsequent effects that mean old policies may no longer be valid. These old policies need to be reviewed as they can cause a system constraint because the original reasons as to why the policy exists have long since gone but the policy still remains. Once the old policies have been reviewed the first step can be applied – identify the new constraint. An example of a policy that will need to be changed is to process all incoming RMs as this policy prevents the focus on throughput generation. If a new policy was 'all actions or decisions carried out must be in terms of increasing throughput or maximising the constraint performance', Fonterra Kapuni is likely to process more RMs. They are likely to process more RMs as they have increased the constraint's performance that is limiting the organisation as whole to improve its

performance. With regard to Fonterra Kapuni identifying its new constraint, the LP spreadsheet can be amended, and resolved, enabling the process to start from step 1 to allow a process of continuous improvement to occur. This application of LP required TOC's *thinking process tools* to assist in identifying the constraint. A future constraint may be outside of the manufacturing process and as result thinking process tools will play a larger role rather than using LP.

3 Discussion

The application showed how TOC and spreadsheet LP could be used to provide solutions by a non-specialist. In fact, I was very much a new practitioner in the use of both models in an applied environment. TOC provided the framework and guidance for me to operate where LP assisted in the processing of the quantitative data. The limitations from the application, as well as implications, will be discussed then concluding with recommendations for future research.

3.1 Implications

This application of TOC, combined with spreadsheet LP, implies that a non-specialist can use these methods to produce significant organisational improvement recommendations. I found at first the idea of using a mathematical optimisation tool somewhat daunting due to the lack of experience I had with such tools. In the application of the tool, and with aid of the papers I mentioned earlier, I found the application relatively straightforward and invaluable in processing the data.

Since the recommendations were not implemented at the completion of this research, it is possible to challenge whether the recommendations would improve an organisation's performance as no results have been generated to support the statement. However, a presentation was held at Fonterra Kapuni to engineers, sales staff, operators, managers etc. where it was agreed that the dual focus was a core problem and the mills were the constraint. The importance of focusing attention on the constraint was agreed to be a logical move although there was some debate over the generation of throughput. Fonterra Kapuni viewed throughput as the generation for product rather than the generation of product sold. TOC defines throughput as the rate at which the system generates money through sales (Goldratt, 1992). It is important that there is agreement on this definition as the focus would no longer be on producing products that generates money but producing products with fastest production times where the product would likely end up in storage.

There were four main lessons learnt from this application of the TOC to Fonterra Kapuni:

- Buy-in and commitment from managers is necessary.
- A system view is necessary in the application of the TOC.
- It is best for employees dealing with the problem to be involved in identifying what to change and the creation of the solution.
- A linear programming model should be used in combination with TOC to get the best results.

These lessons learnt are all in the TOC literature, but in this research I found out for myself the significance of these learnings through the process of double-loop learning. The importance of buy-in and commitment from managers is because they set the measures to how their employees perform and the level of commitment to the TOC philosophy – mainly a focus on throughput generation. The system view is necessary to identify the one main constraint limiting the organisation from improving its performance. Without the system view the constraint would not have been so identified and the uncertainty behind the focus would not have been identified as the core problem. Employees need to be involved in the entire change process as they are involved with the problem and it is their intuition that is likely to provide the best solution. As an outsider I found it very difficult to capture the employees' knowledge of the problem. Employees expressing their own problem that exists in their reality would be a more correct representation of their reality than an outsider's perception. The best approach is to have employees create their own logic diagrams as they are closest to the problem and it also gains their buy-in to the change initiative. Finally, a linear programming model alone would not have improved Fonterra Kapuni's performance as it did not take into account the problems behind the focus. This highlights the importance of using the CRT to identify soft system type problems such as the issues underlying the uncertainty of focus.

3.2 Limitations

The main limitation for this research was the inability to implement recommendations due to time constraints of the research. However, the objective was still achieved, as solutions to the problem were provided. In addition, the application was an invaluable experience for me as I learnt how to apply TOC combined with LP to a "real organisation" where I provided solutions for improvement. There most likely are other methods or tools that could have solved the problem Fonterra Kapuni presented to me. I was biased towards TOC as this was a methodology I had been taught in my fist year of my Masters degree that I thought would best solve the problem.

3.3 Future Research

There are still only a small number of published applications of TOC in NZ manufacturing organisations. Further applications are required in order to be able to determine whether NZ manufacturing will achieve similar results as overseas organisations. In addition to applications of TOC the same can be said with TOC combined with spreadsheet LP, specifically with non-specialists. For NZ manufacturing to benefit from such an application the research should not be left to specialists alone as the easier it is to use a new or different model/tool the less opposition there is likely to its adoption.

4 Conclusion

The TOC in combination with LP has the potential to improve an organisation's performance. By applying the TOC with LP to a New Zealand manufacturing organisation, (Fonterra Kapuni), the methodology was tested on its ability to improve an organisation's performance. In this application an analysis was carried out where the core problem and constraint were identified and subsequently, recommendations were provided. Although the recommendations were not implemented at the completion of this research, the potential to improve Fonterra Kapuni's performance was identified.

A "helicopter" view is used in TOC to focus on the important without the unnecessary detail. The helicopter view was used in this research combined with the TOC methodology to identify Fonterra Kapuni's constraint and core problem that was limiting it from improving its performance. Once the constraint was identified spreadsheet LP was used to sift through the data in order to provide information for improvement with relation to the constraint. In conclusion, the combination of TOC with spreadsheet LP complimented each other in its application by a non-specialist to a New Zealand manufacturing organisation.

5 Acknowledgements

In completing this thesis, I would like to acknowledge and thank numerous people for their support during this operational research endeavour.

I have to thank my family, Associate Professor Victoria Mabin, Fonterra and FRST for providing the Enterprise Scholarship and the opportunity for me to apply this research to a real organisation. I would also like to thank my fellow MMS students and the staff from the Victoria Management School who teach on the MMS programme.

6 References

Cox, J.F., J.H. Blackstone and J.G. Schleier. 2003. *Managing Operations: A Focus on Excellence,* North River Press, Great Barrington, MA.

Dettmer, H.W. 1997. *Goldratt's Theory of Constraints: A Systems Approach to Continuous Improvement*, ASQ Quality Press, Milwaukee, WI

Goldratt, E.M. 1990a. *The Haystack Syndrome: Sifting Information Out of the Data Ocean*, North River Press, Croton-on-Hudson, NY.

Goldratt, E.M. 1990b. *What Is This Thing Called Theory of Constraints and How Should It Be Implemented?*, North River Press, Croton-on-Hudson, NY.

Goldratt, E.M. 1994. *It's Not Luck,* North River Press, Great Barrington, MA.

Goldratt, E.M. 1997. *Critical Chain,* North River Press, Great Barrington, MA.

Goldratt, E.M. and J. Cox. 1992. *The Goal,* 2nd rev. edn., North River Press, Croton-on-Hudson, NY.

Goldratt E.M. and R.E. Fox. 1986. *The Race,* North River Press Croton-on-Hudson, NY.

Ho, M. 2001 "An Exploratory Study into the Implementation of TOC on New Zealand Manufacturing", MMS thesis, Victoria University of Wellington, Wellington.

Hussey, J. and R. Hussey. 1997. *Business Research: a practical guide for undergraduate and postgraduate students*, Macmillan, London.

Lepore, D. and O. Cohen. 1999. *Deming and Goldratt: The Theory of Constraints and the System of Profound Knowledge*, North River Press, Great Barrington, MA.

Mabin, V.J. 1995. *Using Spreadsheet Optimisation Facilities as a Decision Aid, Within the Theory of Constraints Framework* Graduate School of Business and Government Management, Wellington, Working Paper, May: 9/95

Mabin, V. J. and S.J. Balderstone. 2000. "The World of Theory of Constraints: A Review of the International Literature", St. Lucie Press, Boca Raton, FL.

Mabin, V. J. and S.J. Balderstone. 2003. "The Performance of the Theory of Constraints Methodology: Analysis and Discussion of Successful TOC Applications", *International Journal of Operations and Production Management,* **23**:8:568-95.

Mabin,V.J and J.O, Gibson. 1998."Synergies from Spreadsheet LP used with the Theory of Constraints: A Case Study", *Journal of the Operational Research Society*, **49**:9:918-927.

Scheinkopf, L. J. 1999. *Thinking for a Change: Putting the TOC Thinking Processes to Use,* St. Lucie Press, Boca Raton, FL.

Schragenheim, E. 1999. Management Dilemmas: The Theory of Constraints Approach to Problem Identification and Solutions, St. Lucie Press, Boca Raton, FL.

Schragenheim, E. and H.W. Dettmer. 2001. *Manufacturing at Warp Speed: Optimizing Supply Chain Financial Performance,* St. Lucie Press, Boca Raton, FL.

Surfacing Simmering Discontent

Natalie J. H. Bourke, Emma J. Daniell & Tonje Hinze
Victoria Management School
Victoria University of Wellington
New Zealand
bourkenata@student.vuw.ac.nz
danielemma@student.vuw.ac.nz
hinzetonj@student.vuw.ac.nz

Abstract

> [Organizations] often maintain a myth of harmony when everyone knows there is simmering discontent, […] it is often only when pressure is applied to the business that the conflict comes into the open – and that is the worst possible time to address it
>
> (Gascoigne, 2003, p14).

As managers, we are making decisions everyday. The pressing need for fast answers makes it easier to settle with what is only a satisficing response to a situation. This paper explores how the Theory of Constraints can be applied when solving a problem of cleanliness within a shared office environment. More specifically, it shows how the Theory of Constraints can be used to identify the root cause of the problem through surfacing and challenging assumptions.

This study applied the Thinking Processes tools to the Victoria University of Wellington's post-graduate study area at Pipitea Campus. It offers ways to resolve the problem as well as plausible and pragmatic suggestions for instigating improvement within the organization. The application of the Theory of Constraints prompted the realisation that there is a need for logical thinking processes when solving decision making problems.

1 Introduction

On a daily basis we are confronted with problems which need to be solved using our innate decision making tools. It is easy to become complacent, due to this type of regularity, and make choices using irrelevant and illogical thinking processes; to make connections between ideas and actions that do not make sense when properly considered. For managers who are making decisions and solving problems it is imperative that the processes they use do not encourage 'programmed thinking' (Patrick, 2005). It is through adopting the tools utilised through the Theory of Constraints that managers (and students alike) can begin to break through the routinised and taken-for-granted thinking processes used to solve problems.

When approaching our 'problem', the cleanliness of Level One in Rutherford House, the close proximity of the researchers to the problem and predetermined judgements as to the appropriate solutions needed to be considered when framing the situation. What became apparent from the beginning were the assumptions of how the situation was labelled a problem and thus the implications of this needed to be kept in mind when considering the actual core conflict.

The following exploration into Victoria University of Wellington's (VUW) post-graduate study area at Pipitea Campus identifies the core problem as to why the goal of the organization was not being achieved. Using the Thinking Processes approach, we establish what needs to be changed, what this needs to be changed to, and how the change can happen.

2 The Situation

The goal of VUW is to provide an effective learning and working environment for post-graduate and post-experience students, as well as staff and tutors who are also using Level One. It became apparent that the allocated workspace for post-graduate students was less than conducive to effective learning. This was affecting the ability for students to work productively in the study environment which had been provided.

The post-graduate students are sources of income for VUW and for that reason are consumers of the service that VUW provides. Consequently, it is imperative that VUW maintains a high standard of service to ensure students return and their grades and successes adequately reflect the university.

Previous to this study there had been numerous complaints circulating among users of Level One, and some of these comments had filtered through to management. As well as management being notified, the situation had intensified and there were groups emerging within Level One creating a sense of mute conflict.

Initially the situation was framed as a systemic problem. Within the frame of use it was established that the key constraint within the system were the cleaners, responsible for emptying rubbish bins, vacuuming and replenishing toilet paper and hand towels. However, it so transgressed, that as a result of interviews with cleaners the problem was reframed as being caused by both the action and inaction of students. It became apparent that the cleanliness of Rutherford House Level One was not a system, but rather a situation that was commonly recognized in many workplaces. The symptoms present in this situation can also be seen on other levels, not just with students but staff also, and thus we recognise that this problem cannot be isolated to just this one area. It was also apparent that the situation could exist in many other organizations that have communal facilities.

Through talking with a number of people within the Level One community, there were four key issues that arose:

1. The lack of cleanliness of the facilities (kitchenette, computer labs and bathrooms);
2. Some students create their own 'territories', by leaving their personal resources around computer and study areas;
3. The expectations of roles, in regards to cleanliness within the community are not explicit and thus students have varying ideas of what these are, and;
4. There is a lack of open modes of communication within the community.

The results of 45 surveys administered within the Level One community overwhelmingly reinforced the four key issues that had been identified. Following this, two students within the community and a member of staff responsible for Level One were interviewed. The four interviews revealed the conflicts that were present in each of the three common areas and prompted further exploration as to what were the reasons behind the problems that were leading to the undesirable effects.

3 The Thinking Processes Tools

Goldratt identified several tools, within his novel *The Goal* (1984), which when effectively utilised can help cut through a manager's predetermined approach to solving problems, with the intention that managers will become 'non-programmed thinkers' and subsequently develop new and more effective initiatives than they previously would have. The tools that Goldratt developed

> are comprised of five distinct logic trees and the 'rules of logic' that govern their construction. The trees include the Current Reality Tree, the 'Evaporating Cloud', the Future Reality Tree, the Prerequisite Tree, and the Transition Tree
>
> (Dettmer, 1997, p.22).

The rules used to hold the logic of the Trees together are the Categories of Legitimate Reservation (CLR). The CLR consists of "eight rules, or tests, of logic that govern the construction and review of the trees", these include "clarity, entity existence, causality existence, cause sufficiency, additional cause, cause-effect reversal, predicted effect existence and tautology" (Dettmer, 1997, p.26). The rules of CLR can be used to analyse, review and improve trees in order to broaden managers' problem solving abilities, and their ability for verbalising differing ideas as to the best action to take to reach a desired effect. Each of Goldratt's five tools of logic, the trees, can be used independently but gain strength when combined as "an integrated 'thinking process'" (Dettmer, 1997, p.26).

The Thinking Processes are based predominantly on what Goldratt has determined to be rational thinking. He labels these modes of thinking 'sufficiency logic' and 'necessity logic'. When managers explore their decisions using sufficiency logic, which uses the key phrase of "if…then…" (Patrick, 2005), they often find that the results do not always justify the actions taken. The statement "because…" often comes out of necessity logic, extending sufficiency logic, and this can fully explain the reasons *why* an action needs to be taken. This can often help to reveal any predispositions or assumptions managers are relying on when solving a problem (Patrick, 2005).

An important idea is the iterative nature of problem solving and decision making. In this, a significant consideration to remember is that no solution or action is 'straight forward'. Using the metaphorical symbol of 'clouds', 'trees' and 'branches' denotes the underlying organic nature behind TOC, with the continual use of logical thinking.

4 Application of the Tools

4.1 Evaporating Cloud

The first means through which managers can break through their pre-programmed responses to a problem is by using The Evaporating Cloud. Necessity logic is used as a building block for this particular tool. The Evaporating Cloud is often referred to as a "conflict cloud, a dilemma cloud, or a conflict resolution diagram" (Patrick, 2005). The aim when using the Evaporating Cloud is to break through subconscious assumptions managers have, which are stopping them from finding a solution to an ongoing problem. The Evaporating Cloud is always presented in the same definitive way: In order to have [A], we must [B] (Kendall, 1998, p.38).

As a way of beginning to solve the problem, three Clouds were developed to represent separate ‘problem areas’ within the total. From these Clouds, which focused on the cleanliness problems in the kitchenette, bathrooms and computer labs, a Core Conflict Cloud was formed as a way of addressing the core conflict of the problem and subsequently why the overall goal of the community was not being achieved. The three areas were combined to develop a core conflict cloud as outline in Figure 1.

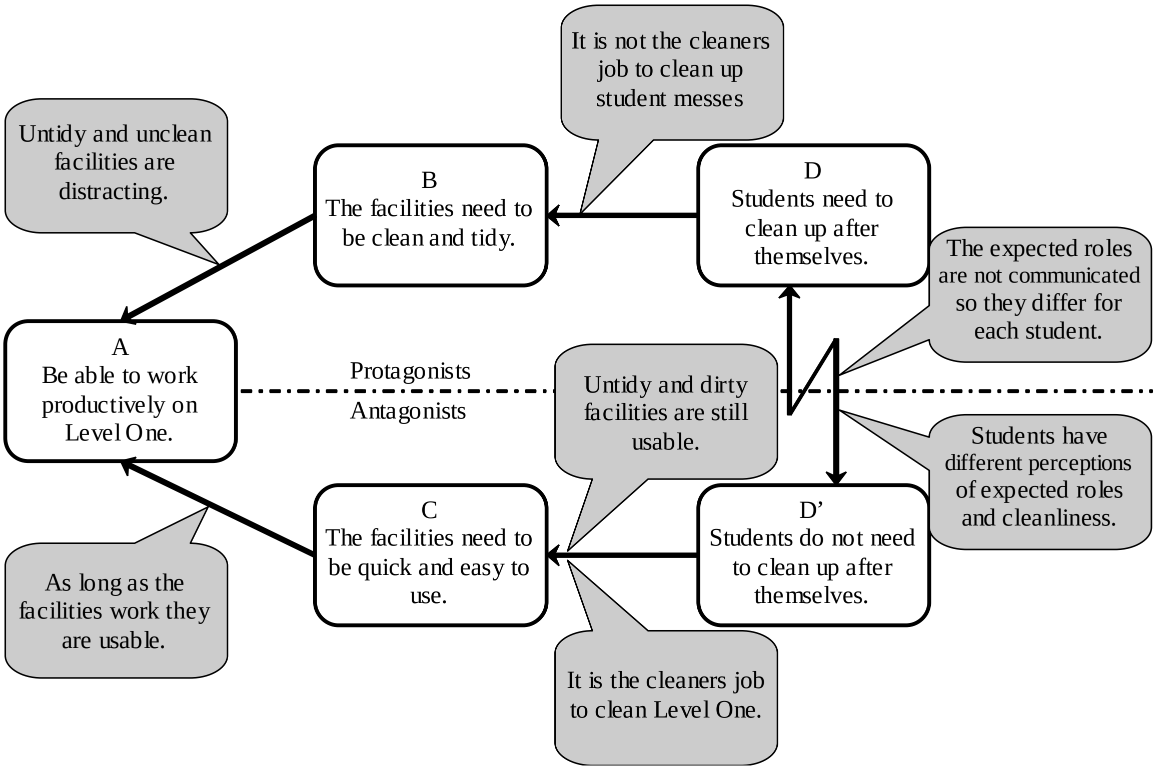


Figure 1. Core Conflict Cloud

The Core Conflict Cloud (CCC) surrounded the cleanliness of the whole Level One and it demonstrated the underlying problem that caused all of these conflicts to emerge. It reads as follows:

AB In order for the students to (A) be able to work productively, students need (B) the facilities to be clean and tidy.

BD In order for (B) the facilities to be clean and tidy, (D) students need to clean up after themselves.

AC In order for the students to (A) be able to work productively, students need (C) the facilities to be usable.

CD’ In order for (C) the facilities to be usable, (D’) students do not need to clean up after themselves.

DD’ On the one hand, students must clean up after themselves while on the other hand, students must not clean up after themselves.

D and D’ are mutually exclusive, they cannot occur simultaneously.

Key assumptions for each linkage were identified (shown in the shaded boxes). The clouds continually enabled the assumptions to be challenged. This proved to be a vital step in forming the foundation for the rest of the Thinking Processes tools.

4.2 Current Reality Tree (CRT)

The second tool, the CRT, can be formed using the Core Conflict Cloud (turned sideways) as the trunk or base of the tree. The CRT uses sufficiency logic, and is an effective tool for detailing the problem being faced by the manager. When a CRT is formed using the CCC as the base, it will become apparent to the manager whether the identified problem is in fact the core problem of the situation or whether ‘reframing’ around the problem is required. This tree begins to look at answering one of the fundamental questions of the Thinking Processes and that is ‘what to change?’ – “the one simplest change to make that will have the greatest effect on [the] system” (Dettmer, 1997, p.23).

The CRT became the most important tool. Through the construction of the CRT, assumptions were surfaced with regard to finding the solution to the problem. Initially ‘differing perceptions of cleanliness’ was identified as the core problem, but this was later found to be only an assumption. The problem was found to have grown out of the lack of role definition for the students using the facilities.

The CRT uses sufficiency logic, which is read as “If (cause) then (effect)” (Cox et al., 2003, p. 78). The arrows in the CRT demonstrate the connection between the cause and effect and indicate which direction the condition is taking. Often there are two causes that together lead to an effect. The two causes can either be dependent on each other or they can individually contribute to the effect. Two causes that are both needed to bring out an effect are linked together with an ellipse and the sufficiency logic structure is read as follows: “If (cause entity) and (entity) then (effect entity)” (Cox *et. al.* 2003, p.85).

Eight examples from our CRT illustrate how this logic is applied:

1. If (12) students need to eat and drink to sustain themselves then (15) students prepare food and drink.
2. If (55) paper towels block toilet drainage system then (58) toilets are not usable.
3. If (76) students do not remove unwanted food and containers from the refrigerator then (78) the refrigerator gets overfilled.
4. If (46) students use toilet paper and (47) toilet paper is not always replenished frequently by the cleaner then (51) toilet paper runs out.
5. If (61) some students leave personal messes in the computer labs and (62) some students leave personal resources in the computer labs then (65) cleaners are unaware of what is rubbish and what is not.
6. If (84) the refrigerator is left open and (85) student’s food must be refrigerated then (87) student’s food goes off.
7. If (106) some students have bad relationship with each other and (107) bad relationships create stress then (113) some students become stressed.
8. If (115) students need effective learning environment at university to perform and (116) Level One is an ineffective environment and Level One is the only space available for the students, then (117) students under perform.

The CRT was a huge task that was constantly evolving. However the process of creating this tree enabled an appreciation of how complex the situation really was and it provided the platform for developing solutions that broke the underlying assumptions.

4.3 Negative Branch Reservation (NBR)

In order to successfully implement new action plans, all the negative possibilities that may arise from it need to be considered so that preventative action can be taken before these occur. The NBR is an important tool in this part of the Thinking Process, where the negative outcomes can be 'trimmed' or 'pruned' from the branch through the injection of future precautionary action.

It was identified that to begin to implement a solution to the problem of cleanliness on Level One all students needed to be made aware of the situation and their role in acting out the solution. A group meeting, held during class time, was proposed as a way of addressing this need. This allowed for the opportunity of speaking to a large number of users at one time.

In order to identify any of the difficulties, or negative outcomes, that may arise from such a proposal an NBR was constructed. This helped to establish what the possible negative outcomes of the meeting could be. From this, preparation could be made for any eventuality that could arise from the meeting to ensure that preventative action could be taken to avoid any potential negative effects. Once the solution had been identified, the elements needed to implement it were considered. The NBR "takes advantage of people's natural propensity and ability to point out why something can't be done"(Patrick, 2005). Often this step is the most important when implementing a new initiative as people are inclined to voice their reservations after the fact rather than before. If as many possible obstacles are identified before the proposed action is to take place then the plan of 'how to change' will be stronger and better utilised.

Similarly to the CRT, the NBR also uses sufficiency logic when developing the negative consequences that could lead from an injection. A better understanding of all the possible problems that could follow the proposed injection could be found through an NBR. Thus, methods of preventing the negative outcomes from occurring could be established:

1. If (23) people have limited available time and (24) people may think they have to do all the cleaning then (27) people may feel too much effort is required to use Level One.
2. If (21) people feel isolated then (26) the meeting will not achieve forming a cohesive group.

The NBR specifically addressed a proposed solution and identified the possible consequences, helping to gain a better understanding of a possibly sensitive situation without naively jumping into it.

4.4 Future Reality Tree (FRT)

When addressing the issue of 'what to change?' managers need to fully examine any proposed corrective actions, these can be analysed through the creation of a Future Reality Tree. The FRT also uses sufficiency logic and helps managers to identify the future outcomes of new actions, including any negative consequences. When both these tools, the NBR and the FRT, are used in conjunction with each other the 'what to change to?' part of the Thinking Process is explored and resolved.

The Undesirable Effects (UDEs) from the CRT are translated to desirable effects (DEs) in the FRT through the exploration of these UDEs in the NBR.

The FRT was used in conjunction with the NBR to help establish a protocol of the best way to interact with the group. It was a key tool that triggered buy in from the community and thus provided the best possible immediate outcome.

4.5 Prerequisite Tree (PRT)

The PRT emerged from the resulting positive reaction received from the group meeting. This resulted in the first desired affect of enthusiastic group buy in, but a means of sustaining these effects over the long term needed to be developed. Subsequently a PRT was constructed in order to find ways in which the resulting exceptional conduct of the students using Levels One could be transformed into habitual cleaning behaviour.

The challenge was no longer in finding out ‘how to change’, but rather how to make the change last. The PRT assisted in identifying the obstacles that prevented ongoing action with regard to the solution. The aim was to create the habit of cleanliness among students and this issue was addressed through the PRT.

The PRT identifies steps which construct a plan of action and can later be translated into a critical chain in order to exploit the full potential of the proposed solutions. Obstacles that would block the implementation of each injection are uncovered that could prevent the achievement of the Ambitious Target: ‘Ensure the community remains proactive with maintaining the cleanliness of Level One’. In the PRT Target Objectives were identified that would cause the Ambitious Target if implemented.

The PRT is based on necessity condition logic when addressing the different obstacles. Because the PRT was the last tool implemented, it was easy to dismiss as a less important element of the Thinking Processes. However, through its use it was found to be a keystone to the whole integrated process whereby the cyclical nature of the Thinking Processes was realised. This, in turn, created the most effective ongoing change.

5 Present and Future Reality

Dettmer outlines the notion of the ‘Silver Bullet Fallacy’ whereby there is often not one solution that will solve the core problem and it is “much more likely that we [will] need to make several changes, in combination, to achieve our desired effect” (Dettmer, 1998, p 121). It is therefore important to introduce other measures to reinforce and ensure that the problem is correctly addressed.

The placement of cleaning tools, such as toilet brushes in the toilets, on Level One, as well as things like hooks for hanging wet cloths and tea towels has encouraged students to be proactive about taking care of the facilities themselves. A desired effect was in building a sense of ownership and respect among students for the facilities on Level One, so that taking responsibility for cleanliness became normal not exceptional behaviour. In order to prevent any unwarranted stress, friction or discontent amongst students the emphasis was placed on group effort being a key in maintaining the cleanliness of Level One.

Some of the initial verve for keeping Level One tidy has come about through the articulation of management intention to remove resources and facilities if the problems continued. Though the possibility of confiscation remains, it was intended that this would not be a distraction for the whole community. The threat of removal would, however, serve as a motivator towards preventative action (such as cleaning up ones own messes).

Cleanliness in any communal environment, such as this, is an ongoing issue. So rather than trying to enforce rules, it would be advised to promote the ‘if you make a mess, clean it’ message. Signs were erected which were constantly changed, that used humour as a way to gain continued support and buy in from the community. A community culture with the users of Level One was fostered through group gatherings that helped to create a better sense of ownership over the facilities and resources, as well as respect for other users. One has to be aware of the reasonably high turnover of occupants in Level one, simply because students stay only to complete their degrees, so effort needs to put into ensuring that the community spirit remains strong. It would be appropriate to introduce the etiquette surrounding cleanliness on Level On during the orientation for new students entering post-graduate studies so as to avoid future confusion. The reliance over the years on informal modes of information sharing among new and current students allows for too much inconsistency in the levels of responsibility each student is taking for their own cleanliness. When used in conjunction with management led directives both forms of communication can add strength to the existing culture of cleanliness on Level One.

It is intended that the habit of cleanliness among students will be further developed. Although the time spent here is prolonged for some and intermittent for others, it is obvious that in order for the space to be an effective learning environment there needs to be a cohesive community culture. Situations such as cleanliness can aggravate the community, so it is important that cleanliness is routine and expected from students and other users on Level One.

The fiery situation on Level One was extinguished. Tensions growing directly from the core problem as well as peripheral issues were relieved by bringing the issue out into the open offering students the ability to be part of solving the problem.

Because of the nature of Level One and the inclusion of both part time and full time students, it was difficult to convey the message to everyone. It has since been noticed that when a mess is left in a common area, students are quick to sort out who left it and also inform the perpetrator of the new ‘policy’; whereby everyone contributes and cleans up after themselves.

It is vital to see the link between this study and other organizations, whereby the small issues often escalate into more challenging situations. This work demonstrated how the often simple nature of solving problems, where something as effortless as communication, was needed in place of fire fighting (Bohn, 2000), and the result was a much more favourable outcome. It is also obvious to see the reduced costs in approaching a problem in this way.

It is proposed that this policy will continue to be communicated among students in the years to come so that the problem does not escalate into similar proportions. Therefore initial optimism about the future on Level One remains high.

The cost of the resources needed in order to utilise the tools and the time taken, are minimal in comparison to that which would be exhausted on fire fighting. Organizations must be committed to understanding the core problem in order to establish valid and applicable solutions. We found that the Theory of Constraints was an effective and valuable tool when surfacing the simmering discontent on Level One.

6 Bibliography

Bohn, R. 2000. “Stop Fighting Fires”. *Harvard Business Review*, (Jul-Aug): 83-91.

Cox, J., Blackstone, J. & Schleier, J. 2003. *Managing Operations: A focus on Excellence*, North River Press, Great Barrington, M.A.

Noreen, E., Smith, D. & Mackey, J. 1995.*The Theory of Constraints and its Implications for Management Accounting*, North River Press Publishing Corporation, Great Barrington, MA.

Dettmer, H. 1997. *Introduction to the Theory of Constraints. A System Approach to Continuous Improvement*, ASQC Quality Press, Milwaukee.

Gascoigne, C. 2003. “How to deal with trouble and strife: Family business”, *Financial Times (London) Limited*, (Jan 23).

Goldratt, E. 1989. *The Goal: A Process of Ongoing Improvement,* Gower, Aldershot, England.

Kendall, G. 1998. *Securing the Future: Strategies for experimental growth using the TOC,* St Lucie Press, Florida.

Patrick, F. 2005. "Achieving Focused Performance Through the Application of the Theory of Constraints, [Web Document], Available: www.focusedperformance.com, [April 14, 2005].

Unit Commitment at Southdown Power Station

A. Downward
Department of Engineering Science
University of Auckland
New Zealand
adow031@ec.auckland.ac.nz

Abstract

This paper examines the electricity market from the point of view of a small generator. We consider both price taking and price making scenarios and investigates how best to operate the plant. The price taking scenario models price processes using a time-inhomogeneous Markov chain within a stochastic dynamic programming recursion. This creates a policy, which specifies the best decisions in expection, for all possible states. In the price making scenario, we perform market simulations, using offers and loads, to calculate the effect on price that this generator can have and whether this should be taken into account when making generation decisions.

1 Background

1.1 Southdown Power Station

Southdown power station, owned by Mighty River Power, is located in the Auckland suburb of Southdown, near Otahuhu. This plant supplies up to 118MW at full load, which means it can supply approximately 2% of the country's total peak demand of 6500MW. (Mighty-River-Power 2005)

1.2 Combined Cycle

Southdown is a combined cycle plant, consisting of two gas turbines and a steam turbine. The steam turbine is mainly powered by waste heat generated by the gas turbines. Heat is produced as a by-product when the gas turbines burn fuel to produce electricity; to be efficient, the waste heat is used to boil water to produce low pressure steam. This steam is then pressurised and fed into the steam turbine. (Mighty-River-Power 2005)

Gas Turbines The two gas turbines at Southdown can be operated independently of each other – this allows for greater flexibility in the plant's configuration. As the gas turbines are ramping they use 12% more fuel, per megawatt of power they produce, than in steady operation. This means that they are not as efficient during the ramp up process as when they are running at a constant generation.

There is no economical way to continuously run the gas turbines at partial loads. This means that the turbine must either be at full power or off – unless it is ramping.

Steam Turbine The steam turbine is powered by the heat of the gas turbine and thus it costs almost nothing to run. This means that the steam turbine should always be running in conjunction with the gas turbines. The steam turbine needs a steam pressure of 45 bar to operate; it takes approximately one hour to achieve this pressure when running one gas turbine.

Duct Burners The steam turbine's power output can be increased if the duct burners are switched on. These burn gas directly within the ducts leading to the steam turbine. Using these reduces the efficiency of the plant somewhat, however it boosts the amount of power generated. There are two sets of duct burners, one set in each of the ducts connecting the gas turbines to the steam turbine. They can be operated independently of each other. If one gas turbine is operating, only one set of duct burners can be used. However, if both gas turbines are being used, either or both of the duct burners can be used, which allows a little flexibility in the power output. Switching the duct burners on and off is a slow manual process and it is desirable that they are not switched on and off too frequently.

1.3 Electricity market

On October 1st 1996, a reformed New Zealand Electricity Market began operations. Prior to this, there had existed a monopoly in all the main areas of electricity delivery – generation, transmission, distribution and retailing. The new wholesale electricity market, known as the NZEM, consists of a number of generators and retailers, creating competition in the sector. (Wikipedia 2005)

Each day is divided into 48 half-hour trading periods. In each period, the generators offer electricity into the grid at injection points and the retailers bid for electricity off-take at grid exit points. The injection and exit points on the grid are known as nodes. The transmission grid consists of 244 nodes across the country, where generators and retailers sell and buy electricity respectively. The price of electricity varies over the grid; each node experiences a different price of electricity, which depends heavily on the offers of the generators and the bids of the retailers. The price of electricity also changes over each day; this is due to changes in demand and changes in the generator's offer stacks. (Wikipedia 2005)

1.4 Offers & Dispatches

Each electricity generator connected to the grid is required to submit an offer stack containing up to five tranches, for each trading period. Once all offers have been made, an optimisation program, minimising the cost of electricity, is run to calculate how much each generator should produce and how much they will be paid.

Offer Stack Each generator's offer must be formatted in a specific way; they may submit the offer in up to five tranches. Each tranche is an offer of a quantity of electricity at a chosen price. The tranches must be ordered in the stack such that the price is always non-decreasing. Figure 1 gives an example of a five tranche offer stack, and shows how the generator could be dispatched. The dispatch can occur anywhere on the offer stack and the generator will be paid the price, p/MWhr, for the total quantity, q, of power for which it is dispatched. The area of shaded region ($q \times p$) is the amount the generator will be paid per hour.

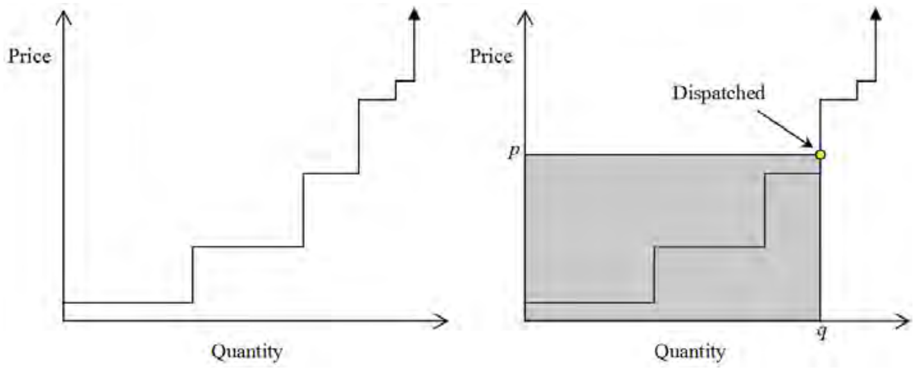


Figure 1: Example of Five Tranche Offer Stack and Possible Dispatch

The price of electricity begins to rise very steeply as demand approaches capacity. This offer structure can cause very high electricity prices in times of shortages.

Price Takers and Price Makers In the electricity market there are two types of generator – price takers and price makers.

Price Takers Price takers tend to have a small capacity compared to the total size of the market; these generators' offers are unlikely to affect the price of electricity at other nodes in the system. As competitor generators in the market are not impacted by a price taker's offer, it is a valid assumption that the price taker will not influence any competitor's offer strategy.

Price Makers Price makers, on the other hand, tend to have large generation capacities. These generators have significant impact on the clearing price, at their node and others. Price makers do not always have to be large generators; if the demand for electricity is close to the capacity of the network, even a small generator can affect the price – in so much as the price of electricity increases markedly for small changes in supply, as demand approaches system capacity.

1.5 Scheduling, Pricing and Dispatch

The Scheduling, Pricing and Dispatch model is used by Transpower to calculate the price of electricity at all the nodes on the grid. This optimisation model tries to minimise the cost of electricity, based on the generators' offers, while taking into account all the constraints of the physical network. The physical constraints include, line capacities and Kirchoff's laws, which are electrical laws. Every five minutes, the SPD optimisation model is rerun and the price of electricity at all the various nodes is recalculated. (Transpower 2005)

2 Markov chain price taking model

An important element in making profitable generation decisions is the ability to anticipate the price of electricity in future trading periods. As it is impossible to perfectly predict the price of electricity in future periods, a model to estimate future prices is required. The model in this chapter assumes that the quantity Southdown offers has very little impact on the price they or other generators receive.

2.1 Markov Chains

A time-inhomogeneous Markov chain was used to approximate the distribution of prices throughout each day. The transition matrices for this were generated based on historical data.

There are 336 trading periods in a week. When attempting to fit a time-inhomogeneous Markov chain for price to the trading periods, it is useful to divide the week into two types of day – weekdays and weekends. The weekdays can be approximated by a set of 48 weekday transition matrices, and weekends can be approximated by a set of weekend matrices. This assumes that the electricity prices for all weekdays follow identical Markov processes; an analogous approximation is made for weekend periods. The difference between the average prices over a weekday compared to a weekend shown in figure 2 supports the use of separate price processes.

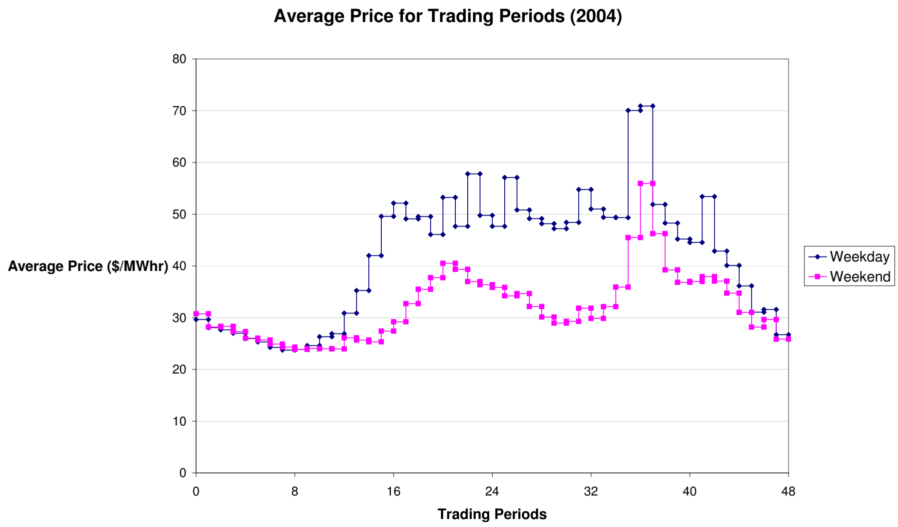


Figure 2: Price Trends within a Day

Price Bins When choosing how to set up price bins, a number of options need to be considered. The first is how many price bins are needed for approximating the price process. The second question is where the bin boundaries should be placed. These two options are crucial in trying to reliably model the change in price over the day.

The conditional expectation of each price bin is recorded; this is the mean of all prices which fall into that bin. This conditional expectation is used instead of the midpoint of the price bin as it takes into account the asymmetrical distribution of prices within the bin. Asymmetry would be expected as some prices are more common than others, skewing the value of the bin. This represents the expected value of a price given that the price is within that bin.

Historical Prices Using historical prices, and a set of bin boundaries, it was possible to generate a Markov price transition matrix for each of the 48 weekday and 48 weekend trading periods. From the data, conditional expectations for each period and price bin combination are calculated.

2.2 Stochastic Dynamic Programming

States of the Dynamic Programme The dynamic programme has three state variables:

- The day and period.
- The price-bin in which the price fell in the previous period.
- The configuration of the plant at the start of the period.

With these state variables, the optimal decision policy can be determined.

Implementing Markov Chain for Prices The stochastic element of the dynamic programme comes from using a Markov chain to approximate the price process. This means that in the dynamic programme's backward recursion, all the possible price transitions need to be considered in determining the optimal decisions.

Modelling Plant Configuration Each possible plant configuration is represented as a state of the dynamic programme, these state are linked together to represent possible ramping of the plant. For example, the state representing both gas turbines on with no duct burners, has five possible ramping decisions, which relate to which equipment is switched on or off. Each of the ramping decisions is a state in the dynamic program, and these in turn have their own subsequent decisions.

Creating a Policy for Southdown

- $V(t, i, c)$ is the expected value of being in period t and price bin i, with the plant configuration in state c.
- P_{ij}^t is the probability of the price transition starting from bin i in period t into bin j in period $t+1$.
- S_c is the set of all possible subsequent configurations, following c.
- g_{cd} is the average power in MW generated, if at the start of period t the plant is in configuration c, and it ramps into configuration d over the period.
- h_{cd} is the cost of the transition from configuration c to configuration d over the period.
- p_i^t is the conditional expectation of price within price bin i at time t.
- n is the number of price bins.

In order to formulate this problem a backward recursion must be used. This is due to the price of electricity being revealed over time.

$$V(t,i,c) = \max_{d \in S_c} \left\{ \sum_{j=1}^{n} P_{ij}^{t} \left(g_{cd} \times p_j^t - h_{cd} + V(t+1,j,d) \right) \right\} \tag{1}$$

Equation 1 defines the recursion that must be solved to create a policy for Southdown's generation decisions. The timeframe of the recursion is taken to be 336 consecutive periods, one week; this consists of the first 240 periods being treated as weekdays, and the remaining 96 periods being treated as weekends.

2.3 Results

In order to test the quality of the policies generated from the stochastic dynamic programme, it is necessary to define some benchmark plans which the profit can be compared against. Specifically, an upper and lower bound on profit are necessary in order to ascertain exactly how good a certain policy is.

Upper Bound In order to determine how well a policy is performing, it must be put in perspective. This can be done by creating an upper bound on profit. One way to create this is to use perfect foresight, in other words, to optimise Southdown's decision for each period, based on the knowledge of future prices; there is no stochasticity involved, the result is purely deterministic. This results in the maximum profit which can be made; it is, of course unattainable, as no process with stochastic events can be predicted with total accuracy.

Lower Bound It is also useful to place a lower bound on the results. A way to do this is to apply a simple optimisation policy to the decisions, which could be considered adequate, but not exceptional. This allows one to see how much better the more advanced optimisation policy performs, and even to see whether the method has merit.

In Sample and Out of Sample Testing When testing the performance of an optimisation model under uncertainty, it is useful to perform in sample and out of sample testing. An in sample test assesses how well the model is able to represent the underlying data; whereas out of sample testing can assess how well the policies perform when the data used to create the model are independent of the data the model is tested on.

Comparison These four methods above were compared to analyse the effectiveness of the Markov chain based method, the prices which occurred during the period between April 1st and June the 30th was used for this comparison:

To calculate the profit of the method with perfect foresight, a dynamic programme was run without any random elements as the price of each period is known. The profit from each period was calculated and recorded and converted to a running total over time.

For the in-sample test, eight-bin Markov price transition matrices were calculated based on prices which occurred between April and June 2004. Whereas for the out of sample test, eight-bin Markov price transition matrices were calculated based on historical prices between April and June 2002. Policies were created using this data, by using the recursion from equation 1. These policies were applied to days between April and June 2004, by revealing the price data over time and carrying out the action specified by the policy.

The lower bound was created by optimising the decisions each day, based purely on the price realisation on the previous day. In other words, a plan is created which would be optimal if the previous day's prices occurred in exactly the same manner on the current day.

Figure 3 shows the profit generated over the three month by each of the decision methods. The chart shows that the policy from the out of sample stochastic dynamic programme is considerably more profitable than the optimisation method which assumes the previous day's prices. However, as would be expected, it is not as good as the optimal plan, which is an unattainable upper bound. The positive aspect is that the out-of-sample profit is only 9% less than the in-sample profit. The quality of the in sample solution is not degraded significantly when creating the out of sample solution, even though underlying Markov chain matrices are altered to reflect a previous year.

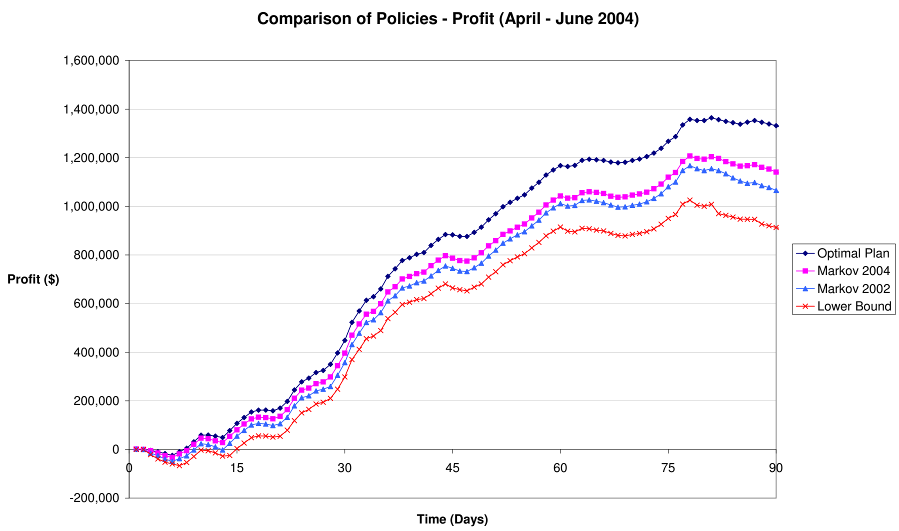


Figure 3: Comparison of Profit from Different Policies

3 Southdown's Effect on Electricity Price

3.1 Market Distribution Function

The market distribution function or ψ function was developed by Anderson and Philpott(Anderson and Philpott 2002). The market distribution function $\psi(q, p)$ represents the probability that a generator is not fully dispatched when the generator offers in q MW of electricity at a price of $\$p$. As Southdown is a relatively small generator, for the purposes of this analysis, we will assume Southdown has no impact on the decisions of its competitors.

3.2 Boomer

In order to generate market distribution functions for Southdown, one can simulate the market under various scenarios.

Boomer is a simulator of the New Zealand electricity market. It is able to process demand information for each node as well as each generator's offer stack to simulate prices. Effectively, it is a smaller version of the Scheduling, Pricing and Dispatch optimisation model, as Boomer only contains 20 generator/demand nodes. Boomer generates AMPL/cplex model and data files, which are optimised to calculate the prices at each node.

As price and offer quantity are continuous variables are continuous, they must first be discretised into a quantity/price grid, so that Boomer can handle them. Boomer then runs simulations for each quantity of electricity that Southdown can offer and returns a price.

We used Boomer to investigate what, if any, impact Southdown's offer has on the price of electricity.

3.3 Dynamic Programme Recursion

Unlike the price taking dynamic programme, the price making model does not have a price state variable. Equation 2 shows the recursion used.

$$V\left(t,c\right)=\max_{d\in S_c}\left\{g_{cd}\times E\left[p^t_{g_{cd}}\right]-h_{cd}+V\left(t+1,d\right)\right\} \tag{2}$$

Where $E\left[p^t_{g_{cd}}\right]$ is the expected price that will be received if g_{cd} MW is offered into the market during period t. These parameters for the recursion are generated using Boomer.

As this recursion does not depend on any price states, a fixed plan is created instead of a policy. This means that nothing observed during the day will affect the dynamic programme's decisions.

3.4 Results

Comparison of Price Making and Price Taking Plans. A comparison of price making and price taking plans will show whether the impact Southdown has on the price of electricity will significantly change their optimal offer strategy. To test this, an empirical distribution for demand was set up using historical data, and three sets of offers stacks (from all the generators) were taken to be different offer realisations. Each demand/offer pair became an equally likely realisation, which was run through Boomer to find the prices which would have eventuated under that scenario. From each pair, a distribution for price was able to be built up for each period and each of Southdown's offers. This set of distributions was able to be run through the dynamic programme to develop an optimal plan (in expectation) for the price making situation. We, however, also created a plan assuming that Southdown did not affect the price of electricity. This was done by taking the price distribution for each period based on the 46MW Southdown offer; this meant that Southdown's offer had no impact on the price. The revised dynamic programming recursion is shown in equation 3.

$$V\left(t,c\right)=\max_{d\in S_c}\left\{g_{cd}\times E\left[p^t_{46MW}\right]-h_{cd}+V\left(t+1,d\right)\right\} \tag{3}$$

Figure 4 shows the two plans. In the price making plan, less electricity is generated as the plan takes into account the drop in price Southdown would receive if they were to increase their offer. The price taking plan does not consider the potential drop in price.

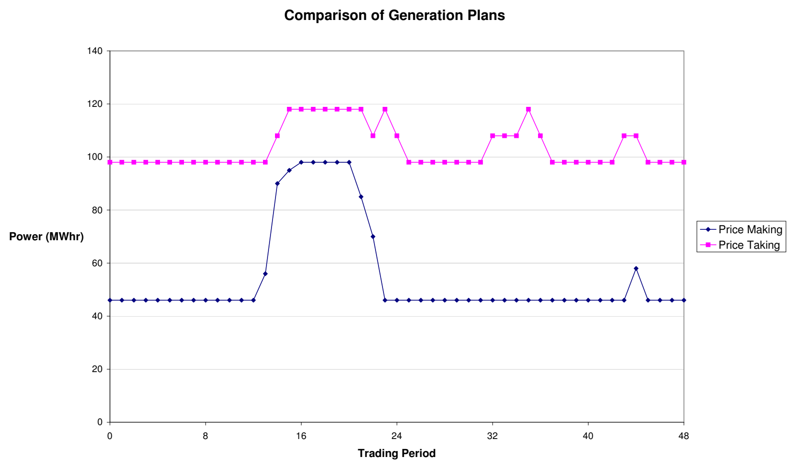


Figure 4: Price Making and Price Taking Plans

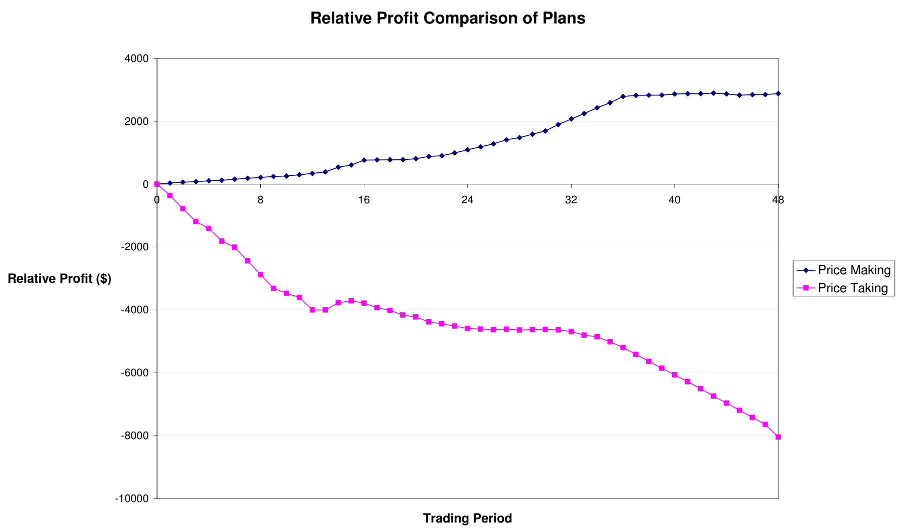


Figure 5: Profit from Plans Normalised against Actual Profit on December 8

Out of Sample Test These two plans were then compared on another day, which was not included within the sample data used to construct the plans; this is an out of sample test. Figure 5 shows the profit Southdown would make on this day, this has been normalised against the profit Southdown actually made. Therefore a positive result means the plan performed better and a negative result means the plan performed worse. The price making strategy does not offer as much electricity as it knows that the more it offers, the lower the price it will receive. For this reason it performs significantly better than the price taking scheme.

From this, we can see that Southdown can in fact be a price maker. The very fact that different offer strategies are created when comparing price making and price taking, shows that Southdown's offer must have some effect on price.

4 Conclusions

This project has found that treating the prices which Southdown received to be based upon a time-inhomogeneous Markov chain is a reasonably valid approximation. We found the plant is able to ramp appropriately when observing prices, even when the Markov transition matrices are created from a different year. The stochastic dynamic programme policy also performed markedly better than a simple optimisation based on the previous day's prices.

In comparing the different optimal plans for price making and price taking, there was a clear difference in the optimal offer strategy, from this we can state that Southdown is at least sometimes a price maker, affecting the price at their node. However, more investigation is required to see if this is always true.

Acknowledgments

I would like to thank my supervisors Professor Andy Philpott and Doctor Golbon Zakeri, for their guidance and support throughout this work. Their tireless efforts are much appreciated.

I also would like to thank Mighty River Power for giving me the opportunity to create a model of their plant; without the data which they provided, we would not have been able to create such an accurate model, and none of the results would be reliable. I would like to specifically thank Vinesh Naicker who provided much of the information to us.

References

Anderson, E., and A. B. Philpott. 2002. "Optimal offer construction in electricity markets." *Mathematics of Operations Research* 27:82–100.

Mighty-River-Power. 2005. Co-Generation. World Wide Web, `http://www.mightyriverpower.co.nz/generation/aboutus/cogeneration/default.aspx`.

Transpower. 2005. Scheduling and Dispatch. World Wide Web, `http://www.transpower.co.nz/?id=4474`.

Wikipedia. 2005. New Zealand Electricity Market. World Wide Web, `http://www.answers.com/topic/new-zealand-electricity-market?method=5&linktext=New%20Zealand%20Electricity%20Market`.

Strategic Production Plan Model for the Hunua Quarry

Oliver L. Tompkins
Department of Engineering Science
University of Auckland
New Zealand
olivert@ihug.co.nz

Abstract

In early 2004 Winstone Aggregates approached the University of Auckland about developing a strategic model for decision planning at their largest quarry. The main pit at the Hunua Quarry has only seven to tens years of operational life left, and the development of an adjacent pit at Symonds Hill is being investigated. Winstone Aggregates wanted a long term model to assist with evaluating the optimal way to handle the development and transition to the new pit. The development represents a significant capital commitment and brings with it associated uncertainties, this tool provides a basis with which to control some of the uncertainty.

The Strategic Model was developed in VBA and AMPL and is solved through a two stage process making use of the free solver, LPSOLVE. Each development scenario is evaluated and ranked based on NPV and a risk factor associated with the ability to handle short term variation. The model was verified with test data and has been implemented by Winstone Aggregates.

1 Overview

1.1 Problem Background

In early 2004 Winstone Aggregates approached the Engineering Science Department about developing a model to help with production decisions at their largest quarry, the Hunua Quarry. This approach was prompted because of depleting resources at the main Hunua Pit and the need to investigate when an adjacent pit, Symonds Hill, should be opened.

Aggregate products are the crushed rocks, gravel or sand produced by quarries (Winstone Aggregates, 2005). They are vital for the infrastructure of our cities and are used in the construction of roads, motorways, buildings, homes, and drainage and water treatment systems. With such a wide variety of uses, aggregate demand is high. Total aggregate protection in the Auckland Region is currently growing at a rate of 12% per annum and with the current economic boom and the recent increased investments in roading, projections are that this growth will continue (Ministry of Economic Development, 2005).

This increasing demand has placed pressure on the all the quarries in the Auckland region. Extraction per year, however, is capped by Resource Consents, and so to meet the current demand, most of the quarries are already running at, or near, their legal capacities.

In the Auckland region, Winstone Aggregates have a nearly fifty percent market share (Winstone Aggregates, 2005). The Hunua Quarry, a Greywacke quarry situated in the Hunua Ranges, is the company's largest quarry. However the main pit at this quarry now has only between seven and ten years of resources left and so Winstone Aggregates is investigating the opening of an adjacent pit which will secure the operation of the quarry for a further thirty years. Because of the pressure on supply in the Auckland region this development is not only financially important to Winstone Aggregates, but is also important to ensure an adequate aggregate supply for the region.

1.2 Quarry Development

Quarry Development in New Zealand is governed in part by the Resource Management Act 1991 (RMA). When planning the development of a new quarry, or the extension of an existing one, quarrying companies must ensure that all aspects of the proposed quarry meet the requirements of planning objectives, rules and guidelines set out in the RMA. The main objective of the RMA is to manage resources while:

> **Section 5 c)** Avoiding, remedying, or mitigating any adverse effects of activities on the environment.

The process of gaining the necessary consents can be long and exhaustive, and the onus, at all points, is on the quarrying company to provide detailed information on all aspects of the quarry. The outcome often places conditions on developments and it therefore contains a large degree of uncertainty. Any changes imposed on the development will affect the financial viability of continuing and need to be investigated.

1.3 Model Requirements

Winstone Aggregates wanted a model to help them investigate the financial and operational implications of major decisions. Two key issues were raised by them as being of particular importance:

- How should the Symonds Hill pit development be handled.
- How well will the quarry respond to future demand variations and, if required, how early do mitigating actions need to be taken.

To ensure the model would get continual use, it also needed to be simple to set up and operate, yet flexible enough to be continually updated for changing market conditions and/or legal circumstances.

Simultaneously with my work, a geological consultant was contracted by Winstone Aggregates to determine different ways in which the Symonds Hill Pit could be developed. His work focused on different contours or routes that could be taken through the pit, and also looked at different years in which it could be opened. The development options he produced make resource forecasts and the purpose of this model is to compare these development options to determine which the best is.

1.4 Model Type

My initial investigation into production planning in the aggregate industry suggested that it quite closely resembled a hierarchical production planning structure. A hierarchical structure for the analysis of a decision making process was first proposed by Robert N. Anthony (1965). He demonstrated that for an ongoing system the decision processes can be categorized into three distinct categories, with each one corresponding to different time horizons, and generally also to different levels of management. The requirements, and time frame of this model, clearly mark it as a strategic decision.

1.5 Related Work

Within the Department of Engineering Science at the University of Auckland, two previous production plan projects have been done for Winstone Aggregates. These were done by Mitchell (1997) and Elangasinghe (2001). Both of these projects, however, were tactical models whereas this project is a strategic model. For this reason, those works were only consulted for background information. No similar work within this area was found outside the University of Auckland.

Significant research has, however, looked into the usage of operational research techniques in the optimising of long term production scheduling for open pit mining. Mathematical programming is well suited to this problem (Hochbaum & Chen 2000, Ramazan, Dagdelen & Johnson 2005) and OR techniques have become common in commercial packages used for planning open pit mines. An aggregate quarry is one form of open pit mine.

Virtually all the research into this area has focused on high grade mineral extraction rather than on aggregate extraction. Aggregates are a commodity and production is not greatly affected by the market price which is fairly stable. Hence most of the research into operations research in open pit mining applications is not applicable for this problem.

2 Problem Description

2.1 Resource Forecasts

All rock in a quarry is graded into one of three quality categories based on the degree of weathering (break down due to exposure to the elements) that the rock has been subjected to. The more weathered the rock is the lower its quality. Rock in the Hunua Quarry is graded into blue for the highest quality, blue-brown for medium quality and brown for low quality.

To extract rock from a quarry, a pit is quarried along vertical faces which generally contain differing amounts of each of the three grades. It is not possible to extract just one rock quality as the whole face must be quarried at once. Towards the top of a pit the faces will generally contain more brown rock and as you move deeper into the pit, the proportion of blue rock increases.

This distribution means that, as the pit ages, the proportion of unweathered to weathered rock increases. The changing proportions make the move to a new pit problematic. As the old pit runs down, large amounts of high quality, blue rock are extracted, and then, when the new pit opens, large amounts of lower quality, brown rock are extracted. Winstone Aggregates need a constant supply of each of the three resource types. For this reason both pits will need to be quarried simultaneously for a period of time, this is expensive and so the crossover period needs to be minimised.

The actual proportions of each of the three qualities that are extracted from the quarry fluctuations significantly over the short term but remain fairly constant over the long term.

2.2 Processing

Once the rock has been quarried it is processed into end products. This processing is done by crushing the rock and screening it into size groupings. A large crushing machine crushes the rock from boulders into smaller particles. The crushed rock is then screened into different particle sizes and either re-crushed into smaller pieces, or processed in secondary plants (RJ Maxwell, 2005).

The Hunua Quarry produces over thirty five different products and these products are made by setting up the crushing and screening machines to run different production modes. The cost of running the plant is dependent on the production mode being used. This cost changes due to the number of people required to operate the plant, the wear and tear on the plant equipment and utility costs.

For the purposes of this strategic model the 35 products can be grouped into four product groupings: Concrete, Asphalt, GAP and Hardfill, where each grouping is defined by the proportions of each quality resource it requires. The usages of each product and the ideal proportions of the resources in them are summarised in Table 2.1.

Product	***Usage***	***Blue***	***Blue Brown***	***Brown***
Concrete	Construction	1	0	0
Asphalt	Road sealing, construction	0.22	0.78	0
GAP	Roading base course layer	0.15	0.85	0
Hardfill	Roading sub-base course layer, filler	0	0.08	0.92

Table 2.1: Product groupings description.

While each product grouping has defined ratios for each type of resource quality, these ratios are continually changed up and down, by small amounts, to ensure that all demands can be met from the fluctuating resource supply.

The ranges in which each product grouping should be produced have not been explicitly defined by Winstone Aggregates and they have primarily used this flexibility as a tool to handle supply variation. To work out the ideal, minimum and maximum percentage bands on the resource inputs for each product grouping I analysed historical production data and talked to Winstone Aggregate managers about what they felt were acceptable minimum and maximum bounds as well as what they thought to be the "perfect" resource usage.

2.3 Demand Forecasts

Because the demand comes predominantly from long term contracts, accurate forecasts can be made for the demands of the quarry. In most instances, these contracts are legally binding and therefore the quarry is essentially demand driven. The legally binding nature of the demands is the reason why Winstone Aggregates want to be able to investigate their ability to handle future variations in demand or supply.

2.4 Capital Requirements and Variable Operating Costs

Quarrying is a capital expensive process and the decision when to commit capital is the main strategic decision facing management. The capital costs can be roughly lumped into two groups: capital commitments that are fixed in time and must be made regardless of the strategic decisions made, and capital commitments which can be committed at varying times depending on how the quarry is developed.

The first group of capital costs, those fixed in time, have no bearing on the decision process for this model as they can not be altered. This capital, however, does have a significant impact on the cash flow of the quarry and so to ensure the model's accuracy it is important to include them:

- **Plant Capital:** All capital required to maintain the processing plant. There is only one plant for the two pits so this is independent of which pit is being quarried.
- **Other Capital:** The operation and development of the quarry requires fixed capital injections for purposes which may not be covered in other categories. These capital injections can often occur suddenly and unexpectedly.

The second group of capital costs, those which are only required if a pit is being quarried, is the important group for the decision process. The point in time in which these costs will be committed will differ for each development option.

- **Operating Capital:** The physical process of extracting the rock from the ground is capital intensive. Expensive machinery must be purchased, or hired, and this machinery will have associated depreciation costs.
- **Consenting Costs:** The development and operation of a quarry is governed by the RMA (see § 1.2). Consents must be obtained before a pit can be developed.
- **Site development:** The development and quarrying of a pit requires large capital expenditure in site development. Roading and other infrastructure must be developed and maintained between the pit and the processing plants.

Forecasts for all of the required capital can be made accurately based on previous experience and have been done by Winstone Aggregates in an initial investigation into the viability of the Symonds Hill Pit.

In addition to the capital costs outlined above, the quarrying process has associated variable operating costs. These relate to both the extraction of the resource and the processing into end products.

The extraction cost per cubic metre will be different for the two pits, making these costs important in the decision process.

2.5 Price Forecasts

The market price for aggregate products can be estimated by using historical trends and other forecasting techniques. Because products have been aggregated into product groupings, an average price has been used for each grouping.

The actual pricing is not going to affect the decision of the model but, again, it will affect the accuracy of NPV calculated.

2.6 Information Required by Winstone Aggregates

In order to make the best decision about when to open Symonds Hill, Winstone Aggregates wanted to be able to investigate the financial implications of different options. This information is important to them because of the long lead time required when making this decision and also for dealing with set backs that may result from the resource consent process.

The most important factor in evaluating the financial implications for all possible quarrying scenarios is, “can it meet all required demands?” For the scenarios that can meet all of the demands, the NPV of each scenario provides a way of evaluating the financial implications of committing the development capital at differing times and provides a basis for comparing each scenario.

The Net Present Value is a simple accounting technique that calculates the total value of the quarry taking into account the time value of money.

The formula for the NPV is:

$$NPV = \sum_{t=0}^{N} \frac{C_t}{(1+i)^t}$$

Where C is the net cash flow in period t and i is the internal rate of return.

A secondary consideration for Winstone Aggregates when evaluating each development scenario is, under each scenario, how well the quarry will be able to handle variations in demand or supply.

To determine this, the amount of deviation away from the ideal that is required in the proportions of each resource grade for each product, can be viewed as the degree of risk inherent in each scenario. The more deviation required implies that the quarry is being stretched closer to its operational limits and therefore, there is less flexibility left to deal with short term demand or supply fluctuations. There is a greater risk that the quarry will become infeasible at some point in time. An ability to quantify this risk and to compare it among different scenarios helps Winstone Aggregates make prudent strategic decisions.

3 Model Formulation

The key piece of information that Winstone Aggregates must be able to get from the model is - which development scenario provides the best financial return? To answer this, multiple scenarios must be evaluated to ensure that they are feasible, meet all demands, and for those that do, the optimal development scenario is the one with the best NPV.

The second piece of information that must evaluated is - within each year, what proportions of each resource should be used in each product? For feasible scenarios, this information can be used to develop a risk profile, while for infeasible scenarios, it can be used to analyse where mitigating actions need to be made to achieve feasibility. For this decision, constraints are put in place on the allowable ranges of each resource for each product and a solution that minimises the deviation away from the ideal resource proportions described above is deemed the best solution.

3.1 Approach

To formulated the model as a two stage model solving for the NPV and feasibility in Visual Basic for Applications (VBA) and solving for the quality deviations in AMPL. This method was chosen as it provided significant advantages in both the simplicity of the model and also in the amount of information that could be taken from the model.

By splitting the problem into the two stage process using VBA and AMPL, I was able to avoid binary and nonlinear constraints and the size of the problem was reduced significantly. This simplified the model enough so that it solved very quickly using the free solver, LPSOLVE, and it also ensured that, when expanded, the model grew linearly.

3.2 Model Structure

The first stage of the model is primarily concerned with checking the feasibility of development options and then solving for the deviations away from the ideal proportions that are required in order to meet all demands. The second stage calculates the NPV for all feasible data sets and ranks the feasible data sets accordingly.

The feasibility is checked in VBA. This is done as the information is needed to determine whether to solve for the entire scenario data set or to only solve for part of it. If a scenario is feasible on every year then the entire data set is deemed feasible and the data set is passed onto AMPL. If a scenario is not feasible during any year then the scenario is deemed infeasible and all years up to the infeasibility are passed to AMPL to be solved. This is allows trends to be analysed.

3.3 Approximations

In formulating the model it was necessary to make certain approximations. Two that warrant specific explanation are: firstly, stockpiling constraints are ignored; and secondly, resource surpluses are discarded between years. These approximations decoupled the model between each year and significantly reduced the complexity of the problem.

Because of the long term nature of the model it was acceptable to ignore stockpiling. Stockpiling is a short term technique used tactically to allow the rate of production to remain consistent when demand oscillates, and to allow for production modes to run for longer thereby reducing downtime. The ability to meet the demands in this model, however, is considered on a yearly basis in the model, therefore the demand fluctuations do not need to be considered and it is acceptable to ignore stockpiling.

The discarding of resource surpluses has a greater impact on the accuracy of the model and, if not monitored, could potentially result in skewed results. Calculated surpluses are discarded for two reasons. Firstly, surpluses can not be rolled over year to year as this will result in resource forecasts that can not be physically achieved. Secondly, when considering technical aspects of the model, not rolling the surpluses over decouples the problem by removing linkages between the years. This makes the model significantly simpler to solve.

3.4 AMPL Model

The linear programme model formulated in AMPL is solved for each scenario in each year to find the best possible resource proportion mix for the final products. Deviation costs are assigned to force deviations away from the ideal proportions to occur first in the lower grade resource and in the lower value products; this is what is done currently by Winstone Aggregate quarry managers.

The model was formulated as a production planning model with the decision variable being the resource allocation for each product. The objective function was chosen to minimise the deviations away from a defined ideal and the constraints ensure that all demands are satisfied and that resource proportion deviations remain within the specified upper and lower bounds.

Indices

i = resources: Blue, BlueBrown, Brown.

j = products: Concrete, Asphalt, GAP, Hardfill.

k = time period: $S,\ldots, E$.

Parameters

S = first year of the solve period.

E = *last year of the solve period.*

Cij = *cost of deviating away from the ideal proportion of resource i in product j.*

Rik = *amount of resource i available in time period k.*

Djk = demand of product j in time period k.
Iij = ideal proportion of resource i in product j.
Mnij = minimum proportion of resource i in product j.
Mxij = maximum proportion of resource i in product j.

Decision variables

z_{ij} = surplus resource *i* in time period *k*.
d^{+}_{ijk} = deviation of resource *i* in product *j* and time period *k* away from the ideal proportion towards the maximum bound.
d^{-}_{ijk} = deviation of resource *i* in product *j* and time period *k* away from the ideal proportion towards the minimum bound.

Production Plan Model

1) Minimise $\sum^{3}_{i=1} \sum^{4}_{j=1} \sum^{E}_{k=S} C_{ij} (d^{+}_{ijk} + d^{-}_{ijk})_j$.
2) $I_{ij} + d^{+}_{ijk} - d^{-}_{ijk} \geq Mn_{ij}$ for all i, j, k.
3) $I_{ij} + d^{+}_{ijk} - d^{-}_{ijk} \leq Mx_{ij}$ for all i, j, k.
4) $\sum^{3}_{i=1} [x_{i,j} + d^{+}_{ijk} - d^{-}_{ijk}] = 1$ for all j, k.
5) $\sum^{4}_{j=1} D_{jk} [x_{i,j} + d^{+}_{ijk} - d^{-}_{ijk}] + z_{ij} = R_{ik}$ for all i, k.
6) $x_{i,j}$, d^{+}_{ijk} , $d^{-}_{ijk} \geq 0$.

Explanation

1) The objective is to minimise the total deviation cost to get the best resource allocation.
2) The proportion of each resource in each product must be greater than the minimum allowable.
3) The proportion of each resource in each product must be less than the maximum allowable.
4) The proportions of the resources in each product must total 1 for consistency.
5) The total product supplied to the market plus the surplus must be equal to the total resource available.
6) All variables are positive.

The C_{ij} costs were choosen to force deviations to occur firstly in the lower quality, lower value products. The relative costs used in the model were found, in consultation with Winstone Aggregates, through observation of what produced the most desirable allocation of the resources. Deviations in either direction away from the ideal were penalised equally.

3.5 Risk Assessment

The output from the AMPL model was used to develop a risk profile for each scenario. This is important information for the evaluation of each scenario, as the higher the required deviations are, the closer to its operational limit the quarry will have to operate at. The risk profile therefore shows the ability the quarry has to absorb short term variations and/or set backs.

To display this information in a concise and simple to interpret way, a risk count was calculated. For each scenario, a count is made of the number of year, product, and resource triples in which the deviation is more than fifty percent of the maximum. The size of this count between different resource scenarios represents the relative risk difference between each scenario.

The standard and average deviations for each product resource pair are also calculated and displayed as a percentage of the allowable deviations. This information can be used for further investigation into the risk associated with each scenario.

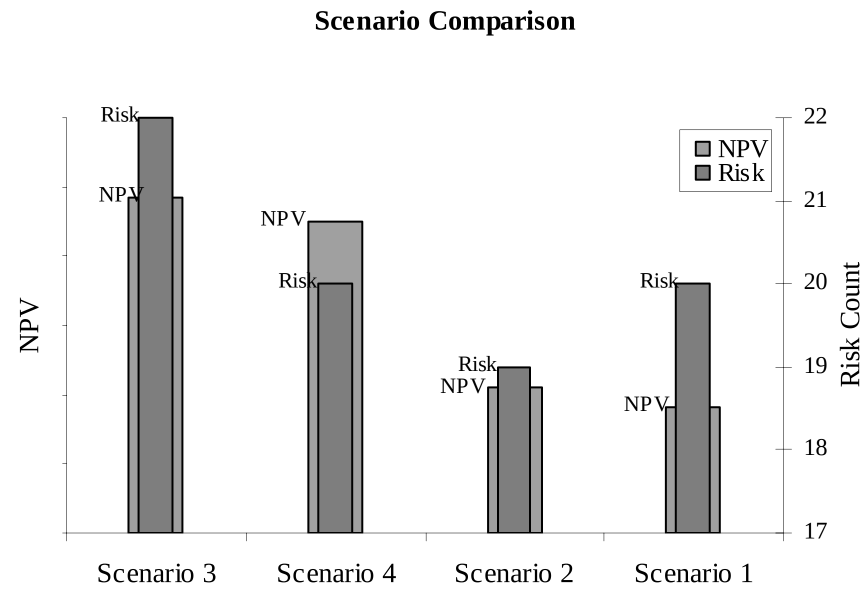


Figure 3.1: Results summary chart produced automatically by the model. The NPV values are blacked out for commercial confidentiality.

4 Implementation

A key issue that I identified early on in this project was the need to make the model very simple to use. For a tool such as this to have lasting usefulness it must be possible for any person with a reasonable understanding of the quarrying process to be able to set up the model and evaluate the results. To achieve this, two issues needed addressing:

- The interface needed to be simple and easy to navigate.
- The interface needed to be simple and not time consuming to set up.

To achieve both these goals I choose to use an Excel workbook and based its layout on spreadsheets provided to me by Winstone Aggregates. I split the workbook into six input sheets, one corresponding to each of the physical components of the problem (§2), and one results sheet and a results chart. At Winstone Aggregates' request, I also highlighted in a consistent colour all cells where the user is required to input data. This significantly increased the ease of use.

5 Model Validation

The purpose of developing this model was to provide Winstone Aggregates with a tool which they could use in an ongoing manner to continually undertake long term planning for the Hunua Quarry. To validate the model I formulated five resource forecasts, each based on different development options. I formulated these in such a way that the correct NPV ranking was obvious and I used all the demand, price, cost and capital data provided to me by Winstone Aggregates. Each forecast runs the Hunua Pit resource down at a different rate and opens the Symonds Hill Pit at a different time and at a different quarrying rate.

When the model was solved using the test resource data sets, the results produced were as expected. This validated the accuracy of the model.

A summary of the results with the NPV expressed as percentage of scenario 1 is presented below in Table 5.1. Because of the scale of the quarry operation, small percentage increases represent significant increases in value.

Scenario Name	***Symonds Hill Open Year***	***NPV***	***Risk Count***	***Rank***
Scenario 1	2009	100.00%	20	4
Scenario 2	2010	100.17%	19	3
Scenario 3	2011	101.78%	22	1
Scenario 4	2011	101.58%	18	2
Scenario 5	2012	-	No	-

Table 5.1: Validation results.

6 Conclusions

A strategic model was developed for Winstone Aggregates to model the long term performance of the Hunua Quarry. The model provides an accurate representation of both the physical quarrying operations and the capital commitments required for this operation and in doing this, it provides Winstone Aggregates with detailed information on the long term performance of the Quarry.

The model gives Winstone Aggregates a flexible, easy to use tool, with which they can investigate both of the key issues raised by them, namely:

- When should Symonds Hill Pit be opened and at what rate should quarrying begin.
- How will the quarry respond to future demand variations, and when should any mitigating actions begin.

To address these issues it investigates any required number of different quarrying scenarios and provides the following information for each one:

- The NPV of the quarry
- The required deviations from ideal resource proportions for each product and a risk profile derived from these deviations

To ensure the accuracy and flexibility of the model, all parameter information is controlled by Winstone Aggregates and can be changed by them at any point in time. This ensures the model is of continuing use.

Winstone Aggregates have implemented the model and have begun using as an integral part of their long term planning.

6.1 Limitations

The model was developed for use as a Strategic Production Plan Model and is intended for the long term planning of the Quarry development. It is not intended to be used for short or intermediate term planning.

The model has been highly aggregated into just four product groupings and necessary approximations were made. These simplification steps are necessary for the development of a long term model as they focus attention towards the important long term decisions, however, it is important that this is understood and that the model does not get used as a substitute for current tactical and operational planning procedures. It would be preferable if complementary optimisation models dealing with the short and intermediate time frame decisions were developed in the near future.

7 Acknowledgments

I would like to thank Dr Stuart Mitchell and Dr. Cameron Walker, my project supervisors, for all the help and guidance they provided to me through this project. I would also like to thank Winstone Aggregates for the opportunity to work on such a project, and for the time they spent gathering information and discussing the problem with me. My special thanks to Bernie Chote, Georgia Manning and Mark Rippey.

8 References

Anthony, R. N. 1965. “Planning and Control Systems: A Framework for Analysis.” *Harvard Business School.*

Elangasinghe, M. 2001. *Development of Strategic Planning Software at Winstone Aggregates.* University of Auckland Engineering Science Fourth Year Project.

Hochbaum, D. S., Chen, A. 2000. “Performance analysis and best implementation of old and new algorithms for the open-pit mining problem.” *Operations Research.* 48 (6) 894-914.

Mitchell, S. 1997. *Optimal Production Planning At Winstone Aggregates.* University of Auckland Engineering Science Fourth Year Project.

New Zealand Minerals, Ministry of Economic Development. 25 August 2005. www.crownminerals.govt.nz/minerals/facts/2003/region.html.

Quarrying Process and Quarrying Products, RJ Maxwell. 3 September 2005. www.rjmaxwell.com/education.

Ramazan, S., Dagdelen, K. and Johnson, T. B. 2005. “Fundamental tree algorithm in optimising production scheduling for open pit mine design”. *Mining Technology (Trans. Inst. Min. Metall. A).* 114 45-54.

Resource Management Act 1991, New Zealand Government. 19 September 2005. www.legislation.govt.nz/

Winstone Aggregates. Between 20 April 2005 and 25 August 2005. www.winstoneaggregates.co.nz.

Capacity modelling of the South Island chicken operation at Tegel Foods Ltd

Tim Steer and Richard Bramley
Department of Management
University of Canterbury
New Zealand
tps27@student.canterbury.ac.nz
rmb76@student.canterbury.ac.nz

Abstract

Tegel seeks more accurate estimates of the capacity of its nationwide chicken operations. A prototype decision support tool for the South Island chicken operation was developed as a starting point in achieving this goal. The model calculates capacity and utilisations of the different areas of the operation as well as overall capacity. The capacity model aids strategic decision making and promotes understanding of different areas of the operation. Emphasis is placed on presenting relevant information in a user-friendly way, with the objective of adding maximum value in a limited timeframe. This paper describes the capacity model created and the logic behind the modelling process. The model is validated using 2006 forecasted demand. Extensions are recommended which will increase the accuracy and benefit of the model.

1 Introduction

Tegel Foods Ltd is a privately owned organisation specialising in the production and processing of poultry. It is the leading supplier to the New Zealand poultry market and employs over 500 people in the South Island alone. Tegel's high level of vertical integration means that it has control over most facets of a chicken's lifecycle.

Tegel frequently makes strategic decisions concerning production targets, new product introduction and facility expansion. These decisions are often made based on the experience of individual area managers with the support of little numerical evidence. This leads to what staff describes as time consuming and inefficient communication and does not promote the diffusion of knowledge amongst the company. Area managers at Tegel make operational decisions based upon demand forecasts produced by senior management. These forecasts describe demand in terms of whole birds per month and the prescribed average weight of birds.

Tegel seeks to improve its understanding of the capacity of its nationwide chicken operations. As the first step to achieving this, a decision support tool for the entire South Island chicken operation has been developed. It is a capacity model that, if it proves successful with managers, may be adapted for use nationwide.

Two main groups are expected to use the model. Senior management will explore the effect of strategic changes on capacity and the feasibility of production plans. Individual area managers will use the model to learn about the effect of changes on other areas of the operation and consider where bottlenecks lie. Interestingly, the managers who originally considered the model to be of least use to them are now most supportive and enthusiastic about its results. It demonstrates that the usefulness of model can be substantially improved if the misconceptions and threats of the model are broken down.

The model consists of a Microsoft Excel® spreadsheet with the use of Visual Basic for Applications. The use of Excel® ensures it is simple and user friendly to all users.

The capacity model is a prototype of a more extensive decision support tool that should be developed in the future. To make it a final product, product differentiation and improvement of estimates at the main plant should be undertaken. To allow future extensions and adaptation to other operations in the North Island, the model is designed to be adaptive and easy to control.

The model has been validated using production plans for the financial year of 2006. The capacity utilisations estimated by the model are close to those hypothesised by managers.

2 Problem description

The South Island Tegel chicken operation begins at the breeding of eggs and ends after goods are processed and distributed, as shown below in figure 1. At eleven breeder farms spread out across Canterbury, chickens are raised to produce eggs that are then transported to Tegel's hatchery in Hornby, where they are incubated and hatched. The day-old chicks are then sent to one of thirty broiler farms situated close by. These farms consist of large sheds where chickens are raised until they reach a desired weight and are gathered for processing. They are taken to the main processing plant in Hornby, where they are killed and transformed in to saleable product. Tegel also owns and operates the largest feed mill in the South Island, which produces the feed for both the breeder and broiler farms.

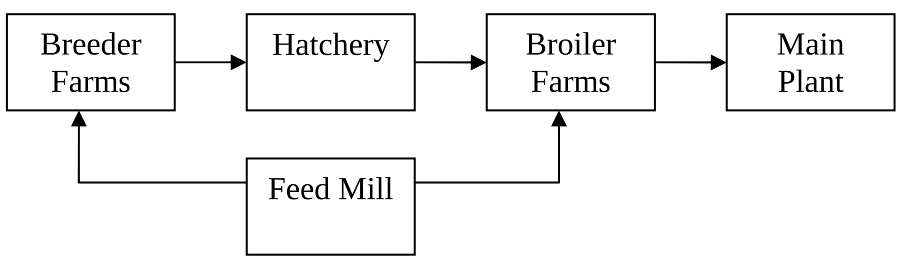


Figure 1. Tegel's South Island chicken operation

There are no models at Tegel that incorporate all areas of the operation and that are dedicated to estimating capacity. Most of the models in use at Tegel relate to the planning and scheduling of livestock.

Developing a capacity model presents many modelling challenges because of the nature of Tegel's chicken operation. Tegel has a long internal supply chain and consequently all parts of the operation must be considered to gauge the true capacity of any one area. Another challenge of developing a holistic capacity model is that the different areas of Tegel's chicken operation are highly dependent. Tegel is aware that being able to describe the interrelationships in a single model will be more beneficial than detailed independent models.

An illustrative example of the relationships between areas of the chicken operation is the decision to process larger but fewer birds at the main plant, in order to meet total kilograms demanded per month. Fewer eggs must be laid and hatched, requiring fewer breeder hens. Fewer birds must be grown at the broiler farms and processed at the main plant. However, birds must be grown for longer which requires more of the broiler farms in a given month. The processing speed of the plant may also slow down. Based on intuition alone, the net effect on overall capacity is unknown. Despite the same kilograms of product being supplied, the interaction between inputs is such that the effects on capacity can be very different.

3 Objectives

The objectives of the project were determined together with Tegel. These were to:

- Develop a capacity model that encompasses the livestock, feed mill and processing aspects of the South Island Tegel chicken operation. It should provide a high-level understanding of the capacity constraints of the operation and should illustrate the interdependencies between the different areas. It should be useful to all individual livestock, feed mill and processing managers by providing quantitative evidence of what they already intuitively understand about capacity and by providing understanding of relationships between different areas.
- Identify and develop key summary measures of capacity utilisation for the livestock, feed mill and processing areas.
- Develop a means to conduct ‘what if’ analysis, such as facility expansion or a change in operating procedure. This should be useful to senior management as an aid for strategic planning by allowing them to consider the quantitative implications on capacity from new scenarios.

4 Users of the model

The capacity model developed for Tegel Foods will be used by a variety of managers, each for a different purpose. The needs of the users are important because they determine the design of the content, the time frame and the form of the model.

4.1 Primary users

There are two primary users of the model.

The first primary user is the ‘Business Integration Manager’ at Tegel Foods, who plans production nationwide based on demand forecasts received from marketing. He will use the model to determine what capacity is for a given month or year, and consequently determine whether production plans appear feasible. It will save him considerable time and allow him to plan more efficiently because he no longer has to rely solely on guesstimates by managers and planners.

The second primary user is the ‘South Island Manager’. The model is of benefit to him by allowing him to explore long term strategic changes to the South Island operation. In particular, he can assess the effect on capacity when hypothetical changes are made in the model. This might include the addition of new broiler sheds, the conversion of a breeder farm or the introduction of a better performing feed.

4.2 Secondary users

The secondary users of the model are the individual area managers. These include the breeder farms manager, the hatchery manager, the feed mill manager, the livestock planner and the main plant manager. Among the secondary users, two different responses to the project have been observed:

In the first group of responses, managers are eager that a highly detailed daily capacity tool be developed that allows them to plan capacity through scheduling and other optimisation techniques. It is of advantage to each manager that their particular area be modelled, yet not everybody can be satisfied in the time available. The conflicting interests in the project made it important to be clear about the strategic aims of the model and balance the detail of the individual area models with those aims.

The second group of responses to the project consists of managers who did not view an overall capacity model as being useful to them; that it does not contribute any further knowledge than they already have. It was crucial that the managers understand that the model aids their understanding, not replaces it.

5 The model

In all areas of the Tegel chicken operation collating relevant information is time consuming and challenging. In an operation of this size, there is extensive technical detail and data, only some of which is relevant to describing monthly capacity. It is a common requirement to collect data from staff who work in a detailed part of an area, such as a chicken farmer or mill worker, who are unaware of the project's objectives.

The design of the model was based upon six criteria taken from Little (1984). These criteria were that the model be simple, robust, easy to control, adaptive, complete and easy to communicate with.

4.1 Microsoft Excel capacity model

The capacity model is developed in Microsoft Excel®. This was chosen because it is widely used amongst managers at Tegel and allows them to alter and extend the model with ease, thus keeping it simple, adaptive and easy to communicate with. Features, such as the use of cell protection and undo options are included in the model to ensure that it is robust and valid. The planning spreadsheets currently in use at Tegel, described by staff as difficult to manipulate and understand, are large and contain unnecessarily complex macros. The capacity model is kept user-friendly by keeping the use Visual Basic for Applications macros to a minimum and by only using them for display features of the model. Likewise, formulae are kept as simple as possible.

The model consists of six worksheets, five for the areas of the operation and one to summarise the information of the model. The summary sheet includes overall capacity and utilisation and is designed for the use of senior management who do not understand the areas in depth.

4.2 Describing capacity

Hill (2003) defines capacity as "the maximum rate of output of a process, measured in units of output per unit of time". Correspondingly, the model seeks to describe the maximum number of whole birds that can be processed in a month by Tegel's South Island chicken operation.

Each area at Tegel Foods is made up of a number of parts, including machines, storage and transport. The capacity of an area is the minimum of the capacity on the individual parts. This is understood intuitively, a production line can operate no faster than the speed of its slowest process. Calculating capacity at the main plant is more complex because birds have different pathways through the factory depending upon what product it will become.

The capacity of the whole operation is calculated as the minimum of the capacities of the different areas. This is valid because the flow between areas is linear, as discussed below.

4.2.1 Measuring capacity

The model incorporates five different areas of operations (main plant, broiler farms, hatchery, breeder farms and feed mill), some of which operate in different units. For example, the main plant inputs live birds and outputs kilograms of product, whereas the hatchery processes eggs (unhatched to hatched) and the feed mill outputs tonnes of feed.

The model expresses the capacity of each area in its unique units because they are most meaningful to that area's manager. Area capacities are calculated in comparable units so that overall capacity can be determined. Monthly production at Tegel is planned in terms of whole birds processed by the main plant (specific product demand is derived from this). Senior management is accustomed to communicating in these units. As a result, the summary sheet of the model displays the capacity of each area in terms of the number of whole birds entering the main plant in a month. Expressing area capacities in these terms is also new information to area managers and is very beneficial.

A benefit of measuring monthly capacity rather than daily or weekly is that it removes most uncertainties from capacity estimation. Over a day or a week, bird growth rates and machine processing speeds can vary significantly, but production data shows that over a month variation is minimal. Capacity is accurately described by using averages.

4.2.2 Theoretical vs. demonstrated capacity

The model distinguishes between machine (theoretical) capacity and achievable (demonstrated) capacity. Capacity estimates are used to determine the feasibility of product plans, and therefore it is important that what can actually be achieved is calculated. In the main plant machines cannot operate at one hundred percent efficiency. The maximum output that can be realistically produced will not be the same from day to day due to normal factors such as staffing, productivity and equipment breakdowns.

Capacity is first described in the model in terms of the maximum rate of output that can be produced if machines operate non-stop. Users of the model can view theoretical machine capacity and be certain of its accuracy. To describe reality, factors such as staffing and breakdowns are complex and require more analysis to accurately model.

Calculation of achievable capacity attempts to approximate these affects by estimating 'maximum machine utilisation'. This is the maximum percentage of machine capacity that can be utilised when fully loaded. This was estimated through discussion with managers and shop floor workers, and by observing demonstrated capacity in times of peak demand. The use of achievable capacity ensures the model is easy to communicate with. It is in accordance with Tegel Food's operating policy that machine efficiency is the priority and that staff can always be employed to operate the machinery at its optimum. This operating strategy is the result of Tegel's high cost of capital investment and expansion; staffing costs are relatively cheap in comparison. Tegel also seeks to use maximum machine utilisation as a production target in the near future.

4.2.3 Qualitative aspects

Qualitative aspects of capacity play a significant role in determining the level of accuracy required when estimating capacity. These include the physical and financial feasibilities of increasing capacity. For example, if a machine is relatively cheap and can be installed in a short period of time, the benefit of modelling it in detail is small. Realistically, if the machine becomes a constraint on capacity it can be easily relaxed and an extra five percent of accuracy in estimating its utilisation will be of little benefit. The decision to model is a balance between the time involved and the benefit gained.

4.3 Factories

Among the areas of the chicken operation there are, by nature, two different types – the machine driven factories and the livestock farms. Each was modelled in different ways.

The machine driven factories include the hatchery (machines are used to hatch the eggs), the feed mill and the main plant. To use 'overall capacity is the minimum capacity of the parts', it was essential that the flow of product (eggs/feed/birds) was linear. This means that product flows sequentially along a line, it does not flow down an alternative path when one direction reaches capacity. This was definitely true of the feed mill and hatchery which are both relatively simple processes.

Capacity calculations also assume that the flow of product in these factories is constant. The reality is that, in the example of the main plant, a series of smaller birds may arrive at the factory and will consequently all be processed down the same line, whilst machinery dedicated to larger birds sits idle. To incorporate this into the model would require scheduling, which was deemed beyond the scope of the project; therefore capacity is calculated under the assumption that flows are constant. This is accounted for to some extent by maximum utilisation of machinery. The addition of scheduling is a possible extension to the model.

4.3.1 The hatchery

Modelling the hatchery is relatively simple due to linear flows and stable process times. Some difficulty arose regarding the fact that because of the hatchery's operating hours; eggs can only be put in and taken out of machines on certain days. This will always occur due to the availability of staff and is included because it affects achievable capacity (it causes downtime on the machines).

4.3.2 Modelling the feed mill

The useful information required by users of the model with respect to the feed mill is the amount of feed required and also when a type of feed is required by livestock. Capacity is shown, however utilisation is not important because the feed mill can reduce its production of non Tegel related feed if need be.

4.3.3 Modelling the main plant

The main plant is significantly more complex than the other areas of the operation. Production is made up of an automated processing line, approximately five different types of machinery in secondary processing and finally three refrigeration units. It is challenging to model the capacity of the plant in summary form.

Determining the flow of products along different paths was complex. Flow is driven by product demand and therefore the flow of fresh/frozen, whole birds/portions and a-grades/down-grades will vary from month to month. The same total demand, in terms of kilograms of product, may result in different capacities at the main plant dependent upon which specific products are produced. The model does not differentiate between products and, consequently, flows had to be approximated using the average proportions that are processed at respective machines. Historical data is used to make these approximations.

The flows at the main plant are non-linear. Circumstances arise when flow is directed elsewhere to alleviate a bottleneck on the line. However, these redirections are rarely planned or consistent and were too uncertain to model in the time available. The model therefore assumes linear product flows. Without a definite set of production rules, accounting for non-linear flows becomes very challenging and is a possible extension of the model.

Although capacity at the main plant lacks accuracy because of the approximations, the benefit is that it remains simple and easy to understand. With the exception of the main plant manager, the users of the model know little about how the main plant functions and need to be able to understand the consequences of altering key inputs and parameters. A more detailed daily capacity model of the main plant, which includes scheduling, would be extremely useful to the main plant manager but would not be appropriate for the other users. Ideally a summarised form of a more detailed main plant model could later be expressed in this capacity model.

4.4 Livestock farms

The livestock growing farms of the chicken operation at Tegel are the breeder and broiler farms. Capacity at these farms is calculated based on shed space and growth rates rather than machine processing speeds. The challenge when determining monthly capacity of these farms is that unlike the factories, not all resources are used to supply a single month's demand. In other words, it takes longer than a month for a bird to grow to full size, or for a breeder to reach its oldest breeding age. Therefore, in the example of the grower farms, some sheds must be growing flocks to supply demand two months from now, whilst others grow for next months demand. To determine capacity of these farms the notion of a 'growing cycle' or 'production cycle' was used. This is the length of time from the beginning of growing/breeding till the next flock of broilers/breeders can begin. It will include both growing/breeding of a flock and time for cleaning and bio security checks. Once an accurate estimate of the average growing/production cycle is determined, and given the total time available in a month, the proportion of total capacity available for a single month is easily calculated.

4.4.1 Modelling the breeder farms

The breeder farms provide a challenge because they consist of separate facilities, one type for rearing and another for the production phase of the breeding hen's lifecycle. The two processes have distinctly different purposes and operating methods and are therefore modelled as independent parts. The relationship between the ages of the hen and number of eggs produced had to be modelled in the production farms.

4.4.2 Modelling the broiler farms

Tegel contracts 30 broiler farms in the Canterbury area to supply the main plant. Each grows birds from day old chick to a weight at which they are ready to be killed. The two key constraints on the capacity of the broiler farms are the total shed area and the stocking density limits, as defined by Tegel's nationwide policy (within welfare restrictions).

The most difficult aspect of the broiler farms is the significant variation between farms in the daily growth rates of birds. Standard deviations and other means of incorporating uncertainty in the model were considered, but are not necessary because all farms are treated equally by Tegel and therefore daily uncertainty balances out over the course of a month. It is a contractual agreement that all farms be placed with an equal spread of birds over the course of a year; it is not the case that some poorer performing farms only grow smaller birds whilst others grow the larger birds. Over half of the thirty farms grow for a specific month which showed to be sufficiently enough such that the predicted total growth in a month is close to what is observed in reality.

