Reservoir management with risk aversion

Andrew Kerr and Grant Read
Department of Management
University of Canterbury
New Zealand
a.kerr@mang.canterbury.ac.nz

John Kaye
School Of Electrical Engineering
University of New South Wales
Australia
j.kaye@unsw.edu.au

Abstract

We consider the problem faced by a manager planning the operation of a mixed hydro/thermal system, where the manager controls the reservoir release made in each week (single reservoir), as well as the generation from other sources. Demand is deterministic and must be met in each period, while the inflows experienced in each week are uncertain. Stochastic Dynamic Programming (SDP) is a technique often applied to reservoir management problems, with the scheduling horizon divided naturally into discrete time periods, storage as the state variable, and release as the decision variable. If the objective is to minimise the expected annual cost, then a large number of observations with low costs can cancel out a few observations with large costs (because all outcomes are weighted equally). In reality, the manager might want to be able to trade-off a reduction in extremely good outcomes for an improvement in extremely bad outcomes ie, putting more 'weight' on the bad outcomes, but this invalidates the standard dynamic programming recursive relationship. We describe a SDP formulation which accommodates a non-linear end-of-horizon utility function by augmenting the state space. We describe some simple algorithmic modifications which significantly reduce the computational requirements of the optimisation, and illustrate the impact of risk averse attitudes on system performance.

1 Introduction

Reservoir management is concerned with the planning of reservoir releases, and the resulting hydro generation. It is an interesting and complex problem because water is storable commodity, so there is a continuous process of deciding whether to release it now, or to store it and release it at a later date, where the time frame for these decisions can range from minutes to months.

In New Zealand, for example, approximately 70% of the annual national energy demand is supplied from hydro sources, so detailed planning of reservoir operation is crucial because of uncertainty about the level of natural inflows, the fact that the aggregate storage capacity is only 6 weeks (approximately) of national demand, and because thermal generation is relatively expensive.

Reservoir planning models (e.g. Read [4]) typically minimise expected thermal and shortage costs, though they do not account for the variability of these costs, so the impact of a large cost, which may be a major concern of the decision maker (DM), can be cancelled out by a large number of observations with low costs. Although heuristic procedures are often employed to modify release schedules so as to avoid excessive costs, there has been no analytical technique which can trade-off storage and cost in a manner consistent with a DM's attitude towards these costs.

Stochastic Dynamic Programming (SDP) is a technique often applied to reservoir management problems (e.g. Archibald et al [1]) where the scheduling horizon divides naturally into discrete time periods (the stages of the SDP) and storage is the state variable which links the stages. Ranatunga [3] developed an approach termed Risk Averse Stochastic Dynamic Programming (RASDP) that incorporates a DM's end-of-horizon utility into decisions made over time. He applied RASDP to the purchase and sale of forward contracts, so a risk averse DM owning thermal plant with significant intertemporal risks could hedge against price uncertainty.

In this paper we describe an implementation of RASDP to a medium-term reservoir management problem (§2). Some algorithmic modifications are detailed in §3, and in §4 we briefly illustrate the impact of the DM's risk aversion on storage and end-of-horizon performance.

2 Modelling approach

We consider a scenario where a DM owns all the generation sources, and is trying to plan generation and reservoir release (for a single reservoir) over a one year planning horizon (T) divided into weeks ($t \in T$). In each week the benefit depends on the level of release, the storage level, and the level of demand. Weekly inflows are uncertain, and the volume of water able to be stored in the reservoir, and the volume of release in a period, are bounded. The objective is to maximise the DM's expected utility of end-of-horizon net wealth (NW) which has 2 components:

- end-of-horizon accumulated wealth, w_{T+1}, which is defined as the sum of the weekly benefits from release. The weekly benefit function, $B(q_t, d_t)$, is a function of release (q_t) and demand (d_t), and is defined as the negative of the cost of satisfying demand, so $B(q_t, d_t) \leq 0 \ \ \forall t$ and $w_t \leq 0 \ \ \forall t$.

- value of end-of-horizon storage, $V(s_{T+1})$, which reflects the value of holding water at the end of the planning horizon.

