Can Agent-Based Models Forecast Spot Prices in Electricity Markets? Evidence from the New Zealand Electricity Market

David Young[1], Stephen Poletti[2], Oliver Browne[2]

24th January 2012

Abstract

Modelling price formation in electricity markets is a notoriously difficult process, due to physical constraints on electricity generation and flow. This difficulty has inspired the recent development of bottom-up agent-based models of electricity markets. While these have proven quite successful in small models, few authors have attempted any validation of their model against real-world data in a more realistic model. In this paper, we take one of the most promising algorithms, the modified Roth and Erev algorithm, and apply it to a 19-node simplification of the New Zealand electricity market. Once key variables such as water storage are accounted for, we show that our model can mimic short-run (weekly) electricity prices at these 19 key nodes quite closely.[3]

Keywords: Agent-based modelling, electricity markets

---

[1] Electric Power Research Institute, Palo Alto, California, corresponding author. Research for this paper was largely conducted while this author was at the University of Auckland Business School Energy Centre.

[2] University of Auckland Business School Energy Centre.

[3] This research was partly funded by a University of Auckland FDRF Grant #9554 / 3627082. The authors would like thank Andy Philpott, Golbon Zakeri, Anthony Downward, and participants at the EPOC Winter Workshop 2010 for their helpful comments.

1. Introduction

Modelling the strategic behaviour of firms in electricity markets is surprisingly difficult. Of the two standard approaches in the literature, analytical game-theoretic models and computational competitive models, neither is fully satisfactory. Game-theoretic models allow for full strategic behaviour by firms. However, keeping the models tractable requires considerable simplifications and omissions of many of the key features of electricity network architecture and markets, such as locational pricing, line losses, and reserves, all of which impact on final prices and dispatch. As a result it is not clear how robust the intuitions derived from stylised game theoretic models are. Competitive models are solved numerically, which has the advantage of allowing for realistic networks and detailed representations of generation technologies. On the other hand, the assumption of perfect competition makes these models unsuitable to investigate policy questions where market power is an issue.[4]

Given the difficulty of modelling the strategic behaviour of firms with realistic electricity networks it is not surprising that researchers are exploring different approaches. In the last ten years, an alternative has appeared in the academic literature. This is computer simulation modelling using algorithmic 'agents' and a market dispatch solver. Instead of bidding at cost, an agent places bids for each plant, using an algorithm that aims to choose bids to increase profit. Weidlich and Veit (2008), and Guerci et al. (2010) are two good surveys of this agent-based modelling literature.[5] The potential advantage of this approach over cost-based modelling is readily apparent since it allows firms to act strategically over realistic electricity networks. The key disadvantage is equally apparent. Since agent-based models typically predict neither competitive prices, nor Nash equilibrium prices, there is no theoretical basis against which to validate the results.

In the absence of theory, the standard way to validate such a model is to compare the simulated data against real-world data, in this case, wholesale electricity prices of an actual electricity market. If an agent-based model were to consistently predicts prices across a range of market conditions, that would be a powerful validation of the underlying behaviour of the agents in the model. However, Guerci et al. (2010) observe

> "...the few researchers who have performed an empirical validation at a macro level, that is compared simulated prices with market prices, have often limited their comparisons to verbal or graphical considerations. No paper has tackled a statistical analysis at an aggregate level to prove the statistical significance of the computational results, to the best of the authors' knowledge." [p280]

Modelling nodal prices is challenging as it can only be attempted within a framework that includes realistic loop flows, line constraints, and line losses. While some papers have considered zonal pricing, see for example Rastegar et al. (2009), Sun, J. and Tesfatsion, L. (2007), Veit *et al.* (2006), and Veit et al. (2009), it is the case, as Weidlich and Veit (2008) observe , that "The large majority of models neglect transmission grid constraints." [p1753] To our knowledge there have been no attempts to model nodal prices against a real network using agent-based models, and thus no empirical validation of the behaviour of agent-based algorithms.

[4] In principle firms could be modelled as strategic players submitting bids based on their conjectures on the other players' strategies and the market equilibrium searched for numerically. However in practice problems emerge even for relatively simple networks with either multiple Nash equilibria or else no Nash equilibria at all.

[5] Weidlich (2008) also contains much useful material for readers interested in an overview of the subject.

In this paper we aim to fill this gap in the literature, by creating a detailed agent-based model that lets us compare simulated prices with prices from an actual market to validate the performance of the algorithm. We create an agent-based model based on the well-known modified Roth and Erev algorithm. We separately create a highly detailed 19-node simplification of the 244 node New Zealand electricity market. We show empirically that our agent-based model is capable of consistently predicting short-run prices in this market across a range of market conditions. This we argue, validates the behavioural assumptions underlying this particular algorithm. It also shows that these models are capable – given sufficient market data – of predicting short-run prices, something we believe has not been demonstrated previously.

We have chosen the New Zealand market for this comparison for several reasons. First, the market is small enough that we can model every significant generator in the country. Second, the New Zealand Electricity Authority maintains a publicly available dataset containing key variables such as price, dispatch, and actual bids. Third, the New Zealand market is one of the least regulated electricity markets in the world. It is one of the purest examples of an energy-only market, with no price cap, and no capacity market. This is important, as it means that generators' bids in the wholesale market are driven by profit from selling electricity, not by profit from making capacity available. Against these advantages, the New Zealand market has the additional complexity of being dominated by hydro generation. This is a significant challenge to energy modellers, as the limited hydro storage capacity means there may be periods of months up to a year with prices significantly higher than usual. In 2001, 2003, and 2008 for instance there were fears that the hydro storage lakes would run dry, which would result in forced outages. In these years, much less hydro was dispatched than normal and prices were extremely high at times, even higher than the marginal cost of thermal generation (Wolak, 2009). We address this issue in some detail, and try and distinguish between the results that are driven by water value, and those that would occur in a system with more thermal generation.

The aim of the present work is thus three-fold. First, we aim to develop a realistic agent-based model of the New Zealand Electricity Market. This involves collecting and aggregating network and plant data to create a database upon which the agent algorithms can draw. Particularly important are the transmission constraints. Then we aim to select and calibrate an algorithm in conjunction with the network data with the aim of having the model output realistic prices. Finally, we aim to verify that the model can actually predict prices in the New Zealand market across a wide range of market conditions. Aside from demonstrating the underlying algorithms have reasonable behavioural assumptions, such a model could also prove very useful for policy applications.

In Section 2 we critically review recent papers that simulate existing electricity markets and pick up a few key points from recent survey articles. Section 3 introduces and describes the model. In Section 4 we carefully and systematically calibrate the model, followed by an extensive validation procedure in Section 5, where we simulate prices for the NZEM over a complete year. Finally in Section 6 we summarise our conclusions.

2. Literature Review

Two recent surveys of the agent-based wholesale electricity market model literature were conducted by Weidlich and Veit (2008), and by Guerci et al. (2010).[6] Here we wish to focus on some important issues they identify, as well as briefly review key recent papers that attempt to model existing electricity markets.

[6] Weidlich (2008) also contains much useful material for readers interested in an overview of the subject.

There have been a number of recent agent-based studies of the German electricity market by researchers at the University of Karlhusse using the PowerACE model. (Möst and Genoese (2009), Genoese et al. (2007), Sensfuß et al. (2008)). Their approach has generator agents offering bids into the spot market based on their costs plus a mark-up that varies according to the fixed costs of the plant and a capacity scarcity factor.[7] In our view, the mechanical nature of this mark-up rule limits the usefulness of the model, as it cannot explain how changes in market design or conditions will alter the market power of participants. The strength of their approach is that they have tested their model using empirical data. Despite having no realistic networks in the model they find that the model gives a good description of prices in the German market. Presumably the mark-up rule they use includes any price impacts caused by congestion or line losses.

Rastegar et. al. (2009) model the Italian electricity market with a network of 11 zones and 10 transmission lines with realistic line constraints; probably the most complex network modelled to date. Generators are paid at the zonal price, but buyers pay the unique national price which is the demand weighted average of the zonal prices. In their model firms bid a constant mark-up factor over cost[8] for all their generation, which is determined through a modified Roth and Erev algorithm (Nicolaisen, 2001). The fact that the mark-up is determined through a learning rule is, in our view, a major advantage of their methodology compared to the PowerACE model discussed above. A possible weakness of their approach is that the linear mark-up rule means that the supply functions offered in by firms may not have enough curvature[9] to generate high enough prices when there is scarcity. The model is empirically validated at the aggregate level for prices over a single day. Using a graphical comparison of predicted versus actual national prices they conclude that the model "is in good correspondence with historical values except for some peak hours when simulated prices are significantly lower than historic values" [p6] .

Sueyoshi and Tardiparthi (2007, 2008) take quite a different approach to modelling electricity markets. Generators update their bids each turn by submitting a slightly higher capacity offer at a slightly higher price if successfully dispatched in the previous period. If unsuccessful they bid less capacity at a lower price. The authors introduce an adaptive sigmoid decision rule that updates each time period and determines how much higher or lower the new bids are. They have modeled both the PJM market and the California market with considerable success. The one drawback of their model is that it is complicated and hard to interpret intuitively. In common with other models it does less well when market power issues are at the fore, such as in California during 2001.

Our approach here has been greatly influenced by the work of Weidlich and her colleagues based at the University of Mannheim (Weidlich (2008), Veit and Wedlich (2008a, 2008b), Veit *et al. (2006), and* Veit *et al.* (2009)). Weidlich (2008) uses a learning algorithm to model firm behaviour in the German wholesale market. Firms offer in each generator at a price discovered through trial and error and learning. Although there is no realistic attempt to model the network explicitly, her work is, in our view, one of the best examples of careful modelling, calibration and validation. She examines a

[7] The markup during the peak period is a factor of ten higher than that for the off-peak period. It is not clear how they arrive at the mark-up rule. For example Möst and Genoese (2009) state "The mark up function is based on various publications in the field of electricity prices" [p59].

[8] To be precise if the marginal cost for firm *"i"* of generator *'j"* is $CM_{i,j}$ and m_j is the mark-up factor then the firms offer in the full capacity of each generation plant at an offer price $= CM_{i,j} + m_j * CM_{i,j}$

[9] That is typically observed offers by firms will have the last few tranches offered in at very high prices to take advantage of supply shortfalls or high demand. Mostly theses tranche won't be dispatched but when they are the market price will be extremely high.

number of different algorithms and finds that both the modified Roth and Erev algorithm (used here) and the Q-learning algorithm work well.[10] Veit et al. (2009) modify the agent-based model of electricity systems (AMES) developed by Sun and Tesfatsion (2007) where firms use a modified Roth and Erev algorithm to choose one of 16 linear offer curves to maximise profit. It is one of the few papers which include line-losses and network constraints explicitly with locational marginal prices, although it is not validated empirically.

One of the key themes we identify from our reading of the recent literature is that the jury is still out as to how well computer agents can realistically model actual firm behaviour in electricity markets. A common feature of almost all the agent-based models of wholesale markets that we have seen is that they struggle to reproduce the variation in prices found in actual markets. Typically off-peak prices are higher than those observed whilst peak prices are lower. Thus agent-based models tend to underestimate market power when the potential for firms to exercise market power is high. For example Guerci et al. simulate peak prices of about €100/MWh compared to the €200/MWh observed, and a similar pattern is seen in the Weidlich (2008) simulation of February prices in the German market. Another example can be found in Bunn and Day (2009) who develop an agent-based model of the UK market and find that for high demand levels the simulated market supply curve lies well below the observed supply curve. One possible interpretation they put forward is that this may be evidence of "tacit collusion" on the part of the firms.

Another point is that there has been no work which simulates nodal prices on realistic networks which may have hundreds of different locational prices. However there has been some progress simulating zonal prices with a limited number of zones. (For example, Rastegar et al. (2009), Sun, J. and Tesfatsion, L. (2007), Sueyoshi and Tadiparthi (2007), Veit et al. (2006), and Veit et al. (2009)).

Turning now to the survey articles by Weidich and Veit (2008) and Guerci et al. (2010), we wish to highlight a few key points they make. Weidlich and Veit found that the large majority of models ignored transmission grid constraints and further noted that in some papers model description was less than adequate, a point that our analysis of the more recent literature bears out. They also observe that much of the literature focus is on generic market design questions using model networks with relatively little work done modelling empirical prices. This point is picked up by Gurerci et al. (2010) in their comprehensive review of computer agent-based modelling of electricity wholesale markets. They state

> "As a final remark, it is important to note that the majority of these papers are purely computational studies, that is, empirical validation is seldom addressed. This is a critical aspect that needs to be addressed by researchers to assess the effectiveness of their modeling assumptions." [p246]

Our model picks up on some of the key themes emerging from our reading of the literature. We attempt here to carefully explain, calibrate and validate our model of the NZEM with the aim of seeing how a realistic agent based model behaves under quite different market conditions. We validate our model systematically using statistical comparisons to actual market prices. It is the first agent-based model to explicitly analyse a hydro-dominated system with shadow prices for water. It is one of the most complex and realistic networks examined to date. The combination of realism combined with large differences in the ability of firms to exercise market power as hydro lake storage vary enables us to consider the question carefully of whether agent-based models do indeed

[10] Weidlich finds that the Q-learning algorithm performs slightly better.

systematically underestimate market power. We aim to convince the reader that the agent-based approach has an important role to play which complements analytic models.