5 Validation of the model

It is essential that the model is a realistic representation of capacity so that users can consider the feasibility of a production plan or explore strategic changes with confidence. This includes ensuring that the model is robust and does not produce confusing solutions.

Triangulation was used as a method of validation in the factories. The model's estimates were compared to both historical data during peak demand and managers' expectations. The model determined that at the main plant, current production levels could only be achieved by running overtime, which is exactly what is occurring in reality. Although the results show that the models estimates are very close to observed capacity, it is unsurprising given that the same sources were used to determine the parameters of the model in the first place.

Validation of the livestock farms is more credible because modelling does not rely on manager's estimates. The growing and production cycles at the broiler and breeder farms determined by the model are very close to what is currently being observed.

Overall, the model shows what many staff at Tegel predicted. Given the production plan for 2006, almost all areas need to operate at high utilisation, although the main plant and the hatchery are of most concern because, unlike the livestock farms, they cannot easily be expanded.

6 Extensions

6.1 Developing a final product

The capacity model is a prototype of a complete decision support tool. A final product was not possible in the time available given the size and complexity of the operation. Project contacts at Tegel understood from the beginning that, whilst there remained many possibilities for modelling, not all were possible in this project.

The model, in its current form, is still useable and will be of high benefit to staff at Tegel Foods. It is designed to be as accommodating as possible to additions to the model. The extensions listed below are additions that would be expected of a complete product and should be undertaken in the short term future:

- The current model only deals in terms of whole birds. Converting this into specific products be useful to users and will improve estimation of the product flows between machines at the main plant. This may require the integration of the further processing plants (on site at Hornby) into the model.
- Some parameters currently used in the main plant model are estimates by staff that have not been validated thoroughly. Improvements include estimation of refrigeration capacity using cooling formulae and investigation into production rules.

6.2 Future modelling

There are a number of other possible extensions to the model, some of which will require significant undertaking. Users of the model will have to assess the costs and benefits of each of the extensions listed below:

- Many of the parameter and input changes that can be made in the model are not necessarily cost effective. Planners could better assess the benefits of strategic changes if both the capacity and costs were quantified in the model.
- The model estimates monthly capacity. This best suits the primary users of the model by aggregating the whole process to a level that can be easily understood. Tegel also needs more detailed capacity models in the main plant and feed mill so that the area managers can undertake capacity planning.
- Scheduling is a method of capacity planning that would be of high benefit to Tegel. The breeder and broiler farms are currently scheduled manually, sometimes in an ad-hoc manner. This method has been adopted to deal with the uncertainty in the growth rates of livestock and the production of the main plant. Despite this, there is definitely opportunity to improve their spreadsheet scheduling model by incorporating automated heuristics or optimisation techniques. There is also a potential to implement scheduling in the main plant.
- The scope of the model currently excludes the capacity of the supplier to Tegel's breeder farms and of the distribution and storage of products. Estimating these capacity constraints will be useful to users.

o Adapt this model for use in the two North Island chicken operations.

7 Conclusion

A capacity model of the entire South Island chicken operation at Tegel Foods Ltd has been developed in Microsoft Excel®. It includes integrated worksheet models of the main plant, broiler farms, feed mill, hatchery and breeder farms at Tegel.

The model will aid senior management in strategic decision making, particularly through the use of 'what if' scenarios built into the model. The model will aid area managers by providing them with a capacity model of their area and by increasing their understanding of the interrelationships between areas of the chicken operation.

Forecasted demand for 2006 was used to validate the model. The model's results for this period were close to managers' estimates of capacity utilisations and showed that the hatchery and the main plant are the bottlenecks of concern, as predicted by many managers.

The model, although useable, is not a final product and should be extended by differentiating between products and improving the accuracy of estimated parameters at the main plant. Further modelling will be very beneficial to Tegel, in particular the automated scheduling of the livestock

8 Final comments

The development of the capacity model at Tegel Foods illustrates the value of a basic, user-friendly model that is unthreatening to managers. There remains large benefit to be gained from Operations Research methodology/modelling at Tegel at relatively low cost. Tegel is certainly not the only large scale manufacturer in New Zealand who could benefit from Operations Research. The challenge for consultants is to develop models that are not necessarily ground breaking in their technical processes, but that will continue to be used effectively by managers, such as the one developed in this project.

9 Acknowledgments

We would also like to thank staff at Tegel Foods Ltd, in particular Ed Campion, Jim Kuo, James Bradford, Wayne Miller and Gordon Fitzgerald for assisting the project and making data available. We would like to acknowledge our project supervisors Venkat Pulakanam, Shane Dye and Nicola Petty from the Department of Management, University of Canterbury for well-needed guidance.

10 References

Little, J.D.C. April 1970. "Models and Managers: The Concept of a Decision Calculus". *Management Science*, vol 16 no 8: 466:485.

Hill, A.V. 2003. "Encyclopedia of Operations Management Terms". http://www.poms.org/POMSWebsite/EducationResources/omencyclopedia.pdf

Maximising Manufacturing Performance with TOC

Simon Bleach-Wood
Kathy Goddard
Dean Harding
Carolyn Kildare
Steve Messenger
Master of Business Administration Students
Victoria Management School
Victoria University of Wellington
New Zealand
Stevem@Printstop.co.nz

Abstract

The mechanistic approach to solving complex manufacturing problems considers that individual local process improvements add together to make a global improvement –or, "the whole is the sum of the parts". The Theory of Constraints provides a systems thinking methodology that challenges this view, recognising that the local processes are interdependent and have bi-directional causality, or "the whole is greater than the sum of the parts".

This paper shows how measures and policies caused behaviours that maximised local departmental efficiencies in a New Zealand printing company that suboptimised overall company performance. The case analysis demonstrates the process of increasing existing production throughput and the identification and implementation of new business initiatives, using the Theory of Constraints Thinking Tools.

"This project created a fundamental shift in the way we think, unlocking potential we had never realised existed". Lawrence Evans, Managing Director, Astra Group.

1 Introduction

"Our bindery system is a constraint," announced a Manager at Astra Group. "We need to invest in expensive new equipment to enable us to match the capacity of the bindery department to the output of the printing presses."

Astra Group had spent $3.0m on new printing presses, which increased the overall print capacity of the business. But the bindery department couldn't match the processing speed of the new presses, and the printing presses were being held up waiting for printing plates from the pre-press department. Did Astra need to invest more money in capacity in another area of the production process, or was there another way to look at the problem?

From a very early age, we are taught to break apart problems, to fragment the world. This apparently makes complex tasks and subjects more manageable, but we pay a hidden, enormous price. We can no longer see the consequences of our actions: we lose our intrinsic sense of connection to a larger whole (Senge, 1990). Astra subscribed to a mechanistic approach in measuring the “productivity” of their operating system. Each workstation focusing on its own set of operating measures: machine hours, utilisation and production ratios. The system goal appears lost in the detailed complexity[1]. The relationships between each workstation in the overall process were not well understood.

This case study describes the application of the Theory of Constraints (TOC)[2] to a Wellington printing company’s operating system. TOC to many people means the unconventional management text written by Dr Eliyahu M. Goldratt: “The Goal”[3]. The TOC methodology is a work in progress moving it far beyond its beginnings when “The Goal” first took us through Alex Rogo’s plant difficulties in 1984[4]. In the case of Astra, TOC enabled the identification of system constraints that affected the whole process – and the constraints, and the solutions, went much wider than simply buying a new piece of equipment.

2 The theory of constraints

Goldratt said that: “Constraints determine the performance of a system.” Every process has a constraint[5]. If it didn’t, it would have infinite performance (profit in the case of a business). The implication of Goldratt’s statement is that to improve the whole process, one needs to focus only on the constraint (Cox et al., 2003).

TOC is a multi-faceted systems methodology that has been progressively developed to assist people and organisations to think about problems, develop breakthrough solutions and implement those solutions successfully (Mabin and Balderstone, 2003). TOC provides a set of tools that lead the user to answer 3 key questions: What to change? What to change to? How to cause the change? The major component of TOC that underpins all the other parts of the methodology is the TOC thinking processes (Mabin et al., 2001).

2.1 The theory of constraints thinking process tools

The thinking processes provide the basis for ongoing improvement in any environment (Cox et al., 2003). This is not gradual change, the way most improvements are caused in organisations, but breakthrough change (Cox et al., 2003). The thinking processes comprise a suite of 5 logic diagrams (4 trees and a "cloud") and a set of logic rules. The diagrams use 2 different types of logic. The current and future reality trees and the transition tree use sufficiency logic. The other 2 tools, the evaporating cloud and the prerequisite tree, use necessary condition logic. The 5 tools can be used individually or in concert depending on the complexity of the situation that is being faced (Mabin et al., 2001).

3 Case study: Astra Group

This paper is drawn from a research project presented to Victoria University of Wellington in fulfilment of the requirements for a Master’s degree in Business Administration.

3.1 Company background

The Astra Group (Astra) operates in the New Zealand print services market. This market has annual sales of $1.6bn and is supplied by approximately 1,200 enterprises employing 10,700 full time equivalent workers (Statistics New Zealand, 2003). The environment for print services is highly competitive. The print industry is moving from a number of small to medium sized print shops owned and managed by a craft printer to a smaller number of large, diversified, professionally-managed New Zealand-based businesses and several international competitors.

The origins of Astra can be traced back to 1910 when Brooker & Friend began trading as a publishing company in College Street, Wellington. In 1993 what is now Astra Print was split off from the parent company Brooker. Today, the Astra Group is a Wellington-based firm with revenues of \$20m[6] and 150 permanent employees, providing offset and digital printing services for paper-based substrates.

The operations process at Astra is a "job shop" in which small batches of a large number of different printing jobs are processed. Print jobs move in a sequence between specialist areas, sometimes flowing back the way it came to a previous area before continuing on in the process because of the dependency between 1 area and another. Each step in the process must be completed before the next step can commence. A simplified process is shown in Figure 1 below.

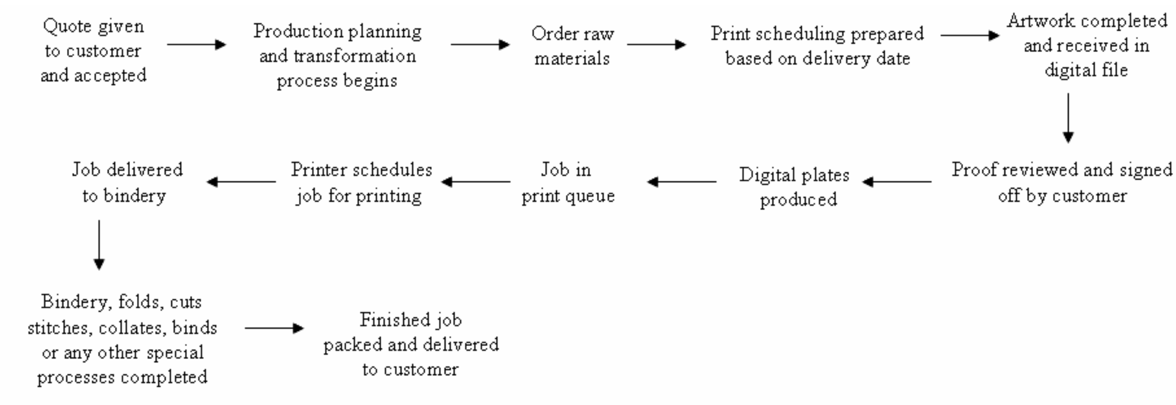


Figure 1. A print job process flow

It is a discrete process in which the process is made up of individual parts. There are many resources combined and assembled into the finished product, such as ink, paper, printing plates, machinery, and human labour.

4 The application of the Theory of Constraints

4.1 Astra's problem

The overall goal for Astra was to satisfy its customers. They considered the key criteria for customer satisfaction were *on-time delivery* and *acceptable finished product quality.*

Astra developed measures to monitor each workstation's contribution to this goal for each print job. The measures were primarily aimed at monitoring productivity, on the assumption that maximum productivity at each workstation would lead to minimum time spent on each job. Astra measured productivity at each workstation in terms of conventional measurements such as machine hours and percent utilisation. Each job was processed by each workstation as soon as it was received in order to meet the delivery date requirements. Yet in front of each workstation there were large volumes of work in progress. Each workstation saw the previous workstation as the cause of inefficiency, as each workstation was often being provided with work that did not match the customer's specification. The end result was that internal deadlines and customer delivery dates were often missed.

Astra's Management thought the measurement system supported the goal. They believed that that was no more capacity in the operating system, and that the only way to increase output was to invest in more equipment. At the time of this project Astra were considering purchasing more bindery capacity. As bindery is the last station in the process for many of Astra products it was assumed that more capacity at this point would improve on-time delivery.

4.2 Analysis methodology

The first stage is to identify *what to change?* The diagnosis stage uses a 3-cloud process to identify the generic cloud – or underlying system issue – that forms the base of the current reality tree. The future reality tree takes these ideas for change and ensures the new reality created (*what to change to?*) will resolve the unsatisfactory system conditions and not cause new Undesirable Effects (UDEs). The prerequisite tree determines obstacles to implementation and ways to overcome them (*how to cause the change?*) (Mabin et al., 2001).

4.2.1 What to change: The Generic Cloud

The Generic Cloud (GC) is created from merging three individual (evaporating) clouds (EC), each of which is based on a single UDE (Mabin et al, 2001). The GC is a logic-based tool for surfacing assumptions related to a core conflict or systemic issue (Cox et al., 2003). It is this conflict that is manifested in the UDE.

To develop a GC for Astra, staff from production (prepress, printing and bindery) and business support (sales and finance) were interviewed using standardised interview questions. From these interviews we developed 15 individual UDEs (3 for each department). A UDE was selected from each area to be further developed into 1 of 5 individual EC. The individual ECs were then reduced to 3 and then distilled into a single GC shown below in figure 2.

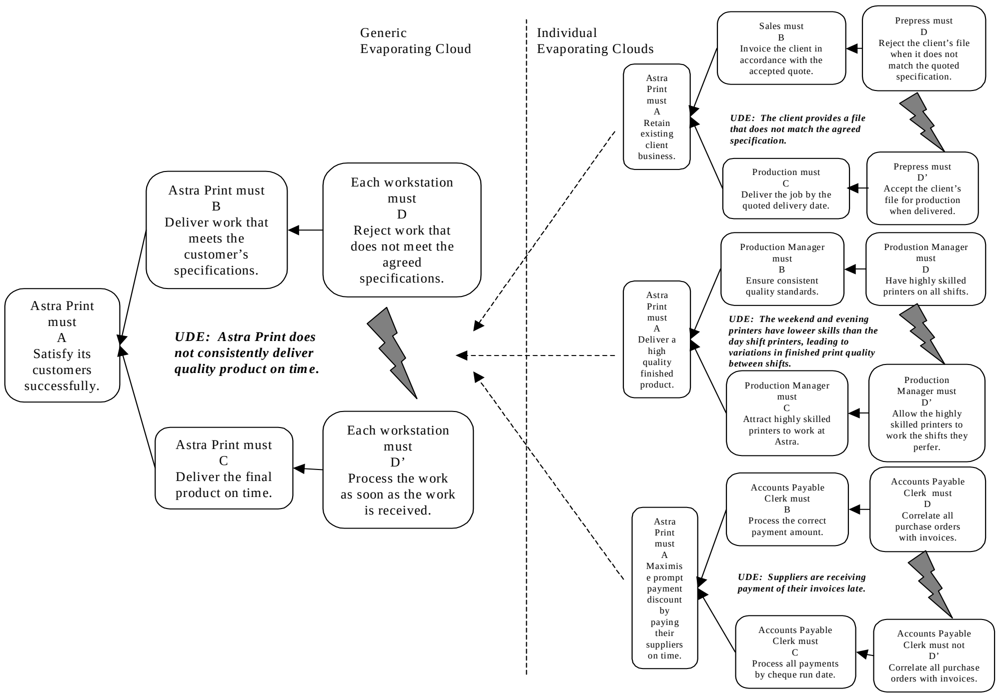


Figure 2. Individual and Generic Clouds for Astra

The clouds are built on *necessary condition* logic – the cause is necessary or required to create the result (Cox, 2003). Using the Generic Cloud as an example, the clouds are read as follows:

Our objective is to satisfy customers successfully, and in order to achieve this, we must deliver work that meets customer's specification. In order to deliver work that meets customer's specification, we must reject work that does not meet specification. On the other hand, in order to satisfy customers successfully, we must deliver the final product on time. In order to deliver the final product on time, we must process the work as soon as the work is received.

The generic cloud demonstrates how an individual UDE stems from a generic conflict. It is clear that there is a conflict present in this operating system. It is not possible to deliver the work on time (efficiency) and deliver quality work (effectiveness) at the same time if there are problems with the specifications of the work delivered to the workstation.

For each EC or GC the next step is to surface the assumptions underlying each connection (arrow). This is read for each connection A:B, B:D, A:C, C:D' as follows:

In order for Astra Print to A, satisfy its customers successfully, Astra Print must B, deliver work that meets the customer's specifications because;

AB1 Astra Print and the customer have agreed the specification.
AB2 The files supplied by the customer are correct.
AB3 The customer will reject work that does not meet specification.

In order for Astra Print to B, deliver work that meets the customer's specifications, each workstation must D, reject work that does not meet the agreed specifications because;

BD1 Processing work that does not meet agreed specification necessitates rework.
BD2 The next workstation cannot process work until the defect is corrected.

The surfaced assumptions represent the "conventional wisdom" of the operating system. This process surfaced 78 assumptions for the EC, reducing to 12 generic assumptions for the GC. Each assumption was tested for its validity by asking what action could be taken to invalidate the assumption and solve the conflict. Determining which assumptions were necessary (valid) provided the basis for the injections developed in the Future Reality Tree.

4.2.2 Sufficient conditions: The Current Reality Tree

The generic cloud forms the basis of the Current Reality Tree (CRT). The CRT represents the most probable chain of cause and effect, given a specific, fixed set of circumstances (Mabin et al., 2001). It is designed to provide the basis for understanding complex systems and helps to identify policies, measurements, and behaviours that contribute to the existence of the UDEs (Mabin et al., 2001).

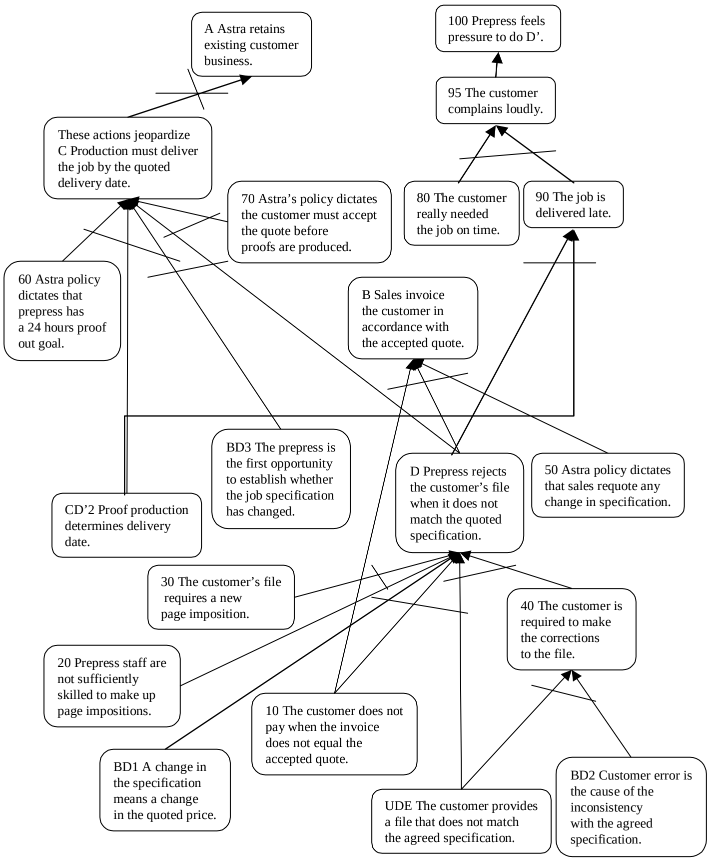


Figure 3. A Current Reality Branch for Astra's Prepress

The CRT is built on sufficient condition logic, the cause is sufficient to cause the results (Cox, 2003). Using a single branch of the CRT as an example, the tree is read bottom up using "if…and….then" logic:

If the customer provides a file that does not match the agreed specification, and the customer error is the cause of the inconsistency with the agreed specification, then the customer is required to make the corrections to the file.

Building the CRT identifies whether the casual logic is not sufficient and allows logic blocks to be added as required. The tree shows how doing 1 action D creates pressure on the operation to do the conflicting action D'. This process clarifies why the system UDEs occur, the Future Reality Tree tests possible solutions.

4.2.3 What to change to: The Future Reality Tree

The future reality tree (FRT) originates in injections (solutions) and ends in desirable effects (DE: positive system outcomes) that really reflect the opposite of the 'problems' (UDEs) in the current reality tree. The FRT determines whether proposed system changes will produce the DEs without creating negative side effects. It effectively tests new ideas before committing resources to implementation and serves as an initial planning tool (Mabin et al, 2001).

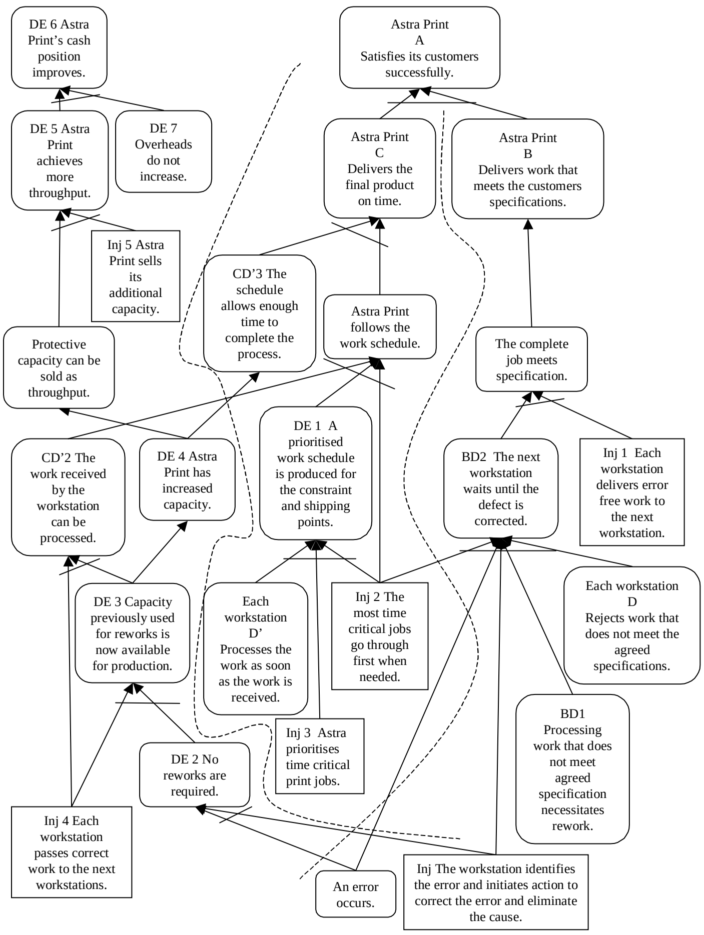


Figure 4. A Future Reality Tree for Astra

The FRT is also built on sufficiency logic. Using a section of the FRT as an example, the tree is as with the CRT read bottom up using if…and….then logic. This analysis provides a view of what the system needs to change to. The analysis developed 6 primary injections to achieve the goal, that Astra satisfies its customers successfully.

When the injections are uncovered they look like commonsense. In fact, it is hard to convince management or the employees that the 'simple' solutions (i.e. injection 1: each workstation delivers error free work to the next workstation), are the result of considerable analysis. If the problems identified, and the solutions developed, are handed to the organisation as a result of a consulting engagement it can be difficult for the organisation to accept the new "wisdom". Our approach was to present our findings to the organisation and demonstrate the effect of maximising individual workstations to the overall system's detriment by having them play the "penny game[7]". The penny game established the credibility of the analysis, as it proved that processing faulty work decreased throughput, and increased lead-time and inventory in the system.

4.2.4 How to cause the change: The Prerequisite Tree

The prerequisite tree (PRT) is a logic-based tool for determining the obstacles that block the implementation of a solution. Once the obstacles have been identified, objectives for overcoming obstacles can be determined (Cox, 2003).

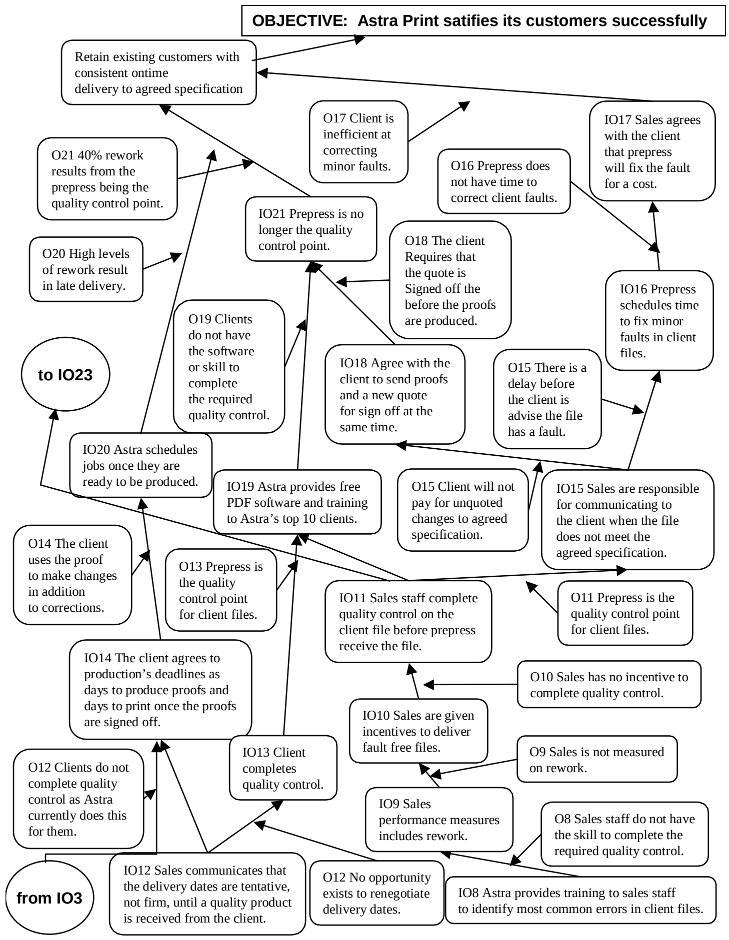


Figure 5. A Prerequisite Tree for Astra

The PRT is also built on necessary condition logic. Using 1 section of the PRT as an example, the tree is read from the bottom: In order to overcome the obstacle (O), the lower intermediate objective (IO) must be achieved, which allows the upper IO to be addressed:

In order to overcome the O 8 sales staff do not have the skill to complete the required quality control, the lower IO 8 Astra provides training to sales staff to identify most common errors in client files, must be achieved. This allows the upper IO 9 sales performance measures include rework to be addressed.

A main advantage to the TOC thinking tool approach is that it does not stop at suggesting solutions. Having convinced Astra that there was an opportunity to improve the existing operating system and that hidden capacity existed, the PRT allowed us to anticipate the obstacles to successfully implementing the solutions developed in the FRT.

4.3 Astra's results

As a result of this research project immediate changes were made, incorporating:

- Quality control: the sales staff identified "common" errors before the customer's file was accepted for processing.
- Job docket routing: the pre-press department pre-flighted jobs before they entered the queue for estimating.

- Scheduling: Astra initiated a priority scheduling system at the Raster Image Processor (RIP).

These changes required no additional staff or investment in equipment, but the reduction in the amount of “faulty” work substantially increased the available capacity. Astra was able to process 30% more plates in the month following implementation. Lawrence Evans, Astra’s Managing Director reported in a follow up to the implementation that:

“Astra hit a record month in July for sales, gross profit and net profit! Pre-press would normally process between 1800 and 2000 plates per month. In the month of July they were coping with the workloads and keeping ahead of the presses! They processed 2600 plates in July. We have now put a 4th shift on the 8 colour press and still seem to be keeping up. It is amazing how much the few small things we have done so far have affected our throughput capability.”

4.4 Theory of Constraints: Lessons learned

TOC is a way of thinking about the operating system and can produce breakthrough change. It is not for the faint hearted as a system review similar to the analysis and recommendations developed for Astra requires common acknowledgement that there is a problem, the commitment of all staff to solving the problem, and a significant amount of time.

5 Conclusions

TOC is a powerful analysis method that not only provides a process for identifying process problems, but also provides a framework that allows the communication to the layman of the problem and the solution.

When we presented the results of our investigation and analysis to Astra’s management and staff, the staff appeared complacent and bored. As we used the tools to outline the logic of how the parts of the process fit together to create a system-wide problem, they became more and more engaged. By the end of the presentation, the management and staff clearly understood that each individual workstation was part of a wider system. As a communications tool, TOC proved its worth within Astra.

TOC also represented a rigorous analysis tool. The analysis did take a considerable length of time, but by taking a systems view of the wider supply chain, a full analysis was facilitated.

As with any process improvement project, it is the application of the solutions resulting in profitability improvement that is the ultimate judge of success. Astra has been able to use the capacity identified in this study as the basis for a new business development venture, which has the potential to dramatically grow the business. The implementation of TOC has literally been worth millions of dollars to Astra.

6 Acknowledgments

The authors would like to take this opportunity to thank the large number of people who provided their time and commitment in completing the TOC project described here. In particular Jim Cox for his Socratic teaching style, Vicky Mabin for her persistence, Lawrence Evans and his staff for allowing us the opportunity to use these tools at Astra.

7 References

Astra Group: www.astra.co.nz

Atwater, J.B., and Chakravorty, S.S., 2002. “A Study of the utilization of capacity constrained resources in Drum-Buffer-Rope Systems.” *Production and Operations Management* 11, 2, 259-273

Cox III, J.F., Blackstone Jr., J.H., and Schleier Jr. J.G., Managing Operations: A Focus on Excellence, *North River Press, Great Barrington, MA* (2003)

Goldratt, E.M., Cox, J., The Goal: A process of ongoing improvement, 2nd Ed., *North River Press, Great Barrington, MA* (1992)

Mabin, V.J., and Balderstone, S.J., 2003.“The performance of the theory of constraints methodology: Analysis and discussion of successful TOC applications.” *International Journal of Operations and Production Management* 23, 5/6, 568-595

Mabin, V. & Davies, J., 2004. “A Case of Ethical Dilemma: A Multi-Framing Approach.” *International Business Trends: Contemporary Readings*, Fullerton, S & Moore, DL (eds), Academy of Business Administration: Ypsilanti, MI

Mabin, V.J., Forgeson, S., and L. Green. 2001. “Harnessing resistance: using the theory of constraints to assist change management.” *Journal of European Industrial Training* 25, (2/3/4), 168-191

Senge, P.M., *The Fifth Discipline: The Art and Practice of The Learning Organisation*, Currency Doubleday Publishing, (1990)

Sheu, C., Chen, M-H., and Kovar, S., 2003. “Integrating ABC and TOC for better manufacturing decision making.” *Integrated Manufacturing Systems* 14, 5, 433-441

Statistics New Zealand 2003 annual enterprise survey

---

[1] There are 2 types of complexity. Detail complexity: the sort of complexity in which there are many variables and dynamic complexity: situations where cause and effect are subtle, and where the effects over time of interventions are not obvious (Senge, 1990).

[2] TOC was originally conceived in the 1970s as a scheduling algorithm. In 2003 400 published works were found (Mabin and Bladerstone, 2003).

[3] Now in its third revised edition (Goldratt and Cox, 2004)

[4] Goldratt produced a number of texts including *The haystack syndrome: sifting information out of the data ocean* (Goldratt, 1990), *Its Not Luck* (Goldratt, 1994), *Critical Chain* (Goldratt, 1997), *Necessary But Not Sufficient* (Goldratt et al., 2000).

[5] Constraints may be physical constraints, market constraints or policy constraints (Cox et al., 2003).

[6] For the year ending 31st March 2005.

[7] This is a more elaborate version of the matchstick game from “The Goal” (Goldratt and Cox, 1992).

Forest Harvesting with Wildlife Corridors

Alastair McNaughton
Department of Mathematics
Tamaki Campus
University of Auckland
Private Bag 92019
New Zealand
a.mcnaughton@auckland.ac.nz

Xenia Chen
Clinical Trials Research Unit
School of Medicine
Tamaki Campus
University of Auckland
Private Bag 92019
New Zealand
x.chen@ctru.auckland.ac.nz

Abstract

It is required to optimize timber output subject to the maintenance of wildlife corridors. These allow endangered species the opportunity to traverse the forested area when they wish to. We show how it is possible to meet this requirement without substantially reducing the present net worth of the forestry estate.

1 Problem review

Large commercial forests are often situated close to the habitat of endangered species such as the New Zealand kiwi. These animals may well inhabit pockets of native bush on the fringes of the plantation. It is often the case that these pockets have become isolated. The small communities remaining are very vulnerable and face a grim future. However, the adjacent forest, while not favoured as a habitat, may provide a safe environment for travel between these pockets. It is often the case that these animals will travel freely through a forest provided there is a mature canopy over their heads. They will not traverse open ground or land that has been recently felled.

So the practical requirement is to modify the harvest scheduling in such a way that there is always at least one safe corridor of continuous stands of mature trees between any two pocket habitats. The model given here shows how this can be done without substantially reducing the present net worth of the forestry estate.

For the present study we use a hypothetical forest of 400 blocks with random size, establishment age, croptype, and adjacency. We assume operational factors force each block to be harvested in one complete fell. The forest is approximately square-shaped. We consider there are two areas of kiwi habitat on opposing sides of the forest. We want to produce a harvesting schedule that guarantees the existence of a corridor between these two habitat areas at all times. The corridor is not required to remain the same throughout the planning horizon. We will assume a felled area will remain unsuitable for crossing for 4 years after harvest. The output will indicate the size of the loss in revenue associated with this harvesting requirement and it will also allow the actual path to be observed from year to year.

2 The harvesting model

The basic harvesting model is the same as that used in all of my forestry work. It is a column generation model in which the decision variables each represent a road harvest plan. It is assumed that each block lies on a particular road. Each variable represents a harvesting plan for all the blocks on a certain road and encompasses all the time periods in the planning horizon. For example, a typical harvest plan on road 10 could be to harvest block 3 in time period 5 and block 4 in time period 8, with the other blocks on road 10 left unharvested. The integer variable x_{jn} represents the n-th plan generated for road j, and takes value 1 if this plan is implemented, and 0 otherwise. The objective coefficient associated with this variable is $-c_{jn}$. Here c_{jn} represents the yield and consequent revenue to be obtained from this plan along with the related harvesting and road maintenance costs. These are discounted so that they represent the present net worth. During the column generation phases, which occur repeatedly during the solution algorithm, new improved harvesting plans are constantly being devised, and new variables x_{jn} added to the model.

In order to model the corridor requirements, extra variables are included. We define y_{kmt} as a 0/1 decision variable which takes value 1 if a path lies from block k to block m in time period t. These are directed paths, so y_{kmt} is distinct from y_{mkt}. There is no objective coefficient associated with these variables.

Using a minimisation format, the model can be stated as:
Minimise

$$\sum -c_{jn}x_{jn}, \tag{2.1}$$

subject to

$$\sum_{\substack{k \\ x_{jn}\in H_{jkt}}} a_k x_{jn} + s_t \leq A_t, \quad t = 1, t^* \tag{2.2}$$

$$\sum_n x_{jn} = 1, \quad j = 1, R, \tag{2.3}$$

$$\sum_m \sum_{k\in S} y_{kmt} \geq 1, \quad t = 1, t^* \tag{2.4}$$

$$\sum_m \sum_{m \in E} y_{kmt} \geq 1, \qquad t = 1, t^*, \tag{2.5}$$

$$\sum_m y_{kmt} - \sum_m y_{mkt} = 0, \qquad t = 1, t^*, \quad k = 1, 400, \quad k \notin S, E \tag{2.6}$$

and $$\sum_{\bar{t}=t-g+1}^{t} \left(\sum_{x_{jn} \in H_{jk\bar{t}}} x_{jn} \right) + \sum_m y_{kmt} + \sum_m y_{mkt} \leq 1, \qquad t = 1, t^*, \quad k = 1, 400, \tag{2.7}$$

where
x_{jn} = the 0/1 decision variable associated with harvest plan n on road j,
y_{kmt} = the 0/1 decision variable associated with a possible path from block k to block m in time period t,
c_{jn} = the present net worth associated with plan n on road j,
$c_{jn}{}'$ = the cost of maintenance for road j associated with plan n,
a_k = the area of block k,
A_t = the area , in hectares, to be harvested in time period t,
s_t = one or more appropriately bounded slack or surplus variables,
H_{jkt} = the set of harvest plans on road j in which block k is harvested in time period t,
S is the set of blocks adjacent to the wildlife habitat on one side of the forest,
E is the set of blocks adjacent to the wildlife habitat on the other side of the forest,
g = the green up, that is the length of time which must elapse before the replanted trees are sufficiently mature as to provide acceptable cover for the wildlife,
R = the number of roads, and
t^* = the planning horizon.

The first two groups of constraints are basic to all my harvesting models. The constraints given in Equation 2.2 ensure there is a non-declining yield. Provided the forest management programme incorporates automatic replanting, these constraints also impose an age distribution requirement on the residual forest.

The constraints given in Equation 2.3 are called the plan constraints. They ensure that only one road harvesting plan may be included in the solution for each road in the forest. This ensures each block is harvested at most once during the planning horizon, which in this paper will be not more than one rotation.

The last four groups of constraints are particular to the corridor problem. The constraints given in Equation 2.4 ensure that there is a path which commences at a block adjacent to one wildlife habitat in every time period.

The constraints given in Equation 2.5 ensure that there is a path which ends at a block adjacent to the other wildlife habitat in every time period.

The constraints given in Equation 2.6 ensure that if a path leads into block k, then the path will also leave block k. The blocks adjacent to the wildlife habitats are excluded from this constraint. The purpose of these constraints is to ensure that a continuous path will always lie from set S to set E. This path will possibly meander about. It is because of this constraint that the path variables need to be directed.

The constraints given in Equation 2.7 ensure that a path may only enter or leave a block provided the block has not been recently felled. The duration of the period during which the land is too bare to provide cover for the wildlife is called the *green up*.

3 Performance levels from simulation trials

These trials involved a 400 block hypothetical forest as described in the introduction.

time periods in horizon	length of green-up 0	1	2	3	4	objective value (million$)	time (seconds)
4	0					8.5	33
4		1				8.5	36
4			2			8.6	35
4				3		8.6	22
4					4	8.5	48
6	0					11.7	55
6		1				11.7	72
6			2			11.7	70
6				3		11.7	77
6					4	11.7	73
8	0					14.2	92
8		1				14.2	113
8			2			14.2	144
8				3		14.2	167
8					4	14.2	151
10	0					16.2	169
10		1				16.2	189
10			2			16.1	283
10				3		16.2	202
10					4	16.2	200

Table 1: Output from simulation trials.

In Table 1 the category of 0 green-up represents the relaxed problem in which no path is required. This permits a simple comparison of objective values and computational time. It is immediately apparent that the requirement to provide a wildlife corridor has negligible impact on the objective value. It does, however, require the investment of a little more computational time. The output also includes a specific path detailed block by block, for each year. These paths will be shown in the talk. It is a pleasant surprise that in each case just one path is generated, even though the programme merely requires at least one.

4 Conclusions

It appears from the numerical trials that an operational solution has been found for a significant conservation problem. It is especially encouraging that this has been achieved without a significant decrease in forest revenue. This indicates a research programme which should be continued, ideally in the form of practical implementation. This offers a means of accomplishing a very significant conservation goal with no cost to forest management other than a little more time and care spent in the planning process.

References

[1] Chen, X., 2002. *Safe Corridors for Kiwi Birds*, Fourth year research project for B.Tech. (I.M.), University of Auckland.

[2] McNaughton, A.J., M. Rönnqvist and D.M. Ryan. 2000. *A Model which Integrates Strategic and Tactical Aspects of Forest Harvesting.* In System Modelling and Optimization, Methods, Theory and Applications, Edited by M.J.D. Powell and S. Scholtes, Kluwer Academic Publishers Boston, pp189-208.

[3] McNaughton, A.J., G.D. Page and D.M. Ryan. 2001. *Adjacency Constraints in Forest Harvesting*, proceedings of the ORSNZ, 2001, pp 9-15.

[4] McNaughton, A.J. 2002. *Optimisation of Forest Harvesting Subject to Area Restrictions on Clearfell*, proceedings of the ORSNZ, 2002, pp 307-313.

[5] McNaughton, A.J. 2003. *Adjacency constraints and adjacency branches*, proceedings of the ORSNZ, 2003, pp 13-20 .

[6] McNaughton, A.J. 2004. *Recent Progress on the Area Restriction Problem of Forest Harvesting* , proceedings of the ORSNZ, 2004, pp 139-143.

[7] Murray, A. 1999. *Spatial Restrictions in Forest Scheduling*, Forest Science 45(1), pp 45-52.

[8] Nelson, J., J.D. Brodie. 1990. *Comparison of a Random Search Algorithm and Mixed Integer Programming for Solving Area-based Forest Plans*, Canadian Journal of Forestry Research 20, pp 934-942.

[9] Torres-Rojo, J.M., Brodie, J.D., Sessions, J., 1991. *The use of relaxation to solve the Habitat Dispersion Problem*, Forest Research Laboratory Report, Oregon State University.

[10] Yoshimoto, A., Brodie, J.D., 1994 *Short- and Long-term Impacts of Spacial Restrictions on Harvest Scheduling with reference to Riparian Zone Planning*, Oregon State University.

E-lections and the price of democracy

John Paynter and Gabrielle Peko
Department of Information Systems and Operations Management
University of Auckland, New Zealand
j.paynter@auckland.ac.nz; g.peko@auckland.ac.nz

Abstract

The triennial general elections are arguably the largest logistics exercise in the country. The Chief Electoral Office staff expands from 15 to over 16,000 on Election Day (E-day). A complex schedule for finance, supplies, personnel, processes and training is mapped out. In this paper we identify processes that may be improved, and how the transition to electronic voting may enable this. With tight deadlines and untrained staff clerical errors occur. An analysis of these can indicate where additional resources should be allocated and processes improved. A transition to electronic elections can minimise the number of special votes that need to be cast, improve the scrutiny of the rolls and finally achieve faster completion of the count.

1 Introduction

The general election is one of the largest logistics exercises in the country. 16,000 people are employed on Election Day (E-Day) in 62 general electorates (where the votes are also collated for the five Maori electorates) with roughly 30 polling places per electorate (more for rural ones). Approximately 10 tonnes of supplies are required per electorate, with more delivered around E-day to provide the physical infrastructure (chairs, tables, polling stands and signage). The exact date of the election is not known until approximately six weeks out from the day announced by the Prime Minister; so much of the recruitment and training can not be done until then. The ballot (voting) papers are not finalised (i.e. who is standing and where) until three weeks out and the roll of electors is published at the same time (writ day). Supplementary (additional) and post-writ deletion rolls are published later. These all cause potential problems for the election processes.

A testament to the organisation and processes devised by the Chief Electoral Office is that things go smoothly and the results are known so quickly (bar the outcome of the Special Votes, SV) is. With approximately 80% of the population voting and 90% of this by way of casting Ordinary Votes (OV), the aim is to have the results reported by 9:15PM, just over 2 hours after the close of poll. In the following week the roll is scrutinised, and the Ordinary Votes recounted in the official count. In the meantime the SV declarations are catalogued and validated (qualified) and in the following week, integrated with the ordinary vote roll and counted. In this paper we examine some of the processes involved in order to highlight potential improvements.

2 Background

The country is divided into 62 general and 5 Maori electorates. A general election is held every 3 years to determine the party or parties that can form a government. Since 1999 the method used is Mixed Member Participation (MMP), prior to that the electoral method was First Past the Post (FFP). Under MMP each elector gets two votes – one to determine the local candidate who will represent them, the other to determine the overall party share in Parliament.

A census is held every five years (the next one is March 9 2006) and following this electorate boundaries are redrawn to give every electorate a similar population base. This can cause a change in Polling Place structures.

After writ day the ballot papers are printed and distributed (just) in time for advance voting to begin. Each electorate will have a number of advance polling places where voters who will be overseas, otherwise absent from the electorate or unable to vote on polling day can cast their vote. Potential electors can also change their enrolment details up to the day before the election. Because of this supplementary rolls (for new electors and those moving into an electorate) and post-writ deletion rolls have to be produced.

On voting day, those in the electorate, or in a neighbouring electorate that has a shared polling place, can cast an Ordinary Vote. These are simple to process and can be counted on the night (ordinary Advance Votes are counted on the afternoon of the election). Advance votes from outside the electorate, overseas and defence force votes, and special votes from other electorates or from the electorate where the voter does not appear on the printed roll can also be cast. All these are required to be qualified prior to them being counted. Conversely those votes that are disallowed (e.g. the voter has died or otherwise become ineligible) have to be withdrawn from the count. Where the voter has voted in an electorate from which they have subsequently moved (for more than one month prior to the election) then they are only entitled to a party vote (i.e. the ballot paper issued was not for a candidate in their current electorate). There is also the possibility that a voter has voted more than once (i.e. has a dual vote). Each of these cases has to be investigated and one or more of their votes potentially disallowed.

Ordinary Votes are recounted in the Official Count. The sum of the number of votes cast for each party must equal the sum of the votes cast for each candidate at each Polling Place (PP). The number of votes counted must not exceed the number issued at that polling place. Any discrepancies of more than a few votes must also be investigated. On election night (E-Day) some tolerance was allowed (a schedule is printed of the tolerance based on the number of votes cast). If the counts were outside the tolerance then the votes had to be reconciled again and recounted. If after the second count the tolerance was still exceeded the results were called to the electorate HQ rather than delaying the overall count.

These processes must be completed in the two weeks following the election. A detailed operations manual outlines all the steps that must be followed before, on E-day and the period after the election. In addition, emails and Action Notes may be sent out advising of changes of procedures or giving clarifications and further advice.

3 Method

The authors had worked in the 2002 General Election as Polling Place Managers, both having participated in other roles in previous (pre-MMP) elections. In addition they had undertaken research on the potential for electronic elections.

For the 2005 election they worked for the Mt Albert Electorate as trainers (training Election Day workers and advance vote facility staff), managers (Mobile and Polling Place) and team leaders; one team scrutinising the rolls and the vote count, the other processing special vote declarations. Mt Albert represents a typical urban seat – a breakdown of the peculiarities of each electorate is available at www.elections.org.nz. Voting and other statistics from past and the present elections are also available at www.electionresults.org.nz. Thus the observations made are taken from a detailed analysis of a single urban electorate and the overall patterns from the country as a whole. While other measures are taken, they are not available for analysis. For example, a statistics officer is assigned to each electorate - they visit a PP on the day and also record the incidence of Informal Votes during the official count.

4 Results

Electorate Number:	27	**Final:**	**Yes**	
Polling Places Counted:	44 of 44 (100.0%)	**Votes Counted:**	32,342	
Winning Candidate:	CLARK, Helen Elizabeth (LAB)	**Majority:**	14,749	
Parties		**Candidates**		
		BAGNALL, James	IND	83
New Zealand First Party	1,089	BATCHELOR, Julian	NZF	746
Green Party	2,985	CARAPIET, Jon	GP	1,485
Labour Party	17,501	CLARK, Helen Elizabeth	LAB	20,918
United Future New Zealand	649	GORDON, Tony	UFNZ	529
National Party	8,488	MUSUKU, Ravi	NAT	6,169
Direct Democracy Party	10	PONGA, Howard	DDP	30
		RAVLICH, Anthony George	HR	47
ACT New Zealand	651	SEYMOUR, David	ACT	746
		TAYLOR, Erik John	IND	29
		WHITMORE, Daphna Kaye	ACAP	79
Destiny New Zealand	157	WILLIAMSON, Anne	DNZ	337
Jim Anderton's Progressive	356	WILSON, Jenny	JAP	233
99 MP Party	6			
Alliance	22			
Aotearoa Legalise Cannabis Party	43			
Christian Heritage NZ	40			
Democrats for Social Credit	3			
Libertarianz	19			
Mâori Party	168			
New Zealand Family Rights Protection Party	20			
OneNZ Party	0			
The Republic of New Zealand Party	5			
Party Informals	130	**Candidate Informals**		316
TOTAL	32,342	**TOTAL**		31,747

Table 1: Voting results for the Mt Albert electorate, 2005 General Election

While the Mt Albert candidate (Helen Clark) majority (the second highest in the country) would suggest it is not indicative of the overall population in terms of voting patterns, an analysis of the votes may be indicative of any problems in the processing. Indeed, having a candidate and party with a high count leads to problems in the counting and handling of the votes. E.g. during the Official Count one set of votes got in the wrong pile and the count had to be redone. Anecdotes of other problems were heard from different electorates. We will now consider the processes that occur within a general electorate in chronological order from pre-election to post-election.

4.1 Supplies

It appeared that some issuing points (i.e. the individual poll clerks) were not given the correct supplies. E.g. for some shared polling places, those issuing ordinary votes did not have the corresponding ballot papers and/or rolls for the second electorate. Most of the missing supplies were delivered before the start of polling or within a short time of the polls being opened. No voter would be disadvantaged by the supply omissions, as they could be either redirected to another polling place issuing those (ordinary) votes or be given a special vote for that electorate.

Difficulties arose during scrutinising the rolls and ensuing processes, as the books of empty and part-used ballot papers did not necessarily have the correct issuing point stamped on the cover. This made it harder to reconcile the number of votes issued against those counted. It also made it harder later to find the stubs for those votes to be reallocated as post-writ deletions or apparent duals.

4.2 Early votes

Advance voting starts seventeen days before the election (on Aug 31st compared to Sep 17th). There is a rush to send out supplies and little time to train the advance polling place staff. Consequently disproportionately more errors appeared to be made for votes issued under these circumstances. It should be stressed that this does not necessarily lead to the votes being invalidated, rather it requires more effort to track, document and rectify any apparent anomalies in the votes. Mistakes occur in marking off the roll and in issuing special takeaway votes.

4.3 Election Day

Shared electorates
At the shared polling places errors included issuing the wrong ballot paper. E.g. A voter may be given an Auckland Central rather than Mt Albert voting paper at shared issuing points. In the Tamaki electorate, for example, there may be confusion between the general 'Tamaki' electorate and the Maori electorate 'Tamaki Makaurau' whose boundaries overlap. Where this happened the ballot papers had to be found and withdrawn from the official electorate count for that PP. Subsequently they are put into the Party-Vote only count.

Takeaway votes
It appeared that some polling place managers, who are authorised to issue takeaway votes – (where the voter is not present) marked the voters off the (ordinary) roll. Hence when the Special votes were scrutinised they were recorded as apparent duals that had to be investigated and documented. This occurred in some of the Election Day polling places and the advance vote facility places.

Apparent Duals
Although there is provision to question a voter at the time the vote is issued and to set aside any such vote for investigation in the apparent duals (Envelope 2), this is rare.

More likely the second and subsequent votes are identified during the scrutiny of the rolls (as in the case of the takeaway votes above).

Roll marking

The issuing of Easy Vote cards (used by 85% of the voters) should make this process simple, with fewer errors. That is, the voter brings an easy vote card that details the roll on which they appear, their name, page and line number. Despite this most investigations of apparent duals are likely to be clerical errors in marking the rolls. It is not possible to record if easy vote cards were used, in cases where clerical errors were made, or if the person stated (or wrote down) her name.

Completing the stubs

As is the case for the rolls, the wrong line and page number may be written on the stubs of the ordinary votes. Other stubs were left blank, twinked out or ripped off when the ballot paper was issued. All these make it impossible to investigate apparent duals and/or remove duals and post-writ deletions. In some instances the ballot paper number on the ballot paper was torn off so it could not be read.

Reconciliation

As detailed earlier, the number of ordinary votes issued should equal those counted in the election night count. Discrepancies in the count included votes being in the wrong ballot boxes (these should be sorted before the count begins). In some instances, voters may walk out with ballot papers. Hence the number issued will exceed the number counted. Discrepancies in the number issued can occur for a number of reasons. Votes may be issued from the back of a partly-used pad and not be counted. The number of votes issued, rather than the number of votes unissued may be subtracted from the number in a ballot book. Special vote pads may be included in the count of ordinary votes issued or conversely, one or more pads of issued ordinary votes may be omitted from the reconciliation. Some special votes may have been issued instead of ordinary votes for the same electorate.

Spoilt votes

Very few votes were spoilt. Most issuing points had none, some had one or two. These occurred when the issuing officer pulled out more than one ballot paper, pulled out one for the wrong electorate or the voter miss-marked the ballot paper and asked for another.

Informal votes

Some voters made their intentions clear despite not completing the ballot papers properly. Others had chosen two candidates and/or two parties (perhaps this is because they are told that they have ‘two votes’) or left them blank. As some voters only selected a party or only a candidate, we assume that the choice was deliberate. An analysis is made by the statistics officer and is out of the scope of this paper.

4.4 Scrutiny of the rolls

Scrutinising the roll is the most time consuming and apparently error-prone of the post-election day processes. In 2002 a system to scan the rolls was used that reportedly resulted in 39,000 apparent dual votes. The 2005 election reverted to a manual process. On the Sunday following the election the post-writ deletions (D) and transfers (TRF) are marked in black on the master roll. This is repeated for subsequent changes on the following Wednesday (E-day+4). On the Monday six teams of two people each, work with a subset of the master roll (four teams of two for the Maori rolls). The first team start with the first alphabeticised portion of the roll and pass on each roll in turn (there may be 130 rolls in the electorate – one for each ordinary issuing point) to the next pair. The master roll is marked in blue for each voter found on a certified (issue-point) roll.

Although in theory teams should work at the same pace, it is necessary to move faster pairs to the first positions. The names are not distributed equally throughout the electorate. That is, different ethnic groups congregate in different suburbs. Hence Somalis may be in Owairaka with a preponderance of names starting in A or M; Chinese in Avondale (W-Z), Indians in Sandringham (P and S) and Islanders in Morningside, Kingsland and Arch Hill. Hence it was not unusual for two pairs to be idle while an earlier pair work through their section of the roll.

When a roll was found to contain a voter who had been already marked as voting, it was necessary to look back through all the earlier rolls until the one marked earlier could be found. Each day during the scrutiny process this step was repeated (with more rolls to potentially search over time). This was repeated for the Wednesday Post-Writ deletions and again when the special votes were added to the master roll. To save time during the scrutiny of the special votes, two rolls were used. In some instances the voter was recorded on both due to clerical error. Over 100 apparent dual voters had to be investigated and the ballots for over 60 post-writ deletions found. A number of post-writ deletions existed in the special votes, but the corresponding ballot papers were in the envelopes attached to the declarations and did not have to be pulled from the count at that stage.

In the case of apparent dual ordinary votes, confirming that the vote existed, by checking the stubs of the ballot papers caused a new set of problems. Due to the problems issuing supplies not all the IPs had the correct rolls and books of ballot papers. In some cases it was necessary to search through all the stubs in the ballot books used in the polling place. Some stubs had not been completed, despite the mantra of "Page, Line, Sign, Sticker, Stamp" being drilled in during training the issuing officers. In some instances the ballot paper number had been torn off the stub when the ballot paper was issued. Thus it was not possible to confirm that a ballot paper had been issued to the voter and the number of the ballot paper.

Searching for the ballot papers within the votes sorted by Candidate and Party for each Polling Place was time consuming. The ballot paper could be checked for the correct IP stamp (assuming it was recorded correctly). To confirm the ballot paper number one could try to read the reverse side of the ballot paper (faster) or take the black sticker off that covers the number on the ballot paper. Dual votes were pulled out and filed under section 177 and post-writ deletions placed in another envelope for the Party-only count.

4.5 Special votes

At the time of writing an analysis of the special votes is not available to the authors for the country as a whole. The results for Mt Albert are shown in Table 2. The results for the 2002 election showed that the major cause for the vote being invalid was because the voter was not on the roll. This was not surprising as interactions with people wanting to vote in the polling place on election day revealed that some did not know their name (or rather the name under which they were registered on the roll), their electorate and whether they were on the General or Maori roll. When the declarations were checked and the voter could not be found on the roll the declaration was sent to the Registrar of Electors (RoE) for each electorate for validation. The RoE did a commendable job in finding such voters on the roll. Often they came back with a completely different name than that on the declaration form, the match being made by address and date of birth. The return from the RoE shows a voter registration number. This is not shown on the certified roll, or on the easy vote card.

Not enrolled anywhere	311
Vote received late	6
Not signed by authorised witness	5
Not signed by voter	0
No ground stated	1
Dual votes	4
Post writ deletions	0
No declaration enclosed	6
No ballot paper enclosed	5
Total special votes disallowed	**338**

Table 2: Special declaration Votes Disallowed Report for Mt Albert (2005)

4.6 Investigations

Each apparent dual vote is recorded on a spreadsheet with one line for each apparent vote. In addition it is necessary to record the investigation of the circumstances on a M169-INV in each instance. The phone book was used to look up the voter's phone number and the phone number of a voter with a similar name who had not voted. In most cases, there was no phone number under that name. It was then necessary to look up the habitation index to find other people from that address and look up the phone book for those surnames. The other voter or the apparent dual voter could then be asked for details of where they voted. A checklist is provided to guide this process. Guidelines also suggest looking on the roll for voters on the lines above, below and on the opposite column and opposite page for errors when marking the roll.

Perhaps the best clue to determining clerical and similar human errors is to look at the patterns of the apparent duals sorting by polling place and issuing point. From this it is possible to judge where errors are occurring. For example, some polling places may be issuing takeaway votes (a clue is to the occupation – retired – and age of the voter) and mistakenly marking them off the roll as an ordinary vote. To be fair though, finding several entries under one person may indicate a busy polling place and issuing point. Hence care must be taken in making such generalisations.

4.7 The Official Count.

This process is fairly straightforward and satisfying. Care must be taken when the number of votes for Polling Place or other large counts (e.g. the Advance Count or Specials issued on the day) are high.

It is difficult to count and physically manage the number of ballot papers, particularly when they are unevenly distributed by candidate and / or party. A high number in the case of a PP will also hold up the count on election night. Consideration can be given to establishing a nearby PP to reduce the load.

5 Discussion

Several avenues for improving the current situation can be suggested based on the procedures used and the error rates encountered. These include improved training, new and improved procedures and the use of technology for both the existing system and the future of electronic voting.

5.1 Training

Many problems are due to human error, despite for the most part clear instructions in the procedures manual. It is unlikely that additional training would overcome problems such as incorrectly completed ballot paper stubs and mistakes in marking off the roll. Clearly the polling place managers (PPM) on the day should check, particularly at the beginning of the day, the work done by the issuing officers. Other errors such as those encountered in early votes can be corrected by better training (but time is usually the essence here, as there is a tight schedule and the real versions of the papers are not available until the day). Looking at such problems as encountered in the takeaway votes would enable the personal instruction manuals (PIMS) for EOs and PPMs to include these under trouble shooting and what not to do.

5.2 Procedures

The procedures for scrutiny of the roll should be revised and experiments run to determine the most efficient ways to conduct this process. For instance, improved technology solutions to electronically scanning the roll should be trialled. Another approach would be to electronically enter the page and line number for each roll marked. The resulting name, address and occupation should be checked by one of the pair. The other advantage of this approach is that it would be easy to find the roll on which the first of the apparent duals had been recorded or the roll pertaining to the post writ deletion report available on the Wednesday.

In the case of the special votes the current system requires much data entry. Some of this could be avoided by electronic links to the registrar of electors so that the data returned by them did not have to be rekeyed. The data entry forms too appear to be archaic and related to the days of dumb terminals and 23 lines of type per screen.

5.3 Electronic elections

It is clear that electronic voting offers the best long-term solution to speeding up the vote count, minimising errors and the costs of running the elections. Often new processes and technologies are trialled in local government elections first (e.g. Single Transferable Vote, STV in 2004).

Local government is currently reviewing the existing on-line voting systems and it is envisaged that such system will be ready for 2007 elections and e-polling. This will require the suppliers and local government to examine, develop and recommend security, communication and accessibility standards for a regulatory regime for Internet voting (LGOL, 2004).

However, an analysis of the capabilities of local bodies suggests that they will not be able to meet the target for electronic elections by 2007 (Dunayev, 2005). Fung (2004) too found that the majority of sites provided detailed information and good content organisation (73.1% in 2003 and 74.1%% in 2004). However, not many of the sites provide customer-focussed activities such as e-subscription and e-consultation. Only one council, Auckland Regional Council, provides an on-line forum for discussion and sharing of ideas. None of the sites provide any e-democracy although some sites have put up information about Elections 2004.

Other impediments to such elections involve changing people’s attitudes (Sharkey and Paynter, 2003). The lack in New Zealand of a national ID, and entrenchment of this philosophy in the Privacy Act (1993), are also possible impediments. It is interesting to note though that there is a registration number used by the electoral office to record individual voters. As this would be used for the purpose for which it was intended (voter registration and voting) and the procedures are already in place to use it to register voters, using it would be a short step.

A high number (10%) of special votes were disallowed. Other voters could not cast a vote as they did not know their name or electorate seat, or whether they were on the Maori or General roll. Computer technology could be used to confirm the elector's identity. Within the polling places, even without an electronic linkup, CDs containing the details of all voters on the roll could be issued to PPMs with laptops at each polling place. Information about voters could be searched based on their name, address and date of birth and the voter registration number accessed together with their roll details. Either that number or the roll/page/line could be recorded on the ballot paper stub (for that electorate and any shared electorates) or the special vote declaration. In the first instance this could be used by the PPM and/or EO to give the voter their details to take to the issuing officer. The next stage would be to automatically update the roll electronically with this information. Once this was done online, the problems of apparent dual votes and post writ deletions would be largely overcome. This could be a transition towards full electronic voting where the vote as well as the voter was recorded.

6 Conclusions

The 2005 General election ran smoothly and the official count was completed on time, despite the closeness of the result. Nevertheless the compressed time period, the complexity of the processes involving several organisations and the room for human error suggest that improvements can be made to current processes. These include better training and supervision at the beginning of the advance voting and Election Day voting, and refinement of existing processes, particularly in the scrutiny of the roll. E-enabled elections will allow the quicker announcement of the results and will generally improve the efficiency of the electoral process (Raynsford and Beecham, 2002). It would appear that it would be best to make a gradual transition in the processes to accomplish this. First electronic versions of the roll could be made available at the Polling Place, then electronic scrutiny of the roll, then automatic recording of the voter at the Polling Place and finally electronic casting of the ballot papers themselves. There are both technical and user issues to overcome in order to gain acceptance of this process.

Acknowledgements

The authors wish to acknowledge the leadership and careful documentation of procedures that enable the elections to be run so smoothly. This comes from independent commissions of the Department of Justice, such as the Chief Electoral Office and the Registrar of Electors. Credit also needs to be given to electorate and Election Day staff that work long hours under pressure to get the job done. Finally, thanks should be given to the voters who by and large cheerfully do their democratic duty in voting and ensuring a corruption free and independent society.

7 References

Chief Electoral Office, (2005), Official Results Count – Mt Albert, accessed 31 Oct. 05http://www.electionresults.org.nz/electorate-27.html

Chief Electoral Office, (2005), Operations Manual

Dunayev, A. (2005) Electronic Local Government Elections in New Zealand, unpublished Research Essay, Department of Information Systems and Operations Management, University of Auckland

Fung, M. (2004) E-Government in New Zealand: Local Body Strategies, unpublished Masters Thesis, Department of Information Systems and Operations Management, University of Auckland

(LGOL, 2004), Local Government Sector, accessed 31 October 2005, www.lgnz.co.nz

Raynsford and Beecham, (2002) e-gov@local: towards a national strategy for local e-government: consultation, accessed 31 Oct 2005.Office of the Deputy Prime Minister http://www.odpm.gov.uk/index.asp?id=1137927

Sharkey, E. and J. Paynter (2003). Factors influencing the uptake of Online voting in NZ. CHINZ '03: the 4th annual conference of the ACM Special Interest Group on Computer-Human Interaction, New Zealand Chapter Dunedin. 3-4 July 2003, 121-122

Solving Quadratic and Regression Programming

Dong Q. Wang
School of Mathematics, Statistics and Computer Science
Victoria University of Wellington, New Zealand
Dong.Wang@mcs.vuw.ac.nz

Abstract

This paper will discuss the interaction between multivariate analysis, least squares regression and non-linear programming, and the relationship between the quadratic programming and multivariate methods are explored based on statistical theory and optimization techniques. The parameter estimations of two statistical models will be studied by simplex method and the Kuhn-Tucker theorem. The optimization problems in mathematical programming are reformulated and solved using regression and multivariate algorithms.

1 Introduction

Optimization problems of a non-linear programming arise in almost all areas of industry or society, including product and process design, production, logistics, traffic control and strategic planning. There are a few algorithms to solve optimization problems for non-linear programming, for example, separable programming is expressed as a sum of single variable functions to find optimal solutions (known as a modified version of the simplex method); geometric programming is approximated by a generalization of the arithmetic-geometric mean inequality to solve algebraic non-linear programming; Algorithms for solving special forms of non-linear programming have been developed, especially for quadratic programming [5, 6]. This paper we will focus on the quadratic programming (QP) problem.

Quadratic programming is concerned with the problem of minimizing (or maximizing) a quadratic objective function subject to linear inequality (or equality) constraints with non-negative values for unknown variables.

For optimization problems of quadratic programming, we will consider the optimum value (maximum or minimum) of a function $Z(\beta_0, \beta_1, \ldots, \beta_p)$ of $p+1$ parameters $\beta_0, \beta_1, \ldots, \beta_p$. Due to the fact that max $Z = -$ min Z, we need only consider minimization problems. A typical mathematical programming problem (MPP) consists of a single objective function, representing either a profit to be maximized or a cost to be minimized, and a set of constraints that circumscribe the decision variables.

2 Quadratic Programming and Least Squares

Consider a quadratic programming problem

$$\begin{array}{rcl} Minimize \quad Z(\beta) & = & b^T\beta + \beta^T D\beta \\ subject\ to \quad A\beta & \geq & C \\ \beta & \geq & 0, \end{array} \tag{1}$$

where $A_{k\times(p+1)}$, $(k \leq (p+1))$, $D_{(p+1)\times(p+1)}$ are matrices, $C_{k\times1}$, $\beta_{(p+1)\times1}$, $b_{(p+1)\times1}$ are column vectors, rank $(A) = k$ and D is a *symmetric*, positive definite matrix, which can be decomposed as $D = L^T L$ [1], where L is a real upper triangular matrix with positive diagonal elements and can be obtained by the Choleski decomposition of D. The optimal value $Z(\beta)$ and solutions $\hat{\beta}$ of QP (1) can be found based on the Wolfe's method [7] which used the Kuhn-Tucker *Necessary and Sufficient conditions* (NSC), i.e. any point satisfying the NSC is solved for quadratic programming (1) through steps: (a) to write a standard form; (b) to get Lagrange function; (c) to check the Kuhn-Tucker NSC and (d) to use phase I and simplex method to find a feasible solution [4, 1].

In multivariate linear regression (LR) we consider the squares of difference between the predicted and observed values and add up these squared differences across all the predictions, we get a number called the least squares (or sum of squares error (SSE)). From a statistical point of view we want the SSE to be as small as it can possibly be, i.e. minimizing SSE with the constraints of non-negative variables β.

$$Minimize \quad SSE(\beta) \;=\; \sum_{i=1}^{n} \varepsilon_i^2 = (Y - X\beta)^T(Y - X\beta), \tag{2}$$

where $\{\varepsilon_i\}$ is the residuals. A general least squares linear regression problem (i.e. Regression programming) is obtained by

$$\begin{array}{rcl} Minmize \quad Q(\beta) & = & (Y - X\beta)^T(Y - X\beta) \\ subject\ to \quad A\beta & \geq & C \\ \beta & \geq & 0, \end{array} \tag{3}$$

where $\beta^T \in \mathbf{R}^{(p+1)}$ is the unknown and nonnegative vector, $X_{n\times(p+1)}$, $(n \geq (p+1))$, $A_{k\times(p+1)}$, $(k \leq (p+1))$ are constant matrices, and $Y_{n\times1}$, $C_{k\times1}$, $\varepsilon_{n\times1}$ are column vectors. Moreover $\varepsilon \sim N(0, \sigma^2 I)$, $X^T X \geq 0$ and rank $(A) = k$. The solution of the linear regression with constraints (LRWC) is a subject of the Karush-Kuhn-Tucker theorem.

Let the column vector $X = L$ be the explanatory variable and the column vector $Y = -\frac{1}{2}(L^T)^{-1}b$ be taken as the response variable in multivariate regression analysis (3). We know that quadratic programming is concerned with minimizing a convex quadratic function subject to linear inequality constraints. The unique

[1] Matrices D and A have a well defined structure, with D often being a covariance matrix of random variables in statistical analysis.

solution of a quadratic programming problem (QPP) (1) exists and provides that the feasible region is non-empty (the QP has a feasible space), and the relative minimum optimal value is also a global optimal value. Since $\{\beta : \beta \in \mathbf{R}_+^{(p+1)}, A\beta \geq C\}$ is a closed convex set, then $\hat{\beta}^*$ is an unique optimal solution for model SSE (2). A theorem is obtained based on Wang-Chukova-Lai's results [5, 6] as follows

Theorem 2.1: Let $\hat{\beta}^{*T} = (\hat{\beta}_0,\ \hat{\beta}_1^*,\ \hat{\beta}_2^*,\ ...\ \hat{\beta}_p^*)$ be the least squares estimators without constraints in a linear regression model (2) and $\hat{\beta}^T = (\hat{\beta}_0,\ \hat{\beta}_1\ \hat{\beta}_2\ ...\ \hat{\beta}_p)$ be the least squares estimators in the model (3) (or QP (1)), both $\hat{\beta}^T$ and $\hat{\beta}^{*T} \in \mathbf{R}_+^{(p+1)}$, then we have

(i) The relationship of optimal values between QP (1) and LS (3) is given by

$$Z(\hat{\beta}) = Q(\hat{\beta}) - \tfrac{1}{4} b^T D^{-1} b$$

(ii) The relationship of the optimal solutions of SSE (2) and LS (3) is

$$\hat{\beta} = [I - (X^TX)^{-1}A^T[A(X^TX)^{-1}A^T]^{-1}A]\hat{\beta}^* + (X^TX)^{-1}A^T[A(X^TX)^{-1}A^T]^{-1}C$$

The details of the full proof are given in the Appendix.

3 Solving the Quadratic Programming Problem

We will apply the **Theorem 2.1** to solve the QPs (1), (2) and (3).

3.1 Linear Programming to Solve the QP (1) $Z(\beta)$

Solving QP (1) based on the linear equation algorithms, from QP (1) we used slack variable vector $S_{k\times 1} = (s_1, s_2, ..., s_k)^T$ for constraints to get standard form, and getting $A\beta - S = C$, therefore QP (1) is written as

$$\begin{array}{ll} \textit{Minimize} & Z(\beta) = b^T\beta + \beta^T D\beta \\ \textit{subject to} & A\beta - S = C \\ & \beta \geq 0 \quad \textit{and} \quad S \geq 0, \end{array} \tag{4}$$

The Kuhn-Tucker theorem shows that if $\beta^{(0)}$ is an optimal solution for the QP (1) then $\frac{\partial Z(\beta)}{\partial \beta}$ must equal the linear combination of selected rows of matrix A, i.e. there exists a constant vector d so that $\frac{\partial Z(\beta)}{\partial \beta} = A^T d$, where vector $d = (d_1, d_2, \cdots, d_k)^T)$ contains k variables dual to slack variable vector S, and the conditions are satisfied if $s_i = 0$, we have $d_i = 0$; if $s_i \geq 0$, then $d_i \geq 0$, therefore we obtain a condition $S^T d = 0$ and

$$\frac{\partial Z(\beta)}{\partial \beta} = b + D\beta = A^T d. \tag{5}$$

Since matrix D is non-singular from the above equation we have

$$\beta = D^{-1}A^T d - D^{-1} b,$$

putting it into the constraints of QP (4), a linear programming is given by unknown variables S and d,

$$\begin{array}{ll} & AD^{-1}A^Td - S = AD^{-1}b + C \\ subject\ to & S^Td = 0 \\ & S \geq 0 \quad and \quad d \geq 0, \end{array} \tag{6}$$

The following algorithm may be used to solve the QP $Z(\beta)$: firstly solve goal programming problem (6) (with unknown variables S and d) to find the constants d; secondly use the relationship $\beta = D^{-1}A^Td - D^{-1}b$ between β and d to get the optimal values β, then finally the optimal solution is found for QP (1).

3.2 Solving QP (2) SSE(β) Using the Phase Method

Considering QP (2) we differentiate $Q(\beta)$ with respect to each of β and set equal to zero, i.e.

$$\frac{\partial Q(\beta)}{\partial \beta} = X^TY - (X^TX)\beta = 0_{(p+1)\times(p+1)}, \tag{7}$$

this produces the so-called normal equations in multivariate linear regression. Hence normal equations are obtained from (7)

$$(X^TX)\beta = X^TY \quad \text{and} \quad \hat{\beta}^* = (X^TX)^{-1}X^TY.$$

Consider artificial variables $R^T = \{R_0, R_1, ..., R_p\}$, then QP (2) becomes

$$\begin{array}{ll} Minimize & SSE(\beta) = (Y - X\beta)^T(Y - X\beta) \\ subject\ to & (X^TX)\beta + R = X^TY \\ & \beta \geq 0, \quad and \quad R \geq 0 \end{array} \tag{8}$$

According to dual theorem, the dual problem of QP (8) is found

$$\begin{array}{ll} Maximize & SSE^{(I)}(\beta) = -\mathbf{1}^TR = -(R_0 + R_1 + ... + R_p) \\ subject\ to & (X^TX)\beta + R = X^TY \\ & \beta_{(p+1)\times 1} \geq 0, \quad and \quad R_{(p+1)\times 1} \geq 0 \end{array} \tag{9}$$

where $\mathbf{1}^T = (1, 1, ..., 1)$ and $R_j = (\sum_{i=0}^{n} x_{ji}y_i - \sum_{k=0}^{p}\sum_{i=1}^{n} x_{ji}x_{ki})\beta_j$. For given observations X and Y, the initial table is obtained by

BV	β_0	β_1	$\cdots$	β_p	RHS
$SSE^{(I)}$	$-\sum\sum x_{ki}$	$-\sum\sum x_{1i}x_{ki}$	$\cdots$	$-\sum\sum x_{pi}x_{ki}$	$\sum\sum x_{ki}y_i$
R_0	n	$\sum\sum x_{1i}$	$\cdots$	$\sum\sum x_{pi}$	$\sum y_i$
R_1	$\sum x_{1i}$	$\sum x_{1i}^2$	$\cdots$	$\sum x_{1i}x_{pi}$	$\sum x_{1i}y_i$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
R_p	$\sum x_{pi}$	$\sum x_{pi}x_{1i}$	$\cdots$	$\sum x_{pi}^2$	$\sum x_{pi}y_i$

Using the simplex method the optimal solutions can be obtained from above table.

3.3 Using Least Squares to Solve the QP (3) Q(β)

A stepwise algorithm for searching for the solution to a QP (1) is explored on the basis of statistical theory [5]. It is shown that quadratic programming can be reduced to an appropriately formulated QP (3) with equality constraints and non-negative variables. This approach allows us to obtain a simple algorithm to solve a QP. The applicability of the suggested algorithm is illustrated with some numerical examples in [5]. Quadratic programming with zero-one variables is provided in [3], the problem can be reduced to search the extreme points of a zonotope.

4 Examples

Example 1:

Consider the example with $p+1$ and $n = 10$ observations $\{(x_i, y_i), i = 1, 2, ..., 10\}$ which given by the following table.

x	12	18	24	30	36	42	48
y	5.27	5.68	6.25	7.21	8.02	8.71	8.42

Note the matrix

$$X^T = \begin{pmatrix} \mathbf{1}^T \\ \mathbf{x}^T \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 12 & 18 & 24 & 30 & 36 & 42 & 48 \end{pmatrix}$$

The linear regression model (2) can be rewritten as $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ and $\beta^T = (\beta_0, \beta_1) \geq 0$. Using the SSE to minimize $SSE(\beta)$ with respect to β_0 and β_1, QP (2) is obtained

$$\begin{aligned} &\textit{minimize} \qquad SSE(\beta) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \\ &\textit{subject to} \qquad \beta \geq 0, \end{aligned}$$

applying QP (5) and (6) the initial table is obtained by

BV	β_0	β_1	R_1	R_2	RHS	Ratio
SSE^I	-217	-7518	0	0	-1640.04	–
R_0	7	210	1	0	49.56	0.236
R_1	210	7308	0	1	1590.48	0.2176

Using simplex method we have the optimal solution $\hat{\beta}^T = (\hat{\beta}_0, \hat{\beta}_1) = (3.994, 0.103)$.

Example 2:

This example was given in Fang-Wang-Wu in 1982 [2] for $n = 9$, and $p+1 = 6$. In order to manufacture concrete; asphalt, big and small stones, crushed stones, grit, sands and rock powder are required. Denote these variables as $X_0, X_1, \ldots, X_5$ respectively. Different types of sieves are used to filter these elements and the mixture Y, of them. The percentage of each of these variables that pass through the sieves are set to be $x_0, x_1, \ldots, x_5$ and y respectively. Using $\beta_0, \cdots, \beta_5$ to represent the percentage of these six elements in the mixture, they should satisfy $\beta_j \geq 0$, $j = 0, 1, \cdots, 5$ and $\mathbf{1}^T\beta = \sum_{i=0}^{5} \beta_i = 1$.

According to the science of architecture, the total passing rate $\hat{Y}_j = \sum_{i=0}^{5} \hat{\beta}_i x_{ji}$ should fall into the given range and the closer to the middle point $Y_j, j = 1, 2, \cdots, 9$, the better. Therefore the problem becomes

$$\begin{array}{ll} minimize & Q(\beta) = (Y - X\beta)^T(Y - X\beta) \\ subject\ to & \mathbf{1}^T\beta = 1 \\ & \beta \geq 0, \end{array}$$

which is equivalent to the least squares method for a linear regression problem. Firstly we can find the optimal solutions $\hat{\beta}^*$ without the constraints, we obtained

$$\hat{\beta}^* = \begin{pmatrix} 0.07465, & 0.29883, & 0.19637, & 0.19221, & 0.0000, & 0.26120 \end{pmatrix}^T$$

Using (ii) of the theorem the optimal solutions are found as

$$\hat{\beta}^T = \begin{pmatrix} 0.07611, & 0.2941, & 0.2191, & 0.1562, & 0.0000, & 0.2545 \end{pmatrix}$$

5 Appendix

We prove (i) of the theorem 2.1,

$$\begin{aligned} Q(\beta) &= (Y - X\beta)^T(Y - X\beta) \\ &= Y^TY - 2Y^TX\beta + \beta^TX^TX\beta \\ &= [-\frac{1}{2}(L^T)^{-1}b]^T[-\frac{1}{2}(L^T)^{-1}b] - 2(-\frac{1}{2}(L^T)^{-1}b)^TX\beta + \beta^TX^TX\beta \\ &= \frac{1}{4}b^TL^{-1}(L^T)^{-1}b + b^TL^{-1}X\beta + \beta^TX^TX\beta \\ &= \frac{1}{4}b^TD^{-1}b + b^T\beta + \beta^TD\beta \\ &= \frac{1}{4}b^TD^{-1}b + Z(\beta) \end{aligned} \tag{10}$$

Therefore the relationship of optimal values between QP (1) and LS (3) is obtained by

$$Z(\hat{\beta}) = Q(\hat{\beta}) - \frac{1}{4}b^TD^{-1}b$$

where $X = L$, and $Y = -\frac{1}{2}(L^T)^{-1}b$.

Next we consider proof (ii) in theorem 2.1, the least squares problem with the equality constraints $A\beta = C$ can be expressed by normal equations using the Lagrangian method as:

$$\begin{aligned} A\beta &= C \\ A^T\lambda + X^TX\beta &= X^TY. \end{aligned} \tag{11}$$

where λ are Lagrangian multipliers. normal equations (11) can be rewritten as a matrix form:

$$\begin{pmatrix} 0 & A \\ A^T & X^TX \end{pmatrix}\begin{pmatrix} \lambda \\ \beta \end{pmatrix} = \begin{pmatrix} C \\ X^TY \end{pmatrix}$$

Now we construct an augmented matrix D based on the statistical computing methods in multivariate linear regression:

$$D = \begin{pmatrix} 0 & A & C \\ A^T & X^TX & X^TY \\ C^T & Y^TX & Y^TY \end{pmatrix}$$

Using Gauss-Jordan elimination for matrix D, we obtain

$$\begin{pmatrix} B^{-1} & B^{-1}\begin{pmatrix} C \\ X^TY \end{pmatrix} \\ -(C^T, Y^TX)B^{-1} & RSS \end{pmatrix} = \begin{pmatrix} B^{-1} & \begin{pmatrix} \lambda \\ \beta \end{pmatrix} \\ -(\lambda, \beta) & RSS \end{pmatrix} \tag{12}$$

where

$$B = \begin{pmatrix} 0 & A \\ A^T & X^TX \end{pmatrix}.$$

$$B^{-1} = \begin{pmatrix} [-A(X^TX)^{-1}A]^{-1} & [A(X^TX)^{-1}A^T]^{-1}A(X^TX)^{-1} \\ (X^TX)^{-1}A^T[A(X^TX)^{-1}A^T]^{-1} & T \end{pmatrix}$$

where $T = (X^TX)^{-1} - (X^TX)^{-1}A^T[A(X^TX)^{-1}A^T]^{-1}A(X^TX)^{-1}$, therefore we can obtain solutions of λ and β from equations (12),

$$\begin{aligned} \lambda &= -[A(X^TX)^{-1}A^T]^{-1}C + [A(X^TX)^{-1}A^T]^{-1}A(X^TX)^{-1}X^TY \\ &= [A(X^TX)^{-1}A^T]^{-1}[A(X^TX)^{-1}X^TY - C] \\ &= [A(X^TX)^{-1}A^T]^{-1}[A\beta^* - C] \end{aligned}$$

and note that $\beta^* = (X^TX)^{-1}X^TY$

$$\beta = (X^TX)^{-1}A^T[A(X^TX)^{-1}A^T]^{-1}C + TX^TY.$$

since

$$\begin{aligned} TX^TY &= (X^TX)^{-1}X^TY - (X^TX)^{-1}A^T[A(X^TX)^{-1}A^T]^{-1}A(X^TX)^{-1}X^TY \\ &= \beta^* - (X^TX)^{-1}A^T[A(X^TX)^{-1}A^T]^{-1}A\beta^*, \end{aligned}$$

we have

$$\beta = (X^TX)^{-1}A^T[A(X^TX)^{-1}A^T]^{-1}C + \beta^* - (X^TX)^{-1}A^T[A(X^TX)^{-1}A^T]^{-1}A\beta^*$$

Therefore (ii) is obtained for given the estimators $\hat{\beta}$ and $\hat{\beta}^*$.

References

[1] B.M. Brown and C.R. Goodall, *Application of Quadratic programming in Statistics. Probability, Statistics and Optimization,* Edited by Kelly F.P., ch. John Wiley&Son's Ltd, (1995), pp341-350.

[2] K.T. Fang, D.Q. Wang and G.F. Wu, *A Class of Constrained Regression: Programming Regression*. Computing Mathematics, Vol. 1, pp57-69.

[3] J.A. Ferrez, K. Fukuda and T.M. Liebling, *Solving the Fixed Rank Convex Quadratic Maximization in Binary Variables by a Parallel Zonotope*. European Journal of Operational Research, 166 (1), (2005), pp35-50.

[4] F.S. Hillier, G.J, Lieberman, *Introduction to Operation Research.* McGraw-Hill, Inc. (1995).

[5] D. Q. Wang, S. Chukova and C. D. Lai, *Reducing Quadratic Programming Problem to Regression Problem: Stepwise Algorithm*, European Journal of Operational Research, 164 (1), (2005), pp79-88.

[6] D. Q. Wang, S. Chukova and C. D. Lai, *On the Relationship Between Regression Analysis and Mathematical Programming*, Journal of Applied Mathematics and Decision Sciences, Vol. 8(2) (2004), pp131-141.

[7] P. Wolfe, *The simplex Method for Quadratic Programming.* Econometrica, Vol. 27 (1959), pp382-398.

Airline Revenue Management Over Networks With Unassigned Aircraft

Amir Joshan
Department of Engineering Science
University of Auckland
New Zealand
a.joshan@auckland.ac.nz

Abstract

Revenue management techniques attempt to maximize revenue by controlling the price and availability of diverse products that are produced from scarce resources. An application of revenue management occurs in the airline industry in which the products are the airline tickets and the resources are the seats on flights. Most revenue management models consider the network of flights when the aircraft have been assigned to the legs in the network prior to the start of the booking process. In this article, we present a new formulation of two common revenue management models in the literature, which assigns the aircraft at the same time as the booking limits are determined. We show that not surprisingly, having the option to assign the aircraft during the booking process improves the revenue.

1 Introduction

In every financial enterprise operation, the owner of the firm is trying to make as much money as he can, playing with decisions that he can make: how to sell his product, how much to sell each time, when to sell, when to drop the price, how to price and a number of other questions. Therefore, maximizing profit is an old and common challenge in all businesses.

Revenue management (RM) is a strategy used to maximize the value gained from outputs while dealing with the market and customers.

A short answer to the origins of revenue management is “The Airline Industry”. After the airline deregulation act of 1978, the U.S civil aviation board relaxed control of airline prices. Larger airlines tried to improve their reservation and global distribution systems. At the same time, new low cost airlines were profitably able to price much lower than larger ones for short itineraries. This resulted in significant migration of price sensitive travelers to these new low cost airlines. Consequently, the bigger airlines were faced with the challenge of finding a strategy to recapture that group of passengers while simultaneously not losing their non-price-sensitive customers. They solved this problem using a combination of purchase restrictions and capacity control fares. Taking into account, the discount rates and super-saver prices, the big American airlines noticed that they really needed more intelligent methods and systems to have the optimal capacity-control policy. This led to the development of the Dynamic Inventory Allocation and Maintenance Optimizer System (Dynamo). Dynamo represents the first large-scale RM system in the industry. [Talluri & Van Ryzin 2004]

Littlewood first addressed the airline booking process when he proposed the "Littlewood Rule" [Littlewood 1972]. The rule proposed for only a two-class model, says that low-fare customers should be accepted until their revenue value exceeds the expected revenue of a future high-fare customer. This idea was extended to multiple classes afterward. "Later on, it was shown that under certain conditions it is optimal to accept a request only if its fare level is more than or equal to the difference between the expected total revenue from the current time to the end when respectively rejecting and accepting the request" [Chen & Homem-de-Mello. 2004].

This led to development of methods for estimating the current value of the seats on different legs in the network. These kind of methods resulted in policies called bid-pricing policies. The idea behind these kind of methods is to calculate price of seats in different legs during the time, and then determine a current price for each itinerary simply by adding the current values of the seats that are being used in that itinerary. Then the optimal decision is to accept the arriving customer only if its fare is greater than the determined value (sum of legs values).

However, the original thought of revenue management was in the way of accepting different types of customers up to some determined optimal levels called booking limits. These methods resulted in policies called booking policy.

Most of the early models were built just for single flights. The result of applying these methods into networks was that the policy is only locally optimized and it cannot guarantee global optimization. Later on some models have been developed in which they consider the whole network flights situation in order to determine booking limits for different classes in each individual flight.

Glover et al. 1982 were first to describe the revenue management problem over a network with the assumption of deterministic demand. Using network flow theory, they focused on network aspect of the model. There were a few tries after that on revenue management network problem like Dror, Trudeau and Ladany 1998. The new idea in the latter model was predicting cancellation in the model.

However, perhaps Williamson 1992 first addressed the model that determines booking limits based on linear programming. In this model, she considered the stochastic demand through expected values.

Since in many cases we can have some assumptions about the distribution of demand rather than just expected value, one may want to consider it in the model as well. One way to do such a thing is by resolving the above model. In this case, in the resolving times, we are updating the expectation vector of the left demand by reducing the previous realized demand from the initial estimation of the expectation of demand. Cooper 2002, showed a counter example that resolving this deterministic model in this way may lead to lower total expected revenue.

Another way to consider the demand distribution in the formulation is modeling the problem as a stochastic program. In the simplest case we have only two stages, the beginning of the decision horizon when we decide the booking limits (first stage), and the end of the decision horizon when we realize the revenue gained by those booking limits depending upon the demand realization.

Higle and Sen 2004, and Cooper and Homem-de-Mello 2003, proposed two different stochastic models. The first work has shown that the bid price policy produced from stochastic model, tends to act better than the one produced by deterministic model. The second one produced a hybrid model on which the second stage was the optimal value function of a dynamic program.

Chen and Homem-de-Mello 2004, have shown that modeling the problem with a multi-stage stochastic program may lead to better expected revenue while at the same time it may not be easily solved owing to the lack of concavity. Instead, they proposed a rolling horizon approach in which they resolved the two stage stochastic program several times during the time horizon. They have shown that with this approach the objective function always improves in terms of expected revenue.

All of the above models assume that in the network of flights the aircrafts have been assigned to the legs prior to the beginning of the time horizon. Therefore, the capacities are fixed on each leg. In this paper, it is assumed that the airline has the option to reallocate the aircrafts during the decision process.

The rest of the paper is organized as follows: in section 2 the different mathematical formulations of the problem (deterministic and probabilistic) and their characteristics are presented, in section 3 the new model with the both deterministic and probabilistic states is described. Concluding remarks are presented in section 4.

2 Common models in the literature

In this section, the standard models in the literature of how to allocate seats are described. We have a network of flights, p distinct customer classes and n possible itineraries. Therefore, we have np itinerary-fare classes. The booking process is realized over a time horizon of length T (it is quite common in the literature to assume that the demand arrival process has Poisson distribution) and there is at most one unit of demand per period. The demand for itinerary-class (j,k) over the whole horizon is denoted by d_{jk} and the vector d is comprised of elements d_{jk} . The network contains m legs with the capacities $c \equiv c_i$ and the $(m \times np)$ matrix $A \equiv a_{i,jk}$. The entry $a_{i,jk}$ is equal to one, when class k customers use leg i in itinerary j .

The decision variable x_{jk} is the number of seats to be allocated to class k in itinerary j . Whenever an itinerary class pair (j,k) is accepted, the revenue corresponding to the fare F_{jk} accrues.

These basics are used in the following models in the next parts of this section.

2.1 Deterministic model

Indices:

i : Legs in the network $i = 1,...,m$

j : Possible itineraries $j = 1,...,n$

k : Possible classes $k = 1,..., p$

Parameters:

f_{jk} : Revenue of a ticket itinerary j class k

$a_{i,jk}$: The elements of matrix A . $a_{i,jk}$ is one if class k customers use leg i in itinerary j and zero otherwise.

$A \equiv a_{i,jk}$: Matrix contains $a_{i,jk}$.

$c \equiv c_i$: is the vector of capacities. Each c_i is the capacity of the leg i in the network which represents the capacity of the assigned aircraft to that leg.

$d \equiv d_{jk}$: is the vector of demand for different itinerary classes. Each d_{jk} represents the demand for class k of itinerary j .

Decision variables:

$x \equiv x_{jk}$: is the vector of allocated seats to different classes. Each individual x_{jk} is the number of seats to be allocated to class k in itinerary j.

Model RM:

1- $Max \quad f^T x$

2- $Ax \leq c$

3- $x \leq d \quad w.p.1$

4- $x \geq 0$

This model cannot be solved due to the uncertain demand vector d. Therefore, a simple approach to overcome this constraint is to formulate the problem as below:

Model DLP:

5- $Max \quad f^T x$

6- $Ax \leq c$

7- $x \leq \quad E\ [d] \quad w.p.1$

8- $x \geq 0$

This model called DLP (deterministic linear problem) in revenue management literature. Williamson [7] first used this model to obtain booking limits. In addition, she used the dual variables of the capacity constraint (i.e. 6) to determine the current price of an itinerary. Before describing the probabilistic model, some of the characteristics of this model are discussed. Further information can be found on [1] and [7].

We require solution x_{jk} to be integer. This is easily attainable by rounding any solution of DLP if necessary. In addition, the objective function is not the actual revenue resulting from the policy x^* even if x^* is integer. Because the actual revenue is our expectation of the sold seats that is $E[f^T \min\{x, d\}]$. Moreover, this model ignores any information about the distribution of the underlying random variables.

2.2 Probabilistic model

A common approach to include information about the distribution of the random variable is to write the DLP as a stochastic programming problem:

9- $Max \quad f^T E\ [Min\ \{x, d\ \}]$

10- $Ax \leq c$

11- $x \in \quad Z^+$

Alternatively, we can write the above model as:

Model: SLP

12- $Max \quad f^T x + E\ [Q(x, d)]$

13- $Ax \le c$

14- $x \in \ Z^+$

15- $Q(x,d) = Max\{-f^T y | x - y \le d, y \ge 0\}$

Notice that SLP has the standard form of a two stage stochastic program with simple recourse and if d has a discrete distribution we can easily solve it (i.e. with scenario generation).

Another interesting point is that the objective function of SLP corresponds to the actual expected revenue resulting from any obtained solution. In addition, each rounded feasible solution of DLP is a feasible solution of the SLP. Therefore, the straightforward conclusion is that the optimal policy from SLP is never worse than that of DLP when rounded, in terms of expected revenue.

However, both models assume that the allocation of the aircrafts is fixed. In the next part, it is shown that not surprisingly, if we have the option to assign the aircraft in the same time that we obtaining the booking limits, we may improve the objective function.

3 New model

In both previous models, each leg has a predetermined capacity that represents the capacity of the aircraft that the airline has assigned to that particular leg. In practice, airlines usually decide about the aircraft assignment before determining the booking limits for their customers. The assumption made here is that the airline is still able to change its aircraft assignment during the booking process. This situation can be modeled both with deterministic and stochastic demand. In the next part of this chapter these models are given.

3.1 Deterministic model

Here only the new indices and variables are presented.

Index:

t: The possible aircrafts over the network. This can be greater or equal to the number of legs. Here it is assumed that $t = 1,...,m$

Parameter:

Let A^i denote i row of matrix A.

Variable:

l_{ti}: Binary variable: It is equal to one if we assign aircraft t to the leg i, and zero otherwise.

Then DLP can be extended as follows:

Model: extended DLP

16- $Max \quad f^T x$

17- $\forall\ i : A^i x \le \sum_t l_{ti}\ c_t \qquad i = 1,...,m$

18- $x \le \quad E\ [d] \quad w.p.1$

19- $x \ \ge 0$

20- $\sum_{t=1} l_{ti} = 1 \qquad i = 1,...,m$

21- $\forall t : \sum_{i=1} l_{ti} = 1 \qquad t = 1,...,m$

$x \in \ Z^{+}, l \in \{0,1\}$

The model structure is quite the same as in DLP. The only difference is introduction of the binary variables l_{ti}. We are trying to assign the best aircraft to each leg using the information that currently we have about the demand. constraints number (20) and (21) make sure that each aircraft has been assigned to a legs and that each leg has been assigned an aircraft (note that we assume number of aircrafts is equal to the number of legs. However, one can change it to the desired number of aircrafts by changing constraint (20) and (21)).

