Short-Term Modelling of Electricity Prices in New Zealand

Alan Ansell, Golbon Zakeri and Andreas W. Kempa-Liehr
Department of Engineering Science
University of Auckland

Abstract

Wholesale electricity prices in New Zealand are calculated for each half-hour trading period by solving a complex dispatch problem. These prices are highly volatile and difficult to forecast. This is problematic for consumers who purchase electricity at wholesale rates, as they often cannot react quickly enough to price spikes to avoid incurring significant losses. We used a data science approach to develop a methodology for accurately forecasting electricity prices over the short term.

Electricity consumers have access to the Wholesale Information and Trading System (WITS), a source of real-time aggregate market data. We begin by forming a model for total electricity demand in the two islands. We develop a model which combines the demand predictions produced by this model with the data available through WITS to approximate the behaviour of the full market dispatch problem. We then show how to use our approximate dispatch model to predict prices. Finally we evaluate our price prediction pipeline and find that it achieves a mean absolute error of 15% of the mean price value.

1 Introduction

New Zealand has had a deregulated electricity market since 1996. Generators make offers to supply electricity, while consumers bid for this supply. A complex dispatch problem is solved for each half-hour time period to determine the quantity of electricity to be generated by each generator during that time period, and the “spot” price of electricity for that period.

The vast majority of New Zealand’s residential and commercial consumers purchase their electricity through fixed price contracts with retailers. Large industrial consumers stand to make significant savings in the long run by purchasing on the wholesale market, but this entails considerable risk because these spot prices are highly volatile, a common feature of electricity markets [2]. Many of these consumers would be capable of reducing their consumption given sufficient advance warning of price spikes, but currently lack a reliable means of price forecasting.

We aimed to form a model of NZEM prices which would enable the price of electricity at a given node to be accurately predicted in the short term. We focus on predicting prices 2 hours ahead of time. Such a model would be capable of alerting consumers to upcoming price spikes and enable them to more generally optimise their usage in the short term.

2 Operation of the New Zealand Electricity Market

2.1 Market Structure

The New Zealand Electricity Market (NZEM) is regulated and administered by New Zealand's Electricity Authority (EA). State-owned enterprise Transpower owns the national transmission infrastructure (the "national grid") as a regulated monopoly and is responsible for ensuring security of electricity supply.

The national grid consists of more than 200 *nodes* where electricity is supplied and consumed. These nodes are interconnected by high-voltage alternating current transmission lines, with the exception of the High-Voltage Direct Current (HVDC) link which connects the North and South islands at Haywards and Benmore nodes.

For each trading period, generators submit an *offer stack* specific to the node at which they supply their generated electricity. Similarly, price-responsive consumers submit a *bid stack* specific to their local node. Each stack consists of up to five *tranches*, where a tranche indicates a willingness to generate/consume a certain quantity of electricity at a specified marginal price. Thus each stack takes the form of a step function, as shown in Figure 1.

The objective of the market is to dispatch generators in the cheapest possible way while satisifying the demand at every node. The clearing price and quantity in a hypothetical single-node market would be found at the point where the offer stack and bid stack intersect. In a multi-node market, the need to transmit electricity between nodes adds the following constraints [3]:

- Each transmission line has a capacity which may not be exceeded.
- Resistive losses occur whenever electricity is transmitted. The loss in a line is proportional to the square of the flow in the line.
- The flows in the network must obey Kirchhoff's current and voltage laws.

We will consider a version of the market dispatch problem in which the following simplifications are made:

- Demand is not price-dependent. Few consumers submit price-sensitive bids, so we will treat demand as constant.
- Ancillary generation is ignored. The system operator is required to procure extra "ancillary" generation to ensure security of supply in the event of a failure in the network [5]. This can affect the electricity price, but we do not consider it directly.