3. The Model

Our aim here is to model spot prices in the New Zealand wholesale electricity market (also known as the spot market). Although there is a high degree of vertical integration in the NZEM we focus on the spot market rather than the retail market. Wholesale prices are publicly available and are one of the key indications of the state of the market. Firms adjust many of their retail price offerings infrequently with changes reflecting trends in spot prices over time. Furthermore, in our view market power issues are likely to be reflected by price movements in the spot market. All the four major firms almost always have a net position on the spot markets with retail obligations less and sometime considerably less than their generation capacity offered onto the spot market (Wolak, 2009).

Our model employs computer agents using the Modified Roth and Erev algorithm, with further modifications following Weidlich (2008). Each agent places bids for a firm owning multiple generators. The market takes place over a 19-node simplified version of New Zealand's 244 node network. Electricity flows are modelled using a DC flow model with line losses. Firms bid into a wholesale market in which demand is assumed inelastic. The solver is modelled on New Zealand's market solver, and for given bids, demand, and network parameters, will output dispatch for each generator, prices at each node, and line flows. We now describe in detail how each element of the model was constructed.

3.1 Modelling Philosophy

In choosing the Modified Roth and Erev algorithm to reproduce price formation in the New Zealand market, we make two crucial behavioural conjectures. The first is that a single implementation of the algorithm (with all parameters fixed) should be equally able to predict prices accurately for any configuration of the network, and any level of demand. While we could no doubt get more accurate price predictions by choosing separate parameters for each half hour trading period, the resulting model would be less useful for policy analysis, as for policies that changed the network configuration; the results could not be compared.

The second conjecture is that when network configuration and demand are the same between two periods, the remaining variation in spot prices is caused by trader noise. The market price is essentially set by the decisions of four or five key traders. Small variations in their strategy combined with changes every half hour of demand or supply conditions lead to price variations around the average. Even if the traders face exactly the same market supply and demand conditions the small random variations in their offer strategies is likely to lead to a different market price. Figure 1 below illustrates the spot market half-hourly price over three typical consecutive mid-week days for the Otahuhu node in May 2006 in the New Zealand Market. It is clear from the figure that spot market prices have a clear trend throughout the day and are higher when demand is high. It is equally clear there is quite a bit of variation about the overall trend. Demand profiles for each of the three days are very similar.

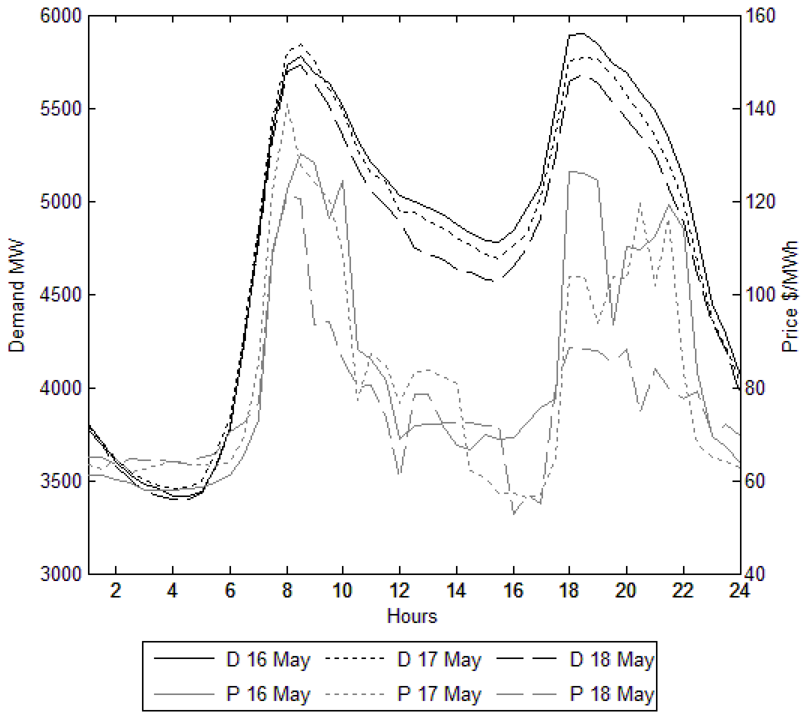


Figure 1: Sample Prices and Demand for three days in the NZEM

Our choice of the agent-based algorithm as discussed above is partly designed to reproduce this volatility. Different random seeds for the computer agents for any particular half hour trading period generate different prices.[11] For this reason we take the average of ten different simulations for each half hour price. The result is a much smoother price path over the course of the day than that seen in the data. If our conjecture is correct then there is a certain random 'trader noise' around the trend which is unpredictable.

Ideally an agent-based model of an electricity market should be robust in the sense that once parameters are chosen then prices can be predicted accurately over a range of market conditions. The more accurately prices can be predicted over a *range* of market conditions as supply and or demand change, the more confident we can be that the model is a useful tool for policy analysis. To re-iterate, a model that needed recalibrating each time market conditions change would have limited policy value as the point of policy is to *change* market conditions.

As will be seen below in the validation section we pay particular attention to evaluating the success of the model in predicting prices over a range of different market conditions.

3.2 Network

The New Zealand wholesale electricity market operates over 244 nodes. Most of these nodes are relatively minor; however, there are a string of key nodes connected to the so-called 'backbone' of the grid consisting of 220kV AC transmission lines that carry electricity up and down the two main

[11] Different random seeds cause the algorithm to choose potentially different actions at the proportional selection stage. This has a cascading effect down the 1200 iterations, often resulting in convergence to a different set of offers in the market and hence different prices.

islands, and a 350kV DC underwater cable across the Cook Strait connecting the North and South Islands.

Our network model is based upon the PEDRO dataset povided by the Electric Power Optimization Centre (EPOC).[12] The PEDRO dataset contains full network data as of 2006 for each of New Zealand's 244 pricing nodes, and the lines connecting them, including line losses, reactance, and capacities. However, PEDRO is too big for serious use in an agent-modelling context. The reason for this is the time it takes to solve the market clearing problem over such a large network. New Zealand's actual market clearing model takes around 60 seconds to solve. This is clearly impractical for agent-modelling, where in the process of 'teaching' the agents, the market will be solved around 10,000 times just to simulate a single half-hour period. From a practical modelling point of view, for a given set of offers the market should take less than ten milliseconds to solve.

In our model, we therefore use a 19-node simplification of the 244 node New Zealand grid. The choice of 19 nodes was based upon two previous modelling attempts. The first was an 18 node simplification used by Simon Young (1997) in his dissertation, with the data originally provided by Transpower, the New Zealand transmission grid operator. As Young (1997) explains in his thesis, the lines between these 18 nodes 'carry the bulk of power in the power system' [p98]. A more recent simplification was a 17 node simplification created by de Silva et al. from New Zealand's Electricity Commission (now the Electricity Authority). This is the regulatory body overseeing operation of the New Zealand electricity market. They wrote an influential report in 2006 on inter-regional power flows to guide new transmission construction. As the authors state in the paper, "This report is intended to identify the major constraints on inter-area transmission capacity in the committed system." For this reason their choice of network was intended to capture major constraints in the network. Obviously such constraints will be highly influential on prices in a nodal pricing market.

These two papers choose almost identical nodes. The only differences are that de Silva et al. drop a couple of nodes to avoid modelling loop flows, but gain one back to model a change in the network that did not exist when Young created his model. The 19 node model we use thus takes de Silva et al.'s 2006 model (which corresponds to our 2006 dataset) and adding back the loop flows used by Young.

The next challenge was to compute resistance, reactance, and capacity for each line in the simplified model. For each line in the simplified model, we went back and drew up a diagram containing all the 220kV lines between the two nodes. We then calculated capacity and admittance by aggregating using Kirchoff's laws, assuming that all demand was concentrated at the two nodes (and not at any of the intervening substations). This approach follows Transpower's computations for Young's 18 node model. Since the PEDRO dataset gives resistance as three piecewise linear components, we were not able to compute aggregated resistance in the same way. However, de Silva et al. have computed resistances in their 17 node model. We used these values, and for the three missing lines, used the values from Young's 18 node model. Since there has been relatively little investment in the grid from 1993 to 2006, these values are likely to be still accurate.

A diagram of the resulting network, along with tables of data, are given in Appendix A.

[12] The Electric Power Optimization Centre (www.epoc.org.nz) is a subgroup of the Department of Engineering Science at the University of Auckland. Their PEDRO dataset was obtained by combining data from Transpower (the New Zealand grid operator) and Contact Energy (one of New Zealand's largest generation companies).

3.3 Demand

In the New Zealand electricity market, retailers and large industrial users are able to bid to buy electricity in the market in much the same way as generators make offers. However, much of New Zealand's demand is retail consumers, who are on fixed price contracts. Wholesale prices do not influence their demand. Modelling demand-side bidders with agents is possible, but more difficult than generators as we do not know the value of electricity to a large industrial user, say, whereas we do know the costs of generation to a reasonable approximation. For these reasons, we model demand as inelastic, and take actual demand data from the Centralized Dataset (CDS).[13]

Castalia[14] (a New Zealand consultancy) has investigated price elasticity of electricity in New Zealand, and found there was relatively little elasticity. They concluded that when the price of electricity reached $300, a demand response of about 2% could be observed. If the price went over $500, emergency conservation measures could result in another 2% reduction. This is a very small reduction, occurring only when prices are abnormally high. We can thus feel confident that assuming away demand elasticity is unlikely to greatly affect the model. It is also possible that the calibration process will partly compensate for the lack of elasticity.[15]

3.4 Firms

There are five large generation firms in New Zealand, each with 5% or more of total generation capacity. These are Meridian, Contact, Mighty River Power, Genesis, and Trustpower. Contact and Trustpower are privately owned; the other three are state-owned enterprises (SOEs) owned by the Government, but required to make a profit. There are also a number of smaller firms who own less than 100MW of generation capacity.

In our model, we distinguish between the big five firms and the smaller firms by requiring the latter to bid at marginal cost. These firms are so small they are unlikely to exercise market power. The big five are all assumed to be profit maximizing, so each of these firms is played by an 'agent' (see below) in our model.

3.5 Must Runs, Reserves, Outages, and the Tiwai Aluminium Smelter Contract

Making the model realistic means that we need accurate information on available generation capacity. In general we abstract away from the long-term contracts, often held by large industrial firms with specific generators. Since this data is not publically available, we cannot directly incorporate it. We make one exception to this: the contract between Meridian and the Tiwai Point Aluminium Smelter. Meridian's Manapouri hydro plant was built specifically to supply the smelter, and there is an existing contract amounting to a continuous supply of 572MW between the two companies. The price is unknown, but likely to be in the $40-$50 range.

If we do not account for this contract in the model, the agent bidding for Meridian will raise prices close to the maximum allowed ($1000). Our assumption that the smelter's demand is inelastic, combined with nearby transmission constraints, imply Meridian is effectively a monopolist over the

[13] The dataset is available on request from the NZ Electricity Authority http://www.ea.govt.nz/industry/modelling/cds/

[14] www.castalia-advisors.com/files/22631.pdf

[15] Furthermore the computer agents who representing the major firms usually set different offer prices for each of their generation plants. Thus they offer an upward sloping supply curve to the wholesale market. Each agent will then "see" a residual demand curve which is already elastic. Including some demand elasticity would likely mean a small change in the residual elasticity and hence a small change in the structural parameters when we calibrate the model.

residual demand of about 100MW in the south of the South Island. Naturally a profit-maximizing agent will raise the price. In reality, the existence of the contract precludes such behaviour, and so we adjust for the contract by subtracting 572MW from Manapouri's capacity, and 572MW from demand at the TIW node. This avoids us having to guess the price for the contract, while accounting for the fact that Manapouri was really built to supply the smelter with cheap electricity.[16]

Long term contracts may reduce a firm's incentive to exercise market power on the spot market but do not eliminate it entirely. Given this, the calibration process should account indirectly for the existence of long term contracts. However it must be noted that the model will only be valid so long as the contract levels are stable. A large change in the contract positions held by the market would alter incentives for firms on the spot market resulting in different prices. However, this change would not be visible in data available to the algorithm, implying that the model would have to be recalibrated.

There are also a number of plants which must run either continuously or as conditions dictate. These include geothermal plants, some hydro[17], and wind. We have accounted for the must-runs by directly searching the CDS for capacity that consistently offers in at a price of approximately zero. Some must-runs were determined by correspondence with electricity traders. A detailed list of must-run generators can be found in Appendix B. The computer agents in the model have no control over the must runs which are always offered in at a price of zero. Must-runs amount to approximately 17% of total capacity.