The straightforward conclusion from the model above is that it will not worsen the objective value comparing with DLP model.

Proposition 1: The policy obtained from the extended DLP yields expected revenue, which is at least that given by the policy obtained from DLP.

Proof: Since every policy obtained from DLP is related with a particular assigned order of capacities to legs, and it is a feasible solution for the extended DLP given that particular order, the extended DLP has an optimal solution, which is better or equal to the optimal solution of DLP in term of expected revenue.

Below is the extension of SLP:

Model: Extended SLP

22- $Max \quad f^{T}x + E \ [Q(x,d)]$

23- $\forall \ i : A^{\ i} x^{T} \leq \sum_{t} l_{ti} \ c_{t}$

24- $Q(x,D) = Max\{-f^{T}y \vert x - y \leq d, y \geq 0\}$

25- $x \in \ Z^{+}$

26- $\forall i : \sum_{t} l_{ti} = 1$

27- $\forall t : \sum_{i} l_{ti} = 1$

$x \in \ Z^{+}, l \in \{0,1\}$

Proposition 2: In the case of having finite number of scenarios for demand, the policy obtained from the extended SLP yields expected revenue, which is at least that given by the policy obtained from SLP.

Proof: Since we can write the second stage recourse function as a linear combination of possible outcome times their probabilities, we will gain a deterministic linear objective function. As we saw in Proposition 1, taking into account the option of changing aircrafts will never worsen the objective function in case of expected revenue.

It is worthwhile saying that in the case of having a distribution with an infinite number of scenarios (i.e. Poisson distribution) we can use the truncated form of that distribution and the result still holds.

In the remainder of this paper, we only discuss the DLP and extended DLP model. The reason for this, as we mentioned above, is the possibility of converting SLP and extended SLP, into DLP and extended DLP models.

3.2 Important issues about the new model

3.2.1 Practical point of view

While it was proven that having the option to change the assigned aircraft during the booking process could lead to a better objective value, one may ask about the possibility of doing it in practice. Is it possible to change the aircraft of a flight with another a while before the departure time?

The answer depends upon several issues. First airlines vary in size, number of aircraft, and the network that they are working over. For a big airline, which is worldwide, it is quite impossible to change the aircraft that have been assigned to an itinerary in one continent, to the one that is due to carry on another itinerary in a different continent. However, for a small airline operating domestically with an integrated network, this would work.

Second, for the bigger airlines, one may be still able to refine the allocation during the booking process by categorizing the aircraft into sets of fixed and changeable aircrafts. It is shown in Corollary 1 below that the objective function could improve if we have the option of changing the order of a subset of aircrafts in both deterministic and probabilistic states. Big airlines usually operate over networks with multiple hubs. Although it may not be possible to swap aircraft operating from different hubs, it may be possible to change the aircraft order during the booking process for those aircraft that departure from same hub.

Corollary 1: Having the option of swapping a subset of capacities over the network of flights will lead to a policy that yields expected revenue, which is at least that of the expected revenue obtained by the DLP model.

Proof: Here again the policy obtained from DLP is a feasible policy for the extended DLP problem with having the option of swapping only a subset of capacities.

3.2.2 Relation between aircraft capacities

The extended models (both probabilistic and deterministic) try to determine the booking limits as well as assigning an appropriate capacity to each leg in the network. If all the aircraft have same size, there is no point to trying to assign them to different legs. However, if the aircraft have different capacities we may be able to save some revenue considering the network of flights and different demands that exist for different itineraries.

3.2.3 Relation between demands and capacities

Again, consider the network of flights. Suppose $E[d_i] \le c_i \ \forall \ i$ in which d and c are the vectors of demand and capacity respectively. It is obvious that again there is no point to use the extended version of DLP since constraint number (2) in DLP model is slack. In fact, if the DLP model holds some of the rows in constraint (2) tight and some of them slack, we may expect to improve our objective function.

4- Conclusion

Two common models in airline revenue management for determining the booking limits over a network of flights are presented. While in the literature all the models assumed the order of the seats are fixed and they have been assigned to the legs, we may obtain a better solution by having the option of creating new order during the time horizon.

This may be helpful especially if we are dealing with price sensitive customers. Since the duals of the capacity constraint is being use as the present price of each leg, by taking into account the best order of assignment, the airline will be able to offer more competitive prices for tickets.

In practice, this may not be applicable all the time for all networks, but for some of them that are reasonably small or by defining a subset of swappable aircraft it could be applied.

The second thing was that the improvement in objective function by applying this model depends on the ration of demand and capacities, and the variety of the capacities that we have in our fleet.

6- References

Birge J.R. and F. Louveaux. 1997. "Introduction to stochastic programming", *Springer-Verlag*, Berlin

Chen L. and T. Homem-de-Mello. 2004. "Multi-Stage Stochastic programming models for air line revenue management", *working paper*, 9 Oct 2005 http://users.iems.northwestern.edu/~tito/list.htm

Cooper W.L. 2002. "Asymptotic behaviour of an allocation policy for revenue management", *Operations research*. 50:720-727

Cooper W.L. and T. Homem-de-Mello. 2003. "A class of hybrid methods for revenue management", *working paper*, 9 Oct 2005 http://users.iems.northwestern.edu/~tito/list.htm

De Boer S. and R. Freling, 2002. "Mathematical Programming for network Revenue Management", *European Journal of Operation Research.* 137: 72-92

Dror M., P.Trudean, and S.P. Ladany, 1998 "Network models for seat allocation on flights", *Transportation science Res*. 22B: 239-250

Glover F., R. Glover, J. Lorenzo and Mcmillan.1982. "The passenger mix problem in the scheduled airlines". *Interface* 12: 73-79

Higle J.L. and S. Sen. 2004. "A stochastic programming model for network resource utilization in the presence of multi-class demand uncertainity" *to be appear in application of stochastic programming, SIAM series on optimization.*

Littlewood K. 1972. "Forecasting and control of passenger bookings", *AGIFORS* Symposium *Proc.* 12, Nathanya, Israel.

Talluri K. and G. Van Ryzin. 2001. "An analysis of bid- controls for network revenue management", *Management Science*, Vol 44, No. 11

Talluri K. and G. Van Ryzin. 2004. "The theory and practice of revenue management", Kluwer *academic Publishers*, Massachusetts

Williamson E.l.1992. "Airline network seat control", *Ph.D thesis*, MIT, Cambridge, Massachusetts

Optimal Traffic Light Control With Pedestrians

Joshua S.J. Koh
Department of Engineering Science
University of Auckland
New Zealand
jkoh006@ec.auckland.ac.nz

Abstract

In this project, we used a simulation software package called Arena to build a realistic simulation model of the traffic control system at the Waterloo Quadrant – Princes Street intersection, which lies within the Central Business District of Auckland City. Our traffic model is an extension to previous work done by Jimin Hong, who modelled the same intersection. Our model attempts to encapsulate both vehicle as well pedestrian activities, in an effort to accurately represent the conceptual traffic model. Validation of our model was performed by comparing its phasing performance to that of the actual system. Through the extensive and repetitive verification/validation process, we managed to detect an inconsistency with the conceptual model and the actual system. Upon consultation with Transit New Zealand, this was attributed to a software fault with the on-site traffic controller, which apparently seemed to be unnecessarily allocating right-turn phases to Princes Street traffic.

We used an optimisation software built into Arena called Optquest, to calibrate any unknown variables in our model, as well as to obtain the optimal values of the controller settings, to minimise vehicle waiting time. Our calibrated model was a good representation of the actual system, with very similar phasing statistics. The optimised model resulted in vehicle waiting times being reduced by as much as 10% on average, during peak periods. This was reduced further if we corrected the software fault.

1 Introduction

Auckland is New Zealand's largest city, with a population growing at the nation's fastest rate of 3.2%. As with any rapidly growing city, traffic congestion has been Auckland's major concern and the demand will become increasingly hard to meet using the existing transportation system.

While the public focus is primarily on improved public transport as a solution, the fact that the level of car ownership is growing at about twice the rate of the population in Auckland means that improved public transport alone will probably not solve its traffic problems.

As an alternative and more practical solution, an advanced traffic management system has been developed in the Auckland region to allow the current roading infrastructure to operate at maximum capacity. This includes ITS (Intelligent Transport Systems) projects and the introduction of SCATS (Sydney Computerized Adaptive Traffic Signal System), an advanced real-time traffic management system.

Traffic signal control is the fundamental element in the traffic management system. It aims to reduce the number of stops and vehicle delay, thus maximizing the throughput. Traffic signal control method varies in complexity, from simple pre-timed systems that utilize historical data to schedule fixed timing plans, to adaptive signal control, which optimizes phasing for a network of signalized intersections, depending on real-time traffic conditions.

In New Zealand, vehicle-actuated signal control with adapted features from Australian standards is used. Vehicle-actuated control employs the use of various controller settings which govern the duration and sequencing of phases. Efficient signal control operation will depend strongly on the value of these settings. Typically, several factors will affect the optimal choice for these settings; these include intersection geometry, traffic composition, pedestrian density, number and types of lanes per phase as well as the location, number and nature of the detectors used. Therefore, it is very difficult to establish a general rule for choosing appropriate values of the controller settings, although some guidelines and rules are available.

This project is aimed at building a realistic simulation model of the Waterloo Quadrant/Princes Street intersection, which lies within the Central Business District of Auckland City. We followed up on previous work done by Jimin Hong, who attempted to model the same intersection, but neglected pedestrian movement as a factor in building her model. For this project, we attempt not only to incorporate pedestrian movements into Hong's original model, but also to further refine and extend it. Finally, we evaluate the efficiency and effectiveness of the current traffic control settings at this intersection, and see if we can improve it.

1.1 Current Signal Control Methods

Currently, there are two main forms of traffic signal control, pre-timed signal control and vehicle-actuated control.

Pre-timed signal control consists of right-of-way assignments based on a predetermined schedule. The right-of-way assignments, or phases, including green period and sequencing for each phase, are fixed and based on historic traffic data.

Vehicle-actuated control consists of right-of-way assignments based on changing traffic demand. This is achieved through the use of vehicle detectors and an actuated controller unit. Green times are continuously adjusted in real-time to meet right-of-way demands registered by vehicles approaching the junction. If required, the phase sequence may be altered as well. In general, vehicle-actuated control tends to reduce waiting delay and increase vehicular throughput over pre-timed control and is thus, more commonly used world-wide.

1.2 Phase Intervals & Controller Settings

When vehicle-actuated control is used, a green phase consists of three parts: *Minimum green interval*, *Rest Interval* and *Extension Green Interval*.

Efficient operation of the signal actuations depends on the values of three controller settings, the *Minimum Green Setting*, *Gap Time Setting* and *Maximum Green Setting.*

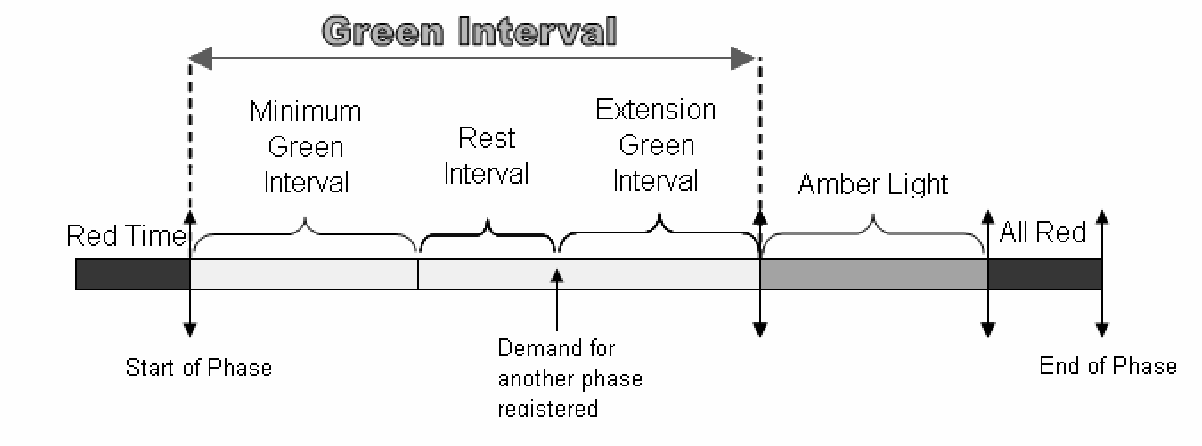


Figure 1: Definition of the Phase Intervals.

The *Minimum Green Interval* has a fixed duration given by the *Minimum Green Setting.* This setting is provided to ensure that on the initiation of a green period, suitable allowance is made for stationary vehicles to start moving.

The *Rest Interval* is an unlimited and un-timed interval which initiates when the *Minimum green interval* expires. During this period of time, the signal controller enters a rest state until demand for another phase is registered. If demand for another phase is registered during the *Minimum green interval,* then the *Rest Interval* is skipped. If no competing demand is detected, then the controller holds the current phase green indefinitely.

The *Extension Green Interval* is a period of variable length. It is determined by the combined effect from the controller extension settings, the *Gap Time Setting* and *Maximum Green Setting.* During this period, the green interval may be terminated.

The *Gap Time Setting* specifies the maximum allowable time between successive detector actuations which will allow the green interval to be extended. If the actual time gap between successive detector actuations exceeds the specified *Gap Time Setting* during the *Extension Green Interval*, then the green interval will terminate. This form of termination is called a *Gap Change.* This is illustrated in Figure 2.

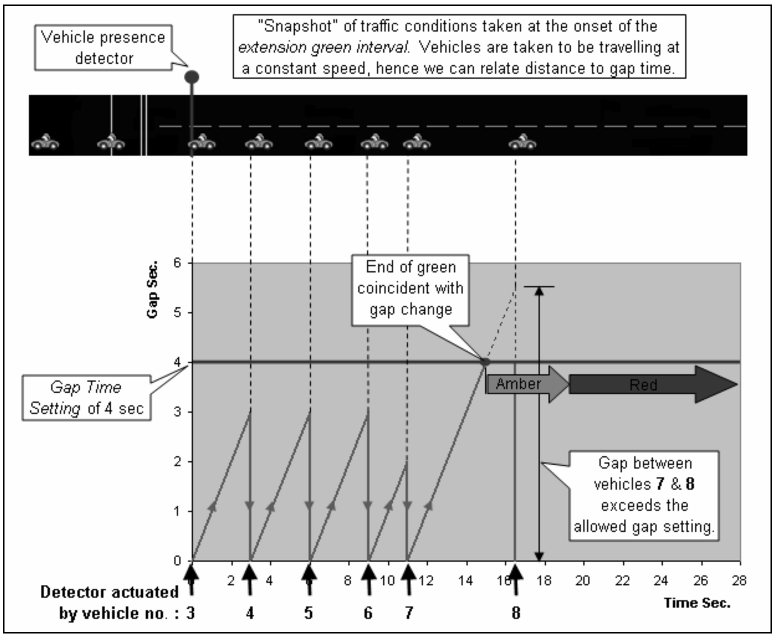


Figure 2: This demonstrates the logic of gap timing and gap change.

In addition to a *Gap Change*, the green interval can be terminated by a *Minimum Change* or a *Maximum Change.* If the actual time gap between successive detector actuations exceeds the specified *Gap Time Setting* and a competing phase demand is registered both during the *Minimum green interval*, the green interval will terminate with a *Minimum Change* once the *Minimum green interval* expires.

A *Maximum Change* occurs when a *Gap Change* has not occurred during the *Extension Green Interval* and the length of the *Extension Green Interval* equals the *Maximum Green Setting.*

1.3 Headway-Waste Control

Headway-Waste is an advanced traffic volume/density control feature currently used in New Zealand and Australia. Headway-Waste control employs two additional extension settings for the controller, *Headway Time Setting* and *Waste Time Setting.* This aims to improve efficiency by allowing the controller to respond to density changes in traffic.

When the time between successive vehicle actuations is in excess of the specified *Headway Time Setting*, the excess gap time between the vehicles is accumulated. This accumulated time is called the *accumulated waste*. When the *Accumulated Waste* reaches the specified *Waste Time Setting*, *Waste Change* occurs. The *Accumulated Waste* is only accumulated during the *Extension Green Interval* when a competing phase demand is registered. This is in addition to *Gap Change.* Figure 3 illustrates this.

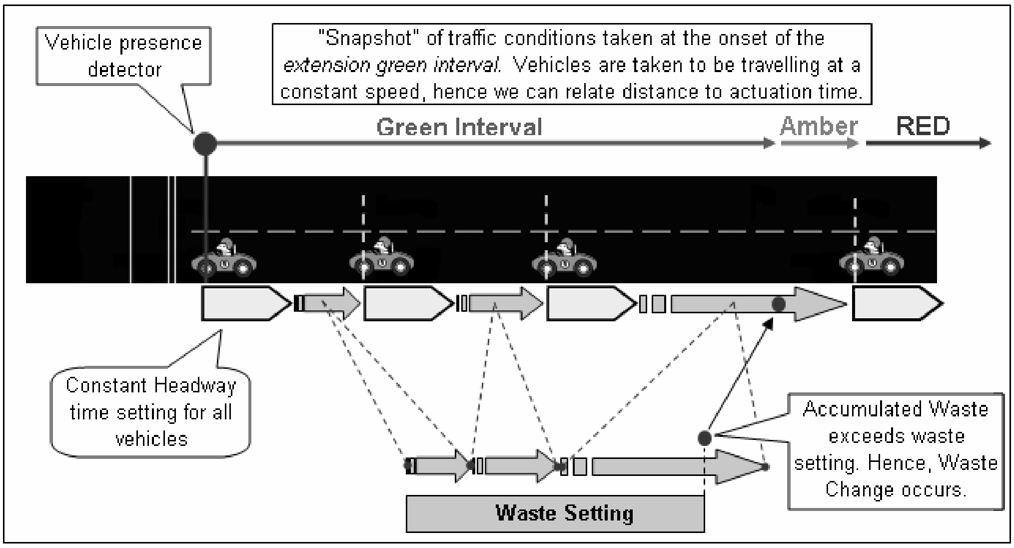


Figure 3: This demonstrates the logic of waste timing and waste change.

1.4 Pedestrian Phases

A pedestrian phase is a phase in which specified pedestrians crossings are given the right-of-way movement for a specified length of time. There are two basic controller settings which determine the length of these pedestrian right-of-way periods, namely, *Walk Time* and *Clearance Time.* These periods are indicated by a pedestrian display showing either a green man or a flashing red man.

Walk Time specifies the length of time which the green man indicator is activated. This is meant to indicate to pedestrians that it is still safe to begin crossing.

Clearance Time specifies the length of time which the flashing red man indicator is displayed. This is meant to indicate to pedestrians that it is not safe to begin crossing as their right-of-way is about to expire.

There are three basic types of pedestrian phases. *Single-Phase Crossings* are when pedestrians begin and finish crossing all in the same traffic phase. This is the most common type of pedestrian crossing. *Inter-Phase Crossings* are where the pedestrian phase operates in two traffic phases. The pedestrians begin crossing in the first traffic phase and finish in the next traffic phase. This form is much less common than Single-Phase Crossings. *Pedestrian Only Crossings* are pedestrian phases which do not run together with a traffic phase. All lights will be red for traffic during this phase. This is used mostly at intersections where high pedestrian density and flow occur.

2 Focus of Project

Delay at signalized intersections is a major source of frustration for drivers. To improve the efficiency of traffic signal control from the driver's perspective, the average waiting time of the vehicles is chosen as our objective.

For the sake of consistency, we define the waiting time to be the amount of time elapsed from when a vehicle arrives and enters the detection zone which begins 13.5m before each stop line, up to when it enters the intersection and arrives at its destination turn-off.

2.1 Objectives

The aim of this project is to revise and improve on Hong's Arena traffic model, and also to account for pedestrian movement at the intersection. We will also use Arena's built-in optimisation package, Optquest, to calibrate unknowns for the system. This will be followed by validating our simulation model against the real system in an attempt to justify the significance of neglecting pedestrians in the modeling process, previously performed by Hong. Finally, we will analyse the effect of altering the controller settings of the validated model to improve the average waiting time of vehicles in the system.

3 Modelling Approach

3.1 System description

For the purpose of this project, we chose to model the intersection where Princes Street, Waterloo Quadrant, Bowen Avenue and Kitchener Street meet.

The reason for this selection is due to the relative isolation of the intersection from surrounding intersections, and hence, interferences from these neighbouring intersections can be considered very small. This is because the signal control at this intersection is independent of those from other intersections unlike most other intersections in the Central Business District, where the signal control is coordinated over a network of multiple intersections and the phasing is interrelated. Therefore, we can justify the assumption that we are analyzing a separate and independent system, which satisfies the condition for Poisson arrivals as mentioned in section 3.2.

Another reason is that since this intersection is located in the Central Business District, it experiences clear and distinct patterns of AM peak, PM peak and daytime off-peak traffic flow. Hence, allowing us to reduce the target time period for our study to these three specifically.

Furthermore, all the pedestrian phases at this intersection are *Single-Phase,* which means that the pedestrians begin and finish crossing in the same traffic phase. This makes the system much easier to model as opposed to *Multi-Phase* crossings.

3.2 Phasing and Detector Details

There are ten approaching lanes and eight outgoing lanes to this intersection. Traffic is regulated by traffic light signals operating in five phases. This is shown in figure 4.

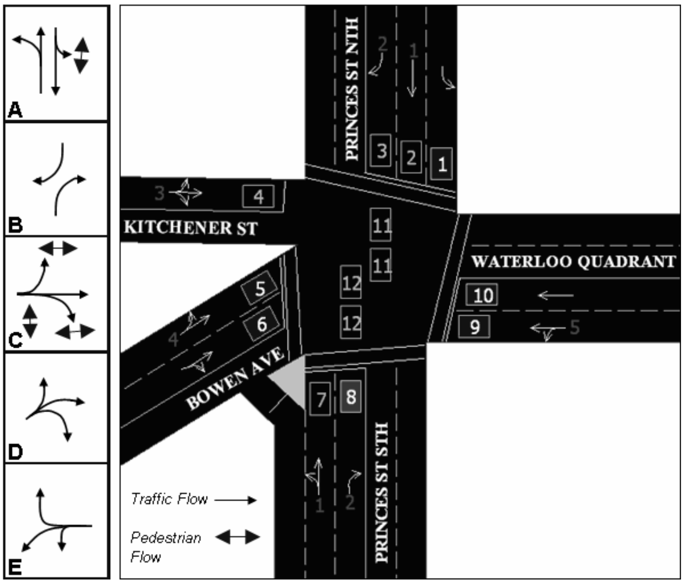


Figure 4: Physical layout of Princes/Waterloo Junction and its 5 signal phases.

Figure 4 shows a screenshot from SCATS used by Transit New Zealand, of the intersection, with roads, detectors (depicted by numbered squares) and phasing schemes. Ten detectors, one for each approaching lane, are positioned just before the stop lines to each approaching lane (numbered 1 to 10 in an anti-clockwise direction). Two additional detectors, numbered 11 and 12, are located in the middle of the intersection to detect turning traffic.

Pedestrian crossings are located across all the roads meeting the junction except for Kitchener Street. Pedestrian call-boxes are located at either end of each pedestrian crossing which pedestrians push to lodge their crossing demands with the controller.

Different groups of detectors, called signal groups, are used for each phase as if they were a single detector, either to extend the green period or to register demands for right-of-way during a red interval. For example, the signal group for Phase A would be detectors 1,2,3,7 and 8 as shown in table 1.

Phase A and B allows traffic approaching along Princes Street to enter the intersection. In Phase A, the through traffic from lanes 2 and 7, is given right-of-way indicated by a green light, while right-turning traffic from lanes 3 and 8, is allowed to filter through gaps in the head-on traffic. Left-turning traffic from lane 1 has to give way to right-turning vehicles via the standard give-way rule. Both traffic from lane 1 and 8, have to give way to pedestrians crossing waterloo quadrant as long as the pedestrian green or flashing red man is still active. Phase B is an exclusive right-turn phase in which right turning traffic from lanes 3 and 8 are given the green arrow right-of-way. Detectors 11 and 12 detect stationary vehicles waiting in the middle of the intersection to turn and register a demand for Phase B.

Phases C, D and E allow traffic access to the intersection from Kitchener Street on lane 4, Bowen Street from lanes 5 and 6 and Waterloo Quadrant from lanes 9 and 10 respectively. In these phases, the associated through and turning traffic have right-of-way with no contesting movement, with the exception of Phase C where vehicles have to give way to pedestrians crossing across Princes Street.

Detectors 1 to 10 are called Stop-line detectors. They have a detection area of 13.5m called the Detection Zone. The traffic-volume, gap time and *Accumulated Waste* are determined from the output of these detectors and queue conditions can be obtained.

The controller responds to and treats calls in different ways depending on the preset detector settings. Phase calls for competing phases can be made in either a locking or non-locking mode of memory. In the locking mode, a call is retained by the controller, even after the vehicle has left the detection zone, until the call demand has been satisfied by a corresponding green phase. In the non-locking mode, a call is retained only while vehicles are in the detection zone, and is dropped by the controller as soon as the vehicle leaves the detection zone. This effectively screens out false calls and reduces delay somewhat. Detectors 1 to 10 are locking detectors, corresponding to Phases A, C, D and E, while detectors 11 and 12 are non-locking, corresponding to Phase B.

3.3 Traffic Data

We were provided with traffic arrival data to the intersection for a 5 day working period, from Monday 6th October 2003 to Friday 10th October 2003, courtesy of Transit New Zealand. This volume data was counted on a lane-by-lane basis for each Stop-line detector, 1 to 10.

It contains the accumulated number of vehicles arriving at each approach lane in 5 minute intervals for a 24 hour period. Effectively, it provides discretised snapshots of traffic volume flow for the intersection over the five days.

The three distinctive time periods of the day, AM peak 7am-9pm, PM peak 4pm-6pm and daytime off-peak 10am-4pm. The start and finish times of these periods have been specifically chosen to match up with the actual phasing statistics from Transit NZ.

3.4 Limitations

Necessary traffic characteristics were derived from Australian traffic surveys since very limited local information was available. Vehicles are also modelled as particles with no physical length. Due to the limited information available on pedestrian flow at the junction, pedestrians are represented in the simulation as probability functions rather than actual individual entities.

3.5 Interaction between system components

A signalized intersection with vehicle-actuated control is a complex system characterized by interaction between vehicle traffic, pedestrians, detectors and the signal controller. Vehicle arrivals and departure determine the output of the detectors, and in turn, vehicle movement is governed by signals from the controller. Detectors provide information on the prevailing traffic conditions, to which the controller responds by extending or shortening the phase duration. Pedestrian demand is registered through call-boxes located throughout the intersection. These demands in turn, determine the nature of the traffic flow for certain manoeuvres as well as the duration of the pedestrian phases, A and C. The following logic flow diagram, figure 5, gives an overview of the interaction between the system components in a typical phase.

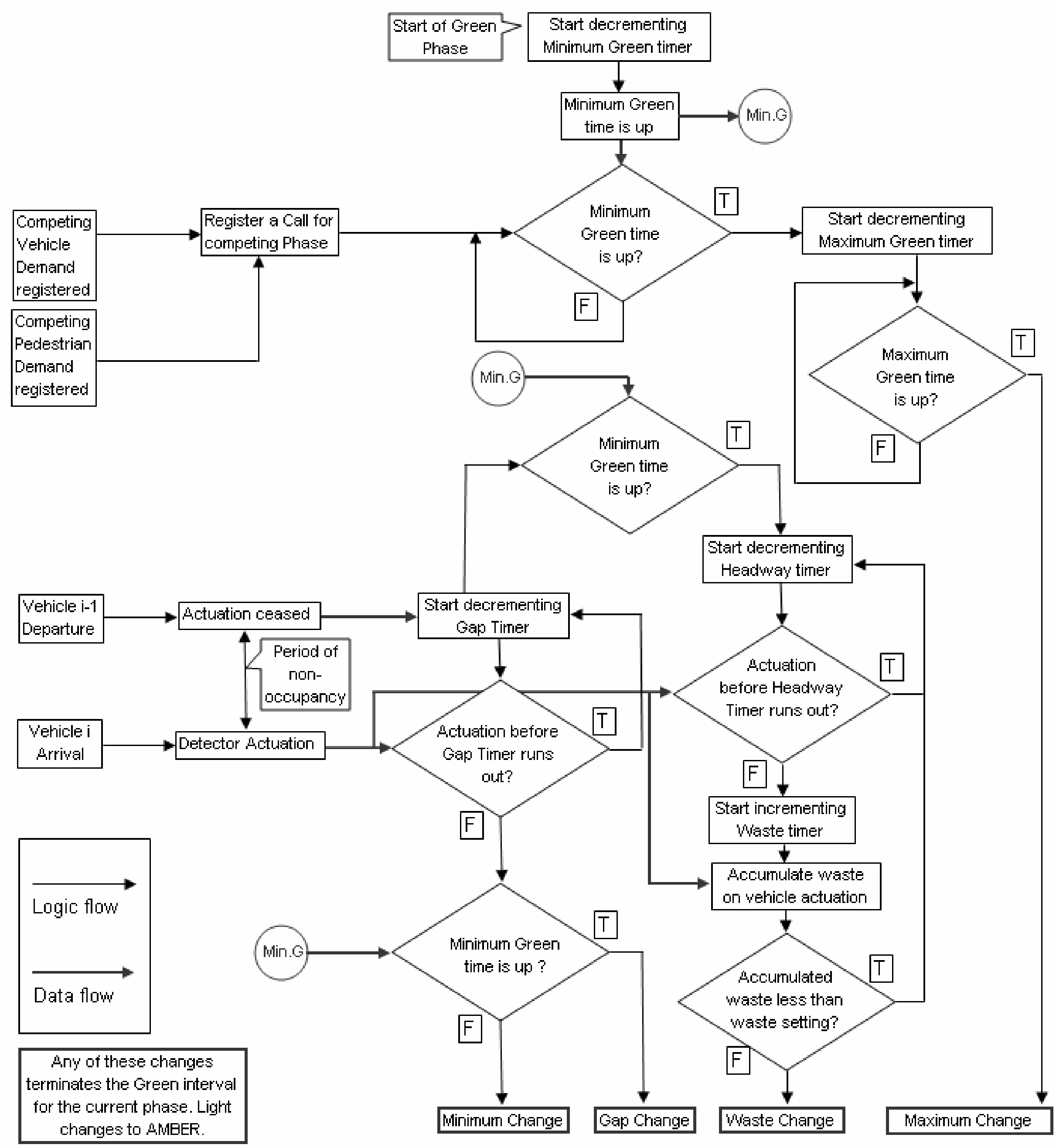


Figure 5: Logic flow diagram illustrating the interaction between system components.

4 Validation and Verification

For our project, there were two ways that we attempted to establish that our model realistically represented the real system.

Firstly, we obtained phasing statistics from Transit New Zealand at the intersection of interest to compare our simulation results against. Secondly, we performed several hours of field observations at the intersection itself. While the former was used primarily in our validation process, the field observations themselves proved invaluable as they helped us to identify inconsistencies with our input data.

4.1 Using Optquest

For the purposes of this project, we defined the time percentage and frequency of each phase as requirements in our Optquest model. These were set according to the

SCATS data which we received from Transit New Zealand. The unknowns in the model, pedestrian crossing times, individual crossing probabilities in Phase C and turning headway times were all set as controls to be calibrated by Optquest. The complete verification, calibration and validation process is illustrated in figure 6.

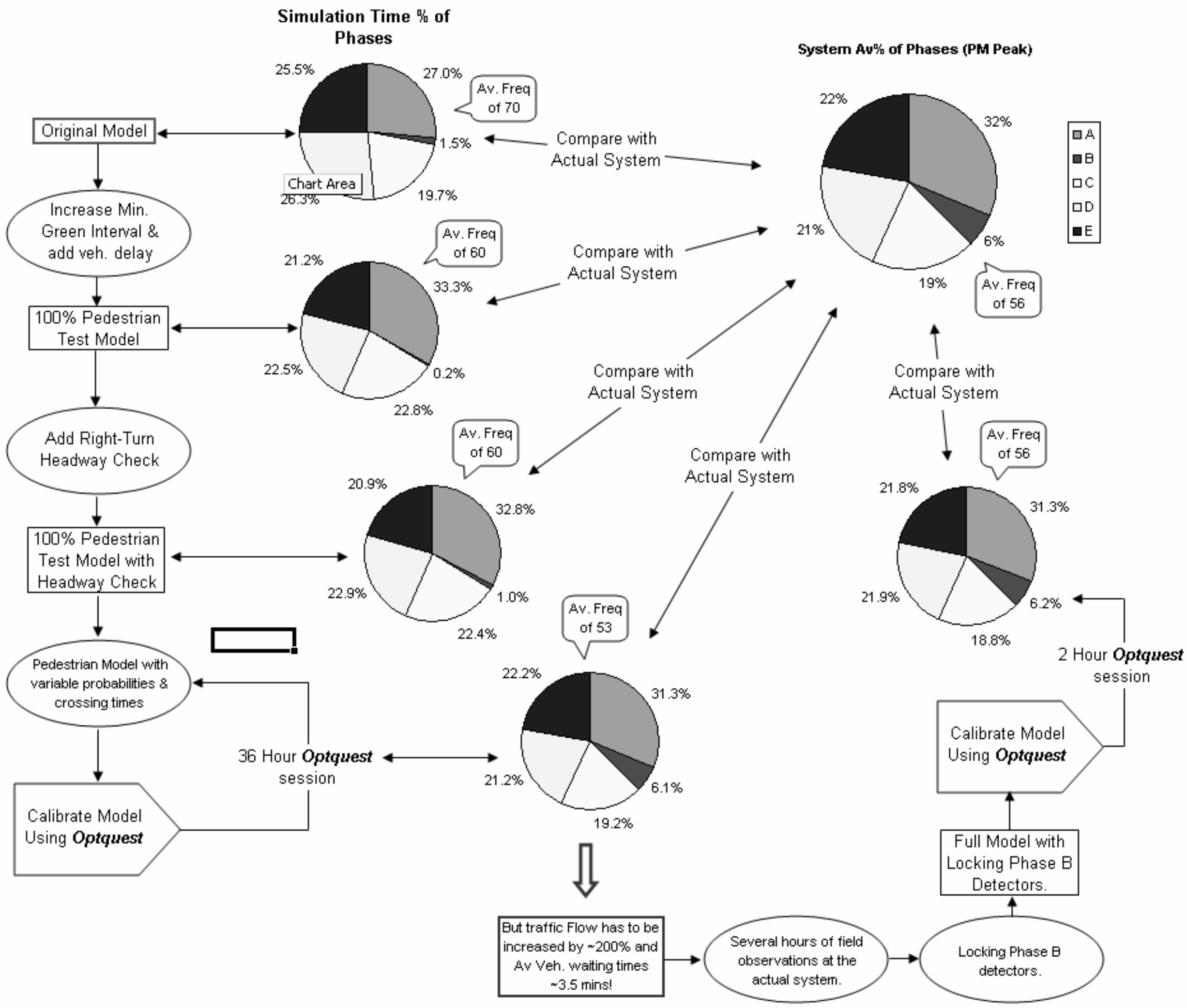


Figure 6: The Incremental Verification/Validation Process.

Optimisation of the objectives, the average vehicle waiting time, the maximum vehicle waiting time and the maximum queue length were eventually performed by specifying them as “Minimise Objectives” in Optquest. For these cases, the controller settings, pedestrian settings and phasing sequencing were set as control variables.

5 Results

5.1 Validation Results

The phasing statistics from 50 simulation replications were compared against the corresponding SCATS data for both the PM Peak period as well as the AM Peak period.

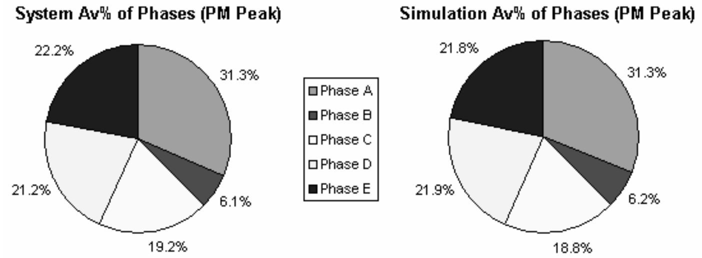


Figure 7: Comparison of % output from the simulation and the real system (PM Peak).

5.2 Simulation Results

Statistic	PM Peak	AM Peak
Minimum Value	2.89	2.98
Maximum Value	141.22	136.36
Mean Value	52.52	47.56
Median	49.22	44.89
Upper Quartile	69.46	58
Lower Quartile	23.5	20.12
Variance	709.69	508.5
Standard Deviation	26.68	22.55
Midspead (IQR)	43.66	37.85
Skewness	0.23	0.51
Number of data values	3012	2423

Table 1: Key statistics on the vehicle waiting times (all values in seconds).

5.3 Optimisation Results

Controller Setting	Phase A	Phase B	Phase C	Phase D	Phase E
Gap Time	1.53	0.11	1.13	1.61	1.58
Waste Setting	4.22	0.94	3.43	4.42	4.27
Headway Setting	0.75	0.2	0.51	0.68	0.71

Average Vehicle Waiting time	46.12

Table 2: Optquest-Optimised controller settings by Phase.

The results from Optquest, shown in table 2 indicate that the optimal controller settings correspond to extremely "tight" extension criteria for Phase B. In conjunction with field observations, this suggests that there is a high degree of inefficiency in having locking Phase B detectors since most instances of Phase B terminate with a *Minimum Change* due to zero detector occupancy. After discussions with Michael Daley at Transit New Zealand, the problem was determined to not to be incorrect configuration data, but rather a software problem with the controller. The controller was retaining Phase B calls even after occupying vehicles had left the detectors. This was not intended in the conceptual design and was thereafter, scheduled to be corrected. If we changed Phase B detectors back to non-locking in our model, with the initial controller settings, the average vehicle waiting time falls to **44.85 seconds**, which is slightly better than using our Optquest-Optimized controller settings with locking Phase B detectors (**46.12 sec**).

6 References

Hong, J. 2004. *Simulation of a Traffic Light Control System.* 4th Year Project Report, University of Auckland, 2003.

Stochastic Model: Warranty Analysis

Tineke Poot
Victoria University of Wellington
New Zealand
tinekepoot@gmail.com

Abstract

An overview of different warranty policies and their cost analysis using renewal theory will be discussed. I will focus on non-renewing and renewing warranties under different lifetime distributions such as Exponential, Weibull, Gamma, Lognormal and Inverse Gaussian. Related cost analysis will be presented.

1 Renewal theory

Renewal theory is important because of its many real life applications. A renewal process can be used to model arrivals (customers to a store), queuing systems and the lifetime of devices, people and animals. Due to its ability to model the lifetime of objects it plays an important role in warranty analysis.

1.1 Counting processes

A stochastic process Z(t) is a collection of random variables which depend on t. If t represents time, Z(t) is the state of the process at time t. A counting process N(t) is a stochastic process which represents the number of events that have occurred up to time t. N(t) is a counting process if it satisfies the following conditions:

a) $N(t) \geq 0$;
b) N(t) is integer valued;
c) If $s < t$ then $N(s) \leq N(t)$;
d) For $s < t$, $N(t) - N(s)$ is the number of events in the interval (s, t].

1.2 Renewal processes

A renewal process is a counting process such that the times between events are independent and identically distributed random variables with an arbitrary distribution. Let the inter-arrival times $\{X_i, i=1,\ldots\infty\}$ have distribution function F(x). Poisson process is a renewal process with $F(x) = 1 - e^{-\lambda x}$, i.e., the inter-arrival time is exponentially distributed. Some other distributions of interest are the Uniform, Erlang, Gamma and Weibull. The occurrence of an event is called a renewal.

Let S_n be the time of the n-th renewal, and it is the sum of n independent and identically distributed random variables X_i's, i.e.,

$$S_n = \sum_{i=1}^{n} X_i, \quad n \geq 1$$

The expected number of renewals in the interval [0, t) is called the renewal function and it is denoted by M(t). The renewal function M(t) satisfies the so called renewal equation, which is

$$M(t) = F(t) + \int_0^t M(t - x)f(x)\,dx.$$

The renewal equation has an analytical solution for a limited number of distributions F(t). Usually in order to solve the renewal equation numerical procedures have to be involved. M(t) is a very important tool in problems related to warranty analysis.

2 Warranty policies

2.1 Non-renewing free replacement warranty

Any item which fails in the period from the time of the first purchase up to time W, the length of the warranty period, is repaired or replaced at no cost to the consumer. The replaced/repaired items are only covered for the remaining time in the warranty period.

Assuming that the warranty claim is rectified by replacement, the expected total cost to the seller is the sum of the cost of supplying the original item and the expected cost of supplying replacements while under warranty. The expected number of replacements is the renewal function M(W), i.e.,

$$E[C_s(W)] = c_s[1 + M(W)].$$

2.2 Renewing free replacement warranty

Any item which fails in the period up to time W is repaired or replaced at no cost to the consumer. The replaced/repaired items also have a warranty period of W, which starts from the time of repair or replacement.

The warranty coverage lasts until a replaced/repaired item or the original item has a lifetime of at least W. The number of replacements required for this has a geometric distribution with mean F(W)/[1 – F(W)]. Assuming that the warranty claim is rectified by replacement, the expected cost will be the sum of the cost of supplying the original item and the expected cost of supplying replacements (c_s multiplied by the mean F(W)/[1 – F(W)]), i.e.,

$$E[Cs(W)] = c_s[1 - F(W)]^{-1}.$$

2.3 Rebate warranty

This warranty is often called a “money-back guarantee”. A consumer receives a complete refund (the full purchase price) for any item which fails in the warranty period, e.g., up until time W from time of purchase.

The expected total cost to the seller is now a function of the buyer’s cost instead of the cost of supplying the item,

$$E[C_s(W)] = c_s + c_bF(W).$$

2.4 Partial rebate warranty

A consumer receives a partial refund, a portion, say α, of the purchase price, for any item which fails in the warranty period, e.g., up until time W from time of purchase. Thus the expected warranty cost is

$$E[C_s(W)] = c_s + \alpha c_b F(W).$$

3 Cost analysis

A number of examples for different warranty policies will be discussed. The length of the warranty period W will be set to 0.5, 1.0 and 1.5. The lifetime distributions used are Exponential, Weibull, Gamma, Lognormal and Inverse Gaussian. For the different examples the mean time to failure μ, which relates to the parameters of the lifetime distributions, is set to 0.5, 1.5, 2.0 and 3.0.

Tables for the value of $E[C_s(W)]/c_s$ for the different lifetime distributions and different values of μ and W under non-renewing an renewing warranties were constructed. More study has been done to address a rebate or partial rebate warranty.

Expected Cost to Seller Under Non-renewing Free Replacement Warranty									
	Life Distribution								
	Exponential	Weibull		Gamma		Lognormal		Inverse Gaussian	
μ		$\beta = 0.5$	$\beta = 2$	$\beta = 0.2$	$\beta = 3.66$	$\theta = 1.339$	$\theta = 0.4915$	$\theta = 0.2\mu$	$\theta = 3.66\mu$
				W = 0.5					
0.5	2.000	3.048	1.624	2.091	1.631	2.542	1.635	1.388	1.298
1.5	1.333	2.019	1.085	1.773	1.056	1.571	1.023	1.264	1.010
2	1.250	1.858	1.048	1.708	1.025	1.430	1.005	1.224	1.002
3	1.167	1.677	1.022	1.627	1.007	1.281	1.000	1.165	1.000
				W = 1.0					
0.5	3.000	4.314	2.638	2.346	2.637	3.808	2.636	1.440	1.474
1.5	1.667	2.570	1.313	1.960	1.301	2.081	1.283	1.348	1.143
2	1.500	2.308	1.184	1.877	1.159	1.835	1.122	1.316	1.062
3	1.333	2.019	1.085	1.773	1.056	1.571	1.023	1.264	1.010
				W = 1.5					
0.5	4.000	5.471	3.637	2.502	3.637	5.021	3.636	1.462	1.497
1.5	2.000	3.048	1.624	2.091	1.631	2.542	1.635	1.388	1.298
2	1.750	2.694	1.385	1.996	1.380	2.199	1.372	1.361	1.186
3	1.500	2.308	1.184	1.877	1.159	1.835	1.122	1.316	1.062

Table 1 Non-renewing Free Replacement Warranty: $E[C_s(W)]/c_s$

In view of the results in Table 1 the following general comments can be made:

1. Under a non-renewing free replacement warranty for lifetimes with Weibull and Gamma distributions with a decreasing failure rate, DFR, ($\beta = 0.5$ and $\beta = 0.2$ respectively) and the corresponding Lognormal distribution there is always higher expected costs than Exponential distribution.
2. Inverse Gaussian is the exception with lower expected costs. This is probably due to a bug in my code.

3. The distributions with increasing failure rates, IFR, always lead to lower expected costs than Exponential distribution.

Expected Cost to Seller Under Renewing Free Replacement Warranty									
	Life Distribution								
	Exponential	Weibull		Gamma		Lognormal		Inverse Gaussian	
μ		β = 0.5	β = 2	β = 0.2	β = 3.66	θ = 1.339	θ = 0.4915	θ = 0.2μ	θ = 3.66μ
					W = 0.5				
0.5	2.718	4.113	2.193	2.261	2.323	3.976	2.482	1.634	1.425
1.5	1.396	2.263	1.091	1.855	1.059	1.785	1.024	1.358	1.011
2	1.284	2.028	1.050	1.775	1.025	1.555	1.005	1.288	1.002
3	1.181	1.781	1.022	1.678	1.007	1.336	1.000	1.198	1.000
					W = 1.0				
0.5	7.389	7.389	23.141	2.597	20.761	8.602	20.471	1.786	1.902
1.5	1.948	3.173	1.418	2.091	1.413	2.801	1.391	1.534	1.167
2	1.649	2.718	1.217	1.985	1.187	2.274	1.139	1.461	1.066
3	1.396	2.263	1.091	1.855	1.059	1.785	1.024	1.358	1.011
					W = 1.5				
0.5	20.086	11.582	1174.5	2.790	311.2	15.492	152.7	1.857	1.988
1.5	2.718	4.113	2.193	2.261	2.323	3.976	2.482	1.634	1.425
2	2.117	3.403	1.555	2.138	1.574	3.080	1.580	1.564	1.229
3	1.649	2.718	1.217	1.985	1.187	2.274	1.139	1.461	1.066

Table 2 Renewing Free Replacement Warranty: $E[C_s(W)]/c_s$

In view of the results in Table 2 the following general comments can be made:

1. Under a renewing free replacement warranty Weibull, Gamma and Lognormal with DFR again have higher expected costs than Exponential distribution. With the exception that this changes for the longer warranty length of W = 1.5 where under Weibull, Gamma and Lognormal for lower values of μ the cost is actually lower.
2. Once again the distributions with IFR led to lower expected costs than Exponential distribution with the exception of μ = 0.5 when W is 1.0 or 1.5 for Weibull, Gamma and Lognormal.

Overall these results were expected. A failure rate function gives the probability that a t-year old item will fail given that it has survived till t. It will have IFR if the failure rate function is an increasing function of t and DFR if it's a decreasing function of t. Therefore for an item with DFR to have the same mean as exponential with its constant failure rate over time, it would have to have a comparably higher failure rate for lower values of t. As t increases this rate decreases and at some point it will become smaller than exponential's failure rate. This is why the pattern of higher expected costs changes for larger values of W. The opposite will be true for IFR.

4 Conclusions

Warranty analysis is a recently developed area of research. As the competition for space in the marketplace increases the warranty will become a selling attribute of the products.

Many companies and industries incur significant losses in covering the warranty claim on their production, which is why warranty analysis is an attractive area of research, both theoretical and applied.

5 References

Beichelt, F.E. and Fatti, L.P. 2002. “Chapter 4: Renewal Processes.” *Stochastic Processes and Their Applications*, Taylor & Francis, London.

Blischke, W. and Murphy, P. 1993. “Chapter 3: Stochastic Processes for Warranty Modeling.” *Warranty Cost Analysis*, Marcel Dekker, Inc.

Blischke, W. and Murphy, P. 1996. *Product Warranty Handbook*, Marcel Dekker, Inc.

Ross, S. 2000. *Introduction to Probability Models*, 7ed., Harcourt Academic Press.

Xie, M. 1989. “On the solution of the renewal-type integral equations,” Commun. Stat.-Simula. **18**: 281-293.

Portfolio Optimization via Stochastic Dominance of Benchmarks

David N. Craigie
Department of Engineering Science
University of Auckland
New Zealand
dcra036@ec.auckland.ac.nz

Abstract

We outline the principles of stochastic dominance with particular emphasis on second order stochastic dominance and its application in portfolio optimization.

We implement Ruszczynksi and Dentcheva's linear programme which seeks to maximize expected return of a portfolio of assets, subject to the portfolio second order stochastically dominating a benchmark. This means the portfolio is preferred over the benchmark by all risk averse investors.

We use this linear programme and the more established Markowitz mean-variance optimization to obtain portfolios using real world historical data. We discuss the various merits and pitfalls of the two portfolio optimization techniques and consider some modelling issues such as the nature of return distributions and benchmark selection.

We conclude that portfolio optimization via stochastic dominance of benchmarks is an appealing theoretical concept as it makes fewer assumptions regarding return distribution and investor preferences than the Markowitz model.

1 Introduction

This paper focuses on the use of second order stochastic dominance constraints in a linear programme for the purpose of portfolio optimization. In particular we consider the differences and the relative strengths and weaknesses of this method compared with the more established Markowitz method of mean-variance optimization. We compare solutions obtained using the two methods in a small numerical example and consider under what circumstances the efficient sets corresponding to each method may coincide.

The primary goal of this paper is not to draw conclusions about investment strategies or the performance of specific assets, but rather to present the concept of stochastic dominance and examine issues with its application in portfolio optimization.

The stochastic dominance rules are objective decision rules that determine whether one asset (or portfolio) dominates another by first, second or third order stochastic dominance. If an asset first order stochastically dominates another then it can be shown that all investors with a non-decreasing utility function will prefer this asset. An investor with such a utility function is non-satiated. An investor with a concave utility function is said to be risk averse. An asset dominates another by second order stochastic dominance if all risk averse investors prefer it. There are higher orders of stochastic dominance, but it is unclear what these infer about investor preferences.

The linear programme outlined in this paper is that of Ruszczynski and Dentcheva (2003). It seeks to maximise the return of a portfolio constructed from a given number of assets, subject to this portfolio second order stochastically dominating some benchmark. In other words, what proportion of our total investment should we invest in each asset, such that we can get the best return possible from a portfolio that will be preferred by all risk averse investors over the benchmark? The benchmark could be some other portfolio consisting of the same assets, say with an equal proportion invested in each, or it could be some suitable market index, or any other meaningful series of returns.

Most portfolio optimization techniques seek to maximise return and minimise risk and as such are bi-criteria models. The stochastic dominance LP doesn't optimise with respect to a risk measure but rather the second order stochastic dominance constraints ensure that the solution is preferred by all risk averse investors over the benchmark. For this reason, the choice of the benchmark is very important as it effectively induces a risk measure. In using this model we are essentially asking "what portfolio should I hold in order to maximise my return while ensuring that my position is less risky than if I had invested in the benchmark?" One of the properties of stochastic dominance is that the dominant portfolio will have a higher expected return than the benchmark. Of course we may not be able to find a portfolio that dominates the benchmark, in which case the solution we obtain is the benchmark portfolio itself.

2 First and Second Order Stochastic Dominance

2.1 First Order Stochastic Dominance

The set of non-decreasing utility functions describes all investors who are non-satiated, or prefer more money. If one asset is preferred over another by all investors with such a utility function, then that asset is said to exhibit first order stochastic dominance over the other.

Mathematically, stochastic dominance occurs among random variables. A random variable has an associated probability density function, the integral of which is its cumulative density function or CDF. A random variable with CDF F is said to first order stochastically dominate (FSD) a random variable with CDF G if:

$$F(\eta) \leq G(\eta) \quad \forall \eta \in [a,b] \quad \text{(First Order Stochastic Dominance)}$$

Graphically, this is depicted below in Figure 1, over an interval $[a,b]$. The return we get from an asset can be thought of as a random variable, thus the distribution of the returns of one asset can dominate another's by the above criteria and hence that asset is said to dominate the other. First order stochastic dominance of F over G is often written as $F \succ_1 G$.

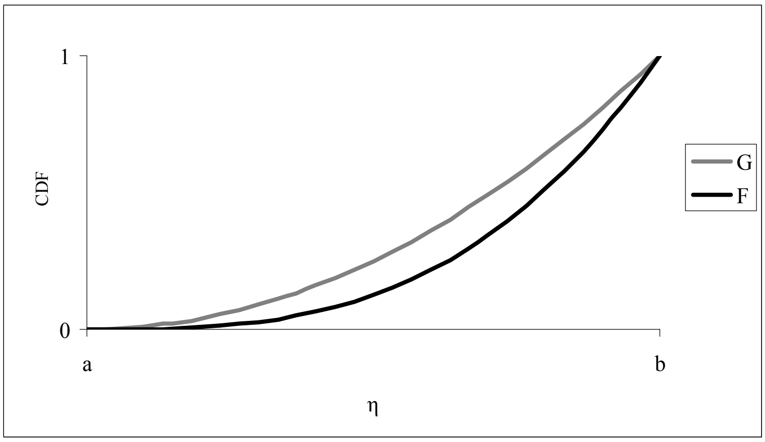


Figure 1: First Order Stochastic Dominance

For an intuitive explanation as to why an asset with CDF F is preferred by all non satiated investors over an asset with CDF G, recall that the value of the asset's CDF at the point η, gives the cumulative probability of obtaining a return less than or equal to η.

If the CDF F is below that of G for all η, then this implies that no matter what return η we consider, the probability that we get a return less than η is smaller for the asset F than for the asset G. Alternatively, the probability that we will obtain a return greater than η is greater for the asset F. Note that if F does not dominate G, this does not automatically imply the dominance of G over F.

2.2 Second Order Stochastic Dominance

In reality first order stochastic dominance of one asset over another is quite rare, as the condition is so strict. But when it does occur, under the assumption of non-satiety we should always choose the dominant asset. However, we can add risk aversion to our assumption of non-satiety and observe that if one asset is preferred by all risk averse investors, over another, it is said to second order stochastically dominate (SSD) the other.

Mathematically, an asset with CDF *F* dominates in the second order an asset with CDF *G* if:

$$\int_a^{\eta}[G(x)-F(x)]dx \geq 0 \quad \forall \eta \in [a,b] \quad \text{(Second Order Stochastic Dominance)}$$

First order stochastic dominance implies second order stochastic dominance but not vice versa. For example the CDFs of assets *F* and *G* cross over in Figure 2, so there is no first order stochastic dominance, but the integral of *F* lies under the integral of *G* in Figure 3, so $F \succ_2 G$

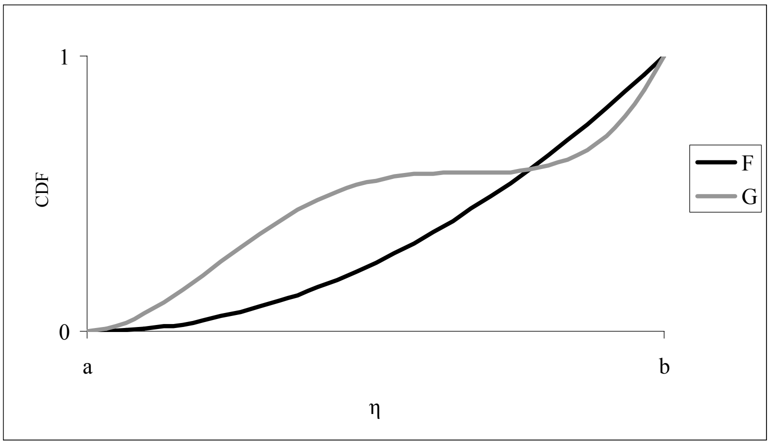


Figure 2: Overlapping CDFs

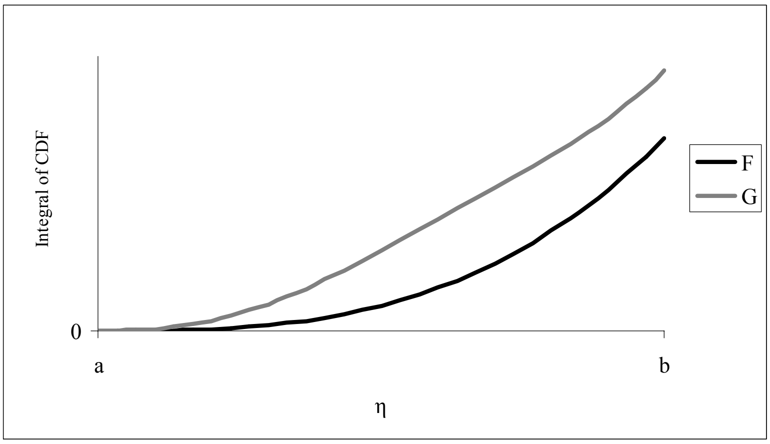


Figure 3: Second Order Stochastic Dominance

F dominates *G* by SSD, since for any negative area ($G < F$) there will be a positive area ($F < G$) which will be greater or equal to the negative area and which will be located before the negative area. This reflects the risk aversion assumption.

3 A Linear Programme with Stochastic Dominance Constraints

3.1 Derivation

Ruszczynski and Dentcheva propose a linear programme with stochastic dominance constraints of the form:

$$\begin{aligned} &\max \text{ Expected Return} \\ &\text{s.t.} \quad \text{SSD Benchmark} \end{aligned}$$

The second order stochastic dominance constraints can be derived as follows:

Recall that a random variable *X* with CDF *F*, stochastically dominates in the second order a random variable *Y* with CDF *G* if:

$$\int_a^{\eta} [G(x) - F(x)]dx \ \geq \ 0 \ \ \forall \eta \in [a,b] \tag{1}$$

When $a = -\infty$ and $b = \infty$ we can write:

$$\int_{-\infty}^{\eta} F(x)\,dx \leq \int_{-\infty}^{\eta} G(x)\,dx \quad \forall \eta \in \Re \tag{2}$$

We can express the integral as an expectation:

$$\int_{-\infty}^{\eta} F(x)\,dx \ = E[(\eta - X)_+] \quad \text{where} \quad (\eta - X)_+ = \max(0, \eta - X) \tag{3}$$

So (1) can be written as:

$$E[(\eta - X)_+] \leq E[(\eta - Y)_+] \quad \forall \eta \in \Re \tag{4}$$

The stochastic dominance constraints must enforce (4).

Let $R_1,\dots,R_N$ be random returns of assets $1,\dots,N$ and $z_1,\dots,z_N$ be the proportions of our total investment to invest in each asset. Total return $X = R_1 z_1 + \dots + R_N z_N$. Let *Y* be some reference random return such as that from an existing portfolio or market index. This gives rise to the following optimization problem:

P:

$$\begin{aligned} \max \quad & f(X) = E[X] \\ \text{s/t} \quad & E[(\eta - X)_+] \leq E[(\eta - Y)_+] \quad \forall \eta \in \Re \\ & X \in C \end{aligned}$$

where *C* is the convex hull of $R_1,\dots,R_N$

This problem has infinitely many constraints. However, if the benchmark *Y* and the portfolio *X* each have a discrete distribution with finitely many realisations *m*, then the formulation simplifies greatly, as shown in the next section.

3.2 Second Order Stochastic Dominance Linear Programme

Parameters:

- m number of realisations or elementary events
- p_k probability of each realisation $(k = 1,...,m)$
- N number of assets
- r_{nk} return of asset n under realisation k $(n = 1,...,N, \quad k = 1,...,m)$
- y_i return of the benchmark under realisation i $(i = 1,...,m)$
- $v_i = E[(y_i - Y)_+]$ shortfall of the benchmark under realisation i $(i = 1,...,m)$

Variables:

- z_n the proportion to invest in asset n (decision variable) $(n = 1,...,N)$
- $x_k = \sum_{i=1}^{N} r_{nk} z_n$ return of the portfolio under realisation k $(k = 1,...,m)$
- $s_{ik} = (y_i - x_k)_+$ shortfalls of the portfolio $(i = 1,...,m, \quad k = 1,...,m)$

The linear programme that seeks to maximize the portfolio's expected return, subject to the portfolio second order stochastically dominating the benchmark Y is given by:

SSD LP:

$$\max \sum_{k=1}^{m} \sum_{n=1}^{N} p_k r_{nk} z_n \tag{5}$$

$$\text{s.t.} \quad \sum_{n=1}^{N} r_{nk} z_n + s_{ik} \geq y_i \qquad i = 1,...,m \quad k = 1,...,m \tag{6}$$

$$\sum_{k=1}^{m} p_k s_{ik} \leq v_i \qquad i = 1,...,m \tag{7}$$

$$s_{ik} \geq 0 \qquad i = 1,...,m \quad k = 1,...,m \tag{8}$$

$$\sum_{n=1}^{N} z_n = 1 \tag{9}$$

$$z_n \geq 0 \qquad n = 1,...,N \tag{10}$$

Note that the linear inequalities of constraints (6), (7) and (8) are equivalent to the second order stochastic dominance condition (4). The constraint labelled (9) simply ensures that the proportions sum to 1, while constraint (10) ensures the proportions must be positive as we assume no short selling is allowed.

The size of the problem is a function of the number of assets N and the number of realisations m. We have N decision variables, for the proportions to invest in each asset, but we also have m^2 shortfall variables that are needed to implement the dominance constraints. The problem has a total of $2m^2 + m + N + 1$ constraints so the problem size tends to a function of m^2 as m increases and the system of equations tends to being square and sparse.

3.3 Efficiency of Solution

Stochastic dominance provides a partial ordering of all the investment alternatives into efficient and inefficient sets. A portfolio in the feasible set is in the first order stochastic dominance efficient set if it cannot be dominated by FSD by any other portfolio in the feasible set. Similarly a portfolio that cannot be dominated by SSD is in the SSD efficient set.

One of the properties of second order stochastic dominance is the mean-necessary condition which states that if $X \succ_2 Y$ then $E[X] \geq E[Y]$. Another property is transitivity which states that if F second order stochastically dominates G and G second order stochastically dominates H, then F second order stochastically dominates H.

A feasible solution to the SSD LP is a portfolio that stochastically dominates the benchmark in the second order. The optimal solution of the SSD LP is the feasible solution with the greatest expected return. This optimal solution must be in the efficient set of portfolios, that is it cannot be second order stochastically dominated by any other feasible portfolio.

To see this, let X^* be the optimal solution to the SSD LP with benchmark portfolio Y. Then clearly $X^* \succ_2 Y$. Suppose there was some portfolio $P \succ_2 X^*$. Then by transitivity $P \succ_2 Y$. Therefore P is a feasible solution to the SSD LP. By the mean-necessary condition $E[P] \geq E[X^*]$. However X^* was the optimal solution to SSD LP, so $E[P] = E[X^*]$.

4 A Numerical Illustration

We implemented an AMPL model of the SSD LP and using the CPLEX solver we tested it on data from Ruszczynski and Dentcheva (2003). This historical data, shown in Table 1 below consisted of the yearly returns of $N = 8$ assets over $m = 22$ years. U.S. three-month treasury bills, U.S. long-term government bonds, S&P 500, Willshire 5000, NASDAQ, Lehmann Brothers corporate bond index, EAFE foreign stock index and gold were the assets 1, though 8 respectively. The return realisations of each year, were assigned an equal probability.

	Table Of Asset Returns (in %)							
Year	Asset 1	Asset 2	Asset 3	Asset 4	Asset 5	Asset 6	Asset 7	Asset 8
1	7.5	-5.8	-14.8	-18.5	-30.2	2.3	-14.9	67.7
2	8.4	2	-26.5	-28.4	-33.8	0.2	-23.2	72.2
3	6.1	5.6	37.1	38.5	31.8	12.3	35.4	-24
4	5.2	17.5	23.6	26.6	28	15.6	2.5	-4
5	5.5	0.2	-7.4	-2.6	9.3	3	18.1	20
6	7.7	-1.8	6.4	9.3	14.6	1.2	32.6	29.5
7	10.9	-2.2	18.4	25.6	30.7	2.3	4.8	21.2
8	12.7	-5.3	32.3	33.7	36.7	3.1	22.6	29.6
9	15.6	0.3	-5.1	-3.7	-1	7.3	-2.3	-31.2
10	11.7	46.5	21.5	18.7	21.3	31.1	-1.9	8.4
11	9.2	-1.5	22.4	23.5	21.7	8	23.7	-12.8
12	10.3	15.9	6.1	3	-9.7	15	7.4	-17.5
13	8	36.6	31.6	32.6	33.3	21.3	56.2	0.6
14	6.3	30.9	18.6	16.1	8.6	15.6	69.4	21.6
15	6.1	-7.5	5.2	2.3	-4.1	2.3	24.6	24.4
16	7.1	8.6	16.5	17.9	16.5	7.6	28.3	-13.9
17	8.7	21.2	31.6	29.2	20.4	14.2	10.5	-2.3
18	8	5.4	-3.2	-6.2	-17	8.3	-23.4	-7.8
19	5.7	19.3	30.4	34.2	59.4	16.1	12.1	-4.2
20	3.6	7.9	7.6	9	17.4	7.6	-12.2	-7.4
21	3.1	21.7	10	11.3	16.2	11	32.6	14.6
22	4.5	-11.1	1.2	-0.1	-3.2	-3.5	7.8	-1

Table 1: Historical Yearly Returns of 8 Assets

Using the same notation as above we define:

$\bar{z}=[1/N,1/N,...,1/N]$	the benchmark portfolio	(11)
$y_k=\dfrac{1}{N}\sum_{n=1}^{N}r_{nk} \quad k=1,...,m$	the return of the benchmark in year k.	(12)

Solving the SSD LP using this data, gives an optimal solution:

$$\hat{z}=[0,\ 0,\ 0.07,\ 0.19,\ 0,\ 0.39,\ 0.23,\ 0.12] \tag{13}$$

which has an expected return (the value of the objective function) equal to 11.0% compared with 10.6% for the benchmark. The maximum expected return obtainable is 14.1%, which comes from investing solely in Asset 7. The risk aversion that is enforced by the dominance constraint results in a portfolio whose expected return is well below the maximum achievable expected return. However, Asset 7 has a worst case realisation of -23.4%, whereas our solution portfolio $\hat{z}$ has a worst case realisation of -4.49%.

We also implemented the Markowitz mean-variance model using the same data as above. Again the model was implemented using the AMPL modelling language and solved using CPLEX. Rather than specifying a benchmark, when using mean-variance optimization, we specify a target expected return for the portfolio. The resulting optimal solution is the portfolio with the smallest variance that will meet the target return.

We know the optimal solution (13) had an expected return of 11.0%. Therefore if we solve a mean-variance optimization with a target return of 11.0% it is of interest to see if our solution is the same portfolio as (13). If so, then we know the solution to the SSD LP lies on the mean-variance efficient frontier (the set of portfolios that are pareto optimal in a mean-variance framework).

Solving the mean-variance problem with target return $\bar{r}_p$ = 11.0% gives an optimal solution:

$$\tilde{z}=[0,\ 0,\ 0,\ 0.32,\ 0,\ 0.30,\ 0.19,\ 0.19] \tag{14}$$

which is clearly different from (13).

Both portfolios have an expected return of 11.0%. However, $\mathrm{Var}(\hat{z})=87.82$, while $\mathrm{Var}(\tilde{z})=83.35$. So our solution to the SSD LP is not on the mean-variance (MV) efficient frontier, as there is a portfolio with the same expected return with a lower variance. Table 2 below shows the return realisations of the two portfolios in the 22 years, ordered from lowest to highest for each portfolio.

Return	SD	MV
(1)	-4.49	-5.45
(2)	-3.64	-5.26
(3)	-2.51	-0.14
(4)	0.37	0.23
(5)	1.21	1.48
:		
(18)	17.73	17.41
(19)	18.57	18.44
(20)	19.83	21.66
(21)	29.05	27.17
(22)	29.67	27.78
mean	11.00	11.00
variance	87.82	83.35

Table 2: Ordered Returns of SD and MV solution portfolios

So although $\tilde{z}$ has a lower variance, its worst case return is lower than that of $\hat{z}$. Also, in reducing the variance, it has penalised the higher returns of $\hat{z}$. Despite $\hat{z}$ not being on the efficient frontier, it appears to be preferable to a portfolio of the same expected return that is on the frontier. Note though, that $\hat{z}$ does not dominate $\tilde{z}$ by SSD.

This illustrates that a portfolio that is efficient in a stochastic dominance framework is not necessarily efficient in a mean-variance framework. But is the reverse true? In other words, is it possible to stochastically dominate in the second order, a portfolio that lies on the mean-variance efficient frontier?

The efficient frontier for the data above was calculated. Each of the mean-variance efficient portfolios was then set as a benchmark in the SSD LP. In nearly all cases it was not possible to dominate the MV efficient portfolios. The exceptions occurred at the low risk–low return end of the frontier. This part of the frontier and the portfolios that stochastically dominate it are shown in Figure 4 where the notation SD (MVX) denotes the portfolio that stochastically dominates the mean-variance efficient portfolio with expected return X.

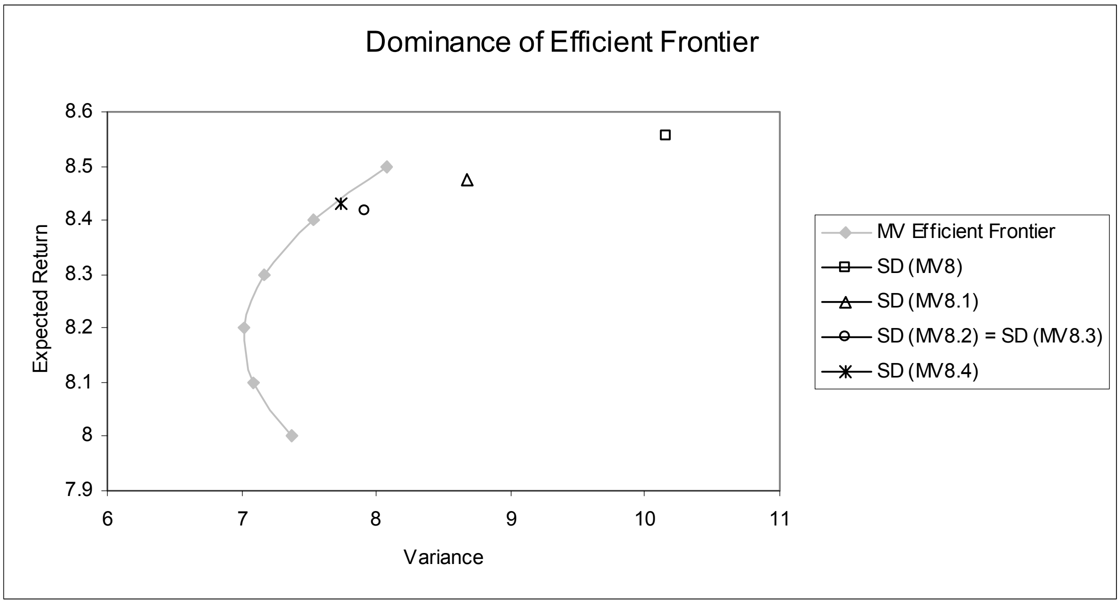


Figure 4: Stochastically Dominating MV Efficient Portfolios

In this case, although MV efficient portfolios at the low end of the frontier could be dominated, there was no stochastic dominance within the frontier. That is, no MV efficient portfolio was able to be dominated by any other MV efficient portfolio.

However, this may not always be the case as the following contrived example illustrates:

Example: Suppose we have only 2 assets and 2 realisations given by:

	Returns (%)	
Period	Asset 1	Asset 2
1	1	2
2	2	4
mean	0.25	1
variance	1.5	3

A portfolio consisting solely of Asset 1 is MV efficient. So is a portfolio consisting solely of Asset 2. Yet Asset 2 $\succ_1$ Asset 1 and Asset 2 $\succ_2$ Asset 1. So it is possible for one MV efficient portfolio to stochastically dominate another.

The stochastic dominance linear programme would seem to be theoretically superior to the Markowitz model in that it makes minimal assumptions regarding investor preferences and no assumptions regarding the shape of the return distribution. The Markowitz model penalises positive variance, which should not be considered as risk. Alternatively, it assumes normality of returns, whereas stochastic dominance makes no assumptions about the return distribution. If the returns do have a normal (or symmetrical) distribution, then the mean-variance efficient portfolios will be efficient in the stochastic dominance framework. In this case solutions to the Markowitz model and the stochastic dominance LP will coincide. However, return distributions are generally not normal, but skewed as large negative returns and large positive returns occur more often than would be expected from a normal distribution.

5 Modelling Issues in Portfolio Optimization

The above example illustrated how we can use historical returns to construct an empirical finite discrete distribution to use in either mean-variance optimization or optimization subject to stochastic dominance constraints. The above example also illustrated why the stochastic dominance solution could be considered theoretically superior to the mean-variance solution. However, for a real world investor, theoretical superiority matters little. Therefore we must consider the limitations of portfolio optimization.

We conducted some out of sample simulation of stochastic dominance and mean-variance portfolios and while the results are too lengthy to be included in this paper, in essence we found their performance to be satisfactory and in some cases very encouraging. However, the fundamental problem with any form of portfolio optimization is that there is no guarantee that future returns will come from the same distribution used to construct the portfolio and as such there is no guarantee that dominance or mean-variance efficiency will hold in an out of sample period.

In testing portfolios in out of sample data, it was often observed that the mean-variance portfolios exhibited more consistency (in terms of returns) than stochastic dominance portfolios although they typically did not reach the same return as the stochastic dominance portfolios by the end of the out of sample period. The mean-variance model seeks a portfolio with minimal variance and hence low return volatility and as such it was encouraging that this was what we observed.

An issue unique to the stochastic dominance method is that of choosing the benchmark. The second order stochastic dominance linear programme effectively partitions the set of portfolios defined by the convex hull of the N assets into two subsets. One subset contains those portfolios that do not second order stochastically dominate the benchmark. The other set contains those portfolios that do dominate the benchmark. We are then effectively solving the maximise expected return problem on this reduced set of feasible portfolios.

Some benchmarks are easier to dominate than others. If a “poor” benchmark is chosen, then it is possible that most of the portfolios that can be constructed from our N assets will dominate it. In the event that all the portfolios that can be constructed dominate the benchmark then our stochastic dominance LP becomes the maximise expected return problem and no longer reflects risk aversion. The following example illustrates what can happen when a “poor” benchmark is selected:

Example: Suppose we solve the SSD LP for the benchmark given in the table below. Suppose the optimal solution is the portfolio P^*, the returns of which are also in the table. Now suppose there is some other portfolio P, that is also a feasible solution to the SSD LP but is not the optimal solution since $E[P^*] > E[P]$. Despite this, the portfolio *P* would seem to be preferable as it has no negative returns and in 4 out of 5 realisations outperforms *P**.

	Returns		
	Bench	*P**	*P*
1	-10	-5	5
2	-5	0	10
3	0	2	13
4	5	5	15
5	10	70	17
Expected Return	0	14.4	12

In this example, a benchmark with a worst case return greater than -5 would make P^* infeasible.

6 Conclusions

We have outlined the stochastic dominance decision rules and seen how the second order stochastic dominance decision rule can be used as a constraint in a linear programme for portfolio optimization. We observed that the SSD LP maximises the expected return of a portfolio of assets subject to the portfolio second order stochastically dominating a benchmark reference return and we have seen an application of this model to some real world data. Furthermore, we compared and contrasted the solutions obtained using the SSD LP with those obtained using the Markowitz model for mean-variance optimization and have found that the mean-variance optimization model penalises positive variance and as such is not an ideal model for dealing with risk.

The stochastic dominance model makes minimal assumptions about investor preferences and return distributions but care must be taken in selecting a benchmark as this effectively induces a risk measure. We know that the mean-variance and stochastic dominance efficient sets can coincide but they can differ also and while we can plot the mean variance efficient frontier it is not yet known how we can obtain all the portfolios in the stochastic dominance efficient set. However, we do know that an optimal solution to the SSD LP will be efficient in a second order stochastic dominance framework.

Finally, we conclude that portfolio optimization is a reasonable practice if we have reason to believe that the distribution of future returns will closely match the distribution of historical returns. However, there is no guarantee that this will ever be the case and consequently it would probably be unwise to use any portfolio optimization method as the sole decision tool when choosing a portfolio. However, even an astute investor who is able to use a mixture of qualitative and quantitative information to determine his or her investment candidates still needs some way of deciding what proportion of the portfolio each candidate should make up. For this purpose portfolio optimization can aid in the decision making.

7 References

Dentcheva, D. & Ruszczynski, A. 2003. “Optimization with Stochastic Dominance Constraints.” *Siam J. Optimization* **14,2**,548-566.

Levy, H. 1998. *Stochastic Dominance : Investment Decision Making Under Uncertainty.* Kluwer Academic Publishers, Massachusetts.

Wind Farm Optimization

S. Donovan
Department of Engineering Science
University of Auckland
New Zealand
sdon032@ec.auckland.ac.nz

Abstract

This paper will formulate integer programs to determine the optimal positions of wind turbines within a wind farm.

The formulations were based on variations of the vertex packing problem. Three formulations are presented, which seek to maximize power generated in accordance with constraints based on the number of turbines, turbine proximity, and turbine interference. These were in the form of budget, clique, and edge constraints.

Results were promising, with turbines exhibiting a tendency to concentrate in areas of high elevation and avoid situations where downstream interference would be significant.

1. Development of the integer programs

Three (mixed) integer programming models are presented. The first two integer program models are vertex packing problems, while the third MIP model is a Generalized Vertex Packing Problem (GVP) [1]. The GVP problem was introduced by Hanif D. Sherali and J. Cole Smith [2].

In these formulations, G = (V, E) denotes a graph with vertices V and edges $E \subseteq V \times V$. The set E is set of vertex pairs between which there exists some relationship. In our case, the vertices V correspond to the locations where turbines can be positioned, and the edges E represent relationships between the vertices, such as turbine proximity and interference.

An appreciation of the relationship between the physical domain and the graph on which the (mixed) integer programs are based is crucial to understanding the material that follows.

The graph is based on an orthogonal grid that is superimposed onto the physical topography. The intersection points of this grid represent the vertices in our graph. The vertex packing problem will thus involve selecting the combination of vertices, or grid points, which generates the most power.

2. Modeling turbine proximity

The first integer program formulation enforced a minimal separation distance between turbines to ensure the blades did not physically clash with one another. The term proximity shall define the area immediately surrounding a turbine in which no other turbine can be built. The grid points that lie within this area are a function of the turbine radius and the physical distance between the intersection points in our grid.

Figure 2.1 demonstrates that a turbine centered on the solid vertex will eliminate the surrounding vertices as potential locations. That is, the vertices connected to the solid vertex by an edge are too close to accommodate another turbine. Those vertices that are not connected to the solid vertex do not impinge on the space required by this turbine.

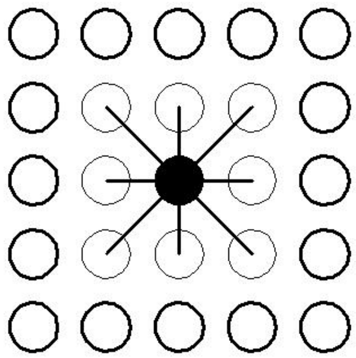


Figure 2.1: Vertex proximity constraint.

To model this proximity requirement we construct the graph G by considering each vertex in turn, and placing an edge between this vertex *i* and any vertex *j*, where the position occupied by *j* violates the area required by a turbine positioned at *i*. For example, in Figure 2.1 an edge would exist between the solid vertex and every surrounding vertex, as shown by the lines. This constraint can be formulated mathematically as:

$$x_i + x_j \leq 1, \forall (i, j) \in E$$

An edge constraint of this form will exist between every vertex in our graph and any other vertex that lies within the required separation distance. In the physical model, this corresponds to a pair of grid points existing too close for a turbine to be located at both positions.

The above "weak edge" formulation can be improved by considering a larger subset of vertices affected by turbine proximity. The structure of the edges on *G*, as well as the relationship between four neighboring vertices Q, is shown in Figure 2.2.

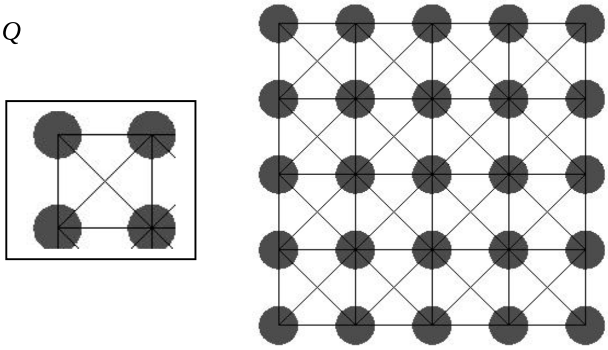


Figure 2.2: Clique structure.

A maximal clique Q is defined as a maximal subset (with respect to cardinality) of vertices whereby all vertices in the subset are connected by an edge to all the other vertices in the subset. For the subset Q shown in Figure 2.2, the clique would involve all four vertices connected by an edge, with the sum of the turbines constructed in that subset constrained to be less than or equal to one. In general terms, the cardinality of a clique will be a function of the turbine radius and distance between the grid points in the x and y direction. Let K denote the set of all maximal cliques in our graph G. Each maximal clique $Q \in K$ is a subset of V.

Let W_v denote the power value associated with a vertex v. Let $x_v=1$ denote a turbine positioned at vertex v, and $x_v=0$ otherwise. A budget constraint restricts the maximum number of turbines to be built in the wind farm to be less than or equal to k.

The integer program can now be formulated as:

$$\begin{aligned} &\text{Maximize} && \sum_{v \in V} W_v x_v \\ &\text{Subject to} && \sum_{v \in V} x_v \leq k \\ &&& \sum_{v \in Q} x_v \leq 1, \quad \forall Q \in K \\ &&& x_v \in \{0,1\}, \quad v \in V \end{aligned}$$

This formulation shall be referred to as IP1.

IP1 led to dense clusters of turbines in areas of high power resource. This stems from the fact that the only constraint on turbine location was the proximity constraint defined by the radius of the turbine blades. To this end, IP1 ensures that turbine proximity is not violated. It does not, however, reflect the influence of interference on the amount of power generated at downstream locations.

An example of a 50 turbine farm optimized using IP1 is shown in Figure 2.3, where the turbine locations are indicated by the dots. Recall this has been optimized for wind flow from the west.

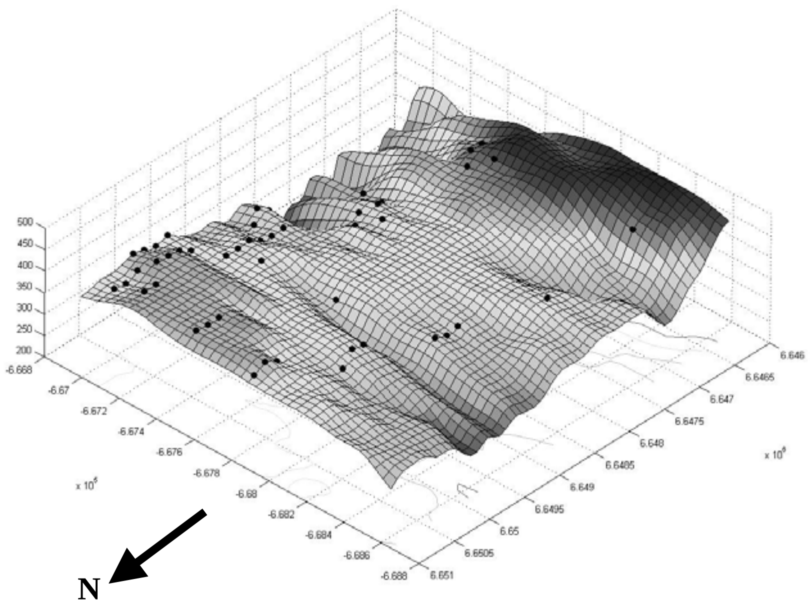


Figure 2.3: IP1 results (50 turbines).

The solution to IP1 has the following features:

1. Turbines are concentrated in areas of high power resource.
2. The proximity constraint prevents turbines from occupying adjacent vertices.

3. Modeling turbine interference

Interference between turbines was not taken into account in the formulation of IP1. The industry standard recommended separation distance will now provide the basis for an integer program formulation that accounts for turbine interference. The distance required between turbines is taken to be 7 turbine diameters when aligned in the predominant direction of flow, and 3 turbine diameters in the other direction [3]. Figure 3.1 demonstrates these separation distances when a turbine is centered at the shaded vertex.

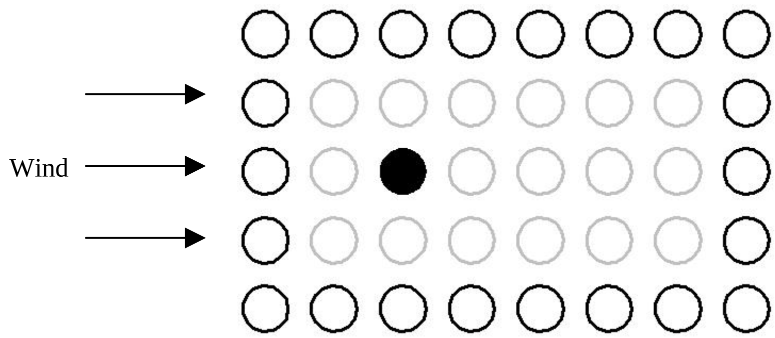


Figure 3.1: Fixed separation distance.

Figure 3.1 shows that a turbine centered at the solid vertex will influence the immediate surrounding vertices as well as vertices further downstream. Vertices outlined with a dark line are unaffected. This model of interference will, therefore, increase the size of the cliques to reflect both turbine proximity and turbine interference. This formulation shall be referred to as IP2.

The results from IP2 varied significantly from IP1, as shown in Figure 3.1.

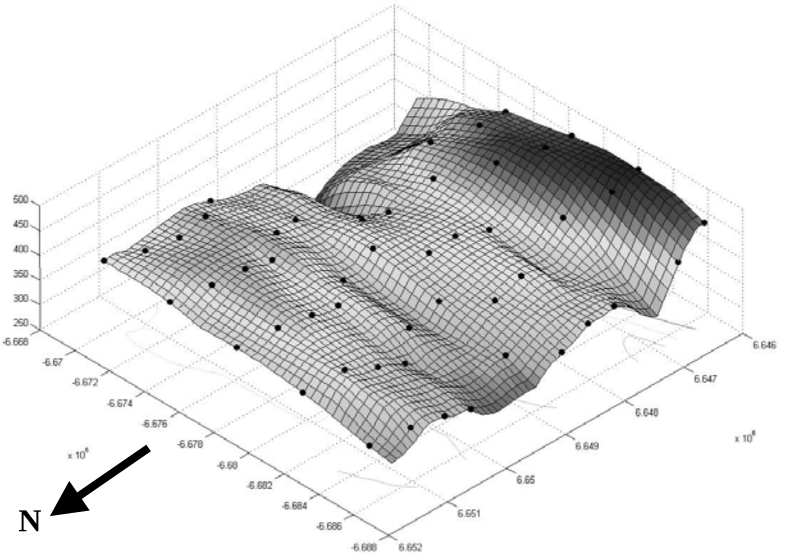


Figure 3.2: IP2 results (50 turbines).

The solution to IP2 has the following unique features:

1. The larger cliques prevent turbines from clustering
2. Turbines are aligned perpendicular to the main direction of flow, reflecting the smaller separation distance imposed in this direction.

IP2 is an improvement over IP1 because turbine interference is taken into account.

4. A better model for turbine interference

While IP2 is an improvement over IP1, the imposition of an arbitrary separation distance between turbines is a blunt approach to turbine interference. Instead, it would be better to position turbines according to the net power gain, which is defined as the amount of power generated less the magnitude of interference.

This separates the area surrounding each turbine into two distinct measures, which reflect:

1. Proximity
2. Interference

These measures are shown in Figure 4.1.

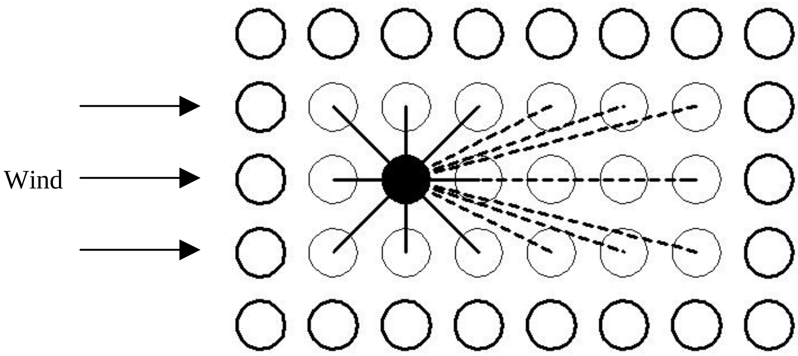


Figure 4.1: Proximity and interference effects.

Figure 4.1 shows that a turbine positioned at the solid vertex will physically obstruct the surrounding vertices, to which it is connected by a solid line. As a result, the set of all cliques K is identical to the proximity constraints formulated in IP1.

However, this model distinguishes itself by introducing edge based on the interference between vertices. The graph G considers each vertex in turn, and places an edge between this vertex *u* and any vertex *v* that experiences interference above a certain magnitude. In Figure 4.1, for example, an edge would exist between the solid vertex and every vertex that is connected to it via a dashed line.

Let E_I denote the set of edges between all vertices *u* and *v* that interfere with each other. Variable $z_{uv}=1$ if a turbine is positioned at *u* and *v*, and $z_{uv}=0$ otherwise. This can be formulated mathematically as:

$$x_u + x_v - 1 \le z_{uv}$$

Let I_{uv} denote the magnitude of the power loss caused by interference between vertices *u* and *v*. The methods used to determine *I* are outside the scope of this paper. If $z_{uv}=1$, which denotes that a turbine is positioned at both x_u and x_v, then the expected value of the power generated will decrease by the amount I_{uv}.

The mixed integer program formulation, which shall be referred to as MIP1, then becomes:

$$\begin{aligned}
\text{Maximize} \quad & \sum_{v \in V} W_v x_v - I_{uv} z_{uv} \\
\text{Subject to} \quad & \sum_{v \in V} x_v \le k \\
& \sum_{v \in Q} x_v \le 1 \\
& x_u + x_v - 1 \le z_{uv}, (u, v) \in E_I \\
& x_v \in \{0,1\}, v \in V, Q \in \theta, z \ge 0
\end{aligned}$$

The decision variables *z* are not constrained to be binary because the formulation enforces them to take binary values of 0 or 1. This follows from *I* being strictly positive. A variable z_{uv} takes a value of 0 unless the corresponding interference constraint forces it to assume a value greater than or equal to 1. In this event, the deleterious impact of *I* on the objective function causes z_{uv} to assume the smallest value possible, which is 1.

The results generated using MIP1 are shown in Figure 4.2.

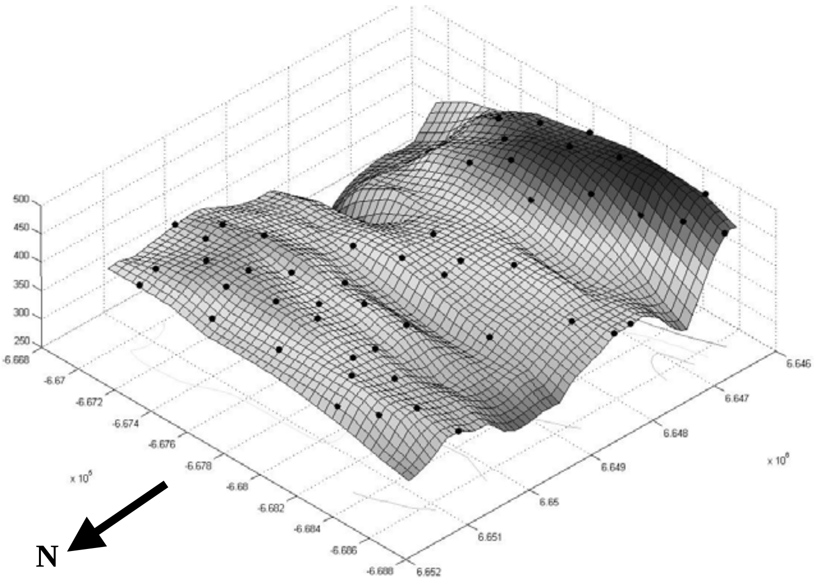


Figure 4.2: MIP1 results (50 turbines).

The major difference between IP2 and MIP1 is in the ability of the latter to evaluate the balance between power generation and interference losses. This means that in situations of high power resource, MIP1 is willing to accept interference if the net power generated will be more than what would result from another position.

Therefore, the MIP1 formulation is distinct from IP1 and IP2 in that it strictly enforces turbine proximity, while accounting for turbine interference using edges. MIP1 can evaluate the net benefit of turbine interference in positions of high power resource, and position turbines accordingly.

5. Analysis of results

All of the aforementioned formulations were modeled in AMPL and solved using CPLEX version 6.6.0 with default settings. AMPL is a text based algebraic modeling language used to program optimization problems. CPLEX is a commercial solver for mixed integer problems.

The difference between the expected power outputs for IP2 and MIP1 reflects the value of compromising on the separation distance between turbines in areas of high power resource. MIP1 outperformed IP2, particularly as the maximum number of turbines k was increased. This trend is shown in Figure 5.1.

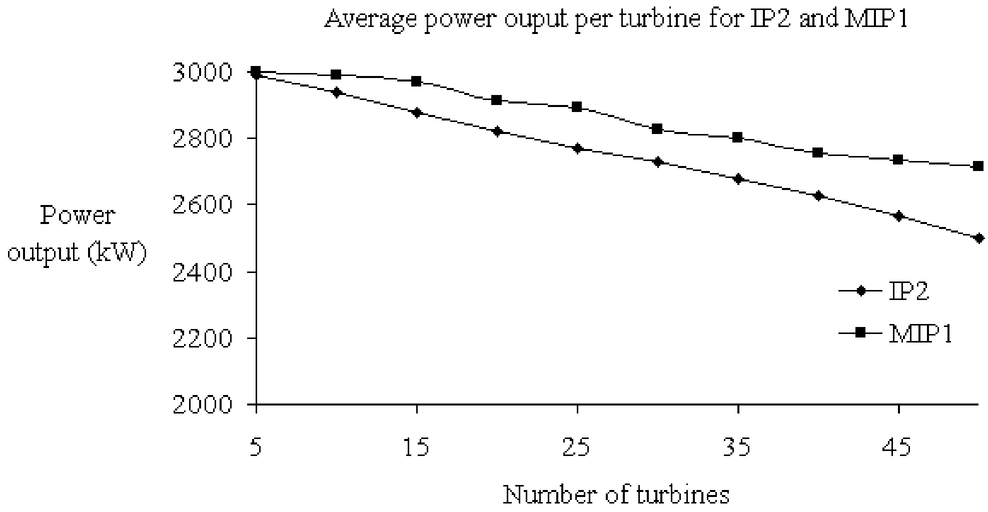


Figure 5.1: Comparison of average power output from IP2 and MIP1.