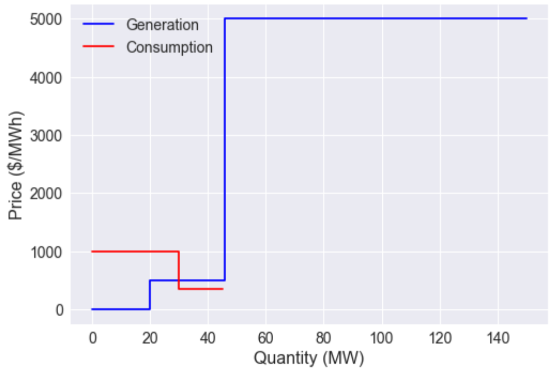

Figure 1: Bid and offer stack at Whirinaki node, 27/06/2017, trading period 33

	Benmore	Haywards	Otahuhu
mean	61.33	65.77	69.23
std	54.47	62.75	68.36
min	0.00	0.00	0.00
25%	42.94	47.48	51.29
50%	55.15	59.46	63.85
75%	73.16	76.80	80.49
max	8016.33	8714.53	9735.81

Table 1: Price summary statistics at reference nodes, 01/01/2013-30/06/2017

The resulting dispatch problem for a trading period can be expressed as a linear programme with the above constraints whose objective is to minimise the total cost of generation [3]. The price at each node is the marginal cost of consumption at that node, which can be found from the optimal solution of the linear programme as the dual of the constraint which forces net inflow to equal demand at that node.

2.2 Overview of Data and Prices

The EA makes available a large amount of data about past bids and offers and other market conditions on its Electricity Market Information (EMI) website.

Due to market regulations, full bid and offer stacks for future and recent past trading periods cannot be made publically available. However some useful information about future periods is available through the EA's Wholesale Information and Trading System (WITS). In particular, WITS makes available the *aggregate* offer and bid stacks for the North and South islands, or in other words the stacks which result from taking the union of all individual nodal offers in each island.

We will focus on the prices at Benmore and Haywards nodes, the endpoints of the HVDC link, and at Otahuhu node near Auckland. Table 1 summarises the price time series ($/MWh) at each of these nodes over the time period 01/01/2013-30/06/2017. The prices have a right-skewed distribution, with the means being greater than the medians, and the maximum values being many times greater than the mean.

The reason for this skew is that the stack which results from aggregating all offers tends to be fairly flat when the quantity generated is small, but the price increases rapidly in the higher regions. These large jumps in price may represent factors such as the high cost of thermal generation relative to renewable generation and the cost of activating idle thermal generation units.

3 Forecasting Demand

WITS makes available the aggregate North and South Island offer stacks, so we focus on forecasting aggregate North and South Island demand. Let d_t^N and d_t^S denote the aggregate North and South Island demand in trading period t. Recall that each trading period is half an hour long. Generators are permitted to alter their offer stacks up until 2 hours prior to the beginning of a trading period, so we will focus on predicting demand 2 hours, or 4 trading periods, ahead of time t. We can use only information available up to the current time, the beginning of trading period $t-4$, so the most recent known demand values are d^N_{t-5} and d^S_{t-5}.

3.1 Benchmarking Demand Prediction

We wish to form a model for demand with the smallest possible mean absolute error (MAE). To get an idea of what constitutes a reasonable level of performance on this problem, we will estimate d_t with the demand 24 hours ago, d_{t-48}. We will call this the Day-Ago Demand Estimator (DADE). Its performance in each island in terms of its MAE as a percentage of the mean demand in that island is shown in Table 2.

3.2 Simple Demand Model

We identified the following as relevant features for forecasting demand in trading period t:

- previous demand values d_{t-5} and d_{t-48},
- The month $M(t)$, day of week $W(t)$ and time of day $P(t)$ of trading period t,
- whether trading period t falls on a national holiday,
- $T(t)$, the temperature during trading period t. Specifically we will use the temperatures recorded in Auckland, Christchurch, Dunedin and Wellington. We do not know these in advance with certainty, but we assume that they can be accurately forecast two hours ahead of time. We have obtained these recorded temperatures from [6], but unfortunately no archived temperature *forecast* data could be found.

We trained two different machine learning models with these features: a linear regression model (implemented using Scikit-learn [7]) and a neural network model with two hidden layers (implemented using Keras [4]). These models obtained the results in Table 2. Note that whenever we train a model, we use blocked temporal cross-validation as recommended in [1], where the folds are individual years.