A couple of new plants were commissioned during our 2006-8 study period, including the Huntly CCGT plant (E3P) and the Kawarau Geothermal plant (KAG). These were accounted for in our dataset. One plant was decommissioned during the period, the New Plymouth OCGT plant (NPL). Since not all plants run at all times, we also drew upon the Maintenance schedule maintained by Transpower to see when each plant had a scheduled outage, and searched the data from the CDS to see if a plant had a unplanned outage.

In the NZEM firms offer into both the generation and the reserve market simultaneously. Optimal dispatch and cleared reserves are solved for simultaneously. Although in principle we could follow a similar procedure, doing so would add a second layer of complexity to the agent's decisions. Since the point of this paper is to demonstrate that an agent-based model can perform well at predicting prices in a wholesale market, we did not want to add complexity that detracted from this goal. We found it simpler to solve only for the spot market and subsequently account for the fact that capacity set aside as spinning reserves cannot be dispatched by reducing the capacity of each plant by the average fraction of cleared reserves. We estimate this from the CDS at 12% of total generation capacity. In this approach we follow Weidlich (2008).

3.6 Water Values

As discussed in the introduction, New Zealand's electricity market is dominated by hydro generation. Since river flows are variable, reliable hydro generation requires water storage – something that New Zealand lacks. The total capacity of hydro lakes in New Zealand is considered to be about 3600GWh, or enough to satisfy about five weeks of demand if the thermal generators run as well.[18] Every winter,

---

[16] The Manapouri hydro plant is located at the MAN node which is linked to the TIW node by three 220kV lines with a joint capacity far higher than that of the plant.

[17] Hydro may be forced to run due to minimum river flow requirements or may be run of river.

[18] This statistic is drawn from the former Electricity Commission's review into the winter of 2008. The report may be found at http://www.ea.govt.nz/our-work/consultations/security-of-supply/review-of-2008-winter/.

when moisture falls as snow rather than rain, the generators start drawing down on their hydro storage. The idea is to make the existing storage last until the end of winter – a period of about 3-5 months depending on how much rain there was during the year. To ensure that they are not dispatched at full capacity, the hydro generators will thus bid higher deliberately reflecting the scarcity value of water. This is despite the fact their nominal cost of water is zero. In fact, these generators have a positive opportunity cost of water, since water used now cannot be used later when it might be even more valuable. To accurately model water values for the national hydro lakes would need a dynamic optimisation program incorporating expectations of future prices, inflows and behaviours which is a daunting task. Instead we use a proxy for the water value suggested by Tipping et al. (2004).

The water storage problem has been well studied in New Zealand, and a serious econometric attempt to calculate the opportunity cost of water given the existing level of water storage was made by Tipping et al. (2004). They posit that the water value for the hydro lake system depends not on the absolute national lake levels but on the difference between the actual and the expected lake levels for the time of the year. They argue that this measure implicitly includes expected annual average patterns of generation and inflows. They estimate an economic model of the New Zealand spot market and find good evidence that including water values in the way they suggest leads to increased accuracy of their spot-price model.

We modify their approach slightly and compute a water value curve as a function of the difference of national storage levels compared to the expected storage level for a given day.[19] Our method differs slightly from Tipping et al., a difference we will describe in more detail in the Calibration section. We will demonstrate in the Validation section that water storage is a critical factor in predicting wholesale electricity prices in New Zealand, and thus we calibrate our model against water storage to compute the opportunity cost of water to hydro generators.

3.7 Marginal Costs

Data on New Zealand's generators was taken from the New Zealand Electricity Authority's (2008) Generation Expansion Model (GEM)[20]. All generators over 10MW in capacity are included in our simulations. Marginal cost (MC) data, including variable operating costs, for each type of generation is summarised in Table 1.[21] Fuel costs for renewable generation are zero, but like the thermal generators' marginal costs, include a small operating and maintenance cost per MWh. Must-run generation are accounted for by fixing their marginal costs at zero. Hydro generators' marginal costs include a water value as discussed above. We restrict the action space of the computer agents so they can only offer generation into the market at a price greater than or equal to marginal cost.

[19] Tipping et al.'s water value is based on the Waitaki storage level only, whereas we take the national storage level. The Waitaki storage capacity is well over half the national storage level. We found during the calibration process that using the total storage capacity of all the hydro lakes gave a better fit to the data.

[20] Information on GEM can be found at http://www.ea.govt.nz/industry/modelling/in-house-models/gem/

[21] Detailed costs on a plant by plant basis are provided in the Appendix.

Generation Type	MW capacity (after reserves)	Percent of total capacity	Comment	MC ($/MWh)
Geothermal	447	6%	Must-Run	7.30
Hydro Must-Run	811	12%	Must-Run	$10+Water Value
Hydro	4032	58%		$10+Water Value
Wind	317.5	5%	Must-Run variable output depending on wind.	2.50
Coal	828	12%		$54.50
CCGT	974	14%		$49-$65
Gas peaking	181	3%		$70 +
Other	156	2%	Smaller Plants	Varies

Table 1: Generation Capacity and MC for NZEM 2008

3.8 Market Solver

We use the Perm market solver provided by EPOC. This is a DC flow model with loop flows and line losses, developed to mimic the real New Zealand market solver. The model is coded in the AMPL Optimization Language, and solved using CPLEX. The optimization problem is a linear program with integer constraints: minimize cost of dispatch subject to satisfying all network constraints. Nodal prices are then extracted as the shadow price of additional generation at each node.

3.9 Agents

The agents used in our model use the a variant of the *Modified Roth and Erev* algorithm, modified again as by Weidlich (2008) to allow for firms owning multiple generators. The Modified Roth and Erev algorithm is well described and well established in the literature. However there is still no universally agreed "best practice" algorithm in the literature. Weidlich (2008) found that both the Q-learning algorithm and the modified Roth and Erev algorithm performed well. Despite this a significant number of researchers have settled on the modified Roth and Erev algorithm. For example Banal-Estañol and Micola (2009) state:

> "We adopt the well-known reinforcement learning method put forward by Roth and Erev (1995). This method has some advantages. Because it is widely used and more parsimonious than other algorithms, our results are more easily comparable to the preceding literature. Moreover, its principles fit some features of energy market trading well. It is based on the law of effect and the law of practice, which are robust properties observed in the literature on human learning. Other simulation algorithms are either completely naive (e.g., zero intelligence) or difficult to interpret in an energy market context (genetic algorithms, Q-learning). One of the main strengths of the [Roth and Erev] method is that one does not need to make assumptions on the information that players have about strategies, history of play, and the payoff structure of the other players. In many cases, energy market players cannot observe one another's current strategies, and only imperfectly infer them from volatile prices. Algorithms like best response, fictitious play, or experience-weighted attraction require agents to have an amount of information that we find difficult to justify. (p. 1818) "

We are sympathetic to this view. Joyce (2010) compared a number of algorithms on a simple eight node network to the behaviour of human agents (Young, 2008) on the same network. She found that the modified Roth and Erev algorithm was better able to capture the dynamics of human agents. Other algorithms converged to Nash equilibria which might be expected from a game theoretic analysis, however the human agents give a much more volatile price series which is captured well by the modified Roth and Erev algorithm. Thus there are strong arguments for using the modified Roth and Erev algorithm which is the approach taken here. We describe here the exact implementation we use, and leave the reader to follow up the references for more information.

Each plant in our model has a rated capacity, and is allowed to bid one price for that capacity. The modified Roth and Erev algorithm requires a discrete action space. We thus divide the action space between 0 and 1000 into 101 options in the set {0, 10, 20, …, 1000}. Each plant can bid anything in this set. If each plant were individually owned and operated (maximizing its own profit), then the modified Roth and Erev algorithm works as follows. The generators bid into the market 1200 times. Each agent *i* has a propensity function $q_{ij}(t)$, defined as the propensity of agent *i* to play action *j* in time *t*. Given known propensities at time *t*, the propensity at time *t + 1* is

$$q_{ij}(t+1) = \begin{cases} (1-r)q_{ij}(t) + R(x)(1-\varepsilon) & if\ \ j = k \\ (1-r)q_{ij}(t) + q_{ij}(t)\frac{\varepsilon}{M-1} & if\ \ j \neq k \end{cases}$$

where ε is the experimentation parameter, *r* is the recency parameter, *k* is the last action chosen, and *M* is the total number of possible actions. *R(x)* is the reinforcement the agent receives from *x* (here *x* is the profit). Roth and Erev define the reinforcement as

$$R(x) = x - x_{\min}$$

where $x_{\min}$ is the smallest possible payoff.

Given the propensities, the action actually chosen in the next round is probabilistic with the probability of choosing a given action *j* given by

$$p_{ij}(t) = \frac{q_{ij}(t)}{\sum_{k=1}^{M} q_{ik}(t)}$$

This is the proportional selection rule, and this choice of action selection rule explains why we cannot admit negative propensities, since this would lead to negative probabilities. Other action selection rules are possible, see Weidlich (2008) for details.

In practice, one firm owns many plants, and will construct a profit-maximizing strategy across all plants. Thus it may be optimal for some plants to make less profit in order to maximize the firm's profit. One way to model this would be to make the firm the agent, with a n-dimensional action space, choosing a bid for each plant in its portfolio at each round. However, this approach is impracticable. If each plant has 100 possible actions, then one firm with two plants has 10,000 possible actions and so forth. Computation time rapidly approaches extremes. Weidlich overcomes this problem by an alternative solution. She introduces the parameter ψ, which is a weighting on how much the plant should consider its own profits versus the firm. The plants remain the agents, but their reinforcement payoff now depends on the firm's profit as well. The new formula for the reinforcement payoff is

$$R(k) = \psi R(k) + (1-\psi)\left(\frac{\sum R(j)}{n}\right),$$

where the sum is over all the plants owned by the firm, and *n* is the number of plants owned by the firm.

Note that Weidlich also weights her profits by the total possible profit, so that all her reinforcement payoffs are between 0 and 1. We do not do this, but increase the starting propensity to compensate.

3.10 Modelling Procedure

The model takes about 600 seconds to obtain simulated prices for a single half-hour trading period in New Zealand.[22] The procedure is as follows. We set a random seed for the model, and then run the modified Roth and Erev algorithm for 1200 repetitions.[23] We call this single run a 'game', and the predicted price from a single game is taken by averaging the prices at the last 100 repetitions at each node. However, to get a predicted price for the model, we repeat the game five times using different random seeds, and take the average of the prices from each game at each node to get our predicted prices. The choice of five repetitions per predicted price was a trade-off between computational speed, our preference that the model replicate some 'trader noise', and the additional variation captured by using additional repetitions. We ran some tests before settling on five repetitions; we discuss these tests in more detail at the end of Section 4.

We initially started with 1500 repetitions per game. However, it became apparent that with our choice of parameters, convergence took place usually within 1000 repetitions.[24] Thus to shave computational time, we chose to cut games to 1200 repetitions.

Although each generator is offered in at a single price[25] the fact that a firm owns a number of generation assets means that they end up offering a supply ladder into the market, albeit one that is less finely graduated than seen in actual offer stacks in the NZEM.

4. Calibration

In this section we calibrate the model by identifying values for the six 'free' parameters such that simulated prices most closely match prices in reality. Of these six, five are parameters of the modified Roth-Erev Algorithm: ε, r, ψ, K, and $s1$. The sixth is water value; the opportunity cost to hydro generators of using water.

Of the authors that have stated their calibration procedure, most have the goal of achieving a single peaked distribution after a given number of runs, i.e. convergence. Weidlich had a looser set of 'rationality constraints' which her algorithms were required to satisfy. Consequently she ended up with a range of parameters. By contrast, our approach is deliberately to choose the parameters that

---

[22] Note that much of this time is taken up reading and writing files to the hard drive as data is passed between C and AMPL. We are in the process of rewriting the program to use a direct connection to the library files offered by CPLEX. We expect this to result in a ten-fold improvement in speed.

[23] The random seed generator is seeded by the choice of random seed, and all draws for the subsequent single run of the algorithm stem from this seed. This allows the results to be replicable.

[24] The algorithm does not usually converge to a single value, since there is always some experimentation. Convergence in this case means that the slope of the learning curve is flat and variation between bids relatively small.

[25] Some of the larger plants are divided into smaller "sub-plants". For example the largest thermal plant is treated as four separate generation plants which implies the firm can effectively offer four price tranches to the market. For details see Appendix 2.

most closely match simulated prices to reality at five key nodes. (The other 14 will have very similar prices differing only by the value of line losses.) By necessity this requires convergence. With the parameters we choose, the agents also display considerable rationality. For instance, if we set transmission parameters to create a load pocket with only one firm available to supply, the algorithm will always have that agent bidding extremely high in that pocket – and only in that pocket.

We perform the calibration exercise in two stages. Our model assumes that when water storage is the same in two different days, the opportunity cost of water in those two days is also the same. Thus we start by choosing a level of water storage that represents '0' water opportunity costs, and calibrating the five algorithm parameters assuming water value is zero. Having fixed these five values, we then run a second calibration exercise to find the functional relationship between water storage and water value that best fits real prices.