Figure 5.1 demonstrates that, for wind farms involving more than 45 turbines, the optimal configuration determined by MIP1 would generate approximately 10 percent more power than IP2. This is a significant gain.

6. Performance of MIP1

General vertex packing problems, which contain vertex packing problems as a specific case, are NP-hard. It is extremely unlikely that there exists an efficient algorithm to solve these problems. Efficient means that the algorithm will run in polynomial time.

This section will assess the performance of MIP1. The number of decision variables involved in MIP1, as determined by the number of vertices and interference edges, is a function of the number of grid points and the methods used to determine interference.

Graphs of the problem size against number of variables and number of non zeros are shown in Figures 6.1 and 6.2.

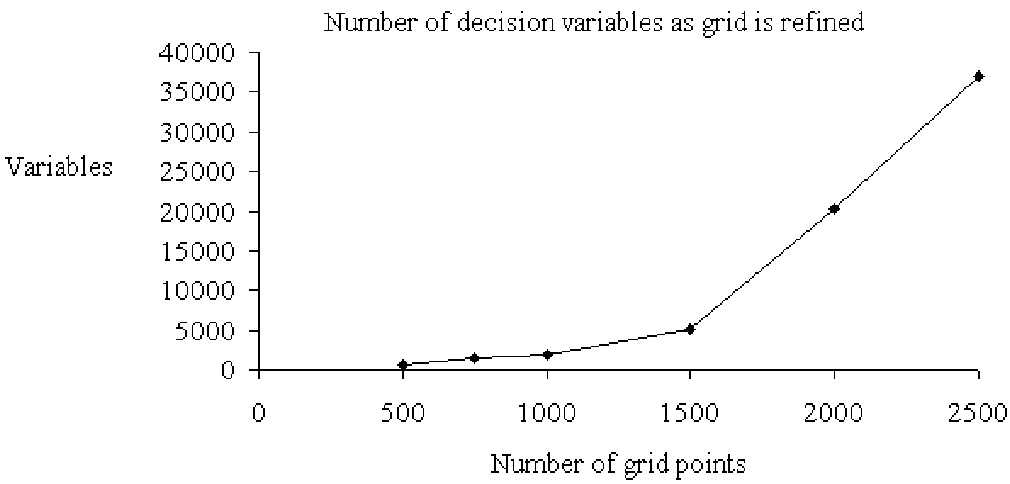


Figure 6.1: Number of variables versus problem size.

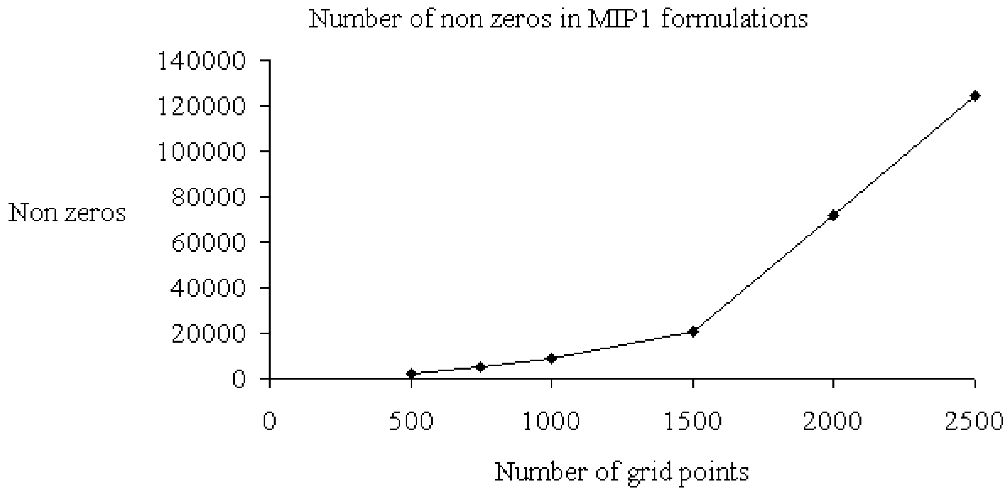


Figure 6.2: Number of non zeros versus problem size.

Moreover, the maximum number of turbines in the farm, *k*, also had a significant impact on the time taken to solve to optimality. For a sample problem involving 2400 grid points, only configurations involving less than 10 turbines could be solved to optimality in less than one hour. This was on a machine operating Windows 2000 with a 3.00MHz Pentium IV processor and 1 gigabyte of RAM.

The gap between the value of the best feasible solution and the best bound found after one hour is plotted in Figure 6.3.

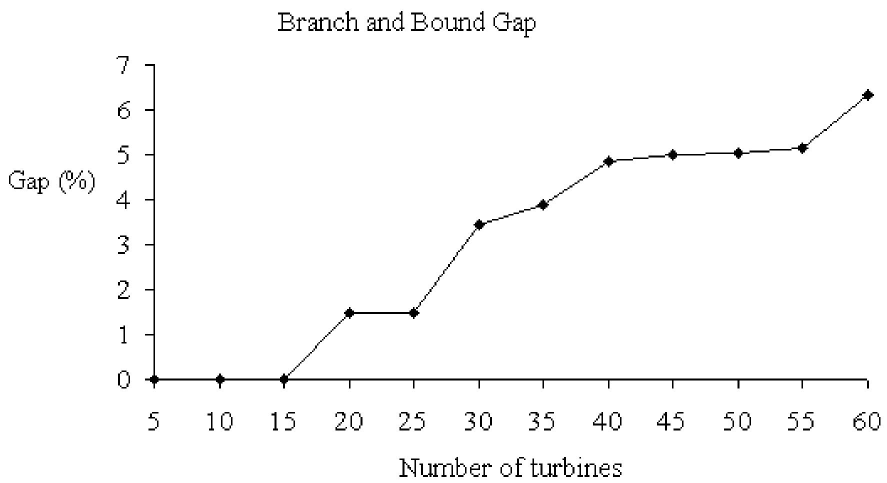


Figure 6.3: Branch and bound gap.

7. Conclusions and future work

This paper has outlined methods for the optimization of turbine locations within a wind farm. Several integer program formulations were presented. The final formulation was designed to maximize the total power generated while observing turbine size and turbine interference. Constraint formulation was based on cliques and edges derived from the underlying vertex graph.

The integer programs could be solved close to optimality in acceptable time. Results were consistent with expectations, with turbines exhibiting a preference for areas of high elevation. The MIP used in this project performed well, even when confronted with relatively large problems. However, the potential exists for the model to be extended beyond its current form.

A potentially important constraint would exclude areas of unduly steep terrain from being selected as a turbine location. This could be incorporated by using information on topographical gradients.

Another interesting constraint would involve relating the total distance between turbines to some construction cost that reflects, for example, the length of trenching required. This would involve another "interference" type sub graph, where cost was linear in variables defined over vertex pairs. More complex interference shapes could be modeled using a set packing formulation.

The budget constraint k could be replaced by a measure of productivity, which controlled the maximum number of turbines that are built. The MIP would then construct as many turbines as possible, while satisfying some minimum output for each turbine. This minimum output could be determined by considering the desired payback period for the wind farm investment.

This means that for every turbine location i, the amount of power generated less the interference experienced must exceed some critical value, denoted by P. This constraint is formulated mathematically as:

$$w_i x_i - \sum_{(i,j)\in E_I} I_{ji} z_{ij} \geq P \quad \forall i \in V$$

There are also opportunities for the application of heuristic algorithms to work towards improved solutions. These may be particularly useful for problems involving nonlinear constraints, such as noise and line of sight, or for improving on a non optimal solution generated by the branch and bound process. The greater complexity of these constraints may well be suited to heuristic, rather than exact, solution methods.

8. References

[1] *A polyhedral study of the generalized vertex packing problem*, H. D. Sherali and J. C. Smith, Math. Programming, to appear.

[2] *A class of web-facets for the generalized vertex packing problem*, H. D. Sherali and J. Cole Smith, in 'Discrete Applied Mathematics', Vol. 146 (2005), pp. 273 – 286.

[3] *Fact Sheet 5: Siting*, New Zealand Wind Energy Association, The Cost of Wind Energy, www.windenergy.org.nz/FactSheets/Siting5, accessed 13/4/05, last updated 13/6/05.

Heuristics for a tree covering problem

Dr Shane Dye
Department of Management
University of Canterbury
New Zealand
shane.dye@canterbury.ac.nz

Abstract

Heuristics are developed for a tree covering problem where the tree is covered by subtrees, pairs of which overlap by at most one node. The problem arises in the implementation of a stochastic programming decomposition algorithm where the decomposition is based on a covering of the scenario tree.

An algorithm for optimal covering is developed. The algorithm is based on repeated optimal bin-packing with a modified objective function. For a class of trees common in stochastic programming (so-called *balanced* scenario trees) the algorithm has a quadratic running time. For general trees, heuristics are developed by using various existing heuristics for the bin-packing steps. Performance results are derived based on the performance of the bin-packing heuristic used.

1 Introduction

An NP-hard tree covering problem is investigated. A tree is covered by subtrees so that each node of the tree is contained in at least one of the subtrees. The number of nodes in each subtree restricted and there is a further restriction that any pair of subtrees can have at most one common node. A cover is sought, which minimises the number of subtrees used.

The problem arises in the context of a stochastic programming algorithm [3], which decomposes the deterministic equivalent problem into subproblems corresponding to the components of a subtree cover of the scenario tree. The size restriction on the subtrees is to limit the size of the subproblems defined. Nodes of the scenario tree which are common to more that one subtree represent variables and constraints which appear in multiple subproblems. 1-overlap covers provide a balance between (arbitrary) subtree covers, which may have too many common nodes, and subtree partitions, which may produce too many small subproblems.

Versions of tree covering problems have been previously studied where the subtrees used in the cover are drawn from an explicit collection [1, 4]. The related subtree partitioning problem is in P[5].

Section 2 formally describes the problem and shows that it is NP-hard. An optimisation algorithm based on optimal bin-packing is described in Section 3. The algorithm involves multiple bin-packing steps where subtrees connected to child nodes are packed together to form subtrees.

Section 4 shows that for so-called balanced scenario trees, the algorithm has running time quadratic in the number of nodes. Balanced scenario trees have common branching degrees within stages. For other trees, heuristics are developed by replacing the optimal bin-packing with various heuristics. The worst-case performance of these heuristics is investigated in Section 5.

2 Minimum 1-overlap cover is NP-hard

Given tree $T = (V, E)$ and positive integer K, a *1-overlap cover*, $\mathcal{D}$, of T with maximum component size K, is a cover of V such that every component, $U \in \mathcal{D}$, has $|U| \leq K$ and induces a *subtree* $S_U = (U, E \cap U \times U)$, *i.e.*, S_U is connected. Further every distinct pair of components, $U, W \in \mathcal{C}$, $U \neq W$, have at most one common node, *i.e.*, $|U \cap W| \leq 1$. Here, the trivial graph with one node and no arcs is considered to be a tree.

MINIMUM COMPONENT 1-OVERLAP COVER (M1C)

INSTANCE: A tree $T = (V, E)$ and positive integer K.

SOLUTION: A 1-overlap cover with maximum component size K, $\mathcal{C}$.

MEASURE: The minimal number of components, *i.e.*, $\min |\mathcal{C}|$.

To aid the discussion we assume that the tree has a root node. If this is not the case an arbitrary node will suffice. The concepts of the parent node, child nodes, ancestors and descendents are used with regard to the root. Also, given the application, the concept of stages within the tree is used. The stage of a node is the number of nodes in the (unique) path from the node to the root. For an arbitrary subtree, S of T, we define the root of S to be the node in S with the lowest stage.

Theorem 1 *M1C is NP-hard.*

Proof: We show that M1C is NP-hard using a pseudo-polynomial transformation from the strongly NP-complete 3-PARTITION [4].

3-PARTITION

INSTANCE: A finite set A of $3m$ elements, a bound $B \in \mathcal{Z}^+$ and a size $s(a) \in \mathcal{Z}^+$ for each $a \in A$, such that each $s(a)$ satisfies $B/4 < s(a) < B/2$ and such that $\sum_{a \in A} s(a) = mB$.

QUESTION: Can A be partitioned into m disjoint sets $S_1, S_2, \ldots, S_m$ such that for $1 \leq i \leq m$, $\sum_{a \in S_i} s(a) = B$?

A tree is constructed from an instance of 3-PARTITION in the following way. Begin with the root node with $3m$ child nodes (called the second stage nodes), each

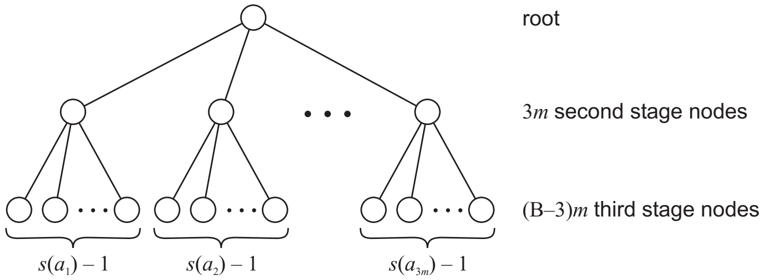


Figure 1: The tree constructed for an instance of 3-PARTITION

one corresponding to a different element of A. For each $a \in A$ the corresponding second stage node will have $s(a) - 1$ children. All of these children comprise the third stage nodes. Figure 1 shows the transformation.

The instance of M1C comprises the constructed tree, $T = (V, E)$ and $K = B+1$. Clearly the transformation is polynomial in m and B and $|V| = mB + 1$. All that remains is to show that the instance of 3-PARTITION will have a yes answer if, and only if, the minimal 1-overlap cover has m components.

First we show that if 3-PARTITION has a yes answer, an m component 1-overlap cover exists. Let $S_1, S_2, \ldots, S_m$ comprise the partition of A corresponding to the yes answer. Construct vertex subsets $R_1, R_2, \ldots, R_m$ as follows. Initially, each subtree contains the root. For each $i = 1, \ldots, m$, if $a \in S_i$ add the second stage node corresponding to a and all its child nodes to R_i. By construction, each R_i induces a subgraph and, since $\{S_1, S_2, \ldots, S_m\}$ covers A, $\{R_1, R_2, \ldots, R_m\}$ covers V. Since $S_1, S_2, \ldots, S_m$ are disjoint each distinct pair R_i and R_j overlap by only the root node. Also

$$|R_i| = 1 + \sum_{a \in S_i} (1 + (s(a) - 1)) = 1 + B = K.$$

Next we show that any m component 1-overlap cover corresponds to a yes answer for 3-PARTITION and that there can be no 1-overlap covers with fewer than m components. Let $\{R_1, R_2, \ldots, R_m\}$ correspond to a subtree cover of T with each component covering no more than $K = B + 1$ nodes. If any R_i does not cover the root node then $|R_i| \leq s(a_i)$ for some unique $a_i \in A$. Without loss of generality assume that the first $j \geq 0$ subtrees do not cover the root node, while the remainder do. Using the individual maximum sizes, define upperbound, UB, on the total component size as

$$\sum_{i=1}^{m} |R_i| \leq \sum_{i=1}^{j} s(a_i) + (m - j)(B + 1) = UB.$$

Since T is covered, with the root node covered $m - j$ times, define lowerbound, LB, on the total component size as

$$\sum_{i=1}^{m} |R_i| \geq mB + (m - j) = LB.$$

The difference between the two bounds is

$$UB - LB = \sum_{i=1}^{j} s(a_i) - jB \leq -jB/2,$$

which is negative unless $j = 0$. So all R_i cover the root node. This also shows that no 1-overlap cover can have fewer than m components. It also follows that the subtree covering a second stage node must also cover all of its child nodes. For each $a \in A$ let n_a be the second stage node in T corresponding to a. Define the following subsets of A for each R_i:

$$S_i = \{\, a \in A \mid n_a \in R_i \,\}.$$

Now, the sets $\{S_1, S_2, \ldots, S_m\}$ partition A and for any $i = 1, \ldots, m$

$$\sum_{a \in S_i} s(a) = \sum_{a \in A | n_a \in R_i} s(a).$$

We then have

$$|R_i| = 1 + \sum_{a \in A | n_a \in R_i} (1 + (s(a) - 1)) = 1 + \sum_{a \in S_i} s(a).$$

Maximum component size is $K = B + 1$, so $\sum_{a \in S_i} s(a) \leq B$. Equality results from

$$mB = \sum_{a \in A} s(a) = \sum_{i=1}^{m} \sum_{a \in S_i} s(a) \leq mB$$

since $\{S_1, S_2, \ldots, S_m\}$ partitions A. □

When the tree has only two stages it is easy to see that the problem may be solved in linear time. When the tree is restricted to three stages, the proof above shows that the problem is still NP-hard since the tree constructed in the proof above has just three stages.

3 An algorithm based on bin-packing

This section gives an algorithm for determining the minimal 1-overlap cover. The algorithm is based on bin-packing using a modified objective function and it allows many results and heuristics from the well studied area of bin-packing to be employed for M1C. The algorithm also provides a basis for heuristics.

In what follows, r is the root node of T, and for any node, x, define $ch(x)$, $V_T(x)$ and $R_T(x)$ as follows. $ch(x)$ is the set of child nodes of x (*w.r.t.* root r), $V_T(x)$ is the set of nodes consisting of x and all nodes which are not in the same graph component as the root node when node x is removed from T. $R_T(x)$ is the subgraph induced by $V_T(x)$; $R_T(x)$ is a subtree. Also for 1-overlap cover $\mathcal{C}$ and any $U \in \mathcal{C}$ define

$$N_{\mathcal{C}}(U) = \{\, j \in U \mid j \notin W, \forall W \in \mathcal{C}, W \neq U \,\}.$$

The bounds on the running time, below, are given in terms of $n = |V|$ and $c = max_{x \in V} |ch(x)|$ which is the maximum number of children of any node.

1-Overlap Cover Algorithm

1. Begin with current tree $T^0 = (V^0, E^0) = T$, current cover, $\mathcal{C}^0 = \emptyset$, and $k = 0$.
2. Label all nodes with $\ell^k(x) = |V_{T^k}(x)|$.
3. If $\ell^k(r) \leq K$, put $\mathcal{C}^{k+1} = \mathcal{C}^k \cup V^k$ and the algorithm is finished, $\mathcal{C}^{k+1}$ is the 1-overlap cover produced.
Otherwise go on to Step 4.
4. If any node, x, has $\ell^k(x) = K$, put $\mathcal{C}^{k+1} = \mathcal{C}^k \cup V_{T^k}(x)$ and $V^{k+1} = V^k \setminus V_{T^k}(x)$ then go to Step 6. Otherwise go on to Step 5.
5. Choose any node, x, with $\ell^k(x) > K$ but all child nodes, y, have $\ell^k(y) < K$.
   (a) Solve a modified bin-packing problem with bins of size $K - 1$ and items corresponding to the child nodes, y, with sizes $\ell^k(y)$. The objective function is to first minimise the number of bins, then (for this number of bins) minimise the space used in Bin 1.
   (b) For each bin except Bin 1 construct a subtree and add it to the collection as follows. Let the items packed in the bin correspond to child nodes, $y_1, y_2, \ldots, y_q$, for some q. Define vertex set $U = \{x\} \cup \bigcup_{i=1}^{q} V_{T^k}(y_i)$. Put $\mathcal{C}^{k+1} = \mathcal{C}^k \cup U$ and $V^{k+1} = (V^k \setminus U) \cup \{x\}$.
6. Increment k by one, put $E^k = (V^k \times V^k) \cap E^{k-1}$ and put $T^k = (V^k, E^k)$. Return to Step 2.

Let $B(c)$ be the worst case running time of the bin-packing step, which cannot be solved in polynomial time unless P = NP and is $O(c!)$. The bi-criteria objective required can be converted to the single objective function $\min Kb + s$ (where b is the number of bins used and s the space used in Bin 1) since saving a bin improves the objective function by more than the maximum saving from Bin 1.

The overall worst case running time is then $O(n^2 + nB(c))$. Note that $c < n$. In general we would expect the running time to be dominated by $nB(c)$. When c is fixed or bounded the algorithm has theoretical running time of $O(n^2)$ (however the constant would be expected to be large).

The following results are needed to show that the algorithm developed solves the problem posed. They provide insight into the structure of an optimal solution.

For tree T and integer $K > 0$, define $\mathcal{M}(T, K)$ to be the collection of all minimum component 1-overlap covers of T with maximum component size K.

Lemma 1 *Let $\mathcal{C} \in \mathcal{M}(T, K)$ for tree $T = (V, E)$ and let $U \in \mathcal{C}$ have the properties that $V' = V \setminus N_{\mathcal{C}}(U)$ induces a (connected) subtree, S, and $|U| - |N_{\mathcal{C}}(U)| \leq 1$. If $\mathcal{D} \in \mathcal{M}(S, K)$, then $\mathcal{D} \cup \{U\} \in \mathcal{M}(T, K)$.*

Proof: $\mathcal{C} \setminus \{U\}$ is a 1-overlap cover of S so $|\mathcal{D}| \leq |\mathcal{C}| - 1$. Also, $\mathcal{D} \cup \{U\}$ is a 1-overlap cover of T since U overlaps V' with at most one node. This means that $|\mathcal{C}| \leq |\mathcal{D}| + 1$. The result follows immediately. □

Lemma 2 *For any $\mathcal{C} \in \mathcal{M}(T, K)$, of all of the subtrees that cover any non-root node x, at most one of them covers the parent of x.*

Proof: This follows immediately from the requirement that for any $U, W \in \mathcal{C}$ with $U \neq W$, $|U \cap W| \leq 1$. $\square$

Corollary 1 *For $\mathcal{C} \in \mathcal{M}(T, K)$, if more than one subtree covers a node x, all but at most one of them have x as the root.*

Corollary 2 *For any $\mathcal{C} \in \mathcal{M}(T, K)$ and any node x, there is at most one $U \in \mathcal{C}$ such that $U \cap V_T(x) \neq \emptyset$ and $U \not\subseteq V_T(x)$.*

Lemma 3 *For tree $T = (V, E)$, if any node x has $|V_T(x)| = K$, there exists $\mathcal{C} \in \mathcal{M}(T, K)$ such that $V_T(x) \in \mathcal{C}$ and $U \cap V_T(x) = \emptyset$ for all $U \in \mathcal{C}$, $U \neq V_T(x)$.*

Proof: Let $\mathcal{C}' \in \mathcal{M}(T, K)$ and put $\mathcal{E} = \{ U \in \mathcal{C}' \mid U \cap V_T(x) \neq \emptyset \}$. If $|\mathcal{E}| = 1$ the hypothesis holds. Otherwise $|\mathcal{E}| \geq 2$ and by Corollary 2 the number of subtrees in $\mathcal{E}$ which cover nodes outside $V_T(x)$ is either none or one. For the latter case let U' be the component of $\mathcal{E}$ which covers nodes outside of $V_T(x)$. For the former put $U' = \emptyset$. Now,

$$\mathcal{C}'' = (\mathcal{C}' \setminus \mathcal{E}) \cup \{V_T(x), U' \setminus V_T(x)\}$$

is a 1-overlap cover with $|\mathcal{C}''| = |\mathcal{C}'| - |\mathcal{E}| + 2 \leq |\mathcal{C}'|$. So $\mathcal{C}'' \in \mathcal{M}(T, K)$ $\square$

Lemma 4 *Let $T = (V, E)$ be a tree with the property that $|V| > K$ and all child nodes, y, of the root have $|V_T(y)| < K$. There exists $\mathcal{C} \in \mathcal{M}(T, K)$ constructed corresponding to the optimal packing of a bin-packing problem with bin size $K - 1$ and items, i_y, corresponding to the child nodes with sizes $s(i_y) = |V_T(y)|$. If the optimal packing uses m bins, then $|\mathcal{C}| = m$.*

Proof: Let r be the root node. For an optimal packing using m bins construct a 1-overlap cover $\mathcal{B}$ in the following way. For each bin, b, the items packed correspond to q_b of the child nodes, $y_{\sigma_b(1)}, y_{\sigma_b(2)}, \ldots, y_{\sigma_b(q_b)}$. Define a component of $\mathcal{B}$, $U_b = \{r\} \cup \bigcup_{k=1}^{q_b} V_T(y_{\sigma_b(k)})$. Each U_b is connected and has

$$|U_b| = 1 + \sum_{k=1}^{q_b} |V_T(y_{\sigma_b(k)})| = 1 + \sum_{k=1}^{q_b} s(i_{y_{\sigma_b(k)}}) \leq K.$$

Now consider any $\mathcal{C} \in \mathcal{M}(T, K)$. From the above we have $|\mathcal{C}| \leq m$. Construct $\mathcal{C}' \in \mathcal{M}(T, K)$ so that each child node of the root, y, has $V_T(y)$ covered fully by a single component as follows. Consider a child node, y, for which $V_T(y)$ is covered by more than one component of $\mathcal{C}$. $V_T(y)$ must intersect some $U \in \mathcal{C}$, which also covers some sibling, z, since otherwise we could replace all components covering $V_T(y)$ by the single set $V_T(y) \cup \{x\}$, violating the optimality of $\mathcal{C}$. Lemma 2 implies only one such U exists. Replacing all components covering $V_T(y)$ with $\{U \setminus V_T(y), V_T(y)\}$, maintains minimality of the 1-overlap cover. Apply this logic repeatedly to generate $\mathcal{C}'$.

From $\mathcal{C}'$ construct a bin-packing using $|\mathcal{C}'|$ bins. Each component, $U \in \mathcal{C}'$, corresponds to a bin containing item i_y if $y \in U$. Assume U contains exactly q_U child nodes, $y_{\gamma_U(1)}, y_{\gamma_U(2)}, \ldots, y_{\gamma_U(q_U)}$.

If $q_U = 1$, the bin contains just one item which is feasible. Otherwise, U must also contain the root and bin capacity holds since

$$\sum_{k=1}^{q_U} s(i_{y_{\gamma_U(k)}}) = \sum_{k=1}^{q_U} |V_T(y_{\gamma_U(k)})| = |U_b| - 1 \leq K - 1.$$

Since this is a feasible packing, $|\mathcal{C}'| \geq m \geq |\mathcal{C}| = |\mathcal{C}'|$. $\square$

Theorem 2 *The 1-Overlap Cover Algorithm solves M1C.*

Proof: The algorithm is applied recursively, removing subtrees from the current tree each time (but possibly leaving their root nodes). Lemmas 1 and 3 imply that it suffices to show that for tree T and any node, x, satisfying the requirements of Step 5, there exists $\mathcal{C} \in \mathcal{M}(T, K)$ with subtrees as constructed (and removed) in this step.

Let $\mathcal{C} \in \mathcal{M}(T, K)$ and let x be any node satisfying the requirements of Step 5. Steps 3 and 4 guarantee the existence of such an x if Step 5 is reached.

By Corollary 2 either one subtree covers both x and its parent node (if it has one) or no subtrees do. First consider the latter case. For this case the covering of $R_T(x)$ can be considered independently of the rest of the tree, so we can assume, without loss of generality, that x is the root node. Lemma 4 shows that in this case the optimal bin-packing corresponds to a minimal 1-overlap cover.

Now consider the case where one component, $U \in \mathcal{C}$, covers both x and its parent node; put $P = U \cap V_T(x)$ and $Q = U \setminus P$. If $P = \{x\}$, Lemma 4 applied to $R_T(x)$ and then Lemma 1 applied recursively give the required result. Otherwise, define $\mathcal{E} = \{\, W \in \mathcal{C} \mid W \cap V_T(x) = \emptyset \,\}$ and $N = |\mathcal{C}| - |\mathcal{E}|$. Use Lemma 4 to construct $\mathcal{B} \in \mathcal{M}(R_T(x), K)$ in such a way that the corresponding optimal packing has least possible space used in Bin 1. Now, $|\mathcal{B}| \leq N$, since $\mathcal{C} \setminus \mathcal{E}$ covers $V_T(x)$ and Q induces a subtree.

If $|\mathcal{B}| \leq N - 1$, then the 1-overlap cover of T, $\mathcal{F} = \mathcal{E} \cup \mathcal{B} \cup \{Q\}$, has $|\mathcal{F}| = |\mathcal{E}| + |\mathcal{B}| + 1 \leq |\mathcal{E}| + N = |\mathcal{C}|$.

If $|\mathcal{B}| = N$, let L be the vertex set corresponding to Bin 1. If $|L| \leq |P|$, consider the 1-overlap cover $\mathcal{G} = \mathcal{E} \cup (\mathcal{B} \setminus \{L\}) \cup \{Q \cup L\}$. Now, $|\mathcal{G}| = |\mathcal{E}| + |\mathcal{B}| = |\mathcal{C}|$, $|Q \cup L| = |Q| + |L| \leq |Q| + |P| = |U| \leq K$ and $Q \cup L$ is connected since $x \in L$.

We are left with the case where $|\mathcal{B}| = N$ and $|L| > |P| \geq 2$. In this case the 1-overlap cover of $R_T(x)$ given by $\mathcal{H} = ((\mathcal{C} \setminus \mathcal{E}) \setminus \{U\}) \cup \{P\}$ cannot correspond to a feasible bin-packing since $|L| > |P|$. Since $|L| > |P|$ there must be some child node of x, y, such that $y \in P$ and $V_T(y) \setminus P \neq \emptyset$; therefore (by Corollary 2) there is another vertex set $W \in \mathcal{C}$ such that $W \subseteq V_T(y)$. In this case we can create $\mathcal{C}'' \in \mathcal{M}(T, K)$ by replacing W with $\{x\} \cup R_T(y)$ and U with $U \setminus R_T(y)$ in $\mathcal{C}$. This can be applied repeatedly (removing at least one vertex from U each time) until either $P = \{x\}$ or $\mathcal{H}$ corresponds to a feasible bin-packing. $\square$

From the above we see that it is possible to form a minimal 1-overlap cover recursively, examining smaller and smaller trees each iteration. Where subtrees overlap, all but one of those subtrees will have the overlapping node as the root and these overlapping nodes can be easily identified. Finally, bin-packing can be used to optimally group the subtrees attached to the child nodes of overlapping nodes. This is the essence of the algorithm described.

4 Branching rate constant for each stage

There is an interesting class of trees which can be covered in $O(n^2)$ time, these are trees for which the branching rate is constant for all nodes at the same stage. Such trees are called "balanced" in the stochastic programming literature [2].

For balanced trees the node labels will be constant at each stage. This means that the choice of node x at Step 5 will be over all nodes at the same stage and the child nodes of each such node will all have the same label (so the items to pack in the bin-packing will all have the same size). The optimal bin-packing is easily generated in linear time and this will be the same for each possible choice of x. Since the labels are only changed on ancestors of x, it is possible to choose all nodes at the same level in sequence to act as node x in Step 5. After all of these nodes have been visited, the tree will once again have common branching rates. This defines an $O(n^2)$ algorithm for solving such trees.

Since the labels are only changed on ancestors of x, the restriction in the order to pick the x nodes of Step 5 is not necessary. So that the algorithm run as written will determine the same 1-overlap cover in the same time as the version just described.

5 Heuristics

For trees without common branching rates, heuristics are investigated.

Heuristics for M1C evolve in a straightforward manner from the optimisation algorithm by replacing Step 5 with an appropriate bin-packing heuristic. We distinguish between two types of heuristics developed in this way. Type A heuristics choose Bin 1 to be the least filled bin. Type B heuristics instead remove the subtrees corresponding to *all* bins (including Bin 1). Performance results are only given for Type B heuristics.

Theorem 3 *Consider the 1-Overlap Cover Algorithm applied as a Type B heuristic, with the optimal bin-packing replaced by heuristic H. Let B^H be the number of bins filled by heuristic H and B^* the minimum number required for instance I. If heuristic H satisfies $B^H \leq WB^* + X$, the corresponding 1-overlap cover Type B heuristic has an absolute worst-case performance ratio of at most $2W + X$.*

Proof: Let $m^*(T)$ be the number of components in a minimal 1-overlap cover of tree T and $m^H(T)$ the number of components in a 1-overlap cover constructed by the heuristic. Each iteration (except for the last) of the two versions of the 1-overlap cover algorithm defines two subtrees which are covered separately. For the optimisation algorithm, label these two trees $B^A(T)$ (for the subtree covered immediately) and $N^A(T)$ (for the remaining subtree). Similarly for the heuristic, label the two trees $B^H(T)$ and $N^H(T)$, respectively. Note that $B^A(T)$ is a subtree of $B^H(T)$.

Both algorithms will perform identically until reaching Step 5 for the first time. For a tree satisfying Step 3, $B^A(T) = B^H(T) = T$ and $m^H(T) = m^*(T) = 1$. For a tree satisfying Step 4, $B^A(T) = B^H(T)$ and $m^H(B^H(T)) = m^*(B^A(T)) = 1$. For a tree that has the root satisfying the requirements of Step 5 the optimisation algorithm will pack Bin 1 fully in the next iteration so that $m^H(T) \leq Wm^*(T) + X \leq (2W + X)m^*(T)$.

The hypothesis is that $m^H(T) \leq (2W+X)m^*(T)$. Assume the hypothesis holds for $N^A(T)$; we show that it holds for T.

From the algorithm description we have

$$m^*(T) = m^*(B^A(T)) + m^*(N^A(T)) \text{ and } m^H(T) = m^H(B^H(T)) + m^H(N^H(T)).$$

Step 5 removes more vertices in the heuristic than in the optimisation algorithm. This means that

$$m^*(N^H(T)) \leq m^*(N^A(T)).$$

Also, if the number of bins used by the optimisation algorithm is b then

$$m^*(B^A(T)) = b - 1 \text{ and } m^H(B^H(T)) \leq Wb + X.$$

Putting this together gives

$$\begin{aligned} m^H(T) &= m^H(B^H(T)) + m^H(N^H(T)) \\ &\leq Wb + X + m^H(N^H(T)) \\ &\leq Wm^*(B^A(T)) + W + X + (2W + X)m^*(N^A(T)) \\ &\leq (2W + X)m^*(T), \end{aligned}$$

since $m^*(B^A(T)) \geq 1$.

Applying the heuristic version of the algorithm generates a finite sequence of subtree pairs. The hypothesis applies to the last tree as shown above; induction then implies that the hypothesis is true for the original tree. $\square$

This worst case analysis is likely to be a very weak bound and the question as to whether a similar version of Theorem 3 holds for Type A heuristics, remains open. Also, Type B heuristics are likely to create a number of unnecessarily small subtrees and one would expect Type A heuristics to perform better in general. Indeed, randomised computational testing found that Type A heuristics performed better than Type B heuristics and all performance ratios were less than 1.5. However, the following example shows that there are circumstances where Type A heuristics perform worse than Type B heuristics.

For the example first fit decreasing will be the heuristic used for the bin packing step. It is useful to consider the following two bin packing instances, both with bin size 21.

$$\begin{aligned} B_1 &= \{13, 9, 9, 5, 5, 5, 5, 5, 3, 3, 3, 3\} \\ B_2 &= \{13, 9, 8, 5, 5, 5, 5, 5, 3, 3, 3, 3\} \end{aligned}$$

The only difference in the item sizes is that the third item is smaller in the second list; however, using the first fit decreasing heuristic the first list uses three bins while the second uses four bins.

The relevant tree is constructed as follows. Create a three stage tree as in the proof of Theorem 1 for the list of items B_1. Call the second stage node corresponding to the third item, t. Choose one child of t, c, and connect to c a single child node, s, with 41 child nodes below it. Figure 2 illustrates the tree.

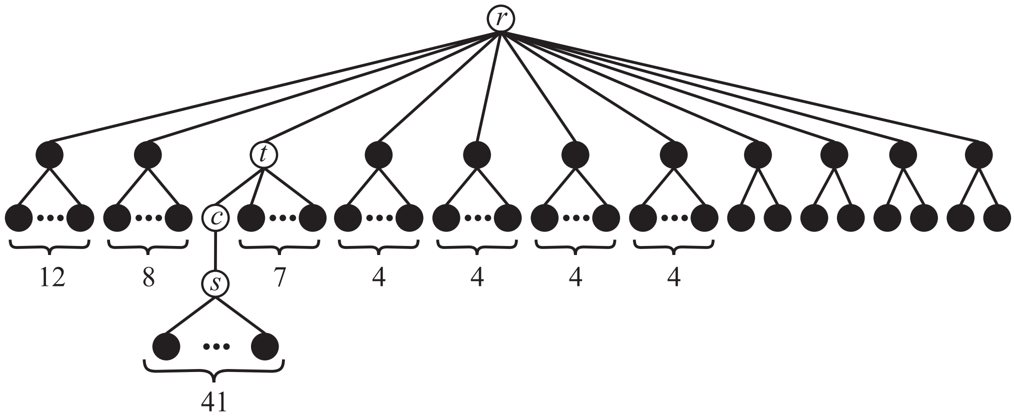


Figure 2: Tree where FFD Type B heuristic outperforms FFD Type A.

The node labels (as constructed in the algorithm) are as follows. The root node, r, will have label $\ell(r) = 115$. All second stage nodes will have a label equal to the corresponding item size expect node t which has label $\ell(t) = 51$. Node c will have label $\ell(c) = 43$ and s will have label $\ell(s) = 42$. All other nodes will have label 1.

With a subtree size of 22 nodes and applying the Type B heuristic, two subtrees are removed below s leaving c but not s. At this point the second stage nodes have labels matching list B_1 which first fit decreasing will pack into three bins, covering the whole tree with five subtrees.

When applying the Type A heuristic with the same subtree size, one subtree is removed below s, then one below c, removing c from the remaining tree. At this point the second stage nodes have labels matching list B_2 which first fit decreasing will pack in four bins, covering the whole tree with six subtrees.

References

[1] I. Bárány, E. Edmonds, and L. A. Wolsey. Packing and covering a tree by subtrees. *Combinitorica*, 6(3):221–233, 1984.

[2] Jitka Dupačová, Giorgio Consigli, and Stein W. Wallace. Scenarios for multistage stochastic programs. *Annals of Operations Research*, 100:25–53, 2000.

[3] Shane Dye. Subtree decomposition for multistage stochastic programs. Working paper, Stochastic Programming e-Print Series, 2003. http://www.speps.info/.

[4] M.R Garey and D.S. Johnson. *Computers and Intractability, a guide to the theory of NP-completeness.* Freeman, New York, 1979.

[5] Sukhamay Kundu and Jayadev Misra. A linear tree partitioning algorithm. *SIAM Journal on Computing*, 6(1):151–154, 1977.

Telecommunication Network Design using Demand Path Enumeration

Gregor Bruce
gregorb@paradise.net.nz

Abstract

Telecommunication networks may be defined in terms of two sets of facts, demands and network resources. A demand is a requirement to transport an amount of information between client systems located at different physical locations. Resources are processes or devices capable of encapsulating and transporting information between physical or logical endpoints. The telecommunication network design is an assignment of resources such that all demands are satisfied.

A goal of this work was to find a solution methodology that was independent of specific telecommunication technologies. This approach is based on the object oriented software paradigm and an abstract object model of a telecommunication network was developed. The root class in the object model implements a binary decision diagram based satisfiability solver. Each object type inherits from this root class and adds rules that reflect the capabilities of a specific resource and its abilities to connect to other types of resources. A collection of object instances, representing a specific problem, can then be created. The problem is then solved by mutual adjustment as the objects interact with each other according to their programmed rules.

1 Introduction

This paper describes a method for designing telecommunication networks at minimum cost. The goal is to implement this in software to automatically produce engineering designs suitable for constructing infrastructure networks spanning a large city, for example Auckland. In effect, this goal determines a level of detail that must be incorporated. A construction specification has to identify all device details including the specific modules required, module installation details, software configuration, the type of cables needed and installation locations. The details for all possible choices of equipment and locations provide the input data for the design.

This data is incorporated into a mathematical network model to calculate the design. A graph based, mathematical model in which the graph's nodes represent telecommunication end-points and edges represent telecommunication links would normally be used for this. The following section introduces an alternative, resource-based network model. It remains a graph based model but uses graph nodes to represent resources and edges to represent connections between resources. The design problem is then formalised as a minimum-cost optimisation using the resource model.

The difficulties of directly solving the formal model are briefly noted and lead into a discussion of the new approach. The information flow, corresponding to a demand, will travel from its origin to destination through a directed sequence, or path of resources. In general there will be multiple alternative paths for any given demand. For example, transmission equipment of different capacities and cost will typically be available as options for each transmission link. So if all possible paths for all demands and all equipment options are enumerated then a selection of paths offering a minimum overall cost will suffice as the network design. This is the essence of the implemented demand path enumeration procedure.

Besides the fine granularity of the model, there is a further dimension to the design problem. The devices to be modelled, particularly the software driven devices, can incorporate complex behaviours that determine the feasibility of demand paths both from a connectivity and performance perspective. Performance is the requirement that the real-time capability to transport demand flows meets latency and rate criteria specified by the demand. The complexity motivated the development of an object-oriented, rules-based software solution. This provides the abstraction necessary to implement the solution procedure while allowing a broad range of simple and complex devices to be modelled. The paper concludes with a brief description of the software package developed by the author to implement the design procedure.

2 Network Design Process

2.1 Informal telecommunication Network Model

Traditional models of telecommunication networks typically identify two distinct types of object; nodes and links. The resource model does not make this distinction. Instead, *links* are treated as multi-port devices; a resource with non zero physical dimensions. Considering that ports in any multi-port device must be physically separated, it is apparent that traditional *links* and *nodes* are distinguished by an arbitrary threshold applied to the distance separating their ports. Although link-type devices still establish the geography of the network, this becomes an implicit property of the link resources and is not explicitly represented in the model topology. In addition to physical resources, there are many intangible or logical resources required to build a network, for example processes such as encapsulation, queueing and routing.

A network if formed by interconnecting resources. A connection or association between two resources identifies that they share a compatible interface and hence are capable of sending and receiving flow. For example, a USB cable will connect to a USB port because they both support the USB physical interface. A more concise understanding of the units of *flow* is also required. These units will typically depend on the type of resource, for example, process resources handling digital data may characterise flow as bits or packets per second, in a network employing optical multiplexing the units of flow at the optical level will be wavelength. This variation is modelled using network resources that transform flow, for example from electrical to optical.

2.2 Mathematical Model

The network, N, is represented by a graph in which nodes correspond to resources, R, and edges, A, correspond to directed associations between resources. The telecommunication demands, W, are a set of origin-destination node tuples: $\{o,d,f\}$ where o and d are resources, f is the required rate of flow.

Each resource will have an associated cost. For example physical equipment will have an installed capital cost. In general the cost for a resource, r, will be a function of the total flow through it F_r, i.e. its capacity; $C_r(F_r)$. Our goal is to find a design that will minimise the total network cost:

$$\sum_{\forall r \in R} C_r(F_r) \qquad (1)$$

subject to the constraints that all demands are satisfied within the capacities of the resources in R. The demands will generate flows along a set of directed paths in N. If x_p is the flow in path p, then

$$F_r = \sum_{\substack{\forall \text{ paths } p \\ \text{containing } r}} x_p$$

and (1) becomes

$$\sum_{\forall r \in R} C_r\Big(\sum_{\substack{\forall \text{ paths } p \\ \text{containing } r}} x_p \Big) \qquad (2)$$

The network design is the aggregate of all resources and associations in the set of selected paths.

2.3 Direct Solution

A Boolean decision variable could be added, along with suitable constraints, to select each path in (2) to formulate the problem as an integer programme (IP). Generally this is not practical as there can be many millions of resources and their cost functions tend to be non-linear. Consequently the resulting IP's become intractably large.

2.4 Path enumeration

As an alternative, a divide-and-conquer approach to the minimum cost problem was explored. Expressing the solution in terms of decisions local to each resource would allow the specific details, the behaviour and capacity, of each type of resource, to be kept local to the resource model's implementation. This would allow flexibility to implement the necessary range of resource types while retaining the ability to achieve optimal designs.

Applying the optimal routing result from Bertsekas and Gallager (1992), if all cost functions are convex, optimal solutions to (2) will have minimum first derivative length. In an optimal network, the demand flows will be routed along paths where the sum of the marginal cost for each resource on the path is a minimum. This suggests it would be possible to obtain an optimal solution by selecting a path, with minimum first derivative length for each demand after determining the marginal cost of all feasible paths through the network. However, real resource cost functions are not convex hence local decisions for each demand cannot assure overall optimality. Also constraints may span feasible paths for multiple requirements. For example, in Figure 1, a constraint that cable c_2 can only connect to one device impacts the path routing for both demands. Hence the design process cannot be fully distributed. It follows that individual resources can make decisions to determine feasible flow paths and routing constraints but the final path selection has to be done as a global optimisation on the whole network.

These ideas are combined in the following optimal path selection procedure.

1. Enumerate all feasible paths through the network for each demand. As a side-effect, all resources are marked with a list of paths in which they appear.
2. Step 1 is repeated. The difference in this second step is that resources are able to make decisions on feasibility and cost given a knowledge of all paths that may be routed through them. At this stage resources will also identify routing constraints.

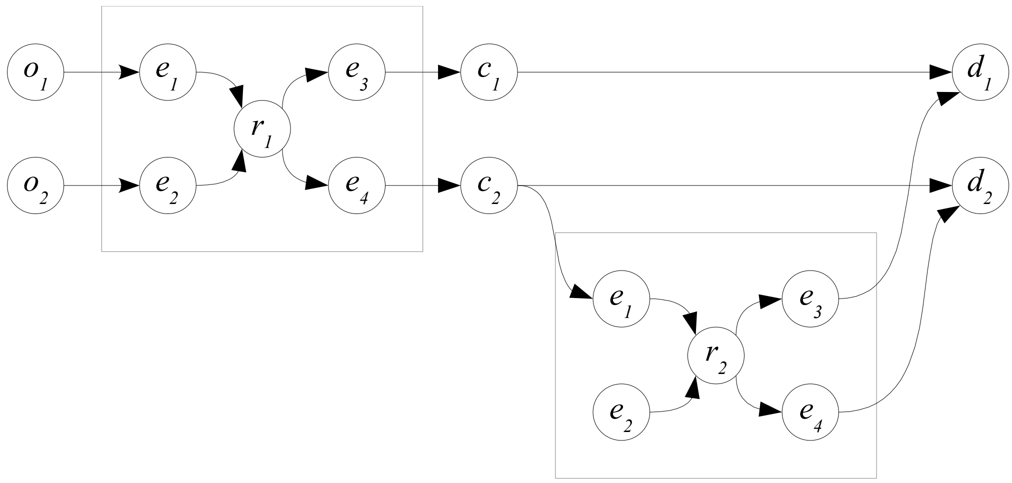


Figure 1. Network example; o_i,d_i are origin-destination demand pairs, r_i, are routers, c_i are cables and e_i, are ethernet interfaces.

3. All feasible paths, including costs and routing constraints, are combined into a minimum cost optimisation to select the design path set.

2.5 Partitioning

It is possible to partition the problem and maintain optimality if paths can be separated into groups with no routing restrictions between paths in separate groups and if cost functions can be assumed to be convex. The common practise of separating the access and backbone network designs is an example of partitioning. It relies on the assumption that the marginal cost of the aggregate flow from the access network into the backbone will be zero for all combinations of access flow and that flows may be independently routed within the two networks.

3 Software Implementation

3.1 Resources

To generate designs for construction, both the physical characteristics and logical behaviour of resources must be modelled. The design needs to specify physical equipment and where it is to be installed, which requires detailed spatial modelling. It also needs to specify settings for all configurable devices to achieve the connectivity and performance required for each demand. Each resource model needs to know all capabilities that affect performance. For example the ethernet connections in Figure 1 are represented as cables. However, there are many characteristics of an ethernet link that determine its performance; the specification for CSMA/CD ethernet (IEEE-802.3 (2002)) is 1,538 pages. There are additional standards depending on various options, such as ethernet bridging. Therefore the method used to model resources needs to model arbitrarily complex behaviour. With this in mind a rule based approach to modelling was chosen. Luger (2002) notes that this approach to representing domain knowledge "*is also one of the most natural and remains widely used*".

3.2 Rules Engine

Initially some prototype code was developed using the Clips rules engine (Clips (2005)). This confirmed the rules based approach is capable of modelling complex telecommunication equipment. However, it suggested there was significant risk that a

Clips solution would not scale up to the number of required resources. Each resource is implemented as its own rules engine so the memory requirements are strongly determined by the requirements of the rules engine. To enable tighter control of memory usage, an alternative binary decision diagram (BDD) based rules engine was developed and integrated into a custom, object-oriented language. The software, called Tnet, is written in C++ and the core application uses Buddy BDD, and Tcl/Tk open source software libraries. The basic idea is that facts are class instances. Classes may inherit from the built-in class *Matching*, as shown in Figure 2, and instances of such classes are unique rules engines, referred to as contexts. Any fact may be asserted in any context. The distinguishing capability inherited from *Matching* is the ability to turn class members into rules by adding a predicate as shown. *Matching* translates predicate expressions into BDD data structures and whenever a fact matching a method's arguments is asserted in the context, the rules engine separately executes that method with each combination of argument values that propogate through the BDD and reach its *true* end node. In the Figure 2 example, all instances of the class *Resource* will gather a set of feasible paths asserted in them. The path objects will be collected in the typed list container *m_feasiblePaths*.

3.3 Performance Requirements

In addition to simple connectivity, for a path to be feasible it must also support the performance specified by its demand. This is a non trivial problem. Flow is a random variable and, in general, random variation of flow is significant. Also, networks will normally incorporate multiple queues. Taking a worst-case deterministic approach to performance evaluation in a queueing network using, for example, the network calculus proposed by Boudec (2001) results in significant under utilisation of network resources most of the time. Effective bandwidth is the most promising analytical alternative. This states that effective bandwidth (or flow magnitude) is a function of space and time parameters, for example see ARTES (2005).

"Parameters s and t are referred to as the space and time parameters respectively. When solving for a specific performance guarantee, these parameters depend not only on the source itself, but on the context on which this source is acting. More specifically, s and t depend on the capacity, buffer size and scheduling policy of the multiplexer, the QoS parameter to be achieved, and the actual traffic mix (i.e. characteristics and number of other sources)."

In other words, effective flow through a resource is a function of the all characteristics of the resource, the flow in question and the aggregate of other flows routed through the resource. Estimating the effective bandwidth of a flow by heuristic methods relevant to the specific network is currently the only practical way to determine if demand performance criteria are satisfied.

3.4 Path Enumeration

The *Matching* root class provides rules that implement demand path enumeration steps 1 and 2. These rules implement a distributed algorithm for calculating all feasible, simple paths between two nodes in a directed graph where cycles are possible. The default behaviour inherited from *Matching* is to mark all demand paths as feasible. However, by overloading inherited methods, specialised resource types can accept or reject path assignments. This is critical to successful path enumeration. As discussed later, resource networks tend to be large so simply enumerating all paths is not generally

```
kclass Resource Matching {
    List<FeasiblePath&> m_feasiblePaths
    method collect-feasible-paths {FeasiblePath& path} {
        predicate {} {
            keval $path m_valid
        }
        kadd $m_feasiblePaths * $ path
    }
}
```

Figure 2. Declaration of a Tnet class to model a Resource. It inherits from the class *Matching* and implements a rule to collect feasible paths (a class defined elsewhere with a property m_valid) as per step 1 of the path enumeration procedure.

possible. However, in practice, the number of feasible paths for each demand tends to be relatively small.

3.5 Path Selection

Step 3 of the design process is implemented using an IP solver. It starts with the set of feasible paths, *P,* collected from demand origins after completing step 2. Then decision variables are introduced to select paths; $z_i = 1$ if path p_i is chosen, 0 otherwise, c_i is the cost of p_i then the objective is to minimise:

$$\sum_{\forall p_i \in P} c_i z_i$$

subject to all demands, *w,* in *W* being satisfied:

$$\sum_{\forall p_i \, implementing \, w} z_i = 1$$

and for each pair of paths *i, j,* containing mutually exclusive associations

$$z_i + z_j = 1 \quad .$$

3.6 Street cabling example

In this section the design procedure is demonstrated by applying it to the simplified example shown in Figure 3. In this example, four new sections require connection to existing telecommunication services accessible in p1 on Main St. This includes the following types of resources; cables c1..c19, pits p1..p3, and ducts u1..u3. There are four demands; *o*1-*d*1, *o*1-*d*2, *o*1-*d*3 and *o*1-*d*4, and four reasonable cable arrangements that will satisfy these demands;

1. each section can be directly cabled from p1 (c1 .. c4),
2. c5 is installed from p1 to p2 and each section is then cabled from p2, (c8 .. c11),
3. c6 is installed from p1 to p3 and each section is then cabled from p3, (c12 .. c15),
4. c7 is installed from p1 to p4 and each section is then cabled from p4, (c16 .. c19),

These are combined to produce the resource and association model shown in Figure 4. Note that in this model there are two resources for each cable which represent each cable end. Applying steps 1 and 2 enumerates four paths per demand. This results in a path selection integer programme with 16 variables.

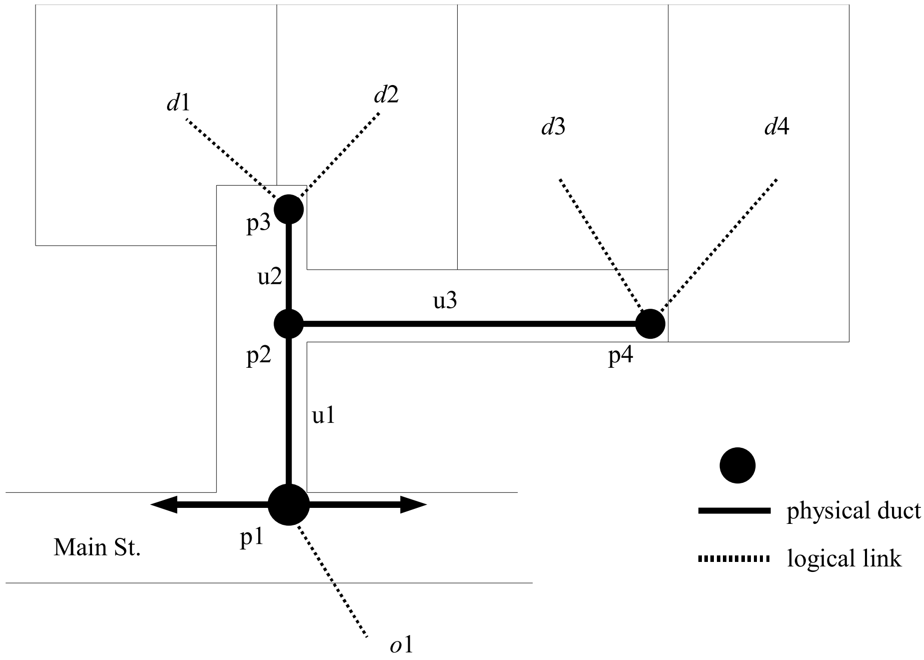


Figure 3. A telecommunications cable example where four new residential sections require connections to an existing cable joint on Main St at p1.

Real problem sizes will typically involve up to 100 demands. They also involve many more resources to accurately model cable cores and joints. However the growth in the corresponding IP is kept roughly linear with respect to demands by applying heuristics to limit the number of enumerated paths. For example, in Figure 3 applying an A* heuristic using geographic length to demands *o*1-*d*1, *o*1-*d*2, paths via p4 can be discarded due to their length relative to cabling via p2 and p3.

This example also illustrates the access-backbone partitioning principle. The cable design is locally optimised on the assumption that the marginal cost of adding the four demands to the network beyond p1 is zero.

3.7 Visualisation

The final design step is to produce design outputs. Typically, these comprise a variety of location diagrams, tabular connection details, equipment configuration files and schematic diagrams. As the model already contains the necessary detail, this step produces data visualisations in appropriate formats. Figure 5, shows Tnet's use of three dimensional spatial modelling to create physical location drawings. It also integrates graph visualisation software to generate schematic diagrams. All output is created in autocad DXF or PDF formats.

4 Application Experience

To date, Tnet's implementation of the resource model has been used to document a broadband network. Path selection optimisation was developed subsequently but was not applied to this network. This broadband network passes approximately 1000 homes (i.e. approx. 1000 demands) and requires a model of approximately 1.5 million resources. This requires approximately 500Mbytes of memory when executed. To create designs for a large city the software will need to scale to around 300,000 demands

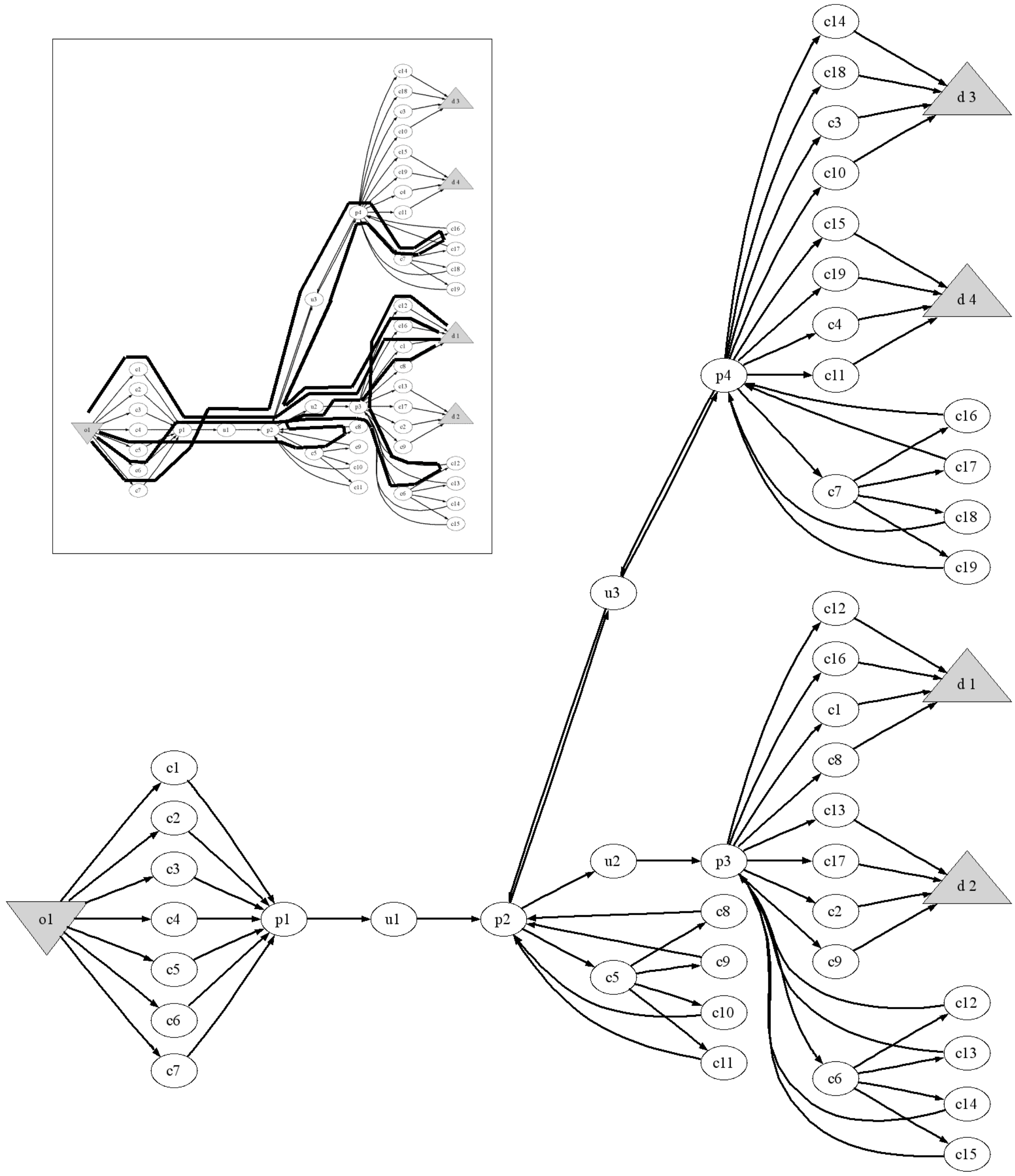


Figure 4. The resource and association model for the Figure 3 example. The insert details the four potential paths for demand o1- d1.

(in the case of Auckland). An expected 300 fold increase in memory would be required which is not practically feasible with the current version of software. An updated version, written in C, is in development to reduce the memory requirements and improve performance. Additional scalability can be achieved by distributing the execution environment across multiple computers.

5 References

ARTES 2005, "Quality of service parameters and link operating point estimation based on effective bandwidths" http://iie.fing.edu.uy/investigacion/grupos/artes/publicaciones/hetnetfinal.pdf (last checked Oct. 2005).

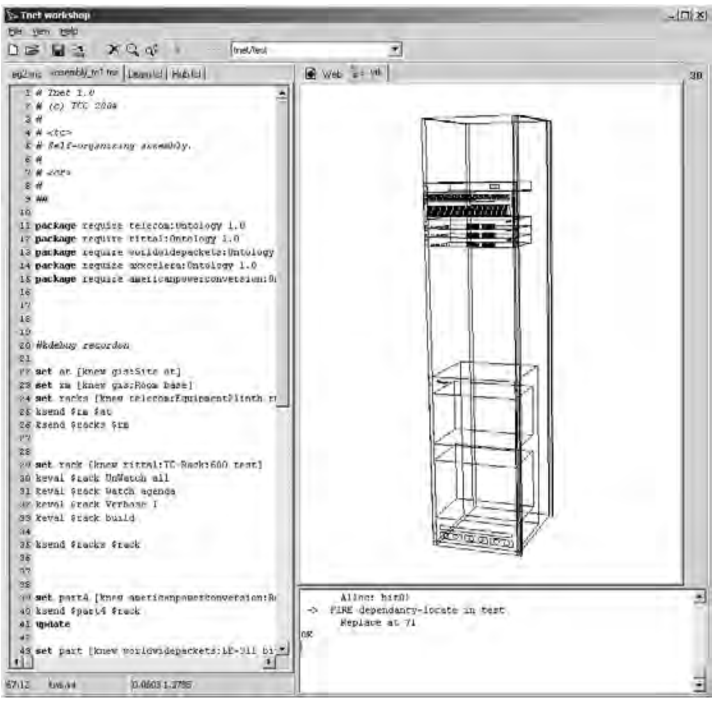


Figure 5. Tnet development environment with example of 3D spatial equipment models.

Bertsekas, D. and R. Gallager. 1992. *Data Networks 2[rd] edition, Prentice-Hall International.*

Boudec, J. and Patrick Thiran. *Network Calculus - A theory of Deterministic Queuing Systems for the Internet. Springer Verlag, 2001*. (can be downloaded from http://lrcwww.epfl.ch/PS_files/NetCal.htm *(last checked Oct. 2005))*

Buddy, "A Binary Decision Diagram library." http://sourceforge.net/projects/buddy. (last checked Oct. 2005)..

Clips "A tool for building expert systems" http://www.ghg.net/clips/CLIPS.html (last checked Oct. 2005).

IEEE-802.3-2002. *IEEE Standard for Information technology Telecommunications and information exchange between systems Local and metropolitan area networks - Specific requirements Part 3*: *Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications.*

Tcl/Tk 2005 http://www.tcl.tk/ (last checked Oct. 2005).

Luger, G.F., 2002. *Artificial Intelligence Structures and Strategies for Complex Problem Solving 4[th] edition, Addison Wesley.*

Determining Knot Location for Regression Splines using Optimisation

Cameron Walker, Michael O'Sullivan & Matthew Gordon
Department of Engineering Science
University of Auckland
New Zealand
cameron.walker@auckland.ac.nz

Abstract

Due to their global nature, polynomial basis functions do not perform well in linear regression when there are abrupt changes in the underlying process or function that we are modelling. In these situations it is better to use regression splines to fit the data. Regression splines can be thought of as a collection of local polynomials that join "smoothly" at knot points. The key decision that needs to be made when fitting regression splines is where to locate the knot points. In this paper we outline and compare one mixed-integer and two non-linear mixed-integer programming formulations for making this decision.

1 Introduction

Polynomial basis functions perform well in regression models when the underlying function that generates the response does not have abrupt changes and is uniformly smooth. However, the flexibility of a polynomial basis is limited due to its global nature. Added flexibility can be incorporated by increasing the degree of the basis, but this can result in a fitted curve that exhibits unrealistic oscillatory behaviour (Silverman, 1985).

Regression splines are made up of a series of piecewise polynomials joining smoothly at breakpoints (knots). Due to the piecewise nature of their construction, regression splines fit locally to the data, allowing a greater degree of flexibility in areas where the underlying function changes abruptly.

Knot specification is a crucial decision when fitting a regression spline, as this determines the extent and location of model flexibility. Typically this decision has been made by eyeballing the data or using local-search heuristic techniques such as forward and backwards regression, although some attempts have also been made to use genetic algorithms. The use of penalised regression splines is viewed by some as a viable alternative to regression splines also. This technique uses a penalty approach which relaxes (to an overstated degree we believe) the need for sensible knot positions. This approach has a global smoothing parameter for the whole surface.

We present here a first attempt at using optimisation techniques to determine the location of knots for regression splines. From hereon we refer to this problem as the Knot Location Problem (KLP).

2 Background

A cubic regression spline is a function with a continuous 2nd derivative, made up of 3rd degree polynomial pieces which meet at knots. It may approximate the smooth function $f(x)$ in $y = f(x) + \varepsilon$ by using:

$$f(x, \underline{\beta}) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \sum_{j=1}^{K} \beta_{3+j} (x - k_j)_+^3 + \varepsilon \tag{1}$$

where $\underline{\beta} = (\beta_0, \ldots, \beta_{3+K})$ is a set of regression coefficients, $(u)_+^3 = u^3 I \quad (u \geq 0)$ or $\mathbf{0}$ $(u < 0)$, and $k_1 < \ldots < k_K$ are fixed knots (MacKenzie, 2004).

If we could place a knot at every x position we could fit any smooth, twice-differentiable function $f(x)$ exactly. Of course this is numerically impossible, so instead we choose a discrete subset of the function domain for knot locations, resulting in what is known as the *Truncated Power Series* form of the regression spline (as given in equation 1).

Figure 1 (overpage) shows a regression spline with knot points at 0.3 and 0.7 fitted to a dataset generated randomly using the following spatially heterogeneous function (Ruppert, 2002):

$$m(x; j) = \sqrt{x(1-x)} \sin\left\{ \frac{2\pi(1 + 2^{(9-4j)/5})}{x + 2^{(9-4j)/5}} \right\}, \ \sigma = 0.3 .$$

The initial cubic ($\beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$) and the terms of the truncated power series are drawn in solid lines, with the resultant regression spline drawn as a dashed line. The coefficients for this fit can be determined with the statictical package R using least squares. The regression output is given below in table 1.

```
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)         0.4023     0.1183   3.399  0.00082 ***
explanatory        -0.9655     2.1375  -0.452  0.65200
I(explanatory^2) -17.8699    10.2021  -1.752  0.08143 .
I(explanatory^3)  42.1220    13.8598   3.039  0.00270 **
C0.3             -76.5642    17.4366  -4.391 1.85e-05 ***
C0.7              99.6748    17.4366   5.716 4.05e-08 ***
---
Multiple R-Squared: 0.4649,     Adjusted R-squared: 0.4511
```

Table 1: Output for regression spline in Truncated Power Series form.

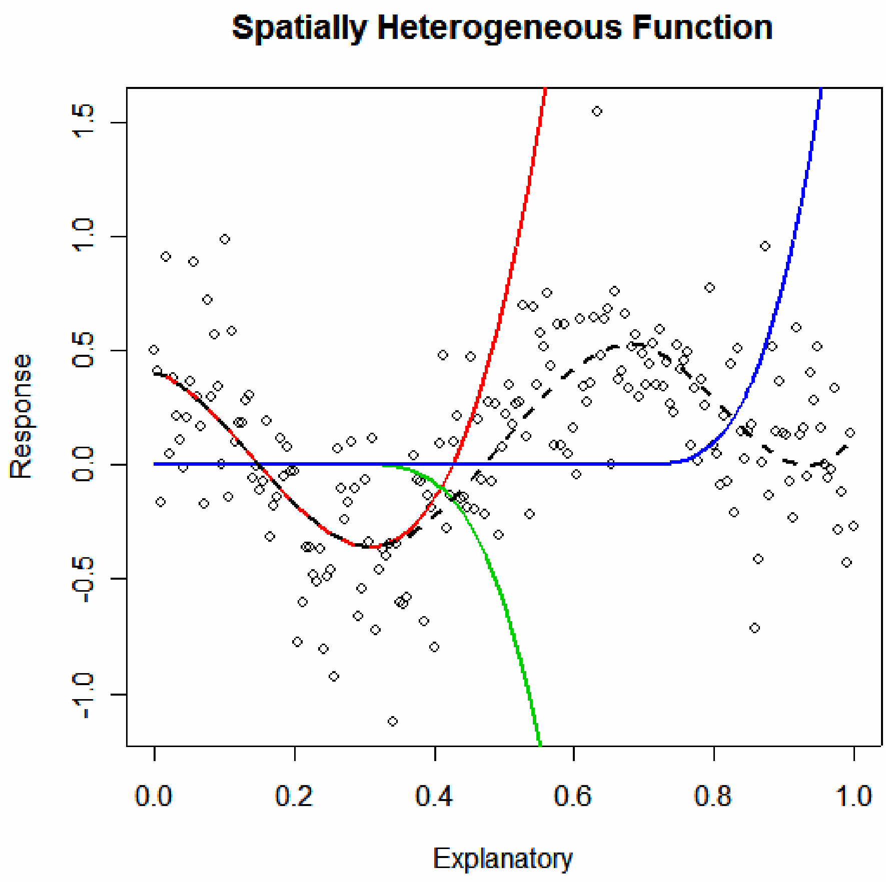


Figure 1: The components of a regression spline in Truncated Power Series form.

In practice the Trucated Power Series form of the regression spline is not used as it is numerically unstable (note in figure 1 how the components of the power series quickly grow very large or very small). Instead regression splines are represented by a tractable additive combination of locally defined low-order polynomials (basis functions). Polynomial B-splines are popular basis splines – they are well-conditioned and yield more stable estimates than the truncated power-series representation. They can be constructed recursively as follows (de Boor, 1987):

$$B_{i,m=1}(x) = \begin{cases} 1 & \text{if } k_i \le x \le k_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

where $i = 1,\ldots, K + 2m - 1$ represents the knot index. Here, m = 1 denotes the order of the *B*-spline. Using this starting basis the following recursion results in bases of any order:

$$B_{i,m}(x) = \frac{x - k_i}{k_{i+m-1} - k_i} B_{i,m-1}(x) + \frac{k_{i+m} - x}{k_{i+m} - k_{i+1}} B_{i+1,m-1}(x)$$

where $i = 1,\ldots, K + 2M - m$. For cubic *B*-splines, $M = 4$, $i = 1,\ldots, K + 4$ and $\mathrm{k} = (k_i)$.

For the example shown in figure 1, the order 1 *B*-splines are shown in figure 2.

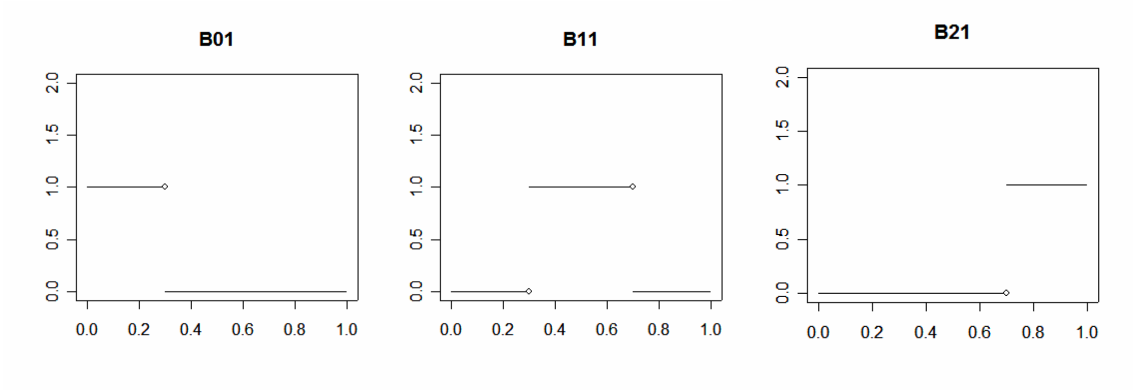


Figure 2: *B*-splines of order 1.

In order to construct higher order *B*-splines we assume that at either end of the domain of interest there are as many knots as required for the recursion. Note that if $k_i = k_{i+1}$ then $B_{i,1} = 0$. Figure 3 illustrates how order 2 *B*-splines are constructed as a weighted average of order 1 *B*-splines.

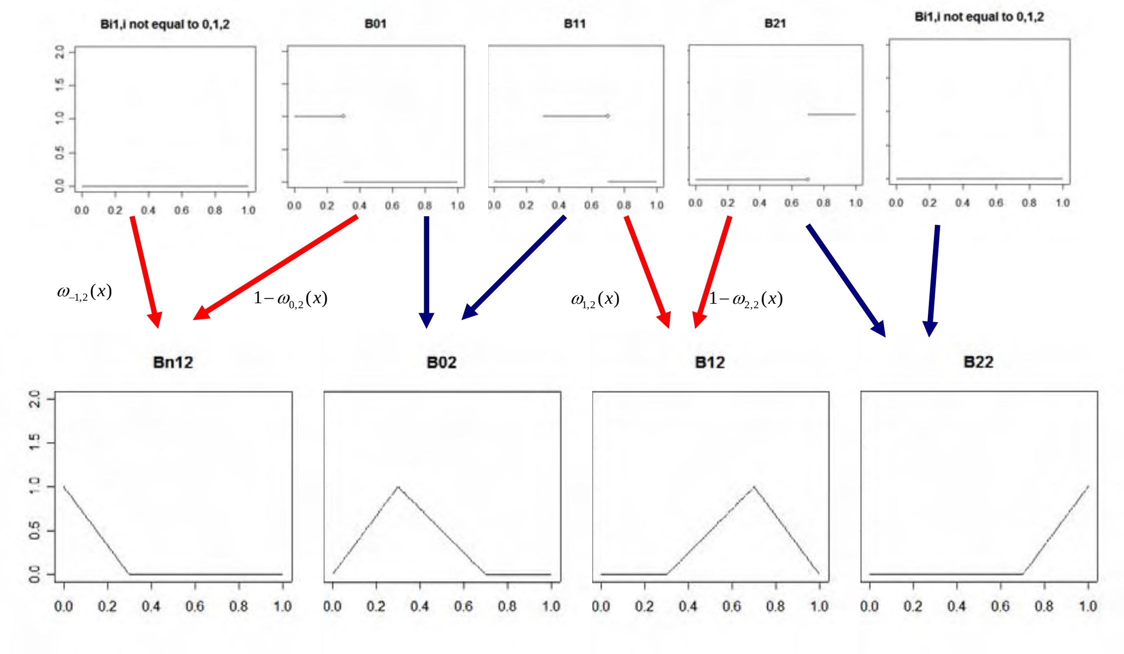


Figure 3: Constructing order 2 *B*-splines from order 1 *B*-splines.

The associated order 3 and 4 *B*-splines are shown in figure 4. We can use the order 4 *B*-splines to fit a regression spline to the data shown in figure 1. Again, the regression coefficients can be determined using least squares, and are given in table 2.

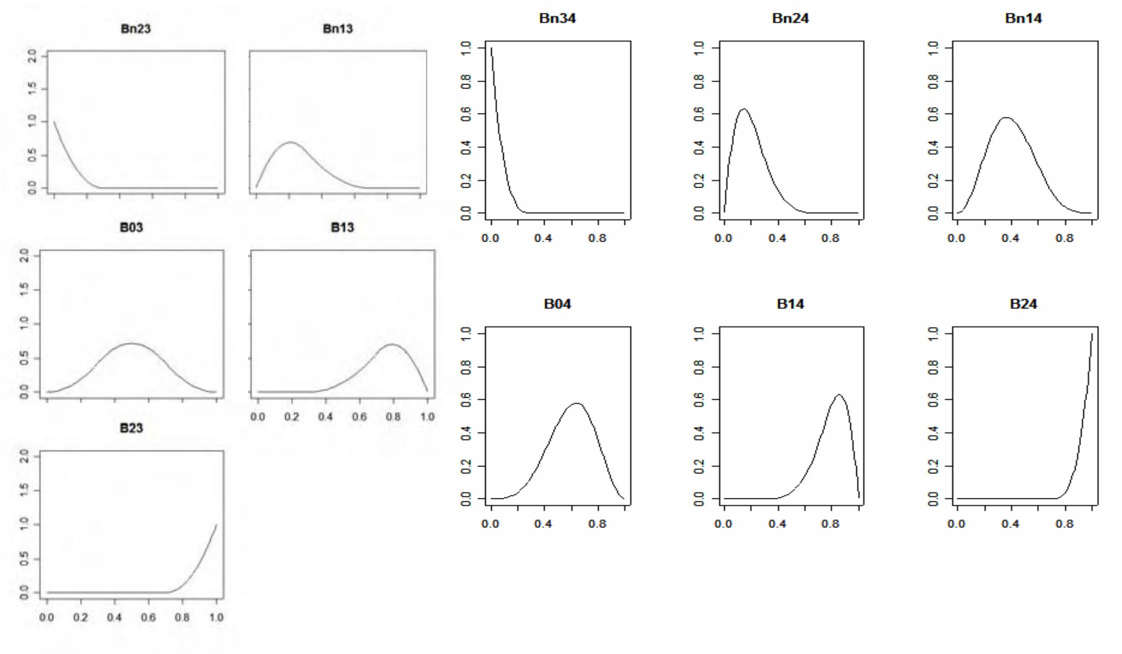


Figure 4: Order 3 and order 4 *B*-splines.

```
Coefficients:
                                         Estimate Std. Error t value Pr(>|t|)
(Intercept)                               0.40225    0.11833   3.399  0.00082 ***
bs(explanatory, degree = 3, knots = knots)1 -0.09655    0.21375  -0.452  0.65200
bs(explanatory, degree = 3, knots = knots)2 -1.57272    0.15127 -10.397  < 2e-16 ***
bs(explanatory, degree = 3, knots = knots)3  0.99446    0.20524   4.845 2.58e-06 ***
bs(explanatory, degree = 3, knots = knots)4 -0.68599    0.16027  -4.280 2.93e-05 ***
bs(explanatory, degree = 3, knots = knots)5 -0.28362    0.17216  -1.647  0.10108
---
Multiple R-Squared: 0.4649,     Adjusted R-squared: 0.4511
```

Table 2: Output for regression spline in *B*-spline form.

The resulting regression spline is the same as that generated using the truncated power series, but the basis components are far more stable (see figure 5).

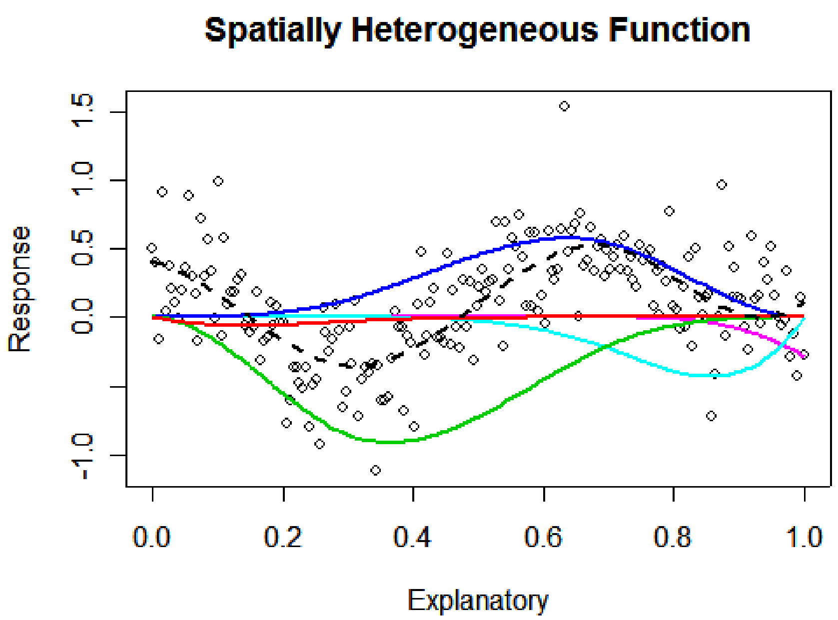


Figure 5: The regression spline (dashed) and its *B*-spline basis components (solid)

3 Optimisation Models

In this section we present three model formulations for determining the knot locations of regression splines. The first is a mixed-integer programme. The last two are non-linear mixed-integer programmes.

3.1 A mixed-integer programming formulation for the Knot Location Problem (MIP)

This initial formulation models the regression spline of degree *P* (order *P* + 1) as a piecewise discontinuous function made up of degree *P* (or less) polynomial sections. This function partitions the domain into *K* parts, each of which is associated with one of the polynomial sections. The points of discontinuity are where we should choose to locate the knots of the subsequent regression spline. In order to maintain linearity we use a minimax objective function, minimising the magnitude of the largest absolute residual.

The model decision variables are as follows:

- the coefficient for the p^{th} term of the k^{th} polynomial segment is a real variable denoted $c_{p,k}$;
- the largest magnitude of the residuals for the fitted values of the k^{th} polynomial is a non-negative real variable e_k;
- the largest magnitude of the residuals across all of the polynomials is a non-negative real variable E;
- if the t^{th} data point is fitted by the k^{th} polynomial then the binary variable $f_{t,k}, f_{t,k+1}, \ldots, f_{t,K}$, are equal to 1. For all other possible *j*, $f_{t,j}$ is equal to 0.

The model also has the following parameters:

- the t^{th} value of the explanatory variable is x_t;
- the t^{th} value of the response variable is y_t.

The maximum magnitude of the residuals for the first polynomial component of our model is computed by constraints MIP1a and MIP1b. The maximum residuals for the other components are computed by constraints MIP2a and MIP2b. The constant *M* is a big-M quantity, ensuring that the constraint is void if point *t* is not fitted by polynomial component *k*.

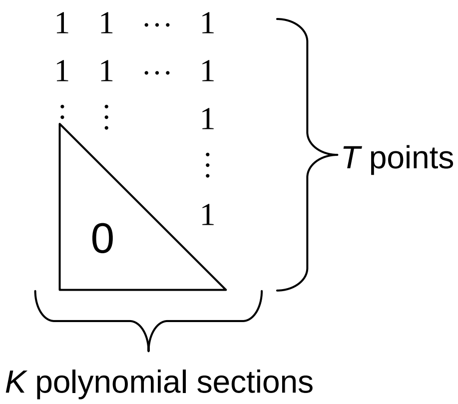
$$\left(\sum_{i=0}^{p} c_{i,1} \times x_t^i\right) - y_t \le \mathrm{e}_1 + (1 - f_{t,1}) \times M \qquad \forall t \qquad \text{(MIP1a)}$$

$$y_t - \left(\sum_{i=0}^{p} c_{i,1} \times x_t^i\right) \le \mathrm{e}_1 + (1 - f_{t,1}) \times M \qquad \forall t \qquad \text{(MIP1b)}$$

$$\left(\sum_{i=0}^{p} c_{i,k} \times x_t^i\right) - y_t \le e_k + (1 - f_{t,k} + f_{t,k-1}) \times M \quad \forall t\ ,\ k > 1 \qquad \text{(MIP2a)}$$

$$y_t - \left(\sum_{i=0}^{p} c_{i,k} \times x_t^i\right) \le e_k + (1 - f_{t,k} + f_{t,k-1}) \times M \quad \forall t\ ,\ k > 1 \qquad \text{(MIP2b)}$$

To ensure our model partitions the data points correctly, we need to put conditions on the binary variables *f*. If we consider them as a *T* by *K* binary matrix *B* with the variable $f_{t,k}$ in position b_{tk}, we force *B* to have the following stepwise structure:



This is done with constraints MIP3a, MIP3b and MIP3c.

$$f_{t,k} \le f_{t,k+1} \qquad \forall t < T \,,\; k < K \qquad \text{(MIP3a)}$$

$$f_{t,k} \ge f_{t+1,k} \qquad \forall t < T \,,\; k < K \qquad \text{(MIP3b)}$$

$$f_{1,1} = f_{T,n} = 1 \qquad \text{(MIP3c)}$$

We also need to fit every data point. This is done with constraint MIP4.

$$\sum_{k=1}^{n} f_{t,k} \ge 1 \qquad \forall t \qquad \text{(MIP4)}$$

Finally, our objective function is to minimise the biggest residual *E*. This quantity is computed from the biggest residuals for each polynomial section using constraint MIP5.

$$e_k \le E \qquad \forall k \qquad \text{(MIP5)}$$

3.2 A non-linear mixed-integer programming formulation for the Knot Location Problem (NLMIP1)

The minimax objective function used in the previous formulation is necessary to maintain linearity. If we allow non-linearity in our formulation we can look to minimise the total sums of squares, as is done in regression. In our first non-linear mixed-integer programme we have the following objective function:

$$\text{Minimize } \sum_{t=1}^{T} e_t^2 \,,$$

where e_t is the difference between the value fitted by the model and the value of the response at x_t.

As our model is now non-linear, we can also dispense with the big-M quantity. Thus we replace constraints MIP1a to MIP2b with NLMIP1. Note also that our new objective makes constraint MIP5 unneccessary, and negative values are now possible for the real variable *e*.

$$\sum_{i=0}^{p} f_{t,1} \times c_{i,1} \times x_t^i + \left(\sum_{k=2}^{n} \sum_{i=0}^{p} \left(f_{t,k} - f_{t,k-1} + \right) c_{i,k} \times x_t^i \right) - y_t = e_t \qquad \text{(NLMIP1)}$$

All other constraints (MIP3a to MIP4) are kept in the model

3.3 A second non-linear mixed-integer programming formulation for the Knot Location Problem (NLMIP2)

The final formulation we present here models the regression spline directly. This is done by including real variables *l* for the knot locations, and including constraints to force the polynomial sections to meet at the knots with the necessary smoothness.

The knot locations are decided using constraints NLMIP2a to NLMIP2c.

$$l_1 \geq x_t - M \times (1 - f_{t,1}) \qquad \forall t \qquad \text{(NLMIP2a)}$$

$$l_i \leq x_t - M \times (1 - f_{t,i+1} + f_{t,i}) \qquad \forall t \qquad \text{(NLMIP2b)}$$

$$l_i \geq x_t - M \times (1 - f_{t,i} + f_{t,i-1}) \qquad \forall t \qquad \text{(NLMIP2c)}$$

If we denote the j^{th} derivative of the k^{th} polynomial section of our fitted model by P_k^j, we can force these sections to join smoothly with constraint NLMIP3.

$$P_k^j(l_k) = P_{k+1}^j(l_k) \qquad j = 1, \ldots, p-1,\ k = 1, \ldots, K-1 \qquad \text{(NLMIP3)}$$

All constraints from formulation NLMIP1 are included in formulation NLMIP2.

3.4 The full model formulations

(MIP) minimise E

s.t.

$$\left(\sum_{i=0}^{p} c_{i,1} \times x_t^i \right) - y_t \leq \mathrm{e}_1 + (1 - f_{t,1}) \times M \qquad \forall t \qquad \text{(MIP1a)}$$

$$y_t - \left(\sum_{i=0}^{p} c_{i,1} \times x_t^i \right) \leq \mathrm{e}_1 + (1 - f_{t,1}) \times M \qquad \forall t \qquad \text{(MIP1b)}$$

$$\left(\sum_{i=0}^{p} c_{i,k} \times x_t^i \right) - y_t \leq e_k + (1 - f_{t,k} + f_{t,k-1}) \times M \quad \forall t\ ,\ k > 1 \qquad \text{(MIP2a)}$$

$$y_t - \left(\sum_{i=0}^{p} c_{i,k} \times x_t^i \right) \leq e_k + (1 - f_{t,k} + f_{t,k-1}) \times M \quad \forall t\ ,\ k > 1 \qquad \text{(MIP2b)}$$

$$f_{t,k} \leq f_{t,k+1} \qquad \forall t < T\ ,\ k < K \qquad \text{(MIP3a)}$$

$$f_{t,k} \geq f_{t+1,k} \qquad \forall t < T\ ,\ k < K \qquad \text{(MIP3b)}$$

$$f_{1,1} = f_{T,n} = 1 \qquad \text{(MIP3c)}$$

$$\sum_{k=1}^{n} f_{t,k} \geq 1 \qquad \forall t \qquad \text{(MIP4)}$$

$$e_k \leq E \qquad \forall k \qquad \text{(MIP5)}$$

(NLMIP1) minimise $\sum_{t=1}^{T} e_t^2$

$$\text{s.t:} \quad \sum_{i=0}^{p} f_{t,1} \times c_{i,1} \times x_t^i + \left(\sum_{k=2}^{n} \sum_{i=0}^{p} \left(f_{t,k} - f_{t,k-1} + \right) c_{i,k} \times x_t^i \right) - y_t = e_t \qquad \text{(NLMIP1)}$$

$$f_{t,k} \le f_{t,k+1} \qquad \forall t < T \ , \ k < K \qquad \text{(MIP3a)}$$

$$f_{t,k} \ge f_{t+1,k} \qquad \forall t < T \ , \ k < K \qquad \text{(MIP3b)}$$

$$f_{1,1} = f_{T,n} = 1 \qquad \text{(MIP3c)}$$

$$\sum_{k=1}^{n} f_{t,k} \ge 1 \qquad \forall t \qquad \text{(MIP4)}$$

(NLMIP2) minimise $\sum_{t=1}^{T} e_t^2$

$$\text{s.t.:} \quad \sum_{i=0}^{p} f_{t,1} \times c_{i,1} \times x_t^i + \left(\sum_{k=2}^{n} \sum_{i=0}^{p} \left(f_{t,k} - f_{t,k-1} + \right) c_{i,k} \times x_t^i \right) - y_t = e_t \qquad \text{(NLMIP1)}$$

$$f_{t,k} \le f_{t,k+1} \qquad \forall t < T \ , \ k < K \qquad \text{(MIP3a)}$$

$$f_{t,k} \ge f_{t+1,k} \qquad \forall t < T \ , \ k < K \qquad \text{(MIP3b)}$$

$$f_{1,1} = f_{T,n} = 1 \qquad \text{(MIP3c)}$$

$$\sum_{k=1}^{n} f_{t,k} \ge 1 \qquad \forall t \qquad \text{(MIP4)}$$

$$l_1 \ge x_t - M \times (1 - f_{t,1}) \qquad \forall t \qquad \text{(NLMIP2a)}$$

$$l_i \le x_t - M \times (1 - f_{t,i+1} + f_{t,i}) \qquad \forall t \qquad \text{(NLMIP2b)}$$

$$l_i \ge x_t - M \times (1 - f_{t,i} + f_{t,i-1}) \qquad \forall t \qquad \text{(NLMIP2c)}$$

$$P_k^j(l_k) = P_{k+1}^j(l_k) \qquad j = 1, \ldots, p-1, \ k = 1, \ldots, K-1 \qquad \text{(NLMIP3)}$$

4 Results

We built models of formulations MIP, NLMIP1, and NLMIP2 in AMPL for the data shown in figure 1. We solved these models for 2 knots with polynomial sections of degrees 1,2 and 3 using CPLEX 9.0 (for MIP) and the NEOS server (for NLMIP1 and NLMIP2). The R^2 for the corresponding (degree 3) regression splines are shown in table 1. Fitting a regression with polynomial basis functions up to degree 3 gives an R^2 of 0.3733.

	Linear	Quadratic	Cubic
MIP	0.4804	0.391	0.3962
NLMIP1	0.4795	0.491	0.3927
NLMIP2	0.4747	0.3735	0.3735

Table 1: R^2 of regression splines.

It is of interest to note that for all three formulations the linear polynomial sections have determined better knot locations than the cubic polynomial sections. For the MIP formulation this is probably because piecewise discontinuous cubic (and even quadratic) sections allow too much flexibility, resulting in a fit which models too much of the randomness (noise) relative to the underlying function (signal). The same appears to be true for the NLMIP1 formulation.

For the NLMIP2 formulation the cubic regression spline has been modelled directly. This means that the (suboptimal) knot locations chosen by the model must be a local optimum. As a first step towards improving this solution we also developed an exchange algorithm to identify a good starting solution for these (and any subsequent) non-linear formulations. With two knots the locations chosen by this heuristic algorithm resulted in a regression spline with R^2 equal to 0.4588. The details of this algorithm will be the subject of a subsequent paper.

A second aspect worth mentioning is the number of knots for which these formulations can solve. The MIP formulation was run for 1 to 17 knots, and solved for all these instances of the problem overnight (the 2 knot case solved in less than 15 seconds for the linear, quadratic, and cubic formulations). The NLMIP1 formulation performed similarly. The branch and bound tree for the NLMIP2 formulation grew very quickly, however, and we could not solve for more than 10 knots in a reasonable timeframe. The heuristic had no such limitations (we can solve 300 randomly generated instances of the problem described in this paper for 1 to 8 knots in less than an hour).

5 Future work

We are currently making improvements to our exchange heuristic, which thus far seems to be the most likely tool to be used for determining knot locations. We will also be investigating the objective function of NLMIP2 more closely in order to determine search strategies for identifying the global optimum. We also intend to apply genetic algorithms to the problem.

Of course the most interesting next step is to investigate the KLP in higher dimensions.

Acknowledgements

We would like to thank Monique MacKenzie for introducing this problem to us and for assistance with fitting regression splines in R.

6 References

de Boor, C. 1987. “B-form basics”. In G. Farin, editor, *Geometric Modeling: Algorithms and New Trends. SIAM.* 131-148.

MacKenzie, M. L. 2004. “Flexible mixed models: Regression Splines and Thin-Plate Regression Splines in a Mixed Model Framework.” *University of Auckland Doctoral Thesis*

Ruppert, D. 2002. “Selecting the number of knots for penalized splines.” *Journal of computational and graphical statistics* **11**, 735-757.

Silverman, B. W. 1985. “Some aspects of the spline smoothing approach to non-parametric regression curve fitting.” *Journal of the Royal Statistical Society, Series B, Methodological* **47**, 1-21.

Some New Sufficient Conditions for Membership in Pedigree Polytope

Tiru S. Arthanari
Department of Information Systems & Operations Management
University of Auckland, Auckland
New Zealand
t.arthanari@auckland.ac.nz

Abstract

The symmetric traveling salesman problem (STSP) polytope is generally studied as embedded in the standard Subtour elimination polytope, SEP_n. The combinatorial objects called *pedigrees* are in 1-1 correspondence with symmetric traveling salesman tours on n cities. And the convex hull of the characteristic vectors of the pedigrees yields the pedigree polytope. The pedigree polytope is embedded in the *MI-relaxation* polytope, arising from the multistage insertion formulation (MI-formulation) of the STSP. The pedigree polytopes are combinatorial polytopes in the sense of Naddef and Pulleyblank [10]. In the context of STSP polytope, testing whether two given tours are non-adjacent is a NP-complete problem. However, non-adjacency testing can be done in polynomial time in the pedigree polytope. In this paper a layered network is recursively constructed and used to provide some new sufficient conditions for membership in the pedigree polytope. Illustrative examples are provided to explain the concepts introduced.