If the DM is risk averse, say, she will wish to adopt a strategy which reduces the chance of achieving extremely bad NW outcomes, and this may be at the expense of reducing the chance of achieving a very good wealth/storage outcomes. The degree to which she wishes to trade off these outcomes can be reflected in a utility curve (U), which will be concave and non-linear if she is risk averse. However, if U is non-linear and B_1 and B_2 are the benefits achieved in weeks 1 and 2, then

$$\mathrm{E}[U(B_1 + B_2)] \neq \mathrm{E}[U(B_1)] + \mathrm{E}[U(B_2)] \tag{1}$$

so the expected utility depends on the benefits achieved in all periods, and cannot be calculated by adding the expected utility of each individual benefit. Therefore, the objective is non-separable, invalidating the use of a recursive relation, and hence dynamic programming.

In order to overcome the problem of non-separability, define another state, w_t, which is the accumulated benefit up to the beginning of period t

$$w_t = \sum_{t=1}^{t-1} B(q_t, d_t) \tag{2}$$

where $w_1 = 0$. End-of-horizon wealth is defined as

$$w_{T+1} = \sum_{t \in T} B(q_t, d_t) \tag{3}$$

where

$$w_{t+1} = w_t + B(q_t, d_t) \tag{4}$$

so we can formulate the problem as a standard SDP problem with a 2-dimensional state space (accumulated wealth and storage).

The end-of-horizon value function, f_{T+1}, is described by calculating the utility of arriving at different points (accumulated wealth and storage levels) in the end-of-horizon state space. The form of utility we use here is

$$f_{T+1}(w_{T+1}, s_{T+1}) = U_W(w_{T+1}) + U_S(V(s_{T+1})) \tag{5}$$

though the method does not require this functional form. The objective to be maximised at each period is

P1

$$f_t(W_t, S_t) = \max_{q_t} \ \mathrm{E}\left[f_{t+1}(W_{t+1}, S_{t+1}, d_t) \mid a_t \right] \tag{6}$$

subject to:

$$w_{t+1} = w_t + B(q_t, d_t) \tag{7}$$

$$s_{t+1} = s_t - q_t + a_t \tag{8}$$

$$s^{\min} \le s_t \le s^{\max} \tag{9}$$

$$q_t^{\min} \le q_t \le q_t^{\max} \tag{10}$$

$$w_1 = 0 \tag{11}$$

The utility for a particular end-of-horizon wealth and storage is calculated in (5), where U_W and U_S are (concave) utility functions for end-of-horizon wealth (w_{T+1}) and value of storage $V(s_{T+1})$. In each week, we choose the release level (q_t) that maximises the expected end of horizon utility (6) given the demand (d_t) and stochastic inflows (a_t) in the period. Equations (7) and (8) describe the state transitions for wealth and storage, where $B(q_t, d_t)$ describes the benefit from release in *t*. Storage and release bounds are defined in (9) and (10), and the initial level of accumulated wealth is 0 (11).

The use of the additional state variable w_t has enabled the separation of the time periods and therefore the ability to apply a dynamic programming principle. The approach described here for solving **P1** is to discretise the state space, so f_{T+1} is defined for discrete storage and wealth values over (W_{T+1}, S_{T+1}). Because the bounds on storage remain the same throughout the optimisation, the storage grid can remain static. This is not the case for accumulated wealth, w_t, which accrues throughout the scheduling horizon. Maximum and minimum values diverge with *t*, so the discretisation grid is different in each period (the number of grid points is kept constant). A discrete range of release decisions is sampled from the continuous set of feasible releases and these remain the same throughout the optimisation (although this is not a fundamental to the technique). Inflows, which are not correlated, are sampled at discrete levels from a pre-defined inflow distribution. Release decisions are made at the beginning of *t* and we assume here that the inflow occurs at the end of *t* in its entirety, so there is no ability to adjust release during the period to account for the level of actual inflow.

The expected end-of-horizon utility for a particular (w_t^m, s_t^n) grid point and release, q_t^k, is calculated as

$$\bar{f}_t^{m,n,k} = g_t(\bar{w}_{t+1}, \bar{s}_{t+1}) \qquad \forall k \qquad (12)$$

where $\bar{w}_{t+1} = w_t^m + B(q_t^k, d_t)$, $\bar{s}_{t+1} = s_t^n - q_t^k$, $m \in M$ (wealth discretisation index), $n \in N$ (storage discretisation index), $k \in K$ (release discretisation index), and g_t is f_{t+1} adjusted for inflow uncertainty. The optimal release, $\hat{q}_t^{m,n}$, is that which corresponds to

$$f_t^{m,n} = \max_k \left[\bar{f}_t^{m,n,k}\right] \qquad \forall k \qquad (13)$$

Starting from the end of the horizon and working backwards in time, we find and store the optimal expected end-of-horizon utility value $\left(f_t^{m,n}\right)$ and the associated optimal release $\left(\hat{q}_t^{m,n}\right)$ for each grid point in $W_t \times S_t$.

