Benchmarking wholesale hydroelectricity markets with risk-averse agents*

Andy Philpott† Ziming Guan‡

November 9, 2018

Abstract

Using a suite of empirical models we study the effects of randomness and risk aversion on wholesale electricity prices that are affected by uncertainty in hydroelectric reservoir inflows. Our models combine stochastic dual dynamic programming with a high fidelity simulation model of the New Zealand wholesale electricity market. Results of the models are given for the calendar year 2012, and compared with wholesale market outcomes in that year. They demonstrate the effects of increasing uncertainty on productive efficiency and wholesale electricity prices. We also compare these results for a subset of models with those obtained using data from 2013.

Keywords: electricity market, hydroelectricity, stochastic inflows, risk-aversion, market power.

1 Introduction

In this paper we study the effects of uncertainty on outcomes in wholesale electricity markets that are dominated by hydroelectric generation. Unlike markets

*We would like to acknowledge support of the New Zealand Marsden fund under contract UOA1520. The work has benefited from modelling and theoretical contributions from many people. We wish to acknowledge in particular the contributions of Geoffrey Pritchard, Golbon Zakeri, Tony Downward, and Vitor de Matos to this work. The views and opinions expressed in this paper are, however, solely attributed to the authors.

†Electric Power Optimization Centre (EPOC), Department of Engineering Science, University of Auckland: a.philpott@auckland.ac.nz.

‡Electric Power Optimization Centre (EPOC), Department of Engineering Science, University of Auckland: z.guan@auckland.ac.nz.

consisting solely of thermal plant, markets with hydroelectricity have an inter-temporal aspect arising from the fact that energy (water) can be stored for later use. This complicates the decision making of hydroelectric generators as their computation of the marginal cost of releasing water must involve some modelling of opportunity cost and possible shortage costs, which depend on uncertain future reservoir inflows.

Previous work by the authors ([23] and [21]) used market simulation models to predict market outcomes under an assumption of perfectly competitive agent behaviour. One aim of these models was to provide regulators with a benchmark against which to compare market performance. The purpose of the current work is to outline a new suite of models that extends those described in [23] and [21]. The new models make use of recent features incorporated into the stochastic dual dynamic programming model used for water valuation, and use a high-fidelity market simulation model to accurately predict market outcomes arising from counterfactual agent behaviour. We compare the results of the simulation under different assumptions on uncertainty governing the reservoir inflow processes and the risk-aversion of market participants. Our demonstration of the simulation models focuses on the historical year 2012. This was noted as a year with low reservoir inflows over the first six months that led to high prices. We will make some comparisons with results from the following year (2013) but leave detailed analysis of other years to a forthcoming companion paper [22] that presents the results of applying our model to data from the historical years 2008-2016.

The perfectly competitive benchmarks that we compute are the result of simulating a system optimal policy. If electricity market participants maximize productive and allocative efficiency, then they will maximize social welfare for the system as a whole, at least in the short term. Maximizing system social welfare over an uncertain future is a challenging problem to define, let alone solve. For the purposes of defining a risk-neutral benchmark, we define the objective of this problem to be the expectation of operating benefits of participants (defined to be revenue minus operating cost for producers and the integral below the demand curve minus payment for consumers), taken with respect to a stochastic process of inflows that is common knowledge. In this setting it is possible to solve an approximation of the welfare maximizing problem and simulate the policy over historical inflow sequences. This provides a benchmark against which historical outcomes can be assessed.

We begin our study of 2012 with a version of the counterfactual model applied to the New Zealand market by Wolak [26]. The Wolak model uses a benchmark (called *fixed hydro* in this paper) in which hydro generation is set to its historical levels, and thermal plant is then offered in each period at its short-run marginal cost. It is argued (incorrectly) in [26] that the demand-averaged marginal cost computed from this experiment will be biased above the truly competitive

marginal cost, so using this benchmark as a counterfactual will give conservative estimates of price markups. In our previous work [23] we discuss the shortcomings of this counterfactual[1], as motivation for developing a benchmark based on multistage stochastic programming. Our study shows how incorporating uncertainty gives different results from fixed hydro.

The aim of benchmarking market performance is to identify inefficiencies, diagnose the causes of these and then devise institutional arrangements and instruments that might be used to reduce them. The causes of inefficiency identified by our work are not clear. Wolak's paper [26] attributes price differences to unilateral exercise of market power, but it is difficult to discriminate between this and the competitive behaviour of risk-averse agents. Indeed, the correspondence between a competitive equilibrium in hydro-dominated markets with risk averse agents, and a socially optimal plan is not well understood. In theory, under very special circumstances (no strategic play and risk neutral players), it is possible for a Walrasian equilibrium to give a stochastic process of prices with respect to which every agent optimizes its own expected benefit (revenue minus variable cost) with the outcome of maximizing total expected welfare. However, as shown by the examples in [3], the stochastic process of prices that yields an equilibrium might be very complicated with none of the stagewise independence properties that make computing optimal policies easy for generators.

When agents are assumed to be risk averse or do not have a common view of the (random) future, a competitive equilibrium is harder to identify. Agents in such an equilibrium (if it exists) will estimate marginal water values based on their risk-adjusted view of the future, and their actions in aggregate will yield equilibrium prices that are then used to form these views. Determining these prices for a multistage equilibrium would be very difficult. Furthermore, if we were to seek an equivalent socially optimal plan then it is necessary to integrate individual risk measures into a system risk measure to be optimized. It is well known that this is not possible for general nonlinear utility functions. On the other hand in the special case where agents optimize coherent risk measures [2], then it is possible to obtain a system risk measure in complete markets, i.e. where all risk can be traded by market participants (see [10], and [19] and [7] for this theory applied in a dynamic setting). Although we cannot expect such complete markets in practice, the formulation and solution of risk-averse centrally planned hydrothermal models in this paper provides a first step towards a competitive equilibrium benchmark of a complete market with risk-averse agents.

A similar problem of market incompleteness in hydroelectricity systems can occur in a risk-neutral setting, as shown by the paper by Lino et al [11]. They demonstrate using a computational example where all agents are located on diff-

[1]There have been several similar critiques of the model used in [26], see e.g. [6].)

ferent river systems that a risk-neutral centrally planned solution gives rise to system marginal prices that will clear the market with an optimal dispatch if each agent optimizes its own objective using these prices. Lino et al [11] then show how inefficiency can result from different agents operating hydro stations on the same cascaded river system. In this case, a market instrument pricing the transfer of water between generating stations is needed to recover the economically efficient solution[2].

The results in this paper show that the production inefficiencies in moving to a market model from a central plan are not that great in relative terms. The differences in prices between these solutions on the other hand are quite large, leading to very different distributions of benefits. As mentioned above it is hard to say whether the price differences we observe are due to unilateral exercise of market power. Many of the conditions for Walrasian equilibrium are missing, and the stochastic price process that one would like to use to clear the market as a stochastic optimization problem is never revealed to agents, but defined by a single realization of prices appearing every half hour out of the system operator's dispatch software. It is possible that the market structure we are working with has some way to go to providing a better approximation to the prices that we would obtain in equilibrium from a Walrasian auctioneer. Nevertheless, the presence of the differences between the benchmark and the historical outcomes give some grounds for deeper investigation.

The layout of the paper is as follows. In the next section we describe the New Zealand wholesale electricity market. We then outline the features of the suite of optimization models that we use in our study. The details of the models and the data that they use are provided in an online companion [8]. In section 4 we apply the models under different assumptions to every trading period in 2012, and compare results between models. We also make a comparison with the same models using 2013 data. The final section draws some conclusions.

2 The wholesale electricity market

Since 2004, New Zealand has operated a compulsory pool market, in which the grid owner Transpower plays the role of Independent System Operator (ISO). In this market all generated and consumed electricity is traded[3]. Unlike most

[2]The only case of this in New Zealand relates to Genesis and Meridian who operate different generating stations on the Waitaki system. Our counterfactual models assume that the generation of these stations is coordinated by contractual arrangements so that they yield the collective benefits that would be delivered by one (perfectly competitive) owner.

[3]Small generating stations with capacity of 10 MW or less are not required to make offers. From 1996-2004 a voluntary wholesale market existed, where approximately 80% of electricity

electricity markets in other parts of the world, the NZEM has no day-ahead power exchange. Bilateral and other hedge arrangements are still possible, but function as separate financial contracts. Trading develops by bids (purchaser/demand) and offers (generator/supply) for 48 half hour periods (called *trading periods*) over several hundred pricing nodes on the national grid. (Although demand side bids are included in the official description of the ISO dispatch model, there is currently very little demand-side bidding in the NZEM, so we will omit them from further discussion.)

The offers of generation made by generators to the ISO take the form of *offer stacks*. These are piecewise constant functions defining the amount of power offered at up to five different prices that may be chosen by the generator m. We can represent the offer stack for generator m by the (step) function $C_m(x)$. In the New Zealand market the generator offer functions C_m are not publicly known at the time of dispatch, but are published the following day. These data are made available as part of a Electricity Market Information (EMI) system supported by the New Zealand Electricity Authority [12].

All the prices in the wholesale electricity market in New Zealand are computed by the ISO using a linear programming model called "Schedule Price and Dispatch" or SPD. This represents the New Zealand transmission network by a DC-load flow model. The full version of SPD includes constraints that ensure voltage support, $N-1$ security for line failures, and meet requirements for spinning reserve that are dispatched at the same time (see [1]). The New Zealand Electricity Authority supports a publicly available GAMS model called vSPD [14]. The EMI system archives historical input files that when run on the model will reproduce historical prices and dispatch exactly. This enables very precise historical simulations to be run using counterfactual assumptions.

The essential features of SPD can be described mathematically using a DC-Load flow model formulated in the generic network model shown in Figure 1. For each node i the set $\mathcal{O}(i)$ defines all the generators at node i, where generator m can supply any quantity $q_m \in Q_m$. The demand at node i is denoted D_i. This gives the following market dispatch model:

$$\begin{array}{lllll} \text{MP1:} & \text{minimize} & \sum_i \sum_{m\in\mathcal{O}(i)} \int_0^{q_m} C_m(x)dx & & \\ & \text{s.t.} & g_i(y) + \sum_{m\in\mathcal{O}(i)} q_m = D_i, & [\pi_i] & i \in \mathcal{N}, \\ & & q_m \in Q_m, \quad m \in \mathcal{O}(i), & & i \in \mathcal{N}, \\ & & y \in Y. & & \end{array}$$

At the optimal solution to MP1, the shadow price π_i on the flow balance constraint at node i defines the locational marginal price. This is the price at which energy is traded at this location. The components of the vector y measure

was traded; the remaining 20% by bilateral contracts.