1 Introduction

Since Khachian [7] showed that the linear programming problem can be solved in polynomial time, researchers were attempting efficient use of linear programming formulations of $0-1$ integer programming problems [5]. They necessitate the study of polytopes whose vertices are $0-1$ vectors. Such polytopes are called $0-1$ polytopes. Problems involving graphs formulated as combinatorial optimization problems (COP) also lead to the study of such polytopes. State-of-the-art concepts, theoretical results and algorithms for COP are presented in a recent book by Korte and Vygen [8]. Symmetric traveling salesman problem (STSP) is one such COP problem. STSP has the standard $0-1$ formulation given by Dantzig, Fulkerson, and Johnson as early as 1954. This formulation has exponentially many constraints that eliminate sub-tours from consideration. The corresponding relaxation of the formulation defines the subtour elimination polytope, SEP polytope [4].

Arthanari and Usha [3] study the properties of a $0-1$ linear programming formulation of the symmetric traveling salesman problem (STSP), called the multistage insertion formulation (MI- formulation) which has polynomially many constraints and variables. They also show that the set of slack variable vectors corresponding to feasible solutions to the MI-relaxation is contained in the SEP-polytope. Arthanari [1] considers integer feasible solutions of the MI- formulation and defines a combinatorial object called *pedigree* and study the properties of the corresponding polytope. Flow problems are defined for a given solution for the MI-relaxation, X, for the purpose of obtaining necessary and sufficient conditions for X to be in the pedigree polytope. A polynomial time nonadjacency testing algorithm for pedigree polytopes is also developed there[1]. In [2] we observe that the pedigree polytope is a combinatorial polytope, in the sense that, given any two nonadjacent vertices of the polytope, the mid point of the line joining these vertices is also the mid point of two other vertices of the polytope. [See [2] for the proof and consequences of this property and other properties of the pedigree polytopes. Research by the author on pedigree polytopes can be found at [12].]

In this paper we prove additional properties of the pedigree polytope to obtain sufficient conditions for a solution to the MI-relaxation to be in the pedigree polytope. Section 2 introduces the notations and defines pedigrees and provides some results on pedigrees from earlier related papers. Section 3 introduces a layered network used in defining a multi-commodity flow problem. Section 4 also defines pedigree packability and states results on pedigree packability of flows in some restricted networks defined. Section 5 briefly indicates the future directions and research in proving the complexity of the computations involved.

2 Preliminaries & Notations

This section gives a short review of the definitions and concepts used in the paper.

Let n be an integer, $n \geq 3$. Let V_n be a set of *vertices*. Assuming, without loss of generality, that the vertices are numbered in some fixed order, we write $V_n = \{1, \ldots, n\}$. Let $E_n = \{(i,j)|i,j \in V_n, i < j\}$ be the set of *edges*. The cardinality of E_n is denoted by $p_n = n(n-1)/2$. We assume that the edges in E_n are ordered in increasing order of the edge label, $l_{ij} = p_{j-1} + i$, for the edge (i,j). Let $K_n = (V_n, E_n)$ denote the complete graph of n vertices. We denote the elements of E_n by e where $e = (i,j)$. We also use the notation ij for (i,j). Notice that, unlike the usual practice, an edge is assumed to be written with $i < j$. For a subset $F \subset E_n$ we write the *characteristic/ incidence* vector of F by $x_F \in R^{p_n}$ where $x_F(e) = 1$ if $e \in F$, and 0 otherwise.

For a subset $S \subset V_n$ we write $E(S) = \{ij|ij \in E, i,j \in S\}$. Given $u \in R^{p_n}$, $F \subset E_n$, we define, $u(F) = \sum_{e \in F} u(e)$. For any subset S of V_n, let $\delta(S)$ denote the set of edges in E_n with one end in S and the other in $S^c = V_n \setminus S$. For $S = \{i\}$, we write $\delta(\{i\}) = \delta(i)$. A subset H of E_n is called a *Hamiltonian cycle* in K_n if it is the edge set of a simple cycle in K_n, of length n. We also call such a Hamiltonian cycle a $n - tour$ in K_n.

[1]Papadimitriou [11] has shown that the corresponding problem of nonadjacency testing for $STSP$ polytope is NP-complete.

Consider a balanced transportation problem, in which, some arcs called the *forbidden* arcs are not available for transportation. We call the problem of finding whether a feasible flow exists in such an incomplete bipartite network, a Forbidden Arcs Transportation (FAT) problem[9]. Bipartite network flow problems are of special interest among maximal flow problems.

2.1 Definition of the Pedigree Polytope

In this sub section we present some definitions and concepts from [1] and an alternative polyhedral representation of the STSP, using the definition of pedigrees.

Let Q_n denote the standard $STSP$ polytope, given by the convex hull of the characteristic vectors of Hamiltonian cycles or $n-tours$ in K_n.

Given $H \in \mathcal{H}_{k-1}$, the operation *insertion* is defined as follows: Let $e = (i, j) \in H$. Inserting k in e is equivalent to replacing e in H by $\{(i, k), (j, k)\}$ obtaining a $k-tour$. When we denote H as a subset of E_{k-1}, then inserting k in e gives us a $H' \in \mathcal{H}_k$ such that,$H' = (H \cup \{(i, k), (j, k)\}) \setminus \{e\}$.

Definition 2.1 *[Pedigree] The vector $W = (e_4, \ldots, e_n) \in E_3 \times \ldots \times E_{n-1}$ is called a* pedigree *if and only if there exists a $H \in \mathcal{H}_n$ such that H is obtained from the $3-tour$ by the sequence of insertions, viz., 4 is inserted in $e_4 \in E_3$ and 5 is inserted in $e_5 \in E_4$ and so on until we insert n in $e_n \in E_{n-1}$ and obtain a $n-tour$, H.*

The pedigree W is referred to as the pedigree of H. Let the set of all pedigrees, corresponding to $H \in \mathcal{H}_n$ be denoted by $\mathcal{P}_n$. For any $4 \le k \le n$, given an edge $e \in E_{k-1}$, with edge label l, we can associate a $0-1$ vector, $\mathbf{x}(e) \in B^{p_{k-1}}$, such that, $\mathbf{x}(e)$ has a 1 in the l^{th} coordinate, and zeros else where. That is, $\mathbf{x}(e)$ is the indicator of e.

Let $\tau_n = \sum_{j=3}^{n-1} p_j$. We can associate a $X = (\mathbf{x}_4, \ldots, \mathbf{x}_n) \in B^{\tau_n}$, the characteristic vector of the pedigree W, where $(W)_k = e_k$, the $(k-3)^{rd}$ component of W, $4 \le k \le n$ and $\mathbf{x}_k$ is the indicator of e_k.

We write, $P_n = \{X \in B^{\tau_n}\ : X$ is the characteristic vector the pedigree $W \in \mathcal{P}_n\}$.

Consider the convex hull of P_n. We call this the *pedigree polytope*, denoted by $conv(P_n)$.

An interesting property of $X = (\mathbf{x}_4, \ldots, \mathbf{x}_n) \in P_n$ is that, for any $4 \le k \le n$, X restricted to the first $k-3$ stages, written as $X/k = (\mathbf{x}_4, \ldots, \mathbf{x}_k)$ is in P_k.

Similarly, $X/k-1$ and $X/k+1$ are interpreted as restrictions of X. We use this notation for any $X \in R^{\tau_n}$ as well. For examples of pedigrees and connections to $STSP$ polytope see [1].

Definition 2.2 *Given $e = (i, j) \in E_n$, we call*

$$G(e) = \begin{cases} \delta(i) \cap E_{j-1} & \text{if } j \ge 4, \\ E_3 \setminus \{e\} & \text{otherwise}, \end{cases}$$

the set of generators *of the edge e.*

Since an edge $e = (i, j), j > 3$ is generated by inserting j in any e' in the set $G(e)$, the name *generator* is used to denote any such edge.

Example: 2.1 *Consider* $n = 5, e = (3,5)$. *Here* $j \geq 4$, *so, we have* $G(e) = \delta(i) \cap E_{j-1}$. *Since* $i = 3$, $\delta(3) = \{(1,3),(2,3,)(3,4),(3,5)\}$, *and* $E_{j-1} = E_4 = \{(1,2),(1,3),(2,3),(1,4),(2,4),(3,4)\}$. *Therefore,* $G(e) = \{(1,3),(2,3,)(3,4)\}$. *Consider* $e = (1,3)$. *Since* $j \leq 3$, *we have* $G(e) = E_3 \setminus \{e\} = \{(1,2),(2,3)\}$.

Definition 2.3 *[Extension of a Pedigree] Let* $y(e)$ *be the indicator of* $e \in E_k$. *Given a pedigree,* $W = (e_4,\ldots,e_k)$ *(with the characteristic vector,* $X \in P_k$*) and an edge* $e \in E_k$, *we call* $(W,e) = (e_4,\ldots,e_k,e)$ *an* extension *of* W *in case* $(X, y(e)) \in P_{k+1}$.

Observe that given W a pedigree in $\mathcal{P}_k$ and an edge $e = (i,j) \in E_k$, (W,e) is a pedigree in $\mathcal{P}_{k+1}$ if and only if 1] $e_l \neq e, 4 \leq l \leq k$ and 2] there exists a $q = max(4,j)$ such that e_q is a generator of $e = (i,j)$.

Definition 2.4 *[*$P_{MI}(n)$ *Polytope] The multistage insertion formulation of the* $STSP$ *problem is called the* MI*- formulation, as given in [3]. With the* $0-1$ *integer constraints relaxed, it is called the* MI*- relaxation. The corresponding polytope consisting of the feasible solutions to MI- relaxation, is called the* $P_{MI}(n)$ *polytope, where* n *refers to the number of cities.*

All X in P_n are extreme points of $P_{MI}(n)$ polytope. But $P_{MI}(n)$ has fractional extreme points as well. Thus the pedigree polytope $conv(P_n)$ is contained in $P_{MI}(n)$ polytope.

Definition 2.5 *[Weight Vector] Given* $X \in P_{MI}(n)$ *and* $X/k \in conv(P_k)$, *consider* $\lambda \in R_+^{|P_k|}$ *that can be used as a weight to express* X/k *as a convex combination of* $X^r \in P_k$. *Let* $I(\lambda)$ *denote the index set of positive coordinates of* λ. *Let* $\Lambda_k(X)$ *denote the set of all possible* weight vectors, *for a given* X *and* k.

Definition 2.6 *[Active Pedigree] Given* $X \in conv(P_k)$, *we call a* $X^* \in P_k$ active *for* X/k, *in case there exists a* $\lambda \in \Lambda_k(X)$ *and an* $r \in I(\lambda)$ *such that* $X^* = X^r$. *In other words* X^* *receives positive weight in at least one convex combination expressing* X.

3 Construction of the Layered Network N_k

In this section we define the layered network $N_k(X)$ with respect to a given $X \in P_{MI}(n)$ and for $k \in V_{n-1} \setminus V_3$. Given $X/k \in conv(P_k)$, this network is used in showing whether $X/k+1 \in conv(P_{k+1})$ or not. Since X is fixed throughout this discussion we drop the X from the notation for the network and write simply N_k.

We define for each k, a layered network, N_k, with $(k-2)$ layers. We denote the node set of N_k by $\mathcal{V}(N_k)$ and the arc set by $\mathcal{A}(N_k)$. Let $v = [k : e]$ denote a node in the $(k-3)^{rd}$ layer corresponding to an edge $e \in E_{k-1}$ in the layered network, N_k. Let $x(v)$ refer to $x_k(e)$, the component of $\mathbf{x_k}$, corresponding to an edge $e \in E_{k-1}$.

Let

$$V_{[r]} = \{v | v = [r+3 : e], e \in E_{r+2}, x(v) > 0\}.$$

Notice that the node $[r+3 : e]$ refers to the stage r and the decision to insert $r+3$ in edge e at that stage.

First we define the nodes in the network N_k, for $k = 4$.

$$\mathcal{V}(N_4) = V_{[1]} \bigcup V_{[2]}, \text{ and } \mathcal{A}(N_4) = \{(u, v) | u \in V_{[1]}, v \in V_{[2]}, e_\alpha \in G(e_\beta)\}$$

where $u = [4 : e_\alpha]$ and $v = [5 : e_\beta]$.

Capacity on a node $v \in V_{[r]}$ is $x(v), r = 1, 2$. Capacity on an arc $(u, v) \in \mathcal{A}(N_4)$ is $x(u)$. Given this network we consider a flow feasibility problem of finding a nonnegative flow defined on the arcs that saturates all the node capacities and violates no arc capacity. We refer to this problem F_4. Notice that the problem F_4 is one and the same as the problem $FAT_4(\mathbf{x}_4)$ defined in [1]. Therefore, F_4 feasibility is equivalent to $FAT_4(\mathbf{x}_4)$ feasibility. So $X/5 \in conv(P_5)$. If F_4 is infeasible we do not proceed further. (We conclude $X/5 \notin conv(P_5)$ and so $X \notin conv(P_n)$ as shown in [1]).

If F_4 is feasible we use an algorithm (such as Gusfield's [6]) and identify the set of arcs with fixed flow in all feasible solutions to the flow feasibility problem considered. We call these arcs, as rigid arcs and denoted the set by $\mathcal{R}$. If there are arcs with zero flow, called dummy arcs, in $\mathcal{R}$ we delete them from $\mathcal{A}(N_4)$ and update $\mathcal{A}(N_4)$. For rigid arc with positive frozen flow we freeze the flow along the arc and colour the arc *'green'*.

Now we say N_4 is *well-defined*.

In general, we say N_{k-1} is well-defined if *we have shown* that $X/k \in conv(P_k)$. (The details of checking the membership of X/k in $conv(P_k)$ is not discussed as yet. Let us assume we can do this.)

Given N_{k-1} is well-defined, we proceed to define N_k recursively. Firstly we define,

$$\mathcal{V}(N_k) = \mathcal{V}(N_{k-1}) \bigcup V_{[k-2]}.$$

Now consider the links between layers $k-3$ and $k-2$. Which of the links can be considered as arcs in the network N_k? A max flow problem defined on a restricted network derived from N_{k-1} and a link L decides this. If the maximal flow in the restricted network is positive, then we can have the link L, with capacity equal to the maximal flow.

Next we define the restricted network which is induced by the deletion of a subset of nodes from $\mathcal{V}(N_{k-1})$.

Definition 3.1 *[Restricted Network $N_{k-1}(L)$] Given $k \in V_{n-1} \setminus V_4$, a link $L = (e_\alpha, e_\beta) \in E_{k-1} \times E_k$, let $e_\beta = (i, j)$ and $e_\alpha = (r, s)$. $N_{k-1}(L)$ is the sub network induced by deleting the nodes (if any) according to the following rules:*

Deletion Rules:

(a) Delete $[l : e_\beta], max(4, j) \leq l < k$.

(b) Delete $[l : e_\alpha], max(4, s) \leq l < k$.

(c) Delete $[j : e], e \notin G(e_\beta)$ if $e_\beta \in E_k \setminus E_3$.

(d) Delete $[s : e], e \notin G(e_\alpha)$ if $e_\alpha \in E_{k-1} \setminus E_3$.

(e) $[k : e], e \neq e_\alpha$.

Let $\mathcal{D}(L)$ denote the set of deleted nodes.

In the sub network induced by the remaining nodes $\mathcal{V}(N_{k-1}) \setminus \mathcal{D}(L)$, the arcs incident with the deleted nodes are all deleted. Any node $[l : e], 4 < l < k$ with in or out degree zero, is deleted from the network. Deletion rule [a] ([b]) ensures that the edge e_β (e_α) is not used for insertion earlier than stage $(k-3)$. Deletion rule [c] ([d]) ensures that the edges not in the generator of the edge e_β (e_α) are deleted, in layer $j-3(s-3)$. Finally [e] ensures that the only sink in $(k-3)^{rd}$ layer is $[k : e_\alpha]$.

Remark: 1

1 *Deletion of a node or arc can be equivalently interpreted as imposing an upper bound of zero on the flow through a node or an arc respectively with respect to a given link (treated as a commodity). This interpretation is useful in considering multicommodity flow through the network N_k.*

2 *Elsewhere a multi commodity flow problem is formulated towards answering the question: Given $X/k \in conv(P_k)$, does $X/k+1$ belong to $conv(P_{k+1})$?*

We consider the problem of finding the maximal flow in $N_{k-1}(L)$ satisfying all the restrictions on nonnegativity, flow conservation and capacity on the available nodes and arcs.

The only sink in the network is $[k : e_\alpha]$ and the sources are the undeleted nodes in $V_{[1]}$. Let $C(L)$ be the value of the maximal flow in the restricted network $N_{k-1}(L)$. We find $C(L)$ for each link L.

Now we are in a position to define the flow feasibility problem, called F_k. Consider a capacited flow feasibility problem with

$$
\begin{array}{llcl}
O -- & Origins] & : & u = [k : e_\alpha] \in V_{[k-3]} \\
D -- & Sinks] & : & v = [k+1 : e_\beta] \in V_{[k-2]} \\
\mathcal{A} -- & Arcs] & : & \{(u,v) \text{ such that } L = (e_\alpha, e_\beta) \text{ is a link and } C(L) > 0\} \\
C -- & Capacity] & : & C_{u,v} = C(L).
\end{array}
$$

If F_k is feasible and $k < n-1$ we identify the set of rigid arcs, $\mathcal{R}$ and the dummy subset of arcs in that. We update $\mathcal{A}$ by deleting the dummy arcs and marking the rigid arcs green with the frozen flow. We finally have $\mathcal{A}(N_k) = \mathcal{A}(N_k) \bigcup \mathcal{A}$. We ensure that N_k is well-defined. If $k = n-1$ we stop.

We shall subsequently prove the main results of the paper, in section 4, which brings out the significance of the construction of N_k and problem F_k we have laboured to define.

Definition 3.2 *[Pedigree path] Consider the network, N_{k-1}. Let path(X^r) denote the path corresponding to a $X^r \in P_k$, given by*

$$[4 : e_4^r] \to [5 : e_5^r] \ldots \to [k : e_k^r]$$

where X^r is the characteristic vector of $(e_4^r, \ldots e_k^r)$.

Definition 3.3 *[Pedigree pack] Consider any feasible flow, f in $N_{k-1}(L)$ for a link $L = (e', e'') \in E_{k-1} \times E_k$. Let v_f be the value of the flow f, that is, v_f reaches the sink in $N_{k-1}(L)$. We say v_f is* pedigree packable *in case there exists a subset $P(L) \subset P_k$ such that [a] $\lambda_r (\geq 0)$ is the flow along path(X^r) for $X^r \in P(L)$, [b] $e^r_k = e'$ $\forall$ $X^r \in P(L)$, [c] $\sum_{r \ni x^r(v)=1} \lambda_r \leq x(v), v \in \mathcal{V}(N_{k-1}(L))$, and [d] $\sum_{X^r \in P(L)} \lambda_r = v_f$.*

We refer to $P(L)$ as a pedigree pack *of v_f.*

Therefore given $P(L)$ every $X^r \in P(L)$, is such that, the corresponding pedigree $(e^r_4, \ldots, e^r_k = e')$ can be extended to $(e^r_4, \ldots, e^r_k = e', e'')$. This follows from the construction of the $N_{k-1}(L)$.

4 Main Results

In this section we state the main results without proofs. Explicit use of $X \in P_{MI}(n)$ is made in proving the pedigree packability of any node capacity at layer $k-2$ given that $X/k \in conv(P_k)$.

Theorem 4.1 *Given $X \in P_{MI}(n)$ and $X/k \in conv(P_k)$, consider the network N_{k-1}. Fix any $[k+1 : e_\beta] \in V_{[k-2]}$. Let $L_\alpha = (e_\alpha, e_\beta)$ be any link. Then*

1. *We have a flow f_α in $N_{k-1}(L_\alpha)$ such that the value of the flow f_α, given by v_{L_α}, is pedigree packable,*
2. *$\sum_\alpha v_{L_\alpha} = x_{k+1}(e_\beta)$, and*
3. *$\sum_\alpha \sum_{X^r \in P(L_\alpha),\ x^r(u)=1} \mu_r \leq x(u), u \in \mathcal{V}(N_{k-1})$,*

where μ_r is the flow along the path(X^r).

In other words, given $X \in P_{MI}(n)$ and $X/k \in conv(P_k)$, we have pedigree paths in N_{k-1} bringing in a total flow of $x_{k+1}(e_\beta)$ into, $[k : e_\alpha]$, for some $e'_\alpha s$ and all these paths can be extended to pedigree paths in N_k, ending in $[k+1 : e_\beta]$.

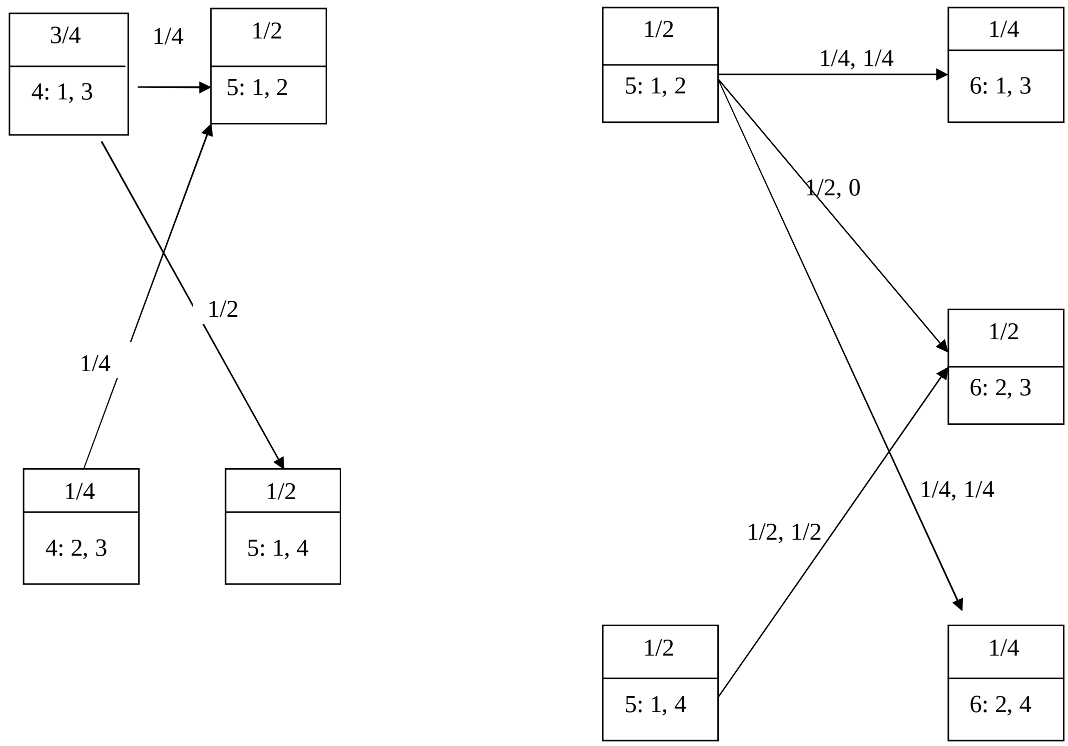
Remark: 2 *Even though we can apply this theorem for any $x_{k+1}(e) > 0$, the simultaneous application of this theorem for more than one e, in general, may not be correct. This is so because, for some paths the total flow with respect to the different e_β may violate the node capacity, $x([l+3:e]) = x_{l+3}(e)$, at some layer l for some e.*

Corollary 4.2 *Given $X \in P_{MI}(n)$ and $X/k \in conv(P_k)$, if $x_{k+1}(e) = 1$ for some e, then $X/k+1 \in conv(P_{k+1})$.*

We next state a result that gives a sufficient condition for $X/k+1$ to be in the pedigree polytope $conv(P_{k+1})$.

Theorem 4.3 *Given $X/k \in conv(P_k)$, and $x_{k+1}(e) = 0$, $\forall$ $e \in E_{k-1}$. Then F_k feasible implies $X/k+1 \in conv(P_{k+1})$.*

Example 4.1 illustrates that F_k feasible does not in general imply $X/k+1 \in conv(P_{k+1})$.



a. Layered Network N_4

b. F_5 is feasible

Figure 1: Layered Network N_4 and Bipartite Network of F_5 for Example 4.1

Example: 4.1 *Consider X given by*

$$\begin{aligned}\mathbf{x}_4 &= (0, 3/4, 1/4);\\ \mathbf{x}_5 &= (1/2, 0, 0, 1/2, 0, 0);\\ \mathbf{x}_6 &= (0, 1/4, 1/2, 0, 1/4, 0, 0, 0, 0, 0).\end{aligned}$$

It can be verified that $X \in P_{MI}(6)$. And F_4 is feasible. And f given by $f_{([4:(1,3)],[5:(1,2)])} = \frac{1}{4}$, $f_{([4:(2,3)],[5:(1,2)])} = \frac{1}{4}$, and $f_{([4:(1,3)],[5:(1,4)])} = \frac{1}{2}$ does it. All the arcs in F_4 are rigid as well. Also $X/5$ is the convex combination of $(0,1,0;1,0,0,0,0,0)$, $(0,0,1;1,0,0,0,0,0)$ and $(0,1,0;0,0,0,1,0,0)$ with weights $\frac{1}{4}$, $\frac{1}{4}$ and $\frac{1}{2}$ respectively. Next via the restricted networks $N_4(L)$ for the links in the layer two, we obtain the bipartite network given in Figure 4.1(b). The numbers along any arc give the capacity and flow respectively. Notice that F_5 is feasible, with f given by $f_{([5:(1,2)],[6:(1,3)])} = \frac{1}{4}$, $f_{([5:(1,2)],[6:(2,4)])} = \frac{1}{4}$, and $f_{([5:(1,4)],[6:(2,3)])} = \frac{1}{2}$.

Suppose X is in P_6, consider any $\lambda \in \Lambda_6(X)$. Then there are pedigrees $X^r, r \in I(\lambda)$ such that $x_6^r((2,3)) = 1$. The total weight for these pedigrees is $x_6((2,3)) = \frac{1}{2}$. But these pedigrees can not have $x_4^r((2,3)) = 1$ as they all have $x_6^r((2,3)) = 1$. So this forces the alternative $x_4^r((1,3)) = 1$ for all these pedigrees. However no pedigree X^r with $x_6^r(e) = 1$ for $e = (1,3)$ or $(2,4)$ can also have $x_4^r((1,3)) = 1$, as $(1,3) \notin G(e)$ for both the e's. And so the pedigrees $X^r, r \in I(\lambda)$ do not saturate the node capacity $3/4$ for $[4:(1,3)]$ as required. Or the supposition X is in P_6 is incorrect.

Example 4.2 illustrates that finding a pedigree pack for each $x_k(e_\beta)$ is possible but simultaneously doing this for all e_β may fail.

Example: 4.2 *Consider X as given in Example 4.1. The unique path* $[4:(2,3)] \rightarrow [5:(1,2)] \rightarrow [6:(1,3)]$ *brings the flow of* $\frac{1}{4}$ *to* $[6:(1,3)]$. *The unique path* $[4:(1,3)] \rightarrow [5:(1,4)] \rightarrow [6:(2,3)]$ *brings the flow* $\frac{1}{2}$ *to* $[6:(2,3)]$, *as required. Similarly for the node* $[6:(2,4)]$, *we have the unique path* $[4:(2,3)] \rightarrow [5:(1,2)] \rightarrow [6:(2,4)]$ *with the flow* $\frac{1}{4}$. *Notice that these paths correspond to the extensions of the pedigrees active for $X/5$. However, we can not satisfy the requirements at nodes* $[6:(2,4)]$ *and* $[6:(1,3)]$ *simultaneously, using the respective paths, as at* $[4:(2,3)]$ *the node capacity restriction is violated.*

Instead, consider X' given by:

$$\mathbf{x'}_4 = \mathbf{x}_4; \mathbf{x'}_5 = \mathbf{x}_5; \;\; but\; \mathbf{x'}_6 = (0,0,0,0,0,1,0,0,0,0).$$

We can check that $X' \in P_{MI}(6)$. The paths $[4:2,3] \rightarrow [5:(1,2)] \rightarrow [6:(3,4)]$, $[4:(1,3)] \rightarrow [5:(1,2)] \rightarrow [6:(3,4)]$, *and* $[4:(1,3)] \rightarrow [5:(1,4)] \rightarrow [6:(3,4)]$ *bring the flows of* $1/4$, $1/4$ *and* $1/2$, *respectively to* $[6:(3,4)]$, *adding up to* 1 *which is $x'_6((3,4))$. As these paths saturate the node capacities in all layers, indeed $X' \in conv(P_6)$, as assured by Corollary 4.*

We saw that F_k feasibility is not sufficient for $X/k{+}1$ to be in $conv(P_{k+1})$. However, we can prove the following necessary condition for membership in the pedigree polytope.

Theorem 4.4 *Given $X \in P_{MI}(n)$, and $X/k \in conv(P_k)$, if the bipartite flow feasibility problem, F_k, is* infeasible *then $X/k+1 \notin conv(P_{k+1})$.*

5 Conclusions & Further Research

In this paper we have constructed a layered network to obtain some new insight into the membership problem for the pedigree polytopes. We have also defined pedigree packability and proved that the flow in the restricted networks $N_{k-1}(L)$ for any link L is pedigree packable. Some sufficient conditions for membership in the pedigree polytope are derived using this result. Also F_k feasibility is shown as a necessary condition for membership in the pedigree polytope $conv(P_{k+1})$.

Current research is directed towards formulating a multicommodity flow problem in the recursively defined layered networks N_k for $k \in \{5, \ldots, n-1\}$ to go beyond the results obtained in this paper regarding the membership problem. It is conjectured that given a feasible solution to the MI-relaxation, X such that $X/n-1 \in conv(P_{n-1})$, then $X \in conv(P_n)$ if and only if there exists a solution to the above mentioned multicommodity flow problem with the total flow value equal to unity.

References

[1] T.S. Arthanari, On Pedigree polytope and Hamiltonian Cycles, paper presented at R. C. Bose Symposium, December 18-23, 2002, Kolkata, [accepted

for publication in the Special Issue of *Discrete Mathematics* devoted to the R.C. Bose Centennial Symposium].

[2] T.S. Arthanari, Pedigree polytope is a Combinatorial Polytope, Chapter 1 in *Operations Research with Economic and Industrial Applications: Emerging Trends* Eds. S.R. Mohan and S.K. Neogy. Anamaya Publishers, New Delhi, India, 2005.

[3] T.S. Arthanari and M. Usha, An Alternate Formulation of the Symmetric Traveling Salesman problem and its Properties, *Discrete Applied Mathematics* 98 (2000) 173-190.

[4] G. Dantzig, D. Fulkerson, and S. Johnson, Solution of a Large Scale Traveling Salesman Problem, *Operations Research* 2 (1954) 393-410.

[5] M. Dell' Amico, F.Maffioli and S.Martello,*Annotated Bibliographies in Combinatorial Optimization*, Wiley-Interscience, John Wiley & Sons, New York (1997).

[6] D. Gusfield, A Graph Theoretic Approach to Statistical data Security, *SIAM. J.Comput.* 17 (1988) 552-571.

[7] L.G. Khachiyan, A polynomial algorithm in linear programming," (in Russian), Doklady *Akedamii Nauk SSSR* 244 (1979) 1093-1096, [English translation: *Soviet Mathematics Doklady* 20 (1979) 191-194].

[8] B. Korte and J. Vygen, *Combinatorial Optimization*, Theory and Algorithms, Springer, New York,(2000).

[9] K.G. Murty, *Network Programming*, Prentice-Hall, NJ, (1992).

[10] D. Naddef and W.R. Pulleyblank, Hamiltonocity and combinatorial polyhedra, *J.Combi. Theory*, Ser. B 31 (1981) 297-312.

[11] C.H. Papadimitriou, The Adjacency relation on the Traveling Salesman Polytope is NP-Complete, *Mathematical Programming* 14 (1978) 312-324, .

[12] http://staff.business.auckland.ac.nz/tart002/index.htm

Robust and Integrated Airline Scheduling

Oliver Weide
Department of Engineering Science
University of Auckland
New Zealand
o.weide@auckland.ac.nz

Abstract

The airline planning process includes the problems schedule design, aircraft routing and crew pairing. In current practice these problems are solved separately. Given a set of flights in a schedule, the presented approach integrates the crew pairing problem and the aircraft routing problem. The goal is to simultaneously find a cost minimal assignment of crew and aircraft to flights in a schedule. We discuss various problems in airline scheduling and present recent attempts to integrate and solve some of the problems. We propose a Dantzig-Wolfe decomposition approach to solve the integrated aircraft routing and crew pairing problem. This concept is compared with the most successful models in the literature. We outline two extensions of the model in order to find robust solutions and incorporate departure time windows.

In airline scheduling a variety of planning problems and operational decision problems have to be solved. First, marketing decisions in the *schedule design problem* determine which flights the airline operates. Each flight is specified by origin, destination and departure time. Given the set of flights in a schedule the solution of the *fleet assignment model (FAM)* determines which flight is operated by which aircraft type. The objective is to maximize profit with respect to the number of available aircraft and other resource constraints. Next, the *aircraft routing problem* seeks a minimal cost assignment of available aircraft to the flights. A route is assigned to each individual aircraft such that each flight is covered by exactly one route. This problem is separable by aircraft type. One aircraft routing problem is solved for each aircraft type in the fleet. Similarly to the aircraft routing problem, the *crew pairing problem* (or *tour of duty problem*) allocates crews to flights in a minimal cost way. A set of generic pairings is constructed with respect to a large number of rules such that each flight is covered exactly once. Under the assumption that the crew is only allowed to operate a single aircraft type, which is the case for pilots for example, the crew pairing problem is also separable by aircraft type. The last of the tactical planning problems is *crew rostering*. Based on the constructed crew pairings a line of work is assigned to each individual crew member. See Klabjan (2005) for a detailed description of the various airline scheduling problems.

In this paper three of the above problems, schedule design, aircraft routing, and crew pairing, are considered. The fleet assignment model is important for large airlines with multiple aircraft types. In the context relevant for this paper, the fleet can be regarded as homogeneous and FAM can be omitted. The crew rostering problem can be viewed as a separate optimization problem with no influence on the cost of the overall solution and is also not considered.

Currently all three problems are solved sequentially. While the solution of the schedule design problem obviously influences the two other problems, the solutions of the latter two are also not independent. After the arrival of a flight some minimal time is required for both aircraft and crew until the next departure is possible. While these times are identical when the crew stays on the same aircraft, the time needed by the crew increases whenever they change aircraft. Solving the aircraft routing problem first restricts the number of possible connections between two flights a crew is able to operate. This might lead to sub-optimal solutions. Conversely, if the crew pairing problem is solved first, the aircraft routing problem might be infeasible.

To eliminate these effects, a model to solve the aircraft routing problem and the crew pairing problem simultaneously is proposed here. The Dantzig-Wolfe decomposition approach (Dantzig and Wolfe 1960) we present links the problems together but preserves their original structures. These structures are well understood and can be exploited in order to solve the individual problems which are already $\mathcal{NP}$-hard. We can extend our model to identify solutions which also behave robust in operations. Robust solutions have a small planned cost and do not cause large additional costs once disruptions occur. The solution approach follows the idea developed by Ehrgott and Ryan (2002). In a subsequent step we incorporate time windows for the departure times in order to find a low cost solution which is also robust.

A number of recent publications address various problems in airline scheduling. Different combinations of problems have been integrated and a variety of robustness measures have been introduced. However, none of the papers exploits the advantages of Dantzig-Wolfe decomposition and no attempt has been made to integrate all three problems, schedule design, aircraft routing and crew pairing, and include a robustness measure.

The paper is organized as follows: In Section 1 the individual problems are formulated. We review the most recent approaches in the literature in Section 2. Section 3 describes the integrated model and in Section 4 the solution approach is presented. Section 5 concludes with two possible extensions of the model.

1 Problem Formulation

In this section the crew pairing problem and the aircraft routing problem are formulated.

1.1 Crew Pairing Problem

Given a flight schedule the *crew pairing problem* is defined as the problem of assigning crews to the set of flights in the schedule such that each flight is operated by exactly one crew. A sequence of flights which can be flown by a crew on one work day is a *duty period*. An alternating sequence of duty periods and rest periods is called *crew pairing* (or *tour of duty*). Any crew pairing must start and end in the same crew base and is restricted by a number of requirements. These are obtained from various rules like rest time regulations or flight time restrictions. Costs are calculated for each crew pairing based on the contained flights. The crew pairing problem seeks to find a cost minimal set of pairings that partition the flights in the schedule. We assume that there are no capacity constraints on the crew bases.

The pairings can be represented as columns of a matrix $A_K \in \{0,1\}^{(m,n_K)}$ where m is the number of flights in the schedule and n_K is the number of possible pairings. The element $A_K(i,j)$ is set to 1 if flight i is contained in pairing j and 0 otherwise. With this matrix representation the crew pairing problem can be formulated as a standard set partitioning model:

$$\begin{array}{lrcl} \text{Minimize} & c_K^T x_K & & \\ & & & \\ \text{subject to} & A_K x_K & = & \mathbb{1} \\ & x_K & \in & \{0,1\}^{n_K} \end{array} \tag{1}$$

The element $c_K(j)$ of $c_K \in \mathbb{R}^{n_K}$ is the cost associated with pairing j. The decision variable $x_K(j) \in \{0,1\}^{n_K}$ is equal to 1 if pairing j is chosen in the solution and 0 otherwise.

Note that to solve the problem the number of pairings n_K is too large to efficiently enumerate all possible pairings. For this reason column generation is employed. Only a small fraction of all pairings is considered initially and new pairings are generated as needed.

To obtain a solution, first the LP relaxation of (1) is solved using column generation. Then a branch-and-bound algorithm is used to obtain an integer solution where additional columns are generated at the nodes of the branch-and-bound tree. See Wolsey (1998) for a general description of these techniques and Barnhart et al. (2003) for a detailed description of the crew pairing problem.

1.2 Aircraft Routing Problem

The *aircraft routing problem* is the problem of assigning aircraft to a given set of flights in a schedule. One seeks to find a routing for each aircraft such that each flight of the schedule is contained in exactly one routing. The number of available aircraft is given and each particular aircraft is assigned to one specific routing. This is also referred to as *tail assignment*. The problem is similar to crew rostering where one assigns a line of work to a particular crew member and in contrast to the crew pairing problem where we seek an unknown number of generic pairings to partition all flights without assigning a pairing to a specific crew.

Similar to pairings, routings can be represented as columns of a matrix $A_F \in \{0,1\}^{(m+a,n_F)}$ where a is the number of available aircraft and n_F the number of possible routings. The first m rows are defined as above: The element $A_F(i,j)$, $1 \leq i \leq m, 1 \leq j \leq n_F$ is set to 1 if flight i is contained in routing j and 0 otherwise. Additionally the element $A_F(m+i,j), 1 \leq i \leq a, 1 \leq j \leq n_F$ is set to 1 if routing j is operated by aircraft i and 0 otherwise. With this matrix representation the aircraft routing problem looks identical to the crew pairing formulation but observe that the matrix definition has been altered:

$$\begin{array}{lrcl} \text{Minimize} & c_F^T x_F & & \\ & & & \\ \text{subject to} & A_F x_F & = & \mathbb{1} \\ & x_F & \in & \{0,1\}^{n_F} \end{array} \tag{2}$$

The element $c_F(j)$ of $c_F \in \mathbb{R}^{n_F}$ is the cost associated with routing j. The decision variable $x_F(j) \in \{0,1\}^{n_F}$ is equal to 1 if routing j is chosen in the solution and 0 otherwise.

As in the crew pairing problem, the number of possible routings is very large and column generation techniques and LP based branch-and-bound methods are used to solve the problem.

2 Related Literature

Airline scheduling problems have been addressed in an extensive number of publications, see Klabjan (2005) for a detailed description of the problems. Crew pairing, fleet assignment and aircraft routing have received most attention in the literature. The most recent publications are listed here.

Schedule design is not discussed as a separate problem but in combination with fleet assignment. Hence, we start with the literature on FAM. A basic FAM model is introduced by Hane et al. (1995). It is also called leg-based fleet assignment model because revenue effects between flights are not modeled. Desaulniers et al. (1997) integrate FAM and time windows. Barnhart, Kniker, and Lohatepanont (2002) and Kniker (1998) describe an enhanced model using demand forecasts for origin-destination pairs. This model is much harder to solve than the basic FAM.

One integrated model for schedule design and fleet assignment is presented by Lohatepanont and Barnhart (2004). They use the origin-destination fleet assignment model. Given a set of mandatory and optional flights, a subset of optional flights is chosen to maximize profit. Further models are described by Lettovský, Johnson, and Smith (1999), Yan and Wang (2001), Yan and Tseng (2002) and Erdmann et al. (2001).

The aircraft routing problem is modeled as a partitioning model using a string formulation for sequences of flights by Barnhart et al. (1998). Cordeau et al. (2001) and Mercier, Cordeau, and Soumis (2003) model the problem as a multi-commodity network flow model with nonlinear resource constraints. Barnhart et al. (1998) also combine fleet assignment and aircraft routing. A model for integrating fleet assignment and crew pairing is proposed by Barnhart, Lu, and Shenoi (1998).

We refer to Barnhart et al. (2003) for a detailed description of the crew pairing problem and a review of the vast amount of literature addressing the problem. See Ernst et al. (2004) for an annotated bibliography of rostering.

Various combinations of airline scheduling problems have been integrated. Here we focus on the integration of crew pairing and aircraft routing problems:

Klabjan et al. (2002) partially integrate aircraft routing, crew pairing and schedule design. They reverse the order of crew pairing and aircraft routing using plane count constraints to ensure the existence of a feasible solution for the aircraft routing problem. Their results are based on a hub-and-spoke network. Schedule design is incorporated by relaxing feasibility parameters and generating a larger set of pairings. They solve the model via an LP based branch-and-bound algorithm.

A model to integrate aircraft routing and crew pairing is proposed by Cordeau et al. (2001) and Mercier, Cordeau, and Soumis (2003). They use Benders decomposition and branch-and-price to solve the model. Employing the crew pairing problem as the sub problem as well as the master problem have been tested, the latter with better success. Both methods add cut inequalities to the original set partitioning polytopes of the problems. They also try to apply the approach presented in Klabjan et al. (2002) to an interconnected network but were not successful in obtaining feasible solutions for the aircraft routing problem.

Cohn and Barnhart (2002) also integrate aircraft routing and crew pairing. They add one variable to the crew pairing problem for each solution of the aircraft routing problem. LP based branch-and-price is used as the solution method but even after reducing the problem size the approach remains computationally expensive.

Sandhu and Klabjan (2004) partially integrate fleet assignment, aircraft routing and crew pairing with a similar approach as Klabjan et al. (2002) and solve the model with both Lagrangian relaxation and Benders decomposition.

In vehicle scheduling Haase, Desaulniers, and Desrosiers (2001), Freling, Huisman, and Wagelmans (2003) and Huisman, Freling, and Wagelmans (2003) propose models to integrate vehicle and crew scheduling.

In order to improve the behavior in operations of the constructed solutions of the crew pairing problem a number of robustness measures have been introduced:

Schaefer et al. (2000) use an expected operational cost for the crew pairings instead of planned cost. Monte Carlo simulation is used to estimate costs. The simulation ignores interactive effects between pairings and uses only a push-back strategy for recovery. Pairings with expensive operational cost are penalized.

Yen and Birge (2000) formulate the crew pairing problem as a two-stage stochastic binary optimization problem with recourse in a computationally expensive approach. The recourse problem measures the cost of delays. Crew switching aircraft are penalized in the objective function.

A similar measure of robustness is introduced by Ehrgott and Ryan (2002). Crew pairings are penalized where crew changing aircraft are likely to increase the effects of disruptions. Robustness is treated in a second objective function in a bi-criteria approach.

Most recently, Shebalov and Klabjan (2004) solve the crew pairing problem first and then maximize the number of move-up crews (possible swaps) without increasing the cost too much. They compare their method with the method of solving the standard crew pairing problem by simulating disruptions with SimAir.

3 Integrated Crew Pairing and Aircraft Routing Problem

If two flights are operable consecutively by the same aircraft or the same crew a *connection* occurs between the pair of flights. The time between arrival of the first and departure of the second flight is called *connection time*. The minimal time required for an aircraft to operate a connection is the *minimal turn time*. The minimal time required for a crew is called *minimal sit time*. The minimal sit time usually exceeds the minimal turn time if the crew is changing aircraft . If the crew stays on the same aircraft both are identical. A connection of flights i, j is called *short* if

$$(\text{minimal turn time})_{ij} \leq (\text{connection time})_{ij} \leq (\text{minimal sit time})_{ij}.$$

Thus, in a feasible solution, short connections are only allowed if the crew stays on the same aircraft. This condition might result in sub-optimal or infeasible solutions if the two problems are solved separately.

By enumerating all possible short connections we can define two matrices $B_K \in \{0,1\}^{(m_s,n_K)}$ and $B_F \in \{0,1\}^{(m_s,n_F)}$ where m_s is the number of short connections. For each pairing j, $B_K(i,j)$ is set to 1 if short connection i is contained in pairing j and 0 otherwise. B_F is defined in an analogous way.

With this matrix representation the *integrated crew scheduling and aircraft routing problem* can be formulated as follows:

$$\begin{array}{llcll} \text{Minimize} & c_K^T x_K & + & c_F^T x_F & \\ \\ \text{subject to} & A_K x_K & & & = \mathbb{1} \\ & & & A_F x_F & = \mathbb{1} \\ & B_K x_K & - & B_F x_F & \leq 0 \end{array} \tag{3}$$

where $x_K \in \{0,1\}^{n_K}$ and $x_F \in \{0,1\}^{n_F}$.

The third set of constraints is coupling the aircraft routing and crew pairing problems. The constraints ensure that short connections which are operated in the solution by some crew are also operated by some aircraft.

4 Solution Approach

The approach in the literature for solving the integrated aircraft routing and crew pairing problem which currently seems to be the most successful uses Benders decomposition (Benders 1962). The integrated formulation is decomposed into a master and a subproblem where the crew pairing problem is treated as the master problem, see Mercier, Cordeau, and Soumis (2003). During the algorithm, cut inequalities are added to the original set partitioning polytopes of both problems, causing slow convergence towards an optimal solution.

We propose using Dantzig-Wolfe decomposition (Dantzig and Wolfe 1960). The aircraft routing problem and the crew pairing problem are both solved as subproblems and the additional coupling constraints are treated in the master problem. This approach does not alter the structure of the original polytopes and is described in the following.

We consider problem (3) and represent the set partitioning polytope $\mathcal{P}_K = \text{conv}(\{x_K \in \{0,1\}^{n_K} | A_K x_K = \mathbb{1}\})$ of the crew pairing problem by its extreme points, i.e. $\mathcal{P}_K = \text{conv}(\{v_1, \ldots, v_k\})$. $V_K = \{v_1, v_2, \ldots, v_k\}$ is the set of extreme points of the set partitioning polytope and each feasible solution x_K is a convex combination of these extreme points. We define $\mathcal{P}_F, V_F$ in the same way and represent problem (3) with relaxed integrality property in the following alternative formulation:

$$\begin{array}{llcllll} \text{Minimize} & c_K^T V_K \lambda & + & c_F^T V_F \mu & & & \\ \text{subject to} & B_K V_K \lambda & - & B_F V_F \mu & \leq 0 & \rightarrow & \overline{y} \\ & \mathbb{1}^T \lambda & & & = 1 & \rightarrow & \overline{y}_K \\ & & & \mathbb{1}^T \mu & = 1 & \rightarrow & \overline{y}_F \\ & \lambda, & & \mu & \geq 0 & & \end{array} \tag{4}$$

This defines the master problem of the decomposition. The first set of constraints are the short connection constraints as above. The other two constraints insure that the solution is contained in the convex hull of V_K and V_F, respectively. The sets V_K

and V_F are unknown, they are constructed iteratively during the algorithm by two subproblems. These are precisely (1) and (2) with altered objective functions:

$$\begin{array}{lrcl} \text{Minimize} & (c_K^T - \overline{y}^T B_K)x_K & & \\ \text{subject to} & A_K x_K & = & \mathbb{1} \\ & x_K & \in & \{0,1\}^{n_K} \end{array} \tag{5}$$

$$\begin{array}{lrcl} \text{Minimize} & (c_F^T + \overline{y}^T B_F)x_F & & \\ \text{subject to} & A_F x_F & = & \mathbb{1} \\ & x_F & \in & \{0,1\}^{n_F} \end{array} \tag{6}$$

The single steps of the algorithm are as follows:

1. The sets V_K and V_F are initialized with artificial variables to ensure the existence of a feasible solution of the master problem. $\overline{y}$ is set to 0.
2. We solve both subproblems. The two optimal solutions x_K and x_F are added to the sets V_K and V_F.
3. The master problem is solved and dual variables $\overline{y}$ are obtained as input for the subproblems in the next iteration. We update the objective coefficients of the subproblems and resolve.
4. This procedure is stopped when the dual variables $\overline{y}_K$ and $\overline{y}_F$ of the master problem associated with the convexity constraints both exceed the objective value of the related subproblem. This proves that no entering column with negative reduced cost exists, and thus optimality of the current solution of the master problem is guaranteed.

To obtain an integer solution the master problem is embedded in a branch-and-bound structure. Note that the master problem with integrality conditions reduces to choose two matching columns, one for each of the subproblems. Both subproblems are solved with column generation and branch-and-bound methods as described above.

5 Model extensions

5.1 Robust solutions

The concept of *robust solutions* is to identify good cost solutions that are also likely to behave well in operations. A more robust solution is understood as a solution where effects of potential delays in the pairings are decreased. These effects incur additional costs, caused by extra necessary crews or required cancellations, which we would like to avoid.

If one flight of an aircraft is delayed, e.g. due to bad weather, and it is operating on a tight schedule the following flights of the same aircraft will also be delayed. If the crew that operates the delayed flight changes aircraft to operate another flight after only minimal connection time this flight will also be delayed. The goal is to identify and avoid solutions with such bad behavior. The example gives reason for the following approach:

Crews who change aircraft will be penalized in the objective function if the connection the crew uses is just slightly longer than the minimal sit time but less than some *restricted time*.

A connection between two flights i, j is called *restricted* if

$$(\text{minimal sit time})_{ij} \leq (\text{connection time})_{ij} \leq (\text{restricted time})_{ij}.$$

In order to include robustness in our model, we add constraints similar to the short connection constraints. The difference is that whenever a crew is changing aircraft short connections are forbidden in any feasible solution while restricted connections are allowed but will incur a penalty in the objective function. The resulting solution contains few connections where crew change aircraft with only minimal connection time and is therefore likely to be robust.

5.2 Time Windows

In this further extension of the model the originally scheduled departure times of the flights in the schedule are allowed to vary in some time window. Hence, we do not consider the total schedule design problem as described above but the so called *schedule perturbation problem*. Allowing some flexibility in the departure times results in a larger set of feasible aircraft routings and crew pairings. The degree of flexibility can be adjusted for each flight separately so that some flights may depart a few minutes earlier or later than scheduled while others must depart at the originally scheduled time. The goal is to identify a slight perturbation of the schedule which enables originally infeasible solutions for the aircraft routing problem and the crew pairing problem to become feasible. We expect to obtain solutions with low cost which are also robust.

In order to integrate the concept of time windows into the model we start with a size of 0 for all windows and thus solve the integrated problem as described above. We then steadily increase the allowed size of the windows. Due to the growing set of possible solutions we eventually find a better cost or more robust solution. The widening of the windows is continued until the maximum allowed size of the windows is reached. We then choose between the original solution and preferable solutions with an acceptable variation of the schedule.

One major difficulty is to synchronize the departure times in both subproblems. One has to prevent the use of departure times in the crew pairing solution and in the aircraft routing solution that contradict each other. This problem is addressed by additional inequalities in the master problem.

6 Conclusion

The approach we presented integrates the aircraft routing problem and the crew pairing problem, considers a robustness measure and partially integrates the schedule design problem. Existing models in the literature do not consider all these problems simultaneously. Furthermore we described the advantage of preserving the original structure of the subproblems by solving the formulation with Dantzig-Wolfe decomposition.

References

Barnhart, C., N. Boland, L. Clarke, E. Johnson, G. Nemhauser, and R. Shenoi. 1998. "Flight string model for aircraft fleeting and routing." *Transportation Science* 32:208–220.

Barnhart, C., A. Cohn, E. Johnson, D. Klabjan, G. Nemhauser, and P. Vance. 2003. "Airline Crew Scheduling." In *Handbook of Transportation Science*, edited by R. Hall. Kluwer Scientific Publishers.

Barnhart, C., T. Kniker, and M. Lohatepanont. 2002. "Itinerary-based airline fleet assignment." *Transportation Science* 36:199–217.

Barnhart, C., F. Lu, and R. Shenoi. 1998. "Integrated airline schedule planning." In *Operations Research in the Airline Industry*, edited by G. Yu, 384–403. Kluwer Academic Publishers.

Benders, J. 1962. "Partitioning procedures for solving mixed-variables programming problems." *Numerische Mathematik* 4:238–252.

Cohn, A., and C. Barnhart. 2002. "Improving crew scheduling by incorporating key maintenance routing decisions." *Operations Research* 51:387–396.

Cordeau, J., G. Stojković, F. Soumis, and J. Desrosiers. 2001. "Benders decomposition for simultaneous aircraft routing and crew scheduling." *Transportation Science* 35:375–388.

Dantzig, G., and P. Wolfe. 1960. "Decomposition Principle for Linear Programs." *Operations Research* 8:101–111.

Desaulniers, G., J. Desrosiers, Y. Dumas, M. Solomon, and F. Soumis. 1997. "Daily aircraft routing and scheduling." *Management Science* 43:841–855.

Ehrgott, M., and D. Ryan. 2002. "Constructing Robust Crew Schedules with Bicriteria Optimization." *Journal of Multi-Criteria Decision Analysis* 11:139–150.

Erdmann, A., A. Nolte, A. Noltemeier, and R. Schrader. 2001. "Modeling and solving the airline schedule generation problem." *Annals of Operations Research* 107:117–142.

Ernst, A., H. Jiang, M. Krishnamoorthy, B. Owens, and D. Sier. 2004. "An Annotated Bibliography of Personnel Scheduling and Rostering." *Annals of Operations Research* 127:21–144.

Freling, R., D. Huisman, and A. Wagelmans. 2003. "Models and algorithms for integration of vehicle and crew scheduling." *Journal of Scheduling* 6:63–85.

Haase, K., G. Desaulniers, and J. Desrosiers. 2001. "Simultaneous vehicle and crew scheduling in urban mass transit systems." *Transportation Science* 35:286–303.

Hane, C., C. Barnhart, E. Johnson, R. Marsten, G. Nemhauser, and G. Sigismondi. 1995. "The fleet assignment problem: Solving a large-scale integer program." *Mathematical Programming* 70:211–232.

Huisman, D., R. Freling, and A. Wagelmans. 2003. "Multiple-depot integrated vehicle and crew scheduling." Technical Report EI2003-02, Econometric Institute, Erasmus University, Rotterdam.

Klabjan, D. 2005. “Large-scale Models in the Airline Industry.” In *Column Generation*, edited by G. Desaulniers, J. Desrosiers, and M. Solomon. Kluwer Academic Publishers.

Klabjan, D., E. Johnson, G. Nemhauser, E. Gelman, and S. Ramaswamy. 2002. “Airline crew scheduling with time windows and plane count constraints.” *Transportation Science* 36:337–348.

Kniker, T. 1998. “Itinerary-based airline fleet assignment.” Ph.D. diss., Massachusetts Institute of Technology.

Lettovský, L., E. Johnson, and B. Smith. 1999. “Schedule generation model.” Edited by R. Darrow, *Thirty-Ninth Annual AGIFORS Symposium.* New Orleans.

Lohatepanont, M., and C. Barnhart. 2004. “Airline Schedule Planning: Integrated Models and Algorithms for Schedule Design and Fleet Assignment.” *Transportation Science* 38:19–32.

Mercier, A., J. Cordeau, and F. Soumis. 2003. “A computational study of Benders decomposition for the integrated aircraft routing and crew scheduling problem.” Technical Report G-2003-48, GERARD.

Sandhu, R., and D. Klabjan. 2004. “Integrated Airline Planning.” Working paper (submitted for publication).

Schaefer, A., E. Johnson, A. Kleywegt, and G. Nemhauser. 2000. “Airline crew scheduling under uncertainty.” Technical Report TLI-01-01, Georgia Institute of Technology.

Shebalov, S., and D. Klabjan. 2004. “Robust airline crew scheduling: Move-up crews.” To appear in Transportation Science.

Wolsey, L. 1998. *Integer Programming.* Wiley-Interscience publication.

Yan, S., and C. Tseng. 2002. “A passenger demand model for airline flight scheduling and fleet routing.” *Computers and Operations Research* 29:1559–1581.

Yan, S., and C. Wang. 2001. “The planning of aircraft routes and flight frequencies in an airline network operations.” *Journal of Advanced Transportation* 35:33–46.

Yen, J., and J. Birge. 2000. “A stochastic programming approach to the airline crew scheduling problem.” Technical Report, University of Washington.

A survey of beam intensity optimization in IMRT

Lizhen Shao
Department of Engineering Science
University of Auckland
New Zealand
l.shao@auckland.ac.nz

Abstract

Intensity modulated radiation therapy (IMRT) is an important modality in radiotherapy. Due to the complexity of IMRT treatment plans, optimization methods are needed to design high quality treatments. Three main problems need to be addressed in IMRT: (1) the beam angle optimization problem, (2) the beam intensity optimization problem, (3) the intensity delivery optimization. To date many approaches have been proposed to solve these problems.

In this paper, we focus on the beam intensity optimization problem. Basically there are four types of models used to formulate the problem; they are linear programming, nonlinear programming, mixed integer programming and multiple objective programming. We will review the literature on the mathematical models of beam intensity optimization developed for IMRT.

Keywords: intensity modulated radiation therapy (IMRT), optimization, mathematical model.

1 Introduction

With the development of computer techniques and the dynamic delivery technique based on computer-controlled multileaf collimators, radiotherapy developed quickly. This led to the emergence of intensity modulated radiation therapy (IMRT). Modern IMRT is generally designed using inverse planning, which uses optimization techniques to shape the dose distribution, with the capability of generating concave dose distributions and providing specific sparing of sensitive normal structures within complex treatment geometries. IMRT has the potential to achieve a much higher degree of target conformity and/or normal tissue sparing than most other treatment techniques, especially for target volumes and/or organs at risk with complex shapes and/or concave regions.

Three problems need to be addressed in IMRT. Firstly, the choice of the beam angle. Radiation treatment is applied by moving the treatment unit around the patient body and stopping at certain positions to deliver radiation for a specified period of time, so the directions selected for delivery of radiation greatly affect radiation dose levels in the tumor and healthy tissue/organs surrounding the tumor. Secondly, optimization of beam intensities. Given the directions of the beams, beam intensities need to be produced to best control the tumor and limit the dose on critical organs. Thirdly, optimization of the intensity delivery. Once the beam intensity maps are determined, one must convert these into MLC leaf sequences that attempt to realize them. Many approaches have been proposed to solve the three problems.

At present, in all clinically relevant optimization approaches for IMRT inverse planning, the beam number and direction are fixed. In practice, they are manually selected before optimization.

In this paper, we will focus solely on the problem of beam intensity optimization. This paper is organized as follows. In Section 2, we give a background of beam intensity optimization, detailing the dose calculation and introducing dose-volume constraints (DVCs). Then, in Section 3, we discuss the mathematical models which are used for beam intensity optimization. We conclude in Section 4.

2 Beam Intensity Optimization

The IMRT beam intensity optimization problem has been extensively studied for a number of years. A survey on this subject from a mathematical point of view has been done by Shepard et al. 1999. Lim, Ferris, and Wright 2002 developed an optimization framework for conformal radiation treatment planning. There are two factors that affect the optimization performance and outcome, they are the form of the mathematical model and the optimization strategy.

In inverse treatment planning of IMRT, the clinical objectives are specified mathematically in the form of an objective function. The objective function measures the goodness of a treatment plan, so the choice of the objective function is crucial for the optimization of a treatment plan. Usually, the objective function is a function of the beamlet intensities.

Two types of objective functions are used: physical models and radiobiological models. Physical models are solely based on dose, while biological models argue that optimization should be based on the biological effects produced by the underlying dose distributions. A common method to express radiobiological objective functions is based on tumor control probabilities (TCP) and normal-tissue complication probabilities (NTCP). The treatment objective is usually stated as the maximization of TCP while maintaining the NTCP within acceptable levels.

However, because the dose response function, which is used to describe TCP and NTCP is not sufficiently understood, biological models have not been widely used in practical optimization. Thus only dose-based physical models are considered in this paper.

2.1 Notations and Dose Calculation

For dose calculation, the patient's 3D volume is divided into $|\mathcal{V}|$ small volume elements or $|\mathcal{V}|$ voxels. The discretisation is based on patient anatomy information, usually CT/MRI scans. For example, for a 5mm thickness CT slice, we could set its voxel size equal to $5 \times 5 \times 5$ mm. Each voxel is assigned to a particular structure (tissue type). Here we use $\mathcal{T}$ to represent tumor, $\mathcal{C}$ to represent critical organs (K

critical organs are represented by $\mathcal{C}_1, ..., \mathcal{C}_K$), $\mathcal{N}$ to represent normal tissue. Delineating the structure, some voxels will be marked as part of the tumor, some as part of the liver $\mathcal{C}_1$, some as part of the stomach $\mathcal{C}_2$, etc. The total number of voxels $|\mathcal{V}|$ is equal to $|\mathcal{T}| + |\mathcal{C}| + |\mathcal{N}|$, where $|\mathcal{C}| = |\mathcal{C}_1| + \ldots + |\mathcal{C}_K|$.

The beam intensity profile at a given gantry angle a is a 2D function. In order to mathematically model beam intensities using MLC, the beam is discretised into a beamlet map of B elements, the edge length of which is determined by the width of an MLC leaf channel and the positions in which the leaf can stop. Let $|\mathcal{B}|$ be the total number of beamlets, which is calculated as the number of beams A delivered multiplied by the number of beamlets B per beam head, so $|\mathcal{B}| = AB$.

The dose in each voxel v is calculated for a set of beamlet intensities as follows:

$$p(v) = d(v)x = \sum_{b=1}^{|\mathcal{B}|} d_{vb} x_b, \text{ for } v = 1, \ldots, |\mathcal{V}|, \tag{1}$$

where $p(v)$ denotes the total dose deposited in voxel v, d_{vb} represents the dose deposited in voxel v due to unit intensity in beamlet b. Since d_{vb} is a function of the geometry between the beamlet and the voxel, the modality and energy of the beamlet and the anatomy involved, it is constant for a given patient, x_b denotes the intensity in each beamlet b. The individual intensity profiles for each beam are composed of the single beamlet values of this beam. In the matrix notation, the dose calculation formula (1) becomes

$$p = Dx, \tag{2}$$

where p is a dose vector consisting of $|\mathcal{V}|$ elements and D consists of all the elements d_{vb}, it is called dose deposition matrix.

In addition, we introduce $D_\mathcal{T}$, $D_\mathcal{C}$, $D_\mathcal{N}$ to represent the dose deposition matrix on tumor, critical organs and normal tissue, they are the sub matrices of D consisting of the elements d_{vb} in voxel set $\mathcal{T}$, $\mathcal{C}$ and $\mathcal{N}$, respectively. Then the dose vector for tumor can be expressed as $p_\mathcal{T} = D_\mathcal{T} x$. We use TG, CG, NG to represent the prescribed doses for tumor, critical organs and normal tissue, they are vectors. TLB, TUB, CUB, NUB are all vectors, they represent lower bounds and upper bounds on the dose delivered to tumor, critical organs and normal tissue, respectively. Correspondingly, TLB_v, TUB_v, CUB_v, NUB_v, mean the bounds for the v-th voxel.

2.2 Dose-Volume Constraints

Most modern IMRT inverse planning systems allow the specification of dose-volume constraints (DVCs). Typical dose-volume constraints limit the relative volume of a structure that receives more or less than a particular threshold. For example, in order to improve tumor dose while avoid serious lung complications, instead of specifying the strict upper dose limit 20Gy on the lung, the planner could specify that "no more than 40% volume of the lung can exceed a radiation dose of 20Gy". This kind of constraint is very useful for the clinic.

3 Mathematical Models

In this section, we briefly review the mathematical models with emphasis on more recent publications. We classify the mathematical models into four categories: (1)

linear programming models, (2) nonlinear programming models, (3) mixed integer programming models, and (4) multiple objective programming models. The advantage and disadvantage of each model will be discussed.

3.1 Linear Programming

The first linear optimization model that was developed to aid IMRT design appeared in the literature in 1968 (Bahr et al. 1968). Since then, many researchers have experimented with linear models. The reader is referred to Shepard et al. 1999 and Lim, Ferris, and Wright 2002 for an overview of these models.

Basically, all these objectives of linear programming models are variations of: (1) minimize average/maximum dose or deviation from upper bounds on dose to the critical organs and normal tissue, (2) maximize average/minimum dose to the tumor, and (3) minimize average/maximum deviation from prescribed dose to the tumor.

The constraints might be: (1) nonnegativity constraints for the beam intensity, (2) upper bound for the critical organ and/or normal tissue, (3) lower and/or upper bound for the tumor, (4) upper bound on the ratio between the maximum beamlet intensity and the average beamlet intensity, and (5) upper bound on the mean dose to the critical organ.

By sensibly combining above objectives and constraints, different models can be formulated. Normally, if the objective is on the tumor, then the constraints should be on the critical organs and normal tissue. On the other hand, if the objective is on the critical organs and normal tissue, then the constraints should be on the tumor. However, nonnegativity constraints are always present as a physical constraint.

A simple example for minimizing the weighted sum of maximum deviation from the prescribed dose on organs subject to nonnegativity constraints can be described as follows:

$$\begin{array}{ll} \min & \omega_T \|D_{\mathcal{T}}x - TG\|_\infty + \omega_C \|D_{\mathcal{C}}x - CG\|_\infty + \omega_N \|D_{\mathcal{N}}x - NG\|_\infty \\ \text{s.t.} & x \geq 0, \end{array} \tag{3}$$

where ω is a vector of weighting factors, in the literature also known as structural importance factors.

The linear programming approach has the advantages of speed and that it is guaranteed to have positive solutions. However, sometimes linear programming can not come out with a feasible solution and the source of infeasibility is unknown. Furthermore, due to simplex algorithms producing an extreme point solution, linear solvers have the problem that physicians' limits are often attained, that means therapy plans narrowly adhere to the prescription, i.e. either portions of the critical organs are to receive their maximum allowable dose or the tumor is to receive the lowest allowable dose, while none of these two results are desirable. Moreover, it is hard to include dose-volume constraints.

Recently, there are some improved methods. Holder 2003 proposes a new linear programming model incorporating elastic constraints. The interior of the feasible set is never empty and when solved with a path following interior point method it terminates with a solution that strictly satisfies as many inequalities as possible.

The elastic model is as follows:

$$
\begin{array}{ll}
\min & \omega \cdot l^T\alpha + u_{\mathcal{C}}^T\beta + u_{\mathcal{N}}^T\gamma \\
\text{s.t.} & \\
& \begin{array}{rll}
TLB - L\alpha & \le D_{\mathcal{T}}x & \le TUB \\
D_{\mathcal{C}}x & \le CUB + U_{\mathcal{C}}\beta & \\
D_{\mathcal{N}}x & \le NUB + U_{\mathcal{N}}\gamma & \\
0 & \le L\alpha & \le TLB \\
-CUB & \le U_{\mathcal{C}}\beta & \\
0 & \le U_{\mathcal{N}}\gamma & \\
0 & \le x, &
\end{array}
\end{array}
\tag{4}
$$

where $\alpha \in \mathbb{R}^{qT}$, $\beta \in \mathbb{R}^{qC}$, $\gamma \in \mathbb{R}^{qN}$, $l \in \mathbb{R}^{qT}$, $u_{\mathcal{C}} \in \mathbb{R}^{qC}$, $u_N \in \mathbb{R}^{qN}$, $L \in \mathbb{R}^{|\mathcal{T}|\times qT}$, $U_{\mathcal{C}} \in \mathbb{R}^{|\mathcal{C}|\times qC}$, $U_{\mathcal{N}} \in \mathbb{R}^{|\mathcal{N}|\times qN}$. The constraints $TLB - L\alpha \le D_{\mathcal{T}}x \le TUB$, $D_{\mathcal{C}}x \le CUB + U_{\mathcal{C}}\beta$, and $D_{\mathcal{N}}x \le NUB + U_{\mathcal{N}}\gamma$ are called elastic constraints because the bounds are allowed to vary with the vectors α, β, and γ, respectively. The matrices L, $U_{\mathcal{C}}$, and $U_{\mathcal{N}}$ define how one measures the amount of elasticity, and l, $u_{\mathcal{C}}$, $u_{\mathcal{N}}$ show how one either penalizes or rewards the amount of elasticity. ω is the weight deciding the importance of the tumor uniformity. As ω increases, we increase the emphasis of finding a plan that achieves a uniform, tumoricidal dose.

Different elastic functions lead to different solution analysis, in particular, the author analyses two collections of elastic functions. One is average analysis, where $l = \frac{1}{|\mathcal{T}|}e$, $u_{\mathcal{C}} = \frac{1}{|\mathcal{C}|}e$, $u_{\mathcal{N}} = \frac{1}{|\mathcal{N}|}e$, $L = I$,$U_{\mathcal{C}} = I$, $U_{\mathcal{N}} = I$. The other is absolute analysis, where $l = 1$, $u_{\mathcal{C}} = 1$, $u_{\mathcal{N}} = 1$, $L = e$, $U_{\mathcal{C}} = e$, $U_{\mathcal{N}} = e$.

This model overcomes the disadvantage of infeasibility. However, the dose-volume constraints are not included.

Attempts to include dose-volume constraints in linear models are found in Merritt and Zhang 2002, Romeijn et al. 2003 and Romeijn et al. 2005.

Merritt and Zhang 2002 use successive linear programming to optimize the beam intensity map. An objective of maximizing minimum tumor dose is used and upper bounds for the tumor and the critical organs are specified. Moreover, dose-volume control is imposed by means of successive relaxation of upper bounds on critical organ dosages until the specified maximum allowable volume has been relaxed. However, the infeasibility problem still exists in this model.

Romeijn et al. 2003 and Romeijn et al. 2005 use a piecewise linear convex function to approximate any convex objective function. This model overcomes the apparent limitations of linear programming. In addition, they use the concept of CVaR as a novel type of dose-volume constraint. A constraint that bounds the tail averages of the differential dose-volume histograms of structures is imposed while retaining linearity as an alternative approach to improve dose homogeneity in the target volumes, and to attempt to spare as many critical structures as possible.

3.2 Nonlinear Programming

At present, a basic nonlinear programming model, the weighted least squares model is one of the most prevalent formulations, see, e.g., Spirou and Chui 1998 and Xing et al. 1998 for reference. Most commercial IMRT systems on the market model their objective as a weighted least squares function, such as Helios, Pinnacle, Focus, CORVUS, and KonRad.

The weighted least squares model is:

$$\begin{array}{ll} \min & \frac{\omega_{\mathcal{T}}}{|\mathcal{T}|}\|D_{\mathcal{T}}x - TG\|_2^2 + \frac{\omega_{\mathcal{C}}}{|\mathcal{C}|}\|D_{\mathcal{C}}x - CG\|_2^2 + \frac{\omega_{\mathcal{N}}}{|\mathcal{N}|}\|D_{\mathcal{N}}x - NG\|_2^2 \\ \text{s.t.} & x \geq 0, \end{array} \tag{5}$$

where ω has the same meaning as in (3). The function calculates the weighted sum of average squared deviation from the prescribed dose for each organ.

Alternatively, we could only penalize the overdose part for the critical organs and normal tissue, this change would turn the objective into:

$$\begin{array}{ll} \min & \frac{\omega_{\mathcal{T}}}{|\mathcal{T}|}\|D_{\mathcal{T}}x - TG\|_2^2 + \frac{\omega_{\mathcal{C}}}{|\mathcal{C}|}\left\|(D_{\mathcal{C}}x - CG)_+\right\|_2^2 + \frac{\omega_{\mathcal{N}}}{|\mathcal{N}|}\left\|(D_{\mathcal{N}}x - NG)_+\right\|_2^2 \\ \text{s.t.} & x \geq 0, \end{array} \tag{6}$$

where $(\cdot)_+ = \max\{0, \cdot\}$.

There are two ways to deal with dose-volume constraints in nonlinear programming. One way is by adding a penalty term to the original objective function, this is the most popular way. See Cho et al. 1998, Spirou and Chui 1998 and Wu and Mohan 2000 for reference. All these models adopt a volume sensitive penalty function.

Another way to fulfill the dose-volume constraints in nonlinear programming models is to include the constraints in the optimization algorithm. Cho et al. 1998 use a method that is based on the theory of projections onto convex sets (POCS) in which the dose-volume constraint is replaced by a limit on integral dose. Furthermore, Starkschall, Pollack, and Stevens 2001 use a dose-volume feasibility search algorithm to find the solution and Dai and Zhu 2003 discuss two techniques to convert dose-volume constraints to dose limits.

Quadratic objective functions have become an accepted standard. In general, they produce satisfactory plans. But the weakness of quadratic objective functions cannot be neglected. Weighting factors ω have no clinical meaning and the choice is quite arbitrary. Therefore, normally several plans for different choices of weighting factors would have to be tried before a final plan is selected. Moreover, for general functions, the optimizer can only guarantee that the solution is locally optimal.

3.3 Mixed Integer Programming

It is very hard to achieve dose-volume control using pure linear programming. However, by introducing binary variables into the model it is very simple to impose dose-volume constraints. By setting these binary variables to 0 or 1, it is possible to enumerate the number of voxels which receive a dose lower than a threshold or higher than a threshold in a specified structure, thus it is possible to model dose-volume constraints. These additional variables turn a linear programme into an MIP.

Here, we will take an example to show how to convert a dose constraint to a dose-volume constraint. In linear programming, a single upper bound on the critical organ may be imposed as follows:

$$d(v)\,x \leq CUB_v \quad \forall v \in \mathcal{C}. \tag{7}$$

Sometimes, such a constraint can bring infeasibility to the problem due to its strictness. By specifying an overdose fraction F and a volume fraction P, this constraint can be relaxed to a dose-volume constraint:

$$\begin{array}{rcl} d(v)\,x & < & (1 + y_v F)CUB_v \quad \forall v \in \mathcal{C} \\ \sum_{v \in |\mathcal{C}|} y_v & < & P|\mathcal{C}|, \end{array} \tag{8}$$

where y_v is a binary variable.

For example, suppose the upper bound for the lung is 15Gy, i.e. $CUB_v = 15$. The dose-volume constraint is that "no more than 40% volume of the lung can exceed a radiation dose of 20Gy", then the overdose fraction F should equal to 33% and P is equal to 40%.

The first MIP incorporating DVH constraints was proposed by Langer et al. 1990. In that work, the authors use MIP to find the beam intensities of wedged and open beams in which dose-volume constraints are considered. Although the application is only for conformal radiotherapy not IMRT, it is straightforward to extend this to IMRT.

Volume-based objective functions can also be used in MIP models. Bednarz et al. 2004 propose volume-based objective function in which the number of under- or overdosed voxels in selected critical structures and/or targets is to be minimized. By minimizing the objective function, the model obtains a better control of the over- or underdosed volumes in critical structures and targets. It also circumvents the interactive and judicial adjustment of the DVH values in order to obtain a specific solution.

Moreover, MIP can be used for solving some combined problem. For example, Lee, Fox, and Crocker 2003 and Wang, Dai, and Hu 2003 optimize beam angle and beam intensity at the same time. Bednarz et al. 2002 and Preciado-Walters et al. 2004 consider beam intensity and beam delivery together.

Undoubtedly, MIP frameworks significantly enhance the model flexibility. The flexibility allows us to built more complicated constraints and objectives, thus we could use it to solve more complicated problems. However, usually MIP is solved by branch and bound algorithms with the simplex method. For the typical size of clinically relevant MIP problems with many thousands of integer variables, finding an optimal solution to such models in an acceptable amount of computational time is very hard. Therefore, execution time can be a problem for MIP approaches.

3.4 Multiple Objective Programming (MOP)

In radiotherapy, the desired dose distribution can not always be obtained, due to physical limitations and to the existence of trade-offs between the various conflicting treatment goals. This multiobjective character of inverse planning has been recognized only in the last 5 years even though radiotherapy has been applied for several decades.

One way to solve a multiple objective (MO) problem is by transforming it into a single objective problem using a specific set of weighting factors for each objective. This is the most prevalent method currently, it is called a priori method for solving MO optimization problems. From this point of view, we can say that the nonlinear models (see (5) and (6)) and linear programming models (see (3) and (4)) are actually a scalarization of multiple objective problems. The problem of this approach is that the weighting factors have no clinical meaning and their relationship to the solution is not known in advance. Therefore, in order to obtain some satisfactory solution, the treatment planner is required to often repeat the optimization with other weighting factors.

Alternatively, an a posteriori method in which an optimization engine will obtain a representative set of the entire Pareto (efficient) set is more preferable compared to the a priori method. Pareto optimal solutions are defined as solutions in which an improvement in one objective will always lead to a worse result in at least one of the other objectives. The Pareto set is used to analyze the trade-offs between the objectives in conflict and then to select a solution that satisfies simultaneously all objectives. Analysis could be done after the optimization and the optimization is not required to be repeated again.

MOP approaches in IMRT can be divided into two categories, multiple objective nonlinear programming (MONP) and multiple objective linear programming (MOLP).

MONP is based on nonlinear programming models. The first MOP model for IMRT is proposed by Cotrutz et al. 2001. The authors obtain a set of optimal solutions by using different weighting factors. A conjugate gradient algorithm is used to solve the problem. To eliminate the negative beam intensity problem, the problem is converted to an unconstrained problem by using the square root of beam intensities as decision variables. Based on Cotrutz et al. 2001, Lahanas, Schreibmann, and Baltas 2003 use L-BFGS to solve the problem and conclude that globally Pareto optimal solutions can be found using L-BFGS by comparing the results with the fast simulated annealing algorithm. Moreover, Lahanas et al. 2003 use the evolutionary algorithms NSGA-II and NSGA-IIc to improve the optimization process.

The objectives in Lahanas, Schreibmann, and Baltas 2003 can be described as follows:

$$\begin{aligned}
F &= (f_{\mathcal{T}}, f_{\mathcal{N}}, f_{\mathcal{C}}) \\
f_{\mathcal{T}} &= \frac{1}{|\mathcal{T}|} \|D_{\mathcal{T}}x - TG\|_2^2 \\
f_{\mathcal{N}} &= \frac{1}{|\mathcal{N}|} \|D_{\mathcal{N}}x\|_2^2 \\
f_{\mathcal{C}} &= \frac{1}{|\mathcal{C}|} \left\|(D_{\mathcal{C}}x - CG)_+\right\|_2^2,
\end{aligned} \tag{9}$$

where $(\cdot)_+ = \max\{0, \cdot\}$, $f_{\mathcal{T}}$ is the average squared deviation from the prescribed dose to the tumor, $f_{\mathcal{C}}$, is the average squared overdose to the critical organ and $f_{\mathcal{N}}$ is the average squared dose to the normal tissue.

Based on linear programming models, researchers formulate the beam intensity optimization using MOLP (see, e.g., Küfer and Hamacher 2000; Ehrgott and Burjony 2001; Hamacher and Küfer 2002; Küfer et al. 2003).

A simple MOLP model can be described as follows:

$$\begin{array}{lll}
\min & F = (f_{\mathcal{T}}, f_{\mathcal{C}}, f_{\mathcal{N}}) & \\
\text{s.t.} & D_{\mathcal{T}}x & \geq TLB\,(1 - f_{\mathcal{T}})\, e_{|\mathcal{T}|} \\
& D_{\mathcal{C}}x & \geq CUB\,(1 - f_{\mathcal{C}})\, e_{|\mathcal{C}|} \\
& D_{\mathcal{N}}x & \geq NUB\,(1 - f_{\mathcal{N}})\, e_{|\mathcal{N}|} \\
& f_{\mathcal{T}}, f_{\mathcal{C}}, f_{\mathcal{N}} & \geq 0 \\
& x & \geq 0,
\end{array} \tag{10}$$

where $f_{\mathcal{T}}$ is the maximal deviation from the prescribed dose for the tumor, $f_{\mathcal{C}}$ is the maximal deviation from the prescribed dose for the critical organ and $f_{\mathcal{N}}$ is the maximal deviation from the prescribed dose for the normal tissue.

Solving MOLP is much easier than solving MONP. There is some efficient method to find a balanced solution for MOLP (Ehrgott and Burjony 2001). Also, a unifying mathematical framework that allows for a comparison of different models via the comparison of the corresponding sets of Pareto optimal plans is proposed by Romeijn, Dempsey, and Li 2004.

Most of the papers use the weighted sum method, which is not very effective. First of all, it is not very efficient because the running time to solve the problem is proportional to the number of weighting factors. For example, if we want to get 30 Pareto optimal plans, it will take around 10 hours to find those solutions if finding one solution needs 20 minutes. Secondly, all models in MONP now use gradient based solution algorithm, which can not guarantee that the obtained solutions are indeed Pareto optimal if there exist local minima. Moreover, because the choice of the weighting factors is quite arbitrary, these discrete solutions may not fully represent the whole Pareto optimal solution space. For solving the problem, i.e. finding a representative set of the Pareto optimal solutions, there is still no effective way. New methods need to be proposed. Additionally, new models which can include dose-volume constraints need to be developed.

4 Conclusion

In this paper, we review the mathematical models used in the beam intensity optimization problem of IMRT. We classify these models into linear programming models, nonlinear programming models, mixed integer programming models and multiple objective models. Recently, there are some improvements in linear programming models addressing the infeasibility problem and introducing new ways to address dose-volume constraints. For nonlinear programming, gradient based solution methods will only achieve local optima, while stochastic algorithms like simulated annealing are slow. MIP frameworks significantly enhance the modeling flexibility, and allow for the incorporation of many different objectives and constraints into the optimization model, but the computation time will be a problem. By analysing these models, we find that the most existing nonlinear, linear programming, and mixed integer programming models are multiobjective in nature due to the weighting factors. In MOP, instead of specifying the weighting factor, a representative set of Pareto optimal solutions should be computed for the planner to choose from. Currently, there are no effective solution methods for MOP and new methods need to be developed.

References

Bahr, G.K., J.G. Kereiakes, H. Horwitz, R. Finney, J.M. Galvin, and K. Goode. 1968. "The method of linear programming applied to radiation treatment planning." *Radiology* 91:686–693.

Bednarz, G., D. Michalski, P.R. Anne, and R.K. Valicenti. 2004. "Inverse treatment planning using volume-based objective functions." *Physics in Medicine and Biology* 49:2503–2514.

Bednarz, G., D. Michalski, C. Houser, M.S. Huq, Y. Xiao, P.R. Anne, and J.M. Galvin. 2002. "The use of mixed-integer programming for inverse treatment planning with pre-defined field segments." *Physics in Medicine and Biology* 47:2235–2245.

Cho, P.S., S. Lee, R.J. Marks, S. Oh, S.G. Sutlief, and M.H. Phillips. 1998. "Optimization of intensity modulated beams with volume constraints using two methods: Cost function minimization and projections onto convex sets." *Medical Physics* 25:435–443.

Cotrutz, C., et al. 2001. "A multiobjective gradient-based dose optimization algorithm for external beam conformal radiotherapy." *Physics in Medicine and Biology* 46 (8): 2161–2175.

Dai, J., and Y. Zhu. 2003. "Conversion of dose-volume constraints to dose limits." *Physics in Medicine and Biology* 48:3927–3941.

Ehrgott, M., and M. Burjony. 2001. "Radiation therapy planning by multicriteria optimisation." *Proceedings of the 36th Annual Conference of the Operational Society of New Zealand*. 244–253.

Hamacher, H.W., and K.-H. Küfer. 2002. "Inverse radiation therapy planing – A multiple objective optimization approach." *Discrete Applied Mathematics* 118 (1-2): 145–161.

Holder, A. 2003. "Designing Radiotherapy Plans with Elastic Constraints and Interior Point Methods." *Health Care Management Science* 6 (1): 5–16.

Küfer, K.-H., and H.W. Hamacher. 2000. "A multicriteria optimization approach for inverse radiotherapy planning." Edited by W. Schlegel and T. Bortfeld, *XIIIth International Conference on the Use of Computers in Radiation Therapy.* Heidelberg, Germany: Springer Verlag, Berlin, 26–28.

Küfer, K.-H., A. Scherrer, M. Monz, F. Alonso, H. Trinkaus, T. Bortfeld, and C. Thieke. 2003. "Intensity-modulated radiotherapy – A large scale multi-criteria programming problem." *OR Spectrum* 25:223–249.

Lahanas, M., E. Schreibmann, and D. Baltas. 2003. "Multiobjective inverse planning for intensity modulated radiotherapy with constraint-free gradient-based optimization algorithms." *Physics in Medicine and Biology* 48 (17): 2843–2871.

Lahanas, M., E. Schreibmann, N. Milickovic, and D. Baltas. 2003. "Intensity Modulated Beam Radiation Therapy Dose Optimization with Multi-objective Evolutionary Algorithms." Technical Report, Department of Medical Physics and Engineering, Strahlenklinik, Klinikum Offenbach, Germany.

Langer, M., R. Brown, M. Urie, J. Leong, M. Stracher, and J. Shapiro. 1990. "Large scale optimization of beam weights under dose-volume restrictions." *International Journal of Radiation Oncology, Biology, Physics* 18:887–893.

Lee, E.K., T. Fox, and I. Crocker. 2003. "Integer Programming Applied to Intensity-Modulated Radiation Therapy Treatment Planning." *Annals of Operations Research* 119:165–181.

Lim, J., M. C. Ferris, and S. J. Wright. 2002. "An Optimization framework for conformal raiation treatment planning." Technical Report, Computer Sciences Department, University of Wisconsin, Madison. http://www.cs.wisc.edu/ ferris/papers/conformal.pdf.

Merritt, M., and Y. Zhang. 2002. "A Successive Linear Programming Approach to IMRT Optimization Problem." http://www.caam.rice.edu/ zhang/reports/tr0216.pdf.

Preciado-Walters, F., R. Rardin, M. Langer, and V. Thai. 2004. "A coupled column generation, mixed integer approach to optimal planning of intensity modulated radiation therapy for cancer." *Mathematical Programming, Series B* 101:319–338.

Romeijn, H.E., R.K. Ahuja, J.F. Dempsey, A. Kumar, and J.G. Li. 2003. "A novel linear programming approach to fluence map optimization for intensity modulated radiation therapy treatment planning." *Physics in Medicine and Biology* 48:3521–3542.

———. 2005. "A new linear programming approach to radiation therapy treatment planning problems." *Operations Research.* to appear.

Romeijn, H.E., J.F. Dempsey, and J.G. Li. 2004. "A unifying framework for multi-criteria fluence map optimization models." *Physics in Medicine and Biology* 49:1991–2013.

Shepard, D.M., M.C. Ferris, G.H. Olivera, and T.R. Mackie. 1999. "Optimizing the Delivery of Radiation therapy to Cancer Patients." *SIAM Review* 41 (4): 721–744.

Spirou, S. V., and C.-S. Chui. 1998. "A gradient inverse planning algorithm with dose-volume constraints." *Medical Physics* 25 (3): 321–333.

Starkschall, G., A. Pollack, and C. W. Stevens. 2001. "Treatment planning using a dose-volume feasibility search algorithm." *International Journal of Rdaiation Oncology, Biology, Physics* 49 (5): 1419–1427.

Wang, C., J. Dai, and Y. Hu. 2003. "Optimization of beam orientations and beam weights for conformal radiotherapy using mixed integer programming." *Physics in Medicine and Biology* 48:4065–4076.

Wu, Q., and R. Mohan. 2000. "Algorithms and functionality of an intensity modulated radiotherapy optimization system." *Medical Physics* 27 (4): 701–711.

Xing, L., R.J. Hamilton, D. Spelbring, C.A. Pelizzari, G.T. Chen, and A.L. Boyer. 1998. "Fast iterative algorithms for three-dimensional inverse treatment planning." *Medical Physics* 25:1845–1849.

Senge meets Goldratt: Combining Theory of Constraints with Systems Thinking

Vicky Mabin, John Davies
Victoria Management School
Victoria University of Wellington
New Zealand
vicky.mabin@vuw.ac.nz

and

James F Cox III
Terry College of Business
University of Georgia at Athens
USA

Abstract

This presentation briefly overviews current research exploring the combination of Theory of Constraints (TOC) with Systems Thinking, in particular TOC's Thinking Process tools with Senge's Systems Archetypes.

We provide examples of TOC being used with systems methods such as Causal Loop Diagrams to explore the nature of Senge's Systems Archetypes. We start by reviewing our work building on Senge's Fix That Fails archetype, using the TOC Thinking Process tools of Evaporating Cloud and Negative Branch Reservation. We then sketch some of the other archetypes, such as Shifting the Burden, Tragedy of the Commons, Success to the Successful and Limits to Growth. In each case we indicate how TOC Thinking Process tools may be used to take these problematic situations and start developing more of an understanding about the characteristics of the problem and possible solutions.

Workflow Analysis, Risk Management and Laboratory Information Management Systems for Genome Research Projects

Megan Too, Allen Rodrigo
Bioinformatics Institute
and John Paynter
Department of Information Systems and Operations Managagement
University of Auckland, New Zealand
megantoo@gmail.com, a.rodrigo@auckland.ac.nz; j.paynter@auckland.ac.nz

Abstract

Genome research is a growth field in biological science. The knowledge derived may be used as input to other research projects and can be applied to create technology valuable to society. The value of research outcomes depends on the research process. Genome research is subject to uncertainty and so the risk of unplanned events may affect research outcomes.

A genome research project was used as a generalisable model. The techniques applied were information requirements analysis, process modelling and risk analysis. Genome research was found to be collaborative in nature with information users spatiotemporally separated; users share some information, but have different specific information uses. Thus a laboratory information management system (LIMS) should promote information sharing but allow customised user-views. Knowledge management is needed to support knowledge discovery, transfer and application. Workflow analysis revealed many dependencies. The most significant dependency occurred between genome assembly, annotation and sequence-reliant experiments. Risk analysis indicated a variety of risk sources and a complex network of effects, risk impact may be intangible and thus difficult to gauge. Risk however may be controlled through information management, process management and human resource management. A risk management plan was developed for the genome sequence assembly process as an example.

1 Introduction

Genomics is a popular discipline in contemporary biological research. A genome is defined as the complete set of deoxyribonucleic acid (DNA) in a group of chromosomes (Griffiths, Miller, Suzuki, Lewontin, & Gelbart, 2000). The first free-living organism to be genome sequenced was the bacterium Haemophilus influenzae in 1995 (Fleishmann et al., 1995). Since then, many other bacteria and model organisms have been sequenced including the Baker's yeast (Goffeau et al., 1996), nematode worm (The C.elegans Sequencing Consortium, 1998), fruit-fly (Myers et al., 2000), mustard weed (The Arabidopsis Genome Initiative, 2000), mouse (Abril et al., 2002) and rat (Gibbs et al., 2004). Additionally, the human genome was sequenced by two separate groups – the International Human Genome Sequencing Consortium, whom worked on the Human Genome Project (HGP), and private company Celera Genomics. Both parties completed "working draft" versions of the human genome in 2001 (International Human Genome

Sequencing Consortium, 2001; Venter et al., 2001). Completion of these genome projects represents a paradigm shift in biological research. Genomics is based on a fundamental principle whereby knowing the entire genome sequence of an organism will form the basis to understand organism function. Hence in popular press, a genome sequence is often referred to as the “blueprint” of an organism. However, the genome sequence alone is not sufficient to understand function. It is vital to discover how gene expression is controlled and gene products interact in order to produce a phenotype. Thus, genomic research involves genome sequencing and related studies on the structure and function of a genome. It is important then to distinguish between “structural genomics” and “functional genomics”. The goal of structural genomics is to discover the physical composition of all the DNA, ribonucleic acid (RNA) and proteins in an organism (Griffiths et al., 2000). Hence genome sequencing is one form of structural genomics. In contrast, the aim of functional genomics is to decipher how gene products function alone and interact as a system in order to produce phenotype, and thus the function, of an organism (Eisenberg, Marcotte, Xenarios, & Yeates, 2000; Pandey & Mann, 2000). The common goal of all genome projects is to sequence the organism genome whereas the choice of which functional studies to pursue is variable (Vukmirovic & Tilghman, 2000). Moreover, the chosen functional studies will determine which research techniques are applied and the scientific outcomes of the project. Hence the process of genome research and scientific value of outcomes vary between projects.

It is clear that each genome project is centred on genome sequencing and involves some form of functional genomics research. But what makes one genome project more successful than another project? The success of research can be defined in terms of project process and project outcomes. Project process (or workflow) is the set of activities that contribute to the project, and relationships between activities. Workflow depends on the availability of information and physical resources. Moreover, the specific process followed by a research group affects project budget and schedule. Thus a successfully managed process would minimise research cost and time to completion. Timeliness is especially crucial in scientific research because investigators aim to be the first to publish their discoveries. Workflow is also subject to uncertainty and therefore the risk of unplanned events. Hence the plasticity of research and ability to cope with unforeseen events can contribute to project success. Another viewpoint of success relates to project outcomes. The outcome of genome research may be deemed successful if project goals have been achieved and research findings are valuable to wider society. However, outcomes are dependent on project process and so effective project management is required to optimise outcomes.

A microbial ecological genomics (eco-genomics) research project is being conducted at the University of Auckland (UoA). For the purpose of this honours research, the eco-genomics project was used as a model for genome projects in general. The organism of interest is a bacterium called *Acidovorax temperans* (strain CB2) and may be of functional significance in the wastewater-treatment industry. The study is the first genome research project to be conducted at the UoA.

2 Laboratory Information Management System (LIMS)

Information management is an important part of project management. The project process is affected by the information system, because information serves as input to and output from each project task. Additionally, the outcome of research is dependent on information quality (correctness) and usability. Usability describes the extent that information can be used for the intended purpose of the user (Turban, King, Lee, Warkentin, & Chung, 2003) – in the case of scientific research, the main purpose is knowledge discovery. Hence effective information management can ensure quality control and extend the usability of information, thus optimising the research process and outcome.

A LIMS can support the scientific and administrative needs of those involved in research. However, the most suitable LIMS design is one based on the core information requirements of the scientific organisation. Thus it is important to understand information requirements so that a LIMS can be tailored to organisational needs.

3 Methods

Participants of the UoA eco-genomics project were interviewed to gain background information on genome research projects. Interviewing was used as a research method because it allows collection of rich, detailed information and the chance to clarify questions or answers (Hoffer, George, & Valacich, 2002). The people interviewed were: Management-level researchers, who have scientific expertise and oversee strategic direction of the project; Operations-level researchers, who conduct experiments and other tasks required for scientific discovery. Interviewees were asked questions related to: Responsibilities of people, Project activities, Experimental processes, input and output, Recording of information, Sample tracking, Use of information resources, Collaborative links and communication methods.

Examples of input and output from experiments were observed. The technologies used by the eco-genomics team were also examined. The purpose of this was to support interviewing as a means to understand the information and technical resources and to see how users interact with technologies.

Data gathered from interviews and observation were structured in order to filter out irrelevant facts and reorganise the data for analysis. Data were structured and analysed via the following techniques: Stakeholder analysis, Use-case modelling, Process modelling, Risk analysis.

System models were evaluated to allow proposal of project management techniques. The proposed project management techniques address information management, process management, human resource management and risk management.

