Chapter 14

An Electricity Procurement Model With Energy and Peak Charges

A. B. Philpott* and G. Pritchard†

Summary

We describe a model developed to help minimize the energy procurement costs of a New Zealand process industry that is a high user of electricity. The model accounts for stochastic prices that depend on the hydrological state of the electricity system, as well as transmission charges that are incurred during coincident electricity peaks. We describe how these are modelled and derive a stochastic dynamic programming algorithm that is used to arrange production to meet demand while minimizing the expected costs of electricity procurement.

1 Introduction

This paper deals with a practical problem facing many manufacturing industries with reasonably flexible production: when and how should they procure electricity to minimize the cost of production needed to meet some future demand. This problem is particularly important in process industries that are heavy users of electricity (such as aluminium, food processing and the pulp and paper industry). By shifting production into time periods in which electricity is inexpensive, companies may make considerable savings. This type of behaviour is often called "peak shaving" or "demand response" and there is a substantial literature and consulting activity devoted to doing this efficiently (Borenstein *et al.*, 2002).

The rationale for shifting loads out of peaks comes about because utilities and electricity markets extract higher prices during periods of peak

*Department of Engineering Science, The University of Auckland, Private Bag 92019, Auckland, New Zealand, a.philpott@auckland.ac.nz.
†Department of Statistics, The University of Auckland, Private Bag 92019, Auckland, New Zealand, g.pritchard@auckland.ac.nz.

demand. The economic theory underlying such peak-load pricing has a long history (see e.g. Boiteux, 1960, Crew *et al.*, 1995). In this theory, higher prices should be charged during peak times to reflect the higher utilization of capacity at these times. Pricing policies of this type can be devised to maximize welfare.

In the setting of electricity markets, sellers of electricity do not choose prices with the aim of maximizing total welfare. Rather, the prices emerge from some market-clearing mechanism that meets demand given the supply functions chosen by electricity suppliers. In this environment, prices will vary with time of day, being more expensive in peak hours when the market clears at high-price points on the sellers' supply curves. Consumers who are flexible have incentives to move production out of these peaks. In choosing their supply curves, sellers of electricity should account for this consumer response in seeking to maximize their own profit. A simple model of this phenomenon is discussed in Pettersen *et al.* (2005).

Our focus in this paper is, however, not on the sellers' pricing decisions, but on the purchaser's decision problem given these prices. The prices we consider are market pricing outcomes for energy and reserve, and peak-load tariffs imposed by the grid owner. We assume that the purchaser is not strategic, and so acts as a price taker. The purchaser then uses a mathematical programming model to optimize its production facilities. When prices are known, this becomes a deterministic optimization problem that can be attacked with standard mixed-integer programming software. See for example Henning (1998) and Ashok and Banerjee (2001) for models of this type. When prices are uncertain, but can be modelled with scenarios one might adopt a stochastic programming approach to this problem (see e.g. Philpott and Everett, 2002).

In this paper we focus on the purchaser's decision problem in a particular industrial setting in New Zealand in which a single product is made to meet a known demand occurring at known future points in time. The particular industry we have in mind exports products by ship, and enough stock must be on hand to satisfy a given schedule of ship visits. The production problem is relatively simple when electricity prices are known - one simply produces in sufficiently many low priced periods to generate enough stock to meet demand. When prices are random, but can be modelled as a Markov chain, this problem can be solved as a Markov decision process using dynamic programming (see e.g. Ravn and Rygaard (1994) for such a model for meeting heat and power constraints over time). We follow this line of reasoning in developing our model.

The model we discuss in this paper has two new features. Production facilities that are running may be paid a reserve price l to provide an interruptible load in circumstances where there is some contingency (e.g. a generating unit failure). In other words, the system operator pays l in every trading period for the option to disconnect the power supply to the production units if such an event happens. If the industry has agreed to such a contract then this payout effectively decreases the price p that is paid for electricity. (The decision to offer reserve in this way is exogeneous to our model.) Our stochastic price process is therefore estimated from historical sequences of the net price

$$e = p - l,$$

where both p and l are determined for each trading period by the electricity market clearing mechanism.

The second feature of our model that is novel concerns a payment that depends on the regional peak demand. In New Zealand, the grid owner records the peak demands in each region for each trading period (each having duration half an hour) in the 12 months from September 1 to August 31. These are called *coincident peaks* as they relate to all loads, not just those of the electricity consumer we are modelling. On August 31 each year the grid owner sums each consumer's load in the 100 highest coincident peak periods of the past 12 months. The consumers then pay the grid owner a peak charge M for every megawatt-hour (MWh) purchased in those periods. As regional demand increases towards the daily peak, the purchaser of power is faced with a delicate decision problem: should she shut down her plant in anticipation that the next half hour will be one of the highest 100 periods, even though this fact might not be known until many months in the future?

We show how this problem can be solved by dynamic programming producing a threshold-type policy. This policy requires an estimate at each point in time of the probability that the regional load being observed will exceed the 100th highest regional load observed over the 12 months from September 1 to August 31. This estimate is made using a model for these peak demands.

The paper is laid out as follows. In the next section we formulate the decision problem we wish to solve as a dynamic programming problem, and show that a threshold policy is optimal. We then describe models of random electricity prices and coincident peak demand that have been developed to incorporate some of the serial correlation in these data. In the following section we use these models to extend the dynamic programming model, and illustrate its output on an example with fictitious plant data, but real price and peak demand data from the New Zealand electricity market. The paper concludes with a discussion of the implementation of this model in a practical setting.