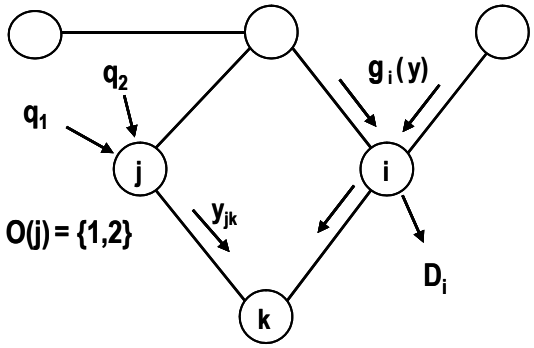

Figure 1: Generic network model illustrating notation

the flow of power in each transmission line. We denote the flow in the directed line from i to k by y_{ik}, where by convention we assume $i < k$. (A negative value of y_{ik} denotes flow in the direction from k to i.) We require that this vector lies in the convex set Y, which means that each component satisfies the thermal limits on each line, and satisfies loop flow constraints that are required by Kirchhoff's Law. The function $g_i(y)$ defines the amount of power arriving at node i for a given choice of y. This notation enables different loss functions to be modelled. For example, if there are no line losses then we obtain

$$g_i(y) = \sum_{k<i} y_{ki} - \sum_{k>i} y_{ik}.$$

With quadratic losses we obtain

$$g_i(y) = \sum_{k<i} y_{ki} - \sum_{k>i} y_{ik} - \sum_{k<i} \frac{1}{2} r_{ki} y_{ki}^2 - \sum_{k>i} \frac{1}{2} r_{ik} y_{ik}^2.$$

In SPD the quadratic losses are modelled as piecewise linear functions of arc flow which enables MP1 to be solved as a linear program (at least when losses are minimized by the optimal solution).

Bids and offers start 36 hours before the actual trading period. Up to 4 hours (pre-dispatch) before the trading period starts, a forecast price is calculated to guide participants in the market. From 4 hours to the start of the trading period every half hour a *dispatch price* is calculated (and communicated). Two hours before the start of the trading period, bids and offers for the period in question are locked in. From that point onwards any new prices reflect the ISO's adjustments in load forecasts and system availability.

During the half hour period the ISO publishes a new real-time price every 5 minutes and a time-weighted 30-minute average price. The real-time prices are used by some large direct-connect consumers to adapt their demand. The above

prices are a guide only, as the final prices are calculated ex-post (normally noon the following day, unless there are irregularities or disputes) using the offer prices as established 2 hours before the trading period, and volumes metered during the trading period.

As mentioned above SPD (and vSPD) include constraints that ensure voltage support, $N-1$ security for line failures, and meet requirements for spinning reserve that are dispatched at the same time as energy offers. We assume that voltage support and security constraints are relatively unaffected by the dispatch, and so in any simulation of a historical trading period we assume the constraints that applied at the time.

Spinning reserve does depend on the dispatch, and can have a large effect on prices, so we attempt to account for this in our counterfactual models. Spinning reserve protects the system from a frequency collapse if a large thermal unit or transmission line fails. At the beginning of every run the system operator takes the current dispatch and runs an AC simulation (called Reserve Management Tool or RMT) to estimate the levels of fast response (6-second) and sustained response (60-second) spinning reserve that would be required should a large unit (or the inter-island HVDC link) fail. The outputs of RMT are levels of freely available reserve and automatic load shedding, and the level of extra reserve that must be supplied by market participants in each island, who offer quantities of reserve to the market at prices of their own choosing. Details can be found in [1].

3 The models

Our study makes use of a similar suite of models as defined in [23]. The major difference in comparison with previous work is that we now use a full representation of the New Zealand high voltage transmission sytem as represented by vSPD. The models we use are:

vSPD:	Dispatch model solved over one trading period;
HydrovSPD:	Daily dispatch model including river chains;
DOASA:	A stochastic planning model solved over one year;

We examine a counterfactual proposal that supposes that the national electricity system is controlled centrally by a system planner who solves DOASA every two weeks in a rolling horizon fashion with updated data. The output from DOASA is used to determine water values for the model HydrovSPD that is solved sequentially over 14 days between solves of DOASA. The outcomes of the model HydrovSPD are then compared with observed outcomes in the wholesale

market as computed in vSPD. The details of this process are defined in publications that can be downloaded from the online companion [8]. We digress only briefly here to give an overview of the process.

3.1 HydrovSPD

To investigate the dispatch of hydroelectricity over the course of a day a national river-chain dispatch and nodal pricing model (HydrovSPD) combines offers from generation plant with river scheduling constraints over 48 half-hour trading periods, $p = 1, 2, ..., 48$. A diagram of the location of the river chains modelled is shown in Figure 2.

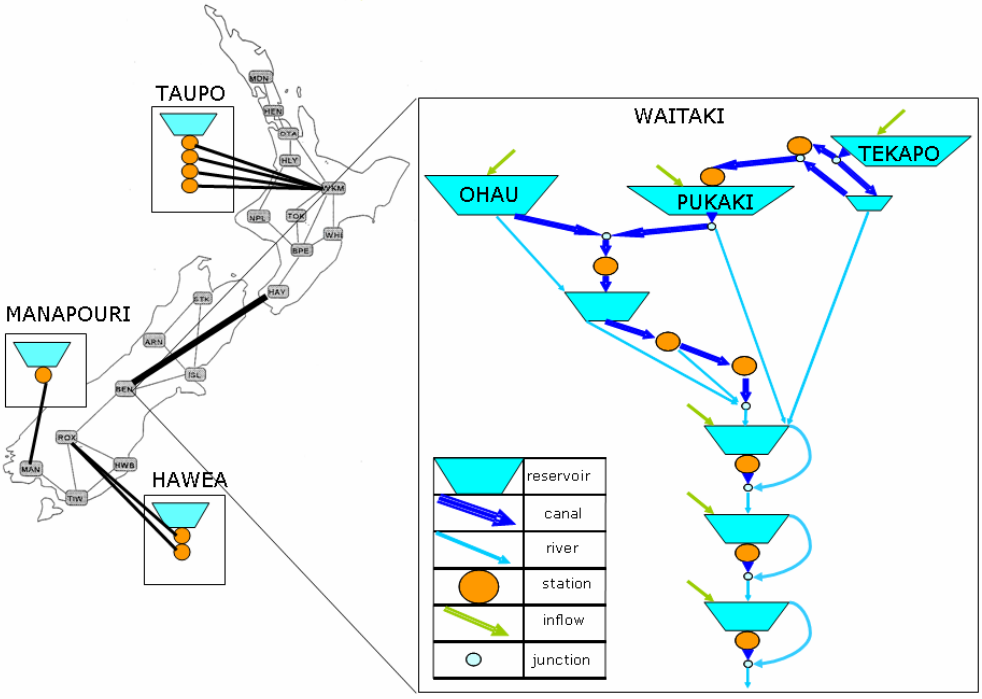

Figure 2: Approximate network representation of New Zealand electricity network showing main hydro-electricity generators

In the model we discriminate between thermal generation f_m, $m \in \mathcal{F}(i) \subseteq \mathcal{O}(i)$, and hydro generation $\gamma_m h_m$, $m \in \mathcal{H}(i) \subseteq \mathcal{O}(i)$. The parameter γ_m, which varies by generating station m, converts flows of water $h_m(p)$ into electric power. We denote the set of trading periods by $\mathcal{P} = \{1, 2, \ldots, 48\}$[4], and add the argument

[4] $\mathcal{P}$ can have 46 or 50 trading periods on days in which daylight saving changes.

p to all variables.

The storage in a reservoir or headpond r is denoted by x_r. The initial storage at the start of period $p = 1$ is given by the vector $\bar{x}$. The water balance constraints in each period are represented by

$$x_r(p+1) = x_r(p) - A_{rm}(h_m(p) + s_m(p)) + \omega_r(p),$$

where $x_r(p)$ is the storage in reservoir r at the start of period p, $s_r(p)$ denotes its spill in period p, and $\omega_r(p)$ is the uncontrolled inflow into the reservoir in period p. All these are subject to capacity constraints. (In some cases we also have minimum flow constraints that are imposed by environmental resource consents.) The node-arc incidence matrix A represents all the river-valley networks, and subtracts controlled flows that enter a reservoir from upstream from those that leave a reservoir by spilling or generating electricity. In other words row r of $A(h(p)+s(p))$ gives the total controlled flow out of the reservoir (or river junction) represented by row r, this being the release and spill of reservoir r minus the sum of any immediately upstream releases and spill.

We differentiate between large storage reservoirs $r \in \mathcal{R}$ and small headponds $r \in \mathcal{S}$. We require small headponds to end the day 50% full, while the marginal value of water in the large storage reservoirs is calculated using a piecewise linear convex cost-to-go function $\theta(x)$, defined by cutting planes

$$\theta(x) = \max_{k \in \mathcal{K}} \{\alpha^k + \sum_{r \in \mathcal{R}} \beta_r^k x_r(49)\}.$$

Here the values of α^k and β_r^k are determined from the output of a longer term model. If at some x we have $\theta(x) = \alpha^\ell + \sum_{r \in \mathcal{R}} \beta_r^\ell x_r(49)$ then $-\beta_r^\ell$ defines the marginal value of water in reservoir r at the end of the day. Marginal water values defined by these cutting planes will be reflected in the counterfactual energy prices $\pi_i(p)$ determined for each trading period and location by solving HydrovSPD($\bar{x}$).