Our modelling hypothesis is that the computer agents are characterised by behavioural parameters which are independent of demand and generator marginal costs. That is, there is one set of parameters that can be applied to any network to get a realistic prediction of nodal prices. Thus we calibrate the agents in situations where lakes are close to capacity and hence zero water value. Having established the structural parameters we then estimate the water value curve as function of the relative lake storage levels.

Note the underlying assumption here is that all the network parameters are accurate, including capacities, costs, and line losses. This is why we have paid very careful attention to getting the most accurate network data in Section 3. However, we cannot get data for water opportunity costs since these are dependent upon a firm's level of risk, which is unobservable. What we can do is figure out when those opportunity costs are likely to be zero, and calibrate upon those days, then use the calibrated parameters to figure out what water opportunity costs should be for other levels of water storage.

4.1 Baseline Calibration

The choice of parameters in Roth and Erev algorithms greatly influences the performance of the algorithm. Different authors have chosen radically different parameters, depending on their exact implementation. There are also several design choices to be made within the framework of the Roth and Erev algorithm. One is how the agents choose a new action from the set of actions. Some authors use proportional selection, others use Softmax selection. We use proportional selection. Authors using the original Roth and Erev algorithm have the option of using 'spillover' – where part of the return on a given action 'spills over' to neighbouring actions. We refer the reader to Weidlich's book (Weidlich, 2008) for a good description of the algorithm variations.

We present a comparison of the basic parameters used by different authors below, cautioning that because each author's implementation of the algorithm is unique, none of them are truly comparable. The models of Veit, Weidlich & Krafft (2009), and Sun & Tesfatsion (2007) in particular have firms bidding linear supply curves rather than step supply functions. Of the sample below, only Weidlich uses the ψ parameter, and her algorithm also allows firms to choose capacity, adding a sixth parameter beyond the five we ourselves use.[26] Our model, while closest in spirit to Weidlich's, is thus not exactly identical to any previous model, and we essentially start our calibration from scratch.

[26] Weidlich had two markets, a day ahead market and a balancing market afterwards. Thus she allowed firms to choose capacity in the day ahead market, but made them submit the rest to the balancing market. Since New Zealand has a single

Paper	Algorithm	ε	r	s1	K	ψ	Runs
Veit, Weidlich, & Krafft 2009	MRE with Softmax (C = 1000)	0.97	0.07	6000	16	n/a	3000
Weidlich 2008[27]	MRE with Proportional	0.2, 0.3, 0.4	0.1, 0.2	1	21 × 5	0.5	7300
	MRE with Softmax (1/tau = 6,7,8)	0.2, 0.3, 0.4	0.1	1	21 × 5	0.5	7300
	RE with Spillover and Proportional	0.2, 0.3, 0.4	0.1, 0.2	1	21 × 5	0.5	7300
	RE with Spillover and Softmax (1/tau = 6,7,8)	0.2, 0.3, 0.4	0.1	1	21 × 5	0.5	7300
Sun & Tesfatsion 2007	MRE with Softmax (C = 1000)	0.97	0.04	6000	10 × 10 ×1 = 100	n/a	422
Nicolaisen et al. 2001	MRE with Proportional	0.97	0.04	1	30	n/a	1000
	MRE with Proportional	0.99	0.02	1	100	n/a	10000
	MRE with Proportional	0.2	0.1	9	30	n/a	1000
Roth and Erev 1998	RE Original with Proportional	0.2	0.1	9	2-5	n/a	~200

Table 2: Calibrated Parameters in the Agent-Based Modelling Literature[28]

We started by choosing four ‘representative days’ in 2007 and 2008 to calibrate the behavioural parameters in the model. We chose one trading period from each of those days (one peak period, one off-peak period, and two shoulder periods).[29] We then carefully adjusted the network for each of these four periods to reflect the real conditions at the time. Specifically, we adjusted the network to compensate for generation outages, transmission outages, and checked actual output from intermittent generation (wind) at the time. As discussed above, we choose days where total New Zealand hydro lake storage was around its 89th percentile (as measured since 1990 using data from the NZEM CDS). We performed further checks to ensure that no single lake was full or spilling on calibration days.

We then performed essentially a grid search procedure across all five parameters to identify which gave the best fit to real prices on those four periods. We paid more attention to some parameters than others. K, for instance, represents only the number of actions each generator can choose from. We set this at 101 for computational reasons, and did not touch it. Similarly, s1 – the starting propensity for each action – is relatively unimportant. The higher it is, the longer it will take the algorithm to

settlement system, it makes sense for us to force firms to bid all their available capacity in the market. However, firms can set whatever price they like on their capacity.

27 Weidlich used an action space that had 5 capacity choices and 21 price choices in the balancing market. Her firms also traded in a day-ahead market that had 6 capacity choices. Weidlich scaled all propensities to between 0 and 1, and thus chose a very low value for s1.

28 Here ‘MRE’ stands for ‘Modified Roth and Erev algorithm’ and ‘RE’ stands for the original Roth and Erev algorithm without the modification made by Nicolaisen et al. (2001).

29 The periods chosen are: period 29 on the 26th February 2007, period 16 on the 19th March 2007, period 40 on the 9th December 2008, and period 5 on the 28th November 2008.

converge. We tried only three values for this (500000, 700000, and 1000000). For ψ, we started with five values in the range [0.3, 0.7], following Weidlich (2008). We judged ε and r to be the most important variables, and did a wide grid search for these.

The performance of the algorithm for each combination of parameters was assessed by looking at the root sum squared error (the l2-norm on the vector of differences between simulated prices and real prices) over five key nodes.[30] These included OTA (Otahuhu), HAY (Haywards), and TWZ (Twizel), which are the three nodes at which hedge prices are set in the New Zealand market. The other two were BPE (Bunnythorpe) and TIW (Tiwai), chosen because they represent nodes at the far side of two major transmission bottlenecks in New Zealand. These five nodes include three in the North Island, and two in the South Island, and are roughly evenly spread over the length of the grid.

After the wide grid search, we did a more targeted grid search on the parameters that performed well in the first search. We increased the fineness of our grid search on ε and ψ as well, and we increased the number of random seeds used per price prediction from five to ten. This search revealed several combinations of parameters that performed well over all four calibration periods. From this we choose the following parameters as the best at matching simulated prices with actual prices.

Parameter	Value
$s1$	700000
ε	0.89
r	0.07
ψ	0.7
K	101

Table 3: SWEM Calibrated Parameters

4.2 Water Values

With behavioural parameters chosen from the first round of calibration, we then turned our attention to calibrating the water value curve. To accurately model water values would need a dynamic program incorporating expectations of future prices, inflows and behaviours[31]. Added to the computational requirements of the existing model, this is clearly unrealistic.

As discussed in Section 3, in our paper we take a similar approach to Tipping et al. (2004). Our aim is to construct a water value curve with water values as a function of the deviation of actual water storage level on that day from the historical average. Specifically, we calculate the historical average and variance (using data back to 1990), then calculate a benchmark by computing the mean minus 1.8 standard deviations. The difference between this and the actual water storage level (denoted by D) gives the value that we use to compare water storage on different days and years.

The approach here is slightly different to that followed by Tipping et al., who construct their water values by looking at the lake levels for the Waitaki hydro chain only. Tipping et al. focus on modelling prices at the Twizel node; the node the Waitaki hydro generators feed electricity into. Thus

[30] Recall that bids were constrained to be between 0 and 1000, whereas there is no maximum price in the NZEM. Final simulation prices were determined by averaging over the last 100 iterations, out of a total of 1200 iterations.

[31] See Philpott et al. (2010) who model optimal hydro dispatch in the NZEM using a stochastic optimisation procedure which is equivalent to setting water values for the much simpler example of a competitive market.

it is natural for them to concentrate on storage levels only for the Waitaki system.[32] There is some merit in their approach. The Waitaki storage level is more than that held by the entire rest of the country. Furthermore, industry participants have informally told us that other hydro chains tend to take their lead from the behaviour of the hydro firms on the Waitaki. However there are also strong arguments for considering the storage level for the country as a whole or even trying to construct separate water values for the three major hydro chains – especially since we are trying to model nodal prices across the entire country. We chose to take an empirical approach and tested all three approaches, before settling on using water values obtained using the combined storage level of all the major hydro chains. Water values are modelled here as a direct increase in the marginal cost of hydro generation.

Calculating the water value as a function of the relative hydro storage level is relatively straightforward, albeit time consuming. We use lake storage data from the New Zealand Electricity Authority.[33] We chose a number of half hourly periods with a range of different lake levels. For each period prices are simulated for a range of different water values between -\$50/MWh to \$300/MWh. From these values, we choose the water value that does the best job of predicting prices for that trading period. This gives one point on the water value curve. Eventually after we have enough points scattered along the water value curve, we fit an exponential function to the data points.

Specifically we choose 2 days from each month in 2007 (a relatively wet year) and 2008 (dry year) and simulate 3 periods from each day[34]. For each of the half hour periods we ran the simulations for 15 different water values. For each day, and for each water value we calculate the error, which is the difference between the predicted price and the actual price. Presuming the water value on any particular day is constant we find the water value[35] which minimises the square root of the sum of the squared error terms over the three half hour periods for that day. Following this procedure for each of the 48 days yields data points scattered around the fitted water value curve. Following Tipping et al. we fitted the data to a negative exponential with two free parameters y and c. We assumed the following functional form, where D is the deviation in water storage as described above.

$$WV = c + we^{yD}$$

Tipping et al. used a more complicated version with an extra quadratic term added within the exponential. Doubtlessly additional free parameters would give better results, however we find our chosen specification works well and has the virtue of parsimony.

One shortcoming in our approach is that it is myopic. In reality water values will depend on expected demand and inflow sequences as well as the lake levels. For example forecast heavy rain would likely reduce the water value even for relatively low lake levels. In general inflow and demand patterns exhibit a strong seasonal trend. Inflows tend to be low until winter when demand is highest. Thereafter from spring to summer inflows increase substantially. Tipping et al. (2004) argue that they expect reservoir management policies to be different for the warm season (September to February) compared to the cold season (March-August) and construct different water value curves for the two

[32] The Twizel (TWZ) node or nodes nearby are the base for most hydro generators in the area that use water from the Waitaki river to generate electricity.

[33] The dataset can be downloaded from the NZ Electricity Authority website at http://www.ea.govt.nz/industry/modelling/cds/hydro-lake-storage/.

[34] The exact days and times were the 12th and 24th of each month at 1am (off peak), 8.30 am (peak) and 3pm (shoulder).

[35] To increase the accuracy we fit a spline function to the root sum of squares error data points as a function of water value. The minimum of the fitted curve is the water value for that day.

periods. We use a similar approach here. The formulas for summer (1 August – 28 February) and winter (1 March – 31 July) are

$$WV_{Summer} = 130e^{-0.0017D} - 45,$$
$$WV_{Winter} = 185e^{-0.0018D} - 28,$$

where WV is the water value, and D is the deviation from mean as described above. The data and fitted curves for summer and winter are depicted on the next page.

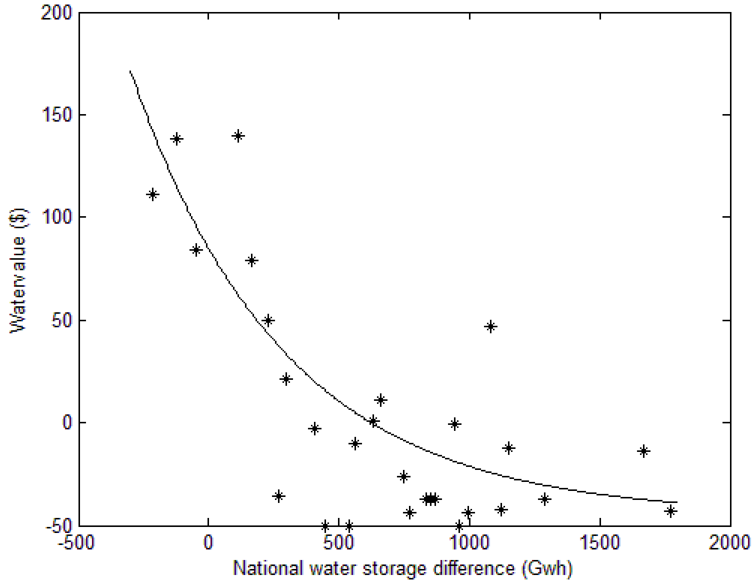


Figure 2: Fitted Water Value Curve (Summer)

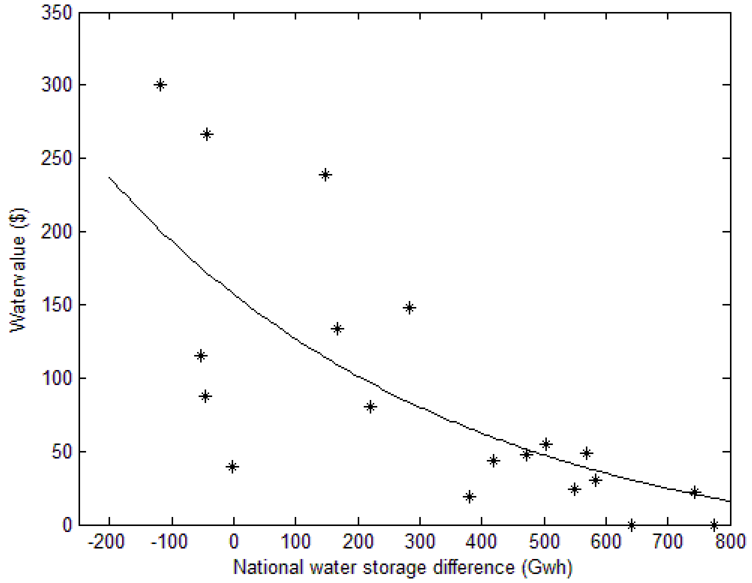


Figure 3 Fitted Water Value Curve (Winter)

The water values in winter show considerable variation for low lake levels, perhaps reflecting the increased importance of inflows during this period. Our function differs from Tipping et al., who fit a curve to the Twizel spot price using the Waitaki storage value. Their water value curves for summer and winter are reproduced below. However it must be remembered that their curves are not directly comparable to ours. They are implicitly assuming the market is competitive and they focus on the Waitaki hydro chain only[36].

$$WV_{Summer}^{Tipping} = -10 + 275e^{-0.58(5+0.01SD)^{0.50}}$$
$$WV_{Winter}^{Tipping} = 23.5 + 390e^{-2.725*(2.5+0.01SD)^{1.33}}$$

While this approach does not capture all the nuances of the opportunity cost dynamics of a hydro-dominated electricity market, it does appear to account for a good deal of the price fluctuation remaining after calibrating the behavioural parameters, and does so in a relatively simple way. We will come back to this point in Section 5 when we verify the performance of the algorithm.