2 A dynamic optimization model

In this section we derive a dynamic optimization model for production planning using machines that consume electrical energy. We assume for simplicity that production levels are binary - either machines are all running at full capacity, or they are all turned off. In the last section of the paper we will discuss how this assumption can be relaxed.

The model we describe has two levels of time discretization. We denote by $t = 1, 2, \ldots, T$, the *stages* of a dynamic decision problem, where for each stage we will compute an optimal action. Within each stage there are *trading periods* denoted $k = 1, 2, \ldots, K$. In the New Zealand application, we choose a stage length of one day, divided up into 48 half-hour trading periods.

The decision to be made in any stage t and any trading period k is to determine whether to run the plant given:

1. the current inventory level of our product (z);
2. the current electricity price (p) and price (l) for interruptible load;
3. the current value of total regional demand (x).

We assume that p, l, and x for the trading period are all known at the time the decision is made to shut down. In practice, p, l, and x will be short-term forecasts based on data recorded in previous trading periods. This of course makes them subject to some forecast error, which we assume is neglible in this description. We will discuss the implication of this assumption in the conclusion of the paper.

We consider first the optimal action to be taken at any given stage t, assuming that z is known at the start of stage t. Suppose that the producer requires a given amount y to be produced in stage t. The optimal decision at stage t will depend on realizations of p, l, and x at every $k = 1, 2, \ldots, K$, which we must treat as random variables $P(k)$, $L(k)$, and $X(k)$. We define the random variable $E(k)$ to be the net price per MWh of procuring energy in trading period k, where $E(k) = P(k) - L(k)$ is the difference between the electricity spot price and the price of interruptible load.

First, consider a situation in which the net price and regional demand $(E(k), X(k))$, $k = 1, 2, \ldots, K$, are i.i.d. random variables with joint density $f(e, x)$ over sample space Ω. For any realization (e, x) we wish to decide whether to run the plant or not. Let $\rho(e, x)$ be an indicator function that is 1 if the plant is run (at full capacity) and 0 if it is shut down. The probability that the plant runs in any trading period is

$$\int_\Omega f(e, x)\rho(e, x)dedx.$$

Suppose that we observe e and x. Then the expected cost per MWh of running in realization (e, x) is

$$e + M\,\mathrm{Pr}(x \geq X_{100})$$

where X_{100} is the (random) 100th highest regional demand, and M denotes the peak period charge per MWh. Suppose we know the distribution $G(x) = \mathrm{Pr}(X_{100} \leq x)$. Then the expected cost per MWh of running in realization (e, x) is

$$e + MG(x).$$

In order to produce y tonnes on day t, the producer seeks an indicator function ρ to solve

$$\begin{aligned} P : \text{minimize } & \frac{b}{a}\int_\Omega (e + MG(x))f(e, x)\rho(e, x)dedx \\ \text{s.t. } \qquad & b\int_\Omega f(e, x)\rho(e, x)dedx = y \end{aligned}$$

where b is the capacity of the plant in tonnes per day, and a is the number of tonnes of product from one MWh of electricity consumption.

The Lagrangian for P is

$$\mathcal{L}(\rho, \lambda) = y\lambda + \frac{b}{a}\int_\Omega (e + MG(x) - a\lambda)f(e, x)\rho(e, x)dedx.$$

Minimizing $\mathcal{L}$ defines a threshold policy

$$\rho_\lambda(e, x) = \begin{cases} 1, \ e + MG(x) \leq a\lambda \\ 0, \ e + MG(x) > a\lambda \end{cases}$$

where λ is chosen so that this policy gives

$$b\int_\Omega f(e, x)\rho_\lambda(e, x)dedx = y.$$

The optimal value of λ can be interpreted as the marginal value of an extra tonne of production and will be an increasing function of the requirement y.

In a multi-stage setting, we wish to adjust the amount we produce in each stage t depending on how much inventory has been accrued. The optimization problem that we wish to solve is to determine whether to shut down the plant when we observe high values of $e + MG(x)$ on a particular day. The peak values usually occur in the morning and early evening when domestic demand is high. Typically the firm would stop production for these periods for sufficiently high values of $e + MG(x)$. Production typically resumes in the late evening, and the procedure repeats the following morning.

For any given day t in the planning horizon, we seek an optimal threshold value λ. Given this value, the plant will shut down completely in every trading period that has $e + MG(x) > a\lambda$. Given a value of λ, the amount of product produced by policy $\rho_\lambda(e, x)$ is

$$w_t(\lambda) = b \int_\Omega f_t(e, x)\rho_\lambda(e, x) dp dx,$$

and its expected cost is

$$c_t(\lambda) = \frac{b}{a} \int_\Omega (e + MG(x)) f_t(e, x)\rho_\lambda(e, x) de dx.$$

Observe that the density $f_t(e, x)$ may now vary with the stage t.

