Within-island Basis Risk

Prepared by Energy Link

for

The Electricity Authority

May 2012

Quality Assurance Information

Name	EA WIBR study May-12 v3.docx
File reference	E-EA-950
Issue Status	Issue 1
Issue Date	6 June 2011
Client	Energy Link Ltd

Definitions

The following terms, abbreviations and acronyms are used in this report. Node abbreviations are not included in this table.

Authority	Electricity Authority
CFD	Contract for differences (a type of swap contract)
Code	Electricity Industry Participation Code
em6	The market information system operated by the Energy Market Services division of Transpower
FTR	Financial transmission right
GIP	Grid injection point
GXP	Grid exit point
Hedge node	The node at which a party has a hedge contract
IIBR	Inter-island basis risk (a.k.a. inter-island LPR)
ICPD	Inter-cluster price difference
INPD	Inter-nodal price difference
LCE	Losses and constraints excess
LPR	Locational price risk
LPRTG	Locational Price Risk Technical Group
NI	North Island
OTC	Over-the-counter. Refers to the market for hedges in which hedging instruments are traded directly between the parties to the hedge (as opposed to being traded on an organised and regulated exchange such as a futures market).
PDS	Pre-dispatch Schedule
Physical node	The node at which a party has an exposure to the spot price by virtue of a contract to buy or sell electricity at the prevailing spot price
POCP	Planned Outage Coordination Process (see pocp.redspider.co.nz)
RMS	Root mean square – used to calculate the change in offered quantities by trading period, and the error in a simple demand forecast
SI	South Island
SPD	Scheduling, Pricing and Dispatch model
SWE	Spring washer effect
WIBR	Within-island basis risk (a.k.a. within-island LPR or intra-island LPR)
WITS	Wholesale Information and Trading System

Contents

Figures

Tables

1 Introduction

The Electricity Authority (the Authority) amended the Electricity Industry Participation Code (the Code), with effect from October 2011, to allow for the introduction of inter-island financial transmission rights (FTRs). An RFP was issued in August 2011 for the role of FTR Manager and Energy Market Services, a division of Transpower, was appointed to this role in April 2012. The inter-island FTRs will hedge the price difference between two key nodes, Otahuhu in the North Island (NI) and Benmore in the South Island (SI), and the FTRs will be funded from the losses and constraints excess (LCE). The FTRs will assist in managing the financial risks associated with inter-island basis risk[1] (IIBR, a.k.a. inter-island locational price risk or LPR).

The Authority is now considering options to develop a solution within the NI and SI for the management of within-island basis risk (WIBR, a.k.a. intra-island LPR). Energy Link was engaged late in 2011 to advise the Locational Price Risk Technical Group (LPRTG) of further analysis required to assess the need for LPR hedge instruments on a within-island basis.

Energy Link was then engaged in January 2012 to analyse WIBR using a three phased approach:

1. **Phase 1: Identify Clusters**
   Determine a set of regions of the grid (or ‘clusters’) within which basis risk is below a specified threshold, and between which basis risk is above the threshold. In essence, a cluster represents a group of nodes between which basis risk is sufficiently low that no WIBR hedging instrument is required within the cluster. WIBR hedging instruments could, however, be considered for use in hedging between clusters.

2. **Phase 2: Process WIBR Drivers**
   Based on the clusters defined in Phase 1, analyse market data to determine the underlying factors (the ‘WIBR drivers’) that may cause basis risk between clusters. Process the WIBR driver data and determine relationships between inter-cluster price differences (ICPDs) and each driver.

3. **Phase 3: Projections**
   Discuss how each driver will change in the foreseeable future and hence draw conclusions about how WIBR may evolve over time.

There are three main outputs from this study. The first is a number of cluster sets, based on half hourly prices over the five year period from 1-Jan-07 through to 31-Dec-11. Each cluster set is derived using a specified threshold of correlation between prices at nodes within each cluster. For example, a threshold correlation of 0.9 gives a set of clusters in each island within which the correlation between half hourly prices over each month in the study period, at each and every node within the cluster, remains at or above 0.9. Clusters were derived using correlation thresholds of 0.7, 0.8, 0.9, 0.95 and 0.99.

[1] Basis risk is a generic term which refers to the risk that changes in two elements of a hedging strategy do not offset each other perfectly.

The second output is a list of the key drivers of basis risk in each island, between the clusters derived with correlation threshold of 0.9.

The third output is a discussion of how the WIBR drivers, and thus WIBR itself, may evolve over the coming years and decades.

This report includes an extensive section 3 which establishes context for the study and describes how various key parameters were selected. Readers who are already fully familiar with issues of risk management and hedging in the electricity market may wish to skip this section and go straight to section 4.

Section 4 gives the results of the cluster analysis and section 5 gives the results of WIBR drivers analysis, primarily through a series of graphs of the frequency and magnitude of high price spread events plotted across the range of each WIBR driver.

Section 6 summarises the results of the WIBR analysis and the implications for WIBR and hedging into the foreseeable future.

2 Summary

An unhedged party can choose whether to hedge or not, but if they do hedge then they are still exposed to some degree of basis risk if their hedge node is not also their physical node.

When hedging at a node distant on the grid, the hedge quantity can be adjusted by the expected (forecast) location factor of the physical node relative to the hedge node, and with this adjustment, the hedge will perform well over a wide range of market conditions as long as the actual location factor remains relatively close to the forecast location factor. Events such as spring washer effect (SWE) that create large and highly unpredictable inter-nodal price differences (INPDs), however, may move location factors far beyond expected values, creating the potential for large, unpredictable and adverse movements in hedge payout.