One point worth noting is that the water values may be negative if the lakes are very full. In the New Zealand market, offers cannot be negative, hence even if the water values are negative the computer agents are constrained to offer in at a price of zero. Negative water values are not unexpected since if the lakes are full and there is a likelihood of further rain it becomes an imperative to release water to prevent flooding and spilling of water.

4.3 Are Five Seeds Sufficient to Predict a Single Price?

In Section 3, we outlined our modelling procedure, which states that each price is the average of the prices of five runs using different random seeds. This raises the question of whether we are losing vital information by only using five seeds. To test this, we selected four periods from 2008, and ran the model 50 times with different random seeds on each period. Recall from Section 3.1 that different random seeds often result in different prices. We then calculated the average of any five of the results across 50 different combinations, and did the same for the average of ten results.

As would be expected, the five seed results showed more variation in price predictions than did the ten seed results. The former had a standard deviation of $4, the latter, a standard deviation of $2. However, if you consider the real NZEM, and take 50 periods with identical demand and supply characteristics, you get a range of prices with a very similar variation. Against the dubious benefits of using 10 seeds per price prediction are the obvious negatives. A simulation of one year's worth of data takes about a week using five seeds. We judged that any potential gains by switching to ten seeds were outweighed by the additional computational time.

Another argument against using more seeds is that our verification procedure averages prices weekly across a year. In this case, capturing period to period fluctuations is less important than the week by week consistency of average predictions against real averaged prices. In this sense, we felt using ten seeds would add very little additional information.

We made one exception to our use of five seeds – in the final stages of calibration, having narrowed the field of parameter candidates considerably, we used ten seed runs to home in on our final choices.

[36] There is a further technical difference in that their variable 'SD' is defined using the smoothed 10th percentile lake level for that day. Here the lower bound lake level is defined as the lake level for that day which is 1.8 standard deviations lower than the mean lake level for that day. We found slightly better results using our definition.

5. Verification

The real test of any model is whether it can accurately replicate historical outcomes. As we discussed in Section 2, few authors have attempted to validate the modified Roth and Erev algorithm or indeed other agent-based algorithms against historical electricity market data. One of the strengths of our model is its relative realism against the New Zealand electricity market. This allows us to use actual historical data from the market as the basis for validating our model's performance.

To validate our model we now use it to predict historical wholesale electricity prices during the year 2006.[37] We run a simulation that replicates four periods every day over the course of the entire year of 2006, specifically periods 8, 20, 32, and 44 (corresponding to 4am, 10am, 4pm, and 10pm respectively). These four periods include one off peak period, two shoulder periods, and one peak period. Simulating over the course of a year aims to show that the model can accurately reproduce the changes in price driven by changes in demand, water value and random outages.

To accurately run this validation, every major line and generation outage over the course of the year needs to be checked and incorporated into the model. This was done systematically using information from the NZEM CDS. A separate network configuration was created for every period depending on the state of the network.

The year 2006 was chosen for validation for several reasons. It is the year that the network data we obtained was taken from, so this choice minimizes inaccuracies caused by network changes. Second, it has a less extreme range of lake levels than the wet 2007 and extremely dry 2008, so there are no inputs for which the model has not previously been calibrated for.

5.1 Water Values

The water values for 2006 are inferred from the lake's storage difference using the relationship calibrated in the previous section. From March-May in 2006, water values are extremely high, due to very low lake levels. After May, lake levels remain below average, but not low enough to generate a large opportunity cost to water.[38] The water values for 2006 and associated storage level compared to the reference low water level bench mark are depicted in Figure 4.

[37] Recall the model was calibrated using 2007 and 2008 data.

[38] Water levels from May onwards were above our reference low storage level benchmark.

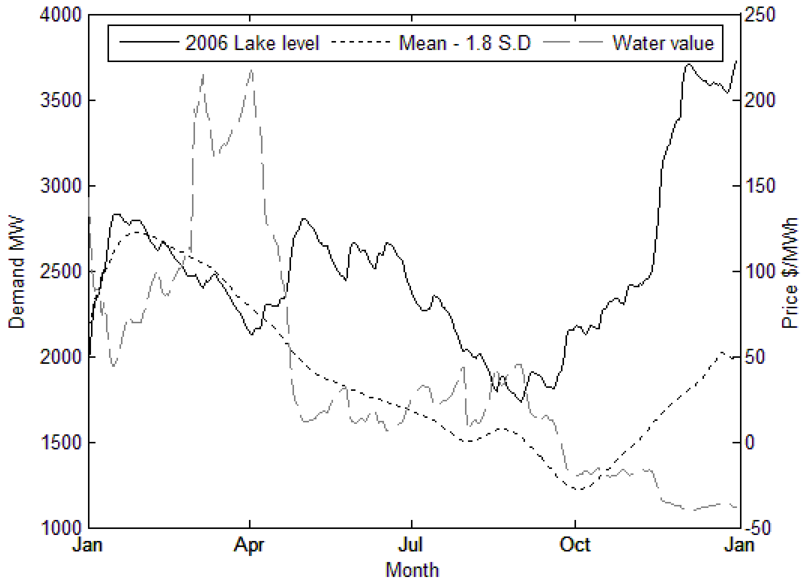


Figure 4: Water Storage and Values in 2006

5.2 Predicted Prices in 2006 – Graphs and Statistics

Using SWEM, we generated price predictions for all 19 nodes for four periods in each day of 2006. Figure 5 and Figure 6 show average weekly prices comparing our simulated prices to the actual prices for two important nodes: Otahuhu (OTA) in the North Island and Twizel (TWZ) in the South Island.[39] Figure 7 displays the average simulated price for the five key nodes we used to calibrate the model.

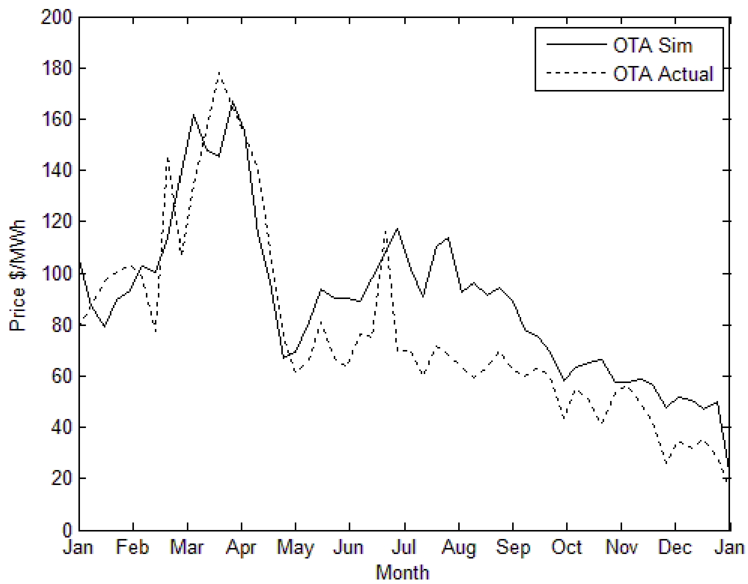


Figure 5: Average Weekly Price Comparisons for OTA

[39] In general the prices in the other North Island nodes are similar to Otahuhu whilst Twizel is representative of other South Island nodes.

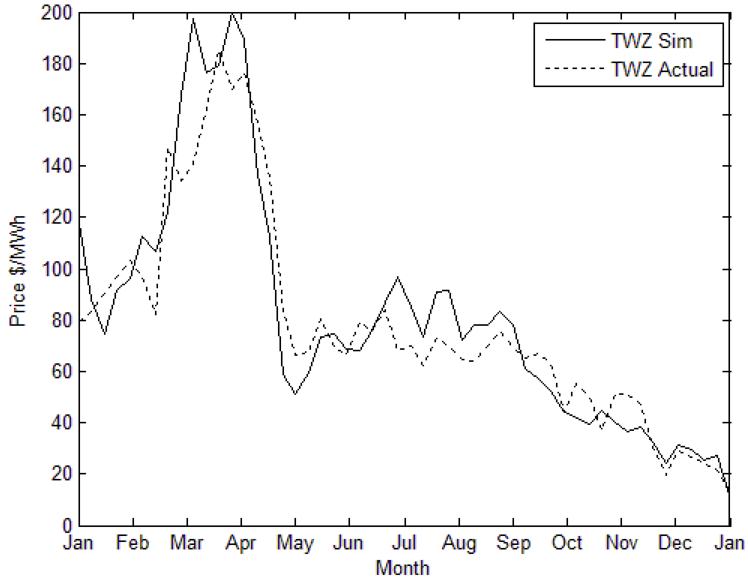


Figure 6: Average Weekly Price Comparisons for TWZ

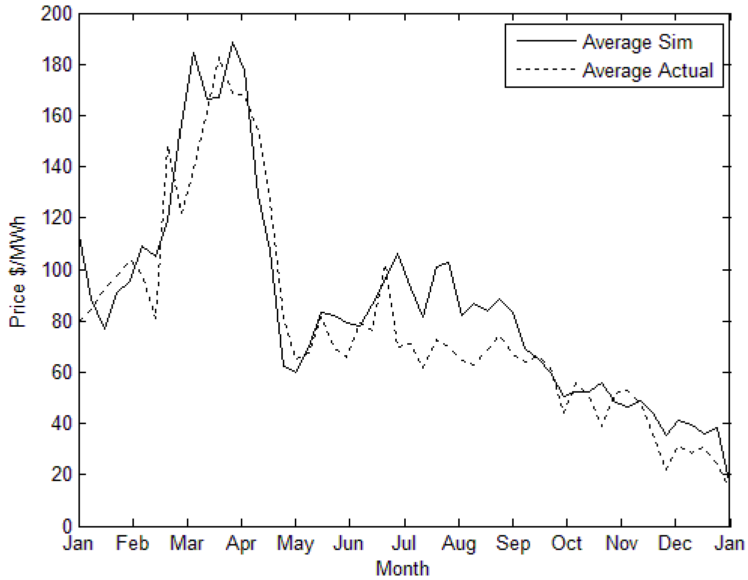


Figure 7: Average Weekly Price Comparison for 5 Key Nodes

A first visual inspection of these graphs indicates that simulated prices track real prices reasonably closely across the entire year with some critical exceptions. For instance, the model clearly does not perform well during the months of July and August. We will discuss possible reasons for the differences later in this section.

The graphs give a visual perspective of the model's pricing performance at selected nodes. To look at the rest of the nodes, we compute summary statistics across the whole network. Table 4 summarizes some of the key statistics of the frequency distributions of the difference between simulated and actual prices for the five key nodes as well as the average of these nodes and the average across all 19 nodes.