This is formulated as:

$$\min \quad \sum_{p\in\mathcal{P}} \sum_{i\in\mathcal{N}} \sum_{m\in\mathcal{F}(i)} \phi_m f_m(p) + \theta$$

$$\text{s.t.} \quad g_i(y(p)) + \sum_{m\in\mathcal{F}(i)} f_m(p) + \sum_{m\in\mathcal{H}(i)} \gamma_m h_m(p) = D_i(p), \quad [\pi_i(p)] \qquad i \in \mathcal{N},\ p \in \mathcal{P},$$

vSPD constraints: e.g. security, spinning reserve, voltage support, loop flow, ramping

$$0 \le f_m(p) \le a_m, \quad m \in \mathcal{F}(i),\ i \in \mathcal{N},\ p \in \mathcal{P},$$

$$x_r(p+1) = x_r(p) - A_{rm}(h_m(p) + s_m(p)) + \omega_r(p),\ r \in \mathcal{R} \cup \mathcal{S},\ p \in \mathcal{P},$$

$$0 \le h_m(p) \le b_m, \quad 0 \le s_m(t) \le c_m, \quad m \in \mathcal{H}(i),\ p \in \mathcal{P},$$

$$0 \le x_r(p) \le w_r, \quad r \in \mathcal{R} \cup \mathcal{S}, \quad p \in \mathcal{P},$$

$$\alpha^k + \sum_{r\in\mathcal{R}} \beta_r^k x_r(49) \le \theta,\ k \in \mathcal{K},$$

$$x_r(49) \ge 0.5 w_r(1), \ \ r \in \mathcal{S}, \quad x_r(1) = \bar{x}_r,\ r \in \mathcal{R} \cup \mathcal{S}.$$

3.2 The medium-term hydro model: DOASA

To investigate the dispatch of hydroelectricity over the course of a year, a hydro-thermal release policy must be determined. This involves the solution of a large-scale stochastic dynamic programming model which is defined as follows. Let $x\,(t)$ denote the reservoir storage at the beginning of week t, and let $C_t(x, \omega(t))$ be the minimum expected fuel cost to meet electricity demand in weeks $t, t+1, \ldots, T$, when reservoir storage $x(t) = x$ and week t's inflow is known to be $\omega(t)$. Here

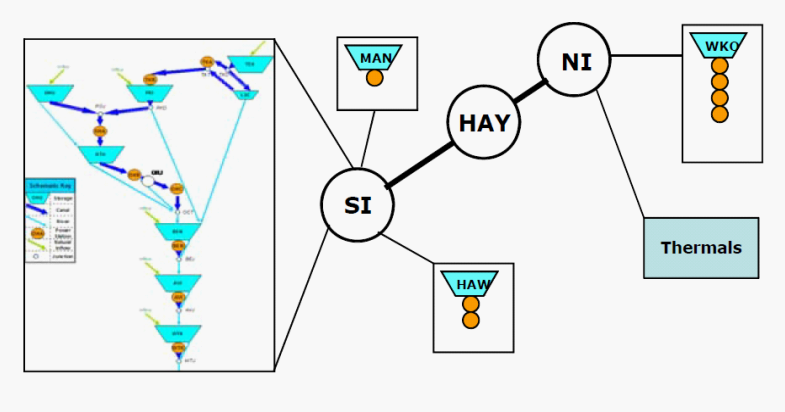

Figure 3: The 3 node transmission network and major generators in DOASA.

$C_t(x, \omega(t))$ is the optimal solution value of the mathematical program:

$$
\begin{aligned}
\mathrm{P}_t(x,\omega): \quad \min \quad & \textstyle\sum_{i\in\mathcal{N}} \sum_{m\in\mathcal{F}(i)} \phi_m f_m(t) + \mathbb{E}[C_{t+1}(x(t+1), \omega(t+1))] \\
\text{s.t.} \quad & g_i(y(t)) + \textstyle\sum_{m\in\mathcal{F}(i)} f_m(t) + \sum_{m\in\mathcal{H}(i)} \gamma_m h_m(t) = D_i(t), \quad i \in \mathcal{N}, \\
& x(t+1) = x - A(h(t) + s(t)) + \omega(t), \\
& 0 \le f_m(t) \le a_m, \quad m \in \mathcal{F}(i),\, i \in \mathcal{N}, \\
& 0 \le h_m(t) \le b_m, \quad 0 \le s_m(t) \le c_m, \quad m \in \mathcal{H}(i), \\
& 0 \le x_m(t) \le r_m, \quad m \in \mathcal{H}(i),\, i \in \mathcal{N}, \quad y \in Y.
\end{aligned}
$$

To solve this, we use the DOASA code [20] which is based on the SDDP technique of Pereira and Pinto [16]. This approximates $\mathbb{E}[C_{t+1}(x(t+1), \omega(t+1))]$ using a polyhedral function defined by cutting planes that is updated using samples of the inflow process.

The DOASA model uses weekly stages. A calendar year is divided into 52 weeks. A plan year is typically a year of 52 weeks with the starting week chosen to be a particular week in the plan year and a fixed number of weeks used. Historical inflows are sampled from a file that records weekly inflows as described below. The New Zealand electricity system is represented as shown in Figure 3.

Weekly demand is represented by a load duration curve with three blocks. These are called peak, off-peak and shoulder. We have chosen peak hours to be 6am-8am and 6pm-8pm weekdays, shoulder hours to be 8am-6pm and 8pm-10pm weekdays, and offpeak hours to be the other hours in the week. The total

demand in MW in each node i is then averaged over these trading periods to get a total demand rate $D_i(b,t)$ for each block. The energy requirement in node i for each block b in week t will be its duration $T(b,t)$ times the average demand rate $D_i(b,t)$ for this block.

The choice of what data to include in demand is a delicate matter. Publicly available demand figures (e.g. those in the EMI Data Set [12]) make various assumptions about what embedded generation and demand is included. These must be carefully studied to ensure that demand is not overlooked or double counted. The DOASA model aggregates demand to three locations (SI, HAY, NI) representing the South Island, the lower North Island, and the upper North Island, and allows transfers between these regions limited by line capacities. This means that aggregating demand in each region will ignore the intra-regional losses, implying that the regional totals of historical demand will underestimate the true demand to be met by generation and net imports to SI, HAY, and NI. So some inflation of total demand is needed to account for these losses. If the aggregation is carried out geographically then ignoring line losses might also bias the generation mix in the dispatch towards geographically close (yet electrically distant) plant.

The aggregate (into SI, HAY, and NI) of historical dispatch of the large generators can be used as a proxy for the demand adjusted for losses. We ignore the generation supplied by small generators as long as demand is adjusted for this, and compute the total generation of large generators in each region (SI, HAY, and NI) in each trading period using vSPD, and then add to this generation the net import minus export of power through transmission lines joining the region to its adjacent region. The result will give the half-hourly demand in the region satisfied by large generators and transfers between the regions. This is then transformed into load blocks for each week and used as the demand to be met by large generators. Of course this means that small fixed generators (not included in the demand estimation) should have their generation fixed at zero in the DOASA input data (as demand has effectively been reduced by these values).

A precise description of the demand calculation is as follows. Index all the generators represented in DOASA by $g \in \mathcal{G}$. Denote the regions (SI, HAY, and NI) by indices S, H and N. The first step in computing demand is to determine what vSPD nodes lie in each region. The boundaries of these regions can be defined somewhat arbitrarily, although Cook Strait is one obvious boundary between HAY and SI. In its optimization stage model, DOASA does not model transmission losses between the regions but does model transmission constraints. So constraints on the HVDC and power transfer between Wellington and Auckland can be modelled.

Let $\mathcal{N}_S$, $\mathcal{N}_H$, and $\mathcal{N}_N$ define the vSPD nodes corresponding to each region. Based on the transmission lines in vSPD, we define four sets of transmission

variables

$$\begin{aligned}
\mathcal{L}_{SH} &= \{\text{transmission lines directed from a node in } \mathcal{N}_S \text{ to a node in } \mathcal{N}_H\} \\
\mathcal{L}_{HS} &= \{\text{transmission lines directed from a node in } \mathcal{N}_H \text{ to a node in } \mathcal{N}_S\} \\
\mathcal{L}_{HN} &= \{\text{transmission lines directed from a node in } \mathcal{N}_H \text{ to a node in } \mathcal{N}_N\} \\
\mathcal{L}_{NH} &= \{\text{transmission lines directed from a node in } \mathcal{N}_N \text{ to a node in } \mathcal{N}_H\}.
\end{aligned}$$

The generators $g \in \mathcal{G}$ in each region are treated as if they are in a single location. We compute the total generation in each region in each trading period in each day of year t, by running vSPD with the GDX file for the trading peiods in that day. This gives us 365 days, each containing 48 periods (except for leap years and daylight savings days.). For each generator g let

$$q_g(p) = \text{generation in MWh computed by vSPD for period } p$$

and

$f_l(p) =$ the line flow variable (MWh) for line l as computed by vSPD for period p.

Here $f_l(p)$ does not include the losses incurred, half at each endpoint of the line. For these we define the flow leaving the node at the start of the line to be $f_l^+(p)$ and the flow arriving at the end of the line to be $f_l^-(p)$. Irrespective of the sign of $f_l(p)$ it follows that

$$\begin{aligned}
f_l^+(p) &= f_l(p) + \frac{\alpha}{2}(\text{flow loss}) \\
f_l^-(p) &= f_l(p) - \frac{\alpha}{2}(\text{flow loss}),
\end{aligned}$$

where α is a parameter we can vary. Then we compute the total generation in a node to be

$$q_r(p) = \sum_{g \in r} q_g(p).$$

Now adding flows we get

$$\begin{aligned}
f_{SH}(p) &= \sum_{l \in \mathcal{L}_{SH}} f_l^+(p) - \sum_{l \in \mathcal{L}_{HS}} f_l^-(p) \\
f_{HS}(p) &= \sum_{l \in \mathcal{L}_{HS}} f_l^+(p) - \sum_{l \in \mathcal{L}_{SH}} f_l^-(p) \\
f_{HN}(p) &= \sum_{l \in \mathcal{L}_{HN}} f_l^+(p) - \sum_{l \in \mathcal{L}_{NH}} f_l^-(p) \\
f_{NH}(p) &= \sum_{l \in \mathcal{L}_{NH}} f_l^+(p) - \sum_{l \in \mathcal{L}_{HN}} f_l^-(p).
\end{aligned}$$