Given this optimization model for each stage t, we can derive a dynamic programming recursion for $t = 1, 2, \ldots, T$. Let $C_t(z)$ be the minimum expected future cost of meeting demand d_τ, $\tau = t + 1, t + 2, \ldots, T$, if there is a stock level of z at the end of day t. Then at the start of day t we seek

$$C_{t-1}(z) = \mathbb{E}\begin{bmatrix} \min_{\rho,y} \dfrac{b}{a} \displaystyle\int_\Omega (e + MG(x)) f_t(e, x)\rho(e, x) de dx \\ \quad + C_t(z + y - d_t) \\ \text{s.t. } y = b \displaystyle\int_\Omega f_t(e, x)\rho(e, x) de dx \end{bmatrix} \tag{1}$$

The stage problem (1) gives Lagrangian

$$\begin{aligned}\mathcal{L}(\rho, \mu) = {} & \frac{b}{a} \int_\Omega (e + MG(x)) f_t(e, x)\rho(e, x) de dx + C_t(z + y - d_t) \\ & + \lambda y - \lambda b \int_\Omega f_t(e, x)\rho(e, x) de dx\end{aligned}$$

and a policy

$$\rho_\lambda(e,x) = \begin{cases} 1, \; e + MG(x) \leq \lambda a \\ 0, \; e + MG(x) > \lambda a \end{cases} \tag{2}$$

where

$$0 \in \lambda + \partial C_t \left(z + b \int_\Omega f_t(e,x)\rho_\lambda(e,x)dedx - d_t \right).$$

Thus λ is the marginal value of storage at $z + b\int_\Omega f_t(e,x)\rho_\lambda(e,x)dedx - d_t$. If we assume that λ is constant over the range of production then this gives a threshold policy defined by (2).

3 Statistical models

The recursion (1) used in the model in the previous section assumes a stagewise independent distribution for (E, X). In reality, (E, X) follows a more complex stochastic process with some stagewise dependence. In the next section, we will derive the dynamic programming recursion for this more complex model, which we now proceed to describe in more detail.

Although we are primarily interested in the electricity price, much of the structure of price series in a hydropower-dominated system (like New Zealand) is derived from the underlying hydrology. This is illustrated in the plot in Figure 1 that shows regional electricity spot prices in New Zealand varying with storage in its largest hydro-electric catchment (the Waitaki system). The first step in modelling prices that depend on hydrology is to develop a process for representing inflows and releases to the hydro catchments. The dynamics of a hydro reservoir are governed by the equation:

$$S_{t+1} = S_t + I_t - R_t$$

where S_t is the stock of water, I_t is inflow and R_t is release of water (for generation and spill). Although releases represent actions of the generators, they are influenced by storage levels to the extent that one can derive reasonable statistical models that represent this (see e.g. Tipping and Read, 2010). In other words high storage levels tend to give large releases and low storage levels give small releases.

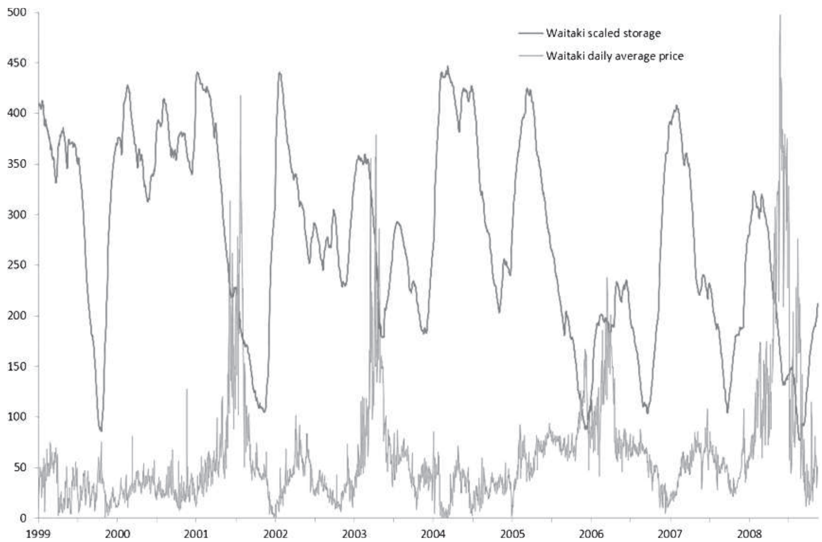

Figure 1 Electricity spot prices and hydro storage levels in the Waitaki hydro system. Prices spike when hydro storage levels drop because of low inflows.

3.1 *Inflows*

We begin our discussion of hydrological processes by discussing inflows. It has been traditional to model hydropower inflows by replaying inflow sequences from past years, starting at the appropriate time of year. This is a generally sound idea, as it reproduces the appropriate distributions, serial correlation structure, etc. But if the sequences are to be regarded as possible scenarios for the immediate future, one should take some account of current conditions. If recent weeks have been dry, one should not use sequences taken from wet years. The following model is an attempt to make a simple adjustment for this.

Our model is:

$$\log I_t = \alpha \log I_{t-1} + T_t^I + \text{error}$$

where I_t is the inflow in week t, and T_t^I is an annual seasonal factor consisting of a second-order trigonometric polynomial with a 1-year period. That is, the log-inflows consist of a fixed seasonal pattern with superimposed red noise.

We can use this in conjunction with historical inflow sequences as follows. The model gives

$$I_t = I_{t-1}^{\alpha} \times \text{(seasonal and random factors)}.$$

If the historical sequence is $h_0, h_1, h_2, \ldots$, and we have already observed this year's inflow for the week corresponding to h_0 to be I_0, then a scenario for the following week's inflow is

$$I_1 = \left(\frac{I_0}{h_0}\right)^{\alpha} h_1.$$

For the week after that, we get

$$I_2 = \left(\frac{I_1}{h_1}\right)^{\alpha} h_2 = \left(\frac{I_0}{h_0}\right)^{\alpha^2} h_2,$$

and similarly,

$$I_j = \left(\frac{I_0}{h_0}\right)^{\alpha^j} h_j. \tag{3}$$

Since $\alpha < 1$, we have $\alpha^j \to 0$ and so $I_j \approx h_j$ after the first few weeks.