The statistics are based on the four simulated prices for each day of the year for 2006. Using the paired t-test, we tested the hypothesis that the simulated prices were significantly different from real prices for each of the 19 nodes across the year 2006, and present the p-values for this test.[40]

NODE	**2006**		
	Average Deviation from Actual Prices ($/MWh) SIM	**Standard Deviation ($/MWh)**	**Paired T-test p-value**
MDN	13.48395	18.39468	0
HEN	12.50246	18.08558	0
OTA	12.21524	17.88627	0
HLY	12.86356	17.54086	0
WKM	9.689394	16.28694	0.0001
NPL	12.52776	16.0065	0
TKU	10.00109	15.89526	0
WHI	10.00931	16.42884	0.0001
BPE	11.42095	16.02676	0
HAY	10.27436	16.26158	0
Average NI Nodes	**11.49881**	**16.64282**	**0**
STK	3.476456	18.3936	0.1798
KIK	3.13215	18.24332	0.2233
IGH	2.311941	18.52769	0.3725
ISL	4.064528	17.61789	0.1020
TWZ	2.935148	16.2549	0.1996
ROX	1.491415	15.90016	0.5016
HWB	1.591476	16.9807	0.5009
TIW	-0.08954	16.59953	0.9687
MAN	3.372776	17.02506	0.1586
Average SI Nodes	**2.476261**	**17.13773**	**0.3026**
Average 5 Key Nodes	**7.35123**	**16.60581**	**0.0013**
Average All Nodes	**7.224969**	**15.80098**	**0.0018**

Table 4: Price Difference Frequency Statistics

Table 4 reinforces the intuition from the figures above. Not just the Otahuhu node, but all the North Island nodes record simulated prices well above actual prices – on average a difference of $11.50/MWh. As such, the paired t-test was able to comprehensively reject the null that the two series had the same mean. On the other hand, we recorded an average difference of at most $3.50 at the South Island nodes, and the paired t-test was unable to reject the null for any of these nodes at the 10% level. The difference between the two islands is significant, and may point to an error in the data or the calibration process. Given the calibration of parameters was based upon the average price of 5 key nodes across New Zealand, we might expect that North Island simulated prices would be higher than actual, and South Island simulated prices would be lower than actual, so that on average

[40] Note that this test is performed on the respective weekly averaged prices for both the simulated and actual price series. As we note previously, the goal of this paper is to try and replicate the weekly average prices. There is a lot of variability in half-hour prices that this model specifically does not capture, which would result in higher standard deviations if all the half-hour prices were used in the test.

simulated price across all five nodes is close to the actual. Instead, we see a persistent positive bias across the country, and getting noticeably larger moving from the southernmost node to the northernmost node.

It can also be seen from the table that simulated prices for any given half hour period show quite a bit of dispersion from actual prices – the standard deviations are between $30-$40/MWh for each node. It is our view that this is to be expected. The actual data shows significant variance. Comparing the same half hour period across a typical week it is not unusual to see similar variations in price which are hard to explain by looking at demand changes or generation availability. It seems to be the case that there is some degree of trader noise which accounts for at least some of the variation of prices seen in the data. The agent based model certainly has a great deal of trader noise. Adding more random seeds for each price point prediction might reduce the noise in the agent-based model somewhat, but as we have already noted (at the end of Section 4), we considered the use of additional seeds and concluded it had little added benefit, if any.

As noted above, there is a particularly large gap in prices during the July/August period. We re-ran the paired t-test excluding these months. The results showed no major differences. North Island simulated prices were still significantly higher than real prices at the 1% level. South Island price differences remained low.

5.3 Normality

One of the assumptions of the paired t-test is normality of each of the data series being compared. Figure 8 shows the distribution of the difference between the actual price and the simulated price across all 19 nodes for all the simulated prices for 2006 – total of 27,740 simulated prices. A standard normal distribution has been fitted to the actual distribution. It is clear that the distribution of the price deviations is similar to a standard normal distribution, albeit with a fatter tail for positive price differences.

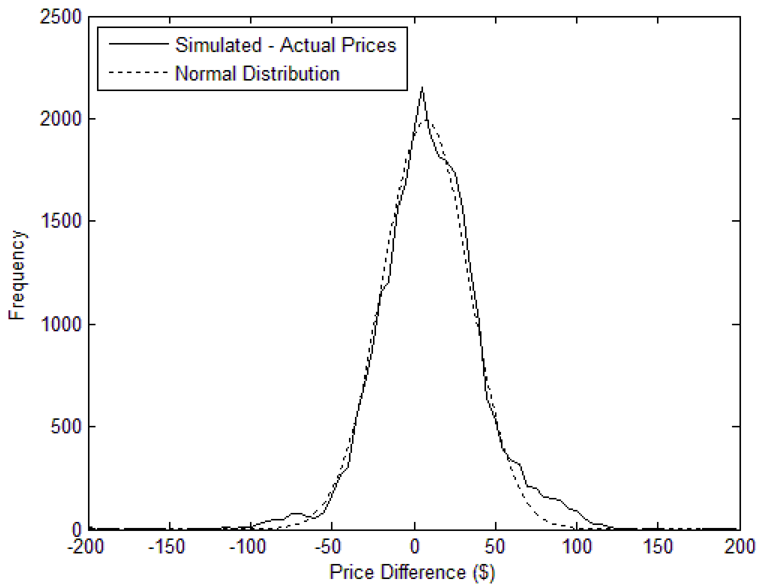


Figure 8: Price Difference Distribution Across All Nodes

We check this formally for all nodes using the Shapiro Test for normality. The test does not reject the null at every node at the 5% level.

NODE	Shapiro Test for Normality – p values	NODE	Shapiro Test for Normality – p values
MDN	0.22798	STK	0.40041
HEN	0.18792	KIK	0.40415
OTA	0.189	IGH	0.44967
HLY	0.1887	ISL	0.09178
WKM	0.38663	TWZ	0.0868
NPL	0.50087	ROX	0.11194
TKU	0.60528	HWB	0.21867
WHI	0.41458	TIW	0.21981
BPE	0.83729	MAN	0.11414
HAY	0.8497		
Average NI Nodes	**0.4545**	**Average SI Nodes**	**0.14635**
Average 5 Key Nodes	0.9729	Average All Nodes	0.95361

Table 5: Shapiro Test Results

5.4 Price Duration Curves

Another way to visually compare the predicted and actual prices is by constructing a price duration curve, where all 1460 simulated prices for the year are ordered from least to highest. Whereas the previous graphs depicted model performance across time, the price duration curves allow to see how the model performs at specific price points. Figure 9 and Figure 10 show the price duration curve for the Otahuhu and Twizel nodes respectively.[41]

[41] One outlier from the actual data has been omitted when the price rose above $1000 for a single trading period.

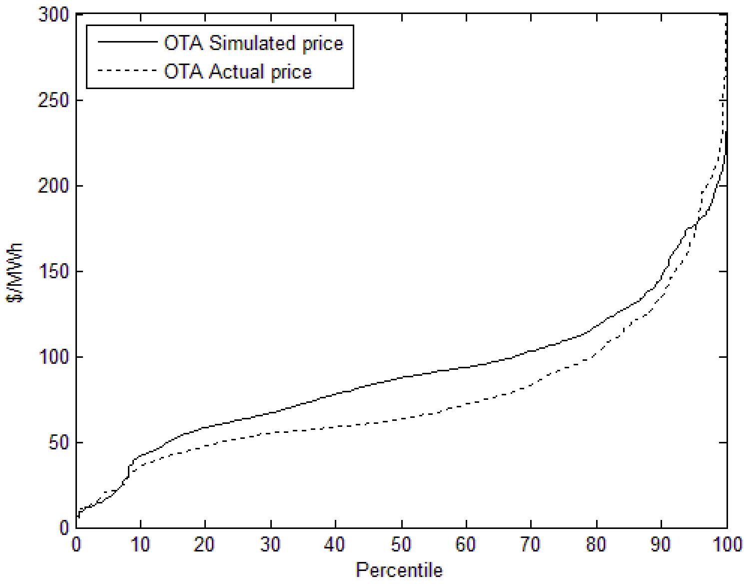


Figure 9: Price Duration Curve for OTA

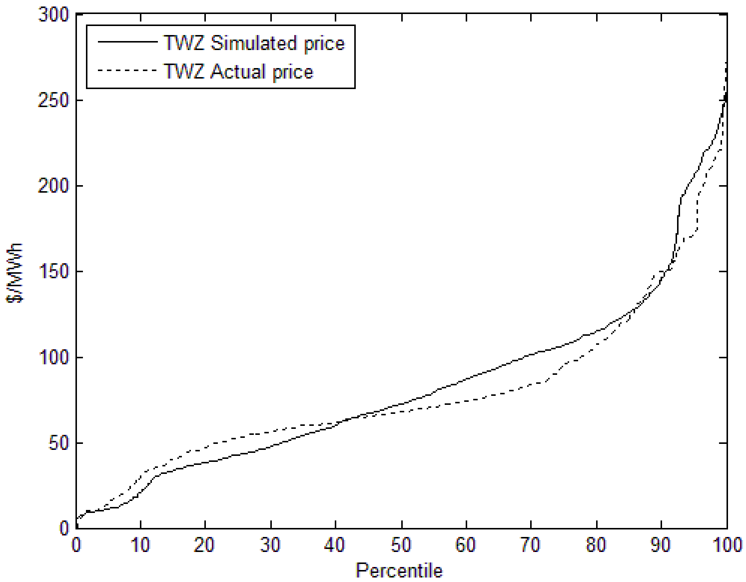


Figure 10: Price Duration Curve for TWZ

Figure 11 depicts the price duration curve for a composite average of the five key nodes.

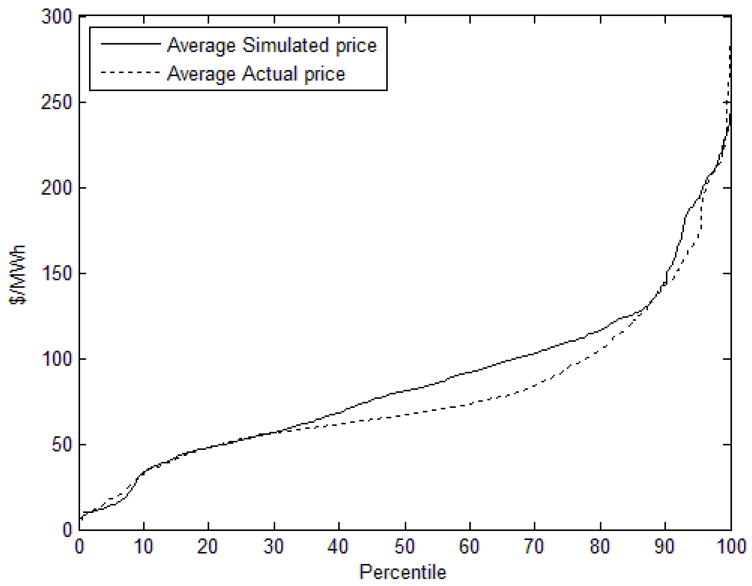


Figure 11: Price Duration Curve for 5 Key Nodes

It is immediately clear from these graphs that the model performs better at some price points than others. Looking at the simulated price duration curve for very low and very high prices, the model does well on average. In the midrange, however, there is a positive bias on average across the five key nodes. This bias can be partially accounted for when we consider the data points used for calibration. Actual prices across all nodes ranged between 31.56 and 72.69 across the four periods chosen for calibration. A quick glance at the graph shows that the model indeed does an excellent job at predicting prices when simulated prices are in this range (recall that the calibration was targeted at the average of the 5 key nodes). It is somewhat too low at OTA and somewhat too high at TWZ).

For the sake of completeness we should point out that the price duration curves are not comparing simulated and actual prices for the same period. Figure 12 below compares the actual same period prices to the simulated price duration curve. For simulated prices between $30-$70 on actual prices across all five nodes are on average $0.72/MWh too high. The actual prices are scattered above or below the simulated price curve.

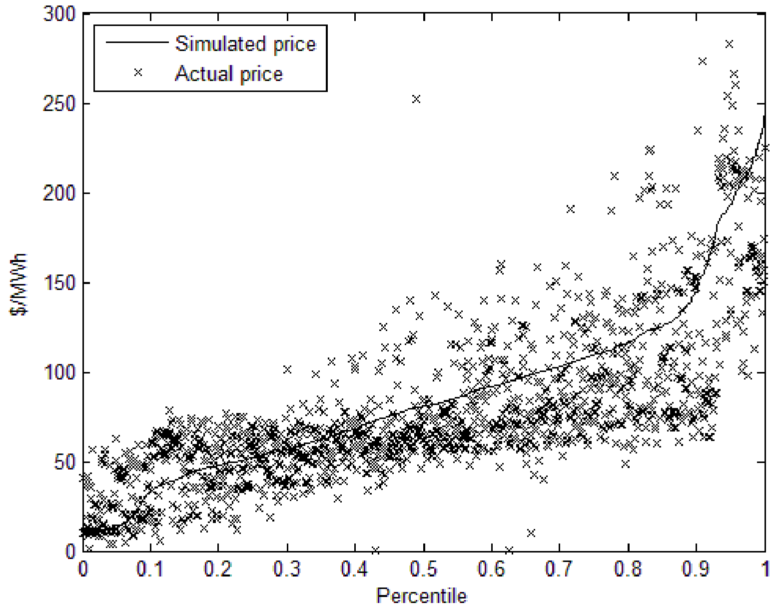


Figure 12: Simulated Price Duration Curve with Actual Prices for 5 Key Nodes

5.5 Discussion

The graphs and statistics above highlight several important points. South Island prices are accurately predicted, but North Island prices are not. There is a positive bias across all predicted prices, which is accentuated as you move northwards across New Zealand. Finally the model's predictive power is significantly weakened in situations where prices rise above $70/MWh.