The process models were illustrated as process diagrams, drawn using Visio™ Professional 2002 software package. (Visio Professional 2002)

4 Processes

Every process can be broken down into sub-processes that describe the flow of events in more detail. The scope of workflow analysis would be huge if every minute process was modelled. Hence for the purpose of this project, only a few key processes were modelled and analysed. Figure 1 shows one such example. The process models show the interaction between genome project activities (Figure 2).

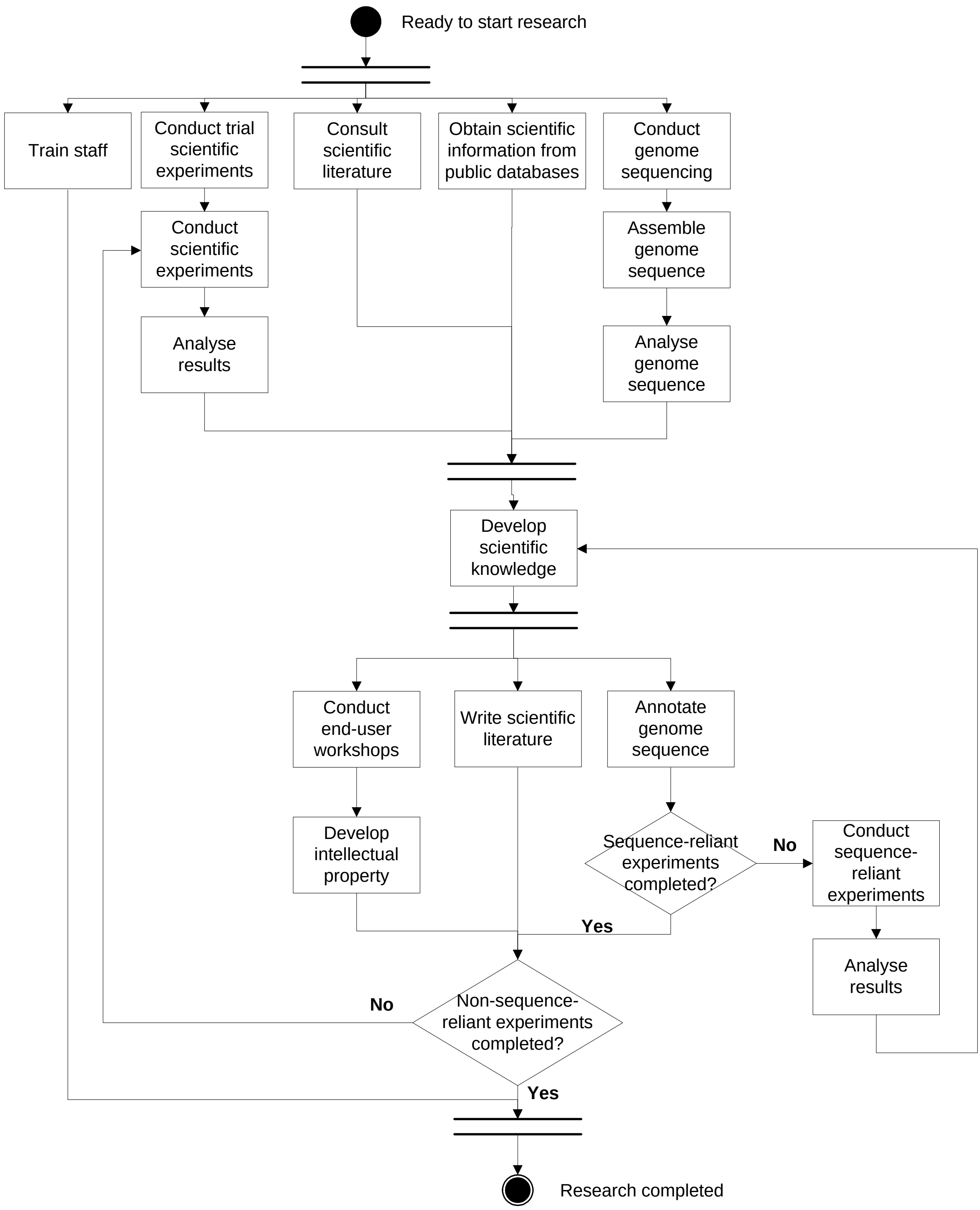


Figure 1: "Perform Research" Process Model

Additionally, the models highlight interdependencies and therefore potential sources of risk. Future research could involve model refinement to describe workflow in more detail. Quantitative analysis and simulation could then be applied. A quantitative model might show how each alternative process pathway affects project cost, schedule or outcomes. Simulation can then be applied to change the order of events and find an optimal process. However, to construct a reliable quantitative model requires a solid data source. Thus it may be best to gather data from experienced genome research institutes

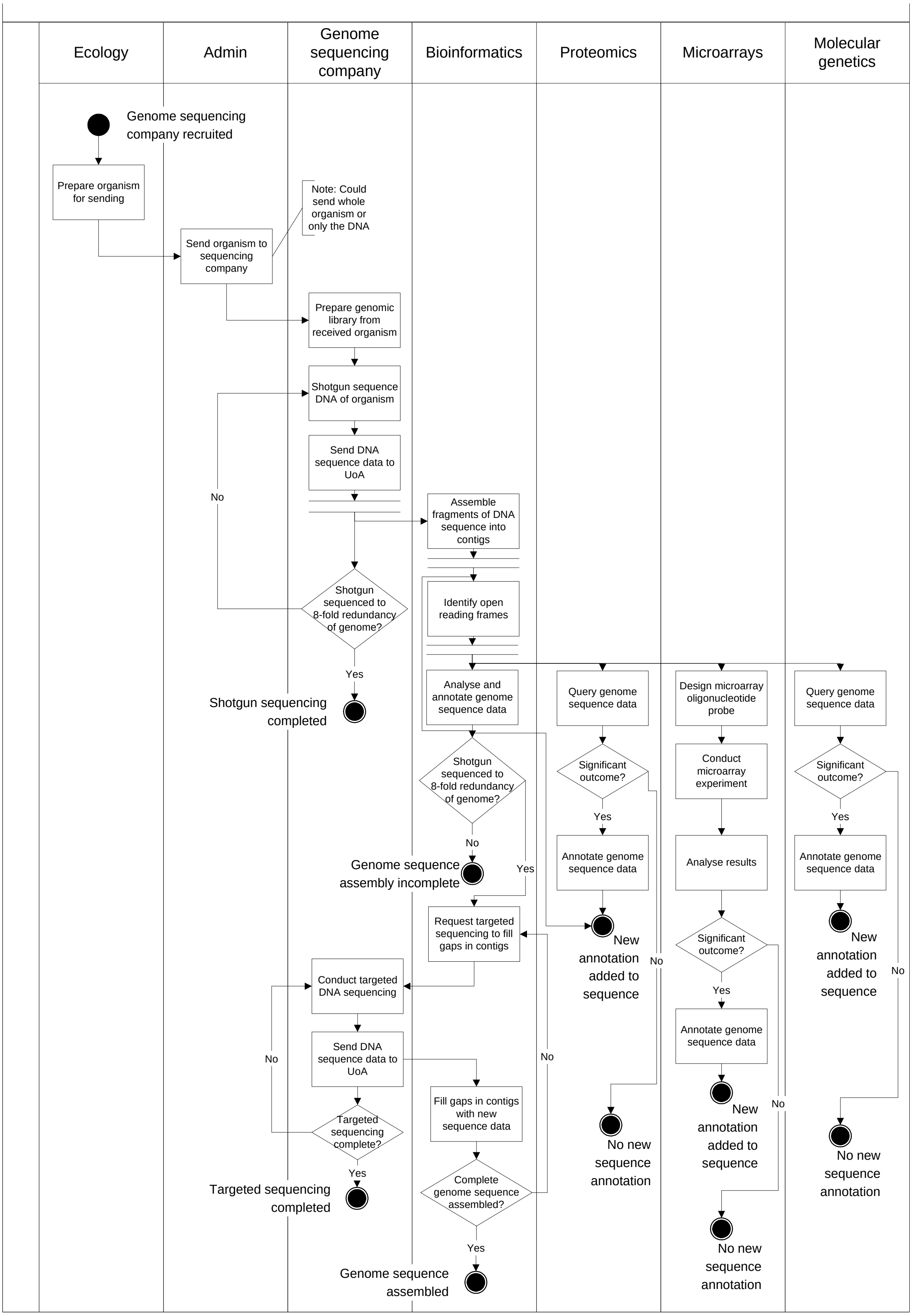


Figure 2: Genome Sequencing, Assembly and Annotation Cross-Functional Unit Process Model

5 Risk Analysis

So far information requirements for genome research were identified and the research workflow modelled. Now, the information requirements are united with process models in order to reveal project risk. Risk occurs when there is lack of information and thus uncertainty (Sommerville, 1996). There is no simple definition of risk; it is often described as the chance that an unplanned event will occur and cause an adverse effect (Christensen & Thayer, 2001). However, risk may also allow the chance to gain, such as when a decision is made and there are a range of benefits for each possible outcome (Frame, 2003). It is important to manage risk because every project is subject to uncertainty and so may encounter an unplanned event. Risk management involves planning for risk, monitoring risk and controlling risk (Frame, 2003). Thus, risk management is a proactive approach that can enhance the ability to cope with risk. The first step in risk management is to perform risk analysis. The purpose of risk analysis is to discover and understand project risk and thus allow development of risk control methods (Christensen & Thayer, 2001). We pay specific attention to the genome sequence assembly process.

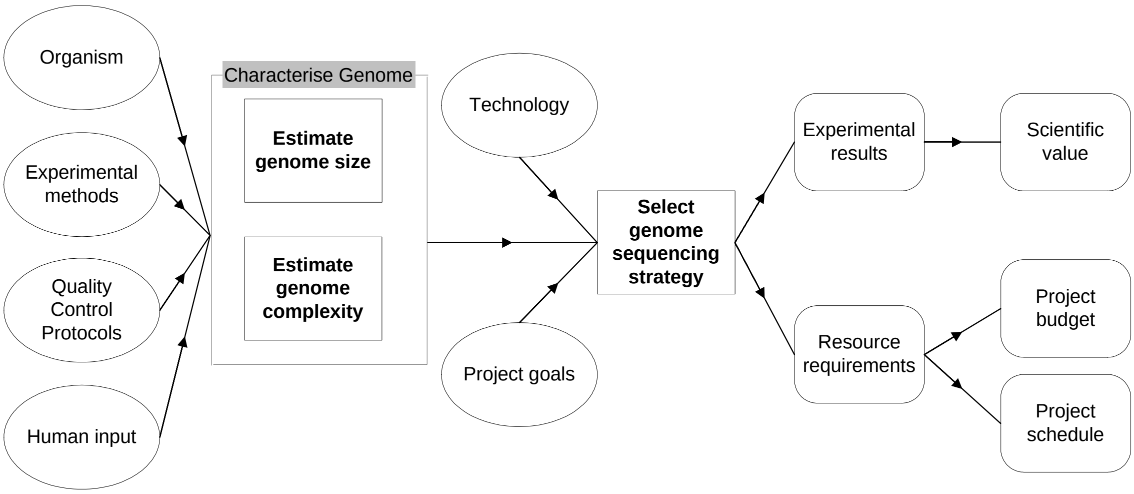


Figure 3: Risk Model Centred on Genome Characterisation Process

Each activity is dependent on many risk input factors. For example, Figure 3 shows that to characterise genome, the activities “Estimate genome size” and “Estimate genome complexity” are conducted. The process and outcome of genome characterisation are dependent on the organism whose genome is characterised, experimental methods, quality control protocols and human input to the activities. Thus, a range of risk management techniques are required to ensure the outcome of genome characterisation is of acceptable quality and the process does not cost more or take longer than planned.

5.1 Risk Models

The following points can be made from the risk models.

5.1.1 Multiple risk input factors

Each activity is dependent on many risk input factors. For example, Figure 3 shows that to characterise genome, the activities “Estimate genome size” and “Estimate genome complexity” are conducted. The process and outcome of genome characterisation are dependent on the organism whose genome is characterised, experimental methods,

quality control protocols and human input to the activities. Thus, a range of risk management techniques are required to ensure the outcome of genome characterisation is of acceptable quality and the process does not cost more or take longer than planned.

5.1.2 Complex network of effects

Risk effect is complex and involves many impact factors. For example, Figure 10 (Too, 2004) illustrates risk factors to the activities “Select organism to study” and “Define scientific objectives”. The outcome of these activities affects the scientific value of research and many other downstream factors, such as the ability to recruit staff or obtain future funding. Furthermore, risk impact factors may act as additional cause of risk because they influence each other. For instance, collaborative partnerships and scientific community profile are mutually dependent. The complexity of risk impact means it may be difficult to evaluate risk in quantitative terms. However, despite the problem with quantification, complexity reinforces the need for a range of risk management methods.

5.1.3 Intangible effects

Some risk impact factors are intangible. For example, Figure 12 (Too, 2004) shows a risk model for the activity “Transfer biological sample”. A biological sample may be transferred from one location to another. The sample transfer activity may result in loss of the sample, causing social implications. For instance, sample loss could result in political arguments over who is to blame for the loss and the arguments may then decrease employee morale. However, social impact and employee morale are intangible and so may not be easy to measure or predict. Thus proxy measures may be required to quantify risk impact on intangible factors.

5.1.4 External risk sources

Risk may originate from a source external to the UoA. For example, Figure 13 (Too, 2004) shows risk centred on the activity “Select genome sequencing company”. The genome sequencing company is an external party that may have control over risk impact factors, such as information quality. Thus the quality of information is at risk because it depends on the genome sequencing company. For instance, should the sequencing company have poor quality control protocols, biological samples may be contaminated and quality of genome sequence information will decrease. The UoA researchers may not be able to prevent risk that comes from an external source. However, a risk management plan can have strategies to mitigate adverse effects should they occur.

5.2 Risk Analysis for Genome Sequence Assembly

Detailed risk analysis was performed for the activity “Assemble genome sequence”. Figure 14 shows the corresponding risk cause-and-effect model. In addition, Table 1 describes each risk input factor and the impact of risk on time taken, cost and fidelity of genome sequence assembly. Fidelity is defined as the degree that the assembled genome sequence is a true representation of the DNA sequence of the underlying organism genome (from which sequence reads were derived) (Schmutz et al., 2004). The higher the fidelity, the greater the similarity between the assembled genome sequence and the true organism genome sequence. Too (2004) outlines in Table 8 a summary of risk management strategies that may be applied to prevent or mitigate risk.

6 Discussion

Genome research is subject to many risks. Risk may come from a variety of sources and have a complex network of effects. Furthermore, several risks require human judgement as an input factor and thus offer the potential to gain. In order to control risk, mitigation or prevention strategies are required. A risk control plan was proposed for the genome sequence assembly process. However, risk management should be an ongoing task that is carried out during the entire research project. Effective risk management requires constant monitoring of risk, application of risk control strategies and revision of the risk management plan as necessary.

Future research could involve quantitative risk analysis for genome projects. Quantitative analysis would describe the probability of risk occurrence and express risk impact in numerical terms. A quantitative model may be used for scenario analysis and risk prediction. Sound historical data or expert judgement is required to define numerical values for a quantitative model. It may also be difficult to quantify risk impact because some risks are intangible. Hence the utility of a quantitative model may depend on which risk factors are addressed in the model.

7 Conclusion

Results from information requirements analysis, workflow analysis and risk analysis are drawn together and recommendations are made for genome project management. Information management is required for successful project management. Information requirements were analysed and LIMS features are proposed to provide an information management strategy. The most significant findings were as follows:

7.1 Stakeholders and Collaboration

A range of people are affected by genome research and they may have direct or indirect contribution to the information system. Thus when a LIMS is designed, it is vital to consider how the LIMS caters to all stakeholders. Genome research is collaborative and spans organisational boundaries. Information system users may be separated in time and location, and may require information privacy. Thus it is important for a LIMS to facilitate information sharing but not compromise the security of information.

7.2 User-interface

A LIMS should have customised user-interfaces in order to address specific needs of diverse users. Furthermore, an ideal LIMS would allow users to interact in a manner that mimics natural research processes. For instance, mobile computers and audiovisual technology would suitably integrate with traditional scientific research methods.

7.3 Knowledge-based systems

The goal of science is to discover new knowledge and in scientific research, knowledge serves as both input and output to the research process. Thus a LIMS should support knowledge capture, use and transfer. In addition, a knowledge base would enable a scientific organisation to retain knowledge derived from its employees, should they leave the organisation.

Risk input factor	Impact of risk eventuating
Human Input	
Number of people doing assembly – Risk choosing too many or too few people to work on genome sequence assembly.	Too many people cause unnecessary cost because employees must be paid. Too few people may decrease efficiency of sequence assembly because the assembly process is dependent on the skill and availability of the employees. Thus without enough employees, assembly may take longer than planned.
Decision required to set quality standards for the completeness, accuracy and contiguity of the assembled genome sequence	The decision required for setting quality standards involves a trade off between cost, time and fidelity. The optimal quality standards would balance between an aim to minimise cost, minimise time and maximise fidelity. Thus, a sub-optimal decision will cause unnecessary cost or time, or decrease fidelity.
Risk of incorrect human judgement in sequence assembly process	Incorrect judgement will decrease fidelity of assembled sequence. Error in assembly may cause error in subsequent analyses that rely on the assembly. May increase time and cost due to need to identify and fix error.
Sequence data content	
Format of sequence data when it is received from sequencing company – risk of format being incompatible with in-house systems.	Sequence data may require reformatting to allow compatibility with in-house systems. This will increase time and cost.
Risk of poor quality sequence data being used in genome sequence assembly process	Poor quality data may decrease fidelity of assembled sequence. May increase time and cost due to need to identify and fix error.
Number of different sizes of genomic DNA libraries used in order to generate DNA sequence data (e.g. Three different insert libraries could be used, with 2kb insert, 10kb insert and 50kb insert sizes)	This risk is a trade off between cost, time and fidelity. A greater number of libraries and types of libraries will increase cost and increase time for sequencing. However, more types of libraries will provide additional information. This may decrease time for assembly and increase fidelity (e.g. by making it easier to resolve problems with repetitive sequences and genome order).
Sequence data delivery process	
Method used by sequencing company in order to deliver DNA sequence data (e.g. copy computer files to compact disc and send by traditional mail, download direct from company website)	Delivery method affects time taken and resources required to obtain data. A suboptimal method will increase time and cost. Opportunity for data intercept may also impact data security.
Quantity and frequency of DNA sequence data delivery in each batch update from sequencing company	Insufficient data quantity and/or infrequent delivery may inhibit ability to assemble genome. Thus, may increase time and cost of assembly.
Target organism of genome project – inherent properties	

Risk input factor	Impact of risk eventuating
Whether organism of interest in the study is a prokaryote or eukaryote	A eukaryote will generally have a more complex genome than a prokaryote. Thus, choosing a eukaryote over a prokaryote may increase time and cost of project.
Size (number of base-pairs) of genome being assembled	In general, the larger the genome, the greater the time and cost of sequence assembly.

Table 1: Risk Input Factors and Risk Impact for Genome Sequence Assembly Process

7.4 Workflow Analysis

Workflow analysis was performed to understand the project process. Several process models were created to illustrate interrelationships between activities. Moreover, process models highlight dependencies that may cause risk. In genome research, the most significant dependency is probably that between genome sequence assembly, annotation and sequence-reliant experiments. Genome sequence data serves as input to sequence-reliant experiments. However, sequence-reliant experiments produce scientific information that feeds back to the genome sequence via annotation. The process loop is a potential source of amplified error that can affect future research and so may be considered as high-risk. Hence assembly and annotation are important issues that should be addressed in risk management. Future research could seek to develop quantitative process models. Quantitative models may then be used in simulation to find optimal processes.

7.5 Risk Analysis

Risk occurs when there is absence of information or lack of control over an event. Several risk models were developed and they show that risk may come from many sources and form a complex network of effects. In addition, many risks in genome research create not only the chance for damage but the opportunity to enhance research outcomes, or reduce research time or cost. Detailed risk analysis was performed for the genome sequence assembly process. Information management, process management and human resource management strategies were given in a risk control plan. However, extensive risk analysis for other project activities and ongoing risk monitoring are required for effective management of the entire genome project.

In summary, genome research can yield valuable information for the scientific community and general public. However, research outcomes depend on the research process, which is subject to uncertainty and therefore risk. With knowledge of information requirements and project workflow, a risk management plan can be developed and applied to manage the research process and thus enhance research outcomes.

8 References

Abril, J. F., Agarwal, P., Alexandersson, et al. (2002). Initial sequencing and comparative analysis of the mouse genome. Nature, 420(6915), 520-562.

Christensen, M. J., & Thayer, R. H. (2001). The Project Manager's Guide to Software Engineering's Best Practices. California: IEEE Computer Society.

Eisenberg, D., Marcotte, E. M., Xenarios, I., & Yeates, T. O. (2000). Protein function in the post-genomic era. Nature, 405(6788), 823-826.

Fleishmann, R. D., Adams, M. D., White, O., et al (1995). Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science, 269, 496-512.

Frame, J. D. (2003). Managing Risk in Organizations: A Guide For Managers. San Francisco: Jossey-Bass.

Gibbs, R. A., Weinstock, G. M., Metzker, et al. (2004). Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature, 428(6982), 493-521.

Goffeau, A., Barrell, B. G., Bussey, H., et al (1996). Life with 6000 Genes. Science, 274(5287), 546-567.

Griffiths, A. J. F., Miller, J. H., Suzuki, D. T., et al. (2000). An Introduction to Genetic Analysis (7th ed.). New York: W. H. Freeman and Company.

Hoffer, J. A., George, J. F., & Valacich, J. S. (2002). Modern Systems Analysis and Design (3rd ed.). New Jersey: Prentice Hall.

International Genome Sequencing Consortium (2004)

Myers, E. W., Sutton, G. G., Delcher, A. L., et al. (2000). A Whole-Genome Assembly of Drosophila. Science, 287(5461), 2196-2204.

Pandey, A., & Mann, M. (2000). Proteomics to study genes and genomes. Nature, 405(6788), 837-846.

Schmutz, J., Wheeler, J., Grimwood, J., et al. (2004). Quality assessment of the human genome sequence. Nature, 429(6990), 365-368.

Sommerville, I. (1996). Software Engineering (5th ed.). New York: Addison-Wesley.

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature, 408(6814), 796-815.

The C.elegans Sequencing Consortium. (1998). Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. Science, 282(5396), 2012-2018.

Too, M. H. F. (2005) Workflow Analysis, Risk Management and Laboratory Information Management Systems for Genome Research Projects, Bachelor of Science (Honours) Dissertation, University of Auckland.

Turban, E., King, D., Lee, J., Warkentin, M., & Chung, H. M. (2003). Decision Support Systems In the 21st Century (2nd ed.). New Jersey: Prentice Hall.

Visio Professional 2002. Retrieved, from the World Wide Web: www.microsoft.com

Vukmirovic, O. G., & Tilghman, S. M. (2000). Exploring genome space. Nature, 405(6788), 820-822.

Timetabling a Grouped Progressive Lunch

Ross J.W. James
Department of Management
University of Canterbury
ross.james@canterbury.ac.nz

Abstract

The Grouped Progressive Lunch Timetabling problem deals with allocating people or families for different courses of a meal to different locations. The objective is to try to get people to meet as many other people as possible while also moving around locations.

This paper investigates two methods of solving this problem. The first using an integer program which is formulated and solved using AMPL and CPLEX, while the other uses a Tabu Search meta-heuristic. These two techniques are compared on small problems and found that although the overall quality of solutions is the same, tabu search is faster.

A larger manually generated solution, that was actually used, was then compared to those that were generated by the integer program and tabu search. The solutions found by these techniques dominated the manual solution, and reduced the overall penalties from between 23-32%. The tabu search also proved to be very robust by generating solutions in 10 seconds that were better than what AMPL/CPLEX could find in one hour in half of the instances.

1 Introduction

A group of families meet once every couple of months to have lunch and socialise. Due to the number of people involved, during the summer months this usually occurs at some outdoor location – for example at a park or zoo. However during the winter months finding an indoor location large enough to fit the number of people involved at a desirable time can be problematic. In order to solve this problem the group have come up with what we will call a "Grouped Progressive Lunch". The concept is that there are several designated "host" families that will entertain several other "visiting" families at their house. The "visiting" families move from one house to another for each course of the lunch (designated as Mains, Desert and Coffee) while the "host" families stay in their own house throughout.

If the number of people who can go to a specific location is limited then this problem is similar to allocating people to tables for each course in a social setting in order for people to meet as many other people as possible.

The key to the success or otherwise of this approach is in the timetabling of which families visit which hosts and in which order. The main objective in this scheduling exercise is to have as many people meet as many other people as possible over the various courses of the lunch. Along with this we do not want the same people being with together all the time and also want people to move around the various locations so that the hosts also have different people to entertain. The problem therefore contains multiple criteria.

The size of the problem being dealt with changes over time with the availability of families and the number of families involved. Initially, however, the problem was to solve the problem with three host families and 11 visiting families over three courses.

Unfortunately there was not a great deal of literature that was particularly relevant to this particular problem. There was some literature on sports draw timetabling and scheduling which is similar (for example Fleurent and Ferland 1993, Russel and Leung 1994, Costa 1995 and Kostuk 1997) however the constraints and objectives are quite different and therefore the relevance of this literature is somewhat limited.

We therefore set about formulating this problem as an integer program and smaller problems are evaluated. This is described in section 2. We then explore the use of Tabu Search to solve this problem in Section 3. The evaluation of the two different approaches and the manual solution for the real sized problem is made in section 4, while our implementation is detailed in section 5. Conclusions are then drawn in Section 6.

2 Integer Programming

This problem can be formulated as an integer programme (IP) as follows:

Let P = number of people/families

Let T = number of time slots

Let L = number of visiting locations

Let α = Penalty if in the same location more than once

Let β = Penalty if a someone meets the same person/family more than once

Let γ = Penalty if two people/families never meet.

Let $i \in 1...P$, $j \in 1...T$, $k \in 1...L$, $b \in 1..\ T\text{-}1$

Let $x[i, j, k]$ = 1 if person/family i at time slot j visits location k, 0 otherwise.

Let $l[i, k, b]$ = extra units of the location penalty incurred if person/family i visits location k, more than b times, otherwise 0.

$l[i, k, b] \geq 0$

Let $ls[i, k, b]$ = units of location penalty avoided if person/family i does not visit k more than b times.

$ls[i, k, b] \geq 0$

Let $i_1 \in 1...P\text{-}1$, $i_2 \in i_1\text{+}1...P$

Let $M[i_1, i_2, j, k]$ = 1 if person/family i_1 meets person/family i_2 at location k and time j, otherwise 0

Let $MT[i_1, i_2]$ = the number of times person/family i_1 meets person/family i_2.

$MT[i_1, i_2] \geq 0$ and integer.

$MZ[i_1, i_2]$ = 1 if person/family i_1 does not meet person/family i_2, otherwise 0.

$\mathrm{MM}[i_1, i_2]$ = number of times person/family i_1 has met person/family i_2 more than once.

$\mathrm{MM}[i_1, i_2] >= 0$ and integer

Objective:

$$\text{Minimize } \alpha\sum_{i=1}^{P}\sum_{k=1}^{L}\sum_{b=1}^{T-1} l[i,k,b] + \beta\sum_{i_1=1}^{P-1}\sum_{i_2=i_1+1}^{P} MM[i_1,i_2] + \gamma\sum_{i_1=1}^{P-1}\sum_{i_2=i_1+1}^{P} MZ[i_1,i_2]$$

Subject to:

$$\sum_{k=1}^{L} x[i,j,k] = 1 \qquad \forall i,j \qquad (1)$$

$$l[i,k,b] = \sum_{j=1}^{T}\left(x[i,j,k]\right) - b + ls[i,k,b] \qquad \forall i,k,b \qquad (2)$$

$$ls[i,k,b] \le b \qquad \forall i,k,b \qquad (3)$$

$$M[i_1,i_2,j,k] \le \tfrac{1}{2}x[i_1,j,k] + \tfrac{1}{2}x[i_2,j,k] \qquad \forall i_1,i_2,j,k \qquad (4)$$

$$M[i_1,i_2,j,k] \ge x[i_1,j,k] + x[i_2,j,k] - 1 \qquad \forall i_1,i_2,j,k \qquad (5)$$

$$MT[i_1,i_2] = \sum_{j=1}^{T}\sum_{k=1}^{L} M[i_1,i_2,j,k] \qquad \forall i_1,i_2 \qquad (6)$$

$$MZ[i_1,i_2] \ge 1 - MT[i_1,i_2] \qquad \forall i_1,i_2 \qquad (7)$$

$$MZ[i_1,i_2] \le 1 - \tfrac{1}{P}MT[i_1,i_2] \qquad \forall i_1,i_2 \qquad (8)$$

$$MM[i_1,i_2] = MT[i_1,i_2] - 1 + MZ[i_1,i_2] \qquad \forall i_1,i_2 \qquad (9)$$

Constraint 1 ensures that each person/family can only be in one place at any one time.

Constraints 2 and 3 are used to calculate the penalty for being in a location more than once. An arithmetic series is used for the penalties as the number of visits to a single location increases. For example if the person/family is at a location twice, the penalty is 1, if they are there three times the penalty is 1+2 = 3, if they are there four times then it is 1+2+3 = 6. The variable *l* in this constraint defines the elements used in calculating this penalty.

Constraint 4 and 5 along with the requirement of M being binary ensure that two people/families have meet only when they are in the same place at the same time.

Constraint 6 simply defines the number of times two people/families have met in the schedule.

Constraint 7and 8 along with the requirement of MZ being binary ensures that MZ is 1 only if the two people/families have never met.

Constraint 9 defines the penalties for meeting people/families more than once. It allows one meeting before penalties are incurred.

To add the restriction on location size, we need to add the following definition to the formulation:

Let *LocLimit*[*k*] = the maximum number of families that can come to location *k* at any time.

And the following constraints:

$$\sum_{i=1}^{P} x[i,j,k] \le LocLimit[k] \qquad \forall k,j \qquad (10)$$

The IP was coded in AMPL and solved using CPLEX 8.2. A time limit of 1 hour was placed on the solver to provide a solution.

Various sized problems were tried, however CPLEX was unable to solve to optimality anything larger than 7 visiting families in the one hour time limit, and even then some 6 visiting family problems could not be solved to optimality. A summary of the problems solved are shown in Table 1.

			Penalties			Unlimited Capacity		Limited Capacity	
Families	Courses	Locations	Location	Meet > 1	Not Meet	Time (sec)	Objective	Time (sec)	Objective
5	2	3	1	1	1	1.83	3	3.12	6
5	2	3	3	2	1	2.79	6	1.33	6
5	2	3	6	4	1	1.70	6	1.20	6
5	3	2	1	1	1	0.67	9	0.68	9
5	3	2	3	2	1	0.77	23	1.26	25
5	3	2	6	4	1	0.76	46	1.24	48
5	3	3	1	1	1	53.70	4	16.83	4
5	3	3	3	2	1	6.97	4	4.93	4
5	3	3	6	4	1	4.24	4	3.30	4
5	3	4	1	1	1	700.29	3	707.21	4
5	3	4	3	2	1	198.49	3	187.75	4
5	3	4	6	4	1	140.03	3	119.11	4
6	2	3	1	1	1	16.07	5	15.80	9
6	2	3	3	2	1	37.92	9	8.68	9
6	2	3	6	4	1	11.10	9	5.27	9
6	3	2	1	1	1	5.39	12	5.57	15
6	3	2	3	2	1	4.95	30	4.00	33
6	3	2	6	4	1	5.87	60	3.73	63
6	3	3	1	1	1	700.87	6	138.84	6
6	3	3	3	2	1	41.33	6	21.05	6
6	3	3	6	4	1	12.93	6	7.86	6
6	3	4	1	1	1	3600.28	***5***	3608.99	***6***
6	3	4	3	2	1	3600.44	***5***	3086.37	6
6	3	4	6	4	1	2123.58	5	1635.40	6
7	2	3	1	1	1	142.30	7	480.67	12
7	2	3	3	2	1	767.30	14	264.89	14
7	2	3	6	4	1	248.97	16	226.55	16
7	3	2	1	1	1	30.04	15	30.27	15
7	3	2	3	2	1	63.91	39	39.40	39
7	3	2	6	4	1	70.67	78	50.10	78
7	3	3	1	1	1	3600.30	***9***	3609.33	***9***
7	3	3	3	2	1	2119.78	12	1718.40	12
7	3	3	6	4	1	604.79	16	620.39	16
7	3	4	1	1	1	3600.51	***7***	3609.31	***12***
7	3	4	3	2	1	3600.57	***8***	3568.39	***12***
7	3	4	6	4	1	3600.51	***8***	3594.59	***12***
8	3	3	1	1	1	3599.81	***13***	3594.58	***13***
8	3	3	3	2	1	3600.35	***18***	3594.62	***19***
8	3	3	6	4	1	3600.38	***26***	3618.66	***27***
8	3	4	1	1	1	3600.08	***10***	3600.75	***16***
8	3	4	3	2	1	3599.47	***11***	3599.59	***16***
8	3	4	6	4	1	3601.77	***11***	3599.55	***16***

Table 1. CPLEX Solution Times and Objectives for Small Problems
(Bold objectives are the best solutions found in the one hour time limit)

From these results it was clear that obtaining an optimal solution to this problem could take a long time, especially when considering full sized problems. The IP could, however, be used to get approximate solutions by limiting the time spent on problems. Trials on the limited capacity problems above showed that a limit of 10 seconds produced results that were on average 0.4% worse than those reported above, while a 1 second limit produced results that were 9.3% worse than those reported above. A further consideration, in terms of selecting a solution technique to solve this problem, was the accessibility of the software required in order to use this technique. Hence it was decided that a heuristic method should be developed as an alternative method for solving this problem.

3 Tabu Search

Tabu Search has been successfully applied to many combinatorial optimisation problems including timetable and scheduling problems (Glover and Laguna, 1997) and hence it likely to be a suitable method for solving this problem. The Tabu search was formulated as follows:

A solution is defined as a feasible allocation of each family at each time to a location. Feasibility implies that the solution must adhere to all limitations placed on each location.

The construction heuristic works by taking a list of families and simply breaking the list up evenly among the locations or according to the location limitations. In order to achieve this, times and families are looped through one at a time and added to a location until the location is deemed “full”. In the next time period, the process is repeated but the process is started at the next location. When the location capacity limits are not binding and the number of families at each location is not the same, the first locations will contain the extra people/families. Essentially this heuristic ignores the people/family mixing objective and simply focuses on the location objective.

The neighbourhood scheme for the tabu search is defined as changing the location of one family at a given time to a different location. If this move causes the new location to be infeasible, due to location limitations, then it also includes a move of a different family from the location with the infeasibilty to another location which will not cause any feasibility violations.

The search uses a first improvement strategy, so that as soon as a solution that is better than the current solution is found the search moves to that solution immediately.

The tabu list records the family(ies), time(s) and location(s) from which a move has been made from. The tabu criteria is that a move is tabu if it involves moving a family at a given time to a location and time that is currently on the tabu list. The tabu list implemented is a static tabu list that does not vary in size over time.

The aspiration criteria is that a tabu move is accepted if the move is better than the best solution so far.

The tabu search was coded using Microsoft Visual C .NET and run on an AMD Athlon XP 1800+ 1.53 Ghz computer with 1 Gb RAM. The search was run for 10 seconds on each of the small problems above. Tabu list sizes of 15, 20, 25, 30 and 35 were tested. In all except once instance (that of a tabu list size of 35 and problem with 6 families, 3 time, 3 locations, all penalties being 1 and no capacity constraints) the search produced solutions exactly the same results as those produced by CPLEX. For the search with no capacity limitations, the average time to find the best solution was just 0.0026 seconds, and for the capacity limited problems the average time was just 0.0069 seconds

From these tests it is clear that the Tabu Search can produce very good solutions to this problem very, very quickly.

4 Evaluation

In order to see if these techniques would be able to produce results better than what was done manually, the manual schedule, shown in Figure 1, was evaluated.

Course\Where	Loc1	Loc2	Loc3
Lunch	F1 F2 F3 F4	F5 F6 F7	F8 F9 F10 F11
Dessert	F5 F9 F10 F11	F1 F4 F8	F2 F3 F6 F7
Coffee	F6 F7 F8	F2 F9 F10 F11	F1 F3 F4 F5

Figure 1. Manually Generated Solution

Analysing this solution we can see that we have:

1 Location penalty (F3 being in Loc3 for both Desert and Coffee)

13 Mix Penalties (Families that meet twice [1 penalty] - F1 and F3 meet Lunch and Coffee, F2 and F3 meet Lunch and Desert, F3 and F4 meet Lunch and Coffee, Families that meet 3 times [2 penalties] - F1 and F4, F6 and F7 F9 and F10, F9 and F11, F10 and F11)

23 Combinations that do not meet – (F1 and F6, F7, F9, F10, F11; F2 and F5, F8; F3 and F8, F9, F10, F11; F4 and F6, F7, F9, F10, F11; F5 and F8; F6 and F9, F10, F11; F7 and F9, F10, F11)

The objective value will depend on the values associated with each of the types of penalties. For a penalty, in the form of (α, β, γ), of (1,1,1) the total penalty would be 37, for (3,2,1) a total penalty of 52, and for (6,4,1) a total penalty of 81.

Solving this problem using Tabu Search using a list size of 20 for 10 seconds and with the AMPL/CPLEX model limited to one hour run time produced the results in Table 2.

Penalty Values			Tabu Search					AMPL/ CPLEX (1 hour time limit)
				Penalty Incidents				
Location (α)	Meet > 1 (β)	Not Meet (γ)	Time	Location	Meet >1	Not Met	Objective	
Unrestricted								
1	1	1	0.04	7	6	12	25	26
3	2	1	0.01	0	10	20	40	43
6	4	1	0.01	0	10	20	60	60
Restricted								
1	1	1	0.17	9	4	14	27	28
3	2	1	0.00	0	10	20	40	40
6	4	1	0.01	0	10	20	60	60

Table 2. Tabu Search and CPLEX results on an Actual Problem

Independent of the penalty values we can see that the solution produced manually is dominated by the solution obtained when the penalties are not all the same. This amounts to a savings of between 23% - 32% in penalties for the weight combinations explored here.

These computational tests also show that the tabu search is far more efficient at producing good solutions than using AMPL/CPLEX. Even though CPLEX had one hour to solve the problem, compared to 10 seconds for the tabu search, tabu search still produced solutions that were better than CPLEX in half of the cases and of the same quality in the other half of the cases.

5 Implementation

As the tabu search was able to produce very good quality solutions very quickly, this was chosen as the technique to implement. In order to make the solver accessible, an Excel spreadsheet user interface was developed. This allowed the user to enter names of families, courses and locations and produce a final allocation, which is ready for printing and distribution – along with an analysis of the penalties involved with the solution.

Screen shots of these spreadsheets are shown in Figures 2 and 3.

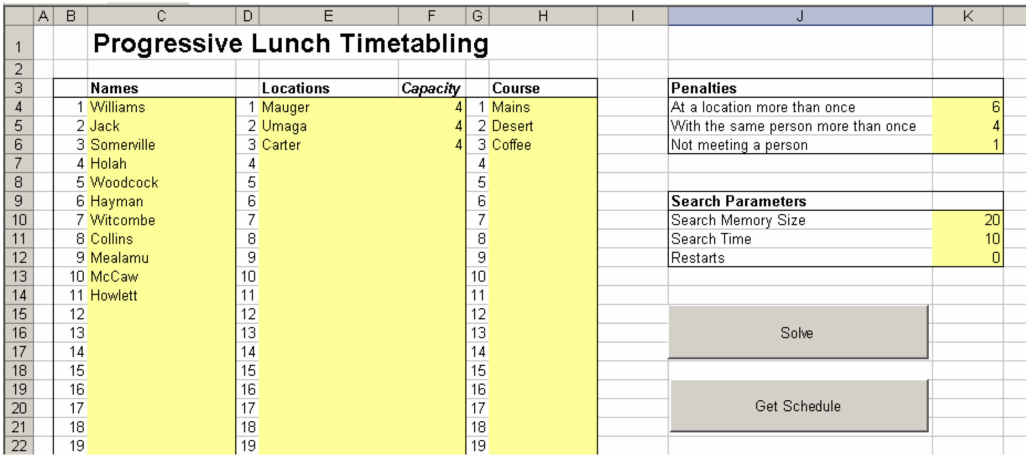


Figure 2. Excel Spreadsheet Data Entry Screen

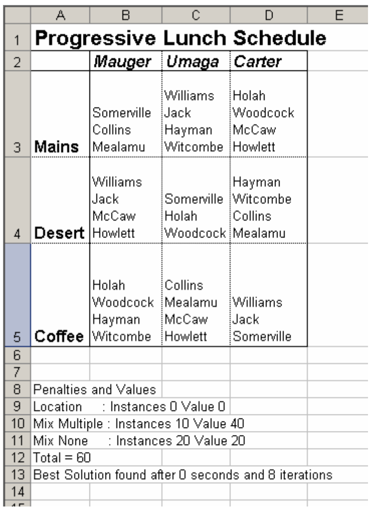


Figure 3. The Final Schedule

The Solve button on the spreadsheet takes the data on the spreadsheet and writes an input file of the problem parameters to a file. It then calls the tabu search program and which writes to another file the best solution found and its analysis. Pressing the “Get Schedule” button on the spreadsheet, reads in this file and creates the final schedule in a form that is ready to be printed, as shown in Figure 3.

6 Conclusions

In this paper we have solved the Grouped Progressive Lunch problem using both integer programming and tabu search. The tabu search approach has proven to be much faster than using integer programming and it was still able to producing solutions that were as good as, and in many cases, even better than was possible from an IP with limited computational time. The advantage of using these methods over doing the allocation manually were demonstrated with a saving of over 20% in the penalties, while taking a lot less human effort to produce an acceptable schedule.

The spreadsheet user interface makes the tabu search solver accessible for this solver, and enables a user friendly schedule to be produced.

7 References

Costa D. 1995. “An evolutionary tabu search algorithm and the NHL scheduling problem”, *INFOR*, **33**: 161-178

Fleurent C. and J.A. Ferland. 1993. “Allocating Games for the NHL using Integer Progamming”, *Operations Research*. **41**:649-654

Glover F. and M. Laguna. 1997. *Tabu Search*, Kluwer Academic Publishers, Boston

Kostuk K.J. 1997. “A Descision Support System for a Large, Multi-event Tournament”, *INFOR*, **35**:183-196

Russel R.A. and J.M.Y. Leung. 1994. “Devising a Cost Effective Schedule for a Baseball League”, *Operations Research*. **42**:614-625

Multiple Farm Crop Harvesting Scheduling

C B Basnet
L R Foulds
University of Waikato
New Zealand

J M Wilson
Loughborough University of Technology
United Kingdom

chuda@waikato.ac.nz
lfoulds@waikato.ac.nz
J.M.Wilson@lboro.ac.uk

Abstract

We discuss the harvesting of renewable resources from an operations scheduling viewpoint. We report on a practical harvesting scenario arising in the agricultural context, involving a commercial contracting enterprise that travels from farm to farm, harvesting crops. The paper is an extension of previous work by two of the authors from the one-farm to the multi-farm case. In both cases: the duration of each operation is dependent upon the combination of constrained resources allocated to it, equipment and worker allocation is restricted, and minimum or maximum time lags on the start and completion of operations may be imposed. The present case incorporates harvesting at more than one farm and thus inter-farm travel times must be taken into account. We report on a harvesting operations scheduling model and solution procedures designed specifically for the multi-farm case. The computational times experienced in solving general instances of the model of small to medium practical size by a commercial integer programming package are encouraging. We have also devised greedy and tabu search heuristics which are capable of solving problems of relatively large dimensions in reasonable computational time. The authors believe that the model and the solution techniques developed represent a useful addition to the farm crop contractor's tool kit.

1 Introduction

We focus on the scheduling of the operations that are carried out by contractors who harvest crops at various farms. The duration of each operation is dependent upon the combination of constrained resources allocated to it, equipment and worker allocation is restricted, and minimum or maximum time lags on the start and completion of operations may be imposed. The present case incorporates harvesting at more than one farm and thus inter-farm travel times must also be taken into account. As far as the authors are aware, none of the solution techniques reported to date in the open literature appear to be applicable to large-scale numerical instances of the multiple-farm harvesting scenario considered in the present paper. This is because the scenario analysed here involves an unusual combination of: resource allocation-dependent operation durations, farm-to-farm travelling time, and time lags.

The presence of a minimum time lag for a given operation implies that the operation cannot start before a specified time has elapsed after the completion of one of its predecessor operations. The presence of a maximum time lag for a given operation implies that an operation must be started, at the latest, a certain number of time periods after the completion of one of its predecessor operations. The set-up times for various types of machines, and the drying and inspection time necessary between certain pairs of operations sometimes create minimum or maximum time lags in any feasible schedule.

2 Scheduling Operations for the Harvesting of Crops

The scenario involves a contracting company that travels from farm to farm to harvest crops. The primary objective of the harvesting process to be discussed in this paper is to specify the activity of each actual machine of each type, and its crew, in each time period in order to minimize the duration of the entire harvesting process at a given set of farms. Thus, set up, processing, and inter-farm travel times must be taken into account. However, in many harvesting operations scheduling scenarios, such as the present one, the paucity of dedicated machines and their trained crews that are capable of carrying out certain operations often give rise to binding constraints on the duration of the overall harvesting process. We assume that: the set of farms to be visited is known, all inter-farm travel times are known, and the inter-farm transportation of any actual machine may involve the skipping of certain farms. (That is, each physical machine is not necessarily transported to each farm.) Also, for each farm to be visited: the harvesting operations are performed in a given sequence, the set up of all machines of each type that are actually utilized are carried out simultaneously, without interruption, while any operation (apart from the last) is being performed, the machines to be used for its succeeding operation can either be set up at the same time or once the operation is completed, each operation must be started immediately after the completion of the set up of its associated machines, each operation must be performed from start to completion without interruption, and any time lag that is imposed constrains the start of each operation in relation to the completion of its immediate predecessor.

The input data needed for the analysis of a harvesting scheduling scenario include: the number of machines and their crews that are available to perform each operation, the set-up time for each type of machine, all inter-farm travel times, the sequence of the operations to be performed, and for each farm: the imposed time lags between operations, and the resource level allocation-operation duration relationships. (That is, how much, if at all, operation processing times are reduced by the allocation of additional machines.) The matters just mentioned are made more precise in the model given in Section 4.

3 A Literature Survey

Extensive literature is available on the use of traditional project scheduling models for scheduling activities in a resource-constrained environment by according priorities to different activities (e.g. Foulds and Wilson, 2004). The scheduling of the operations used in the harvesting of renewable resources, especially in forestry, has been the focus of much research (e.g. Weintraub et al., 1994; McNaughton et al., 1998).

However, there are a number of factors arising in farm crop harvesting that cannot be easily addressed by the previously mentioned models and software. These include:

maximum time lags, operation interdependency, time and activity-dependent criticality, conflicting priorities, and the mutual exclusivity, partial allocation, sharing, the substitution of resources, and inter-site transportation. Thus, there is much scope for development of project scheduling techniques that are effective for practical industrial problems. Attempts to address some of these issues, by reporting on exact and approximate methods for crop harvesting at a single farm, have recently been reported by Foulds and Wilson (2005) and Corner and Foulds (2005). A version of the multi-farm case, with stochastic activity durations and the possibility of machine failure, has been examined by Foulds (2004) who presented a probabilistic dynamic programming procedure which addresses the issue of machine breakdown. The present paper represents an extension of the last three papers mentioned to large-scale deterministic instances of the farm-to- farm case.

Approaches to scheduling resource-constrained projects comprise exact procedures (often based on integer programming), and heuristic procedures. A number of branch and bound algorithms have been proposed for various resource-constrained scheduling problems (e.g. Dorndorf et al., 2000). Heuristics based on priority-rule methods leading to improved results have been reported by Neuman et al. (2003), who deal, for the first time in the open literature, with the additional complication of having the resources constrained by maximum and minimum inventories of cumulative resources.

Some of the literature referenced above deals with minimum time lags (e.g. Brucker et al., 1999). There is less literature available on the additional requirement of maximum time lags than there is on minimum time lags (e.g. Neumann and Zimmermann, 2000). Resource levelling (RL) involves attempting to minimize the fluctuations in requirements for resources, so that they will be utilized as uniformly as possible throughout the process. RL methods that aim to produce optimal schedules have been constructed by Easa (1997). Heuristic methods based on either simple shifting and priority rules have been reported by Takamoto et al. (1995).

4 Model and Solution Technique Development

We begin our description of the model by introducing the necessary notation. Later we report on heuristics that were developed from the model.

Parameters

$H =$	The total number of farms to be serviced,
$n =$	the number of machine types/operations,
$m_i =$	the number of individual machines of type i available, $i = 1, 2, \ldots, n$,
$(1, 2, \ldots, n) \sim$	the sequence in which the operations must be performed (the same at all farms),
$M =$	A relatively large constant.

Input data

$q_{gh} =$	the time to transport any machines from farm g to farm h; $g, h = 1,2,\ldots H$,
$d_{ih}^{\min} =$	the minimum time lag between operations i and $i+1$ at farm h; $h = 1,2,\ldots H$,

$d_{ih}^{\max}$ = the maximum time lag between operations i and $i+1$ at farm h; $h = 1,2,\ldots H$,

p_{ijh} = the time needed to carry out operation i at farm h if j machines of type i are used,

t_i = the set up time of machines of type i.

Dec. variables

s_{ih} = the start time of operation i at farm h, $i = 1, 2, \ldots, n$; $h = 1,2,\ldots H$,

e_{ih} = the completion time of operation i at farm h, $i = 1, 2, \ldots, n$; $h = 1,2,\ldots H$,

x_{ijgh} = = 1, if machine j of type i is sent from farm g to h;
= 0, otherwise, $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, m_i$; $g, h = 1,2,\ldots H$,

y_{ijh} = =1, if j machines of type i are used at farm h,
= 0 otherwise, i = 1, 2, …, n; j = 1, 2, …, mi ; h = 1,2,…H .

Z = The objective function value.

A dummy farm, $h = 0$, is introduced. Initially all machines are stored at farm zero, but can be moved instantly to any farm.

4.1 Model CropHarvest

Minimize Z (4.1)

Subject to:

$Z \geq e_{nh} \quad h = 1,2,\ldots,H$ (4.2)

(The makespan is no smaller than the end time of the last operation at all farms)

$$\sum_{h=1}^{H} x_{ij0h} = 1,\ i = 1,2,\ldots,n;\ j = 1,2,\ldots,m_i \quad (4.3)$$

(All machines are sent from farm zero to some farm.)

$$\sum_{\substack{g=0\\ g\neq h}}^{H} x_{ijgh} - \sum_{\substack{l=1\\ l\neq h}}^{H} x_{ijhl} \geq 0,\ i = 1,2,\ldots,n;\ j = 1,2,\ldots,m_i;\ h = 1,2,\ldots,H \quad (4.4)$$

(Machines sent from farm h must have arrived there previously.)

$$e_{ig} - s_{ih} + (q_{gh} + t_i + M)x_{ijgh} \leq M,\quad i = 1,2,\ldots,n;\ j = 1,2,\ldots,m_i;\ g = 1,2,\ldots,H;\ h = 1,2,\ldots,H, \quad (4.5)$$

(The start time of an operation at a farm is determined by the completion time at farms that send machines to it.)

$$\sum_{j=1}^{m_i} \sum_{\substack{g=0\\ g\neq h}}^{H} x_{ijgh} - \sum_{j=1}^{m_i} j y_{ijh} \geq 0,\ i = 1,2,\ldots,n;\ h = 1,2,\ldots,H \quad (4.6)$$

(The number of machines available for use at farm h cannot exceed the number that have arrived at h.)

$$e_{ih} - s_{ih} - t_i - \sum_{j=1}^{m_i} p_{ijh} y_{ijh} \geq 0 \quad i = 1,2,\ldots,n;\ h = 1,2,\ldots,H. \quad (4.7)$$

(The completion time of an operation at a farm is determined by the start, the setup, and the processing times.)

$$s_{i+1,h} - e_{ih} - d_{ih}^{\min} \geq 0,\ i = 1,2,\ldots,(n-1);\ h = 1,2,\ldots,H. \quad (4.8)$$

(The start time of each operation must allow for the minimum time lag after the completion of the previous operation.)

$$s_{i+1,h} - e_{ih} - d_{ih}^{\max} \leq 0,\ i = 1,2,...,(n-1);\ h = 1,2,...,H. \tag{4.9}$$

(The start time of each operation must allow for the maximum time lag after the completion of the previous operation.)

$$\sum_{j=1}^{m_i} y_{ijh} = 1,\ i = 1,2,\ldots,n;\ h = 1,2,\ldots,H \tag{4.10}$$

(A positive number of machines must be used for each operation at each farm.)

$$x_{ijgh} \in \{0,1\};\ i = 1,2,...,n;\ j = 1,2,...,m_i;\ g = 0,2,...,H;\ h = 1,2,...,H,\ g \neq h. \tag{4.11}$$

$$y_{ijh} \in \{0,1\};\ i = 1,2,...,n;\ j = 1,2,...,m_i;\ h = 1,2,...,H \tag{4.12}$$

It is clear that the above model expressed by (4.1) - (4.12) is an integer program comprising mainly zero-one variables. For numerical examples of the size normally encountered in practice, it can be solved to optimality using current commercial IP codes. However, relatively large numerical instances may require heuristic techniques. Analysis of the model suggests the following simple, greedy heuristic, which considers the farms and operations, one at a time, in the order of the given sequences.

4.2 A Greedy Heuristic

For this heuristic, the farms are renumbered so that the sequence: $h = 1,2,...,H,$ corresponds to a Traveling Salesman Problem (TSP) path through all farms. Given the dimensions of numerical problems commonly encountered in practice, the farthest insertion procedure can be used for this purpose in reasonable computational time. The reader is referred to a description of the TSP, and solution procedures for it, by Applegate et al. (1988).

Each farm h, in the order: $h = 1,2,...,H$:
For each operation i, in the order $i = 1,2,...,n$:
Allocate the maximum available number of machines of type i.
Begin setting up the allocated machines and have them begin operation i in the earliest feasible time periods.
If $h < H$:
As soon as operation i is completed, transport all machines involved to farm $h + 1$.
If $h = H$ and $i = n$, terminate.

The number of machines assigned to each operation i, $i = 1,2,...,n,$ at each farm, is set at m_i, the maximum number of machines available.

We now develop a procedure based on tabu search that is more useful than the greedy heuristic just mentioned.

4.3 A Tabu Search Procedure

Notation: Scheduling order, $S = \{I_1, I_2, \ldots, I_H\}$ is an ordering of the indices of the farm (1, 2, .. H), such that when $I_k = h$, farm, h is the kth farm to be scheduled. Machine allocation, u_{ih} = the number of machines allocated to farm h for operation i, $i = 1, 2, .., n$; $h = 1, 2, \ldots H$.
Schedule generation: Given a scheduling order S of farms and an allocation of machines u_{ih} to all farms for all operations, a non-delay schedule can be developed in the following manner.

1. A Gantt chart is built up chronologically for each machine of each type, starting from time zero. This includes the set up time, operation time, idle time, and transport times for that machine.
2. As soon as any machine is free, it is allocated to a farm following the scheduling order of the farms, until the allocation of machines of that type to that farm is met. Machines that can arrive at the farm in question earliest are given preference for allocation. An allocated machine is scheduled to be transported to the allocated farm without delay after its current operation.
3. Setting up and operation of a particular machine in a farm can start only after the allocated number of machines of that type has arrived at a farm. The transportation, set up, and operation is carried out without inserting delays. Delays are inserted only to the extent necessary to enforce the minimum time lags.
4. If an operation is delayed beyond the maximum time lag, because the machines for that operation are not available in time, the earlier operations for that farm are delayed to the required extent so that the maximum time lag constraint is satisfied. This will have a follow-on effect on the schedule of earlier operations at the farm.
5. The schedule is built up, one farm at a time, in the scheduling order S, and within a farm, one operation at a time in order of the technological constraints. This chronological build-up of the schedule continues until all the operations of all the farms are completed.

The search space: Given a scheduling order of farms and a particular allocation of machines to farms, it is straightforward to show that the above schedule generation procedure always results in a feasible schedule. The search space in the TS search is the set of all schedules generated by the set of scheduling orders of farms and the set of machine allocations. A move transforms an element of the search space to another element. In the TS procedure there are two possible moves:

1. A change in the scheduling order, so that two farms which are next to each other in S swap their priorities. Only farms next to each other (I_k, I_{k+1}) are considered for a swap. The total number of moves of this type to be considered at each iteration is (H -1).
2. A change in the allocation of machines for an operation in a farm so that the number of machines allocated to a particular operation (u_{ih}) increases or decreases by exactly one. This allocation must be at least one, and at most it can be the total number of machines for that operation. Since an allocation can go up or down, the maximum number of moves to be considered at each iteration is $2nH$.

It is straightforward to show that it is always possible to reach the optimal schedule within the search space via the two incremental moves described above.

4.3.1 Tabu search

The objective function Z (s) is the makespan: the completion time of the last operation on the last farm. As all the generated schedules are always feasible there are no penalty parameters in the TS procedure. The procedure is carried out in the normal TS manner, examining all the possible moves, and making the move that yields the lowest objective function, until a given number of iterations have been performed. The length of the tabu list is updated dynamically, and an intensification strategy is followed. In evaluating a move, the schedule does not always need to be generated *ab initio* – if a move involves farm I_k only, the schedule of farms I_k , I_{k+1} ..I_H must be regenerated. The TS procedure requires as input: an initial sequence in which the farms are visited and the number of machines to be allocated to perform each operation at each farm. If no better input is known, the output of applying the a greedy heuristic can be used.

4.3.1.1 Intensification strategy

The intensification strategy comprises a prohibition of all moves associated with particular farms. When the objective does not improve for a given number of iterations, the farms are randomly partitioned into sub-sets of two farms and moves associated with a sub-set are prohibited for a fixed number of iterations. The prohibition is applied to all the sub-sets turn by turn. The intensification strategy is stopped as soon as the objective improves. If the objective does not improve, the partition size is increased to 4, 6,…, until the partition size exceeds half the number of farms. If the objective still does not improve, the normal TS procedure is resumed.

5 The Testing of the Solution Procedures

The IP approach can solve problems involving up to 10 farms and 4 operations fairly rapidly, but becomes cumbersome on larger problems. Hence the greedy and TS procedure were tested on larger numerical instances. Ten problems were generated for each size (up to 20 farms and 5 operations). Problems were randomly generated with the following distributions: Setup time: uniform, between 2 and 5; No. of machines for an operation: uniform, between 2 and 4; Minimum time lag: uniform between 3 and 5; Maximum time lag: uniform between 5 and 10; and One-machine-operation durations were uniformly distributed between 2 and 15. The operation time for using a number of machines was generated by dividing the duration by the number of machines.

Farm locations were generated on a Cartesian co-ordinates grid, with the distance from the origin being distributed uniformly between 0 and 10 along both axes. Distances between farms were calculated as the Euclidean distance. The solution found by the greedy heuristic was used as the input initial solution for the TS procedure. The number of iterations was set to be equal to (the total number of machines)*(the number of farms)*(10). The results are illustrated in Table 1.

Problem size (number of farms, number of operations)	Greedy Heuristic		Tabu Search	
	Average makespan	Average run time, seconds	Average makespan	Average run time, seconds
5, 3	70.4	0	60.0	0.6
10, 4	127.0	0	116.2	9.9
10, 5	142.8	0	133.3	17.6
15, 4	168.0	0	144.2	44.6
15, 5	178.1	0	160.6	80.5
20, 4	209.2	0	180.2	85.9
20, 5	222.0	0	202.5	178.8

Table 1: Computational Results Using the Heuristics

As can be seen in Table 1, the TS procedure is able to improve on the greedy heuristic to the extent of about 10% of the makespan, at a cost of significantly higher (but still modest) run time. The larger problems require about 3 minutes each for the TS procedure, which was programmed in Microsoft Visual C++ and run on a PC with 300HZ and 64 MB RAM.

6 Summary and Conclusions

We have discussed the scheduling of contractors' farm-to-farm crop harvesting operations. We conclude that the scheduling of such harvesting operations is a significantly different scenario from those represented by the scheduling models currently available in the literature, due to the facts that: the duration of each operation is dependent upon the combination of constrained resources allocated to it, equipment allocation is restricted, minimum or maximum time lags on the start and completion of operations can be imposed, and the fact that inter-farm travel times must be taken into account. An IP model of harvesting scenario was developed and greedy heuristic and tabu search procedures were developed. The computational times experienced in solving general instances of the model of relatively large dimensions by the tabu search procedure are encouraging. Thus, the authors believe that the models and procedures reported represent useful additions to the harvesting contractor's toolkit.

7 References

Applegate, D., R. Bixby, V. Chvatal, and W. Cook, 1998. *Doc. Math. J. DMV Extra Volume ICM III* (1998) 645-656

Brucker, P., A. Drexl, R. Moehring, K. Neumann, and E. Pesch, 1999. "Resource-constrained project scheduling: notation, classification, models, and methods." *European Journal of Operational Research* **112**:3-41.

Corner, J L., and L. R. Foulds, 2005. "Scheduling the Harvesting Operations of a Forest Block: A Case Study." *Asia Pacific Journal of Operational Research* **22**:1-14.

Dorndorf, U., E. Pesch, and T. Phan Huy, 2000. "A time-oriented branch-and-bound algorithm for resource constrained project scheduling with generalized precedence constraints." *Management Science* **46**:1365-1384.

Easa, S., 1997. "Resource levelling in construction by optimisation." *Journal of Construction and Engineering and Management* **115**: 302-316.

Foulds, L. R., and J.M. Wilson, 2005. "Scheduling operations for the harvesting of renewable resources." *Journal of Food Engineering* **70**:281-292.

Foulds L. R., 2004. "Hay harvesting operations scheduling subject to probabilistic activity duration and machine failure", in *Proceedings of the 39th Annual Conference of the Operational Research Society of New Zealand*, Auckland, New Zealand (2004), pp 251-262.

Glover, F. and M. Laguna, 1997. *Tabu Search*, Kluwer Academic, Boston.

McNaughton, A.J., M. Ronnqvist, and D.M. Ryan, 1998. "The Scheduling of Forest Harvesting with Adjacency Constraints." *Proceedings of the 33rd Annual Conference, Operational Research Society of New Zealand,* The University of Auckland, Auckland, New Zealand.

Neumann, K. and J. Zimmermann, "Procedures for resource levelling and net present value problems in project scheduling general temporal and resource constraints." *European Journal of Operational Research* **118**: 425-443.

Takamoto, M., N. Yamada, Y. Kobayashi, and H. Nonaka, 1995. "Zero-one quadratic programming algorithm for resource levelling of manufacturing process schedules." *Systems and Computers in Japan* **26**:68-76.

Weintraub, A., F. Barahona, and R. Epstein, 1994. "A Column Generating Algorithm for Solving General Forest planning Problems with adjacency Constraints." *Forest Science* **40**:142-161.

XPRESS-MP. *Dash Optimization*, Blisworth NN73BX, UK.

The 0-1 knapsack problem with shared costs

S. Mitchell and H. Waterer
Department of Engineering Science
The University of Auckland
{s.mitchell, h.waterer}@auckland.ac.nz

Abstract

Consider a 0-1 knapsack problem involving a set of items, a collection of subsets of the set of items, and a knapsack of a given capacity. Placing an item in the knapsack returns a profit. If one or more items of a given subset are placed in the knapsack then a shared cost is incurred. That is, the first item of this set placed in the knapsack incurs a cost. Subsequent items from this set do not incur any further costs associated with that set. Thus, the cost of placing items from this set in the knapsack is shared among those items, the cost per item decreasing as more items from the set are included.

A set of items is said to be feasible if they do not overfill the knapsack. The problem is to find a feasible set of items to place in the knapsack that maximizes the total profit less the shared costs incurred by the items.

This problem occurs as a subproblem in a model of a forestry application. The LP relaxation of the integer programming formulation can be very weak. Preliminary results of an investigation into the use of strong valid inequalities to solve this problem will be presented.

Emergency Vehicle Trip Analysis using GPS AVL Data: A Dynamic Program for Map Matching

A.J. Mason
Department of Engineering Science
University of Auckland, New Zealand
www.esc.auckland.ac.nz/Mason
a.mason@auckland.ac.nz

Abstract

We outline a new dynamic programming approach for determining the most likely route taken by a vehicle given a set of GPS locations and times recorded by an AVL (automatic vehicle location) system. By using optimisation approaches, this map matching algorithm can solve a complex maximum likelihood problem that takes into account data such as the vehicle's location, heading, and speed. The optimisation approach allows routes to be determined even when the GPS data points are widely spaced in time and/or distance. Preliminary results are reported using data collected from ambulances.

1 Introduction

As automatic vehicle location (AVL) systems become more common, a wealth of vehicle tracking data is being collected and stored. For example, many ambulance organisations have now equipped each vehicle with a GPS device that continuously tracks the location of the vehicle. This location data is transmitted back to the headquarters where it is plotted and stored, and typically never looked at again. As part of our work in developing an ambulance simulator [3], we wish to extract as much information as possible from this vehicle GPS data. In particular, we would like to determine road speeds for both standard and 'lights and sirens' travel, and also to extract accurate trip timings from the data. In this paper, we discuss a new dynamic programming approach to analyse the GPS data and extract from it detailed route and timing information that can help address these requirements.

In the literature, the problem we face is referred to as 'map matching' [2]. The basic map matching problem is to translate a sequence of GPS locations into a sequence of arcs that define a route on some underlying road network. If each GPS location was known exactly, and the arc locations defining the road network were also known exactly, then this would be a simple problem of determining which arc each GPS point is located on. As long as the GPS points were sufficiently close, this would then define a sequence of arcs that would, in turn, define a route on the road network.

Unfortunately, the problem is typically not so simple. The GPS system relies on triangulating positions from satellite locations, and thus small errors in a reported location can occur if the number of visible satellites is low, or if atmospheric disturbances introduce transmission timing errors. Large errors or gaps in the data can occur if the signal is blocked or reflected, such as occurs in tunnels or in the 'urban canyons' created by tall buildings and narrow streets. Gaps in the data can also occur at the start of a trip when the GPS unit spends time searching for and then locking onto the satellites. Furthermore, the data used to define the road network will typically contain errors arising from geometry simplification or digitisation difficulties. This combination of uncertainties can complicate the mapping of GPS points to arcs. For example, in Figure 1, it is not clear from the GPS data whether the vehicle turned left or right at the intersection.

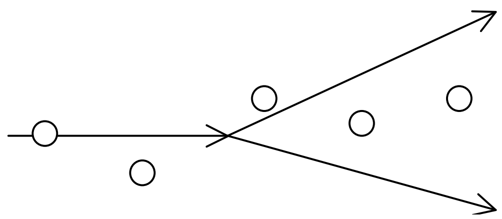


Figure 1: A possible sequence of GPS points (circles) generated by a vehicle travelling on a road network.

A further issue that complicates the problem is the frequency of GPS data transmissions. In most of the previous work, GPS data has been provided at intervals of around 10 seconds or less (e.g. the German and Swedish data used by [1] and [4] was mostly collected at 1 second intervals). This means that there will typically be a number of GPS points available to determine and then confirm that a vehicle is travelling along some proposed road segment. In this case, one can use a back-tracking approach in which some candidate route is constructed by heuristically choosing a 'most likely' next arc at each intersection, and then back-tracking and altering this choice if the successive GPS data points stray too far from the proposed route. However, this approach becomes more problematic if, as in our case, the GPS data points are widely spaced, and thus large segments of the route have to be determined before the next GPS data point is encountered.

1.1 Data Available

In this work, we assume we have a set of GPS data points generated by a vehicle. (In our ambulance data, these points are typically generated using rules that result in a new GPS data point being recorded 'every 500m or 5 minutes' or 'every 333m or 15 minutes'.) Each data point has a time and a location, and, depending on the particular system being used, may also contain information on the vehicle's speed, heading, and the distance travelled since the time the last data point was recorded.

We also assume that we have some detailed road network available represented as a set of nodes and directed arcs with associated travel speeds. Each arc is defined by a sequence of 2 or more locations joined by straight line segments, thus allowing the arcs to contain bends. Each node in the network represents the location of a traffic intersection (where the driver has a choice of next arc), a merging of the traffic flows on two or more arcs, or the end of a dead-end road. We assume we have some efficient algorithm for finding fastest paths in this network.

2 Model Description

We are fortunate in that, as well as having GPS data records available, we also have vehicle status information that tells us when a vehicle begins or finishes a trip. The first step in our data processing is to use this status information to break the sequence of GPS data points into a sequence of trips. Each trip is defined by a set of GPS data points that represent the vehicle travelling from some start point (e.g. an ambulance base) to some destination (e.g. the scene of an accident). As well as using this status information, the trip processing algorithm also constructs new trips whenever it finds a long sequence of data points at the same location. This can occur, for example, when a vehicle makes a lunch stop on the way back from a hospital.

We consider a trip T to be defined as a sequence of n GPS data points, $T = \{g_1, g_2, ..., g_n\}$, where a GPS data point g_c, $c = 1, 2, ..., n$ is represented by a tuple $g_c = (t_c^{gps}, p_c^{gps}, v_c^{gps}, d_c^{gps})$ giving the vehicle's position $p_c^{gps} = (x_c^{gps}, y_c^{gps})$ at time $t_c^{gps} > t_{c-1}^{gps}$, along with its velocity v_c^{gps} (given as a speed and heading (s_c^{gps}, h_c^{gps})) and the distance travelled d_c^{gps} since the last data point was recorded.

In our algorithm, we will be mapping each GPS data point g_c to a corresponding *network position* p_c^{net}, being a point defined on the road network by the pair $p_c^{net} = (a_c, o_c)$. This position, also denoted (x_c^{net}, y_c^{net}), is defined by an arc a_c^{net} and an offset $0 \leq o_c^{net} \leq 1$ where o_c^{net} defines a fraction along arc a_c^{net} measured from the start of arc a_c^{net}. A network position p_c^{net} is typically said to be a 'directed location' in the sense that the vehicle is travelling in the direction of arc a_c^{net}. However, the position may also be an undirected position if the direction of a_c^{net} is not relevant, and thus the vehicle may either be travelling in the direction of a_c^{net} or in the opposite direction, i.e. in the direction of arc a_c^{net}'s reverse arc (if such an arc exists). Network positions will be directed unless stated otherwise.

Our algorithm starts by defining, for each GPS data position $p_c^{gps} = (x_c^{gps}, y_c^{gps})$, an associated set of discretised candidate network positions $\mathrm{P}_c^{net} = \{p_{c,i}^{net}, i = 1, 2, ..., |\mathrm{P}_c^{net}|\}$ where $p_{c,i}^{net} = (a_{c,i}, o_{c,i})$ represents a possible position for the vehicle at time t_c^{gps} when the c^{th} GPS data point g_c was generated. Figure 2 shows one such set of positions that could be generated for a GPS data point. Heuristics can be used to limit the set P_c^{net} of candidate network positions by, for example, limiting the maximum distance allowed between the network position $p_{c,i}^{net}$ and the GPS point p_c^{gps}. Another useful simplification is to restrict the candidate positions to no more than one per arc, thereby converting the generation of candidate positions to the problem of constructing a set of nearby arcs and then finding that point on each arc that is closest to the GPS location.

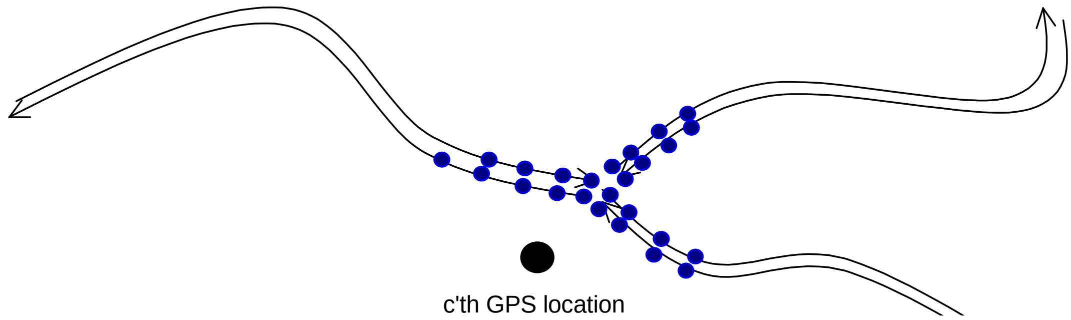


Figure 2: Discretised candidate vehicle locations (network positions) for a GPS data point.

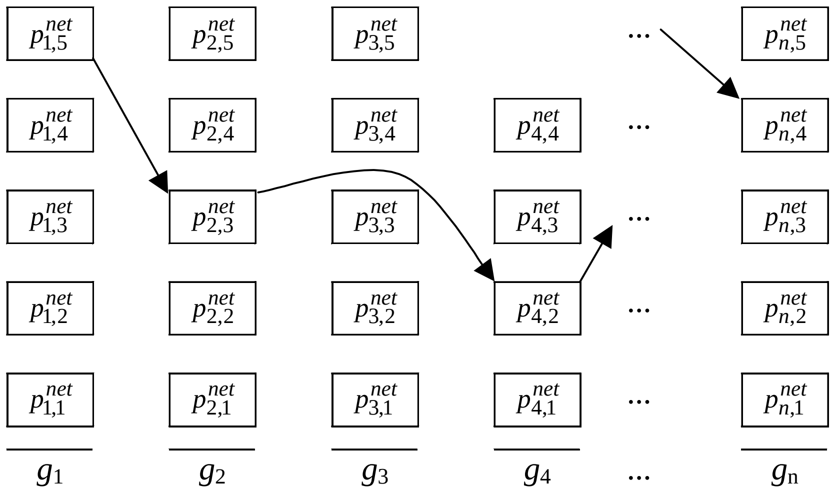


Figure 3: A solution (vehicle route) can be represented as a path through a state space of candidate network positions, where no more than 1 candidate position is chosen for each GPS data point.

2.1 Solution Representation

We can define a possible solution to the map matching problem by choosing a sequence of positions from our sets of candidate network positions. Figure 3 shows a representation of the candidate network positions $p^{net}_{c,i}$ defined for each GPS data point $g_c, c = 1, 2, ..., n$. We define a solution S as a sequence of candidate positions formed by choosing (up to) one candidate position for each GPS data point, along with the fastest paths between each consecutive pair of chosen candidate positions. A possible solution is shown by the path in Figure 3, while Figure 4 shows a segment of the route that might be specified by a consecutive pair of candidate positions. Note that in our model, we allow some of the n GPS data points to be rejected because they have very large errors, and thus we assume that the solution S contains $n_S \leq n$ network positions, denoted $p^{net}_{c_1,i_1}, p^{net}_{c_2,i_2}, ..., p^{net}_{c_{n_S},i_{n_S}}$. In this example solution, we have $c_1 = 1, i_1 = 5$; $c_2 = 2, i_2 = 3$; $c_3 = 4, i_3 = 2$ (skipping data point g_3), ..., and $c_{n_S} = n, i_{n_S} = 4$. For convenience, we will also denote a solution by $p^{net}_{[1]}, p^{net}_{[2]}, ..., p^{net}_{[n_S]}$ (where $p^{net}_{[j]} = p^{net}_{c_j,i_j}, j = 1, 2, ..., n_S$) with associated GPS data points $g_{[c]} = (t^{gps}_{[c]}, p^{gps}_{[c]}, v^{gps}_{[c]}, d^{gps}_{[c]})$.

The solution S states that the vehicle was at location $p^{net}_{[1]} = p^{net}_{c_1,i_1}$ at time

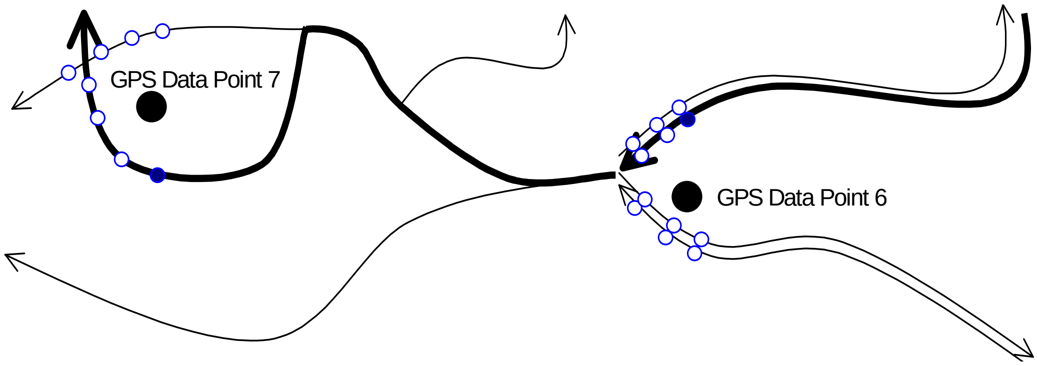


Figure 4: A partial vehicle route (shown in bold) is defined by a pair of candidate network positions (shown as solid circles).

$t^{gps}_{[1]} = t^{gps}_{c_1}$ (when it generated a GPS data point at location $p^{gps}_{[c]} = p^{gps}_{c_1}$), and then at location $p^{net}_{[2]}$ at time $t^{gps}_{[2]}$, and so on.

The measure that we use to distinguish solutions is the probability of the vehicle having been at the locations $p^{net}_{[1]}, p^{net}_{[2]}, ..., p^{net}_{[n_S]}$ given the GPS data points we have observed. Thus, we wish to know the probability (denoted by $P(p^{net}_{[1]}, ..., p^{net}_{[n_S]}|g_1, ..., g_{n_S})$) of these vehicle positions having occurred given the recorded GPS data points. We can estimate this as follows:

$$\begin{aligned}
& P(p^{net}_{[1]}, p^{net}_{[2]}, ..., p^{net}_{[n_S]}|g_1, g_2, ..., g_n) \qquad (1)\\
\approx\ & P(p^{net}_{[1]}|g_1, g_2, , ..., g_n) \times P(p^{net}_{[2]}|g_1, g_2, , ..., g_n) \times ... \times P(p^{net}_{[n_S]}|g_1, g_2, , ..., g_n)\\
\approx\ & P(p^{net}_{[1]}|g_{[1]}) \times P(p^{net}_{[2]}|p^{net}_{[1]}, g_{[2]}) \times P(p^{net}_{[3]}|p^{net}_{[2]}, g_{[3]}) \times ... \times P(p^{net}_{[n_S]}|p^{net}_{[n_S-1]}, g_{[n_S]})\\
& \times P_{\text{skip}}(g_1, g_2, , ..., g_n, g_{[1]}, g_{[2]}, ..., g_{[n_S]})
\end{aligned}$$

In this expression we have decomposed the probability function into a product of probabilities $P(p^{net}_{[c]}|p^{net}_{[c-1]}, g_{[c]})$ that each depends only on successive pairs of vehicle locations/data points, and a term $P_{skip}(g_1, g_2, , ..., g_n, g_{[1]}, g_{[2]}, ..., g_{[n_S]})$ that incorporates the impact of skipping some GPS data points. The main assumption in this simplification is that, at each step, we can ignore any impact on the route of the destination that is implied by the later GPS data points.

The general term $P(p^{net}_{[c]}|p^{net}_{[c-1]}, g_{[c]})$ in our product is the probability that the vehicle was at location $p^{net}_{[c]}$ at time $t^{gps}_{[c]}$ given that the GPS data point $g_{[c]}$ was generated at this time and that the vehicle was at location $p^{net}_{[c-1]}$ at time $t^{gps}_{[c-1]}$. We need to estimate this probability, and the $P_{skip}(g_1, g_2, , ..., g_n, g_{[1]}, g_{[2]}, ..., g_{[n_S]})$ factor, by taking a number of considerations into account.

1. *Normally Distributed GPS Errors*: We need to develop an estimate for the probability $P_{\text{location}}(p^{net}_{[c]}, p^{gps}_{[c]})$ of the vehicle being at location $p^{net}_{[c]}$ given the reported GPS location of $p^{gps}_{[c]}$. Now, the position errors in the GPS system are generally assumed to be normally and independently distributed in the x and y directions. Thus, if $p^{net}_{[c]}$ is a network position at coordinates $(x^{net}_{[c]}, y^{net}_{[c]})$ that corresponds with the GPS data point at $p^{gps}_{[c]} = (x^{gps}_{[c]}, y^{gps}_{[c]})$, then

$$P_{\text{location}}(p^{net}_{[c]}, p^{gps}_{[c]}) \quad = \quad P(p^{net}_{[c]}|p^{gps}_{[c]})$$

$$
\begin{aligned}
&= \frac{P(p^{gps}_{[c]}|p^{net}_{[c]})P(p^{net}_{[c]})}{P(p^{gps}_{[c]})} \\
&\propto P(z_x = x^{net}_{[c]} - x^{gps}_{[c]}) * P(z_y = y^{net}_{[c]} - y^{gps}_{[c]}),
\end{aligned}
$$

where we have assumed all network positions $p^{net}_{[c]}$ are equally likely, $P(p^{gps}_{[c]}) = 1$, and $z_x \sim N(0, \sigma)$, $z_y \sim N(0, \sigma)$, with σ being provided by the GPS equipment operators.

2. *GPS Speed & Heading*: A calculation of the probability $P_{\text{speed}}(p^{net}_{[c]}, v^{gps}_{[c]})$ of the vehicle having a particular speed given by $v^{gps}_{[c]}$ at location $p^{net}_{[c]}$ can be made if speed information is available in the GPS data. A similar calculation can be made for the probability $P_{\text{heading}}(p^{net}_{[c]}, v^{gps}_{[c]})$ of obtaining the observed vehicle heading.

3. *Observed Travel Time and Distance*: We assume that the vehicle travels by the fastest path between points $p^{net}_{[c-1]}$ and $p^{net}_{[c]}$. We can then define $P_{\text{time}}(p^{net}_{[c-1]}, t^{gps}_{[c-1]}, p^{net}_{[c]}, t^{gps}_{[c]})$ to be the probability that a fastest-path trip that starts at location $p^{net}_{[c-1]}$ at time $t^{gps}_{[c-1]}$ will arrive at location $p^{net}_{[c]}$ at time $t^{gps}_{[c]}$. (This calculation would have to assume some distribution for travel times; this is an area of on-going research.) Similarly, we let $P_{\text{distance}}(p^{net}_{[c-1]}, t^{gps}_{[c-1]}, p^{net}_{[c]}, d^{gps}_{[c]})$ be the probability that this trip will have length $d^{gps}_{[c]}$.

4. *Directness*: Drivers typically have a preference for faster more direct routes. To recognise this, we wish to give faster routes a higher probability score than slower routes. This is important because, in our travel model, we assume that U-turns occur only at intersections, and thus the directions associated with network locations $p^{net}_{[c-1]}$ and $p^{net}_{[c]}$ can significantly impact the travel time $T(p^{net}_{[c-1]}, t^{gps}_{[c-1]}, p^{net}_{[c]})$ between these locations. Clearly, some choices of direction at these locations will give much longer travel times, and so we wish to penalise these 'unlikely' choices of direction. To achieve this, we introduce a penalty 'probability' factor $P_{\text{directness}}(p^{net}_{[c-1]}, t^{gps}_{[c-1]}, p^{net}_{[c]}) \propto \exp(-k{*}T(p^{net}_{[c-1]}, t^{gps}_{[c-1]}, p^{net}_{[c]}))$ which decreases with increasing travel time. (Increasing the parameter k makes short routes look increasingly more favourable.) This function behaves correctly in the sense that if a trip is broken into three sub-trips, the product of the penalties for the 3 sub-trips is the same as the penalty computed for the original full length trip.

5. *Skipped Points*: A solution S comprised of n_S network positions for a trip with n GPS data points will have skipped $n_S - n$ of the GPS data points. Clearly, skipping data points is not a good use of the data available, but is required to determine routes in the presence of large GPS errors. Therefore, we assume, with some low probability, p_{invalid}, that any GPS data point is invalid, and incorporate $P_{\text{skip}}(g_1, g_2, , ..., g_n, g_{[1]}, g_{[2]}, ..., g_{[n_S]})$ by introducing a term $P_{\text{skip}}(n, n_S) = (p_{\text{invalid}})^{n-n_S}$ that allows data points to be skipped. (To see this more formally, we assume the probability distribution $P(p^{gps}_{[c]}|p^{net}_{[c]})$ is a weighted sum of a normal distribution centred on $p^{net}_{[c]}$ and a second

distribution that is uniformly distributed over all possible (x, y) locations in the geographical area of interest.)

Taking these terms together, we see that we can write our objective function (1) in the form

$$P(p_{[1]}^{net}, p_{[2]}^{net}, ..., p_{[n_S]}^{net} | g_1, g_2, ..., g_{n_S}) \propto P_{\text{skip}}(n, n_S) \times P_{\text{location}}(p_{[1]}^{net}, p_{[1]}^{gps}) \times P_{\text{heading}}(p_{[1]}^{net}, v_{[1]}^{gps}) \times P_{\text{speed}}(p_{[c]}^{net}, v_{[c]}^{gps}) \times \prod_{c=2}^{n_S} P(p_{[c]}^{net} | p_{[c-1]}^{net}, g_{[c]}) \quad (2)$$

where

$$P(p_{[c]}^{net} | p_{[c-1]}^{net}, g_{[c]}) = P_{\text{location}}(p_{[c]}^{net}, p_{[c]}^{gps}) \times P_{\text{heading}}(p_{[c]}^{net}, v_{[c]}^{gps}) \times P_{\text{speed}}(p_{[c]}^{net}, v_{[c]}^{gps}) \times P_{\text{duration}}(p_{[c-1]}^{net}, t_{[c-1]}^{gps}, p_{[c]}^{net}, t_{[c]}^{gps}) \times P_{\text{distance}}(p_{[c-1]}^{net}, t_{[c-1]}^{gps}, p_{[c]}^{net}, d_{[c]}^{gps}) \times P_{\text{directness}}(p_{[c-1]}^{net}, t_{[c-1]}^{gps}, p_{[c]}^{net})$$

and, as before, $g_{[c]} = (t_{[c]}^{gps}, p_{[c]}^{gps}, v_{[c]}^{gps}, d_{[c]}^{gps})$. Note that the first term $P(p_{[1]}^{net} | g_{[1]})$ in (1) has been handled differently as it has no dependence on an earlier point.

A fundamental assumption of this model is the use of fastest paths to interpolate between GPS data points. This gives rise to an unavoidable inconsistency in our model in that we assume vehicles travel by fastest paths between data points, but that the trip as a whole does not necessarily follow the fastest path. This conflict is made explicit in our directness penalty. If a large weight is placed on this, then paths that are short but skip many GPS data points will score more highly than those paths that faithfully follow the GPS data, even when the date contains errors. Careful parameter selection can help avoid this.

Our probability model provides a flexible framework that can be extended in a number of ways. For example, we note that the probability of a GPS unit generating an invalid data point is much greater in some areas than in others. For example, points generated upon entering or leaving a tunnel often give locations many kilometers away from the correct location. Thus, the handling of skipped points could be improved by computing an 'invalid data point' probability that would depend on the location of the vehicle when the GPS data point was generated. This would involve calculating a path-specific 'skipped points' probability for each of the route segments $p_{c_j,i_j}^{net} \longrightarrow p_{c_{j+1},i_{j+1}}^{net}$ that contain skipped points (i.e. have $c_{j+1} > c_j + 1$). This calculation would require predicting a vehicle location for each of the times $t_{c_j+1}^{gps}, t_{c_j+2}^{gps}, ..., t_{c_{j+1}-1}^{gps}$ associated with the skipped GPS data points, and then using these locations to compute location-specific probabilities of erroneous GPS locations being generated.

2.2 Dynamic Program

Consider now the problem of constructing a solution S that best matches the given GPS data. Once the candidate network positions have been constructed for a trip, the problem becomes that of choosing one candidate position for each GPS data point (or deciding to skip that data point) so that these positions together give the maximum likelihood locations for the vehicle during the trip based on the probability measure given in (2). The product form of this equation allows us to develop a dynamic programming recursion to solve this problem. Let $f(p_{c,j}^{net})$ be the probability (objective function value) associated with the most likely partial route that ends at network location $p_{c,j}^{net}$ taking into account GPS data points $g_1, g_2, ..., g_c$. Then, for $c = 2, 3, ..., n$,

$$f(p_{c,j}^{net}) = \max_{b=1,2,...,c-1} (p_{\text{invalid}})^{c-b-1} \max_{i=1,2,...,|\mathrm{P}_b^{net}|} f(p_{b,i}^{net}) P(p_{c,j}^{net} | p_{b,i}^{net}, g_c).$$

The initial value is defined by $f(p_{1,j}^{net}) = P_{\text{location}}(p_{1,j}^{net}, p_{1,j}^{gps}) \times P_{\text{heading}}(p_{1,j}^{net}, v_{1,j}^{gps}) \times P_{\text{speed}}(p_{1,j}^{net}, v_{1,j}^{gps})$, $j = 1, 2, ..., |\mathrm{P}_1^{net}|$. The optimal objective function is then given by $\max_{j=1,2,...,|\mathrm{P}_n^{net}|} f(p_{n,j}^{net})$. If we take logarithms of the terms in the objective function, we see that this problem is equivalent to that of finding a shortest path through the state space network illustrated in Figure 3.

3 Implementation and Results

The dynamic program has been implemented and tested within the Siren ambulance simulation system [3] using data for Perth, Australia. The GPS data lacked heading, speed and distance information, and so the terms associated with these in the objective function were dropped.

There are a number of algorithmic efficiencies to consider when implementing this algorithm. The main computational burdens are (1) generating the candidate network positions, and (2) constructing the fastest path routes between each pair of candidate network positions. The first of these steps requires identifying arcs that are close to the GPS data point. To avoid searching through all arcs, we break the area into a rectangular grid of cells, and pre-compute all the arcs that intersect each cell. Given the limited objective function we are using, we have found that recording only one candidate network position for each arc still gives good results. The other expensive step is generating the fastest paths between pairs of candidate network positions. In practice, we do not consider skipping more than 2 successive GPS data points, and thereby restrict the number of fastest paths we need to construct.

Figure 5 shows a sample of the GPS data points collected from the vehicles. (In this plot, the points have been coloured using a simple straight-line estimate of vehicle speed.) The area shown has a tunnel located near the centre of the figure, which is responsible for the scattering of erroneous data points in this area. Figure 6 shows what can go wrong if the GPS data is followed too faithfully, even if the GPS errors are very small, while Figure 7 shows a more typical route that has been successfully re-constructed from the GPS data.

There are a number of efficiency improvements currently being implemented to our system. For example, we note that we can use a one-to-many shortest path algorithm to efficiently compute the fastest paths from a candidate network position to all other next candidate positions. (Only one such tree would be required for all those starting candidate positions that share the same arc.) We could also use such a

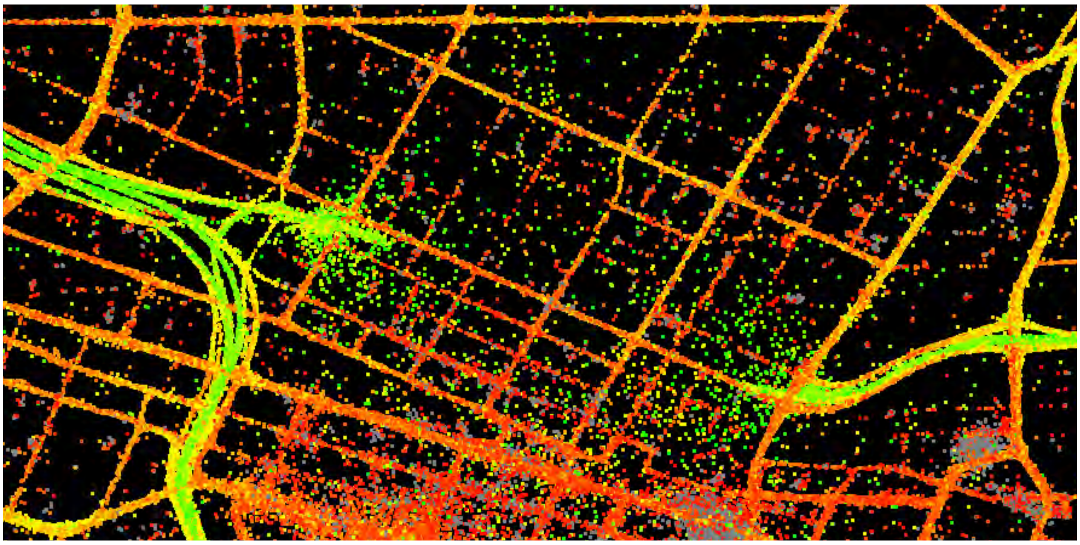


Figure 5: GPS Data Points collected by St John Ambulance in Perth

Figure 6: This figure shows an example of a badly map-matched route with a loop caused by matching to the closest (but incorrect) arc where lanes in each direction are physically separated. This problem was fixed by increasing the importance of 'directness' and increasing the variance in the GPS position estimates.

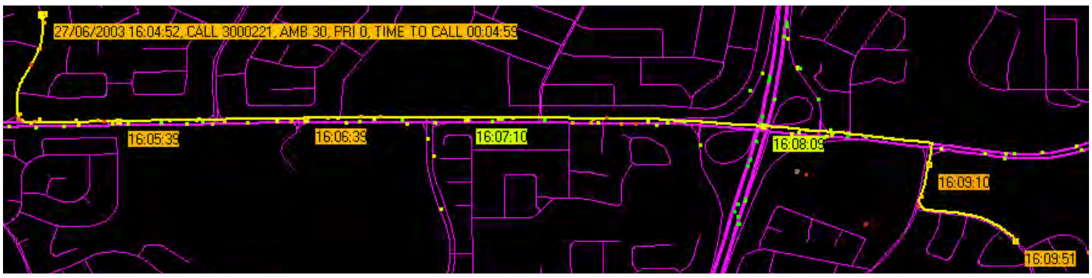


Figure 7: An example of a map-matched route. The GPS data points are labelled with their time stamps.

tree to compute the candidate network positions, perhaps limiting the tree's growth by noting the time when the next GPS data point will be encountered. This last change would make our algorithm a formal generalisation of the algorithms found in the literature (e.g. [2]) that rely on there being only a short distance between successive points.

Conclusions

We have developed a formal dynamic programming approach to the map matching problem that is well suited to the problems faced when analysing the sparse GPS data typically generated by ambulance services. We have validated this approach using data from St John Ambulance in Perth and are currently implementing algorithmic improvements aimed at reducing run times.

Acknowledgments

The author would like to acknowledge his colleagues at Optimal Decision Technologies Ltd (ODT) who supported the research and testing required for this work. The author also thanks Perth's St John Ambulance for providing the data, and both Otto Neilsen and René Jørgensen from the Technical University of Denmark (DTU), and Shane Henderson from Cornell University for their valuable personal communications regarding this problem.

References

[1] F. Marchal, J. Hackney, and K.W. Axhausen. *Efficient Map-Matching of Large GPS Data Sets - Tests on a Speed Monitoring Experiment in Zurich*, volume 244 of *Arbeitsbericht Verkehrs und Raumplanung.* ETH Zürich, Zürich, 2004.