3 Algorithmic issues

We calculate $B(q_t^k, d_t)$ for each release (q_t^k) by integrating the residual demand curve over the range $(q_t^k, q_t^{\max})$, which the cost of generation supplied from other generation sources. Assuming the other stations are loaded in order of their marginal cost and that demand is fixed, $B(q_t^k, d_t)$ is a concave, increasing monotonically as q_t increases. The wealth state transition is defined as $\bar{w}_{t+1} = w_t + B(q_t, d_t)$, so $\bar{w}_{t+1}$ has the same properties as $B(q_t, d_t)$ because w_t is constant when evaluating $\bar{w}_{t+1}$ for a particular wealth/storage pair. The relationship between storage and release is assumed to be linear here[1], so the state transition possibility curve (STPC), which is a function of

[1] An extension would be to assume a non-linear efficiency curve. There curves are concave, so would not affect this analysis, because the STPC would only become more concave if an efficiency curve were to be used.

wealth and storage, has the same properties as $B(q_t, d_t)$. A STPC for a particular wealth/storage point is illustrated in Figure 1.

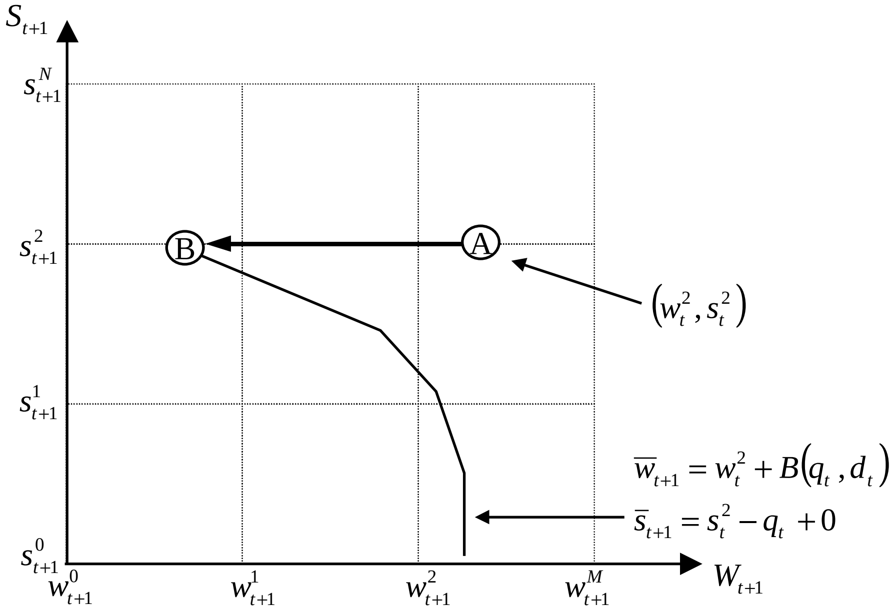


Figure 1: State transition possibility curve

The first point on the state transition possibility curve is point 'B', which corresponds to the first release option, $q^{\min}$, which in this case equals 0MWh. Demand is met by other sources, decreasing $\overline{w}_{t+1}$ to its minimum possible value because the highest possible cost is incurred. As q_t^k increases, $\overline{s}_{t+1}$ decreases because more water is released, and $\overline{w}_{t+1}$ increases because the additional release reduces the quantity of 'other' generation, and hence the total fuel cost. Note that as q_t^k increases, $\overline{w}_{t+1}$ increases at a non-increasing rate because the marginal cost of the generation displaced by the hydro release is decreasing due to the assumption of merit order dispatch.

If U_W, U_S, and V are non-decreasing functions ($U'_W \geq 0$ and $U'_S \geq 0$ where they exist) with non-increasing marginal values ($U''_W \leq 0$ and $U''_S \leq 0$ where they exist), then for $f_{T+1}(w_{T+1}, s_{T+1})$ defined as $U_W(w_{T+1}) + U_S(V(s_{T+1}))$, the contours of g_T (f_{T+1} adjusted for the inflow uncertainty during T) have the form illustrated in Figure 2. High utility is achieved for high levels of storage and wealth, and because U_W and U_S are concave, contours that are equally spaced in terms of their value of utility are spaced further apart as wealth and storage increase because the marginal utility derived from increasing either wealth or storage is decreasing.