We use a single model of this type to represent the combined inflows (in energy-equivalent, i.e. gigawatt-hour, terms) for all the large hydro lakes in the New Zealand power system. (The lakes included are named Tekapo, Pukaki, Ohau, Hawea, Te Anau, Manapouri, Taupo, and Waikaremoana.) The fitted value of α is approximately 0.44.

3.2 *Releases*

We now turn attention to the hydroelectric energy released from storage by the electricity industry. The reservoirs included are the same as those in the inflow model.

The model is:

$$R_t = \beta_1 S_t + \beta_2 S_t^2 + \beta_3 S_t^3 + \beta_4 I_t + T_t^R + \text{error} \tag{4}$$

where R_t is the energy released in week t, S_t the total stored energy at the beginning of week t, and T_t^R is an annual seasonal factor consisting of a first-order trigonometric polynomial with a 1-year period.

After fitting this model the residuals are found to be fairly symmetrical about 0, to be approximately normally distributed, and to have little serial correlation.

3.3 ***Electricity prices***

Electricity prices are less well-behaved than the hydrology, and so the following models are inevitably somewhat more approximate.

3.3.1 *Price duration curves*

The optimization problem that we wish to solve is to determine whether to shut down the plant when we observe high values of $e + MG(x)$ on a particular day. The key data input in this decision is a *price duration curve*,

$$D(s) = \text{the number of trading periods where } e + MG(x) \leq s.$$

The inverse $D^{-1}(s)$ can be thought of a realization of values of $e+MG(x)$ over a 48-period day, ordered from lowest to highest. Such a curve is illustrated in Figure 2.

If one knows the price duration curve, then it is straightforward to find an optimal threshold policy by conducting a line search for λ that yields the desired production amount of y. Recall that there are K trading periods per day, a is the tonnes produced per MWh, and b is the daily production capacity. Then the optimal λ solves

$$b\frac{D(\lambda)}{K} = y.$$

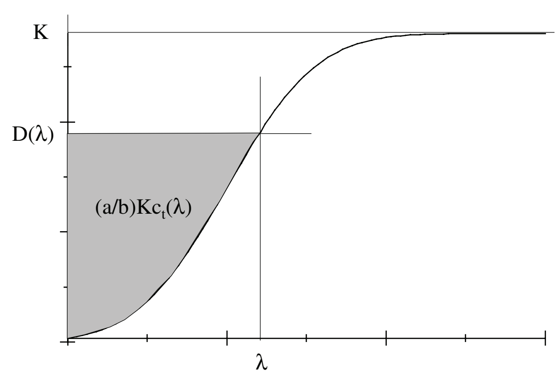

Figure 2 A price duration curve. $D(\lambda)$ is the number of periods with price at most λ. The shaded area when divided by K is the average cost per MWh of policy λ.

The cost of this choice per day is given by the area between the price duration curve and $D(\lambda)$

$$c_t(\lambda) = \frac{b}{aK}\left(\lambda D(\lambda) - \int_0^{\lambda} D(s)ds\right).$$

In our model, the price duration curve is random, having a finite distribution on each day. This distribution is determined by the weekly hydrology models above, and the random curves on consecutive days are not independent but are correlated in a way that we describe below.

We illustrate the modelling approach we have adopted by applying it to the spot price p; a similar process is used to model $e + MG(x)$. To estimate a model for scenarios of duration curves for p, it helps to transform the data. Some very high prices occur, and we first attempt to rein these in by applying the following transformation to all our price values:

$$P = c\left(\left(1 + \frac{3p}{c}\right)^{1/3} - 1\right);$$

the value $c = \$3/\text{MWh}$ is found to be suitable. Of course, this means that when using models for P, we will eventually have to transform back via

$$p = \frac{(P + c)^3 - c^3}{3c^2}.$$

Let $P_{t,k}$ denote the kth largest (transformed) price occurring on day t $(k = 1, \ldots, 48)$. To estimate price-duration curves, we use quantile regression (Koenker and Bassett, 1978). The model equations are:

$$Q_\tau(P_{t,k}) = \sum_{j=0}^{4}\left(\gamma_{1,j}S_t + \gamma_{2,j}\frac{1}{S_t} + \gamma_{3,j}I_t + \gamma_{4,j}H_t\right) f_j(t) + \text{error},$$

where:

1. $Q_\tau(P_{t,k})$ is the τ-quantile of the distribution of $P_{t,k}$;
2. S_t the hydro storage at the beginning of the week that includes t, and I_t the inflow during the week that includes t;
3. H_t a holiday indicator (0 if t is a business day, 1 otherwise);
4. $f_0, \ldots f_4$ are seasonal trigonometric functions with a 1-year period:

$$f_0(t) = 1, \quad f_1(t) = \cos(\omega t), \quad f_2(t) = \sin(\omega t),$$
$$f_3(t) = \cos(2\omega t), \quad f_4(t) = \sin(2\omega t).$$

Fitting these models for $\tau = 0.05, 0.15, \ldots, 0.95$ (and $k = 1, \ldots, 48$) gives a set of ten model price duration curves, each of which is a function of the covariates (S_t, I_t, and H_t, and seasonality) for the day in question.

For some values of the covariates, some of these curves can be slightly non-monotone; if this is important, the prices should be sorted into descending order after they have been computed.

3.3.2 *Transition model*

Each of the ten model scenarios should occur about 1/10th of the time, but this does not happen at random for each day, independently of previous days, as this would ignore the substantial serial dependence. Fitting quantile regressions as above, but for $\tau = 0.1, 0.2, \ldots, 0.9$, allows us to classify each historical price according to which model scenario it most closely resembles (e.g. if a historical price lies between the $\tau = 0.1$ and $\tau = 0.2$ models for the day on which it occurs, it is associated with the $\tau = 0.15$ model scenario). We associate each historical day to a scenario using the classification given by its 6th largest price. This gives a sequence of scenarios which "occurred" historically.