As noted above, the first two issues point to either a flaw in the calibration process, a flaw in the quality of network and/or outage data, or both. There are several explanations. One possibility is that the transmission data used, particularly in the North Island, is too conservative. Conservative estimates for capacity and line losses would inflate the degree of market power and lead to the simulation biased towards higher prices. Such conservatism might arise from our decision to omit any transmission lines smaller than 220kV, thus understating total regional capacity. This argument is supported by the fact that the North Island has far more 110kV and 66kV lines, particularly in the upper North Island, thus the degree of understatement would be higher there. Another possibility is that the simplifying assumptions used to calculate resistance and reactance for aggregated lines causes a systematic downward bias. Again, this effect would be more pronounced in the North Island, where there are far more loop flows and smaller nodes to account for than the sparsely populated South Island.

The representation of the HVDC line between the two islands is particularly suspect, given the big difference in the accuracy of predicted prices between the North and South Islands. The capacity of the HVDC line across the years 2006 onwards is not clear-cut. Ostensibly, the capacity of the line was 1200MW until the breakdown of Pole A mid-2007 lead to a drop in capacity to 500MW. A couple of months later, some switching of the undersea cables to Pole B lead the nominal capacity to rise to 700MW. This last value is the nominal capacity built into the model. The calibration of the model was performed over values in late 2007 and 2008, for which this value is likely 'close enough'. However, for predicting prices in 2006, it might seem reasonable to use the 1200MW figure, minus an allowance for reserves (much of which are in the South Island). Doubt arises in that officials knew that Pole A was outdated, and could have informally set lower limits on the line to minimize possible overheating. These limits of course are not visible.

If some of our line capacities in the model network were too low then model simulations would have more periods with congestion and hence more price separation between nodes than observed. During the period where modelled prices deviate the most from actual prices (July/August) the simulated price in Auckland (OTA) was on average $20/MWh higher than the South Island node of Twizel which lends support to the hypothesis.

Another possibility is that firms are not in practice taking advantage of their natural market power. Statements by generating firms following the 'price spike' event of March 26th, 2011, lend credence to this theory.[42] In effect the generators claimed that they did not take advantage of opportunities to exercise market power when certain transmission lines were out of service. The agents in the SWEM model would of course take full advantage of any market power. The parameters used of course were calibrated to replicate the amount of market power actually exercised in the market, however this was based on the assumption that all firms and plants behave in the same way. If it happened that North Island plants did not take advantage of available market power as much as South Island plants, then the model as stated would not be able to replicate prices in both islands with a single set of parameters.

To some degree, the calibration process should be able to compensate for any of the above noted flaws. For instance, if the capacity of the HVDC line was indeed understated, the calibration should still adjust so that on average (at the 5 key nodes) the predicted prices match the real prices. We see from the price duration curves that this did indeed happen in the range of the calibrated period's actual prices – $30 to $70. We would have expected however, that the resulting prices would be slightly high in the North Island – and slightly low in the South Island. Similarly, if firms did not exercise market power in the North Island (but did in the South Island), the calibration process should have adjusted to have prices somewhat low in the North Island, and somewhat high in the South. We do not see either of these results. Instead we see a positive bias across New Zealand, albeit with a larger bias in the North than the South. This suggests that even if the network data is incorrect, there is still a separate flaw in the calibration process.

As discussed in Section 4, the calibration process took two stages: one to calibrate the behavioural parameters, and one to calibrate the water values. We have already seen evidence that the first stage performed relatively well. Prices in the $30 to $70 range were accurately replicated by the model for the verification year. Significant error crept in when we considered prices outside this range, particularly above $70, and this is likely due to the calibration of water values. As we acknowledged early on, deriving an opportunity cost of water is a difficult and controversial process. We cannot simply extend the first stage of the calibration to incorporate higher prices, because prices higher than $70 are almost invariably associated with higher water values, whereas the first stage of calibration requires an invariant water value across calibrating periods. Similarly, adding more data points to the four chosen would not help greatly, since more points in the $30 to $70 range would only reinforce the part of the model that works well already.

[42] On the 26th March 2011, three major transmission lines into Auckland from the south were taken off-line for maintenance. This left Genesis' Huntly plant as the only plant able to satisfy residual demand in Auckland. Genesis placed offers in the order of NZ$20,000/MWh to dispatch Huntly. This compares to the usual price of around $70/MWh at the Huntly node (HLY) and the previous maximum of about $5000/MWh. Other generators complained, including Meridian who made a Undesirable Trading Situation claim, saying that 'Other generators in the same position have not acted in the same way'. This document can be found at http://www.ea.govt.nz/document/13124/download/act-code-regs/uts/decisions-and-claims/.

Finally we should note two additional factors that impact prices, factors the model simply does not account for explicitly. As noted early on we have not included contract positions explicitly in the model. We account for these implicitly through our choice of behavioural parameters, however this assumption is problematic if contract positions change. Another possible hypothesis is that contract positions may change in a systematic way throughout the year. For example they may be a function of demand which is highest in the winter months. This is a topic we plan to revisit in the future, although as we discussed above contract positions are in general not publicly available.

We also saw that plotted water values for winter are scattered much more than the summer values, which may be due to inflow expectations changing. For example if market players thought that inflows over the next month were likely to be high they may value water less than our model predicts. Thus another possibility is that information which real world agents have has influenced their expectations and hence the value they placed on water. This could systematically bias our prices upwards during such a period. Ideally, the calibration of water values would be extended to include water flows, as well as accumulated water storage, and perhaps searching for a way to calibrate both behavioural and water value parameters simultaneously. However, we feel, with a considerable degree of confidence, that the algorithm as it stands would perform well in a network with low hydro generation, and accurate network data.

6. Conclusions

The goal of this paper was to discover whether a carefully constructed and calibrated agent-based model using the modified Roth and Erev algorithm could indeed predict short-run prices in a realistic sized electricity market. To this end, we constructed a model based on the New Zealand electricity market, using an extensive dataset collected and provided by the Electric Power Optimization Centre and the New Zealand Electricity Authority. We calibrated the key parameters in this model against 2007-8 price data, and then tested the predictive ability of the model against actual 2006 price data.

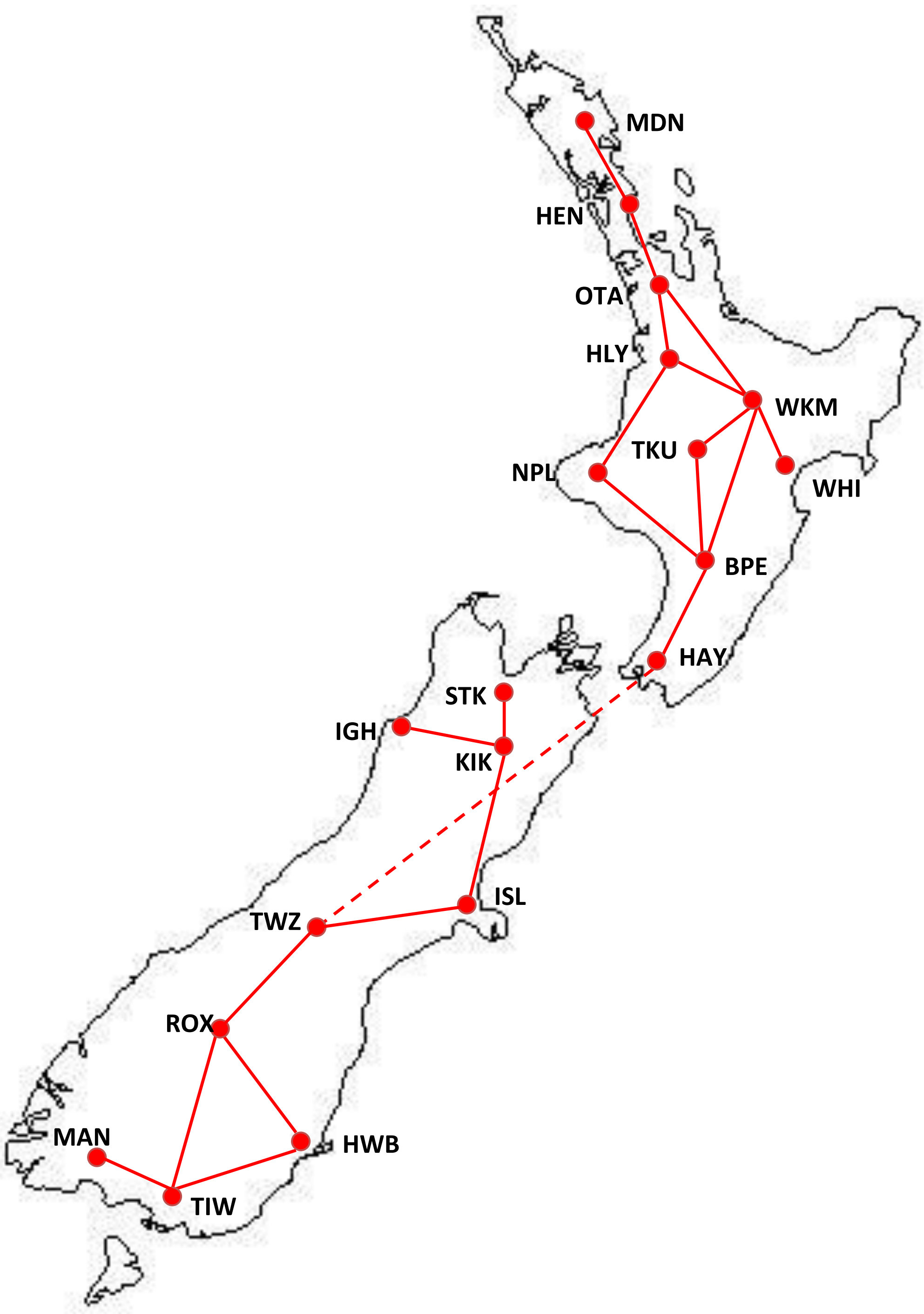
We find that our model can predict the price to a startling degree of accuracy on average at any South Island node, across a larger range of demand and water values. Such variation as exists can likely be explained by trader variation on a day by day basis, which the model cannot capture. However, the model has less consistent success in predicting prices at North Island nodes, with a particularly large blind spot in the July-August timeframe, and a general positive bias across all North Island nodes.

As we discussed in Section 5, many of the differences in the results can be traced to two sources. One source is in the network data – the aggregation of many lines into a few for the model likely led to conservative estimates for line losses and capacity, causing a bias in the output. In particular, it is likely the HVDC was underestimated. Such an issue is in principle, correctable, with a better source of data, or better aggregation procedure. Sensitivity analysis in this area could guide the process of creating a more accurate dataset – perhaps at the least going back to the original to dataset to include the 110kV transmission lines. The other source of pricing inaccuracy can be traced to the calibration of water values. As we pointed out in Sections 3 and 4, our measure of water value relied solely on the water storage in the model, and critically ignored rainfall and river flows, both of which would drive expectation as to *future* storage, and thus determine the opportunity cost of water.

The model worked well when water values were low, at least in replicating average prices across the 5 key nodes (which we recall, was the basis on which the model was calibrated). Based upon a calibration period from 2007-8, the model replicated the average weekly price in 2006 across the 5 key nodes very accurately. This gives us confidence in the underlying approach we have taken here, and supports our conclusion that such models can predict short-run prices far better than has thus far been acknowledged in the literature. We feel that for a thermal dominated network, with good network data availability, this model could accurately reproduce short-run prices across a year. Overall, we conclude that this provides strong evidence that the behavioural assumptions underlying the modified Roth and Erev algorithm are valid.

This research highlights many opportunities for further research, two of which we discuss here. The first concerns the opportunity cost of water. Our approach to estimating the opportunity cost was to use our model assuming no water value to work out what prices would have been in this counterfactual, then search for what water cost caused simulated prices to come closest to actual prices for a given period. Since this process was time independent, we would like to revisit this topic by incorporating a second dimension – future expectations of inflows – into the calibration process. The second concerns the way the algorithm handles portfolios of plants. Due to computational constraints, we followed Weidlich's method for one firm offering multiple plants. There is a small but growing literature in offering supply curves and step supply curves in agent-based models, and we would like to draw upon this to see if it improves firms' ability to offer multiple plants strategically in the model. We hope to revisit both these topics in future work.