Consider one of the contours in Figure 2. Starting from a high wealth/low storage position, if wealth is decreased, utility decreases, so storage must be increased to maintain the same level of utility (i.e., moving up the contour). Because the marginal utility of storage decreases, the increase in storage required to maintain the level of utility increases as we keep reducing wealth, which results in each contour having a steeper slope as wealth decreases and storage increases. Conversely, starting from a high storage/low wealth position on a contour and decreasing storage requires that wealth be increased to maintain the same value of utility, and at an increasing rate. Hence each contour becomes flatter as storage decreases and wealth increases.

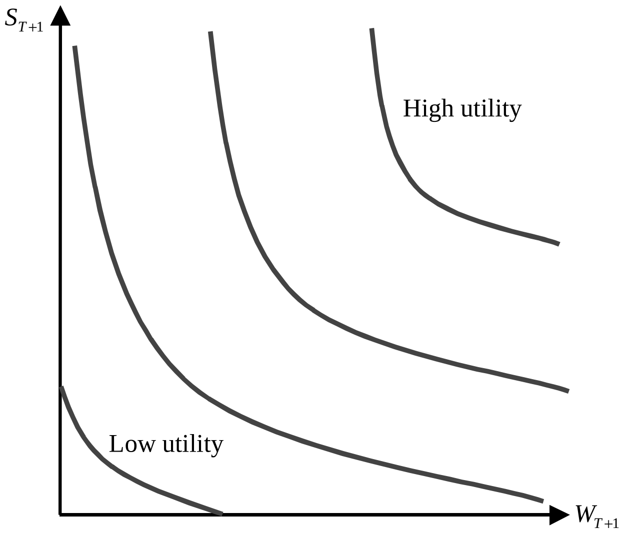


Figure 2: Contours of the utility surface

To determine the optimal release for some $\left(w_T^m, s_T^n\right)$, we map the STPC on to $W_T \times S_T$, then select q^k which maximises value of g_T. This is equivalent to finding the point on the STPC which reaches the highest contour of g_T. Figure 3 shows the STPC for $\left(w_T^m, s_T^n\right)$, indicated by the curve $\left(\overline{w}_{T+1}^m, \overline{s}_{T+1}^n\right)$, which intersects contour c^3 at the (optimal) release level q''. Now consider the STPC from a wealth/storage pair, $\left(w_T^m, s_T^{n+1}\right)$, which has the same level of wealth and a higher storage level than $\left(w_T^m, s_T^n\right)$. The wealth state transition is independent of the level of storage[2], so the mapping of the benefit function in the wealth dimension is the same for $\left(w_T^m, s_T^{n+1}\right)$ as it is for $\left(w_T^m, s_T^{n+1}\right)$ and is shifted upwards in the storage dimension by $s_T^{n+1} - s_T^n$, and the new optimal release level is q''', where $q''' > q''$.

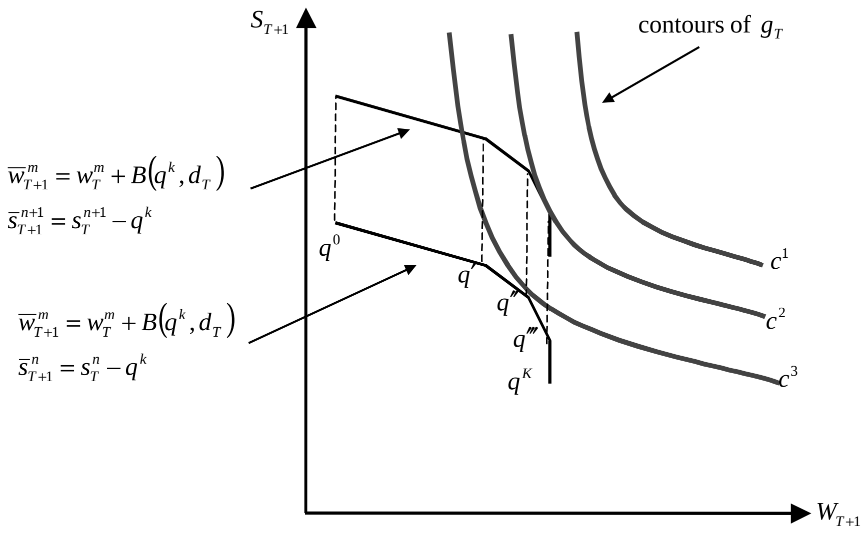


Figure 3: Release ratchet in storage dimension

[2] If head effects were taken into account, this would not be the case.