	DADE	Linear regression	Neural network
North Island	7.9%	4.4%	2.8%
South Island	4.3%	2.7%	1.7%

Table 2: MAE of simple demand estimators (% of mean demand value)

3.3 Normalising Demand

d_{t-5} is not a very good estimator of d_t because there is a pattern of significant fluctuation in demand throughout the day. However d_{t-5} is a newer piece of information than d_{t-48} and tells us something about the demand conditions experienced shortly prior to trading period t. Let us define $\mathbb{T}(t)$ to be the set of trading periods having similar temporal properties to t as follows:

$$\mathbb{T}(t) = \{i : M(i) = M(t), W(i) = W(t), P(i) = P(t), i \in \mathbb{T}^{(data)}\} \tag{1}$$

where

- $M(t)$ is the month during trading period t,
- $W(t)$ is the day of week during trading period t,
- $P(t)$ is the number (1-48) of trading period t in the day,
- $\mathbb{T}^{(data)}$ is the set of all ~60,000 trading periods in our data set.

We can now define the "routine" demand $\bar{d}_t$ in trading period t as:

$$\bar{d}_t = \frac{1}{|\mathbb{T}(t)|} \sum_{i \in \mathbb{T}(t)} d_i. \tag{2}$$

The *demand factor* $\frac{d_{t-5}}{\bar{d}_{t-5}}$ tells us how much higher or lower demand was than its routine value in trading period $t-5$. Assuming that the demand factor stays relatively constant over short periods of time, it seems reasonable to estimate d_t as

$$\hat{d}_t = \frac{d_{t-5}}{\bar{d}_{t-5}} \bar{d}_t. \tag{3}$$

This estimator alone achieves a similar level of performance to the neural network model. We can achieve further improvement if we add the value of this estimator as a feature. A random forest was then the best-performing machine learning model we tried, achieving the error shown in Table 3.

Demand can be broken down into three components: fixed price demand, price-responsive demand and intermittent, low capacity generation such as wind generation, which is modelled as negative demand in the NZEM. We found that certain features were more or less relevant for predicting each of these, and so we tried training models for each of the individual components using the especially relevant features. We will not go into detail about this refinement here, however it achieved the performance shown in Table 3.

	RF with normalised demand	Separated components
North Island	2.4%	2.16%
South Island	1.4%	1.37%

Table 3: MAE of demand models with normalised demand as a feature (% of mean demand value)

4 Benchmark Price Model

Before proceeding to a more complex model, we would like to obtain a benchmark by fitting a simple price model. Suppose we are focusing on an individual node. Let

p_t be the price of electricity in \$/MWh at that node in trading period t. We wish to forecast the price in some trading period t given only the information available 2 hours previously. Recall that this means the most recent data we can use is from trading period $t - 5$.

We will use the following features as inputs to our simple price model:

- previous price value p_{t-5},
- the month $M(t)$, day of week $W(t)$ and trading period number $P(t)$ of trading period t,
- whether trading period t falls on a national holiday,
- the temperatures in Auckland, Christchurch, Dunedin and Wellington,
- the predicted demand values $\hat{d}_t^N$ and $\hat{d}_t^S$.

A neural network with two hidden layers achieves the results shown in Table 4.

Benmore	Haywards	Otahuhu
21.3%	21.8%	21.3%

Table 4: MAE of simple nodal price predictions (as % of mean nodal price)

Figure 2 shows the distribution of errors in predictions for Benmore node made by the simple model according to different ranges of true price value. The mean target value was \$61.75/MWh and the mean predicted value was \$60.98/MWh, so there is a very slight systematic bias. However we can see that there is a slight positive bias when the prices are low and an increasingly negative bias as the true price increases, and also a greater variance in the errors. This is not surprising, as prices above \$150 occur rarely, so the machine learning model's loss function is not penalised much during training for underpredicting them.