[2] Otto Anker Nielsen. Map-matching algorithms for GPS data - methodology and test on data from the AKTA roadpricing experiment in Copenhagen. Technical report, Centre for Traffic and Transport (CTT), Technical University of Denmark (DTU), 2004.

[3] Optimal Decision Technologies Ltd. Simulation for Improving Responses in Emergency Networks (Siren) Software, 2005. www.Optimal-Decision.com.

[4] K.W. Axhausenand S. Schönfelder, J. Wolf, M. Oliveira, and U. Samaga. Eighty weeks of GPS traces: Approaches to enriching trip information. Technical report, IVT / ETH, CH 8093 Zürich, 2003. Presented at the 83rd Transportation Research Board Meeting, Washington, D.C., Jan. 11-15, 2004.

A Set-packing Approach to Routing Trains Through Railway Stations

Rafael Velásquez
Fraunhofer Institute for Industrial Mathematics
Kaiserslautern, Germany
velasque@itwm.fhg.de

Matthias Ehrgott, David Ryan
Department of Engineering Science
The University of Auckland
Auckland, New Zealand
{m.ehrgott,d.ryan}@auckland.ac.nz

Anita Schöbel
Institut für Numerische und Angewandte Mathematik
Universität Göttingen
Göttingen, Germany
schoebel@math.uni-goettingen.de

Abstract

Routing trains through a railway station consists in the assignment of a set of trains to routes that pass through a railway station or railway junction. This problem occurs on the strategic, tactical, and operational planning levels with different goals. A set-packing model is proposed and column generation and a constraint branching technique are used to solve it. Good results were obtained for a test example from the literature.

1 Introduction

In many European countries railways play an improtant role in the public and freight transport sytems. Extensive rail networks exist and major stations and railway junctions have a complex infrastructure. A problem that arises in this context is that of assigning a set of trains to routes through a railway station or junction over a period of time. The route assignment requires that no pair of trains is in conflict, i.e. that no pair of trains is scheuled to use the same track section at any point in time.

The layout of a railway station consists of a set of track sections which can be used by the trains. Trains enter the railway station through track sections which are part of a group of *entering points* and they leave through track sections which belong to a group of *leaving points*. Within these points, the railway infrastructure consists of *normal track sections* and *platform track sections*, a sequence of which defines a *route*. This route is associated to a timetable – which for the routing problem is assumed to be given – that consists of an *arrival time* and a *departure time* at a platform track section.

The train routing problem (TRP) occurs in three planning levels (see [9]). Depending on the planning goal, it can be of either *strategic, tactical*, or *operational* nature. The strategic planning level gives an answer to future capacity requirements of a station: given a current set of trains being routed through a station, the *feasibility problem* is solved for an expected traffic increase in future years. The feasibility problem consists in determining whether all trains scheduled in a timetable can be assigned a route through a station with its given layout. The tactical planning level deals with the present time and the actual generation or validatin of timetables for the trains that go through the railway station. Finally, at the operational level, day-to-day disturbances (such as delayed trains) and their effect on current timetables are considered and resolved such that all scheduled trains make their way through the railway junction.

Previous approaches to deal with the train routing problem include the research done by [3] who solve it as a graph coloring problem on special graph classes. Graph coloring and integer programming strategies can be found in the work of [2]. Closer to the contents of this paper is the modelling of the TRP as a set-packing problem (SPP) which [9] make use of at the strategic planning level, while [4] and [8] apply it to the tactical planning level.

The main objective of this paper is to model the problem in a way that is analogous to set partitioning models for crew rostering problems. This representation allows us to solve the problem in all three planning level hierarchies. In the next two sections, the formulation will be described and the solution approach will be presented.

2 The Train Routing Problem

The main difference between the model of [4] and [8] and the approach presented in this paper consists in the way conflict situations are modelled. In the TRP conflict is a scenario in which two or more trains compete for a a track section at the same moment in time. [4] and [8] define tuples of train-route pairs which are in conflict. By allowing only one of the two conflicting routes of trains to be assigned, a solution to the TRP will be conflict free. In the formulation presented in this section, the conflict is avoided by allowing each track section to be part of at most one route for one train at any one time. In fact, time is discretized to some reasonably chosen interval length (e.g. one second, 15 seconds, or one minute).

For this, let $\mathcal{T}$ represent the set of trains that need to be routed through the railway station. Furthermore, let $\mathcal{S}$ represent the set of track sections given by the layout of the railway station, and $\mathcal{H}$ be the set representing the time blocks corresponding to the discretization of the planning horizon. Moreover, let

$$1_{(t,r)^{(h,s)}} = \begin{cases} 1 & \text{if train } t \text{ on route } r \text{ crosses track section } s \text{ at time block } h \\ 0 & \text{otherwise} \end{cases}$$

and define the decision variables

$$x_{(t,r)} = \begin{cases} 1 & \text{if train } t \text{ follows route } r \text{ through the railway station} \\ 0 & \text{otherwise.} \end{cases}$$

With cost coefficients $c_{(t,r)}$ representing a preference for a route r of train t, the mathematical programming model for the TRP becomes

$$\max \sum_{t=1}^{|\mathcal{T}|} \sum_{r=1}^{n_t} c_{(t,r)} x_{(t,r)} \tag{1}$$

$$\text{subject to} \sum_{r=1}^{n_t} x_{(t,r)} \leq 1 \qquad \text{for all } t \in \mathcal{T} \tag{2}$$

$$\sum_{t=1}^{|\mathcal{T}|} \sum_{r=1}^{n_t} 1_{(t,r)^{(h,s)}} x_{(t,r)} \leq 1 \qquad \text{for all } s \in \mathcal{S}, h \in \mathcal{H} \tag{3}$$

$$x_{(t,r)} \in \{0,1\}, \tag{4}$$

where n_t is the number of possible routes for train t.

(1) maximizes preference of train routes (or number of trains if $c_{(t,r)} = 1$) scheduled through the railway station.

(2) are *train constraints* which ensure that only one route is assigned to every train.

(3) are *track section constraints* which guarantee that each track section at a period of time is only used by one train.

(4) defines the binary character of the decision variables.

The analogy to crew rostering models now becomes apparent: In crew rostering, $\mathcal{T}$ would be a set of crew members, $\mathcal{S} \times \mathcal{H}$ would be a set of tasks to be assigned to crew. (3) with equality constraints then guarantees that every task is assigned to a crew member, and (2) with equality constraints ensures each crew member needs to carry out exactly one set of tasks. The variables in the model correspond to sets of tasks that can be performed by a particular crew member, e.g. a line of work consisting of a number of tours of duty in airline crew scheduling [6].

Note that the TRP is an SPP and that the coefficient matrix has a distinguishable structure, both column and row-wise. Row-wise, the two main blocks are the previously mentioned train and track section constraints while column-wise, the coefficient matrix consists of $|\mathcal{T}|$ different column blocks, one for every scheduled train. Finally, notice that the feasibility problem is simply the TRP formulation with $c_{(t,r)} = 1$ and it needs to be checked whether the optimal solution is equal to $|\mathcal{T}|$ or not, thus giving an affirmative or negative answer to the feasibility problem, respectively.

3 Solving the TRP

The SPP is an NP-complete problem according to [5] and thus no polynomial time algorithm is known. The solution approach described in this paper uses linear programming based branch and bound with column generation and constraint branching.

This builds on experience with solving large set partitioning problems in crew rostering [6]. It is known that the block of columns refering to a single train in constraints (2) and (3) defines a perfect matrix, and therefore no fractional solutions can occur with only a single block of variables.

Column generation is used to avoid enumeration of all possible routes for a train. Thus, the LP relaxation of (1) – (4) can be solved with a small set of decision variables. The restricted model starts with any subset of the decision variables of the master LP that contains a feasible solution. A pricing step follows that calculates the reduced cost of candidate columns of decision variables which are not part of the restricted model yet. Letting $a_{(t,r)}$ represent the coefficients for the column of the coefficient matrix of the master problem corresponding to train t and route r, and π be a vector of dual variables obtained after solving the restricted problem to optimality, the reduced cost for that column is given by

$$rc(a_{(t,r)}) = c_{(t,r)} - \pi^T a_{(t,r)}.$$

The column that yields the largest positive reduced cost is then added to the restricted problem and the pricing step repeated. When no candidate column leads to a positive reduced cost, then the solution x^* solves both the restricted and the master LP optimally. For the TRP, the column generation subproblem can be formulated as a shortest path problem.

In the previous step, a solution is found for a problem where the integrality constraints have been dropped. The TRP is, however, a binary program, thus in order to achieve integrality, a branch and bound procedure based on constraint branch is applied. Traditional variable branching is ineffective for our set-packing problems – the resolution of fractional solutions at the optimal solution of a relaxed SPP leads to a very large and unbalanced branching tree (see [7]). The constraint branching developed by [7] consists in identifying two constraints, $\hat{t}$ and $\hat{g}$, such that the following relation holds:

$$0 < \sum_{(t,r)\in J(\hat{t},\hat{g})} x_{(t,r)} < 1. \tag{5}$$

Here, the set $J(\hat{t},\hat{g})$ is defined as the set of columns of the coefficient matrix which have non-zero coefficients for constraints $\hat{t}$ and $\hat{g}$ or $J(\hat{t},\hat{g}) = \{(t,r) | a_{(\hat{t}(t,r))} = 1$ and $a_{(\hat{g}(t,r))} = 1\}$. In the TRP, any fractional solution will have at least one such pair of constraints (see [1]), in fact due to the perfect blocks in the constraint matrix, fractional variables can only occur due to two trains competing for the same track segment at the same time so that one of the two constraints can be selected from the train constraints (2) and the other from the track section constraint (3).

Having selected constraints $\hat{t}$ and $\hat{g}$, branching is enforced through the 1-branch ($\sum_{(t,r)\in J(\hat{t},\hat{g})} x_{(t,r)} = 1$) and the 0-branch ($\sum_{(t,r)\in J(\hat{t},\hat{g})} x_{(t,r)} = 0$), i.e. train $\hat{t}$ uses the track segment at the time or not. With this type of branching, a simultaneous elimination of many variables is achieved on each side of the branch, leading to a smaller tree. Moreover, as shown in [6], if $\hat{t}$ and $\hat{g}$ are chosen such that $\sum_{(t,r)\in J(\hat{t},\hat{g})} x_{(t,r)}$ is maximized (i.e. chosen as close to 1 as possible), a depth-first 1-branch will lead to a fast and good initial integer solution.

4 Tactical and Operational TRP

When the feasibility problem (strategic TRP) returns a negative answer, a simple way of trying to achieve feasibility is by taking individual routes within the current timetable and force them to enter the railway station a certain number of time periods before or after their initial schedule. This leads to the use of a time shift parameter δ, which was introduced by [9], and which describes the number of time blocks the arrival/departure time of a train is rushed or delayed. The maximum number of time blocks a train t is allowed to be rushed or delayed will be given by δ_t.

The previously defined decision variables are thus extended to $x^{\delta}_{(t,r)}$ and take the value of one if route r is selected for train t, shifting the train's given timetable δ units forward. The resulting formulation for the tactical TRP (t-TRP) is given by:

$$\max \sum_{\delta=-\delta_t}^{\delta_t} \sum_{t=1}^{|\mathcal{T}|} \sum_{r=1}^{n_t} c^{\delta}_{(t,r)} x^{\delta}_{(t,r)} \tag{6}$$

$$\text{subject to} \sum_{\delta=-\delta_t}^{\delta_t} \sum_{r=1}^{n_t} x^{\delta}_{(t,r)} \leq 1 \qquad \text{for all } t \in \mathcal{T} \tag{7}$$

$$\sum_{t=1}^{|\mathcal{T}|} \sum_{\delta=-\delta_t}^{\delta_t} \sum_{r=1}^{n_t} 1_{(t,r)^{(h,s)}} x^{\delta}_{(t,r)} \leq 1 \qquad \text{for all } h \in \mathcal{H}, s \in \mathcal{S} \tag{8}$$

$$x^{\delta}_{(t,r)} \in \{0,1\}. \tag{9}$$

The objective function used for the t-TRP aims at giving a preference to a resulting timetable that is as close as possible to the given but infeasible one. It penalizes a schedule proportional to its deviation from the original one as follows:

$$c^{\delta}_{(t,r)} = 1 - (0.1 \cdot |\delta|).$$

In practice, a feasible timetable is not always possible to achieve due to circumstances that arise along the itinerary of a train (e.g. malfunctioning of a locomotive, weather hazards). The resulting delays lead to possible infeasibility of the initially assigned routes of all trains, as they were feasible and optimal only for the scheduled arrival and departure times. Modelling the operational TRP (o-TRP), i.e. the TRP with delayed trains, is in essence a particular case of t-TRP. The outcome desires to achieve feasibility for a new timetable, namely the one composed of all on-time trains with their assigned routes and schedule (e.g. after solving the t-TRP), and the new arrival and departure times for the set of delayed trains.

In this sense, o-TRP consists of two phases: the first is a decision problem in which it has to be determined whether the delayed train can follow its assigned route despite the current delay. Should this yield a negative answer, then a new feasible assignment has to be looked for. This may involve enlarging the delay of already delayed trains, delay on-time trains or relocate routes for some or all trains. This delay is modelled in the same way as for the t-TRP, with the exception that trains can only be delayed with respect to their scheduled arrival/departure time as otherwise (rushing the arrival/departure time) boarding passengers are affected.

5 TRP – A Test Example

As mentioned in the introduction, the TRP deals with three major questions: first, the feasibility of assigning a set of trains through the railway station under a given timetable. Second, should the feasibility problem yield a negative solution, then the trains' timetable will be shifted such that a route can be assigned for all trains. Finally, the third question finds an answer to what the routes of all trains will be like, when a set of trains is delayed.

For the test case of the Pierrefitte-Gonesse railway junction – a big intersection between four major destinations in France – the route a scheduled train follows along the track sections is assumed to be given. Figure 1 shows the layout, track connections and directions a train may take to cross the railway junction.

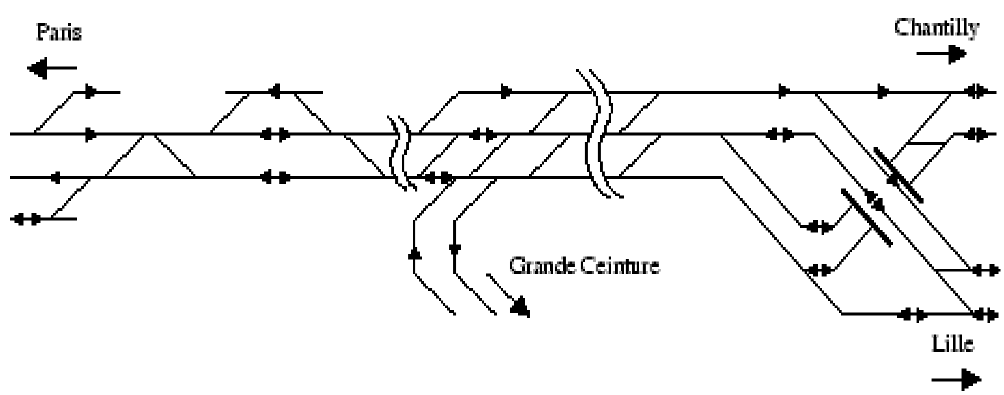


Figure 1: Layout of the test junction Pierrefitte-Gonesse (France).

As described before, the railway junction is subdivided into individual track sections. The 27 track sections for the test junction can be seen in Figure 2. Furthermore, the routes of the considered trains are those connecting Paris Gare du Nord and Lille (over track sections 1-18-2-3-5 and 16-22-17-25-9-14-19-15), Paris Gare du Nord and Chantilly (over track sections 1-18-2-3-4 and 11-20-12-23-10-13-14-19-15) and Grande Ceinture and Chantilly (over track sections 6-21-9-7-10-8-3-4 and 11-20-12-24-25-26-27).

The planning horizon for the test instance corresponds to 18.5 minutes. It is discretized in time blocks of 15 seconds duration, which means that the set $\mathcal{H} = \{0, 1, 2, \ldots, 74\}$ represents the time intervals $[0, 15), [15, 30), \ldots, [1110, 1125)$, with the time units being seconds. The travel time of a train on a particular

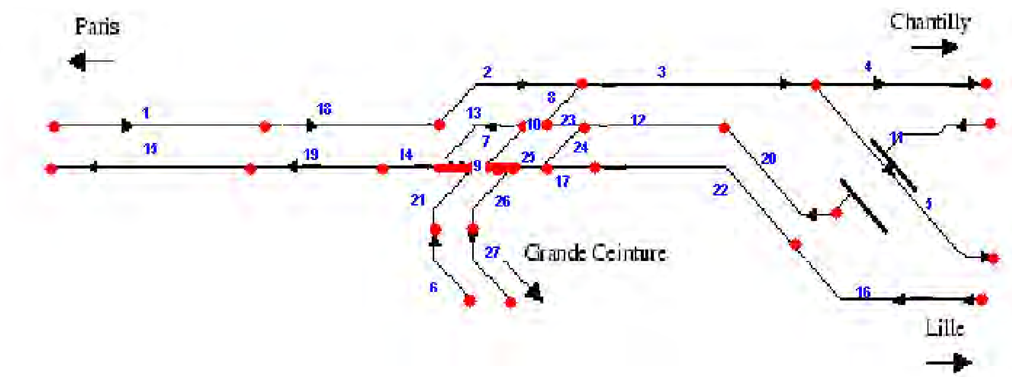


Figure 2: Layout of the junction Pierrefitte-Gonesse subdivided into track sections.

track section is fixed and given according to the speed limit imposed for each track section. It corresponds to the time span (set of time blocks) a track section will be set aside for a train to cross it, along its scheduled route. These values are given in Table 1.

Table 1: Travel times for trains crossing individual track sections of the Pierrefitte-Gonesse junction.

Track Section (No.)	Travel Time (Time Units)	Track Section (No.)	Travel Time (Time Units)	Track Section (No.)	Travel Time (Time Units)
1	6	10	2	19	6
2	6	11	6	20	6
3	4	12	1	21	4
4	4	13	2	22	4
5	6	14	1	23	1
6	4	15	6	24	4
7	4	16	6	25	1
8	4	17	1	26	4
9	1	18	6	27	4

As mentioned before, the TRP assumes that a timetable for all trains that need to be routed is given. Table 2 summarizes this input data for the test instance, identifying the eight trains that will cross the railway junction during the planning horizon as well as the origin and destination of the train. As Pierrefitte-Gonesse is a railway junction and not a station, arrival and departure times correspond to the time the train enters the first track section on its route, while the departure time corresponds to the time period in which it leaves this track section. Finally, the last column of Table 2 denotes the maximum number of time periods a train's scheduled arrival can be shifted either backward or forward. This data is used for the tactical planning level should the feasibility problem yield a negative answer. It provides degrees of freedom to alter, if necessary, a train's schedule.

Table 2: Train routing problem instance.

Train code	Origin (entering track No.)	Destination (leaving track No.)	Arrival Time	Departure Time	Slack time (Time Periods)
D1	Paris (1)	Chantilly (4)	12	18	2
D2	Lille (16)	Paris (15)	22	28	0
D3	Paris (1)	Lille (5)	23	29	6
D4	Chantilly (11)	Paris (15)	26	32	4
D5	Grande Ceinture (6)	Chantilly (4)	30	34	4
D6	Chantilly (11)	Paris (15)	33	39	2
D7	Lille (16)	Paris (15)	34	40	0
D8	Paris (1)	Lille (5)	36	42	2

6 Numerical Experience

All tests were run on a Pentium II PC with a 450 MHz Processor and 256 MB RAM. Given the timetable of Table 2, the feasibility problem was solved. Out of the eight trains, only train D7 could not be assigned to a route through the railway junction. The computational effort required was very small: the solution was obtained after only 0.3 seconds and the LP relaxation had an integer solution. This solution time is much faster than the computational time required by the formulation and solution approach in [4] for a nearly identical instance.

As it was not possible to assign a route to all scheduled trains under the current layout of the railway junction and the given timetable, the next question is whether the timetable can be altered such that the feasibility problem returns an affirmative answer.

To achieve this, the t-TRP was solved using the entries of column six of Table 2 as the maximum slack time or number of time periods which a train could have its schedule advanced or delayed. An optimal timetable that gives an affirmative answer for the feasibility problem was found and the results are given in Table 3. Here, column one states the code of the train, column two shows the arrival time under the original timetable, column three states the corrected arrival time which allows the feasibility problem to have an affirmative answer, and the fourth column lists the deviation of the new timetable from the original one. Recalling that train D7 was the train that led to the infeasibility of the initial timetable, it is interesting to note that in order to establish feasibility, two different trains had to be rescheduled. The computational effort was again very small, as it took 1 second to find the optimal solution. The optimal solution of the LP relaxation for this instance of the t-TRP was already integer, thus not requiring any branching.

After obtaining a feasible timetable from the previous step, two instances of o-TRP were considered to analyze the effect of a delay of trains on the resulting schedule. For the first, train D4 was set to be delayed by four time blocks, while for the second instance, trains D4 and D8 were set to be delayed by four time blocks. For each instance, it was assumed that all trains arriving after the first delayed train could be delayed if it were necessary.

Table 3: Modified timetable for the test instance.

Train	Original timetable (arrival time)	Modified timetable (arrival time)	Change
D1	12	12	0
D2	22	22	0
D3	23	23	0
D4	26	23	-3
D5	30	30	0
D6	33	35	2
D7	34	34	0
D8	36	36	0

The first phase returned a negative answer to the feasibility problem in both instances, meaning that under the current delay scenario, not all trains could be routed through the railway junction. This required solving the second phase of the o-TRP, which can lead to increasing the tobal delay in the system. The results of the second phase can be seen in Table 4, with columns three to five showing the results for the first of the instances and columns six to eight showing the results for the second instance. Columns four and seven show the corrected timetable which will allow all trains to be scheduled through the railway junction, taking into consideration the delays that had occurred before entering the railway junction (see columns three and six). Columns five and eight show the additional delay introduced to the system at the railway junction which is necessary to be able to route all trains through the railway junction. The CPU time for solving the second phase of the o-TRP was 0.3 seconds and again the LP relaxation optimal solution was integer.

Table 4: Results for the two test instances in which delayed trains were assumed.

		Instance A			Instance B		
Train	Scheduled timetable (arrival time)	Train Delay	Modified timetable (arrival time)	Added Delay	Train Delay	Modified timetable (arrival time)	Added Delay
D1	12	-	12	-	-	12	-
D2	22	-	22	-	-	22	-
D3	23	-	23	-	-	23	-
D4	23	4	27	0	4	27	0
D5	30	-	31	1	-	30	0
D6	35	-	39	4	-	39	4
D7	34	-	38	4	4	38	0
D8	36	-	37	1	-	36	0

7 Conclusions

A set-packing model was proposed for the TRP in its three planning levels. A solution method using column generation and constraint branching was implemented. A (small) example from the literature was used to test the proposed method. Determining the maximum number of trains that could be routed through the railway junction and proposing a modified timetable that would allow all trains to be routed was solved much faster than with the approach found in [4]. For the operational problem, two delay scenarios were modelled and solved very fast. Further research is directed towards applying the TRP formulation to larger railway stations or junctions as well as considering a larger planning horizon. Based on the experience gathered from this work and its promising results, solving larger TRP instances to optimality in relatively short time is expected.

References

[1] C. Barnhart, E.L. Johnson, G.L. Nemhauser, M.W.P. Savelsbergh, P.H. Vance, *Branch-and-price: Column generation for solving huge integer programs*, Operations Research 46:3 (1998), 316–329.

[2] A. Billionet, *Using integer programing to solve the train platforming problem*, Transportation Science 37:2 (2003), 213–222.

[3] S. Cornelsen, G. Di Stefano, *Platform assignment*, Proceedings of the 4th Workshop on Algorithmic Methods and Models for Optimization of Railways, Springer Verlag (to appear).

[4] X. Delorme *Modélisation et résolution de problèmes liés a l'exploitation d'infrastructures ferroviaires*, PhD Thesis (in French), Université de Valenciennes (2003).

[5] M.R. Garey, D.S. Johnson, *Computers and intractability: a guide to the theory of NP-Completeness*, W.H. Freemand and Company, San Francisco (1979).

[6] D.M. Ryan, *The solution of massive generalised set partitioning problems in aircrew rostering*, Journal of the Operational Research Society, 43:5 (1992), 459–467.

[7] D.M. Ryan, B.A. Foster, *An integer programming approach to scheduling*, in A. Wren (editor), Computer Scheduling of Public Transport, North-Holland, Amsterdam (1981).

[8] P.J. Zwaneveld, L.G. Kroon, S.P.M. van Hoesel, *Routing trains through a railway station based on a node packing model*, European Journal of Operational Research, 128 (2001), 14–33.

[9] P.J. Zwaneveld, L.G. Kroon, H.E. Romeijn, M. Salomon, S. Dauzèr-Pérès, S.P.M. van Hoesel, H.W. Ambergen, *Routing trains through railway stations: Model formulation and algorithms*, Transportation Science, 30:3 (1996), 181–195.

Markov Chain Monte Carlo Methods

Richard Arnold
School of Mathematics, Statistics and Computer Science
Victoria University of Wellington
New Zealand
richard.arnold@mcs.vuw.ac.nz

Abstract

Markov Chain Monte Carlo (MCMC) methods have become a very widely used means of estimating parameters in statistical models for which such estimation is otherwise intractable.

MCMC methods are particularly appealing in a Bayesian models where the main outcome of modelling is the posterior distribution $p(\theta|\mathbf{y})$ of the model parameters θ given the data $\mathbf{y}$. In many cases there is no closed form expression available for the posterior, or even if such a form exists it may be difficult to draw samples from that distribution. In both of these cases MCMC methods can be used to draw samples from $p(\theta|\mathbf{y})$.

In this paper we give a short introduction to the theory of MCMC sampling using the Metropolis-Hastings Algorithm, and demonstrate its practical application in various different settings. In particular we will demonstrate the special case of Gibbs Sampling, and its implementation in the WinBUGS software program.

Possible Ways of Analysis of Matched Case-Control Data with Multiple Ordered Disease States

Ivy Liu
School of Mathematics, Statistics, and Computer Science
Victoria University of Wellington
New Zealand
iliu@mcs.vuw.ac.nz
and
Bhramar Mukherjee
Department of Statistics
University of Florida
Gainesville, FL 32611

Abstract

In a typical matched case-control study, cases are matched to suitably chosen controls and simultaneous effects of potential risk factors are ascertained through conditional logistic regression. This article considers matched case-control studies where there is one control group, but there are multiple disease states with a natural ordering among themselves. This scenario can be observed when the cases can be further classified in terms of the seriousness or progression of the disease, for example according to different stages of cancer. We adopt a proportional odds model to account for the ordinal nature of the data. The important distinction from a dichotomous conditional logistic regression is that the stratum specific parameters cannot be eliminated in this model via the conditional likelihood approach. This article proposes two choices for analyzing such data. We also illustrate the methods by analyzing data from a matched case-control study on newborns with low birth weight where infants are classified according to "low" and "very low" birthweight and a child with normal birthweight serves as a control. A simulation study compares the methods.

1 Introduction

Case-control studies are important tools in investigating the etiology of rare diseases like cancer. The fundamental goal is to explore association between the disease and its potential risk factors. In modern medicine, with precise characterization of disease states in histological and morphological terms, it is natural to note that the disease state might have more than one category, i.e., we may have subdivisions within the "cases". For example, patients diagnosed with cancer may have cancer of stage-I, stage-II or stage-III at the time of the diagnosis which is an example of ordinal disease categories. There are several popular models for analyzing ordinal response as described by Agresti [1] which may be readily adapted to analyze unmatched case-control data with ordered disease categories.

However, matching is often implemented in a case-control design in order to avoid bias due to potential confounders. A proper statistical analysis should account for the matched design. Breslow *et al.* [5] proposed conditional logistic regression (CLR) for analyzing matched case-control data. Usual unconditional likelihood inference for such finely stratified data with large number of stratum specific nuisance parameters are potentially subject to the Neyman-Scott phenomenon. Conditioning on the complete sufficient statistic for the matched set specific parameters in each matched set, eliminates these nuisance parameters. Moreover, the maximum likelihood equation based on the conditional likelihood provides the optimum estimating equation [7]. Unconditional analysis of matched data generally yield conservative estimates of the relative risk.

Liang and Stewart [11] extended the usual CLR methodology to polychotomous disease or reference categories. They establish that separating the disease or reference states into subgroups and conducting CLR pairwise is less efficient than the polychotomous approach. The polychotomous approach allows for simultaneous estimation of the disease-specific parameters and direct hypothesis testing involving multiple disease or reference categories. Recently, Sinha, Mukherjee and Ghosh [19] consider a Bayesian semiparametric model for analyzing matched case-control data with multiple disease states and missingness in exposure values.

None of the above articles consider the situation when the underlying probability model for the disease incidence acknowledges a possible natural ordering of the disease states when it is present. They all incorporate the multinomial logistic regression model instead of the binary logistic regression model as the underlying probability scheme to handle multiple disease categories. The most commonly used model for ordinal logistic regression is the cumulative logit or the proportional odds model. One major difference between the multinomial logit model and the cumulative logit model when applied to matched samples is that in the latter, the nuisance parameters due to stratification cannot be eliminated by conditioning on the number of cases in each matched set. In absence of any reduction due to sufficiency, for matched pair data with ordinal response, McCullagh [14] and Agresti and Lang [2] propose fitting simultaneous conditional maximum likelihood estimates with binary collapsing of the ordinal response to eliminate the nuisance parameters and then using a weighted average of such conditional likelihood estimates.

In general, there could be several approaches to analyze highly stratified ordinal data. Liu and Agresti [12] propose a Mantel-Haenszel (MH) type estimate for the odds ratio in a cumulative logit model which can be adapted to matched 1:M matched case-control study with ordinal disease states and categorical exposure. Another possible alternative which can handle both categorical and continuous exposure is to consider a random effects model for stratification parameters akin to

Hedeker and Gibbons [8] and maximize the marginal likelihood to obtain relative risk estimates. To the best of our knowledge, there has not been any thorough investigation and comparison of these methods for analyzing matched case-control data with multiple ordered disease states. The goal of the paper is to explore some easy to use options that can be readily implemented by the practitioner and arrive at specific recommendations.

In this article we will consider two methods for analyzing matched case-control data with ordinal disease states: (i) A MH approach with an underlying fixed effect cumulative logit model for the probability of disease incidence, this method is applicable to only categorical exposure; (ii) A random effects cumulative logit model with a normal distribution on the matched set specific nuisance parameters. We compare these two ordinal methods with a choice of *ignoring* the ordinal nature of the data. We ignore the disease subgroups and simply collapse the disease categories into a single disease group and carry out usual binary conditional logistic regression analysis (CLR). A simulation study is conducted to assess the relative performance of these methods under varying degree of stratification.

The example considered in the paper involves a matched case-control dataset coming from a low birth weight study conducted by the Baystate Medical Center in Springfield, Massachusetts. The dataset is discussed in Hosmer and Lemeshow [10] (Section 1.6.2). The data was matched according to the age of the mother. Using the actual birth weight observations we divided the cases, namely, the low birth-weight babies into two categories, *very* low (weighing less than 2000 gms) and low (weighing between 2000 to 2500 gms) and tried to assess the impact of smoking habits of mother on the chance of falling in the two low birth-weight categories. We consider two matched datasets based on this low birth-weight study: a 1:1 matched dataset with 56 matched sets and a 1:3 matched dataset with 29 matched sets. Table 1 shows the data for both datasets, where the row variable is the smoking status (S: smoker and N: nonsmoker) and the column variable is the baby's weight with 3 levels (1: very low, 2: low, and 3: normal).

The rest of the paper is organized as follows. Section 2 considers MH estimation in stratified contingency tables with ordinal response when applied to matched case-control data with ordinal disease states. Section 3 considers a random effects model on the stratum specific parameters. The analysis of the real dataset by various methods is considered in each section separately. Section 4 presents the simulation study.

2 Mantel-Haenszel Estimation for Highly Stratified Ordinal Response

Classical methods for making inference on the odds ratio for pair-matched data with a dichotomous exposure are based on conditioning on the marginal totals of the resultant 2×2 tables and uses only the *discordant* pairs. Tests for hypothesis of no association between the disease and exposure in matched samples were proposed by McNemar [16]. The seminal paper by Mantel and Haenszel [13] furnished the MH estimate of the summary relative risk for stratified data and showed a chisquared test of conditional independence. Models for ordinal data started receiving attention in the 1960s and 1970s [20] [4]. However, the most popular model for ordinal data was inspired by McCullagh [15] by modeling the log odds corresponding to the cumulative probabilities, called the *cumulative logits* described as below. Let us consider a $1 : M$ matched case control study with n matched sets. Let Y_{ij}, $i = 1, \cdots, n$, $j = 1, \cdots, M + 1$, denote the disease status (for our example, the birth weight category) of subject j in matched set or stratum i. Also let Y_{ij} be a c-category ordinal variable with categories scaled from $1, \cdots, c$, with reference category c denoting the control group (newborns with normal birth weight) and the sub divisions within the diseased group denoted by category $1, \cdots, c - 1$. Let X_{ij} denote the observed categorical exposure observed for subject j in matched set i. For our example X_{ij} is a binary exposure denoting the smoking status of mother. The cumulative logit model for disease incidence is given by

$$\text{logit}[P(Y_{ij} \leq k|X_{ij})] = \alpha_k+\gamma_i+\beta X_{ij},\ i = 1, \cdots, n,\ j = 1, \cdots, M+1, k = 1, \cdots, c-1, \tag{1}$$

where $\alpha_1 < \alpha_2 < \cdots < \alpha_{c-1}$ and $P(Y_{ij} \leq c) = 1$. The parameters $\{\alpha_k\}$, called *cut points*, are usually nuisance parameters of little interest. The parameters $\{\gamma_i\}$ are stratum specific nuisance parameters, where the total number of these nuisance parameters increases with sample size. This particular type of cumulative logit model, with exposure effect β the same for all k, is often referred to as a *proportional odds model* [15]. For example, let $X_{ij} = 1$ for a smoker mother and $X_{ij} = 0$ for a non-smoker mother. The model assumes that the odds that Y_{ij} falls below level k for a smoker mother are $\exp(\beta)$ times the odds for a non-smoker mother, for $k = 1, 2$. We refer to $\theta = \exp(\beta)$ as the *cumulative odds ratio* for the conditional association of weight of newborn and smoking status of mother.

When the model holds, but n is large and the data are sparse, the maximum likelihood (ML) estimator of β tends to overestimate the true log odds ratio and leads to biased and inconsistent estimation. This is a phenomenon observed in models where the number of parameters (such as $\{\gamma_i\}$ in (1)) grows at the same rate as the sample size [17]. For instance, for a 1:1 pair matched case-control study, the ML estimator of β converges to double the true value [3]. Similar phenomenon occurs for the proportional odds model (1) when $c > 2$ [12]. For $c = 2$ as a remedy, conditional likelihood methods are implemented to eliminate the nuisance parameters. For $c > 2$, however, the proportional odds model (1) has no reduced sufficient statistics for the nuisance parameters. Therefore, the conditional ML method can not be applied. Based on the proportional odds model (1), Liu and Agresti [12] proposed a MH-type estimator of an assumed common cumulative odds ratio ($\exp(\beta)$), which is an extension of the ordinary MH estimator. Their estimator has behavior similar to the MH estimator of a common odds ratio for several 2×2 tables. It is consistent even when the data are sparse, that is, even when the number of matched sets (strata) increases proportional to the sample size, which is the case in an individually matched case-control study. The method is described below.

For a matched case-control study with c ordered disease categories and a binary exposure, we can cross-classify each matched stratum into a $2 \times c$ table, where the row variable is the exposure and the column variable is the disease. In total, there are n matched strata. Therefore, there are n such $2 \times c$ tables. We use notations

Z_{rki} to denote the cell counts for the row r, column k, and stratum i, where $r = 1, 2$, $k = 1, \ldots, c$, and $i = 1, \ldots, n$. Let n_{ri} be the total number of subjects in row r and stratum i. Let us also denote the cumulative counts in row r and stratum i by $Z^*_{rki} = Z_{r1i} + \ldots + Z_{rki}$. Also, let the total number of subjects in the ith stratum be denoted by $N_i = \sum_r \sum_k Z_{rki}$. Then Liu-Agresti's MH-type estimator equals

$$\hat{\theta} = \frac{\sum_{i=1}^{n} \sum_{k=1}^{c-1} Z^*_{1ki}(n_{2i} - Z^*_{2ki})/N_i}{\sum_{i=1}^{n} \sum_{k=1}^{c-1} (n_{1i} - Z^*_{1ki}) Z^*_{2ki}/N_i}\,. \tag{2}$$

For model (1), the same cumulative odds ratio occurs for all collapsings of the ordinal disease categories into the binary category ($\leq k$, $> k$), $k = 1, \ldots, c-1$. Suppose we naively treat $c - 1$ different 2×2 collapsed tables of each stratum as independent. The estimator (2) is simply the ordinary MH estimator of the common odds ratio for $n(c-1)$ separate 2×2 tables. However, the variance estimator of $\hat{\theta}$ needs to take the dependency of the collapsing tables for each stratum into account. The form of the variance estimator is given by Liu and Agresti [12].

Next, we will apply the MH-type estimator (2) to our example. For our first dataset, the 1:1 matched case-control study in Table 1, $\log(\hat{\theta}) = 0.969$ has a standard error estimate of 0.438. For each matched pair, the estimated odds of having a baby weighing below any fixed level for a smoker mother are $\exp(0.969) = 2.64$ times higher than the odds for a nonsmoker mother. In our second dataset, the 1:3 matched case-control study, $\log(\hat{\theta}) = 0.928$ (Odds Ratio (OR)=2.53) with a standard error estimate of 0.460.

3 Proportional Odds Models with Random Effects

As mentioned above, the unconditional ML estimator of β for the proportional odds model (1) performs poorly because the number of parameters increases at the same rate as the sample size. To reduce the number of parameters, one possible approach is to use random effects terms to describe the strata effects. It is natural when the strata are a sample, such as a sample of matched subjects in our case-control examples.

We now consider a random effects version of model (1) that treats the matched sets as a random sample from a larger population of matched sets. Assume $\{\gamma_i\}$ are independent observations from a $N(\gamma, \sigma^2)$ distribution. The independence assumption can be relaxed by considering a n-variate normal distribution for $(\gamma_1, \cdots, \gamma_n)$ with a known or unknown variance-covariance components, for simplicity we will assume that stratum effects are independent across different matched sets.

Identifiability of the mean of the random effects distribution requires a constraint such as $\alpha_1 = 0$. This model is an extension of the random-intercept logistic-normal model for binary data [18]. It is also a version of multivariate generalized linear mixed models that include random effects terms in a multivariate generalized linear model setting [6]. We will abbreviate this random effects model as REM.

To obtain the likelihood function we construct the usual product of multinomials. Without loss of generality, we assume that the first subject in each stratum is a case and the rest are controls (i.e., $Y_{ij} = c$, for $j = 2, \cdots, M+1$.) The likelihood for the proportional odds model (1) has the form

$$
\begin{aligned}
L(\beta, \alpha_k, \gamma_i) &= \prod_{i=1}^{n} \prod_{j=1}^{M+1} \left[\prod_{k=1}^{c} (P(Y_{ij} \le k | X_{ij}) - P(Y_{ij} \le k-1 | X_{ij}))^{I[Y_{ij}=k]} \right] \\
&= \prod_{i=1}^{n} \left[\prod_{k=1}^{c-1} \left(\frac{\exp(\alpha_k + \gamma_i + \beta X_{i1})}{1 + \exp(\alpha_k + \gamma_i + \beta X_{i1})} - \frac{\exp(\alpha_{k-1} + \gamma_i + \beta X_{i1})}{1 + \exp(\alpha_{k-1} + \gamma_i + \beta X_{i1})} \right)^{I[Y_{i1}=k]} \right. \\
&\quad \left. \prod_{j=2}^{M+1} \left(1 - \frac{\exp(\alpha_{c-1} + \gamma_i + \beta X_{ij})}{1 + \exp(\alpha_{c-1} + \gamma_i + \beta X_{ij})} \right) \right] \qquad (3)
\end{aligned}
$$

We then integrate out the random effects $\{\gamma_i\}$ with respect to the random effects distribution. This gives rise to the integrated likelihood.

Since the integral does not have closed form, it is necessary to use some approximation for the likelihood function. We can then maximize the approximated likelihood using Gauss-Hermite quadratures [8] [9]. Although there exists a variety of algorithmic approaches for approximating the integral, for our model with a single random effects term, it is straightforward to use the Gauss-Hermite quadratures. To fit the model, we use PROC NLMIXED in SAS. NLMIXED is not naturally designed for multinomial samplings, but one can use it for such models by specifying the form of the likelihood.

For the first dataset (1:1 case, Table 1), the ML estimate of β using REM is 1.80 (OR=6.05) with standard error of 0.550 and for the second data set (1:3 case), the ML estimate of β using REM is 1.087 (OR=2.94) with standard error of 0.476. The result is similar to the MH-type approach for the 1:3 case-control example, but the two approaches yield quite different results for the 1:1 case, potentially because of high sparsity. We also noticed in the process that the estimates obtained by REM for 1:1 matched data lack robustness property and are very sensitive to changing observations even in one matched set. A summary of the results is given on Table 2.

4 Simulation Study

In order to assess relative performance of the existing methods for analyzing matched case-control data with ordered subcategories we conduct two different sets of simulations. One with 1:1 matched setting and the other with 1:3 matched setting.

We consider two disease subcategories, $Y = 1$ and $Y = 2$ and, one control group $Y = 3$. The disease incidence model follows (1) with $c = 3$. For all simulation studies we set $\alpha_2 = 1.2$ and $\beta = 1.0$ ($\alpha_1 = 0$ for identifiability of the models) which are close to the estimates of the parameters obtained by analyzing the real data. For both 1:1 and 1:3 situations, we consider two independent scenarios: i) when there is varying stratum effect $\gamma_i \overset{iid}{\sim} N(-3, 4)$ and ii) when the stratum effect is constant, i.e., $\gamma_i = -3$ for all i.

We simulate 500 datasets under each scenario, and calculate parameter estimates by applying the MH method and the random effect method (REM). The results are presented in Table 3. The final estimates of the parameters are obtained by taking mean across all 500 replications. The mean squared error to measure departure from the true value of $\beta = 1$ is also calculated.

To obtain an ad hoc estimate of the proportion of true rejections of $H_0 : \beta = 0$, among these 500 runs, we first calculate the z-score by dividing the estimate with standard error and then calculate the proportion of times this z-score exceeds the value of 1.96. These proportions serve as an ad hoc estimate of the power of the procedure at the alternative $\beta = 1$.

The results indicate that for 1:1 matched case-control data, REM produces biased results with a large MSE. With large estimates of β, the estimated powers are also high. This is possibly because of presence of so many nuisance parameters (56 in this case) introducing numerical instability in the estimates. The MH method performs reasonably in terms of MSE. For 1:3 matched dataset, the performance of REM improves. The estimated power for MH estimate improves dramatically when compared to 1:1 matched setting.

The CLR estimates lose the information and interpretation of multiple disease categories. Suppose that the cumulative logit model is true. If we ignore the ordinal categories and use CLR, we might lose some power, but the parameter estimate of β should be similar to the other methods. The loss of information from collapsing to a binary response has been discussed by Whitehead [21]. From the simulation results, it seems that there is not much loss on power by ignoring the ordering for the 1:1 cases. But when the sample size within each stratum increases (such as 1:3 cases), the power based on CLR is smaller than the power based on other methods, especially for the situation of the constant stratum effect.

References

[1] AGRESTI, A. (2002). *Categorical Data Analysis*, Second Edition. John Wiley, New York.

[2] AGRESTI, A. and LANG, J. B. (1993). A proportional odds model with subject-specific effects for repeated ordered categorical responses. *Biometrika* **80**, 527–534.

[3] ANDERSEN, E. B. (1980). *Discrete Statistical Methods with Social Science Applications*. New York: North-Holland.

[4] BOCK, R. D. and JONES, L. V. (1968). *The measurement and prediction of judgement and choice*. Holden-Day, San Francisco.

[5] BRESLOW, N. E., DAY, N. E., HALVORSEN, K. T., PRENTICE, R. L., AND SABAI, C. (1978). Estimation of multiple relative risk functions in matched case-control studies. *American Journal of Epidemiology* **108**, 299–307.

[6] FAHRMEIR, L. and Tutz, G. (2001). *Multivariate Statistical Modelling Based on Generalized Linear Models*, Second Edition, Springer.

[7] GODAMBE, V. P. (1976). Conditional likelihood and unconditional optimum estimating equations, *Biometrika*, **63**, 277-284.

[8] HEDEKER, D. and GIBBONS, R. D. (1994). A random-effects ordinal regression model for multilevel analysis. *Biometrics* **50**, 933–944.

[9] HEDEKER, D. and GIBBONS, R. D. (1996). MIXOR: A computer program for mixed-effects ordinal regression analysis. *Computer Methods and Programs in Biomedicine* **49**, 157–176.

[10] HOSMER, D.W. and LEMESHOW S. (2000). *Applied Logistic Regression*, Second Edition. John Wiley, New York.

[11] LIANG, K-Y and Stewart, W. (1987). Polychotomous logistic regression methods for matched case-control studies with multiple case or control groups. *American Journal of Epidemiology*, **125**, 720–730.

[12] LIU, I.-M. and AGRESTI, A. (1996). Mantel–Haenszel–type inference for cumulative odds ratios. *Biometrics* **52**, 1222–1234.

[13] MANTEL, N. and HAENSZEL, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute* **22**, 719–748.

[14] McCULLAGH, P. (1977). A logistic model for paired comparisons with ordered categorical data. *Biometrika*, **64**, 449-453.

[15] McCULLAGH, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society, Series B*, **42**, 109–142.

[16] McNEMAR, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, **12**, 153–157.

[17] NEYMAN, J. and SCOTT, E.L. (1948). Consistent estimates based on partially consistent observations. *Econometrika*,**16**, 1–22.

[18] PIERCE, D. A. and SANDS, B. R. (1975). Extra-Bernoulli variation in regression of binary data, Technical Report 46 (Department of Statistics, Oregon State University, Corvallis, Oregon.

[19] SINHA, S., MUKHERJEE, B. and GHOSH, M. (2004). Bayesian analysis of matched case-control studies with multiple disease states, *Biometrics*, **60**, 41–49.

[20] SNELL, E. J. (1964). A scaling procedure for ordered categorical data. *Biometrics* **20**, 592–607.

[21] WHITEHEAD, J. (1993). Sample size calculations for ordered categorical data. *Statistics in Medicine*, 12:2257-2271.

		weight				weight				weight				weight				weight		
		1	2	3		1	2	3		1	2	3		1	2	3		1	2	3
1:1 case control study																				
1	S	0	1	0	2	0	0	0	3	0	0	0	4	0	1	0	5	0	1	1
	N	0	0	1		0	1	1		1	0	1		0	0	1		0	0	0
6	S	0	1	0	7	0	0	0	8	0	0	0	9	0	0	1	10	0	1	1
	N	0	0	1		0	1	1		0	1	1		0	1	0		0	0	0
11	S	1	0	0	12	0	0	0	13	0	1	1	14	1	0	0	15	0	0	0
	N	0	0	1		0	1	1		0	0	0		0	0	1		0	1	1
16	S	0	1	0	17	0	1	0	18	0	0	1	19	0	1	0	20	0	1	0
	N	0	0	1		0	0	1		0	1	0		0	0	1		0	0	1
21	S	0	0	0	22	1	0	1	23	0	0	0	24	0	0	1	25	0	0	0
	N	0	1	1		0	0	0		1	0	1		1	0	0		0	1	1
26	S	0	1	0	27	0	1	0	28	0	1	0	29	0	0	0	30	0	1	0
	N	0	0	1		0	0	1		0	0	1		1	0	1		0	0	1
31	S	0	0	0	32	0	1	1	33	0	1	0	34	0	0	1	35	0	0	0
	N	0	1	1		0	0	0		0	0	1		1	0	0		1	0	1
36	S	1	0	0	37	0	0	0	38	0	1	0	39	0	0	1	40	0	0	0
	N	0	0	1		0	1	1		0	0	1		1	0	0		1	0	1
41	S	0	0	0	42	1	0	0	43	0	0	0	44	0	0	0	45	0	1	1
	N	1	0	1		0	0	1		0	1	1		0	1	1		0	0	0
46	S	0	0	1	47	0	0	1	48	0	1	0	49	0	0	1	50	1	0	0
	N	0	1	0		0	1	0		0	0	1		0	1	0		0	0	1
51	S	0	1	0	52	0	0	0	53	0	1	1	54	0	1	1	55	1	0	0
	N	0	0	1		1	0	1		0	0	0		0	0	0		0	0	1
56	S	1	0	0																
	N	0	0	1																
1:3 case control study																				
1	S	0	0	1	2	0	1	1	3	0	0	0	4	0	0	2	5	0	1	3
	N	1	0	2		0	0	2		0	1	3		0	1	1		0	0	0
6	S	1	0	2	7	0	0	2	8	0	1	0	9	1	0	1	10	0	1	1
	N	0	0	1		0	1	1		0	0	3		0	0	2		0	0	2
11	S	0	1	1	12	0	0	2	13	0	0	1	14	0	1	2	15	0	1	1
	N	0	0	2		1	0	1		0	1	2		0	0	1		0	0	2
16	S	0	0	0	17	0	1	0	18	0	0	0	19	0	0	1	20	0	0	1
	N	1	0	3		0	0	3		1	0	3		0	1	2		1	0	2
21	S	1	0	0	22	0	0	1	23	0	1	2	24	1	0	1	25	0	1	1
	N	0	0	3		0	1	2		0	0	1		0	0	2		0	0	2
26	S	0	0	1	27	0	1	0	28	0	1	1	29	1	0	1				
	N	1	0	2		0	0	3		0	0	2		0	0	2				

Table 1: Case control studies with ordinal categorical variable. Note: Baby's weight is scaled from "very low" to "normal" (1–3). The "S" is smoker and "N" is non-smoker.

1:1 case control study			1:3 case control study
		β	β
MH	Estimate	0.9694	0.9279
	SE	0.4378	0.4596
	CI	(0.1113,1.8275)	(0.0207,1.8287)
REM	Estimate	1.8001	1.0865
	SE	0.5502	0.4758
	CI	(0.7217,2.8785)	(0.1539,2.0191)
CLR	Estimate	1.0116	0.9616
	SE	0.4129	0.4389
	CI	(0.2023,1.8209)	(0.1014,1.8218)

Table 2: Analysis of low birth-weight studies: 1:1 matched data and 1:3 matched data. The SE stands for standard error and CI denotes the 95% confidence interval

		Constant stratum effect $\gamma_i = -3\ \forall i$	Varying stratum effect $\gamma_i \stackrel{iid}{\sim} N(-3, 4)$
1:1 case control			
MH	Estimate	1.1392	1.1131
	MSE	0.5189	0.4662
	Power	0.34	0.302
REM	Estimate	2.069	2.004
	MSE	2.7668	2.4604
	Power	0.63	0.618
CLR	Estimate	0.9728	1.1349
	MSE	0.2697	0.2886
	Power	0.60	0.69
1:3 case control			
MH	Estimate	1.0783	1.0940
	MSE	0.2895	0.2792
	Power	0.60	0.592
REM	Estimate	1.2046	1.2032
	MSE	0.3697	0.3475
	Power	0.656	0.68
CLR	Estimate	0.9694	1.0366
	MSE	0.2429	0.3128
	Power	0.52	0.60

Table 3: Results of the simulation study for 1:1 and 1:3 matched case-control studies with $\beta = 1$. Here Estimate denotes the average value of the estimated parameter of interest over 500 replicates and MSE denotes the average squared deviations of the estimates from the true value of 1. Ad hoc power estimates by reporting the proportion of times the null hypothesis of $H_0 : \beta = 0$ is rejected in the 500 runs. We calculate whether the approximate z-statistic obtained by divining the estimate of β by the estimated standard error exceeds 1.96.

Structural Change and New Zealand Visitor Arrivals: the Effects of 9/11

John Haywood[1] and John Randal[2]
[1]School of Mathematics, Statistics and Computer Science
[2]School of Economics and Finance
Victoria University of Wellington, New Zealand
John.Haywood@mcs.vuw.ac.nz

Abstract

We analyse monthly short-term visitor arrival time series for New Zealand, to assess the effect of the 11 September 2001 terrorist attacks on the number of visitors to New Zealand. Somewhat misleading reports from the media concerning these data are highlighted. In particular, we demonstrate that while some historical events have had a marked structural effect on trends in visitor arrivals to New Zealand, 9/11 was not one of these. Our conclusion is drawn on the basis of an initial nonparametric analysis, followed by a new iterative approach to fitting parametric structural break models, motivated by iterative methods of seasonal adjustment.

Key words: Break dates; Endogenous structural changes; Iterative fitting; Multiple breaks; Nonparametric estimation; Seasonality; Terrorism effects; Trend extraction.

1 Introduction

There seems little doubt that the terrorist attacks of 11 September 2001 have had a pronounced influence on world events since that time. For example, see US Department of State (2004), for a summary of 100 editorial opinions from media in 57 countries around the world, commenting on the three years following September 2001. Those terrorist events and their subsequent effects have been used to explain apparent movements in many time series, and in this paper we focus on a particular example: the number of short term visitor arrivals to New Zealand.

As we shall illustrate, in fact there is actually little to suggest that the September 11 incidents had much effect on New Zealand visitor arrivals, when viewed in the context of 'normal' historically observed movements. In contrast, we identify some other historical events which do appear to have affected visitor arrivals to New Zealand quite markedly. We also illustrate features of the arrivals data which highlight the importance of graphical displays when examining time series.

In Section 2 we highlight some reports from the media concerning these arrivals data, which motivated the need to look at claims of movements in a context of either structural or irregular change, allowing for observed stochastic variation. Section 3 explores and discusses some of the apparent sources of variability for New Zealand visitor arrivals. Section 4 presents a parametric model that allows separate structural changes in the trend and seasonal components. We use a new iterative fitting procedure, motivated by the classical Macaulay cycle method for the decomposition of seasonal data. Finally in Section 5 we give some concluding comments.

2 Motivation: the reported effects of 9/11 on short-term visitor arrivals to New Zealand

As an illustration of the perceived effects of the 9/11 incidents on New Zealand short-term visitor arrivals, we quote some short passages from The Dominion Post, one of New Zealand's widely read daily newspapers. All the quotes appeared around the first anniversary of 9/11. Much of the reported analysis was conducted by specialists within the tourism industry or the government. Article headlines are in bold, and our first example appeared as the lead story on the front page:

> **Sept 11 costs Kiwis $1 billion**
> The main costs incurred *as a direct result of the attacks* include: $64 million in lost income from tourists.
> (The Dominion Post, 7 September 2002, our emphasis)

A second example concerns a supposedly-causal link from the effects of the September 11 events to changes in arrivals numbers:

> **September 11 hits tourist numbers**
> Uncertainty about travelling on September 11 contributed to a 3 per cent drop in visitor arrivals *in August compared with the same month last year.* Visitors from New Zealand's biggest market, Australia, dropped 12 per cent while Japanese tourist numbers fell 6 per cent. Tourism Holdings general manager Shaun Murray said, "[w]e are *very sure* that speculation that people were wary of travelling around the time of 11 September was in fact quite correct." Tourism New Zealand chief executive George Hickton agreed.
> (The Dominion Post, 21 September 2002, our emphasis)

As just noted, comparisons are often made with the 'equivalent' period from the preceding year. Such comparisons may not always correctly reflect observed features of the data. For example, the previous year may have been 'high', rather than 'normal'; should a decrease then be viewed pessimistically? The 9/11 incidents have also been used to account for forecast errors, although such errors occur every year:

> Tourism Ministry figures show there were 26,000, or 1.4 per cent, fewer tourists in total last year *than forecast*, and much of the shortfall *can be attributed to September 11.*
> (The Dominion Post, 7 September 2002, our emphasis)

In fact, the number of visitor arrivals from 1 September 2001 to 31 August 2002 was 1,959,886, an increase of 42,102 (or 2.2%) on the previous year. The wish to explain precisely why a forecast of a stochastic series was wrong is a failure to acknowledge naturally occurring variability: any forecast will be wrong, but a sensible quantification of the likely size of the error is a more realistic goal.

Explanations for variations in seasonality or trend that attribute movements to any one cause, such as the 9/11 incidents noted above, are almost always too simplistic. For monthly data, the 'equivalent' period from the preceding year often means 'same month'. Yet such a simple comparison ignores known calendar effects like Easter – e.g., Zhang, McLaren and Leung (2001) discuss an approach for an Australian Easter effect. Another more obscure 'calendar effect' in these data is a trans-Tasman rugby effect, identified in Haywood and Randal (2004) and due to changes in the month in which New Zealand's All Blacks hosted the Australian Wallabies for a rugby test match. The All Blacks' home game was in August 2001, but moved to July in 2002. The net result is to observe an increase from 2001 to 2002 in visitors from Australia in July but a corresponding decrease in August. Attribution of the observed August 2001-2 decrease in visitors from Australia to a 9/11 effect, as made above, is clearly not the whole story. Neither is a trans-Tasman rugby effect, but there is more than one reason for the observed movements.

Our purpose here is not to provide reasons for all observed changes in arrival numbers, from all source countries. Rather, we hope to illustrate that appreciation of stochastic variation is necessary in order to view apparent movements in an appropriate context; that is structural (longer-term) or not.

3 EDA of short-term visitor arrivals to New Zealand

The economic importance of tourism to New Zealand has recently increased considerably. As Pearce (2001) notes in his review article, international visitor arrivals increased by 65% over the period 1990 to 1999 while foreign exchange earnings increased by 120% (in current terms). More recently, for the year ended March 2002 tourism expenditure was $14.6 billion (Statistics New Zealand 2003). In that year, the tourism industry made a value added contribution to GDP of 9%, split evenly between direct and indirect contributions, and 89,711 people (FTE) had work that was directly engaged in tourism: 5.5% of the total employed workforce. Also in that year, tourism's 14.3% contribution to exports was slightly less than that of dairy products (16.9%) but greater than the contribution of meat products (10.3%), which in turn was greater than the contributions from wood and wood products, or seafood. So tourism is now one of the most important industries in New Zealand.

The short-term New Zealand visitor arrival series is one direct and easily recorded measurement of the international tourist contribution to the New Zealand economy; see Statistics New Zealand (2003) for the most up to date comprehensive breakdown of recent expenditure by domestic and international tourists. On a monthly basis, Statistics New Zealand releases official monthly totals, and these are commonly reported in the media, often with comparisons to the same month one year ago.

We analyse monthly data from January 1980 to December 2004, looking at the total short-term arrivals and also those from the most important countries of origin, ranked by current proportion of the total: Australia, UK, USA, Japan, Korea, China, Germany (see Figure 1). A "U"-shaped seasonal pattern is common, with visitor numbers reaching a local maximum in the summer months December to February, and a local minimum in the winter months June and July.

Several prominent features are evident in Figure 1. Australian and UK arrivals appear relatively stable, with the USA data less so. A large downturn in arrivals from the USA is evident in the late 1980s, a period which immediately followed the stock market crash of October 1987. Japanese arrivals level off in average numbers over the last 15 years. Arrivals from Germany show a clear change from exponential growth prior to the early 1990s to a more stable pattern in recent times. Arrivals from China contain perhaps the most visible short term effect in these series, which is due to the SARS epidemic that virtually eliminated international travel by Chinese nationals during May and June 2003. Some other series, including Other and Total arrivals, show a SARS effect similar to but less prominent than that seen in the Chinese arrivals. The effect of the Asian financial crisis of 1997 is evident especially in the Korean data, with visitor numbers dramatically reduced after this event. One of the more obvious shifts in the aggregate 'Total' series appears to be linked to the Korean downturn, and can be attributed to the Asian financial crisis.

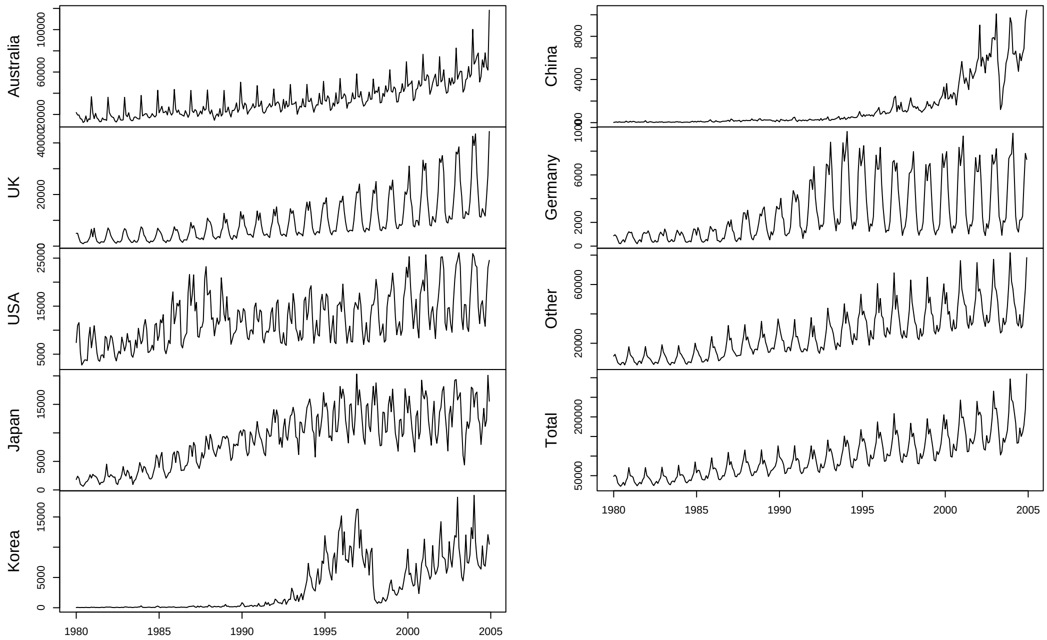


Figure 1. Monthly short term visitor arrivals to New Zealand, by origin, from January 1980 to December 2004. The vertical scales are not equal.

Table 1 shows that Australia is by far the single biggest source of visitors to New Zealand, accounting for almost exactly one-third of visitors in the 2000-2004 five year period. Visitors to New Zealand from Australia made up 33.8% of the total visitors to New Zealand over the data period. The maximum proportion in a month was 58.8% in June 1985, and the minimum was 21.8% in February 1997.

Table 1. Summary statistics for the monthly proportion of visitors to New Zealand, by origin. The final three columns give proportions of the Total for the entire 25 year sample period, and the five-year periods 1980-1984 and 2000-2004, respectively.

	Min	LQ	Median	UQ	Max	80-04	80-84	00-04
Australia	21.8	30.0	35.9	41.6	58.8	33.8	44.9	33.3
UK	3.4	6.3	8.0	10.6	18.3	9.8	7.6	11.8
USA	6.3	10.3	13.0	16.3	29.4	12.4	16.7	10.0
Japan	2.8	7.1	9.1	11.0	17.8	9.2	5.9	7.8
Korea	0.0	0.2	1.0	4.3	10.5	3.4	0.2	4.8
China	0.0	0.2	0.4	1.2	4.7	1.4	0.1	3.1
Germany	0.8	1.5	2.2	3.4	7.5	2.9	1.8	2.6
Other	17.9	23.4	26.1	28.6	34.3	27.0	22.8	26.7

An Australian influence is notable in the total arrivals, because as the nearest neighbour to an already geographically isolated country, the Australian arrivals exhibit variation not seen in the remaining data. As seen in Figure 1, the Australian data has a regular seasonal pattern which is quite different from that of any other country. Closer analysis indicates four peaks per year aligned with the Australian school holidays, one of which typically encompasses the Easter festival. An Easter effect is apparent in both the 1996–1998 and 2001–2003 periods, during which Easter alternated between April and March. In 1997 and 2002 when Easter fell in March, March arrivals were high and April arrivals low. Year to year comparisons for the Australian series are affected by Easter movements. Since Australia is such a large contributor to total arrivals, this also feeds into year-to-year comparisons of totals.

The arrivals data can be decomposed into unobserved components: a trend, representing the general level of the series; a seasonal, encapsulating regular deviation from the trend on a within-year basis; and an irregular, which is the residual, or unexplained variation in the data. One way of estimating these components is to use a robust, nonparametric technique such as STL (Cleveland, Cleveland, McRae and Terpenning 1990). This procedure consists of an iterated cycle in which the data is detrended, then the seasonal is updated from the resulting detrended seasonal-subseries, after which the trend estimate is updated. At each iteration, robustness weights are formed based on the estimated irregular component; these are used to down-weight outlying observations in subsequent calculations.

A typical STL decomposition is shown in Figure 2 for the natural logarithm of the Total arrivals. The plot shows an evolving seasonal pattern, an upward trend with several changes in slope, and a relatively small irregular component. A vertical line indicates September 2001; there is no obvious (structural) change in the trend at or about this month. More prominent is a cluster of negative irregulars immediately following 9/11, the largest of which is the third largest negative irregular in the sample period. Jointly though, these irregulars are smaller and less persistent than those occurring at the time of the SARS outbreak in 2003. Our exploratory analysis with STL thus suggests that while the events of 9/11 may have had a moderate short-term (irregular) effect, there is nothing to suggest that a longer-term (structural) effect occurred. We investigate this hypothesis more formally in Section 4.

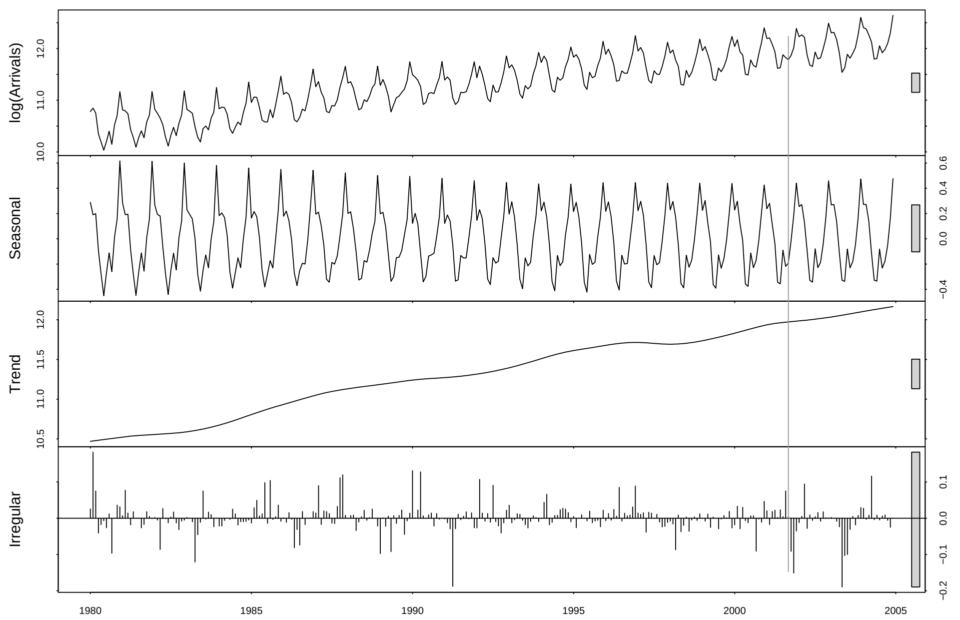


Figure 2. The STL decomposition of the log aggregate monthly visitor arrivals to New Zealand from January 1980 to December 2004. The vertical grey line is at September 2001, and the solid bars on the right hand side of the plot are all the same height, to aid comparisons.

A relevant confounding effect is the collapse of Ansett Australia, which occurred just three days after the terrorist attacks of 9/11; hence it is impossible to separate these two effects with monthly data. The termination of flights on 14 September 2001 certainly affected capacity and timing of arrivals to New Zealand. In addition, in the following week, strike action targetted at Air New Zealand occurred at Melbourne and Perth airports (Air New Zealand had acquired control of Ansett Australia during the year preceding its collapse). Those strikes required the cancellation of all Air New Zealand trans-Tasman flights operating from Melbourne and Perth. So considerable limitations were imposed on arrivals from Australia during September 2001, for reasons other than the 9/11 events. This would have noticeably influenced Total arrivals, due to the large number of visitors from Australia.

4 Modelling the arrivals, allowing for structural breaks

The seasonal variation of the arrivals series typically increases with the level (Figure 1). This suggests a log transformation, yet that is not universally appropriate here. Instead we estimate a power transformation, identified using the robust spread-vs-level plots described in Hoaglin, Mosteller and Tukey (1983). For each individual series we calculate the median and interquartile range (IQR) of the monthly arrivals for each of the 25 calendar years, then regress log IQR on log median. The appropriate transformation is $x^{1-\text{slope}}$, and the transformed series are shown in Figure 3, with the estimated powers. Confidence intervals for the slopes in these regressions support the use of logs only in the case of the UK, USA and Total arrivals (i.e. a slope of one). In the case of Germany the estimated power is negative, so $-x^{1-\text{slope}}$ is used to preserve order. All further analysis is conducted on the transformed data.

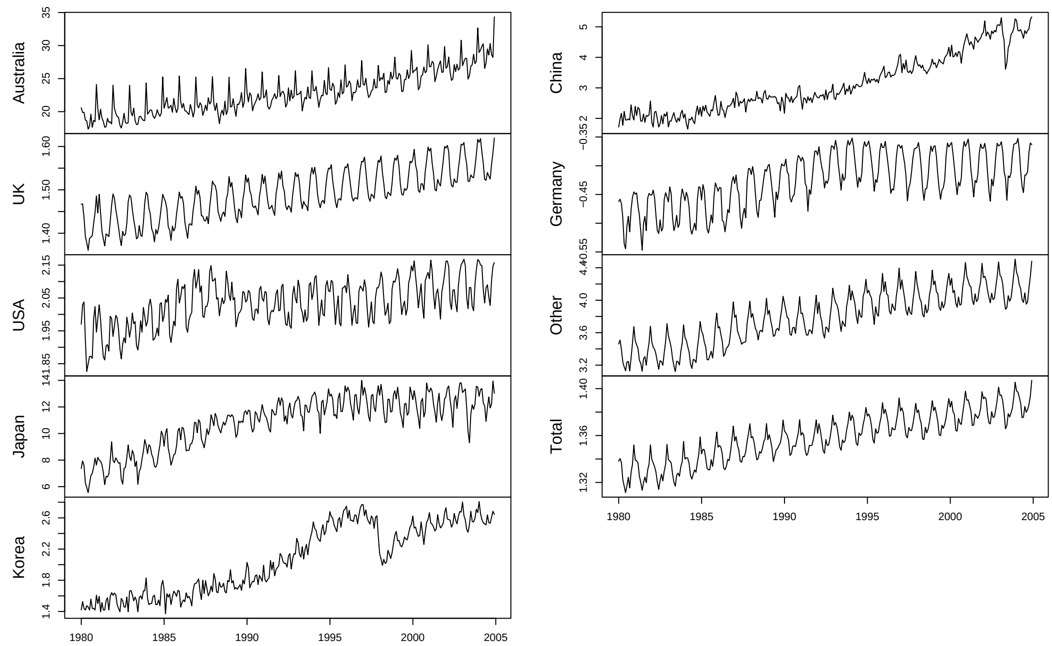


Figure 3. Power transformed monthly short term visitor arrivals to New Zealand, by origin, from January 1980 to December 2004. The power transformations are: Australia 0.3, UK 0.05, USA 0.08, Japan 0.27, Korea 0.11, China 0.18, Germany -0.11, Other 0.13, and Total 0.03.

We now consider modelling the transformed visitor arrivals series using a piecewise linear trend, with a superimposed piecewise constant seasonal pattern. Bai and Perron (1998, 2003) present a methodology for fitting a regression model with structural breaks, in which the break points, i.e. the time at which the parameters change, are determined optimally. Given a sample of T observations, the selected break points are estimated consistently, with rate T convergence of the estimates.

The maximum number of break points is determined by the number of observations relative to the number of parameters in the model. Figure 3 indicates that for most series a linear time trend would need breaks. While the seasonal patterns generally have constant variation over the length of the series (due to the power transformations) we do not wish to preclude seasonal changes over the data period. The parameter-rich trend-plus-seasonal model would severely limit the ability to estimate a good fitted model, since the large number of seasonal dummies (and the stable seasonal patterns) would reduce the possible number of break points.

To address this concern we estimate the trend and seasonal components separately, using a new iterative approach motivated by the Macaulay cycle seasonal decomposition method and the iterative method of STL. This allows more flexible structural break estimation than directly fitting a complete model. We assume that data can be decomposed into a piecewise linear time-trend and a piecewise constant seasonal pattern. Each component is then estimated using the methodology of Bai and Perron (1998, 2003), implemented in R using the software of Zeileis *et al.* (2003). We estimate the trend of the transformed visitor arrivals data using a piecewise linear regression model

$$Y_t = \alpha_j + \beta_j t + \epsilon_t \qquad t = T_{j-1}+1, \dots, T_j$$

for $j = 1, \dots, m+1$, with Y_t the deseasonalised, transformed monthly arrival series and T_j, $j = 1, \dots, m$ the unknown trend break points. We use the convention that $T_0 = 0$ and $T_{m+1} = T$ (Bai and Perron 1998). Once the trend has been estimated, we estimate the seasonal component as a piecewise constant function, given by

$$W_t = \delta_0 + \sum_{j=1}^{11} \delta_j D_{j,t} + \nu_t \qquad t = T'_{j-1}+1, \dots, T'_j$$

for $j = 1, \dots, m'+1$, with W_t the detrended, transformed monthly arrival series, $D_{j,t}$ seasonal dummies and T'_j, $j = 1, \dots, m'$ the unknown seasonal break points. We take $T'_0 = 0$ and $T'_{m'+1} = T$. This estimation process is then iterated to convergence. We are thus able to estimate a trend which, due to its parsimonious representation, is able to react to obvious shifts in the general movement of the data. If required, we are able to identify important changes in the seasonal pattern separately.

The estimated parametric trends and break points (with 95% confidence intervals) are shown in Figure 4, along with nonparametric trends estimated by STL. The intervals have been formed with heteroscedasticity and autocorrelation consistent estimates of the covariance matrix (Andrews 1991). Note the intervals are not symmetric about the breaks, indicating that movement of the break date within the interval is often possible in only one direction. Note too the intervals for the first two estimated breaks in the Total series overlap, indicating possible redundancy. September 2001 is included in four confidence intervals, indicating the possibility that the terrorist events of 9/11 may be linked to a structural break in the trend of arrivals for those four origins: Australia, UK, USA and Other. 'Other' is difficult to interpret of course, given its composite nature. For UK and USA, not only are the intervals very wide on the side that contains 9/11, but also the break points have an increase in both intercept and slope of the trend. Hence if those break points were associated with 9/11, the terrorist events had a positive effect on tourist numbers to New Zealand from UK and USA.

In the case of Australia, a break is estimated close to 9/11. This results in an increased trend slope but a decreased intercept. We cannot attribute this fall solely to the 9/11 terrorist events, due to the 'Ansett effects' already noted in Section 3.

The limit on passenger numbers due to those other effects is one explanation for a decrease in intercept, while the increase in slope mirrors the increases in arrival numbers from both UK and USA.

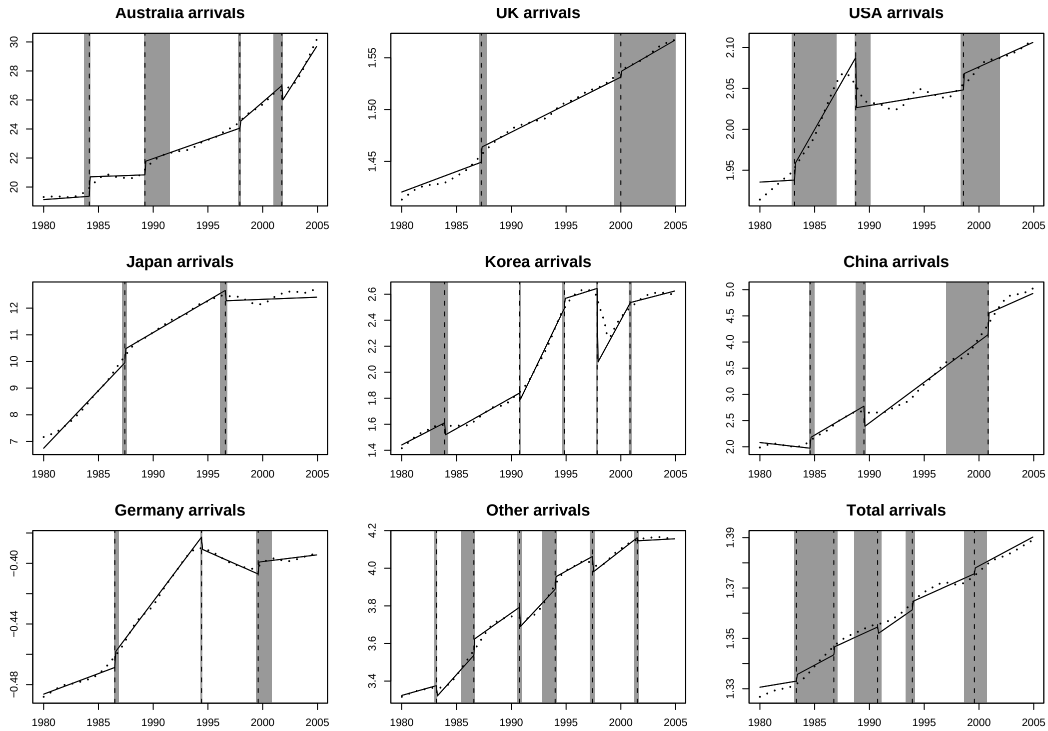


Figure 4. Estimated trends and trend break points for the transformed monthly visitor arrivals to New Zealand, by origin, from January 1980 to December 2004. The solid line is the piecewise linear time trend, while the dotted line is the estimated STL trend. The vertical dashed lines and grey regions respectively indicate the fitted break points and their 95% confidence intervals, estimated using a HAC estimate of the covariance matrix.

Table 2 gives the estimated seasonal break points for the transformed arrivals. As the power transformation has effectively stabilised the seasonal variation, any changes in the seasonal pattern more likely reflect behavioural changes in the time of year when visitors arrive. For example, in Australia's seasonal pattern the peaks have moved, reflecting a shift from a three-term school year to a four-term year in New South Wales in 1987 (NSW Department of Education 1985). The placement of the seasonal break point coincides exactly with the final month under the old three term system, with the first holiday in the new sequence occuring in July 1987.

Table 2. Estimated seasonal break points for the transformed monthly visitor arrivals to New Zealand, by origin, from January 1980 to December 2004. The middle column gives the estimated break points, while the first and third columns give the lower and upper 95% confidence limits respectively, estimated using a HAC estimate of the covariance matrix. Korea, China, Germany and Other have no estimated seasonal break points.

Australia	1987(1)	1987(6)	1987(10)	Japan	1987(10)	1988(6)	1988(12)
UK	1986(4)	1986(12)	1988(5)	Total	1984(11)	1985(2)	1985(4)
USA	1995(1)	1995(4)	1995(11)		1996(12)	1997(5)	1997(9)

Finally, we compare the trend estimates obtained from our new iterated approach to the trend obtained fitting a complete 13-parameter model. In Figure 5 we include trends for the Korean arrivals and those from 'Other' origins. We also show sample autocorrelation functions for the residuals of the complete model, our new iterated approach, and the nonparametric STL trend. The trends are all similar, but more so for the iterated approach and STL. Some differences are clear though, especially at the end of the series. Based on the three approaches forecasts would differ markedly, but the best agreement would be between those from the iterated approach and STL. Further, the iterated technique estimates the well-motivated break point in the Korean data precisely. The irregular components also favour iteration, as the residuals for the other approaches are highly autocorrelated, especially at low lags. In contrast, the residuals of the iterated method exhibit far less autocorrelation, showing a better overall decomposition.

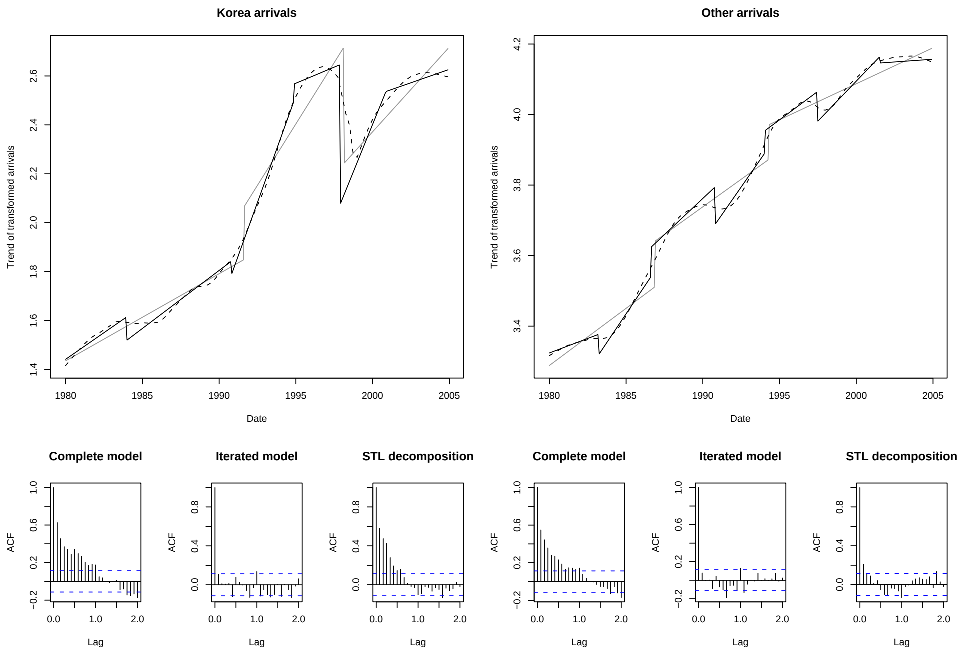


Figure 5. Trend estimates for the Korean arrivals and those from 'Other' origins. The trend estimates are based on the complete model (grey), the new iterated approach (black) and STL (dashed). Also shown are sample autocorrelation functions for the residuals from the three methods.

5 Discussion

The growth in the number of visitor arrivals to New Zealand was lower than expected in late 2001, yet there is no conclusive evidence to attribute this forecast error solely to the terrorist events of 9/11. The termination of flights by Ansett Australia on 14 September 2001 certainly affected capacity and timing of arrivals from Australia to New Zealand, which would have affected Total arrivals noticeably as well. A further plausible cause for the lower than forecast number of visitors is the US recession dated March 2001 (Hall *et al.* 2001), along with the world wide flow-on effects from a slow down in the US economy. The recession predates 9/11 by six months but that is consistent with observed features of the arrivals data. In particular, March 2001 corresponds exactly to the minimum in the second derivative of an STL trend, indicating a maximum decrease in the slope at that time.

While it seems clear that the events of 9/11 did not make much difference to the numbers of short term visitors to New Zealand, our analysis identifies other events which have had structural effects on the trend in these data. In particular, the stock market crash of October 1987 preceded a dramatic decline in arrivals from the USA, followed by a sustained period of only moderate growth. Similarly, the Asian financial crisis of 1997/8 precipitated a massive drop in Korean arrivals. The SARS epidemic affected arrivals from China in a different way, with a very short-lived but large reduction, which we class as temporary and not structural.

Estimation of structural breaks was facilitated by a new implementation of Bai and Perron's (1998, 2003) work. Use of an iterative approach to estimate the trend and seasonal components separately enabled us to locate structural breaks in the data, and to attribute these to either changes in the trend or the seasonal pattern. Estimating these components simultaneously did not achieve the same flexibility in the estimated components, nor in the location of the break points. The agreement between the estimated parametric trends from the iterated approach and the nonparametric STL trends is especially pleasing, as is the lack of residual structure around those parametric trends when compared to other trend estimates.

References

Andrews, D.W.K. 1991. "Heteroskedasticity and autocorrelation consistent covariance matrix estimation." *Econometrica* 59(3): 817–858.

Bai, J., and P. Perron. 1998. "Estimating and testing linear models with multiple structural changes." *Econometrica* 66(1): 47–78.

———. 2003. "Computation and analysis of multiple structural change models." *Journal of Applied Econometrics* 18: 1–22.

Cleveland, R.B., W.S. Cleveland, J.E. McRae, and I. Terpenning. 1990. "STL: A Seasonal-Trend Decomposition Procedure Based on Loess." *Journal of Official Statistics* 6: 3–73.

Hall, R., M. Feldstein, B. Bernanke, J. Frankel, R. Gordon, and V. Zarnowitz. 2001. "The Business-Cycle Peak of March 2001." Technical Report, Business Cycle Dating Committee, National Bureau of Economic Research, USA. Available at `http://www.nber.org/cycles/november2001/`

Haywood, J., and J. Randal. 2004. "Stochastic variation, New Zealand visitor arrivals, and the effects of 11 September 2001." Technical Report 04-01, School of Mathematical and Computing Sciences, Victoria University of Wellington.

Hoaglin, D.C., F. Mosteller, and J.W. Tukey, eds. 1983. *Understanding Robust and Exploratory Data Analysis.* New York: Wiley.

NSW Department of Education. 1985. "1987: Four term future." *Perspectives: looking at education* 8(4) (April).

Pearce, D. 2001. "Tourism." *Asia Pacific Viewpoint* 42: 75–84.

Statistics New Zealand. 2003. *Provisional Tourism Satellite Account 2000-2002.* Wellington: Statistics New Zealand.

US Department of State. 2004. "Three Years after 9/11: Mixed Reviews for War on Terror." Available at http://www.globalsecurity.org/security/library/news/2004/09/wwwh40915.htm

Zeileis, A., C. Kleiber, W. Krämer, and K. Hornik. 2003. "Testing and dating of structural changes in practice." *Computational Statistics and Data Analysis* 44: 109–123.

Zhang, X., C.H. McLaren, and C.C.S. Leung. 2001. "An Easter Proximity Effect: Modelling and Adjustment." *Australian & New Zealand Journal of Statistics* 43: 269–280.

Optimal Warranty Repair Strategy in $\mathbb{R}^2$

S. Chukova and M. R. Johnston
School of Mathematics, Statistics and Computer Science
Victoria University of Wellington
New Zealand
Stefanka.Chukova@mcs.vuw.ac.nz, Mark.Johnston@mcs.vuw.ac.nz

Abstract

For repairable products, the warrantor has options in choosing the degree of repair applied to an item failed within the warranty period. We focus on a particular warranty repair strategy, related to the degree of the warranty repair, for non-renewing, two-dimensional, free of charge to the consumer warranty policy. This study is an extension of (Iskandar, B.P., Murthy, D.N.P., and Jack, N. 2005). A comparison between our strategy, the (Iskandar, B.P., Murthy, D.N.P., and Jack, N. 2005) strategy, and the standard strategies, all warranty repairs are minimal repairs or all warranty repairs are complete repairs, is provided.

1 Introduction

Study of warranty systems are of interest to manufacturers for several reasons. Paid claims are a cost of doing business and a liability incurred by the manufacturer at the time of sale. For these reasons forecasting warranty expenses is of interest. Also, warranty data provide information about the durability of products in the field and, therefore, is of interest to engineers. Furthermore, warranty coverage can be regarded as a product attribute that affects buying decisions. Manufacturers may wish to change warranty coverage to attract more buyers, and they would want to estimate how much the changes would cost. See (Robinson, J. and McDonald, G. 1991) for further discussion on these points. Here we emphasise two-dimensional warranties, like automotive warranties. The warranty guarantees free repairs subject to both age and mileage limits. The most common limit is now 36 months or 36,000 miles, whichever comes first. Age is known all the time for all sold vehicles because sales records are retained. But odometer readings are only collected in the dealership at the time of a claim. So automotive warranties are a case where two usage measures are of interest and information on one of them is incomplete. Thus, in automotive industry the warranties are two-dimensional and guarantee free of charge to the consumer repair subject to time and usage restrictions.

Overall, a product warranty is an agreement offered by a producer to a consumer to repair or replace a faulty item, or to partially or fully reimburse the consumer in the event of a failure. The form of reimbursement to the customer on failure of the product or dissatisfaction with service, is one of the most important characteristics of warranty. The most common forms are a lump-sum rebate, a free repair, a repair provided at reduced cost to the buyer or a combination of these forms (Blischke, W.R. and Murthy, D.N.P. 1993; Blischke, W.R. and Murthy, D.N.P. 1996).

Regarding the mechanism of the warranty coverage, there are two types of warranty policies used in the marketplace and studied in the literature, namely, non-renewing warranty (NR) that covers a newly sold item for some calendar time of fixed duration and renewing warranty (R), which starts anew at every warranty repair. Most domestic appliances, such as vacuum cleaners, refrigerators, washing machines and dryers, televisions, are covered by non-renewing free repair warranty (NRFRW). For example, the light bulbs are covered by renewing free repair warranty (RFRW). A light bulb has an initial warranty period of thirty days and if it fails during this period it is replaced by a new bulb and the warranty starts anew. The replacement can be considered as a particular type of repair, namely, complete repair. Usually renewing warranty is assigned to inexpensive products.

The evaluation of the parameters (e.g., warranty period or price) of the warranty contract can be obtained, by using appropriate models, from the point of view of the producer, seller or buyer. Often these parameters are solutions of an appropriate optimisation problem and their values result from the application of analytical or numerical methods. Due to the complexity of the models, it is almost always necessary to resort to numerical methods, since analytical solutions exist only in the simplest situations. A general treatment of warranty analysis is given in (Blischke, W.R. and Murthy, D.N.P. 1993), (Blischke, W.R. and Murthy, D.N.P. 1996) and (Chukova, S., Dimitrov, B., and Rykov, V. 1993). For a recent literature review see (Murthy, D.N.P. and Djamaludin, I. 2002).

The evaluation of the warranty cost, or any other parameter of interest, in modeling warranties depends on the failure and repair processes and on the assigned preventive warranty maintenance for the items. The repairs can be classified according to the degree to which they restore the ability of the item to function (Pham, H. and Wang, H. 1996). In this study we focus on minimal repairs, i.e., repairs that have no impact on the performance characteristics of the item, and on complete repairs, which are equivalent to a replacement of the faulty item by a new one, identical to the original. A possible approach to model imperfect repairs is studied in (Chukova, S., Arnold, R., and Wang, D. 2004) and it is extended from Bayesian prospective in (Chukova, S. and Hayakawa, Y. 2005).

The outline of this papers is as follows. In section 2, the model formulation is given, along with the description of the process of failures and the warranty repair strategy. Section 3 provides an analysis of the proposed model. A numerical example illustrates the ideas in section 4 and section 5 offers some conclusions.

2 Model formulation

Consider a product sold under two-dimensional warranty, i.e., warranty which is described in terms of two measures. Usually the two-dimensional warranty is associated with a region $\Omega \subset \mathbb{R}^2$, where the warranty measures are well defined. For more on the different shapes of Ω and related warranty policies see (Blischke, W.R. and Murthy, D.N.P. 1996). For convenience, the warranty measures of interest will be called age and usage. Moreover, we focus on NRFRW with $\Omega = K \times L$, i.e., the warranty expires either when the product's age, denoted T, exceeds K, or the total usage of the product, denoted X, exceeds L which ever occurs first. For example, the typical automotive warranties are limited by age as well as mileage and $\Omega = 3$ years $\times$ $36,000$ miles.

2.1 Modeling failures

We will assume that the repair time is negligible with respect to the operating time of the product, i.e., the repairs are instantaneous.

Denote by $T(t)$ the virtual age of the product at calendar time t, and denote by $X(t)$ the virtual usage of the product at calendar time t. Assuming that the warranty repairs are minimal or complete will imply that

$$T(t) \leq t, \qquad \text{for any } t \in [0, K].$$

Also, we assume that at the beginning of their lifetime, products (and in particularly cars) accumulate usage approximately linearly with their age, i.e.,

$$X(t) = R\, T(t), \tag{1}$$

where R is the usage accumulation rate. For a particular customer, R is assumed to be a constant over Ω. The usage accumulation rate is assumed to be a positive random variable with known cumulative distribution function $G(r) = P(R \leq r)$ and corresponding probability density function $g(r)$. In view of equation (1), we model the failure/repair process of the product as a point process with an intensity function $\lambda(t|r)$,

$$\lambda(t|r) = P(\text{product operational up to time } t \text{ fails in } [t, t + \delta t) \mid R = r).$$

In view of equation (1), we model the failure/repair process of the product as a point process with an intensity function $\lambda(t|r)$, i.e., $\lambda(t|r)\, \delta t = P($a product operating up to time t fails in [
$\delta t) \mid R = r)$. Related to $\lambda(t|r)$ cumulative failure rate function is

$$\Lambda(t|r) = \int_0^t \lambda(x|r)\, dx.$$

Knowing $\lambda(t|r)$, the cdf of the time to first failure of the product, say $T_{1|r}$, is the same as the lifetime of the product,

$$F_{T_{1|r}}(t) = 1 - \exp\left(-\int_0^t \lambda(x|r)\, dx\right).$$

If all warranty repairs are complete repairs, conditional on $R = r$, the corresponding failure/repair process is a renewal process generated by $F_{T_{1|r}}(\cdot)$, whereas under minimal warranty repairs, again conditional on $R = r$, the corresponding process is a non-homogeneous Poisson process.

In order to identify the structure of the model, the intensity function $\lambda(t|r)$ has to be selected. Overall, $\lambda(t|r)$ has to be a positive increasing function of r, $T(t)$ and $X(t)$. If appropriate data is available, an approach for obtaining an empirical estimate of $\lambda(t|r)$ is given in (Rigdon, S.E. and Basu, A.P. 2000). For the numerical example in this study, following (Iskandar, B.P., Murthy, D.N.P., and Jack, N. 2005), our choice for $\lambda(t|r)$ is

$$\lambda(t|r) = \theta_0 + \theta_1 r + \theta_2\, T(t)^2 + \theta_3\, X(t)\, T(t). \tag{2}$$

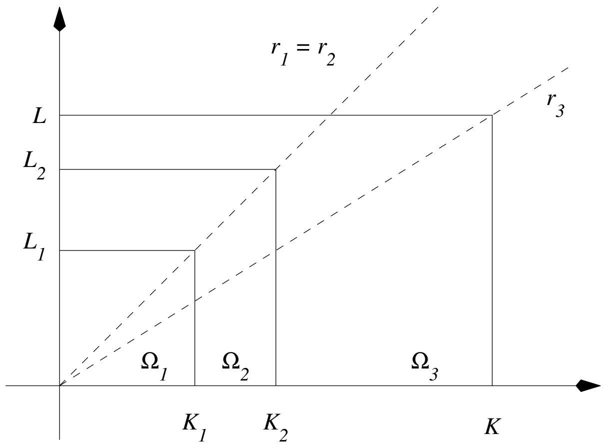


Figure 1: Restricted strategy ($r_1 = r_2$)

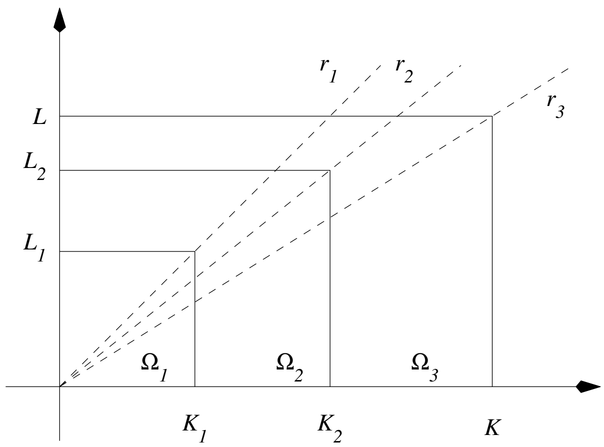


Figure 2: Unrestricted strategy

2.2 Repair strategy

In (Iskandar, B.P. and Murthy, D.N.P. 1999), the given warranty region Ω is divided into two subregions Ω_1 and Ω_2, such that $\Omega_1 \cup \Omega_2 = \Omega$ and $\Omega_1 \cap \Omega_2 = \emptyset$. Two warranty-repair strategies are considered. Under the first strategy, in Ω_1, which covers the early stage of the warranty, the warranty is NRFRW with complete repairs followed, in Ω_2, by NRFRW with minimal repair. Under the second strategy, the types of warranty repairs within Ω_1 and Ω_2 are interchanged.