The slopes of the contours increase as storage increases, but the form of the STPC (and the slopes of the segments) remains the same, so if $\hat{q}_T^{m,n}$ is optimal for $\left(w_T^m, s_T^n\right)$ then for $\left(w_T^m, s_T^{n+1}\right)$

$$\hat{q}_T^{m,n+1} \geq \hat{q}_T^{m,n} \qquad \forall m \qquad (14)$$

where $s_T^{n+1} \geq s_T^n$, which we term the 'release ratchet' in the storage dimension. So for some level of wealth, and given that the marginal utility derived from holding more storage decreases as more water is stored, it will be optimal to release more at a higher storage level than at a lower storage level and move to a position of lower storage and higher wealth. A similar result occurs when shifting the STPC in the wealth dimension and holding storage constant, so if $\hat{q}_T^{m,n}$ is optimal for $\left(w_T^m, s_T^n\right)$, then for $\left(w_T^{m+1}, s_T^n\right)$

$$\hat{q}_T^{m,n+1} \geq \hat{q}_T^{m,n} \qquad \forall n \qquad (15)$$

From these two results it also follows that if $\hat{q}_T^{m+1,n}$ is optimal for $\left(w_T^{m+1}, s_T^n\right)$ and $q_t^{m,n+1}$ is optimal for $\left(w_T^m, s_T^{n+1}\right)$, then for $\left(w_T^{m+1}, s_T^{n+1}\right)$

$$\hat{q}_T^{m+1,n+1} \geq \max\left(\hat{q}_T^{m+1,n}, \hat{q}_T^{m,n+1}\right) \qquad \forall m,n \qquad (16)$$

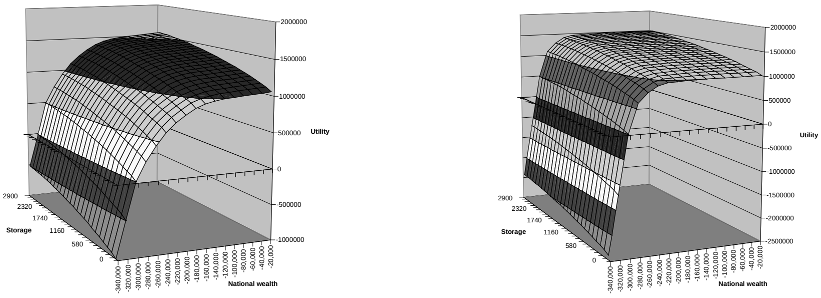
The ability to 'ratchet' the release level as the state space is searched means that the search for the optimal release for some $\left(w_t^m, s_t^n\right)$ can be started from the ratchet release level rather than at $q^{\min}$, significantly reducing the computational requirements for the search.

In addition, if we assume the other stations are loaded in order of their marginal cost, $B(q_t, d_t)$ is a concave function that decreases monotonically as q_t increases, and it is therefore possible to terminate the search for the optimal release when $\bar{f}_t^{m,n,k+1} < \bar{f}_t^{m,n,k}$ because there is a unique utility maximising intersection of f_{T+1} and *B*.

The concavity of the expected utility function is preserved by the form f_{T+1} and *B* (as they are defined here), and is not affected by the uncertainty adjustment, so the modifications described above can be applied for all *t*, resulting in a decrease in the execution time by approximately 99%.

4 Results

To examine the impact of risk aversion, optimisations were performed using four 'risk averse' end-of-horizon utility surfaces (RA1, RA2, RA3, RA4) which increased in curvature (risk aversion) in the wealth dimension; the utility of storage curve remained the same. The 'risk neutral' case (RN), where the utility curve has a constant slope, was also solved and used as a base case against which the risk averse results were compared. Figure 4 illustrates the RA2 and RA4 end-of-horizon utility surfaces.



U_W = RA2 $\qquad$ U_W = RA4

Figure 4: End-of-horizon utility surfaces

The output from the optimisations are optimal release surfaces for each week which describe the optimal release should a wealth/storage pair be arrived at. An example of one of these surfaces is shown in Figure 5. The release is non-decreasing in both the storage and wealth dimensions. For low storage, the release is the same for all wealth values, so the marginal value of storage dominates any wealth/storage trade-off. For higher storage levels, though, release increases as wealth decreases because the marginal value of wealth increases relative to the marginal value of storage.