Observe that

$$\begin{aligned} & f_{SH}(p) + f_{HS}(p) + f_{HN}(p) + f_{NH}(p) \\ = & \sum_{l \in \mathcal{L}_{SH} \cup \mathcal{L}_{HS} \cup \mathcal{L}_{HN} \cup \mathcal{L}_{NH}} \left(f_l^+(p) - f_l^-(p) \right) \\ = & \ \alpha \sum_{l \in \mathcal{L}_{SH} \cup \mathcal{L}_{HS} \cup \mathcal{L}_{HN} \cup \mathcal{L}_{NH}} \text{flow loss.} \end{aligned}$$

The regional demand in period p is now estimated for each region to be

$$\begin{aligned} d_S(p) &= q_S(p) - f_{SH}(p) \\ d_H(p) &= q_H(p) - f_{HS}(p) - f_{HN}(p) \\ d_N(p) &= q_N(p) - f_{NH}(p). \end{aligned}$$

Observe that with this definition, the national demand in period p is

$$\begin{aligned} \sum_{k \in \{S,H,N\}} d_k(p) &= \sum_{k \in \{S,H,N\}} q_k(p) - f_{SH}(p) - f_{HS}(p) - f_{HN}(p) - f_{NH}(p) \\ &= \sum_{k \in \{S,H,N\}} q_k(p) - \alpha \left(\text{total flow loss from inter-regional flow} \right). \end{aligned}$$

The rationale behind this choice is that our backtest is intended to compute an optimal dispatch to meet some observed demand. The demand used in vSPD makes various assumptions about embedded generation and wind that are often difficult to verify, especially for past years. One approach is to backtest the allocation of generation amongst the large generators. If transmission losses are ignored in DOASA (the default) then we choose $\alpha = 0$ in the above analysis which yields national demand in period p equal to $\sum_{k \in \{S,H,N\}} q_k(p)$. DOASA will possibly reallocate generation amongst the generators $g \in \mathcal{G}$ to be cheaper to meet this demand. This might give different transfers between regions than those observed in vSPD.

If transmission losses between regions are modelled in DOASA then we choose $\alpha = 1$, which yields national demand in period p equal to $\sum_{k \in \{S,H,N\}} q_k(p)$ minus national losses. Now DOASA will possibly reallocate generation to be cheaper to meet this demand. This might give different transfers between regions than those observed in vSPD, to not only allow cheaper generation but potentially give lower transmsission losses. Once $d_k(p)$ has been computed for the three regions for every trading period, we take $\sum_{k \in \{S,H,N\}} d_k(p)$ for each p in this week, and allocate the demand to one of three blocks peak, shoulder, and offpeak. The assignment of trading period to block is made a priori and fixed. Thus computing the energy (MWh) in each block in each week is obtained by summing $d_k(p)$ over periods p corresponding to the block.

In meeting demand, in case of supply shortages, load shedding (in MW) is allowed at high costs. The costs depend on the type of customers and amount of reduction (in \$/MWh). Load in each node is divided into three sectors to represent different types of customers, which are industrial, commercial and residential, and each sector has some distribution in each island. The default proportions are the proportions of consumption in 2015 adjusted to higher commercial and residential proportions in the North Island due to a denser population, and to a higher industrial proportion in the South Island due to an aluminium smelter. Although these proportions change over the years, for simplicity we have assumed they are constant. Each sector is then divided into three segments to represent the amount of reduction, namely 5%, 5% and 90%. The third segment represents unplanned interruption of power supply. The cost for load shedding is called the *value of loss load*, or VOLL, in the electricity industry. The VOLLs for the industrial sector are set to be lower than the other two and the VOLLs increase over segments in each sector. We assume that up to 10% reduction in load can be achieved at a relatively low cost, but the value of unplanned interruption is very high (\$10,000/MWh)[5].

The DOASA model assumes that six reservoirs, Manapouri, Hawea, Ohau, Pukaki, Tekapo and Taupo, can store water from week to week. The release of this water through generating stations is controlled. The hydroelectric stations in other parts of the system are treated as run-of-river plant with limited intra-week flexibility. It is important to note that we assume inflows to the main catchments are stagewise independent. These are sampled from the historical weekly inflow series available on the EMI site [12][6]. Full details of the DOASA model for this study can be found in the online companion [8] to this paper.

The solution to $\mathrm{P}_1(x_1, \omega(1))$ defines a set of thermal plants to run and a set of linear functions (or *cuts*) whose pointwise maximum approximates $\mathbb{E}[C_2(x(2), \omega(2))]$. Indeed the DOASA code yields an outer approximation to $\mathbb{E}[C_{t+1}(x(t+1), \omega(t+1))]$ at each stage t, and so this defines a policy at this stage by setting $\bar{x} = x(t)$ and solving HydrovSPD($\bar{x}_r$) in which the constraints

$$\alpha^k + \sum_{r \in \mathcal{R}} \beta_r^k x_r(49) \leq \theta, \quad k \in \mathcal{K},$$

are determined by the cuts defining $\mathbb{E}[C_{t+1}(x(t+1), \omega(t+1))]$. In our experimental

[5]The value of \$10000/MWh is open to some debate. The NEM in Australia applies a VOLL that is indexed to inflation. In 2018-19 the value was \$14,500 [25]. Our choice of \$10,000 is based on the capital cost of approximately \$1M/MW for peaking plant [15] that would be required 5 hours per year over 20 years.

[6]In any year y we select inflows for each catchment in the years from $y - 35$ to $y - 1$ as equally likely random outcomes in each week. Thus for any year we have 35 (vector) outcomes per stage giving a stagewise independent scenario tree for DOASA of 35^{51} scenarios.

setup we use the cuts defining $\mathbb{E}[C_2(x(2), \omega(2))]$ for HydrovSPD($\bar{x}_r$) over the first week of the fortnight, and the cuts defining $\mathbb{E}[C_3(x(3), \omega(3))]$ over the second week of the fortnight. The experiments that we carry out below use 3000 cuts per solve.

In summary the experimental procedure is as follows:

EMBER simulation

Given reservoir levels $x(1)$ solve a 52 week hydrothermal scheduling problem using DOASA.

1. Set $t = 1$.
2. Solve a hydrothermal scheduling problem over weeks $\{t, \ldots t + 51\}$ using DOASA.
3. Set $\bar{x} = x(t)$.
4. For $d = 1$ to 7,
   (a) Select α^k and β^k, $k \in \mathcal{K}$, from the cut intercepts and slopes approximating $\mathbb{E}[C_{t+1}(x(t + 1), \omega(t + 1))]$;
   (b) Solve HydrovSPD($\bar{x}_r$);
   (c) Set $\bar{x} = x_r(49)$.
5. For $d = 8$ to 14,
   (a) Select α^k and β^k, $k \in \mathcal{K}$, from the cut intercepts and slopes approximating $\mathbb{E}[C_{t+2}(x(t + 2), \omega(t + 2))]$;
   (b) Solve HydrovSPD($\bar{x}_r$);
   (c) Set $\bar{x} = x_r(49)$.
6. Set reservoir levels to $\bar{x}$, set $t = t + 2$, and go to step 2.

4 Market comparison

We now decribe a set of experiments that were carried out using data from 2012. Given costs per MWh of gas, diesel, and coal generation it is possible to compute the cost of fuel required to generate the electricity dispatched by the wholesale market in each historical half hour. This cost can be compared with the same cost as optimized by a central plan.

There are several difficulties with such an approach. The first of these concerns dispatch that has limited control. Examples of such dispatch is that from cogeneration, geothermal plant, run-of-river hydro and wind. Although these have low marginal cost, their availability is subject to the vagaries of inflows and wind, and so we cannot centrally dispatch these in a counterfactual. We choose to fix all cogeneration, geothermal generation wind generation, embedded generation, run-of-river generation and small hydro plant at their historical levels. This leaves the large hydro systems (Manapouri, Clutha, Waitaki and Waikato) available for control along with the major thermal plants (Huntly (4 units plus e3p and P40), Otahuhu, Stratford, and Whirinaki). These are the only generators that we allow to offer energy within our model.

In reporting costs, our measure will be the cost of fuel burned by these five plants. The fuel used in a thermal power station is coal, natural gas or diesel, as shown in Table 1. Coal is supplied from stockpiles that are restocked under long-term contracts. Coal costs are assumed to be constant at \$4/GJ. Natural gas is supplied by take-or-pay contracts. It is assumed that in social planning the supply can be secured and the costs are wholesale prices. The quarterly average prices of natural gas for wholesale use are available from the Ministry of Business Innovation and Employment (MBIE) at webpage [13]. The quarterly average prices of diesel for commercial use in [13] are used as the costs of diesel. The quarterly average prices are converted into real dollars in December 2015 and the costs of CO2 emissions (based on the current CO2 price expressed in 2015 terms[7]) are added. This gives the fuel and carbon cost of coal, diesel and gas as shown in Table 2. The weekly costs in the four 13 weeks in each year are the prices of the four quarters in the year respectively. The short-run marginal cost for any plant can be obtained by multiplying the heat rate (see Table 1) by the fuel cost, and adding a variable operations and maintenance cost. These SRMC values are similar to those assumed by other authors (e.g. [5, page 6, Table 2]).

Power station	Heat rate (GJ/MWh)	Fuel
Huntly main 1-4	10.3	coal
Huntly e3p	7.2	natural gas
Huntly peaker	9.8	natural gas
Otahuhu B	7.45	natural gas
Stratford peakers	9.5	natural gas
Taranaki Combined Cycle	7.6	natural gas
Whirinaki	11	diesel

Table 1: Thermal power stations and heat rates

[7]The costs of CO2 permits are automatically adjusted in our model for regulatory relaxations (e.g. 1 for 2 schemes) that were applied in some months.

($/GJ (2015))		MBIE fuel costs		
year	qtr	coal	diesel	gas
2012	1	$ 4.35	$31.27	$ 6.72
2012	2	$ 4.30	$30.65	$ 6.85
2012	3	$ 4.22	$31.33	$ 6.82
2012	4	$ 4.12	$30.88	$ 6.28
2013	1	$ 4.19	$30.73	$ 6.93
2013	2	$ 4.18	$28.45	$ 6.53
2013	3	$ 4.28	$30.11	$ 6.50
2013	4	$ 4.32	$29.28	$ 6.57

Table 2: MBIE fuel and CO2 costs for thermal generation in December 2015 NZ dollars (Source [13]).