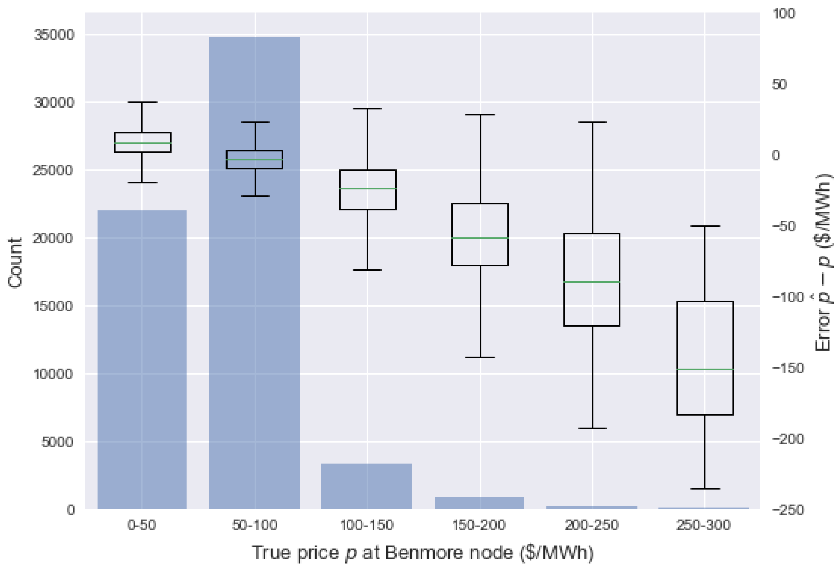

Figure 2: Error (right axis) in simple model predictions at Benmore by price range (range counts on left axis)

5 An Approximate Dispatch Model

We will now consider the problem of combining the aggregate offer stacks made available by WITS with our aggregate demand forecasts to produce a price estimate. Recall that nodal prices are calculated from the solution of the dispatch problem described in Section 2.1. Ideally we would be able to produce price estimates by solving a version of this problem with the North and South Island as nodes and the HVDC link as the single arc. This approach is unlikely to be successful however, as it ignores internal losses and capacity constraints within the islands. The precise effect of these factors is unknown, as we do not know the demand and generation offers at each individual node. However we can attempt to learn their effect using machine learning techniques.

5.1 Modelling Satisfaction

Suppose we are considering how much generation g_t^N and g_t^S to dispatch in the North and South Island in trading period t to satisfy the aggregate demands d_t^N and d_t^S. We say that "satisfaction" is achieved if the quantity of generation dispatched is sufficient to meet demand after line losses and capacities are taken into account. Satisfaction is not well defined if we know only the aggregate generation and demand, because it depends on the location at which the demand and generation occur. Therefore it makes sense to think of satisfaction as a probability. Let the *satisfaction function* $S(g^N, g^S, d^N, d^S)$ give the probability that North and South Island demands d^N and d^S could be satisfied with aggregate generation quantities g^N and g^S. We assume that S is independent of t, though such a dependence may exist, for instance because of electricity network maintenance and upgrades which cause its configuration to change over time.

We can think of the problem of learning S as a two-class supervised machine learning classification problem, where an example is defined by values g^N, g^S, d^N and d^S and has associated target value 1 if the generation is certainly sufficient to satisfy the demand and 0 if it is certainly insufficient. We know the true values of g_t^N, g_t^S, d_t^N and d_t^S in every trading period in our data set. Generation is dispatched in such a way as to just meet demand, so we could be confident that if the quantity of generation in any trading period had been slightly less than the true amount then it would have been insufficient to meet demand. Similarly, if it had been slightly greater, it would have satisfied demand comfortably. This observation enables us to create one negative and one positive training example from each trading period in our data set by imagining a 0.5% reduction or increase in the generation in each island:

$$S(0.995g_t^N, 0.995g_t^S, d_t^N, d_t^S) = 0 \quad \forall t \in T \tag{4}$$

$$S(1.005g_t^N, 1.005g_t^S, d_t^N, d_t^S) = 1 \quad \forall t \in T \tag{5}$$

Before training a machine learning model with these examples, we generate a few more features of interest. Suppose we have an example $[g^N g^S d^N d^S]$. Then we generate the following additional features:

- $g^{extra} = \frac{g^N+g^S}{d^N+d^S} - 1$, the proportion by which total generation exceeds total demand.

- $HVDC^{NS}$ and $HVDC^{SN}$, the estimated net North-South and South-North flow respectively, and $HVDC^{max} = \max\{HVDC^{NS}, HVDC^{SN}\}$, the total flow in the HVDC in either direction.
- $(d^N)^2$, $(d^S)^2$ and $(HVDC^{max})^2$. We include these quadratic terms because line losses are proportional to the square of the flows in the lines.

A logistic regression model trained with the above features yields an accuracy of 90.7%. Recall that one positive and one negative example was generated for each time period, so a trivial model which guessed that every example was positive could achieve only 50% accuracy.