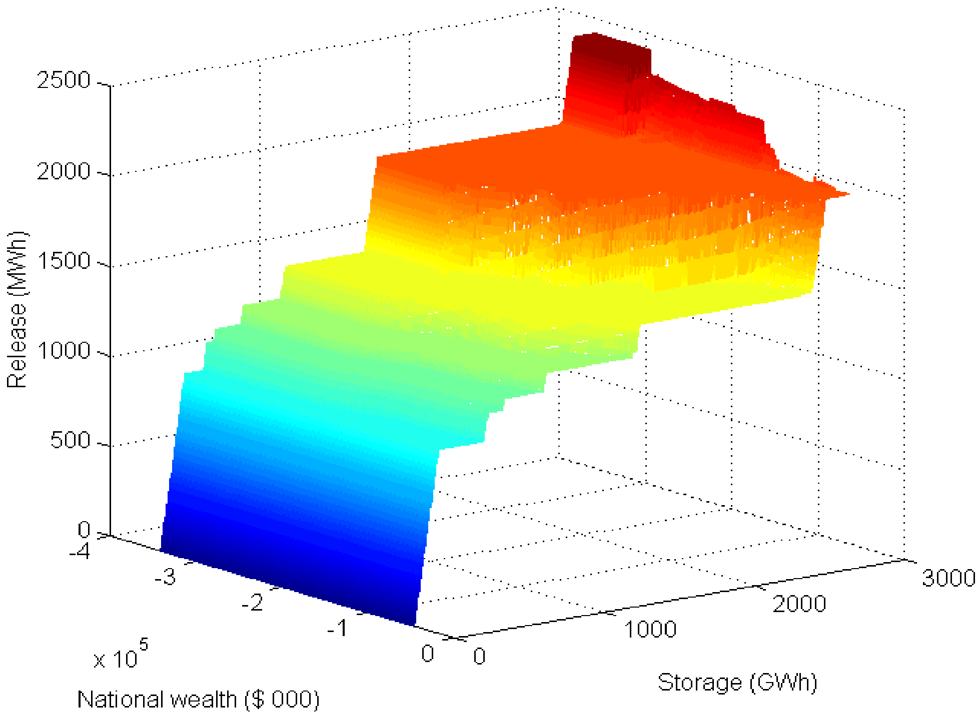


Figure 5: RA4 optimal release surface for *t*=26

Demand is highest over the middle part of the year, so there is the potential for large costs (large decreases in national wealth) to be incurred because the high demand will require the use of more expensive thermal stations. We would expect hydro release to be higher over those periods, then, because low release will result in higher thermal costs. However, as risk aversion increases, the incentives for avoiding low wealth outcomes are ‘stronger’ and the incentives for achieving high wealth outcomes are ‘weaker’, so it will be more ‘acceptable’ to incur a larger cost should the future uncertainty be such that it is desirable to store more water as a mechanism to hedge against that possibility. Figure 6(a) shows the mean simulated hourly release for each week. The impact of risk aversion is to release less, which implies that more water must be held in storage so the mean storage directories are higher then the RN trajectory, as illustrated in Figure 6(b).

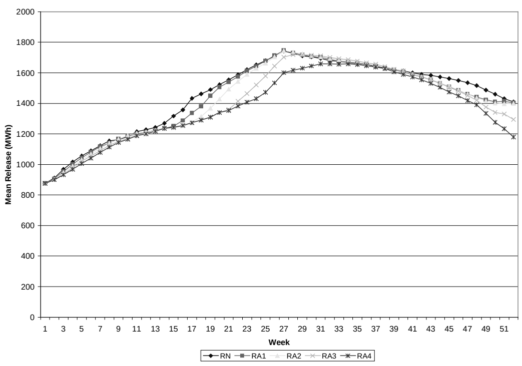


Figure 6(a): Mean release

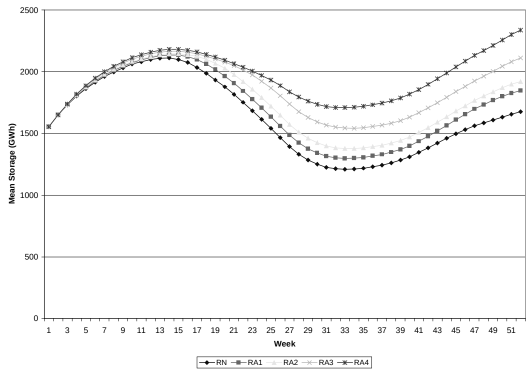


Figure 6(b): Mean storage

The impact of risk aversion on end-of-horizon net wealth (NW) is to decrease the frequency of very high (good) and very low (bad) NW. This is illustrated in Figure 7, which shows the RN and RA4 probability density functions for NW.