The serial dependence in this sequence of scenarios is much stronger than for a Markov chain with the same one-step transition matrix M. To represent this, we use the following "sticky" random model. Let $X_t \in \{1, \ldots, 10\}$ be the scenario used on day t, and $V_t \in \{1, \ldots, 10\}$ a "background state" pertaining to that day. Then (X_t, V_t) follows a Markov process given by

- X_t is chosen according to the V_{t-1}th row of M;
- $V_t = X_t$ with probability δ, otherwise V_{t-1};

with these random choices being made independently of those on previous days. The value chosen for δ is currently 0.10. This allows for a moderate degree of dependence between X_t and X_{t-r} when r is small (as they are likely to have been picked from the same row of M), which does not disappear too quickly as r gets larger.

3.4 ***Peak demand***

The final ingredient in our model is regional peak demand. Figure 3 shows an example of peak demand from September 1, 2006 to August 31, 2007

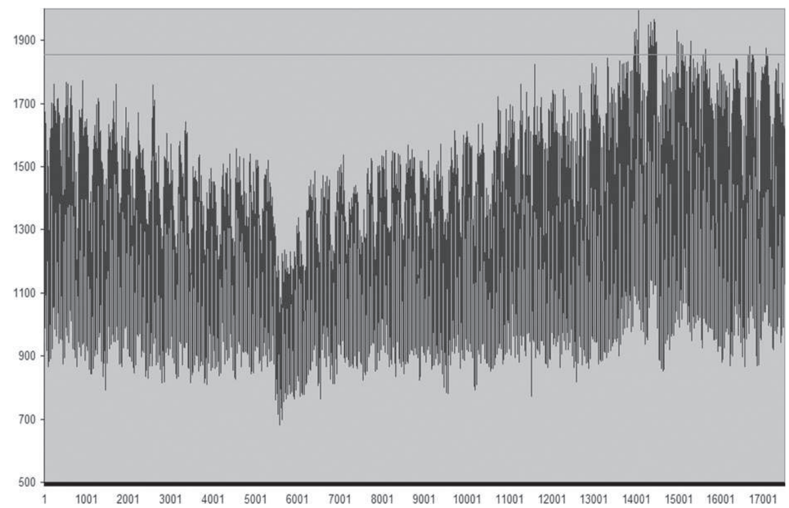

Figure 3 Regional half-hourly demand in 2006–2007. The 100th highest demand realization is shown by the solid horizontal line.

for a region of New Zealand. Observe that most of the peaks occur in the period May 1 to August 31. The horizontal line in the figure is set at the 100th largest demand (1855 MW).

The basis of our peak-load model is a suite of statistical models giving the expected highest, second-highest, etc. load each morning and evening of the winter (April–September). These are simply functions of time (the number of days before or after July 1) and the day of the week (allowing for public holidays). An offset (additive constant) to be added to each model is estimated using recent historical load data beginning on April 1. To this is added a random process chosen so that its serial and cross-correlations between adjacent mornings and evenings are the same as in the data. This allows peak loads to be simulated arbitrarily far ahead. The random process is initialized with a value for the current day predicted from the previous day's loads and air temperatures. The loads are then simulated from the current day until August 31, and the simulation repeated 10,000 times in order to obtain the statistics shown.

Thus for any temperature data provided the model delivers probabilistic estimates of what the 100th largest half-hourly load of the pricing year will be. It gives P5, P50, and P95 estimates (i.e. values with 5%, 50% and 95% probability, respectively, of exceeding the threshold) as well as an estimate of the distribution $G(x)$.

4 Dynamic programming recursion

We now discuss the dynamic programming recursion for determining an optimal policy over the time stages $t = 1, 2, \ldots, T$. The curse of dimensionality precludes a high-dimensional state space. By experimentation we have identified the key states being a national reservoir storage level (represented by $u \in \{1, \ldots, U\}$), electricity price state and background state (represented by $v \in \{1, \ldots, 10\} \times \{1, \ldots, 10\}$) and stock on hand of product ($z \in \{0, 1, \ldots, L\}$).

Based on this structure we have developed for fixed M, a statistical model for duration curves of

$$A = e + M \Pr(X_{100} < x),$$

for all possible values of u and v, as well as probabilities for transitioning between states (u, v) at each stage. We look at each of these in turn.

At each future stage t, we estimate a distribution for inflows at time t, based on the observed inflow at time 0, and simulations using the model represented by (3). This model induces a stochastic process on the storage levels and therefore on the releases according to (4). Thus, at any stage t and reservoir state u, we can estimate a transition probability $q(u, u')$ to state $u' \in \{1, \ldots, U\}$ at the next stage by computing a release decision for each inflow realization using (4) and a sampled error term. The inflows used at each stage in this process are assumed to be statistically independent from those in the preceding stage.