To enable a fair comparison with market outcomes, we have de-rated stations at which plant have been removed for planned maintenance. The weekly de-rating of a generator is taken by default to be the outage amount given in the POCP database [24]. If no data are provided for an offering generator in [24], we de-rate its plant capacity in a given week of the year by the difference between its nominal capacity and the average total offer quantity made by the plant in the same week in previous years. The schedule in POCP defines the starting and end time of scheduled maintenance for generators, which includes the offering generators and all small and run-of-river generators that we consider as fixed (e.g. Tokaanu, Rangipo and Waikaremoana). The HVDC line capacity is treated as fixed in DOASA, but will be assigned its vSPD value during simulation.

As discussed above we also make use of costs for unserved load. These depend on the type of customer and the amount of load reduction as shown in Table 3.

($/MWh)	Up to 5%	Up to 10%	VOLL	North Island	South Island
Industrial	$ 1,000	$ 2,000	$ 10,000	0.36	0.58
Commercial	$ 2,000	$ 4,000	$ 10,000	0.27	0.17
Residential	$ 2,000	$ 4,000	$ 10,000	0.37	0.24

Table 3: Load reduction costs ($/MWh) and proportions of each load that is industrial, commercial, and residential load.

The last two columns of Table 3 show the proportion of load of each type in each island. This shows that (rounded to the nearest percentage) 58% of South Island load is industrial, 17% commercial, and 24% is residential. The costs (in NZ$/MWh) of shedding load are also shown in Table 3. We assume that up to 10% reduction in load can be achieved at a relatively low cost, but the value of unplanned interruption (or reduction above this level) is very high ($10,000/MWh).

Therefore if, for example, load in the South Island was 1000 MW, we could shed up to 5% of 580 MW at \$1000/MWh and at \$2000/MWh, we could shed 5% of 410MW (170MW commercial and 240 MW residential) plus a further 5% of 580MW industrial load.

5 Experiments

We now present the results of applying DOASA and HydrovSPD to data from the calendar years 2012 and 2013.

5.1 Risk neutral agents

The first set of experiments assumes that all agents are risk neutral, i.e. $\lambda = 0$. The generation-weighted average prices (GWAPs) in 2012 are shown for the South Island in Figure 4 and the North Island in Figure 5. The plot in purple shows historical GWAPs (expressed in 2015 dollars). The blue plot is generated by the *fixed hydro* counterfactual: this fixes all hydroelectric generation at its historical level and then runs vSPD in each trading period of the year, with spinning reserve constraints turned off, and thermal plant offers set to their historical maximum offer quantities offered at SRMC. The assumptions of fixed hydro essentially mirror those of the Wolak report [26]. One can see that the prices follow the marginal cost of thermal plant (when this is dispatched) and occasionally drop to very low values if thermal dispatch is off for much of the day. As mentioned in the introduction, fixing hydro generation at historical levels ignores the high costs of shortage that produce high prices when shortage is anticipated, even if it does not eventuate.

As mentioned above, in DOASA we choose to fix some small generation plant at their historical level. This is true for small hydro plant, geothermal generation, and wind generators. We also need to use historical Manapouri generation in HydrovSPD. This is because Manapouri has complicated nonconvex environmental constraints on its operation. Ignoring these leads to infeasible water releases that might underestimate the true costs imposed by the constraints.

To test the effect of fixing hydro generation at Manapouri to historical levels, we performed two runs. The first run, called *fixed Manapouri*, optimizes in DOASA the large storage catchments except for Manapouri that has its dispatch fixed at historical levels in DOASA (as well as in HydrovSPD). This gives the prices shown in the red curve. This shows some increases in price over fixed hydro. The green curve (*no reserve*) relaxes Manapouri generation in DOASA (using a simple decision rule to determine its release) but imposes its historical generation level only in HydrovSPD. This gives some further increase in prices.

Both increases happen in the later part of the winter when reservoir levels are low. Observe that there is no such increase in the fixed-hydro prices. Henceforth we assume in all runs that Manapouri generation in DOASA is determined using the decision rule, and fixed in HydrovSPD.

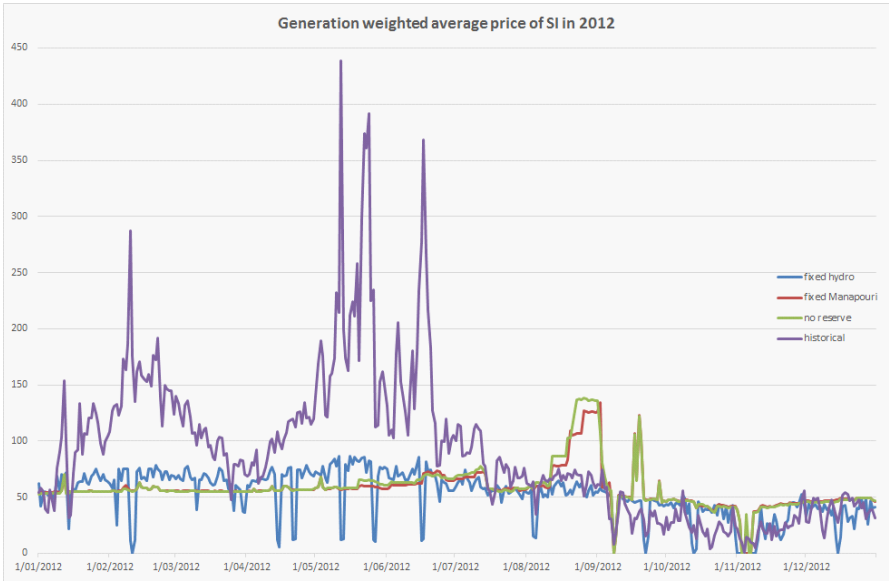

Figure 4: Risk neutral results for South Island in 2012.

It is possible that a further price rise might be obtained by including reserve offers and constraints in HydrovSPD. The difficulty here is that we do not have access to the approriate reserve parameters (from RMT) for the counterfactual models. Our assumption for reserve was that the risk was set by a failure of the HVDC link[8], and HVDC rampup was set to its historical level, and that all historical reserve was offered at zero cost. The results of cost and Ricardian rent (in 2015 dollars) with this reserve included are shown in Table 4. The rents and costs here are calculated using generation quantities and nodal prices only for those stations that are owned by the five largest electricity companies.

Some care is needed in computing prices and rents when spinning reserve is included. In some trading periods there is insufficient reserve offered in the counterfactual model to cover the HVDC risk. The amount of offered reserve required in the historical dispatch can be less because of more net free reserves. To overcome this we increase the amount of net free reserve available to HydrovSPD. The value of net free reserve supplied to HydrovSPD is the minimum amount required to meet an HVDC risk requirement if that is chosen to set the risk in a corresponding historical vSPD run. A reserve shortfall can give very high

[8]We assume that both DCCE and DCECE risks apply.

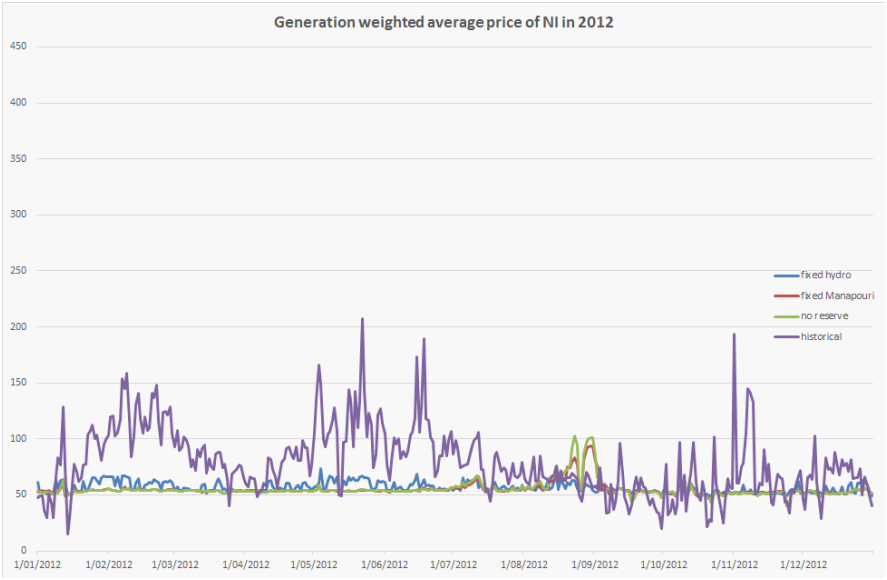

Figure 5: Risk neutral results for North Island in 2012.

prices ($500,000/MWh) in the counterfactual model to signal the infeasibility. In practice, such an outcome would be resolved by relaxing the reserve requirement, until the dispatch was feasible. We model this in the counterfactual as follows. The nodal prices in each run are capped at $10,000/MWh (our choice of VOLL) in trading periods with demand violation. If there is no demand violation in a trading period then the nodal price is capped at $1,000/MWh (99.8% of historical trading periods without demand violation have prices below $1,000/MWh).

		2012		2013	
		cost	rent	cost	rent
Fixed hydro	No reserve	$ 510	$ 1,509	$ 438	$ 1,421
$\lambda = 0.0$	No reserve	$ 456	$ 1,622	$ 361	$ 1,357
$\lambda = 0.0$	With reserve	$ 459	$ 1,677	$ 388	$ 1,641
Historical	With reserve	$ 517	$ 2,498	$ 444	$ 1,885

Table 4: Operational costs and Ricardian rents for risk neutral runs. (All figures in millions of 2015 dollars)

Observe that the rent in 2012 has increased from $1509 for fixed hydro to $1622 when optimizing hydro in DOASA to $1677 when reserve is included. In 2013 the corresponding figures are $1421 for fixed hydro to $1357 when optimizing hydro in DOASA to $1641 when reserve is included. It is intriguing that the fixed hydro model gives higher rents in 2013 than a model with more uncertainty. We shall discuss this later in the report.