The accuracy of the learned satisfaction function does not tell us very much about how useful it is for predictive purposes. Instead we performed an experiment in which we used the satisfaction function to predict an unknown generation quantity. Suppose we know the true values of d_t^N, d_t^S and g_t^S and we wish to recover g_t^N. We can estimate g_t^N as

$$g_t^N \approx \min\{g : S(g, g^S, d^N, d^S) \geq 0.5\}. \tag{6}$$

This value is easy to find, as S is expected to be increasing with g^N. The MAE of estimating g^N in this way was 13 MW, or 0.54% of the mean g^N. The error seems sufficiently low to validate the use of the learned satisfaction function.

5.2 Aggregate Dispatch Problem

Recall from Section 2.1 that the objective of the market dispatch problem is to minimise the cost of all dispatched tranches. The WITS aggregate offer stacks give us the quantity and price of all offered tranches in the North and South Islands. Let $G_t^N(c)$ and $G_t^S(c)$ be the quantity of generation that can be procured in the North and South Island respectively at a cost of c in trading period t. This can be calculated as the quantity g such that area under the relevant offer stack over the interval $[0, g]$ is c. The objective is to choose generation quantities such as to minimise total cost while achieving satisfaction. Thus we can write the *aggregate dispatch problem* in trading period t as

$$\begin{array}{ll} \min & c_t^N + c_t^S \\ \text{s/t} & S(G_t^N(c_t^N), G_t^S(c_t^S), d_t^N, d_t^S) \geq 0.5 \\ & c_t^N, c_t^S \geq 0 \end{array} \tag{7}$$

where the decision variables c_t^N and c_t^S are the costs allocated to procuring generation in the North and South Island respectively. This is an optimisation problem in two variables, with a linear objective and a constraint governed by the non-linear black-box function S. We consider a pair of costs (c_t^N, c_t^S) feasible if the probability of satisfaction with those costs is $\geq 50\%$.

Figure 3 shows the feasible region of this aggregate dispatch problem for the trading period 7:30-8:00am, May 22nd 2017. Every cost allocation (c_t^N, c_t^S) above the boundary is feasible. The feasible point marked with a red cross minimises total cost.

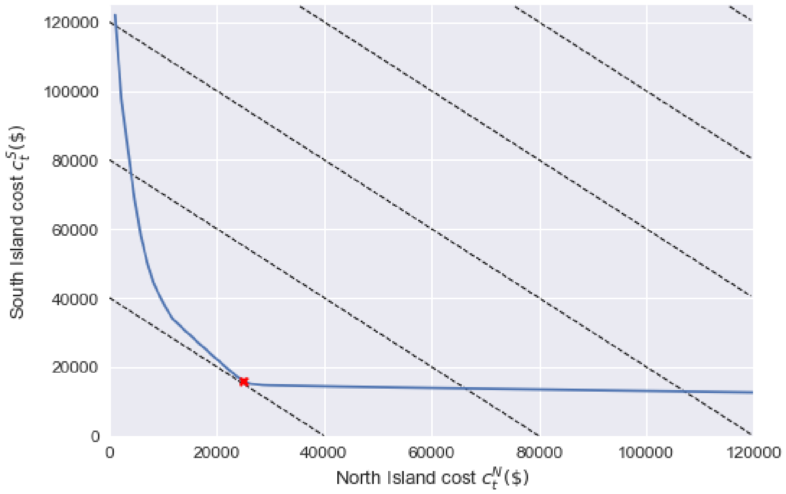

Figure 3: Feasible region of the aggregate dispatch problem

We will briefly outline a solution procedure for the aggregate dispatch problem. We make the following observations:

1. Suppose we can achieve satisfaction for some total cost $C = c_t^N + c_t^S$. Then we should be able to achieve satisfaction at any cost greater than C. If we can answer the question of whether satisfaction can be achieved at a total cost of C, we can thereby solve the aggregate dispatch problem using bisection or some similar univariate search method.

2. Satisfaction S should be increasing with g_t^N and g_t^S. Since $G_t^N(c_t^N)$ and $G_t^S(c_t^S)$ are increasing, S should also be increasing with c_t^N and c_t^S.