Appendix A: Network Map



Appendix B: Generator and Plant Details

Owner	Name	Node	Type	Capacity (MW)	Must Run Capacity (MW)	Fuel Cost ($/MW hr)	Operating Cost ($/MW hr)
Contact	Clyde	ROX	Hydro	380	90	0	10
	New Plymouth	NPL	Gas - open cycle	88	0	61.75	8
	Ohaaki	WKM	Geothermal	88	88	0	7.3
	Otahuhu A	OTA	Gas - combined cycle	0	0	41	5
	Otahuhu B	OTA	Gas - combined cycle	400	0	49.35	4.25
	Poihipi	WKM	Geothermal	52	45	0	7.3
	Roxburgh Unit 1	ROX	Hydro	111	40	0	10
	Roxburgh Unit 2	ROX	Hydro	194	40	0	10
	Stratford	NPL	Gas - combined cycle	370	0	43.8	4.25
	Te Rapa Cogen	HLY	Cogeneration	50	0	45.5	5
	Wairakei	WKM	Geothermal	176	176	0	7.3
Genesis	Huntly Unit 1	HLY	Coal - conventional	207	0	45.5	9
	Huntly Unit 2	HLY	Coal - conventional	207	0	45.5	9
	Huntly Unit 3	HLY	Coal - conventional	207	0	45.5	9
	Huntly Unit 4	HLY	Coal - conventional	211	0	45.5	9
	Huntly Unit 5	HLY	Gas - combined cycle	379	0	46.02	4.25
	Huntly Unit 6	HLY	Gas - open cycle	0	0	61.75	8
	Rangipo	WKM	Hydro	111	30	0	10
	Tokaanu	TKU	Hydro	211	0	0	10
	Anchor Diary Plant	HLY	Cogeneration	24	0	45.5	5
	Waikeremoana	WHI	Hydro	121	12	0	10

Meridian	Aviemore	TWZ	Hydro	194	20	0	10
	Benmore	TWZ	Hydro	475	30	0	10
	Whareroa	NPL	Cogeneration	34	0	45.5	5
	Manapouri	MAN	Hydro	114	0	0	10
	White Hill	TIW	Wind	1	0	0	2.5
	Ohau A	TWZ	Hydro	232	24	0	10
	Ohau B	TWZ	Hydro		20	0	10
	Ohau C	TWZ	Hydro	187	20	0	10
	Tekapo A	ISL	Hydro	22	0	0	10
	Tekapo B	TWZ	Hydro	141	0	0	10
	Te Apiti	BPE	Wind	0	0	0	2.5
	Waitaki	TWZ	Hydro	79	65	0	10
MRP	Waikato 1	WKM	Hydro	69	26	0	10
	Ohakuri	WKM	Hydro	260	98	0	10
	Maraetai	WKM	Hydro	602	288	0	10
	Kapuni Geothermal	WKM	Geothermal	90	90	0	7.3
	Southdown Unit 1	OTA	Cogeneration	131	0	57.75	4.25
	Southdown Unit 2	OTA	Gas - open cycle	42	0	62.25	8
	Rotokawa	WKM	Geothermal	33	33	0	7.3
Todd	Whareroa (Todd)	NPL	Cogeneration	35	0	45.5	5
	Kapuni	NPL	Gas - combined cycle	16	0	41	5
	Aniwhenua	WKM	Hydro	11	8	0	10
	Mangahao	BPE	Hydro	18	0	0	10
	Branch River	STK	Hydro	11	0	0	10
Trustpower	Highbank	ISL	Hydro	25	0	0	10
	Tararua Wind Farm 1	BPE	Wind	11	0	0	2.5
	Waipori	HWB	Hydro	82	0	0	10

	Cobb	STK	Hydro	28	0	0	10
	Coleridge	ISL	Hydro	34	0	0	10
	Patea	NPL	Hydro	19	0	0	10
	Kinleith	WKM	Cogeneration	40	0	0	11.8
	Kumara	IGH	Hydro	6	0	0	10
	Bay of Plenty	WKM	Hydro	125	0	0	10
TUAR**	Mokai	WKM	Geothermal	98	98	0	7.3
EC**	Whirinaki	WHI	Oil	156	0	298.62	9.6
NZWF**	New Zealand Wind	BPE	Wind	1	0	0	2.5
ALNT**	Glenbrook	OTA	Cogeneration	54	0	45.5	5

**These firms bid at marginal cost in the SWEM model.

Appendix C: Network Details

Nodes		Number piecewise approximations to line losses	Capacity (MW)	Reactance (Ohms)	Resistance (Ohms)	Flow Type
MDN	HEN	3	589	0.00007	0.04289	AC
HEN	OTA	3	1492	0.000014	0.0097	AC
OTA	HLY	3	1687	0.0000095	0.01848	AC
OTA	WKM	3	808	0.000104	0.05005	AC
HLY	WKM	3	404	0.000335	0.08895	AC
HLY	NPL	3	754	0.000228	0.12056	AC
WKM	TKU	3	616	0.000083	0.03132	AC
WKM	WHI	3	953	0.000036	0.0413	AC
WKM	BPE	3	239	0.00014	0.21519	AC
NPL	BPE	3	710	0.000085	0.04762	AC
TKU	BPE	3	404	0.000167	0.0748	AC
BPE	HAY	3	1338	0.000035	0.02312	AC
STK	KIK	3	478	0.000039	0.02238	AC
KIK	IGH	3	118	0.00027	0.19075	AC
KIK	ISL	3	350	0.000095	0.06731	AC
ISL	TWZ	3	1100	0.00005	0.03894	AC
TWZ	ROX	3	896	0.000036	0.04569	AC
ROX	HWB	3	770	0.000045	0.05763	AC
ROX	TIW	3	461	0.000099	0.03203	AC
HWB	TIW	3	504	0.00023	0.08616	AC
TIW	MAN	3	936	0.000029	0.02228	AC
HAY	B*	6	700	0.000098	0.01	DC
TWZ	B*	6	700	0	0.01	DC

*B is a dummy node along the HVDC to allow for both Pole A and Pole B to be represented.

Appendix D: Simplified Network matched with Actual Substations.

Swem Node	NZEM Substations that act as Grid Exit Points
MDN	BRB0331 DAR0111 KEN0331 KOE0331 KTA0331 MDN0141 MPE0331 MTO0331 WEL0331
HEN	ALB0331 ALB1101 HEN0331 HEP0331 HPI2201 SVL0331 ;
OTA	BOB0331 BOB1101 GLN0331 GLN0332 MER0331 MNG0331 MNG1101 OTA0221 OTA1101 OTA1102 PAK0331 PEN0221 PEN0331 PEN1101 ROS0221 ROS1101 TAK0331 WIR0331
HLY	HLY0331 TMN0551 TWH0331
WKM	CBG0111 EDG0331 HAM0111 HAM0331 HAM0551 HIN0331 HTI0331 KAW0111 KAW0112 KAW0113 KIN0111 KIN0112 KIN0331 KMO0331 KPU0661 LFD1101 LFD1102 MAT1101 MTI2201 MTM0111 MTM0331 NAP2201 OKI0111 OKE1101 OWH0111 ROT0111 ROT0331 TGA0111 TGA0331 TKH0111 TMI0331 TMU0111 TRK0111 WAI0111 WHU0331 WKM2201 WKO0331 WRK0331
NPL	CST0331 HUI0331 HWA0331 HWA0332 MNI0111 MRA0111 NPL0331 OPK0331 SFD0331 WVY0111
TKU	TKU0331
WHI	FHL0331 GIS0501 RDF0331 TUI0111 WHI0111 WHI2201 WRA0111 WTU0331
BPE	BPE0331 BPE0551 BRK0331 DVK0111 LTN0331 MGM0331 MHO0331 MTN0331 MTR0331 NPK0331 OKN0111 ONG0331 TNG0111 TNG0551 TWC2201 WDV0111 WDV1101 WGN0331 WPW0331
HAY	CPK0111 CPK0331 GFD0331 GYT0331 HAY0111 HAY0331 KWA0111 MLG0111 MLG0331 MST0331 PNI0331 PRM0331 TKR0331 UHT0331 WIL0331 WWD1102 WWD1103
STK	ARG1101 BLN0331 MOT0111 MPI0661 STK0331
KIK	KIK0111 MCH0111
IGH	DOB0331 GYM0661 HKK0661 KUM0661 IGH1101 WMG1101 WPT0111 ORO1101 ORO1102 RFN1101 RFN1102 ATU1101
ISL	ABY0111 ADD0111 ADD0661 APS0111 ASB0331 ASB0661 ASY0111 BRY0111 BRY0661 CLH0111 COL0111 CUL0331 HOR0331 HOR0661 ISL0331 ISL0661 KAI0111 KKA0331 MLN0661 MLN0664 OTI0111 PAP0111 PAP0661 SBK0331 SPN0331 SPN0661 TIM0111 TKA0331 TMK0331 WPR0331 WPR0661
TWZ	BPD1101 BPT1101 NSY0331 OAM0331 STU0111 TWZ0331 WTK0331
ROX	CML0331 CYD0331 FKN0331
HWB	BAL0331 BWK1101 HWB0331 HWB0332 PAL0331 PAL1101 SDN0331 TMH2201
TIW	BDE0111 EDN0331 GOR0331 INV0331 NMA0331 TWI2201
MAN	

References

Banal-Estañol, A. and Augusto Rupérez Micola. Composition of electricity generation portfolios, pivotal dynamics and market prices. *Management Science* 55(11):1813-1831, 2009

Bunn, D., and Day, C. Computational modeling of price formation in the electricity pool of England and Wales *Journal of Economic Dynamics and Control* 33:363-376, 2009

de Silva, R., Archer, M., and Hume, D. Inter-Area Transmission Capacity, *A Report prepared for the New Zealand Electricity Commission Transmission Advisory Group*, 2006. Available at http://www.ea.govt.nz/document/3742/download/

Erev, I. and Roth, A. Predicting how people play games: reinforcement learning in experimental games with unique, mixed-strategy equilibria. *American Economic Review* 88 (4):848-881, 1998

Guerci, E., Rastegar, M. and Cincotti, S. Agent-based modeling and simulation of competitive wholesale electricity markets. *In Handbook of Power Systems II*, pp241-286. Springer-Verlag, Berlin, 2010

Genoese, M., F. Sensfuß, D. Möst, and Rentz, O. Agent-based analysis of the impact of CO_2 emission trading on spot market prices for electricity in Germany. *Pacific Journal of Optimization* 3:401-423, 2007

Joyce, S. Comparing Agent Learning Algorithms in a Transmission Constrained Electricity Network, *University of Auckland Department of Economics Master's Thesis*, 2010

Möst, D. and Genoese, M. Market Power in the German wholesale electricity market. *The Journal of Energy Markets* 2(2):47-74, 2009

Nicolaisen J., Petrov, V., and Tesfatsion, L. Market power and efficiency in a computational electricity market with discriminatory pricing. *IEEE Transactions on Evolutionary Computation* 5(5):504-523, 2001

Philpott, A.B., G. Zakeri, J. Khazaei , and Z. Guan. Productive inefficiency of electricity markets with hydro generation, Utilities Policy, 18, 4, 174-185, 2010

Rastegar, M.A., E. Guerci, and S. Cincotti. Agent-based Model of the Italian Wholesale Electricity Market. *In Proceedings of EEM 2009: 6th International Conference on the European Energy Market*, pp.1-7, 27-29 May 2009. Available at http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5207128&isnumber=5207110

Sensfuß, F., M. Ragwitz, and M. Genoese. The merit-order effect: a detailed analysis of the price effect of renewable electricity generation on spot prices in Germany. *Energy Policy* 36(8), 3076-3084, 2008

Sueyoshi, T., and G.R. Tadiparthi. Agent-Based Approach to Handle Business Complexity in U.S. Wholesale Power Trading. *IEEE Transactions on Power System* 22 (2), p 532–542, 2007

Sueyoshi, T. and G. Tadiparthi. An Agent-based Decision Support System for Wholesale Electricity Markets. Decision Support Systems, 44(2), p. 425–446, 2008

Sun, J. and Tesfatsion, L. Dynamic testing of wholesale power market designs: An open-source agent-based framework. *Computational Economics* 30:291-327, 2007

Tipping J, Read E. G., McNickle D. The incorporation of hydro storage into a spot price model for the New Zealand electricity market. Presented at the Sixth European Energy Conference: Modeling in Energy Economics and Policy. Zurich, 2004. Available at http://www.mang.canterbury.ac.nz/research/emrg/

Veit, D.J., A. Weidlich, J. Yao, S.S. Oren. Simulating the dynamics in two-settlement electricity markets via an agent-based approach. *International Journal of Management Science and Engineering Management 1*(2)83-97, 2006

Veit, D. and A. Weidlich, and J. A. Krafft. An agent-based analysis of the German electricity market with transmission capacity constraints. *Energy Policy* 37(10):4132–4144, 2009

Weidlich, A. Engineering Interrelated Electricity Markets. Physica-Verlag, Heidelberg, 2008

Weidlich, A. and Veit, D. A critical survey of agent-based wholesale electricity market models. *Energy Economics* 30:1728-1759, 2008(a)

Weidlich, A. and Veit, D. Analyzing interrelated markets in the electricity sector -The case of wholesale power trading in Germany. Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE pp.1-8, 20-24 July 2008(b). Available at http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4596728&isnumber=4595968

Wolak, F. An Assessment of the Performance of the New Zealand Wholesale Electricity Market. *Report to the New Zealand Commerce Commission*, 2009

Young, D. Kirchoff's Laws in Electricity Markets: Can Experimental Subjects Imitate Generators? *Report to the National Science Foundation*, 2008

Young, S. The Effect of Embedded Generation on the New Zealand Electricity Market System. *University of Auckland Thesis 97-252*, 1997