5.2 Results for risk averse agents

The results above are for risk neutral agents. DOASA also allows us to compute risk-averse policies with varying levels of risk aversion. Risk is modelled using a nested dynamic risk measure (see [17]), in which the one-step risk measure is a convex combination of the expectation and worst case outcome of future fuel and shortage cost. In other words we use the one-step risk measure

$$\rho(Z) = (1-\lambda)\mathbb{E}[Z] + \lambda\mathbb{W}[Z]$$

where $\lambda \in (0,1)$, Z represents the random future cost, and

$$\mathbb{W}[Z] = \max\{Z(\omega)\}.$$

The dynamic version uses a nested form of ρ, where the risk averse certainty equivalent of a random stream of costs, say Z_1, Z_2, Z_3, is computed using a nested formulation, which would be $\rho(Z_1 + \rho(Z_2 + \rho(Z_3)))$ in this example. A straightforward procedure for implementing this within SDDP algorithms is described in [18]. If there are M scenarios, the measure $\rho(Z)$ is equivalent to weighting all scenarios with equal probability $\frac{(1-\lambda)}{M}$ except for the most expensive scenario which receives weight $\frac{1}{M}(M\lambda - \lambda + 1)$.

The DOASA model also assumes that inflows are stagewise independent, but modified by an adjustment to account for stagewise correlation. This is called *inflow spreading* by some modellers. We call it Dependent Inflow Adjustment (DIA). DIA may not be sufficient to produce the prices that arise from market expectations of shortages. We can use risk aversion to give policies that treat sequences of low inflows as more likely than what one would expect from stagewise independence.

To test this, we ran DOASA with increasing levels of risk aversion, choosing $\lambda = 0.5$ and $\lambda = 0.9$. These correspond to mild and extreme risk aversion respectively. To put this in context, a value of $\lambda = 0.9$ implies that the decision maker each week believes with 90% probability that the worst inflow observed in this week in the last 35 years will occur. This means more than a 50% chance that the next six weeks will be a sequence of the driest weeks out of the last 35 years. The cost and rent results from HydrovSPD (all including reserve) are in Table 5.

Risk averse &	2012		2013	
reserve	cost	rent	cost	rent
$\lambda = 0.0$	$ 459	$ 1,677	$ 388	$ 1,641
$\lambda = 0.5$	$ 484	$ 1,665	$ 406	$ 1,645
$\lambda = 0.9$	$ 497	$ 1,680	$ 416	$ 1,632
Historical	$ 517	$ 2,498	$ 444	$ 1,885
Difference	$ 33	$ 833	$ 38	$ 240

Table 5: Operational costs and Ricardian rents for risk averse runs

Risk aversion increases operating costs (as more fuel is burnt) in 2012 and 2013. Risk aversion increases rent in 2012, but has an ambiguous effect on rent in 2013. We plot the aggregated reservoir storage in 2012 and 2013 in Figure 6 and Figure 7. In 2012 risk aversion increases storage in both North and South Islands. The high risk-averse graph (red) tracks historical storage closely in the South Island, but hits its maximum (spilling) in the North Island for large parts of the second half of the year. In 2012, the mild risk-averse case uses water more aggressively in the South Island, but ends up with more water at the end of the year.

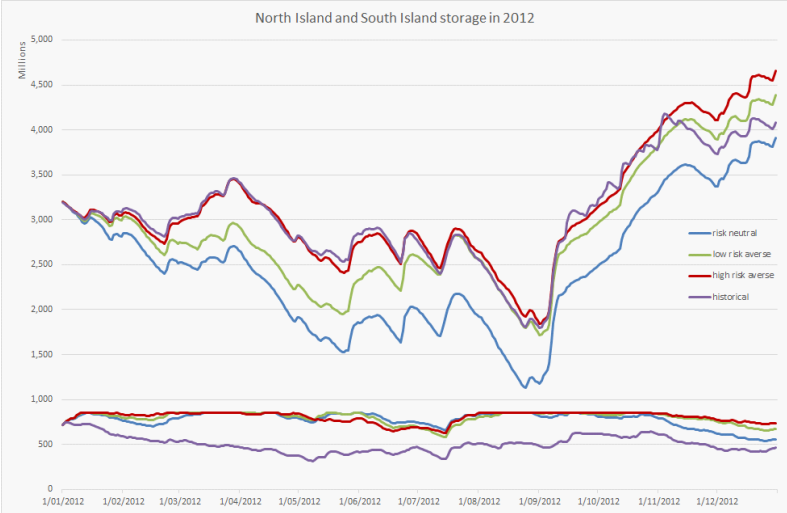

Figure 6: Reservoir storage in 2012. Risk neutral corresponds to $\lambda = 0$, low risk averse to $\lambda = 0.5$, and high risk averse to $\lambda = 0.9$.

We plot the generation-weighted average prices over 2012 in Figure 8 (South Island) and Figure 9 (North Island) to compare with the risk-neutral simulation. South Island prices have some increase earlier in the year for the mild risk-averse case (green) and more deviations in February and March in the extreme risk-averse case (red). It is difficult to explain these deviations especially in the early months of the year. The year 2012 was one of the driest on record. The wholesale market survived this year without requiring a savings campaign, leading to some commentators concluding that the market reforms of 2010 had been successful. Of course 2012 also had reduced demand (particularly from Christchurch). Taking this into account gives the counterfactual price trajectories shown in Figure 8 and Figure 9 which are well below historical prices, and similarly result in no shortages (as shown by the South Island and North Island storage trajectories in Figure 6). It is important to note that our results use a system risk measure,

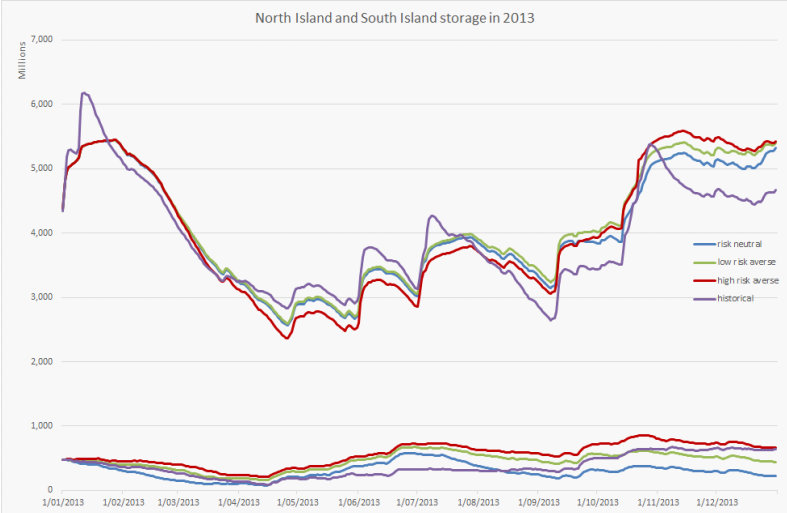

Figure 7: Reservoir storage in 2013. Risk neutral corresponds to $\lambda = 0$, low risk averse to $\lambda = 0.5$, and high risk averse to $\lambda = 0.9$.

which would be appropriate in a setting where agents can trade risk (see [19]). If the contracts that make this possible are not available or thinly traded, then one might expect some deviation from the system optimum, even if the market was perfectly comeptitive in other respects.

The results for the North Island show a similar pattern to the risk neutral results. Observe that North Island counterfactual prices are very volatile in the second half of 2012.

A final set of experiments were carried out with increased fuel costs to enable a comparison with the benchmarks based on the gas costs reported in [13]. The use of gas and coal by thermal plants is complicated by take-or-pay contracts. If a thermal generator holds a gas contract for more than they need then they might offer below a nomial fuel cost to burn the excess at an apparent loss. On the other hand a generator who is short of gas might regard the opportunity cost of gas to be higher than what was paid in a take-or-pay contract. Gas cost is further complicated by ownership. In 2012 the Genesis group had a 31% stake in the Kupe field[9]. This makes reported payments for gas for the group significantly lower than they would be otherwise, where the cost of gas as an operating expense for Huntly power station is interpreted as an opportunity cost, i.e. the foregone value of not selling it elsewhere. To use as a comparison, we were provided costs by First NZ Capital Securities Ltd[10] that result in higher variable costs as shown in Table 6.

[9]This increased to 46% in November 2016.

[10]We are grateful to Nevill Gluyas for providing us with these estimates of fuel prices.

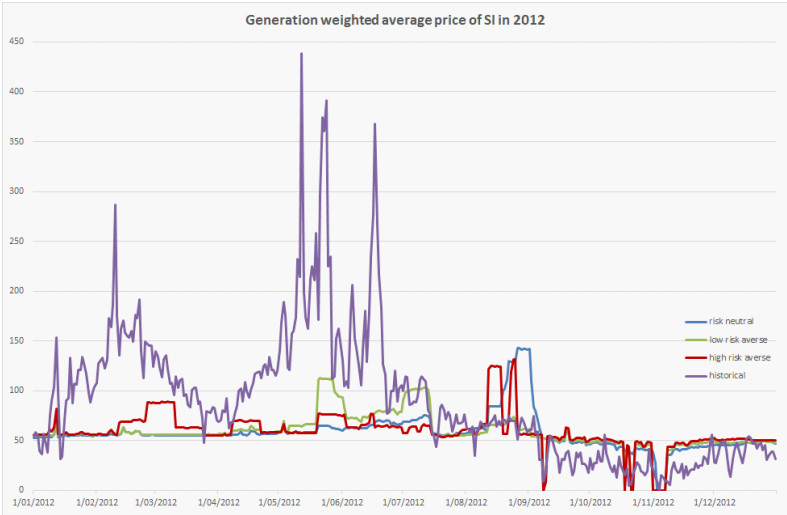

Figure 8: South Island generation weighted average prices in 2012

($/GJ (2015))		high fuel costs (generators)		
year	qtr	coal	diesel	gas
2012	1	$ 5.13	$ 31.27	$ 8.94
2012	2	$ 5.08	$ 30.65	$ 9.13
2012	3	$ 5.01	$ 31.33	$ 9.11
2012	4	$ 4.91	$ 30.88	$ 8.39
2013	1	$ 5.01	$ 30.73	$ 9.65
2013	2	$ 5.00	$ 28.45	$ 9.10
2013	3	$ 5.11	$ 30.11	$ 9.03
2013	4	$ 5.15	$ 29.28	$ 9.13

Table 6: Fuel plus CO2 costs ($/GJ) as estimated by FNZC.

With the larger fuel costs, HydrovSPD gives the storage trajectories shown in Figures 10 and 11. There appears to be little difference between these and the plots for Figure 6 and Figure 7, except that the high-cost counterfactual policies burn less gas earlier in the year.