3. Suppose there is a fixed total cost C. Then we have $c_t^S = C - c_t^N$. The total generation procured G_t^T can be expressed as a function of c_t^N: $G_t^T(c_t^N) = G_t^N(c_t^N) + G_t^S(C - c_t^N)$. It can be shown that G_t^T is concave in c_t^N.

4. We define the satisfaction function with fixed cost C as

$$S_t^C(c_t^N) = S(G_t^N(c_t^N), G_t^S(C - c_t^N), d_t^N, d_t^S). \tag{8}$$

   It makes intuitive sense that $S_t^C(c_t^N)$ should be quasiconcave, since satisfaction is closely related to total generation procured. It appears from our experiments that $S_t^C(c_t^N)$ is indeed quasiconcave in general.

Assuming $S_t^C(c_t^N)$ is quasiconcave, it is easy to find its maximum. If its maximum value is ≥ 0.5, then satisfaction can be achieved with total cost C. As per observation 1 above, this allows us to solve the aggregate dispatch problem in full.

From the optimal cost allocation (c_t^N, c_t^S), we can estimate the electricity price in the North Island as the price of the most expensive tranche used if the North Island tranches are dispatched in ascending order of price until a cost of c_t^N is reached, and similarly for the South Island. We call these estimators the *marginal tranche prices* in the two islands.

6 Prediction Pipeline

These are the steps of our price prediction methodology for trading period t:

1. Use the demand prediction methodology to find estimated demands $\hat{d}_t^N$ and $\hat{d}_t^S$ (see Section 3).
2. Given the estimated demands, solve the aggregate dispatch problem to find the estimated generation quantities $\hat{g}_t^N$ and $\hat{g}_t^S$ (see Section 5.2).
3. Use the solution of the aggregate dispatch problem to find a price estimate $\hat{p}_t$ at the node of interest. We can use the marginal tranche price in the relevant island or a more sophisticated technique which we will discuss in Section 6.2.

6.1 Evaluation

Table 5 shows the error between the marginal tranche prices $\hat{m}_t^N$ and $\hat{m}_t^S$ predicted by this pipeline and the true nodal prices at our reference nodes. Note that in producing these predictions we have used blocked temporal cross-validation in all stages of the pipeline: when making a prediction for a given year, we have never used data from that year that we would not have access to at the time.

	$p_t^{(BEN)}$	$p_t^{(HAY)}$	$p_t^{(OTA)}$
$\hat{m}_t^N$		15.8%	17.0%
$\hat{m}_t^S$	19.7%		

Table 5: MAE between predicted marginal costs and nodal prices (as % of mean nodal price)

Comparing with Table 4, we see that the performance of this pipeline is significantly better than that of our simple model.

6.2 Improved Price Predictions

We would like to find a better estimator of the nodal prices. The marginal tranche price has the following flaws:

- The tranches are not necessarily dispatched strictly in price-order within each island. A different order is sometimes preferable due to line losses and capacities.
- The further a load is from the source of generation satisfying it, the greater the line losses, and therefore the higher the *true* marginal cost of consumption at the load point.
- The full market dispatch problem requires the procurement of reserve, which has an effect on the true marginal cost of consumption. Our model does not currently consider this.

We have very little idea however of what effect factors such as location and reserve procurement will have on nodal prices, as WITS does not give us the location at which each tranche is offered, nor complete information about the reserve market. Therefore we will resort to machine learning to attempt to learn from the past what effect these factors have. For each node $\mathcal{N}$ of interest, we will attempt to learn a function $P^{\mathcal{N}}$ such that

$$p_t^{\mathcal{N}} \approx P^{\mathcal{N}}(\boldsymbol{x}_t) \quad \forall t \in T, \tag{9}$$

where $\boldsymbol{x}_t$ is some vector of features relevant to the nodal price at $\mathcal{N}$ at time t. We will use the following features:

- $\hat{m}_t^N$ and $\hat{m}_t^S$, the estimated marginal tranche prices in the two islands.
- $\hat{d}_t^N$, $\hat{d}_t^S$, $\hat{g}_t^N$ and $\hat{g}_t^S$, the estimated quantities of demand and generation in the two islands.
- $\widehat{HVDC}_t^{NS}$ and $\widehat{HVDC}_t^{SN}$, the estimated HVDC flows as defined in Section 5.1.