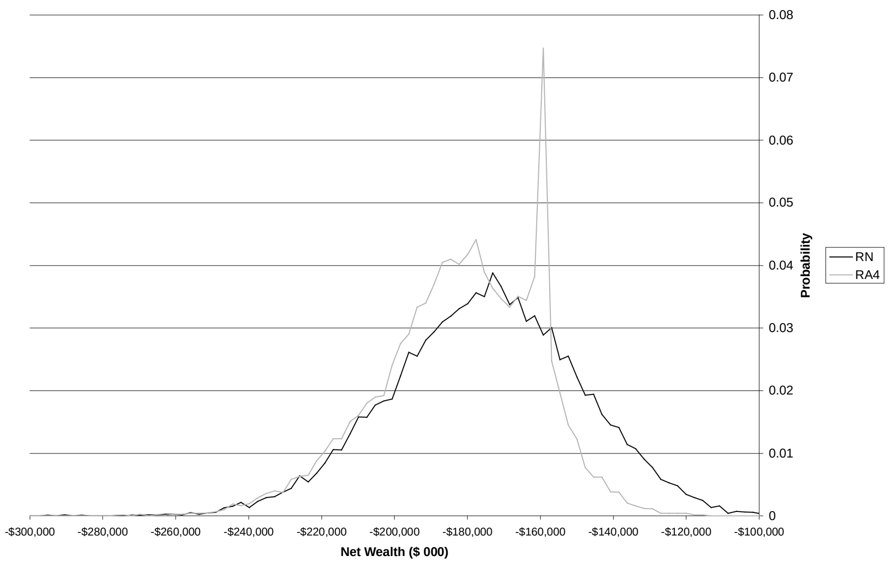


Figure 7: EOH-NW PDFs for RN and RA4

Risk aversion also decreases mean NW and decreases the standard deviation and semi-standard deviation of NW. See [2] for a fuller discussion of these results.

5 Summary

This paper has described a SDP formulation of the medium-term reservoir scheduling problem with inflow uncertainty and a 2-dimensional state-space. By defining the accumulated wealth as a state, we can define a non-linear utility function to reflect the decision maker's attitude towards end-of-horizon wealth and end-of-horizon storage outcomes. For each week, the optimisation determines the optimal release decision for discrete values of national wealth and storage that maximises the expected end-of-horizon utility given stochastic inflows and deterministic demand. Simulations were performed using 5 different utility functions.

For the problem as stated, the different risk averse utility functions had a reasonable impact on the distribution of end-of-horizon net wealth outcomes compared to the risk neutral case. The impact was more pronounced for the ‘good’ outcomes, which had lower probabilities of occurring when the degree of risk aversion was higher. In terms of system operation, the average storage and average release trajectories showed significant changes, with risk aversion resulting in less water being released and hence more water being stored. Future research will explore the impact of risk attitudes on system performance under different system configurations.

Acknowledgements

The authors would like to thank ECNZ for supporting this research.

References

[1] Archibald, T. W., K. U. M. McKinnon, and L. C. Thomas, *An aggregate stochastic dynamic programming model of multireservoir systems*, Water Resources Research, 33(2), p333-240, 1997.

[2] Kerr, A. L., E. G. Read, and R. J. Kaye, *Stochastic dynamic programming applied to medium-term reservoir management: Maximising the utility of a system supply cost minimiser*, EMRG Working Paper EMRG-WP-97-03, Department of Management, University of Canterbury, New Zealand, 1997.

[3] Ranatunga, R. A. S., *Risk averse operation of an electricity plant in an electricity market*, ME dissertation, School of Electrical Engineering, University of New South Wales, 1995.

[4] Read, E. G., *A new variant of stochastic DP for multi-reservoir release scheduling*, Proceedings 21st ORSNZ Conference, 4-7, 1985.