The model for A is very similar to the model for price duration curves described in subsection 3.3.1, except that now we construct duration curves for A rather than p. These curves are estimated a priori and stored for all possible values for t, u and v. We now write

$$\begin{aligned} D_t(s, u, v) = {} & \text{the number of trading periods in stage } t \text{ and state}(u, v) \\ & \text{where } A \leq s. \end{aligned}$$

Given such curves and a choice of threshold λ, we can compute the expected daily production $w_t(\lambda, u, v)$ of the plant, and the expected daily cost of production $c_t(\lambda, u, v)$ that results. Formally

$$w_t(\lambda, u, v) = \frac{b}{K} D_t(\lambda, u, v)$$

and

$$c_t(\lambda, u, v) = \frac{b}{aK} \left(\lambda D_t(\lambda, u, v) - \int_0^{\lambda} D_t(s, u, v) ds \right).$$

We also estimate transition probabilities $r(v, v')$ for moving from state v at time t to state v' at time $t+1$. (These are independent from the transitions in u.) As described in subsection 3.3.1 we use a sticky random model where the second component of v is a background state that changes in only about 10% of the transitions, and $r(v, v')$ depends only on this second component.

Now we can write down the dynamic programming recursion that we use.

$$\begin{aligned} C_{t-1}(z,u,v) &= h_t(z) + \min_{\lambda}\{c_t(\lambda,u,v)) \\ &\quad + \gamma \sum_{u'} \sum_{v'} q(u,u')r(v,v')C_t(z + w_t(\lambda,u,v) - d_t, u', v')]\} \\ C_T(z,u,v) &= V(z) \end{aligned}$$

where

$$\begin{aligned} C_t(z,u,v) &= \text{future cost at end of stage } t \text{ if in state } (u,v) \\ &\qquad \text{and inventory is } z \\ h_t(z) &= \text{holding cost for inventory } z \text{ over stage } t \\ d_t &= \text{demand for product in stage } t \\ V(z) &= \text{terminal future cost with } z \text{ in stock} \\ \gamma &= \text{daily discount factor} \end{aligned}$$

The dynamic programming recursion gives a threshold value $\lambda = \lambda^*(t, u, v, z)$ at each stage t and state (u, v, z) to determine at what (adjusted) price the company should reduce load. To determine what action to take, the company needs to estimate the current state. This is straightforward for u, but to estimate v (the price state) one requires the current duration curve for A at stage $(t-1)$ which can be estimated from the previous time period's observations.

To determine whether a shutdown is necessary in trading period k at stage t, the company must estimate the current value of $A(k)$ to compare it with the threshold value. This is computed using the current observation of $x(k)$ and sample values of X_{100} computed by the peak pricing model from subsection 3.4. If the sample values of X_{100} are sorted and compared to $x(k)$ then $\Pr(X_{100} < x(k))$ can be estimated as $x(k)$ is observed. Thus to determine an optimal action in each trading period, the peak pricing model must be solved at the start of the day to provide the appropriate estimate of $\Pr(X_{100} < x(t))$.

We do not allow backlogging of stock, so z is penalized from becoming negative. In fact, given constraints on production capacity, there is a minimum stock level that must be held at each stage to ensure that there will be enough stock on hand to meet demand (which is known a priori). Deviations below this minimum stock level are heavily penalized, as we can guarantee that they will lead to a shortage in demand. Moreover, in our computations we need only consider values of z that lie above this riskzone level. An example of such a riskzone level is shown in Figure 4.

The form of the solution to the dynamic programming recursion is best illustrated by an example. Figure 5 shows a contour plot of the optimal threshold values (with different shadings between contours) for state $(u, v) = (5, 10, 10)$ and varying z levels over the same period covered by the ship visits shown in Figure 4. In this example we have used 10 reservoir states, so $(5, 10, 10)$ corresponds to an average reservoir storage state (5), but high electricity price and background states (10). The shaded region at the top of the figure corresponds to the range $\lambda^*(t, u, v, z) \in [0, 20]$.

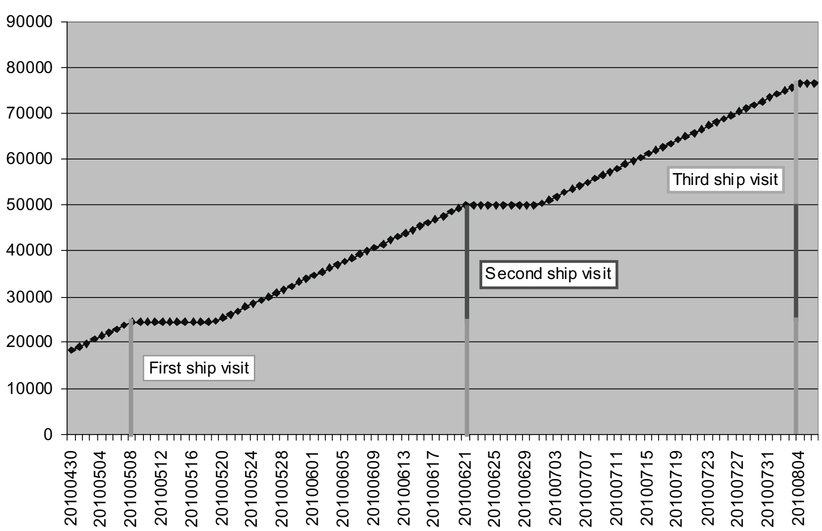

Figure 4 A typical riskzone that includes demand. Ship visits occur on May 8, June 21 and August 4. The model must maintain a cumulative production above the blue line to ensure that the demand from these ship visits can be met. The sloping sections grow at the maximum production rate.