Prices that are generated by the counterfactual policies and the rents accruing are shown below. The generation-weighted average prices are shown in Figure 12 for the South Island and Figure 13 for the North Island. These show generally higher values in the mild risk-averse and extreme risk-averse cases.

The values of these prices over each of 2012 and 2013 are compared in Table 7. We compute a comparison of time-weighted average prices in Table 8.

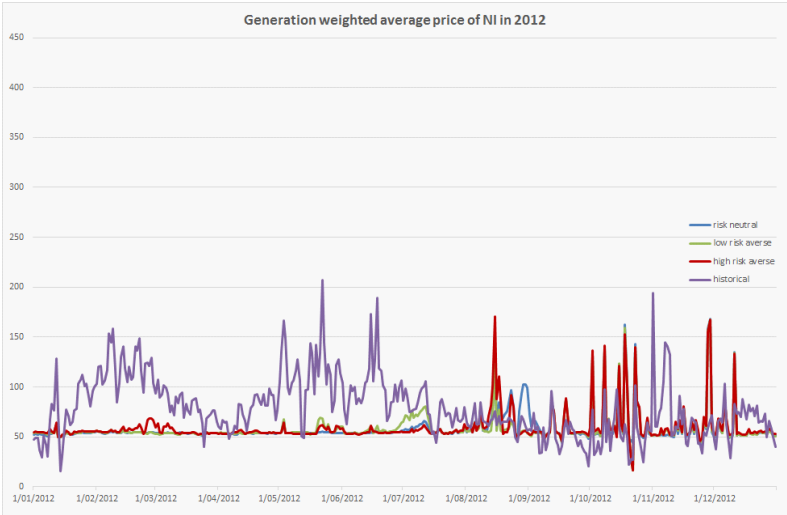

Figure 9: North Island generation weighted average prices in 2012

GWAP	low cost		high cost	
$/MWh	2012	2013	2012	2013
$\lambda = 0.0$	$ 57.77	$ 55.57	$ 73.20	$ 71.04
$\lambda = 0.5$	$ 58.09	$ 56.15	$ 73.12	$ 72.69
$\lambda = 0.9$	$ 58.85	$ 55.91	$ 77.06	$ 73.97
Historical	$ 81.38	$ 64.01	$ 81.38	$ 64.01
Difference	$ 23.29	$ 7.86	$ 8.26	-$ 8.68

Table 7: Differences in generation-weighted average prices for three counterfactual models compared with the market (Historical). Left-hand tables use MBIE fuel costs while right hand table uses NZFC estimated fuel costs. Difference is Historical minus Mild aversion.

			2012			2013	
Price ($/MWh)		OTA	HAY	BEN	OTA	HAY	BEN
low cost	$\lambda = 0.5$	$60.06	$61.50	$58.48	$71.52	$69.45	$41.36
low cost	$\lambda = 0.9$	$61.30	$62.38	$59.02	$80.54	$77.36	$42.23
high cost	$\lambda = 0.5$	$76.99	$76.86	$72.47	$90.09	$86.96	$56.69
high cost	$\lambda = 0.9$	$79.10	$80.38	$78.61	$89.41	$86.54	$59.90
Historical		$78.88	$81.47	$85.80	$67.55	$66.44	$58.73

Table 8: Time-weighted average prices at Otahuhu (OTA), Haywards (HAY), and Benmore (BEN) grid exit points. Counterfactual models with varying assumptions are compared with the market (Historical). Low-cost figures correspond to MBIE fuel costs while high-cost figures correspond to NZFC estimated fuel costs.

Finally the costs and rents are compared in Table 9.

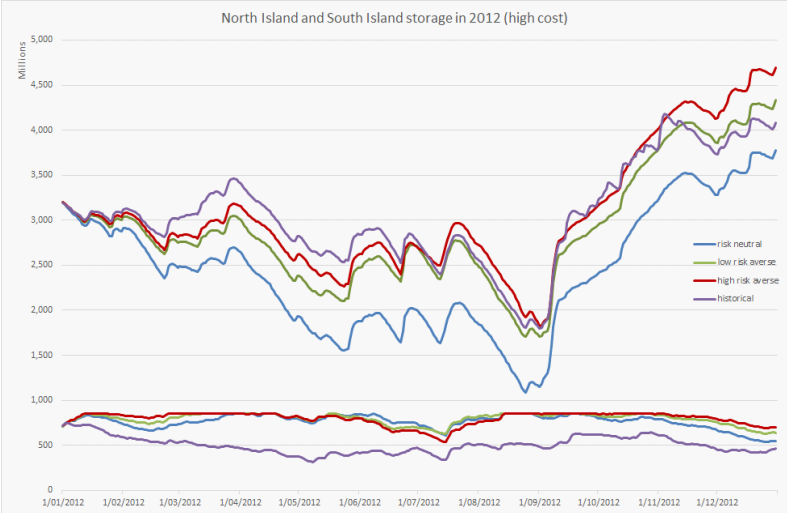

Figure 10: Reservoir storage 2012 with high fuel costs.

Risk averse &	2012 low cost		2013 low cost		2012 high cost		2013 high cost	
reserve	cost	rent	cost	rent	cost	rent	cost	rent
$\lambda = 0.0$	$ 459	$ 1,677	$ 388	$ 1,641	$ 557	$ 2,151	$ 478	$ 2,112
$\lambda = 0.5$	$ 484	$ 1,665	$ 406	$ 1,645	$ 588	$ 2,117	$ 499	$ 2,153
$\lambda = 0.9$	$ 497	$ 1,680	$ 416	$ 1,632	$ 610	$ 2,241	$ 520	$ 2,176
Historical	$ 517	$ 2,498	$ 444	$ 1,885	$ 651	$ 2,364	$ 583	$ 1,746
Difference	$ 33	$ 833	$ 38	$ 240	$ 63	$ 248	$ 84	-$ 407

Table 9: Operational costs and Ricardian rents for risk averse runs.

The rent differences with high fuel prices are smaller, and negative in 2013. The counterfactual models make substantially more rent in the high fuel -cost results compared with MBIE fuel costs, as higher nodal prices yield higher revenues, and the more expensive thermal generation is reduced by a social plan. This yields cost savings of $63 M and $84 M in 2012 and 2013 respectively. Observe that the historical time-weighted average Otahuhu and Haywards prices are lower than the counterfactual prices in 2013. There is evidence in the early part of this year that thermal generators were offering below marginal cost. This could be an artifact of being overcontracted in gas supply, or having large contract positions. As discussed in Ruddell et al [2], generators holding contracts have incentives to offer below marginal cost up to their contract quantities. If water turns out to be more plentiful than forecast, then thermal plant will often be dispatched below their contract quantities, thus depressing spot prices.

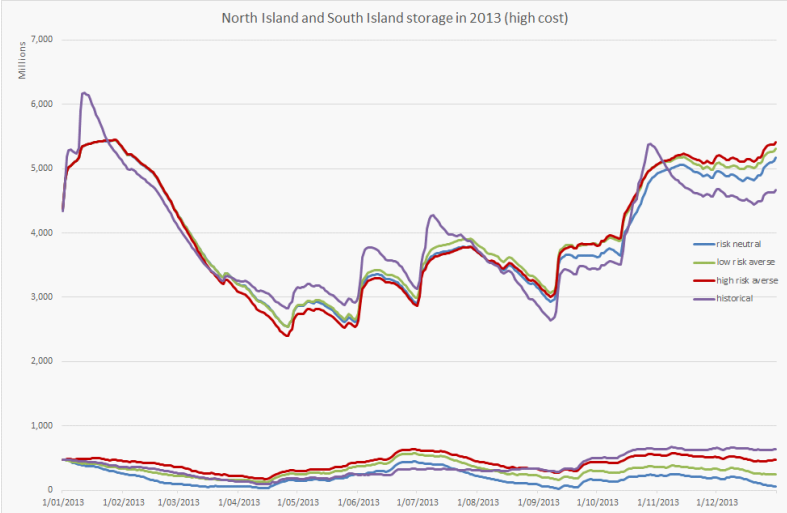

Figure 11: Reservoir storage 2013 with high fuel costs.

6 Discussion

In this paper we have described some experiments with stochastic optimization models of the New Zealand wholesale electricity market that provide counterfactual outcomes for competitive markets. The results arising from these models indicate that wholesale market outcomes have deviated from perfectly competitive benchmarks (at least in 2012 and 2013). When this happens, there is a loss in efficiency, as can be seen in the differences in cost reported in Table 5. Observe that these cost differences represent fuel and carbon costs only. As shown in Figures 10 and 11, the risk-averse counterfactual models leave large reservoirs in both islands fuller than what they were historically. So the historical market dispatch uses more water and more expensive thermal plant is run more often, and satisfies exactly the same demand every half hour as the counterfactual models.

It is important to recognize that our simulations are carried out (in HydrovSPD) using the full representation of the New Zealand wholesale electricity market. The policies that we simulate (apart from fixed hydro) do not anticipate future inflows, and so the outcomes represent an implementable counterfactual. Furthermore, in 2012 the counterfactual simulations require no savings campaigns. Commentators (see e.g [9]) have asserted that the lack of savings campaign in 2012 indicates that the wholesale market was working well by signalling potential shortages with high prices early in the year. Our models show that demand in 2012 can be met using much more modest price signals.