Training a neural network model with two hidden layers using these features, we obtain the results shown in Table 6. Note that in order to make the best use of the training data, the cross-validation folds are individual months rather than years.

	MAE (\$/MWh)	MAE (% of mean)
Benmore	8.99	14.6%
Haywards	9.64	14.8%
Otahuhu	10.30	14.9%

Table 6: MAE of improved nodal price predictions given predicted feature values

Comparing Table 6 with Tables 4 and 5, we see that this model produces predictions which are somewhat more accurate than the estimated marginal tranche prices and significantly more accurate than those produced by our benchmark model. Furthermore if we compare the residual plot for our improved predictions in Figure 4 with the corresponding plot for our benchmark model in Figure 2, we see that there is still a tendency to underpredict when prices are high, but this is much reduced.

7 Summary and Conclusions

We trained a machine learning model to predict aggregate North and South Island electricity demand two hours in advance. This model achieved a mean absolute error of 2.8% of mean demand in the North Island and 1.7% in the South Island. We found that we could improve our model by normalising the demand during a period against mean demand value for historical periods having similar temporal features. The error of our final demand model was 2.2% in the North Island and 1.4% in the South Island.

We used similar techniques to train a benchmark model to predict nodal electricity prices. This model achieved an error of 21-22% on our reference nodes. We found that this simple model had a tendency to heavily underpredict when the true price was greater than $100/MWh.

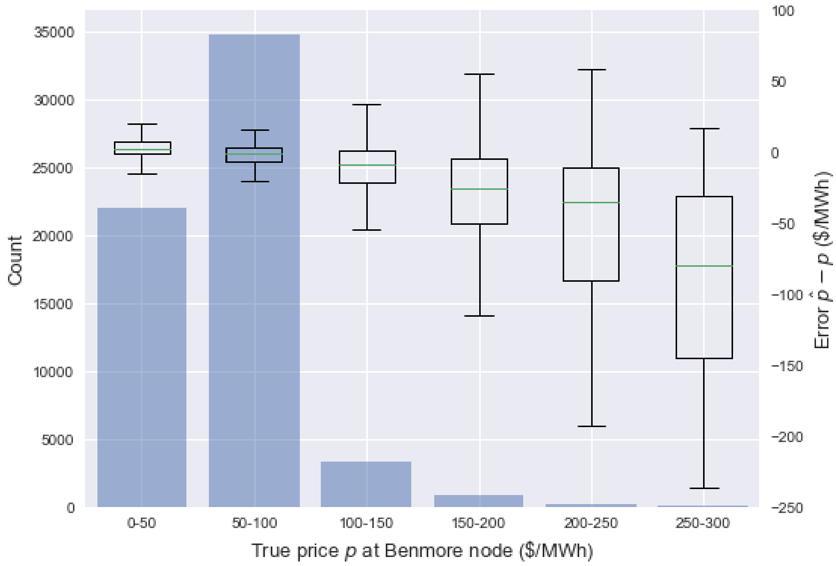

Figure 4: Error (right axis) in improved predictions at Benmore by price range (range counts on left axis)

We showed that given only information about aggregate North and South Island demand and generation offers, it is possible to approximate the full market dispatch problem with a cost-optimisation problem constrained by a machine-learned "satisfaction function." We outlined a technique for solving the resulting aggregate dispatch problem and thereby estimating nodal prices. The resulting price model achieved an error of 14-15% . It exhibited a significantly reduced tendency relative to the benchmark model to underpredict when prices are high.

References

[1] C. Bergmeir and J. Benítez. "On the use of cross-validation for time series predictor evaluation". In: *Information Sciences* 191 (2012), pp. 192–213.

[2] H. Bessembinder and M. Lemmon. "Equilibrium Pricing and Optimal Hedging in Electricity Forward Markets". In: *The Journal of Finance* 57.3 (2002), pp. 1347–1382.

[3] B. van Campen et al. *A Guide to the New Zealand Electricity Market*. 2009.

[4] F. Chollet et al. *Keras*. 2015. URL: `https://github.com/fchollet/keras`.

[5] J. Dunn. "Binary Interruptible Load Optimisation". Part IV project. Department of Engineering Science, University of Auckland, 2013.

[6] NIWA. *The National Climate Database*. URL: `https://cliflo.niwa.co.nz/`.

[7] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.