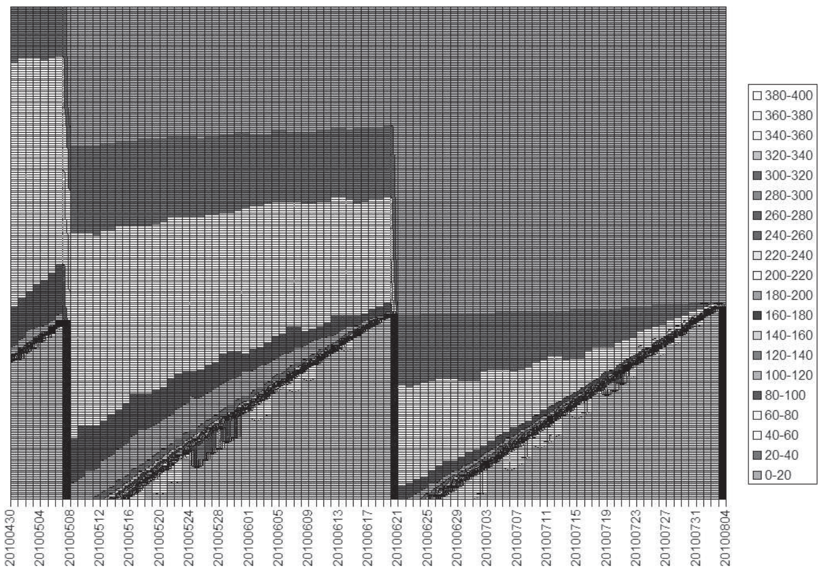

Figure 5 Contours of threshold values $\lambda^*(t, u, v, z)$ for state $(u, v) = (5, 10, 10)$. Shadings of (t, z) indicate different ranges of threshold values. The contours are vertical at times where demand deliveries must be made.

One can see from the plot that for fixed t, $\lambda^*(t, 5, 10, 10, z)$ decreases as the storage z increases. In the top region we have more than enough stock to meet the ship visits, and to avoid inventory costs a threshold value of close to zero is optimal. In this region the plant should shut whenever A gets above 20. The triangular regions where (t, z) lies in the riskzone are easy to see, and the contours of $\lambda^*(t, 5, 10, 10, z)$ become close together as (t, z) approaches the boundaries of these, since, to avoid a stockout, the optimal value of λ must increase at a higher rate to ensure production is not interrupted.

The most valuable information from our model is not so much the threshold policy that it computes, but the marginal value of storage λ that is obtained. This gives the slope of the expected future cost $C_t(z, u, v)$ of meeting the demand from time t to T, with z in stock, given an optimal threshold policy is followed. The future cost is defined for all possible values of the national storage state (u) and market state (v). It is also defined for all possible levels of storage (z) above the riskzone. An example of such a marginal cost curve for June 5, 2010 is shown in Figure 6.

This information can be used to guide a daily optimization model that incorprates the details of the company's plant operations. To do this a plan

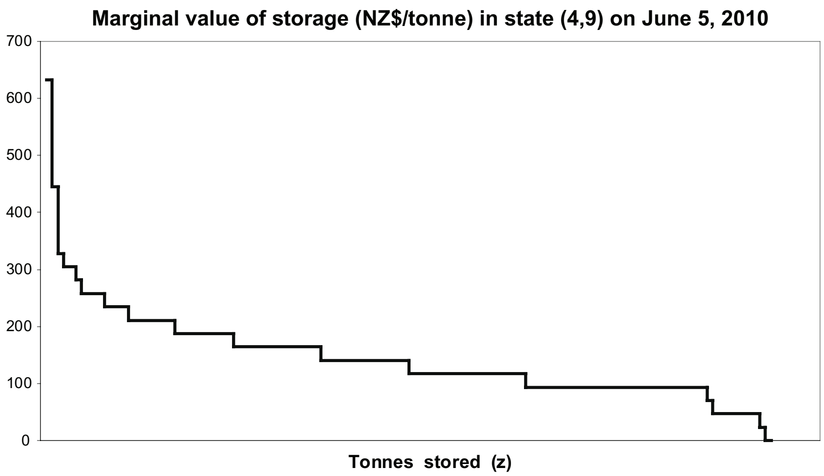

Figure 6 Marginal value of product stored as a function of inventory level (z), for market state (5,10,10) at $t = 2010605$ (as annotated on Figure 5).

is computed for the current day that uses forecasts (or scenarios) for $A(t)$ in the current day, and minimizes today's cost plus the expected future cost as defined by $C_t(z, u, v)$.

The daily optimization can be of higher fidelity than the optimizations used to compute the marginal value of storage, and include several operating units, ramping constraints, startup costs and shutdown costs. The marginal cost curve ensures that a close to optimal trade off is made between running and increasing the amount of inventory stored, and shutting down and avoiding energy and possible peak charges.

An example of such a model is the following mixed integer programming model of the unit commitment type. To see how this would work, assume that we have predicted $A(t)$ over the course of the coming day. This gives the expected cost of purchasing a MWh of electricity (accounting for maximum–demand charges, electricity price and reserve.) Let

$$x(t) = \begin{cases} 1 & \text{if plant runs in period } t \\ 0 & \text{otherwise} \end{cases}$$

Let N be the cost of switching the plant off, and let y denote the total amount of product produced in a day. Suppose the amount of product in inventory is z_0 and we increase this to $z_0 + z$. Let $a_j + b_j(z_0 + z)$, $j = 1, 2, \ldots, J$, be cutting planes representing the future expected savings from producing z over the day. The savings function is then given by the function

$$Q(z) = \min_{j=1,\ldots,J}\{a_j + b_j(z_0 + z)\}.$$

The mixed integer program we solve is:

$$\begin{aligned}
\min\ & \textstyle\sum_{t=1}^{48} Nx(t) + A(t)y(t) - \theta \\
\text{s.t.}\ & x(t) \geq y(t) - y(t+1), \\
& \frac{k}{48}\textstyle\sum_{t=1}^{48} y(t) = z, \\
& \theta \leq a_j + b_j(z_0 + z), \qquad\qquad j = 1, 2, \ldots, J, \\
& x(t), y(t) \in \{0, 1\}.
\end{aligned}$$