It is interesting to speculate on the causes of the differences in prices and rents observed in the market and the counterfactuals above. The Wolak report [26]

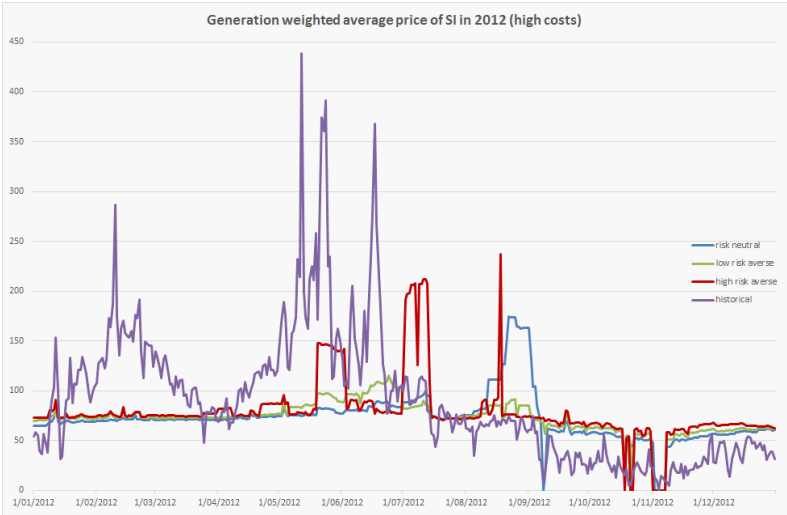

Figure 12: South Island generation weighted average prices in 2012 with high fuel costs

makes a case for strategic exercise of market power as being the primary cause, but uncertainty and risk also play a role that must not be discounted. Electricity spot markets around the world are set up as one shot games, and work well when the time horizon is short enough so that all necessary information is known at the time of dispatch. Uncertainty (such as wind intermittency) causes problems for this model. Valuing the opportunity cost of water is also difficult in this setting. Our counterfactual model is an attempt to remove some of the bias associated with hindsight benchmarks. We have not eliminated this entirely as HydrovSPD admits full clairvoyance of intra-day inflows. In addition, the benchmark model will still have some residual bias from relaxing other information constraints that will be present in a real setting.

As observed above, it is possible (at least in theory) for a Walrasian equilibrium to give a stochastic process of prices with respect to which every agent optimizes its own expected benefit with the outcome of maximizing total expected welfare. Such an equilibrium might give a sample path of prices as observed in Figure 8. As shown by [3], the stochastic process of prices that yields an equilibrium might be very complicated with none of the stagewise independence properties that make computing optimal policies easy for generators. We would contend that optimal electricity market design for systems with large amounts of hydro generation is not as well understood as the theory for purely thermal markets, and more remains to be done to improve the operation of these markets to get closer to welfare-optimizing outcomes. The different solutions adopted in various jurisdictions (e.g. New Zealand, Brazil, Colombia, France) are evidence of the

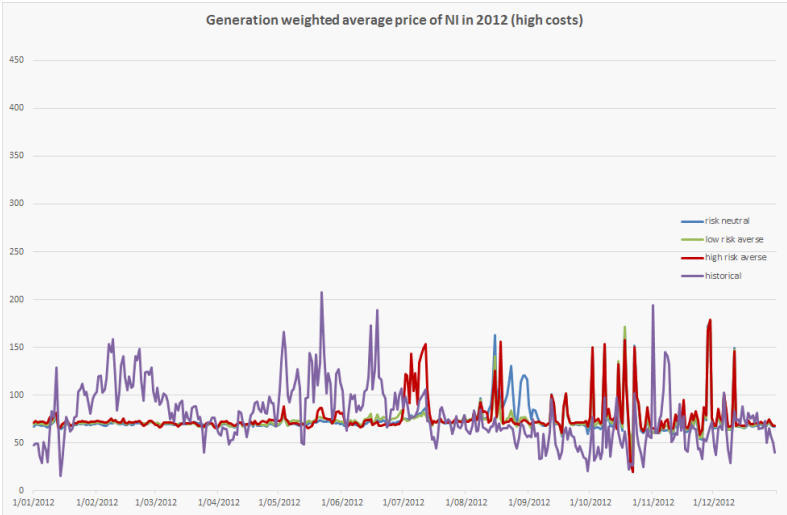

Figure 13: North Island generation weighted average prices in 2012 with high fuel costs

different approaches to tackling this issue.

Our models do not encompass forward contracts for electricity. It can be argued that most electricity produced is sold under such a contract and so it is contract prices that should be used in counterfactual comparisons. Unfortunately it is difficult to do this directly as contract quantities and prices are not in the public domain. There are two important questions to be addressed here. The first of these arises from the fact that our counterfactual models give results that are independent of any contract positions. A contract quantity of Q at contract price f produces a payoff to the seller of $Q(f-p)$ where p is the spot price. A generator who has sold such a contract, and acts as a price-taker would treat the contract payoff as a fixed constant. Even if they were risk averse with a coherent risk measure, the translation invariance of the risk measure means that the generator should act independently of this contract position to maximize (possibly risk-adjusted) profit in the spot market. It is possible therefore to test a hypothesis that agents are behaving as price takers in the wholesale market by focusing only on this market and ignoring contract positions. We have already seen that historical market prices in some periods of the first half of 2013 are lower than marginal costs of dispatched generation, which we explained as a response to overcontracting in gas. Another possible explanation is that offering generators are behaving strategically to keep prices low in case they are dispatched below their electricity contract positions.

The second question relates to generator rents in the spot market. In 2013 we saw that the counterfactual model earned more rent than the historical market.

It would be wrong to deduce that generators made less profit in 2013, since their contract prices f would be likely to be above the average value of p in 2013. In most electricity markets (including New Zealand) contracts are traded at a premium to expected spot prices (see [4]). Thus rents estimated from spot prices would be expected to be underestimates of the contract and spot revenue minus short-run costs of electricity generation.

The inflow processes used in DOASA are assumed to be stagewise independent. This means that a sequence of dry weeks will occur in the model with lower probability than it would in reality. When reservoir levels are low, and low inflows persist, this assumption will tend to produce optimistic estimates of future costs. The marginal water values at low reservoir levels are therefore likely to be lower in our model than in a model with serial dependence. We have attempted to account for this dependence using an inflow adjustment (DIA) and varying levels of risk aversion using a dynamic coherent risk measure. Even with these adjustments, the results of our experiments seem to indicate that risk-aversion to low inflows is insufficient to explain deviations of wholesale market prices from perfectly competitive estimates.

So how might one explain the differences? From the storage trajectories shown in Figures 6 and Figure 10, and matching price trajectories shown in Figures 8 and Figure 12 one can see that there are many different price sequences that will support prudent hydro reservoir releases. Consumers of electricity value it highly, and price has traditionally been a poor instrument to control short-term demand (although more price-responsive demand is emerging through retailers like Flick). Inelastic demand means offer prices early in the year in response to a dry-winter forecast may not lead to much change in consumption or even any change in dispatch. Observed price increases in these circumstances align with broad economic incentives, but this would be true for any price increase. We hope that our models are a first step towards understanding and moderating the forces that lead to prices that might be higher than they need to be to ensure efficient market outcomes.

References

[1] T. Alvey, D. Goodwin, M. Xingwang, D. Streiffert, and D. Sun. A security-constrained bid-clearing system for the New Zealand wholesale electricity market. *IEEE Transactions on Power Systems*, 13(2):340 – 346, 1998.

[2] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

[3] K. Barty, P. Carpentier, and P. Girardeau. Decomposition of large-scale stochastic optimal control problems. *RAIRO-Operations Research*, 44(03):167–183, 2010.

[4] F. Bevin-McCrimmon, I. Diaz-Rainey, M. McCarten, and G. Sise. Liquidity and risk premia in electricity futures. *Energy Economics*, 75:503–517, 2018.

[5] T. Denne. Environmental costs of electricity generation. Technical report, COVEC, 2015.

[6] L. Evans and G. Guthrie. An examination of Frank Wolak's model of market power and its application to the New Zealand electricity market. *New Zealand Economic Papers*, 46(1):25–34, 2012.

[7] M.C. Ferris and A.B. Philpott. Dynamic risked equilibrium. Technical report, Electric Power Optimization Centre, 2018.

[8] Z. Guan. EMBER online companion documents. `http://www.epoc.org.nz/ember2.html`.

[9] C. Hansen. Progress with improving electricity industry performance. `https://treasury.govt.nz/sites/default/files/2014-04/tgls-carlhansen-slides.pdf`, 2014.

[10] D. Heath and H. Ku. Pareto equilibria with coherent measures of risk. *Mathematical Finance*, 14(2):163–172, 2004.

[11] P. Lino, L.A.N. Barroso, M.V.F. Pereira, R. Kelman, and M.H.C. Fampa. Bid-based dispatch of hydrothermal systems in competitive markets. *Annals of Operations Research*, 120(1):81–97, 2003.

[12] EMI-Electricity Market Information System. New Zealand Electricity Authority. `http://reports.ea.govt.nz/emi.htm`.

[13] Ministry of Business Innovation and Employment. Energy statistics: Prices. `https://www.mbie.govt.nz/info-services/sectors-industries/energy/energy-data-modelling/statistics/prices`.

[14] New Zealand Electricity Authority. Vectorized Schedule Price and Dispatch. `https://www.emi.ea.govt.nz/Wholesale/Tools/vSPD`.

[15] Parsons Brinckerhoff. 2011 NZ Generation Data Update. `https://www.mbie.govt.nz/info-services/sectors-industries/energy/energy-data-modelling/technical-papers/pdf-library/2011%20NZ%20Generation%20Data%20Update%20v006a.pdf`, 2012.

[16] M. V. F. Pereira and L. M. V. G. Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming*, 52:359–375, 1991.

[17] A.B. Philpott and V.L. De Matos. Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion. *European Journal of Operational Research*, 218(2):470–483, 2012.

[18] A.B. Philpott, V.L. de Matos, and E. Finardi. On solving multistage stochastic programs with coherent risk measures. *Operations Research*, 61(4):957–970, 2013.

[19] A.B. Philpott, M.C. Ferris, and R.J-B. Wets. Equilibrium, uncertainty and risk in hydro-thermal electricity systems. *Mathematical Programming*, 157(2):483–513, 2016.

[20] A.B. Philpott and Z. Guan. On the convergence of stochastic dual dynamic programming and other methods. *Operations Research Letters*, 36:450–455, 2008.

[21] A.B. Philpott and Z. Guan. Models for estimating the performance of electricity markets with hydro-electric reservoir storage. Technical report, Electric Power Optimization Centre, 2013.

[22] A.B. Philpott and Z. Guan. On the performance of the New Zealand wholesale electricity market 2005-2016. Technical report, Electric Power Optimization Centre, 2018.

[23] A.B. Philpott, Z. Guan, J. Khazaei, and G. Zakeri. Production inefficiency of electricity markets with hydro generation. *Utilities Policy*, 18(4):174 – 185, 2010.

[24] POCP. Database of planned outages. `http://pocp.redspider.co.nz`.

[25] Wikipedia. `https://en.wikipedia.org/wiki/National_Electricity_Market`.

[26] F.A. Wolak. An assessment of the performance of the New Zealand wholesale electricity market. Technical report, New Zealand Commerce Commission, 2009.