In (Iskandar, B.P., Murthy, D.N.P., and Jack, N. 2005), an extension of (Iskandar, B.P. and Murthy, D.N.P. 1999) is considered. The warranty region is divided into three disjoint subregions Ω_1, Ω_2, and Ω_3, such that $\Omega_1 \cup \Omega_2 \cup \Omega_3 = \Omega$ (see Figure 1). They adopted the following repair strategy (S):

1. Any repair in Ω_1 is minimal, costing c_2.
2. The first repair in Ω_2 is complete, costing c_1, and any further repair in Ω_2 is minimal.
3. Any repair in Ω_3 is minimal.

In (Iskandar, B.P., Murthy, D.N.P., and Jack, N. 2005), the analysis of strategy (S) is restricted (as in Figure 1) to the case

$$r_1 = \frac{L_1}{K_1} = \frac{L_2}{K_2} = r_2. \tag{3}$$

They computed the optimal expected warranty servicing cost $EC(\hat{\phi}^*)$ under the restricted strategy (S).

We extend their study to unrestricted strategy (S) by relaxing (3), i.e., allowing $\frac{L_1}{K_1} \neq \frac{L_2}{K_2}$ as well as $\frac{L_1}{K_1} = \frac{L_2}{K_2}$ (as in Figure 2). We identify the optimal repair strategy as the one which minimizes the expected warranty servicing cost per item sold. Thus our main goal is to determine the subregions Ω_1, Ω_2 and Ω_3 so that the corresponding expected warranty servicing cost per item sold is minimum.

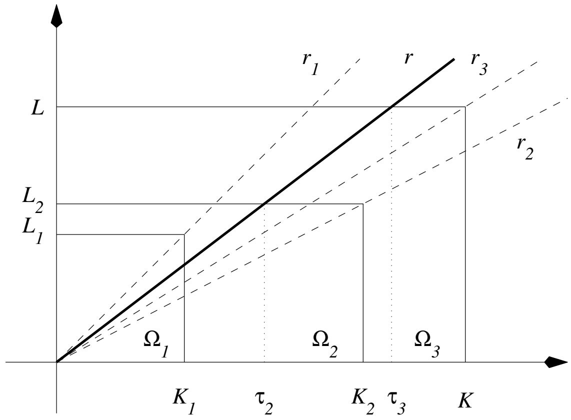


Figure 3: Warranty region for case (E), where $r_2 \leq r_3 \leq r_1$.

3 Analysing the model

3.1 Solving for optimal S

We further assume that the reporting and exercising of claims is immediate and that the repairs are instantaneous. The cost of a minimal repair is c_2, whereas the cost of a complete repair is c_1, with $\mu = \frac{c_2}{c_1}$.

Assuming that the warranty region is predetermined, i.e., Ω is known, there are four decision variables that identify the repair strategy (S), $\gamma = (K_1, K_2, r_1, r_2)$, such that $K_1 \leq K_2$ and $K_1 r_1 \leq K_2 r_2$, as in Figure 2. These four variables determine uniquely the shape of the subregions Ω_1, Ω_2 and Ω_3. Thus we need to select $\gamma^* = (K_1^*, K_2^*, r_1^*, r_2^*)$ so that

$$\gamma^* = \underset{\gamma}{\operatorname{argmin}}\, EC(\gamma).$$

We compute the expected warranty servicing cost per item sold, $EC(\gamma)$, by firstly conditioning on $R = r$ and then removing the condition. Thus our next step is to show how to compute $EC_r(\gamma) = EC(\gamma|r)$. Note that K and L, and thus Ω, is known.

Consider all possible orderings between $r_1 = \frac{L_1}{K_1}$, $r_2 = \frac{L_2}{K_2}$ and $r_3 = \frac{L}{R}$. Thus six cases should be investigated, namely,

(A) $r_3 \leq r_2 \leq r_1$

(B) $r_1 \leq r_2 \leq r_3$

(C) $r_1 \leq r_3 \leq r_2$

(D) $r_2 \leq r_1 \leq r_3$

(E) $r_2 \leq r_3 \leq r_1$

(F) $r_3 \leq r_1 \leq r_2$

and $EC(\gamma)$ computed for each case.

3.2 Detailed analysis of case (E)

We will compute $EC_r(\gamma)$ for case (E) showing all the steps (see Figure 3).

1. Suppose $r \leq r_2$. Then $EC_r(\gamma)$ is a sum of the expected costs $EC_r^{\Omega_i}(\gamma)$ for $i = 1, 2, 3$. We will evaluate each of the $EC_r^{\Omega_i}(\gamma)$ for $i = 1, 2, 3$.

   Over Ω_1, the warranty repair is minimal, thus

$$EC_r^{\Omega_1}(\gamma) = c_2\Lambda(K_1|r). \tag{4}$$

   In order to compute $EC_r^{\Omega_2}(\gamma)$, we need to consider, conditional on $R = r$, the distribution of the time, say $T_{K_1|r}$, to first failure after K_1. Thus we have

$$\begin{aligned} F_{T_{K_1|r}} &= P(T_{K_1|r} \leq t) \\ &= \int_0^t \lambda(x|r)\, e^{-\left(\Lambda(x|r)-\Lambda(K_1|r)\right)}\, dx \\ &= 1 - e^{-\left(\Lambda(t|r)-\Lambda(K_1|r)\right)}, \end{aligned}$$

   where $\lambda(x|r)\,dx$ is the probability of failure at x and $e^{-\left(\Lambda(x|r)-\Lambda(K_1|r)\right)}$ is the probability that there are no failures within (K_1, x). The corresponding pdf is

$$f_{T_{K_1|r}}(t) \quad = \quad \lambda(t|r)\, e^{-\left(\Lambda(t|r)-\Lambda(K_1|r)\right)}$$

   In computing $EC_r^{\Omega_2}(\gamma)$ we consider two possible cases:

   - $T_{K_1|r} > K_2$ with probability

$$\begin{aligned} 1 - F_{T_{K_1|r}}(K_2) &= \overline{F}_{T_1|r}(K_2) \\ &= e^{-\left(\Lambda(K_2|r)-\Lambda(K_1|r)\right)} \end{aligned}$$

   - $T_{K_1|r} \leq K_2$ with probability $F_{T_{K_1|r}}(K_2)$

   Thus

$$EC_r^{\Omega_2}(\gamma) = \begin{cases} \displaystyle\int_{K_1}^{K_2} \left(c_1 + c_2 \int_t^{K_2} \lambda(x-t|r)\, dx\right) f_{T_{K_1|r}}(t)\, dt & K_1 \leq T_{K_1|r} \leq K_2 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

   Similarly,

$$EC_r^{\Omega_3}(\gamma) = \begin{cases} \displaystyle c_2 \int_{K_1}^{K_2} \int_{K_2}^{K} \lambda(x-t|r)\, dx\, f_{T_{K_1|r}}(t)\, dt & K_1 \leq T_{K_1|r} \leq K_2 \\ \displaystyle c_2 \int_{K_2}^{K} \lambda(x|r)\, dx & T_{K_1|r} > K_2 \end{cases} \tag{6}$$

   Finally, using (4), (5), and (6), and simplifying, we obtain

$$\begin{aligned} EC_r^{(1)}(\gamma) &= c_2\Lambda(K_1|r) \\ &\quad + \int_{K_1}^{K_2} \left(c_1 + c_2 \int_t^{K_2} \lambda(x-t|r)\, dx\right) f_{T_{K_1|r}}(t)\, dt \\ &\quad + c_2 \int_{K_1}^{K_2} \int_{K_2}^{K} \lambda(x-t|r)\, dx\, f_{T_{K_1|r}}(t)\, dt \end{aligned}$$

Table 1: Parameters for the numerical example

Usage type	r_ℓ	r_u	$E(R)$	$E(T)$	$E(X)$
Light	0.1	0.9	0.5	1.1118	5559
Medium	0.7	1.3	1.0	0.9575	9575
Heavy	1.1	2.9	2.0	0.7755	15510

2. Suppose $r_2 \leq r \leq r_3$. Using a similar approach, we obtain

$$\begin{aligned} & EC_r^{(2)}(\gamma) \\ = \; & c_2 \left[\Lambda(K_1|r) + \int_{K_1}^{\tau_2} \left(\frac{1}{\mu} + \Lambda(K-t|r) \right) f_{T_{K_1|r}}(t)\,dt + \{\Lambda(K|r) - \Lambda(\tau_2|r)\} \overline{F}_{T_1|r}(\tau_2) \right], \end{aligned}$$

where $\tau_2 = \frac{L_2}{r}$.

3. Suppose $r_3 \leq r \leq r_1$. Then we have

$$\begin{aligned} & EC_r^{(3)}(\gamma) \\ = \; & c_2 \left[\Lambda(K_1|r) + \int_{K_1}^{\tau_2} \left(\frac{1}{\mu} + \Lambda(\tau_3-t|r) \right) f_{T_{K_1|r}}(t)\,dt + \{\Lambda(\tau_3|r) - \Lambda(\tau_2|r)\} \overline{F}_{T_1|r}(\tau_2) \right], \end{aligned}$$

where $\tau_2 = \frac{L_2}{r}$ and $\tau_3 = \frac{L}{r}$.

4. And finally, suppose $r > r_1$. Then we have

$$\begin{aligned} & EC_r^{(4)}(\gamma) \\ = \; & c_2 \left[\Lambda(\tau_1|r) + \int_{\tau_1}^{\tau_2} \left(\frac{1}{\mu} + \Lambda(\tau_3-t|r) \right) f_{T_{K_1|r}}(t)\,dt + \{\Lambda(\tau_3|r) - \Lambda(\tau_2|r)\} \overline{F}_{T_1|r}(\tau_2) \right], \end{aligned}$$

where $\tau_1 = \frac{L_1}{r}$, $\tau_2 = \frac{L_2}{r}$ and $\tau_3 = \frac{L}{r}$.

At last, removing the condition $\{R = r\}$ we obtain

$$\begin{aligned} EC_E(\gamma) \;=\; & \int_0^{r_2} EC_r^{(1)}(\gamma)\,g(r)\,dr + \int_{r_2}^{r_3} EC_r^{(2)}(\gamma)\,g(r)\,dr \\ & + \int_{r_3}^{r_1} EC_r^{(3)}(\gamma)\,g(r)\,dr + \int_{r_1}^{\infty} EC_r^{(4)}(\gamma)\,g(r)\,dr. \end{aligned}$$

Following the same approach, the expected warranty servicing cost $EC(\gamma)$ for the remaining five case in section 3.1 can be derived.

4 Numerical example

We consider the following example, as in (Iskandar, B.P., Murthy, D.N.P., and Jack, N. 2005). The warranty policy is NRFRW with $K = L = 2$ and $r_3 = \frac{L}{K} = 1$, the conditional failure rate is given by (2) with $\theta_0 = 0.1$, $\theta_1 = 0.2$, $\theta_2 = 0.1$ and $\theta_3 = 0.7$, the usage rate R is uniformly distributed over the interval $[r_l, r_u]$, i.e., $R \sim U[r_l, r_u]$ and $r_l \leq r \leq r_u$. Three sets of values for r_l and r_u are considered to represent the usage rates of light, medium and heavy users. These values and the corresponding mean usage rate $E(R)$, mean age to first failure $E(T)$, and mean usage rate at the first failure $E(X)$, are given in Table 1.

Table 2: Expected warranty costs for light usage intensity.

μ	K_1^*	K_2^*	r_1^*	r_2^*	$EC(\gamma^*)$	$EC(\hat{\phi}^*)$
0.1	0.1	0.2	0.4	0.2	**0.3231**	0.3281
0.2	0.1	0.2	0.4	0.2	**0.6427**	0.6469
0.3	0.1	0.2	0.4	0.2	**0.9623**	0.9556
0.4	0.7	1.5	0.8	1.0	**1.1400**	1.1401
0.5	0.6	1.7	1.0	1.0	1.2258	1.2258
0.6	0.6	1.8	1.0	1.0	1.2971	1.2971
0.7	0.6	1.9	1.0	1.0	1.3627	1.3627
0.8	0.6	1.9	1.0	1.0	1.4263	1.4263
0.9	0.5	2.0	1.0	1.0	1.4875	1.4875

Table 3: Expected warranty costs for medium usage intensity

μ	K_1^*	K_2^*	r_1^*	r_2^*	$EC(\gamma^*)$	$EC(\hat{\phi}^*)$
0.1	0.1	0.2	0.4	0.2	**0.3637**	0.3744
0.2	0.1	0.2	0.4	0.2	**0.7274**	0.7319
0.3	0.9	1.8	1.2	0.6	**1.0793**	1.0894
0.4	0.6	1.6	1.2	1.0	**1.2417**	1.2420
0.5	0.6	1.7	1.0	1.0	1.3390	1.3390
0.6	0.6	1.8	1.0	1.0	1.4249	1.4249
0.7	0.6	1.9	1.0	1.0	1.5066	1.5066
0.8	0.6	1.9	1.0	1.0	1.5875	1.5875
0.9	0.6	2.0	1.0	1.0	1.6664	1.6664

4.1 Numerical Results

We assume that $c_1 = 1$ and consider μ varying from 0.1 to 0.9 in steps of 0.1. For every value of μ, we searched for the optimal solution γ^* by performing a coarse grid search with K_1 and K_2 varying from 0.1 to 2.0 with increment of 0.1 and r_1 and r_2 starting from 0.2 incrementing by 0.2 to 3.0. Tables 2–4 summarise our finding for $EC(\gamma^*)$, i.e., the expected warranty servicing cost for the optimal policy γ^*, for unrestricted strategy (S) under light, medium and heavy usage rates. Also, a comparison with the optimal warranty servicing cost $EC(\hat{\phi}^*)$ under restricted strategy (S) in (Iskandar, B.P., Murthy, D.N.P., and Jack, N. 2005) is shown. Bold shows where the results for unrestricted strategy are better than the results for restricted strategy, i.e., where $EC(\gamma^*) < EC(\hat{\phi}^*)$.

4.2 Comparison between different repair strategies

In (Iskandar, B.P., Murthy, D.N.P., and Jack, N. 2005) along with the results for $EC(\hat{\phi}^*)$ shown in Tables 2–4, additional results on strategies consisting only complete repairs or only minimal repairs over Ω are provided. For μ in the range 0.1–0.9, Table 5 summarises the best warranty repair strategy, where MR = always minimal repair, CR = always complete repair, RS = restricted strategy (S), US = unrestricted

Table 4: Expected warranty costs for heavy usage intensity

μ	K_1^*	K_2^*	r_1^*	r_2^*	$EC(\gamma^*)$	$EC(\hat{\phi}^*)$
0.1	0.5	0.6	1.2	1.0	**0.1466**	0.1505
0.2	0.5	0.6	1.2	1.0	**0.2924**	0.2959
0.3	0.5	0.6	1.2	1.0	**0.4381**	0.4412
0.4	1.0	1.4	1.4	1.0	**0.5817**	0.5864
0.5	0.7	1.4	2.0	1.0	**0.7080**	0.7313
0.6	0.6	1.5	2.4	1.0	**0.8168**	0.8384
0.7	0.4	1.6	3.0	1.0	**0.8992**	0.9022
0.8	0.6	1.8	0.8	1.0	0.9500	0.9500
0.9	0.7	1.9	0.6	1.0	**0.9901**	0.9902

Table 5: Optimal warranty repair strategy

μ	light intensity	medium intensity	heavy intensity
0.1	MR	MR	MR
0.2	MR	MR	MR
0.3	MR	MR	MR
0.4	US	US	US
0.5	RS (US)	RS (US)	US
0.6	RS (US)	RS (US)	US
0.7	RS (US)	RS (US)	US
0.8	CR	CR	RS (US)
0.9	CR	CR	CR

strategy (S). From Table 5 the following observations can be made

- For small values of μ (0.1–0.3) the best warranty repair strategy is always to perform minimal repair.

- For $\mu \in [0.4, 0.7]$ the strategy (S) is the best choice. For heavy usage intensity the unrestricted version of (S) is better choice compare to the restricted version of (S).

- For high values of μ (0.8–0.9) the best warranty repair strategy is always to perform complete repair.

5 Conclusions

In this paper, we have extended the results in (Iskandar, B.P., Murthy, D.N.P., and Jack, N. 2005) related to a warranty repair strategy for repairable products sold with two-dimensional warranty. Our strategy is characterised by four parameters, compared to the three-parameter strategy in (Iskandar, B.P., Murthy, D.N.P., and Jack, N. 2005). The additional parameter provides more flexibility and as expected improves the optimal solution in a number of cases. Future work will be to provide more realistic estimation for the intensity function and the distribution of the usage accumulation rates for the three classes of customers, light, medium and heavy users by using information from a warranty database.

References

Blischke, W.R., and Murthy, D.N.P. 1993. *Warranty Cost Analysis*. Marcel Dekker.

———. 1996. *Product Warranty Handbook*. Marcel Dekker.

Chukova, S., Arnold, R., and Wang, D. 2004. "Warranty Analysis: An Approach to modelling Imperfect Repairs." *International Journal of Production Economics* 89:57–68.

Chukova, S., Dimitrov, B., and Rykov, V. 1993. "Warranty analysis. A survey." *Journal of Soviet Mathematics* 67 (6): 3486–3508.

Chukova, S., and Hayakawa, Y. 2005. "Warranty Analysis: Estimation of the degree of imperfect repair via a Bayesian approach." working paper.

Iskandar, B.P., and Murthy, D.N.P. 1999. "Repair replace strategies for two-dimenssional warranty." *Proceedings of the First Western Pacific and Third Australia-Japan Workshop of Stochastic Models in Engineering, Technology and Management*, pp. 206–213.

Iskandar, B.P., Murthy, D.N.P., and Jack, N. 2005. "A new repair-replacement strategy for items sold with a two-dimensional warranty." *Computers and Operations Research* 32:669–682.

Murthy, D.N.P., and Djamaludin, I. 2002. "New product warranty: A literature review." *International Journal of Production Economics* 79 (2): 236–260.

Pham, H., and Wang, H. 1996. "Imperfect maintenance." *European Journal of Operational Research* 94:425–438.

Rigdon, S.E., and Basu, A.P. 2000. *Statistical Methods for the Reliability of Repairable Systems*. John Wiley & Sons.

Robinson, J., and McDonald, G. 1991. Chapter Issues Related to Field Reliability and Warranty Data of *Data Quality Control: Theory and Pragmatics*. Marcel Dekker.

Reflection as a Process; its place and potential in OR/MS education

John Scott,
Department of Management Systems, University of Waikato
(jls@waikato.ac.nz)

Valerie Belton and Helyn Gould,
Department of Management Science, University of Strathclyde

Abstract

This presentation offers encouragement for the increased inclusion of reflection in OR/MS education in keeping with the use of reflection in practice. It will show that good reflective process is fundamental to both learning and to good OR/MS practice. Beyond this we consider the questions:

- What is reflection?
- What frameworks exist for it?
- How can we develop reflective skills?
- What benefits can we anticipate?

We will draw on our own experiences along with those of educators and practitioners from a range of professions, including OR, in seeking to answer these questions.

"The unexamined life is not worth living" (Socrates)

1. Introduction

Simply put, reflection is thought that accompanies action. This is seen in the Learning Cycle, where Concrete Experience(CE) is followed by Reflective Observation(RO), Abstract Conceptualization(AC) and Active Experimentation(AE). This cycle, attributed to Kolb (1983) from earlier work by Kurt Lewin, is seen as being fundamental to Experiential Learning. But it is also very significant for OR/MS practice, because it can be shown to underpin **any** systems thinking or modeling process, hard or soft. For instance, a conventional 'hard' OR/MS model with elements of:

Problem Identification (by CE $\rightarrow$ RO);
Problem Description (RO);
Model Building (AC);
Testing (AE);
Implementation (by AE $\rightarrow$ CE).

Packham and colleagues (1989) presented a good graphic and discussion of this universal underpinning. The Learning Cycle, and the Learning Styles which accompany it, should therefore underpin process and practice courses. Scott (1990) illustrated how useful this could be for OR/MS. The quality of the Conceptualization, or Modeling phase, is dependent on the quality of Reflection phase that proceeds it. Therefore, we need to learn to be good at Reflection.

What do we know about Reflection and how can we stimulate this skill in both our students and ourselves? This presentation sets out to explore this fundamental question, since there has been little explicit reference in the OR/MS literature. We begin with different descriptions of reflection from a range of contexts, showing the scope of the field. Seminal and informative work in the field is introduced, leading into different models of reflection. A core set of activities follows which support the model elements and stimulate growth in the skill of reflection. Three short illustrations show how these activities have been set into an OR/MS learning context. Summary and Conclusions include benefits and potential problems.

2. Reflection Concepts and Contexts

What is reflection?
Boud, Keogh and Walker (1988), pg. 19 describe reflection in the context of learning as "a generic term for those intellectual and affective activities in which individuals (in isolation of groups) engage to explore their experiences in order to lead to new understandings and appreciations." Brockbank, McGill and Beech (2002) sharpen the context and add the aspect of dialogue, saying reflection is "…a process which involves dialogue *with others* for improvement or transformation, whilst recognizing the emotional, social and political context of the learner". Individual reflection is left out to emphasize that it alone is necessary but not sufficient. From our observation, reflection in isolation is a commonly sought after forerunner to reflection through dialogue. Reflection without dialogue looses the richness of different world-views.

The reflection activity involves perceiving relationships and connections between the parts of past & present experiences, iterating between the experience and the relationships being inferred. Creative energy is needed which extracts insights more than exacts analysis, yet both are needed. Extending this, deep reflection requires passive but alert awareness, without purposeful identification, without seeking comparison, without immediate judgment. To understand something systemically, you need to view it passively, with unforced alertness, letting the situation or issue tell its own story. "Someone whose mind watches itself" is Albert Camus' definition of an intellectual.

Donald Schon (1983) is a seminal work. However, the Schon sequel (1987) is more useful here, for the first part both summarizes the background and many of the ideas of the earlier work and directs attention to educating the reflective practitioner, through the concepts of the design studio and reflective practicum, used in the OR/MS modeling studio described later, where the student learns by doing with the help of coaching rather than teaching. Schon's work is thoughtful and well developed. While not without its critics, it is the most referenced work in practical reflection and has inspired many applications. Anyone truly interested in creative reflective teaching could well begin here.

3. Models and Frameworks

We will introduce four frameworks for reflection here that we feel would particularly complement any OR syllabus that has an experiential component.

John Cowan has combined Schon's theory of the reflective practitioner with Kolb's learning cycle in a three part model presented in his book, Cowan (1998). We used this model as the basis for discussion of the featured class, in section 5.2. The three parts of the model are 3 distinct, explicitly planned reflective activities:

Reflection-for-action(R-for-A) – anticipating activity
Reflection-in-action(R-in-A) – analytical reflection during activity
Reflection-on-action(R-on-A) - evaluative, post-activity reflection

Cowan says these forms of reflection should be explicit and a planned part of students' learning. Kolb's learning cycles should still occur between the planned reflective phases, as smaller coils within the reflective spring.

Graham Gibbs (1988) has a reflective cycle of six stages. Parallels with the Kolb cycle can be seen, but Gibbs provides useful cues for reflection couched as a sequence of questions. An emphasis on feelings can be seen, confronting the student to consider both good and bad. Reenacting the situation with new insight allows the student to think what to do "if it arose again".

For those in need of a something simpler, Double-loop learning of Argyris is well worth considering. Double-loop learning asks questions on the efficacy of the process being used, while single-loop learning just considers the usefulness of the result.

Other frameworks we found useful/informative are the Boud, Keogh and Walker (1988), (1993) works, where their extensive experience in education is used to develop a thoughtful generic model for reflection in learning contexts, and the work of Mezirow (1981), where layers of reflectivity surround the object of interest, beginning with the consciousness layers and moving to the three layers of critical consciousness for "becoming aware of our awareness".

We thus see a range of frameworks from the set available and these will feature in the presentation. A common element is in encouraged recognition of the difference between espoused intentions and actions taken. Choosing a useful framework, model or theory comes through choosing one that resonates with both situation and self. It should also be influenced by underlying values.

4. Activities

The models and their elements give us a framework for action. Activities bring a framework alive and we now present a set covering the range that would seem to be useful in common OR/MS activities. References provide the finer detail for helping tune the activity to the context.

<u>Writing</u>: Reflection can be part of any writing, but is more active in a journal. Learning journals are given good coverage in Moon (2000), where the purposes of use include recording experience, encouraging critical thinking and a questioning attitude, enhancing problem-solving skills, and increasing personal ownership of learning. There are unstructured forms, such as free-writing, writing an unsent letter, having an imaginary dialogue with a person or projects - where projects are addressed as if they were people, and double-entry journals (left column for diary / factual entries; right hand for thoughts and reflective comment). Unstructured

forms are often more effective after practice with more structured forms. Structured forms might come as a series of questions to be answered, a Plus / Minus/ Interesting (PMI) list, or perhaps a SWOT analysis.

Talking and Questioning. Questions oil a conversation; interesting ones energise it. Two frequently used aspects of talking and questioning are Advocacy and Enquiry (A&E). Senge (1990) and colleagues recommend balancing A&E – skills for honest conversation they call them - so that conversation can get beyond simply reinforcing the prior views that pure advocacy brings.
Senge (1990, pg. 200) offers the guidelines of Diana Smith for mixing A&E. A sample are:
When advocating your view:
- Say how you arrived at that view,
- Encourage others to explain your view; "Do you see gaps in my reasoning?",
- Actively enquire into others views; "What is your view?"; "How did you arrive at it?".

When enquiring into others views:
- Check your assumptions; "What I believe you to be saying is …"
- Don't bother asking questions if you are not interested in the answers.

These can be incorporated into an A&E script (Dillon (1990), which the students can use on a topic, with R-on-A as follow-up. Aspects of Dillon can be incorporated into A&E, such as Silences and Speakers own question to themselves.

Dialogue is found in several other forms, including: Socratic dialogue (Socrates, the paradoxical teacher who does not teach but acts as provocateur; sometimes provoking self-discovery, sometimes confusion), Structured dialogue (Example 5.2 in John Cowan (1998)), Reflective dialogue as in Brockbank and McGill (1998), and Reflective questions and probing questions as mentioned in Barnett (1995).

Lastly, with dialogue comes the need for attentive listening. Chapter 5 (attentive listening) of Boud, Keogh and Walker (1988) may assist.

Drawing stimulates the right brain. A survey of visual representations available for each stage of problem solving is detailed in Scott (2005). If you allow 2-D matrices (e.g. a threats vs possible controls matrix), coverage was found to be good, although a little thin for finding alternatives and implementation stages. Options include:
- Mind maps, being hierarchies of related concepts (or ideas etc.) fanning out from a central hub concept.
- Hexagons, where brainstormed ideas are written on to hexagons without judgement and then fitted together into patterns.
- Problem structuring maps, Rosenhead (1989), where Cognitive Maps and Rich Pictures are two good examples.

These range from relatively structured (technique rules based), like mind maps, through to relatively unstructured, like rich pictures. Drawing can be seen as an awareness raising tool but can be a problem-solving technique in its own right.

Writing, Talking and Drawing are primary colours. Mixing these we get further activities, each important in their context.

5. Published use in OR

5.1 Reflection-in-action = the Modeling Studio.

Schon (1987) uses the design studio as a concept for coaching architectural design. He refers to “coaching in the artistry of reflection-in-action” where the student cannot be taught but can be coached, gaining intelligence learned from performing under the guidance of an accomplished performer. Schon conjectured that learning professional artistry can best occur in a studio setting, where the students learn by doing in a low-risk environment, undertaking projects that simulate and simplify practice. These ideas were taken up in OR in the “modeling studio” for developing practical skills in the art of modelling. Powell (1995), (1998) describes how this was done including the history, useful resources for developing a modeling studio approach along with practical advice and heuristics.

5.2 Learning Points Essay

A Learning Points Essay (LPE) (Belton, Gould and Scott (2005)) requires reflection from the student but is not a learning journal that seeks to regularly document and reflect on learning and process. Instead a LPE *reports* on the exploration of process and learning outcomes of an activity and the subsequent reflection on this. The LPE also attempts to separate the reflective activity from the assessment mechanism. The students are encouraged to reflect (and write reflectively if possible) on the activity. Once they have explicitly identified a range of learning points the student then compiles their LPE based on the output of this reflection. The student then considers the assessment aspects when selecting the points to include in the LPE, allowing effective reflection unrestricted by assessment to that point. We have become more explicit about this separation as the class has developed and have also provided more support in the form of questions to prompt reflection.

5.3 A Structured Learning Log

Adding reflective activities, using a structured learning log, was used with a production scheduling simulation exercise, in Basnet and Scott (2004). The log was used to record time and summarized detail of the thinking students were using when working with the simulation. Included were four levels of reflection based on single and double-loop learning. The results were fed back to the students in a formal class and briefly discussed *before* they wrote their final essay which required them to consider what they had done and learnt from the experience.

6. Summary and Conclusions

Reflection as seen throught the lens of literature is a generic term for adding understanding and appreciation to an experience. It can be done in isolation but dialogue with others is encouraged. Models and frameworks for reflection have common elements in encouraged recognition of the difference between words & activities. Parallels with modelling steps were seen and can be effectively used, modelling known OR/MS application best practice. Parallels with modelling steps in both hard and soft OR methodologies exist.

Activities bring frameworks and principles alive by giving guidelines for action. Activities are made up of combinations of the three primary colors of writing, talking (including questions and listening) and drawing.

Modelling, by its very nature, is a reflective act. Therefore a general interest in reflection, its forms and practices, should complement modelling, making the modelling act more natural, less mechanical and giving students the benefits of learning in a supportive environment. This is particularly true in situations that are "messy", where the path to follow is not clear. We recommend introducing a reflective framework from the set available that best parallels the OR modelling process the students are working with at the time. With more refined use, it can become part of every action itself, akin to reflection-in-action and Argyris's double-loop learning.

Common pitfalls associated with the explicit encouragement of reflection include: students working through a reflective checklist in a mechanical fashion, a mismatch between the type of reflection chosen and the modeling situation, and treating reflection as if it was an intellectual exercise (Boud and Walker (1998)). Our recommended process for teaching reflection with OR goes some way toward resolving these. Other suggestions will be made in the presentation.

In summary, we are suggesting making reflection parallel the OR model in use & involving students in adaption and refinement. Limited experience in doing this in OR has been published, yet its importance in practice is profound. Hopefully more active, explicit reflective use can be encouraged in OR education.

7. References

Barnett, B. 1995. "Developing reflection and expertise: can mentors make the difference", *Journal of Educational Administration,* Vol. 33, No. 5, pp.45-49.
Basnet, C. and Scott, J. 2004. "A spreadsheet-based simulator for experiential learning in production management", *Australian Journal of Educational Technology*, Vol. 20, No. 3, pp.275-294.
Belton, V., Gould, H. and Scott, J. 2005. "Developing the reflective practitioner: designing an undergraduate course", *Interfaces,* in press.
Boud, D., Keogh, R. and Walker, D. 1988. *Reflection: Turning Experience into Learning,* Kogan Page: London.
Boud, D. and Walker, D. 1993. "Barriers to reflection on experience". In Boud, D., R Cohen and R. Keogh (eds.) *Using Experience for Learning,* pp.73-86, Kogan Page: London.
Brockbank, A. and McGill, I. 1998. *Facilitating Reflective Learning in Higher Education*, SRHE and Open University Press: Buckingham.
Brockbank, A., McGill, I. and Beech, N. (eds.) (2002), *Reflective Learning in Practice*, Gower, Andover. Brooks, AK
Cowan, J. 1998. *On Becoming an Innovative University Teacher – reflection in action,* SRHE and Open University Press: Buckingham.

Gibbs, G. 1988. *Learning by doing: a guide to teaching and learning methods*, FE Unit, DES: London.
Kolb, D. 1983. *Experiential Learning*, Prentice-Hall: Englewood Cliff, NJ.
Moon, J. 2000. *Reflection in Learning and Professional Development – theory and practice*, Kogan Page: London
Mezirow, J. 1981. "A critical theory of adult learning and education", *Adult Education*, Vol. 32. No.1, pp.3-24.
Packham, R., Roberts, R. and Bawden, R. 1989. Our faculty goes experiential. In *Making Sense of Experiential Learning* Weil S. and McGill I. (eds.) Open University Press: Buckingham.
Powell, S. 1995. "The Teachers Forum: Teaching the art of modeling to MBA students," *Interfaces*, Vol. 25, No. 3 (May-June), pp. 88-94.
Powell, S. 1998. "The studio approach to teaching the craft of modelling," *Annals of OR*, Vol. 82, No. 1, pp.29-47.
Rosenhead, J., (ed.).1989. *Rational Analysis for a Problematic World*. John Wiley: Chichester, England.
Schon, D. 1983. *The Reflective Practitioner*, Basic Books: New York.
Schon, D. 1987. *Educating the Reflective Practitioner*, Jossey-Bass: San Francisco.
Scott, J. 1990. "OR methodology and the learning cycle," *OMEGA*, Vol.18, No. 5, pp.551-553.
Senge, P. 1990 *The Fifth Discipline: the art and practice of the learning organization*, Doubleday/Currency: New York.

Is Optimisation a Waste of Time? Recollections of a Retired Lone-Star OR Practitioner

Colin W Walker
Retired Member
Wellington, New Zealand

Abstract

This paper will review my personal experiences in a career that started 40 years ago. Most of this time was spent working alone in the Oil Industry, and I will look at some commercial OR applications including Capacity Planning, large-scale Vehicle Scheduling, and Depot Location.

The transition from having a mechanical calculator on one's desk (with computers far in the background), through programmable electronic calculators, to having your own PC has meant changes in the way problems are tackled.

In a profit-conscious commercial environment there may be pressure to produce "quick" results – hence the title of this paper. I will provocatively challenge the notion that "optimal must be best", and look at the role of non-quantifiable factors.

I promise that my presentation will include more anecdotes than formulae, and should appeal to those who have "been there, done that", and also to those who "have no idea of what things were like in the years BC (before computers)".

1 Background

.

1.1 My First Job

On leaving the University of Otago 40 years ago, where I had studied under Desmond Sawyer and Geoffrey Jowett, I had the choice of well over a dozen job offers. Being young and hard up, but with an expensive hobby (yachting), I took the highest paid offer, as Statistician with a branch of the Meat and Wool Boards. The main tool available was a Monroe mechanical calculator with one memory. I still remember the frustration of spending days extracting eigen-values from 6 x 6 matrix in deriving a discriminant function for the fecundity of ewes. Being the only numerate person in a small office I was also charged with developing a weight gain performance recording system for beef breeders; a world first.

1.2 Changing Jobs and Changing Environment

In 1968, I was persuaded by friends who worked there, to apply for a position with an Oil Company. The big attraction was a new IBM-360 computer due to be delivered. This was an expensive beast, and my introduction to the commercial environment saw “all hands and the cook” working long hours to get all the company’s commercial applications transferred from an old 1401 to the new 360.

1.3 First OR Study

My day soon came with a request to study the delays at a shared tanker loading facility in Auckland. A simulation programme, quickly written in FORTRAN, showed that the cheapest and quickest strategy would be to add an extra pump. This simple job secured my personal credibility, and that for OR, in the Company.

2 More Significant Projects

2.1 Home Heating Oil Delivery Scheduling

A significant problem for the Oil Industry in the late 1960’s was the rapid growth in domestic oil-fired central heating. I was teamed with an engineer and we produced a system which each week forecast the demand for 20,000 customers, and scheduled the tankers to make the required deliveries. The quantity delivered to each customer was fed back into the system to update each customer’s forecasting parameters. Although we had found an accurate forecasting model, it was discarded in favour of a simpler and less accurate model which could be understood by less skilled clerical staff. We used IBM’s Vehicle Scheduling Package (Clarke & Wright, Simple Savings Algorithm) which was offered “free” to major IBM users. This study exposed some interesting intra-company conflicts, such as those between Sales and Operations, or between dispatchers and drivers. The commercial sensitivity of the market meant that I was not allowed to present a paper on the system to the OR Society till well into the 1970’s, after the oil-shock had killed the market.

2.2 LPG Depot Location

Finding offshore oil reserves in the Taranaki Bight led to my involvement in two major studies, where I was a local secondment to a team of overseas “experts”. The first was a large Monte Carlo (again in FORTRAN) risk analysis. The next was a Depot Location study for the distribution of LPG from the Maui field, throughout NZ. The London HO had sponsored a research project and persuaded us to use the Eilon et al heuristic model. This was written in FORTRAN V, which was not compatible with IBM’s FORTRAN IV. The data was punched onto paper tape and telexed to London, who ran the model on their terminal connected to a large UNIVAC in the Midlands. I would telex changes to them every afternoon, and would have the results telexed back to be on my desk next morning. One day I changed one line of the model, and next morning there was only an apology and the explanation that I had overflowed the biggest commercial computer in the UK. Next morning I had my results, with the explanation that the data had been bounced off a military satellite to a larger UNIVAC in Houston. This is claimed to be the World's first commercial application of data transmission via satellite.

2.3 Monte Carlo Method

About 1970 I latched onto Decision Trees as a useful tool, as it provided a concise framework to set out a problem, communicate with the owner of the problem, and elicit estimates of probabilities costs and benefits from people who were meant to know these things. Of course, most of these data were for future time periods, and most experts were reluctant to hazard a single guess, but were quite happy to discuss a range of outcomes, and a "most likely" value. I soon developed a method of anticipating what the underlying distribution may be and was able to construct a Monte Carlo model (again, in FORTRAN) for most problems.

About 1979 I persuaded the Company to buy an Apple II computer, but it was not till a couple of years later when VisiCalc software became available and it was upgraded to Apple IIe that this machine became really useful. This was a time when a HO policy decision meant that all Capital Expenditure proposals had to be subject to a DCF analysis. As I had the only book of Discount Factor tables in the office I was in demand as the resident guru on such matters. Again I found that the cost and cash-flow estimates were most uncertain information, and devised a way to sample the probabilistic variables (this time in Basic) and introduce them into the VisiCalc spreadsheet for the Monte Carlo DCF. With enhancements in software, spreadsheet simulations are a tool I still employ.

2.4 Capacity Sizing

The EOQ formula is one found in most first year OR or calculus courses, and is an example of what I consider to be "useless optimisation". If one actually plots the cost curve, it is common to find that it is virtually flat in the domain of half to double (or more) the optimum. This is where careful communication with the management involved can bring in non-quantifiable preferences, such as human factors. In real life, many of the factors involve discontinuities imposed by constraints such as The Building Code, or Dangerous Goods Regulations. Again, many of the costs may be uncertain, and the problem may be better solved in a Monte Carlo Spreadsheet.

3 Conclusion

In my student days the Treasury had the only computer in NZ. Since then the availability of hardware and software has made many methods, long recognised but impractical to perform, available.

Moving from Tradeoffs to 'Trade-ons'

Vicky Mabin and John Davies
Victoria Management School
Victoria University of Wellington
New Zealand
vicky.mabin@vuw.ac.nz

Abstract

Tradeoffs are a central concern in OR/MS, addressed by methods such as math programming, heuristics, multi-criteria and simulation. The Theory of Constraints 'Evaporating Cloud' method also addresses tradeoffs. This paper explores some typical OR/MS tradeoffs using the Evaporating Cloud, with sometimes surprising results. We argue that the EC is valuable on its own or in conjunction with OR/MS methods, and urge that tradeoffs be 'evaporated' wherever possible rather than simply 'optimising'. This allows us to 'trade on' the best elements of the situation, without compromising.

1 Introduction

Tradeoffs are a central theme of OR/MS approaches where the pursuit of multiple objectives is seen to involve seemingly mutually exclusive choices. The decision to trade off one objective against another implies an acceptance that it is impossible to simultaneously optimise across these multiple objectives, and thus compromise is seen to be unavoidable. OR/MS methods such as mathematical programming, heuristics, multi-criteria and simulation provide practical ways of working with such accepted compromises. The Theory of Constraints 'Evaporating Cloud' method (Goldratt 1990), was designed to address conflict or dilemma situations, which can be viewed as being equivalent to trade-off situations where there is no acceptable compromise. The EC can therefore be used wherever OR/MS methods address a tradeoff. This paper provides an example of examining a typical OR/MS tradeoff using TOC's Evaporating Cloud, with sometimes surprising results.

Tradeoffs are ubiquitous in the real world so it is hardly surprising that they are a central theme/subject of OR/MS approaches. Indeed they are the main reason for the development and use of methods such as mathematical programming for constrained optimisation, multi-criteria methods, simulation, decision trees, and heuristic methods. For some of these methods, the tradeoff is very explicit – others less so.

The Theory of Constraints Thinking process tools, particularly the Evaporating Cloud (EC), have been designed to explore and solve/dissolve/resolve dilemmas. The EC thus provides another way of viewing trade-offs, providing insightful ways of viewing the tradeoffs and leading to alternative ways of dealing with them. Sometimes the trade-off can be avoided completely by reframing the problem. The EC may be used on its own or in tandem with OR/MS methods.

Put simply, a trade-off occurs when we feel pulled in two different directions, on the one hand wanting to take one action, but on the other, we also want to take an opposite action. The advantage of the Evaporating Cloud process is that it makes us explicitly consider and critique the assumptions we make implicitly when we frame the problem as a tradeoff. We deliberately try to challenge some of the assumptions, to create more options for dealing with the tradeoff. Ideally we may be able to satisfy both sides, and hence avoid having to trade-off to meet our desired ends.

There are many examples of tradeoffs in practice and in the OR/MS literature:

- ***Inventory control*** – thousands of papers have been devoted to the EOQ and its variants. The tradeoff here is between incurring set-up costs of multiple production runs and inventory costs, or between the costs of multiple orders versus inventory costs. In the production setting, it is a necessary tradeoff between meeting demand through frequent small production runs, or from stock accumulated through infrequent large production runs, ie producing for stock as a means of supplying customers. At its extremes, it involves producing only to demand incurring costs from frequent set ups and production runs, or meeting demand only from stock accumulated from infrequent large production runs incurring greater inventory costs. The choice can therefore be seen as one between investing in speedy production facilities that can rapidly react to demand, or investing in inventory, so that the trade-off relates to the balancing of aggregate investment costs, rather than the balancing of operating costs. The basic order quantity problem was revisited by Goldratt (1990), demonstrating how the Evaporating Cloud can be used to suggest solution/s that minimise BOTH inventory holding AND order costs simultaneously. A depiction of this tradeoff as an EOQ and as an Evaporating Cloud (after Goldratt, 1990) is provided in the Appendix in the purchase order context.
- ***Centralisation/decentralisation*** - this is a classic dilemma that comes in many forms, under various names, eg Location/allocation. For example balancing economies of scale in warehousing against a need for speedy delivery of supplies to customers ie balancing facility costs vs transport costs for a distribution network; balancing consistency of policy versus tailored responses to demands in different areas for a service organisation; balancing central purchasing power with greater distribution costs for a retail organisation.
- ***Institutional funding vs competitive bid based funding*** – typically a dilemma in scientific, educational and other publicly-funded organisations, balancing a need for certainty in funding to allow planning of research programmes against a need to fund the 'best' programmes as assessed against a set of criteria. The latter is what the funders want, the former is preferred by people being funded.
- ***Scheduling/rostering/ network planning*** - meeting service needs that balance cost to the organisation against the costs to the individual workers arising from anti-social work patterns; or minimising the cost of staffing versus schedule flexibility and/or robustness to disruptions.

Many of these can be seen as a tension between local and wider organisational needs and/or performance measures.

We will now demonstrate how we can reframe a typical OR/MS trade-off through the lens of TOC's Evaporating Cloud, with a view to providing some new ways of resolving such dilemmas.

2 Warehouse/Distribution Example

This example is a typical centralisation /decentralisation case, that is faced by firms wishing to optimise their inventory and distribution systems. In this example, let us assume that the firm provides goods to customers that are sold through its retail stores, but are delivered from the firm's warehouses/depots to the customers' premises where they are installed by the delivery person. The question is how many warehouses/depots should they have and where should they be located. This can be modelled as a facilities location problem, with attention to the inventory costs that would be incurred at the facilities (warehouses or depots) depending on the size of the inventory held, plus the transport costs for the delivery to the customer and installation at the customers' premises.

2.1 The Traditional OR Approach – Optimization

With such tradeoffs, the owner would consider it ideal if they could minimise both warehouse costs and transport costs simultaneously, but the former would require one central warehouse and the latter would require many warehouses. Traditionally, the problem might be modelled and solved as a mixed integer problem to identify the best number of warehouses. This might be repeated for various scenarios eg for different demand patterns and service levels. The results can be plotted (Figure 1) to reveal a similar curve as the EOQ model. This 'optimisation' approach used assumes that minimising both costs cannot happen simultaneously. It assumes there is a satisfactory compromise, and this is found where the Total cost curve is at its minimum value.

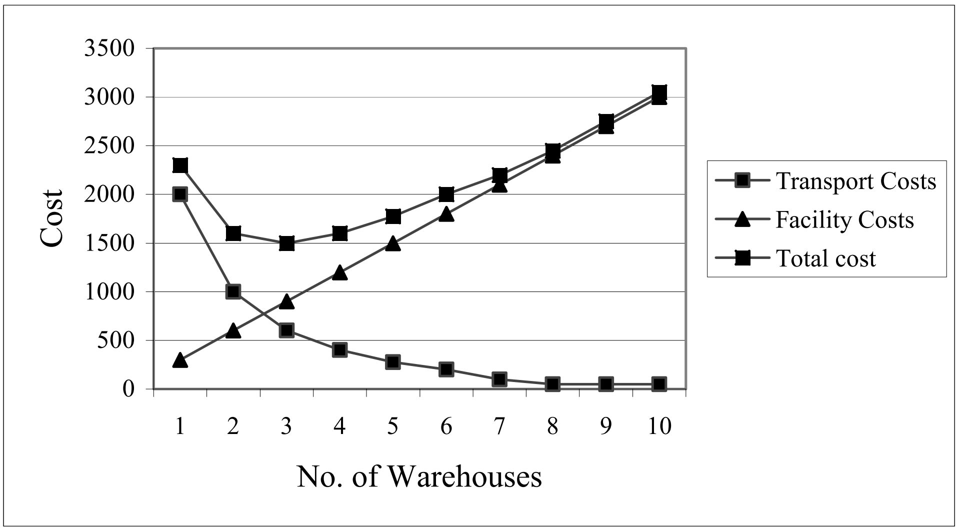


Figure 1: Warehouse/Distribution Tradeoff

2.2 Using the Evaporating Cloud

Let us now reframe this problem using the Evaporating Cloud method first described by Goldratt (1990), and since refined by Goldratt (1994), Dettmer (1998), Cox et al (2003), Cox et al (2005). The EC starts from the premise that a compromise is not the best we can do, and that a win-win solution can be found.

2.2.1 Framing the Tradeoff as an Evaporating Cloud

To frame this tradeoff as a cloud, first we identify the two opposing actions: in this case, having One central warehouse, v having Multiple warehouses, which are shown in boxes labelled D and D' (read as Not D). Then we identify the need that each side is aiming to achieve: on the one side to Minimize facility costs, on the other side to Minimize transport costs, shown in boxes B and C. Finally we identify the common objective in box A, to minimise total costs per unit. B and C are required if A is to be achieved. D and D' are the prerequisites of B and C respectively. Using this method we may derive the following cloud diagram shown in Figure 2.

We read this cloud (Left to Right) as:

> In order for Firms to Minimize total costs per unit, they must Minimize facility costs, and in order to do this, they must have One central warehouse.
>
> On the other hand, in order for Firms to Minimize total costs per unit, they must Minimize transport costs, and in order to do this, they must have Multiple warehouses. Hence the conflict! (denoted by the lightning bolt arrow).

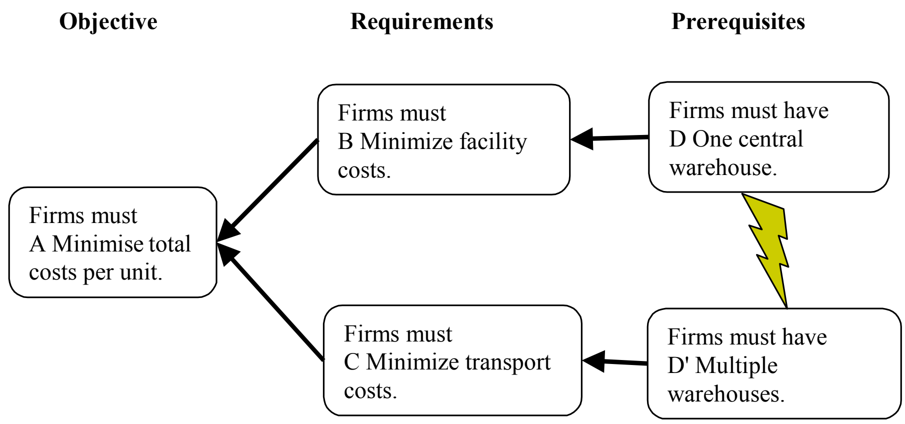


Figure 2 The Tradeoff Reframed as an Evaporating Cloud

2.2.2 Searching for Solutions: Surfacing Assumptions

The purpose of the Evaporating Cloud is, as the name suggests, to evaporate or resolve an undesired situation (a cloud). The search for solutions to evaporate the dilemma can be done several ways. We can generate ideas for solutions (Goldratt calls these 'injections') by examining the cloud directly, or by methodically surfacing assumptions and then seeking to break those. The direct way is done by asking how could we have B without D (or C without D'). Or we may ask, how can we have D AND D' simultaneously.

The methodical way starts with surfacing assumptions that underlie the dilemma. We list our assumptions, which are surfaced by completing the sentence …In order to … we must … because ….. For example:

BD In order to Minimize facility costs, Firms must have One central warehouse because ….

BD1: Each facility incurs costs.
BD2. Total inventories increase with the number of warehouses.
BD3. One central warehouse has least expensive facility costs.

A few more assumptions are shown below (not necessarily complete or comprehensive, but they will provide an indication of how we proceed).

Arrow (Box)	***Assumptions***
Box A	A1. Minimising total costs is desirable.
AB	AB1. Facility costs are a significant part of total costs.
	AB2. Inventory holding costs are a significant part of facility costs.
	AB3. We pay for the facilities.
AC	AC1. Transport costs are a significant part of total unit costs.
	AC2. We pay for transport from facilities to the customer.
BD	BD1. Each facility incurs costs.
	BD2. Total inventories increase with the number of warehouses.
	BD3. One central warehouse has least expensive facility costs.
CD'	CD'1. Transport costs are only reduced if we have more warehouses, because ….
	CD'2. Warehouses can be spread over the region, and because …
	CD'3. We can make use of bulk transport for the main routes, and more suitable local transport for local deliveries.
DD'	DD'1. One central warehouse AND multiple warehouses cannot exist at the same time.

It is pertinent to note that an implicit assumption in the optimisation frame is that there does exist an acceptable compromise, whereas in the EC frame, we challenge this need for compromise, and instead seek a win-win solution, where both sides achieve their needs.

Then we explore ways of resolving the dilemma by challenging the assumptions. Ideas that break assumptions and resolve the cloud are termed 'injections'.

2.2.3 Identifying Some Possible Solutions (Injections)

•Breaking CD': One Central warehouse may actually reduce transport costs if …

– we simplify/streamline inventory holdings and order processing, hold less stock, ship only as needed, …

•Breaking BD: Multiple warehouses could reduce facility costs if …

– each warehouse is simple and efficient, and stocks held are minimal … if the supply chain is excellent.

We would come up with a fuller list of assumptions and injections such as in Table 1.

Normally in TOC we would critique and develop these injections further but we will not do so here. For a complete description of the method, see Cox, Mabin & Davies, 2005.

While the list derived would be specific for the particular situation, the general approach and some of the suggestions are likely to apply more generally.

3 Discussion

This facilities/distribution example was loosely based on the PhD research of one of the authors (Mabin, 1981). As such it provides the opportunity to reflect on practice then and now relating to this general topic. The particular example was tackled back then using a specialised, state-of-the-art MIP modelling package on mainframe computers, and contributed to the field of knowledge at the time. Some of the ideas surfaced in the injections were actually incorporated into the model, but in such a way so as to **fit with** the MIP formulation. It is interesting to note the developments that have since taken place and spawned the field of Supply Chain Management (SCM), and of course many of the injections are now standard SCM. One wonders if we had used the EC back then (had it been known), whether we would have come up with these ideas, or whether our thinking would have remained limited to concepts we already knew about.

We can only postulate a reply based on several years of using the EC to state that it has often allowed us to move beyond the constraints forced by traditional thinking – accepting that we are in the horns of a dilemma – and allowed us to generate new and far better solutions that we had not thought of before, that would have been hidden by traditional thinking.

Our learning from this experience is that we often accept the need to trade off objectives, to accept a compromise, when in fact there is no need to do so. We can have the best of both worlds if we expose our assumptions and challenge them, eg facility and transport costs can be minimised simultaneously through good SCM.

The importance of surfacing implicit assumptions and challenging our assumptions cannot be stressed enough. The mathematical programming approach starts with an implicit assumption that a compromise is both inevitable and acceptable if our outcome measures are comensurate. Breakthroughs in our thinking come by challenging this point. They also arise through challenging the givens. In any of the usual tradeoff situations, there are many givens, like the cost of facilities/ the cost of transport, which the so-called 'optimisation' model then also takes as givens. The EC frame encourages us to question the assumptions concerning these givens, and to discover whether our 'optimisation' is spurious.

The MIP approach requires the construction of a mixed integer program and the collection or modelling of the component costs. Modelling of facility costs requires an understanding of inventory needs for a warehouse of a given size serving a given area, plus the holding costs for such inventory, as well as facility construction, maintenance and operating costs. Before launching into a massive costing and modelling exercise, the EC method would provide a check on the assumptions implicit in such a model. It may well uncover options for further investigation before the MIP is developed. Or it might suggest that a different approach shows promise – like a focus on streamlining the supply chain operations, improving forecasting of demand, and establishing better communication throughout the supply chain so that the same service level can be provided with less inventory.

The time involved in doing this analysis would almost certainly pay off in terms of better understanding of the system to be modelled. It will undoubtedly create much discussion of the system and in doing so produce a clearer understanding between modeller and owner of the system. It may also save considerable modelling time. But more importantly perhaps it may lead to better outcomes by improving the system itself rather than optimising subject to the constraints of the current system.

Even our chosen objective may be spurious. For the EOQ, the aim is essentially to minimise inventory costs. But should we minimise inventory costs, operating costs or investment costs? Or should we maximise profit? Often in mathematical programming we assume that minimising costs is equivalent to maximising profits, but the two yield different results, so shouldn't we maximise profits in organisations that are profit making organisations? Similar questions have been raised by Zeleny (1981) and Gass (1989), and yet we are often blind to such issues in our desire to 'optimise'.

As we saw above, solutions without compromise do exist if we step outside the mathematical model, and although we may come across them in the modelling process, we often ignore them for many reasons:

- Maybe for the sake of the model, research programme, degree or the current paper, we want to continue with the current framework.
- The challenge of figuring out how real-world practicalities can be dealt with within the model is often the 'adrenalin' that good modellers thrive on. Yet this action may block the modeller from realising the opportunities they provide for breakthroughs which would improve, not just optimise, the system.
- Or we may not realise the relevance of significance of the possibilities they offer; or there may constraints to the renegotiation of the focus or scope of study. Both these tend to happen if the modeller has been delegated the task and/or is operating at too low a level in the organisation. We may be limiting ourselves to act in what Eilon (1977) describes as the technician role rather than adviser role.

The meta-decision process is thus seen to be an important determinant of our receptiveness to such opportunities, as argued by Russo & Schoemaker (1990), Nutt (2002).

Other methods besides mathematical programming would also be candidates for reframing using the EC. Multi-criteria decision making is an obvious example where the tradeoff between competing objectives is considered explicitly. Moreover there is an explicit acceptance of the need to compromise in the absence of one alternative that performs best on all criteria. Good MCDA modellers take the discussion following the modelling process to the level of debating how all criteria might be achieved simultaneously. This could be enhanced and possibly earlier in the process by the EC. Similarly simulation and heuristic modelling addresses the desire to do as well as possible on many criteria, under a range of different conditions. The meta-decision should include the circumstances under which all criteria would be maximised simultaneously, and extend to a search for a strategy to achieve such a position. The EC would be an invaluable approach to guide such discussion.

4 Conclusions

In conclusion, this example has shown how the Evaporating Cloud from Theory of Constraints is able to help us explore ways of going beyond compromise solutions. It does this by surfacing and challenging the assumptions that we typically make when framing problems using traditional OR/MS modelling methods. There will inevitably be times when we still need to resort to a math model to find the best tradeoff, but the form of that tradeoff may be very different to the one we started with. By exploring possible solutions through the EC, we can move from a 'trade-off' situation to one of 'trading on' the possibilities uncovered by the EC.

More generally, Soft OR methods are helpful in the problem-structuring phase to explore the implicit interactions and boundaries of the problem, allowing better optimisation methods once the limits of possibility for the system have been defined.

Rothkopf's (1996) editorial argues, "that no operations research model is ever complete. Therefore it is dangerous to stop thinking once the model is built. To be effective, one must use models as aids to further thought." In like fashion, no paper is ever complete, nor argument done: we offer this paper as an aid to further thought.

5 References

Cox, JF, JJ Blackstone, and JS Schleier. 2003. *Managing Operations: A Focus on Excellence*. North River Press, Great Barrington, MA.

Cox, JF, VJ Mabin, and J Davies. 2005. A Case of Personal Productivity: Illustrating Methodological Developments in TOC. *Journal of Human Systems Management 24*, Special Issue on TOC, 39-65.

Daskin MS. 1995 *Network and Discrete Location* Wiley-Interscience: New York.

Dettmer, HW. 1998 *Breaking the Constraints to World-Class Performance*, ASQ Quality Press, Milwaukee, WI.

Eilon, S. 1977. Technician or Adviser? Editorial, *Omega* 5:1, 1-6.

Elshafei, AN and KB Haley, 1974. *Facilities location: Some formulations, methods of solution, applications and computational experience* OR Report No 90, University of Birmingham.

Gass, S, 1989. Model World: A model is a model is a model is a model, *Interfaces* 19/3,58-60.

Goldratt, EM, 1990. *What is This Thing Called The Theory of Constraints and how should it be implemented?*, The North River Press, Croton-on-Hudson, NY. 162 pp.

Goldratt, EM, 1994. *It's Not Luck*. North River Press, Great Barrington, MA.

Mabin, VJ, 1981. *Variability in Distribution*, Ph D Thesis, *Department of Operational Research*. University of Lancaster: Lancaster, UK. 340 pp

Mercer, A, M Cantley and GK Rand. 1978. *Operational Distribution Research: Innovative case studies*, Taylor and Francis, London.

Nutt, PC 2002. *Why decisions fail*. Berrett-Koehler, San Francisco.

Ramos MT and J Saez, 2005. Solving capacitated facility location problems by Fenchel cutting planes, *JORS* 56, No 3, 297 – 306.

Rosenhead, J, ed. 1989. *Rational Analysis for a Problematic World*. John Wiley & Sons: Chichester, 370 pp.

Rothkopf, MH. 1996. Editorial: Models as aids to thought. *Interfaces* 26/6, 64-67.

Russo, JE and P Schoemaker, 2002. *Winning Decisions*. Piatkus, London.

Sridharan R, 1995. The capacitated plant location problem. *Eur J Opl Res* 87: 203 – 213.

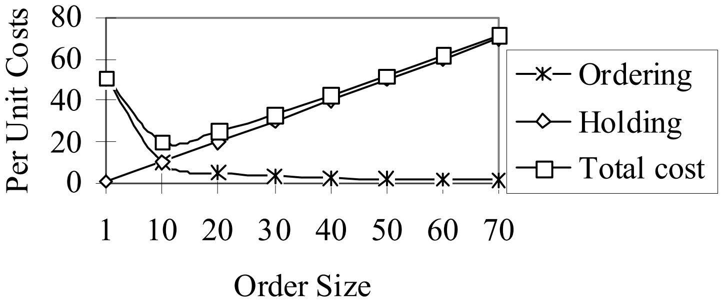
Watson-Gandy CDT. 1972. Warehouse/depot location in Christopher, M and G Wills (eds) *Marketing logistics and distribution planning* Allen and Unwin, London.

Zeleny M, 1981. On the Squandering of Resources and Profits via Linear Programming. *Interfaces* **11** (5), 101-107.

Table 1. Assumptions and Injections for the Evaporating Cloud

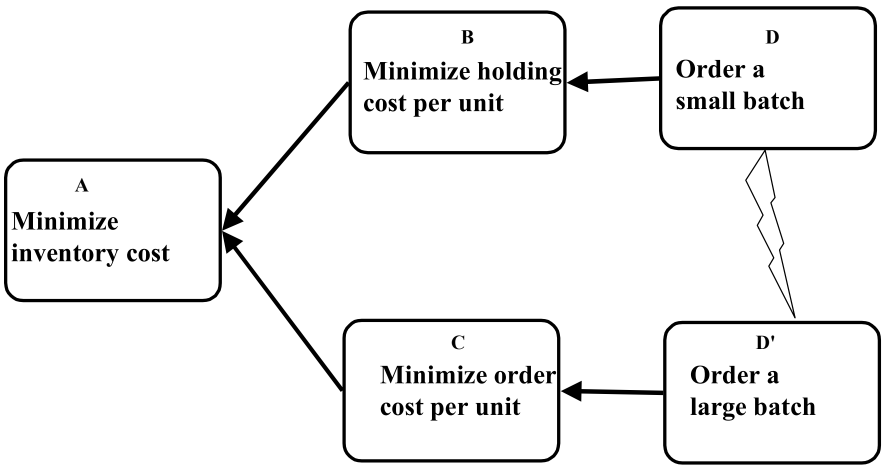
Box	***Assumptions***	***Injections***
A	Minimising total costs is desirable.	Minimising total costs is less desirable than say maximising profits. A focus on minimising costs may harm revenues.
Arrow	***Assumptions***	***Injections***
AB	Facility costs are a significant part of unit costs. Inventory holding costs are a significant part of facility costs. We pay for the facilities.	Reduce facility costs – investigate alternative sites, rental vs owned property. Simplify/streamline order processing/storage/picking systems. Apply buffer management concepts to reduce inventory while improving supply chain performance. Get someone else to pay for the facilities, so they don't count towards unit costs.
AC	Transport costs are a significant part of total unit costs. We pay for transport from facilities to the customer.	Apply buffer management concepts to reduce total transport costs. Consider total costs not just deliveries but also returns, repairs, replacements… Investigate cheaper delivery options – maybe contract out or bring back in-house. Examine service level options. Could get customers to pay for part of the transport. Eg Get customers to uplift goods from a regional depot rather than offer free delivery.
BD	Each facility incurs costs. More warehouses require greater total inventory for the same service level.	Several small facilities may be cheaper than one large one. All facilities need not be the same: can have small depots or even just transfer locations on motorway lay-bys. Smart holding of stock can reduce stock requirements while maintaining service levels.
CD'	Transport costs are reduced if we have more warehouses. Warehouses can be closer to customers. And we can make use of bulk transport for the main routes, and smaller cheaper transport for the local deliveries.	Transport costs can still be reduced by having depots or sites for transfer, not necessarily warehouses. Review service standards instead, relaxing time to deliver. Make product simpler and supply clear instructions so customers can do installation themselves.
DD'	Can't have both one warehouse and multiple warehouses.	Can have one warehouse and multiple transfer sites to minimise both transport and facility costs simultaneously.

Appendix 1. Reframing Classical Inventory Control



We'd like to have low holding costs and low order costs, but we can't have both simultaneously. Hence we derive the EOQ solution, which 'minimizes' the sum of the two.
But is this really the minimum? Why can't we have low holding costs and low order costs simultaneously?

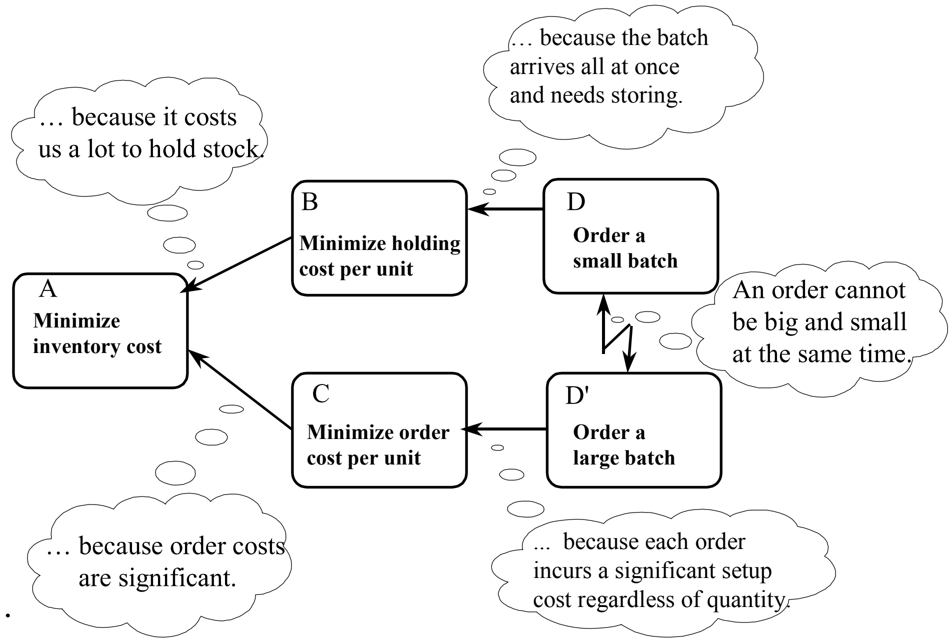
Reframing this as an Evaporating Cloud, we have:



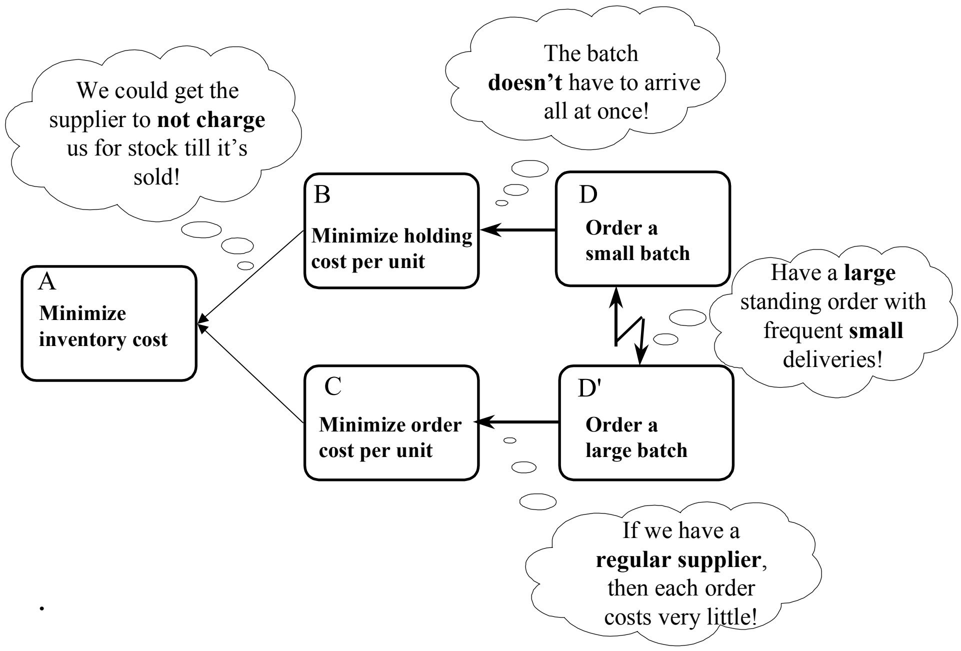
We read this cloud as:
The objective is to Minimise inventory costs. In order to do this, we must Minimise holding cost per unit.
And in order to do this, we must Order a small batch.
On the other hand, in order to Minimise inventory costs, we must Minimise order costs per unit.
And in order to do this, we must Order a large batch. Hence the conflict!
The purpose of the Evaporating Cloud is to resolve or break the conflict, to achieve a win-win solution. So the next question in this frame is, "Are there ways around the conflict?"

We can either break the cloud directly, by looking for ideas (injections) that break the arrows, eg looking for ways to have B without D, or to have C without D'.

Or we surface assumptions and find ways of breaking those.
For example, some assumptions are shown in thought bubbles below:



Which might lead to ideas for solution as follows:



Here we have the idea of a standing order with a regular supplier to minimise order costs coupled with frequent small deliveries to minimise inventory holding costs.

In the production context, an equivalent solution involves setting up the constraint or bottleneck machines to run **large process batches**, while using **small batches** on other processes and for transferring WIP between processes.

Incorporating factors such as Chance and Risk

G A Vignaux
School of Mathematics, Statistics, and Computer Science
Victoria University of Wellington
New Zealand
Tony.Vignaux@mcs.vuw.ac.nz

Abstract

A rambling series of recollections of an old-fashioned OR worker; being a description of personal encounters with decision situations involving chance and risk. We might get round to practical examples of sensitivity analysis, queue theory, simulation, game theory, stock control, investment analysis, decision analysis, dimensional analysis, and genetic algorithms.

Except accidentally, I will not mention any techniques of deterministic OR, classical statistics, or bayesian methods.

A Viability Theory Tutorial: "Satisficing" Solutions to State-Constrained Control Problems Explained

J. B. Krawczyk and K. Kim
School of Economics & Finance
Victoria University of Wellington
New Zealand
Jacek.Krawczyk@vuw.ac.nz
Kunhong.Kim@vuw.ac.nz

Abstract

In some real-life intertemporal decision problems, which include a country's central-bank interest-rate determination, sustainable resource exploitation and aircraft control, optimisation might not be a suitable solution procedure. Routinely, it suggests a unique "optimal" solution yet, the problems might possess many solutions that would be *satisficing*. The latter is Herbert A. Simon's (1978 Economics Nobel Prize laureate) neologism who argued that *satisficing* rather than *optimising* is what economists really need.

It is a solution to a *viability problem* that determines *sets* of initial conditions, for which there exist strategies that maintain the system within a desired state domain. We will use viability theory to show how to compute such sets and how *satisficing* policy rules can be obtained.

1 Introduction

The aim of this paper is to introduce the reader to viability theory and explore its usefulness for the analysis and synthesis of an economic state-constrained control problem.

Herbert A. Simon, 1978 Economics Nobel Prize laureate, talked about *satisficing* (his neologism) rather than *optimising*, as being what the economists really need. We think that economic theory, which follows the Simon prescription, brings modelling closer to how people actually behave. We also think that it is *viability theory*, which is a relatively young area of continuous mathematics (see [1] and [2]), that rigorously captures the essence of *satisficing*. Therefore, viability theory appears to be an appropriate tool of achieving a *satisficing* solution to many economic problems. We aim to demonstrate this by solving a stylised Central Bank macroeconomic problem. The solution will enable us to analyse the system transition trajectory around a steady state (rather than towards the steady state). We believe that an evolutionary analysis enabled by viability analysis gives us a better insight into the system behavour than just an equilibrium analysis.

In this paper, we hope to contribute to the discussion on how a Central Bank can deal with the inflation stabilisation problem. Notwithstanding this macroeconomic application the viability theory can be applied to other problems where uniqueness of the optimal strategy is not of major concern.

In the next section, we provide a brief introduction to viability theory and, in Section 3, we apply it to a simple macroeconomic model[1]. The paper ends with concluding remarks.

Interested? Download the rest of the paper from
http://www.vuw.ac.nz/staff/jacek_krawczyk/somepapers/viab_ORSNZ.pdf .

[1]For a viability theory application to environmental economics see [3]; also, see [12] for a viability analysis of an endogenous business cycle.

Solving the Bidder Problem in a Combinatorial Auction

Natalia Santamaría
RUTCOR, Rutgers University
Piscattaway, New Jersey USA
nsantam@gmail.com

Fernando Beltrán
ISOM Department
The University of Auckland Business School
Auckland, New Zealand
f.beltran@auckland.ac.nz

Abstract

We introduce a model for the bidding problem in a single-round combinatorial auction. In a risk neutral setting, we find analytical expressions for the optimal bidding strategies. Using an alternative definition to reflect the value to the bidder of a bundle of objects, we focus on the two-object case; the alternative approach allows us to gain an insight into the optimal bidding functions when the objects are either substitutes or complements. The optimal bid is a function of the number of competitors (bidders in the auction) and the level of synergy (either substitution or complementarities) the bidder believes her competitors have for the objects.

Keywords: Auctions; Combinatorial Auctions; Bidding Strategy

1 Introduction

Research on combinatorial auctions got a renewed interest due to the intention of the United States Federal Communication Commission to allow package bidding in its ascending auction to allocate a set of spectrum licenses. Research has been mostly focused on studying the winner determination problem ([2], [9], [3]). [7], [8], [9] discuss some rules to reduce the sets available for bidding in a combinatorial auction in order to reduce the complexity of such problem. On the other hand, [1] introduces a an ascending combinatorial auction guaranteeing that the best strategy for a bidder is to report her true valuation so she does not need to solve any problem.

Since a combinatorial auction allows a bidder to express her preferences for the objects through package bidding, the problem that a bidder has to solve in order to determine the bids that maximize her expected utility is more complex than that she faces in other multiple-object auctions, namely, the multiple-round simultaneous auction.

The preferences a bidder has for a set of objects are expressed by the relationship among the objects. For example, in a set with two objects, if winning one object makes a bidder willing to pay more for the other than she would have, had the object been auctioned separately, then the two objects are complements; on the other hand, if getting one object makes her pay less for the second, then the objects are substitutes [6]. Another equivalent definition relates to the marginal value of the objects; given are two sets *A* and *B* not containing an object *x*, with *A* a subset of *B*, which in turn is a subset of the set of objects *M* being considered. If the marginal value of winning $\{B \cup x\}$ is smaller than the marginal value of winning $\{A \cup x\}$, for all *A*, *B* and *x*, then the objects in *T* are substitutes; the objects are complements if the reverse inequality holds [4].

In a single-round "pay your bid" combinatorial auction bidders privately send their bids to the auctioneer. After receiving all the bids, the auctioneer proceeds to allocate the objects in a way that maximizes her revenue. The problem that the auctioneer has to solve is known as the *winner determination problem* and is NP-complete [3].

In this paper we present an approach to the problem that a bidder has to solve in a single-round combinatorial auction; first, we define a representation for the bidder's preferences; then we formally introduce the bidder problem and develop the case for two objects, reporting some results. We show (in a single-round auction with two objects) that the best strategy for a bidder is to offer less than her valuation depending on the number of competitors she is facing, her beliefs about them and her valuation for the objects.

2 Bidder preferences

The value that bidder *i* has for a subset *S* of the set of objects M is given by:

$$v_i(S) = \left[\sum_{u \in S} v_i(S) \right] \cdot k_i \tag{1}$$

- $v_i(S)$: bidder's valuation for the bundle *S*.
- If $k_i \in [0,1]$ then the objects in *S* are substitutes for bidder *i*.
- If $k_i \in (1, \infty)$ then the objects in *S* are complements for bidder *i*.

Our representation of the bidder's preferences is an alternative definition to those presented in [6] and [4]. In contrast with the more traditional use of sub(super)additivity for representing substitution or complementarities, introducing k as a parameter allows us to express how large (or small) is the bidder's gains of bundling when compared to acquiring items in S individually.

We can now state the problem a bidder has to solve:

$$\max \Pi_i(b_i) = \sum_{S \subseteq M} (v_i(S) - b_i(S)) \cdot \mathrm{Prob_i}(S) \tag{2}$$

Where *M* is the set of objects to be auctioned, $\Pi(b_i)$ is *i*'s expected utility when she offers $b_i(\cdot)$, $v_i(S)$ is *i*'s valuation for the set *S*, $b_i(S)$ is *i*'s bid for *S*, and $Prob_i(S)$ is the probability that *i* wins the set *S*.

Now, let O be a subset of M such as $|O|=1$; T be a subset of M, such as $|T|\geq 2$; P_T be the family of partitions of T, and P be a member of P_T, with $P=\{P_1,...,P_w\}$. Then T is allocated to i if the following conditions hold:

$$b_i(T) \geq \sum_{P_l \in P} \max_j b_j(P_l), \qquad P \in \mathrm{P}_T \tag{3}$$

A subset O will be allocated to i if the following conditions hold:

$$\begin{aligned} & b_i(O) \geq \max_{j \neq i} b_j(O) \\ & \max_j b_j(T) \leq b_i(O) + \sum_{P_l \in P, P_l \neq \phi} \max b_i(P_l), \qquad T \supset O, P \in \mathrm{P}_\mathrm{T} \end{aligned} \tag{4}$$

The latter implies that i's problem as stated in (2), for any set T, becomes

$$\max_i \Pi_i(T) = ((v_i(T) - b_i(T)) \cdot \prod_{P \in \mathrm{P}_T} \mathrm{Prob}\left[b_i(T) \geq \sum_{P_l \in P} \max_j b_j(P_l) \right] \tag{5}$$

For a set O, (2) becomes:

$$\begin{aligned} & \max_i \Pi_i(O) = ((v_i(O) - b_i(O)) \cdot \mathrm{Prob}\left[b_i(O) \geq \max_{j \neq i} b_j(O) \right] \cdot \\ & \prod_{T \supset O} \prod_{P \in \mathrm{P}_T} \mathrm{Prob}\left[\max_j b_j(T) \leq b_i(O) + \sum_{P_l \in P, P_l \neq \phi} \max b_i(P_l) \right] \end{aligned} \tag{6}$$

Furthermore, if a bidder is interested in computing the bids for all the subsets in the set M, she must solve 2^m-1-m problems (m=|M|) just as stated in (5) if she is to find the bids for subsets with two or more objects, and m problems as in (6) to find the bids for the individual objects. The number of partitions associated to a set T is known as the *Bell number* $B_{|T|}$. The first Bell numbers for |T|={1,2...,10} are {1, 2, 5, 15, 52, 203, 877, 4140, 21147, 115975}.

3 The two-object case

When two items are to be auctioned in a single combinatorial auction, finding the optimal bidding function with risk-neutral bidders demands the description of the kind of synergies among the items from the bidder's consideration. In this section we study the two-object case as a first step towards understanding what can be learned about optimal bidding strategies when there is an *ex-ante* assessment of the magnitude of substitution and/or complementarities between the two items.

Let us assume that the auctioneer will auction a set M = {a,b} of two objects to n bidders; we also assume bidders are risk neutral and their valuations are independent, identically distributed random variables. The common distribution is uniform on the interval [0,1] and valuations are bidders' private information. If we denote by $\mathrm{Prob}_i(a)$ the probability that i wins a, with similar interpretations for $\mathrm{Prob}_i(b)$ and $\mathrm{Prob}_i(ab)$, then the problems bidder i has to solve in order to determine her bids are:

$$\max \Pi(a) = (v_i(a) - b_i(a)) \cdot \text{Prob}_i(a) \tag{7}$$

$$\max \Pi(b) = (v_i(b) - b_i(b)) \cdot \text{Prob}_i(b) \tag{8}$$

$$\max \Pi(ab) = (v_i(ab) - b_i(ab)) \cdot \text{Prob}_i(ab) \tag{9}$$

3.1 Bidding function for {*a*} and {*b*}

Object *a* is given to *i*, if *i*'s bid is the maximum bid for *a*, and the auctioneer found it best to assign *a*, and only *a*, to *i*; that is

$$b_i(a) \geq \max_{j \neq i} b_j(a) \tag{10}$$

and,

$$\max_j b_j(ab) \leq b_i(a) + \max_j b_j(b) \tag{11}$$

Condition (11) is equivalent to:

$$\begin{aligned} b_i(a) &\geq \max_j b_j(ab) - \max_j b_j(b) \\ b_i(a) &\geq \max_j (b_j(a) + b_j(b)) \cdot k_j - \max_j b_j(b) \end{aligned} \tag{12}$$

The probability that *i* wins *a* is the probability of the event in (10) by the probability of the event in (11). Unfortunately, the value of k_j is *j*'s private information, so *i* does not know k_j. However if $k_i \in [0,1]$ the objects are substitutes to bidder *j* and if $k_i \in (1, \infty)$ the objects are complements to *j*. The latter indicates that conditioned on the value range for k_j, it is known that the items are substitutes or complements, even when the specific value of k_j is not known.

We will also assume that *i* believes that for a percentage p_c of the bidders the objects are complements so, consequently, for the rest of the bidders, (1- p_c), the objects are substitutes.

Let us define $\bar{k}$ as the maximum of all k_j. In addition, we model k_j as a uniform random variable on the interval [0,1] if the objects are substitutes, and, the value of k_j is a uniform random variable on [1, $\bar{k}$] if the objects are complements. Bidder *i* estimates the expected value for k_j as follows:

$$E[k_j] = E[k_j \mid \text{complements}] \cdot \text{Prob}[\text{complements}] + E[k_j \mid \text{substitutes}] \cdot \text{Prob}[\text{substitutes}]$$

$$E[k_j] = \frac{\bar{k}+1}{2} \cdot p_c + \frac{1}{2}(1 - p_c) \equiv k \tag{13}$$

"Complements" and "Substitutes" in the expectation and probability expressions refer to the events that all items being considered are either complements or substitutes.

Using the value of *k* (notice this value is different from k_i which models the individual preferences of bidder *i*) found in (13) as a proxy value for k_j, condition (12) can be rewritten as shown:

$$b_i(a) \geq \max_j b_j(a) \cdot k - \max_j b_j(b) \cdot (1-k)$$
$$b_i(a) \geq \max_j b_j(a) \cdot k + \max_j b_j(b)) \cdot (k-1) \tag{14}$$

Because the bids $b_j(a)$ and $b_j(b)$ are bounded by the valuations, the right side of the inequalities (14) is a random variable with density function:

$$f = \begin{cases} \dfrac{a}{k(k-1)} & \text{if } 0 \leq a \leq k-1 \\ \dfrac{1}{k} & \text{if } k-1 < a \leq k \\ \dfrac{2k-1}{k(k-1)} - \dfrac{a}{k(k-1)} & \text{if } k < a \leq 2k-1 \end{cases} \tag{15}$$

Notice that the offer made by any bidder depends on *k*, that is, on the bidder's beliefs about her competitors. In order to solve problems (7) and (8), we must consider three cases. In the first two cases considered ahead, bidder *i* believes the objects are complements for most of her competitors, so the value of *k* is greater than 1; in the third case, *i* believes the objects are substitutes for most of her competitors.

3.1.1 Case 1: 1<*k*<2

Because the maximum valuation for any object is 1 and the minimum is 0, the support of (15) is [*k*-1, *k*]. Therefore, *i*'s expected utility is:

$$\Pi_i(a) = [v_i(a) - b_i(a)] \cdot \left[\int_0^{b_i(a)} dx\right]^{n-1} \cdot \left[\frac{k-1}{k} + \int_{k-1}^{b_i(a)} \frac{1}{k}\, dx\right]^{n(n-1)}$$
$$\Pi_i(a) = [v_i(a) - b_i(a)] \cdot [b_i(a)]^{n-1} \cdot \left[\frac{b_i(a)}{k}\right]^{n(n-1)} \tag{16}$$

Solving (16), omitting the details, yields

$$b_i(a) = v_i(a)\left[\frac{n^2-1}{n^2}\right] \tag{17}$$

According to (17), bid for *a* is less than the bidder's valuation for the object and it depends only on the number of bidders in the auction; if the number of bidders increases the bid also increases and approaches the valuation. Solving bids for *b* is similar.

3.1.2 Case 2: *k*>2

The support of the density function in (15) is [0, *k*-1]. In this case, *i*'s expected utility is:

$$\Pi_i(a) = [v_i(a) - b_i(a)] \cdot [b_i(a)]^{n-1} \cdot \left[\int_0^{b_i(a)} \frac{x}{k(k-1)}\, dx\right]^{n(n-1)}$$
$$\Pi_i(a) = [v_i(a) - b_i(a)] \cdot [b_i(a)]^{n-1} \cdot \left[\frac{b_i(a)^2}{2k(k-1)}\right]^{n(n-1)} \tag{18}$$

Then, solving (18) yields:

$$b_i(a) = v_i(a)\left[\frac{n-1+2n(n-1)}{n+2n(n-1)}\right] \quad (19)$$

Again, the bid for the object in this case is at most the bidder's valuation and the difference vanishes as the number of bidders increases. Solving bids for *b* is similar.

3.1.3 Case 3: 0<k<1

In this case the condition (11) always holds because the value of *k* is on the interval [0,1] so the bid for {*a*,*b*} is always less than the sum of any bid for *a* and any bid for *b*. Probability of (11) equals 1 and the expected utility is the following:

$$\Pi_i(a) = (v_i(a) - b_i(a)) \cdot [b_i(a)]^{n-1} \quad (20)$$

Solving problem (20) the optimal bidding function is:

$$b_i(a) = v_i(a)\left[\frac{n-1}{n}\right] \quad (21)$$

(21) is the optimal bidding function in a (symmetric) first-price auction. This result suggests that if the objects are substitutes for the majority of bidders bidding is equivalent to auctioning the objects separately. This is also suggested by the results of the experiments reported in [5].

3.2 Bid function for {a,b}

If *a* and *b* were awarded to *i*, *i*'s bid for $b_j(ab)$ would be the maximum of all bids for {*a*,*b*} and the auctioneer would find it best to assign the two objects altogether. These conditions can be expressed as follows:

$$b_i(ab) \geq \max_{j\neq i} b_j(ab) \quad (22)$$

$$b_i(ab) \geq \max_j b_j(a) + \max_j b_j(b) \quad (23)$$

The density function of $max_{j\neq i} b_j(ab)$ conditioned by (22) is:

$$f_1 = \begin{cases} \dfrac{a}{k^2} & \text{if } 0 \leq a \leq k \\ \dfrac{2k-a}{k^2} & \text{if } k < a \leq 2k \end{cases}$$

Similarly, the density function of $max_j b_j(a) + max_j b_j(b)$ conditioned by (23) is

$$f_2 = \begin{cases} a & \text{if } 0 \leq a \leq 1 \\ 2-a & \text{if } 1 < a \leq 2 \end{cases}$$

The probability of winning a and b depends on the value of $v_i(ab)$ as follows:

$$\left.\begin{array}{l} 0 \le v_i(ab) \le k \equiv 0 \le [v_i(a)+v_i(b)]\cdot k \le k \\ k < v_i(ab) \le 2k \equiv k \le [v_i(a)+v_i(b)]\cdot k \le 2k \end{array}\right\} \quad \text{for } f_1$$

and,

$$\left.\begin{array}{l} 0 \le v_i(ab) \le 1 \equiv 0 \le v_i(a)+v_i(b) \le 1 \\ 1 < v_i(ab) \le 2 \equiv 1 \le v_i(a)+v_i(b) \le 2 \end{array}\right\} \quad \text{for } f_1$$

f_1 and f_2 have the same support; thus, the probability that $b_i(ab)$ wins, depends on the value of $v_i(a)+v_i(b)$. We then consider two cases.

3.2.1 Case 1: $0<v_i(a)+v_i(b)<1$

In this case the expected utility is:

$$\Pi_i(a) = (v_i(a)-b_i(a))\cdot\left[\int_0^{b_i(ab)}\frac{x}{k^2}dx\right]^{n-1}\cdot\left[\int_0^{b_i(ab)}x\,dx\right]^{n^2}$$
$$\Pi_i(a) = (v_i(a)-b_i(a))\cdot\left[\frac{b_i(ab)}{2k^2}\right]^{n-1}\cdot\left[\frac{b_i(ab)^2}{2}\right]^{n^2} \tag{24}$$

Solving (24) we found the following bid function:

$$b_i(ab) = (v_i(a)+v_i(b))\cdot k_i\cdot\left[\frac{2(n^2+n-1)}{2(n^2+n-1)+1}\right] \tag{25}$$

Again we can see that i's bid for $\{a,b\}$ is less than her valuation for the set. This difference diminishes as the number of bidders increases.

3.2.2 Case 2: $1<v_i(a)+v_i(b)<2$

In this case the expected utility is:

$$\Pi_i(a) = (v_i(a)-b_i(a))\cdot\left[\frac{1}{2}+\int_{2k}^{b_i(ab)}(\frac{2k}{k}-\frac{x}{k^2})dx\right]^{n-1}\cdot\left[\frac{1}{2}+\int_1^{b_i(ab)}(2-x)\,dx\right]^{n^2}$$
$$\Pi_i(a) = (v_i(a)-b_i(a))\cdot\left[\frac{4kb_i(ab)-3k^2-b_i(ab)^2}{2k^2}\right]^{n-1}\cdot\left[\frac{4b_i(ab)-2-b_i(ab)^2}{2}\right]^{n^2} \tag{26}$$

Bid $b_i(ab)$, is the minimum root of the following polynomial; this value must also be less than or equal to the valuation:

$$b_i(ab)^5(1-2n)+b_i(ab)^4 2v_i(ab)(n-1)-b_i(ab)(2n^2k^2)+2v_i(ab)n^2k^2 = 0 \tag{27}$$

4 Results

In this section we present some numerical results for different values of *k* and *n*.

4.1 Bid function if $0<v_i(a)+v_i(b)<1$

As was determined in the previous section, the optimal bid depends on the value of the sum of the *one-item sets* valuations. Assuming that *i*'s valuations are $v_i(a)=0.5$, $v_i(b)=0.5$, following we show results for $b_i(a)$, $b_i(b)$ and $b_i(ab)$ when $0\leq v_i(a)+v_i(b)\leq 1$.

Table 1 shows results for $b_i(a)$ and $b_i(b)$ with n =10; there is a range of values for the percentage of bidders for whom the objects are complements from 0% to 100%. This percentage changes the value of *k* (the bidder's beliefs about her competitors). Notice that if the percentage of participants to whom the objects are complements increases the bid made by *i* also increases.

k	p_c **(%)**	$b_i( )$
0,5	0	0,450
1,1	30	0,495
1,5	50	0,495
1,9	70	0,495
2,5	100	0,497

Table 1. Bidding values as k varies; n=10

Table 2 shows from considering different values of *k* and two groups of participants, one with 2 bidders and the other with 100 bidders. Notice that bids with 2 bidders are less than bids when there are 100 bidders, with the bid being exactly the valuation when $k\geq 1.5$. Finally, notice that the value of the bids with a fixed number of bidders increases if the value of *k* increases.

n	***k***	$b_i(.)$
2	0,5	0,250
100	0,5	0,495
2	1,5	0,375
100	1,5	0,500
2	2	0,417
100	2	0,500

Table 2. Bidding values as n varies

Finally, the bid for the set $b_i(ab)$, according to (25) and (27), depends on the number of bidders and on the value of k_i. The results for different values of k_i with 2 bidders and 100 bidders are shown in Table 3.

4.2 Bid function if $1<v_i(a)+v_i(b)<2$

Results for $b_i(a)$, $b_i(b)$ and $b_i(ab)$ when $1<v_i(a)+v_i(b)\leq 2$., are shown in the following tables.

k_i	$b_i(ab),n=2$	$b_i(ab),n=100$
0,5	0,455	0,5
1	0,909	1,0
1,5	1,364	1,5
2	1,818	2,0
5	4,545	5,0

Table 3. $b_i(ab)$

When k varies, we obtain the results shown in Table 4 for $b_i(a)$ and $b_i(b)$, assuming $v_i(a)=0.75$, $v_i(b)=0.75$ and the number of bidders is 10. In Table 5 results are shown for $n=2$ and $n=100$.

k	p_c (%)	$b_i()$
0,5	0	0,675
1,1	30	0,743
1,5	50	0,743
1,9	70	0,743
2,5	100	0,746

Table 4. Bidding values as k varies; n=10

In this case the bid function is the polynomial (27). We used *Mathematica 3.0 for Students* to find the bidding values. In all the cases the results had only one real root and the value was less than the valuation. These results are shown in Table 6 ($v_i(ab)=[0.75+0.75]k_i$).

n	k	$b_i(.)$
2	0,5	0,375
100	0,5	0,743
2	1,5	0,563
100	1,5	0,750
2	2	0,625
100	2	0,750

Table 5. Bidding values as n and k vary

According to the results in Table 6, the value of a bid increases when the values of k and k_i increase.

	$b_i(ab)$			
k	$k_i = 0,5$	$k_i = 1$	$k_i = 1,5$	$k_i = 2$
0,5	0,745	1,45	2,14	2,848

1,1	0,749	1,478	2,174	2,867
1,5	0,749	1,486	2,191	2,882
1,9	0,749	1,49	2,204	2,897
2,5	0,749	1,494	2,218	2,917

Table 6. $b_i(ab)$

5 Conclusions

In a combinatorial auction with two objects and risk-neutral bidders, we found that the optimal bids for a set are less than or equal to the real valuation a bidder has for such set. Nevertheless, that difference diminishes when the number of participants increases.

Any participant's bid for a single object (in the two-object case) depends both, on the level of synergy she believes her competitors have for the objects and on the number of competitors she is facing.

Any participant's bid for a set of two objects depends on the number of competitors she is facing, the synergy level she believes they have for the objects and her valuation for the objects.

If the auction mechanism is different to that proposed in [1], there is a need for developing some alternative methods to find the bid function when there are *m* objects and *n* bidders in the auction.

6 References

[1] Ausubel L. and P. Milgrom. 2002. "Ascending auctions with package bidding" *Frontiers of Theoretical Economics*, 1(1):1–42.

[2] CRA and MarketDesign. 1997. "Report 1b: Package bidding for spectrum licenses". Federal Communications Commission (FCC).

[3] de Vries, S. and R. Vohra. 2003. "Combinatorial auctions: A survey". *INFORMS Journal on Computing*, 15(3).

[4] Krishna, V. 2002. *Auction Theory*. Academic Press, First edition,

[5] Ledyard, J., D. Porter, and A. Rangel. 1997. "Experiments testing multiobject allocation mechanisms". *Journal of Economics & Management Strategy*, 6(3).

[6] Milgrom, P. 2004. *Putting Auction Theory to Work*. Cambridge University Press, First Edition.

[7] Nisan, N. 2000. "Bidding and allocation in combinatorial auctions". Working Paper. Hebrew University.

[8] Pekec, A. and M. Rothkopf. 2003. "Combinatorial auction design". *Management Science*. 49(11):1485-1503.

[9] Rothkopf, M., A. Pekec, and R. Hastard. 1998. "Computationally manageable combinational auctions". *Management Science*, 44(8).