This gives a sequence $x(t)$ of ones and zeros that can be used to determine when to shut down the plant. The switching cost N will prevent too many shutdowns in a day. The optimal value z^* obtained will be the total production over the day. The optimal value θ^* from this model will satisfy

$$\theta^* \leq a_j + b_j(z_0 + z^*),\ \ j = 1, 2, \ldots, J,$$

and be as large as possible to minimize the objective of MIP and so

$$\begin{aligned}
\theta^* &= \min_{j=1,\ldots,J}\{a_j + b_j(z_0 + z^*)\} \\
&= Q(z_0 + z^*)
\end{aligned}$$

the future savings from producing z^*.

5 Conclusions

This paper has described a simple but effective model for peak-shaving industrial electricity demand. The model relies on a "top-down" statistical model of electricity prices and regional peak demand. Peak shaving models are well understood in a deterministic framework, but have received little attention in an uncertain environment, and none when peak charges are incurred in hindsight. Our model is a first attempt at constructing a model for problems of this type. Although the model is specifically adapted for use in the New Zealand context, the approach should be readily applicable to other settings, albeit with some structural changes to the statistical models.

The current version of our model assumes that the national inflows are independent from week to week. This assumption is made for computational convenience, and it does have implications for the policies that are computed. In essence, if a week is dry then the model will not assume that the next week is more likely to be dry and so it will not be as conservative in maintaining production as it should be. Similarly if the week is wet, then the model might recommend producing more than needed because it does not account for an increased probability of high inflows in the following week.

The effect of stagewise dependence in national inflows can be investigated by increasing the state space to include the previous day's inflows, with an increase in computational complexity. However the extent of this effect is not as bad as it might seem at first because the dynamic programming model is in some sense "self correcting". This happens because in applying the solution from the model, a sequence of dry inflow weeks and a less conservative policy will eventually give a low national storage level, and lower than needed inventory levels. As this occurs, the optimization model will increase the threshold value of λ until the user starts to shut down less and act more conservatively.

In our model we have assumed that the plant operator can predict e and x immediately prior to the trading period to which these pertain (so that they can shut the plant before incurring the charge $e+MG(x)$). In practice, large firms have data feeds to these parameters, and so the forecasts are relatively precise (in contrast to a prediction that the next period's regional demand will be one of the 100 highest). The model could incorporate some uncertainty in these forecasts with some increase in complexity by estimating probability distributions of e, and x about forecast values $\hat{e}$, and $\hat{x}$. The threshold policy then becomes

$$\rho_\lambda(\hat{e}, \hat{x}) = \begin{cases} 1, \mathbb{E}[e + MG(x) \mid \hat{e}, \hat{x}] \leq a\lambda \\ 0, \mathbb{E}[e + MG(x) \mid \hat{e}, \hat{x}] > a\lambda \end{cases}$$

requiring some additional computation in the dynamic programming recursion.

A futher restrictive assumption in our model is that a single commodity is being produced. Even the process industries that we have in mind in this work admit different product types (e.g. corresponding to purity of aluminium or basis weight of paper). Unfortunately, the curse of dimensionality means that increasing the number of commodities increases the complexity of the dynamic programming problem considerably, unless the production processes can be decoupled so that they do not share costs of electricity procurement. With this proviso, the complexity of the production scheduling process on the day of operation can be considerably more complex than the one we have modelled, as illustrated in the previous section. The model of this paper gives some end conditions on a daily scheduling model that will ensure that the production schedule makes appropriate tradeoffs in electricity procurement.

References

Ashok, S. and R. Banerjee. An optimization model for industrial load management. *IEEE Transactions on Power Systems*, 16(4): pp. 879 –884, 2001.

Boiteux, M. Peak-load pricing. *The Journal of Business*, 33(2): pp. 157–179, 1960.

Borenstein, S., M. Jaske, and A. Rosenfeld. Dynamic pricing, advanced metering and demand response in electricity markets. Technical report, Center for the Study of Energy Markets, University of California Berkeley, 2002.

Crew, M. A., C. S. Fernando, and P. R. Kleindorfer. The theory of peak-load pricing: a survey. *Journal of Regulatory Economics*, 8: pp. 215–248, 1995.

Henning, D. Cost minimization for a local utility through CHP, heat storage and load management. *International Journal of Energy Research*, 22: pp. 691–713, 1998.

Koenker, R. and G. Bassett Jr. Regression quantiles. *Econometrica*, 46(1): pp. 33–50, 1978.

Pettersen, E., A. B. Philpott, and S. W. Wallace. An electricity market game between consumers, retailers and network operators. *Decision Support Systems*, 40 : pp. 427–438, 2005.

Philpott, A. B. and G. R. Everett. Pulp electricity demand management. In *Proceedings of the 37th Annual Conference of ORSNZ*, pp. 55–63, 2002.

Ravn, H. F. and J. M. Rygaard. Optimal scheduling of coproduction with a storage. *Engineering Optimization*, 22(4): pp. 267–281, 1994.

Tipping, J. and E. G. Read. Hybrid bottom-up/top-down modeling of prices in deregulated wholesale power markets. In S. Rebennack, P. M. Pardalos, M. V. F. Pereira, and N. A. Iliadis, eds., *Handbook of Power Systems II*, Energy Systems, pp. 213–238. Berlin Heidelberg: Springer 2010.