Nevertheless, there are ‘clusters’ of nodes on the grid within which the correlation between prices is high (on a historical basis), which means that hedging within a cluster has a good chance of minimising the impact of basis risk.

In the first phase of the study, data from 1-Jun-07 to 31-Dec-11 was processed and tested so that clusters could be identified at a number of levels of correlation. In the second phase of the study we selected the clusters within which prices correlated to the tune of at least 0.9 in each month, which gave enough clusters to facilitate meaningful analysis.

For each half hour in the study period in each island, the average price was calculated for each cluster, and then the spread of cluster prices was calculated for each half hour.

Data was also processed by half hour for each of a number of candidate WIBR drivers, selected from a range of parameters which either cause, or are associated with, changes in dispatch and pricing over time. The processed WIBR drivers included:

- total island demand;

- total capacity of circuits in outage;
- total loading on the grid in the island;
- total MW of generation having an outage within the island;
- total power transfer on the HVDC link, with northward flow defined as positive;
- change in offers during the half hour;
- rates of change of the parameters described above;
- time of day, represented by trading period number;
- time of year;
- capacity margin, equal to the total quantity offered for the half hour less the total demand;
- maximum offer price;
- the total of the SIR and FIR price in the relevant island

The study focused on trading periods in which the price spread exceeded 50% of the average cluster price, thus ensuring that the studied periods were all influenced by a binding constraint of some sort and exhibited significant basis risk. In the NI there were 4,514 (5.6%) periods included, and 2,793 (3.5%) in the SI, all of which were classified as 'high price spread' events.

The frequency and magnitude of the half hourly spreads for all high price spread events was then plotted and scored against each WIBR driver, with the resulting scores shown in the following tables. A score above one indicates a degree of association between the driver and either frequency or magnitude.

NI Driver	Frequency Score	Magnitude Score
HVDC Transfer	11.3	1.9
Demand	10.6	1.0
Time of year (month)	7.5	2.2
Circuit Outage	7.5	2.9
Capacity Margin	7.0	0.5
FIR + SIR Price	5.0	2.2
Grid Overload	4.5	0.5
Group Circuit Outage	3.4	1.0
Maximum Offer Price	2.9	0.4
Trading Period	2.8	2.3
Circuit Outage Change	1.6	2.7
Capacity Margin Change	1.0	0.6
Demand Change	0.9	0.4
Offers Change	0.6	0.7
Generation Outage	0.4	9.6
Generation Outage Change	0.4	9.9

SI Driver	Frequency Score	Magnitude Score
Circuit Outage Change	33.3	3.9
Circuit Outage	25.6	5.6
Group Circuit Outage	19.4	7.4
Generation Outage	15.9	8.8
Capacity Margin	13.4	0.3
Time of year (month)	10.0	3.8
HVDC Transfer	7.1	2.6
Trading Period	6.5	2.5
Generation Outage Change	4.3	5.4
Grid Overload	2.9	0.3
FIR + SIR Price	1.3	0.3
Capacity Margin Change	1.0	0.9
Offers Change	0.2	0.5
Demand Change	0.2	1.2
Maximum Offer Price	0.2	0.5
Demand	0.1	0.3

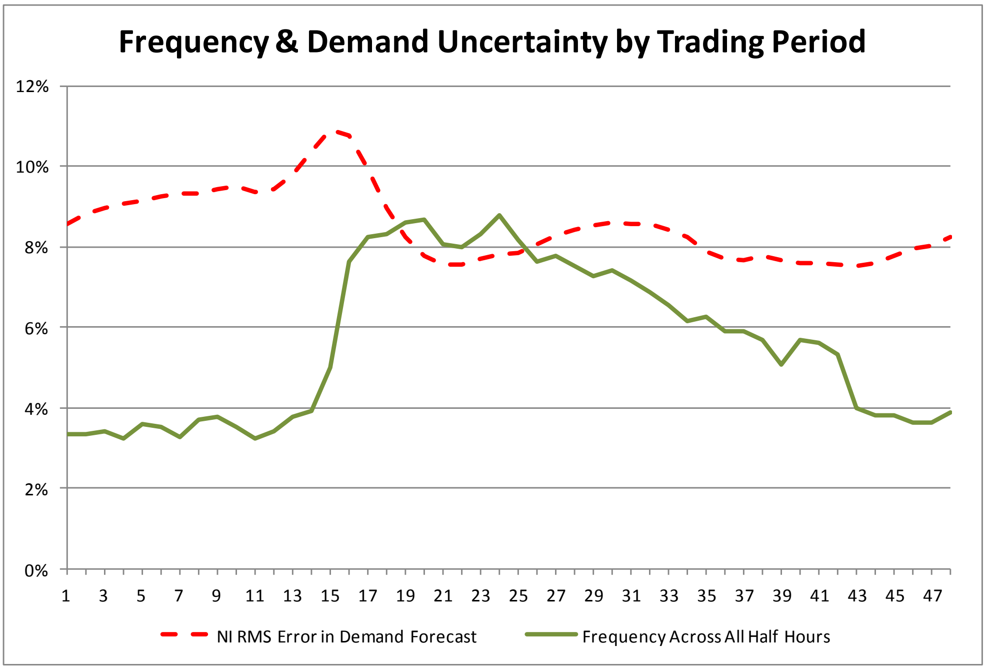
In both islands, the magnitude of high price spread events is considerably less likely than their frequency to be strongly associated with any particular WIBR driver, suggesting that high price spread events, when they occur, have a significant random component in respect of magnitude.

The tables show, however, that a number of WIBR drivers are strongly associated with high price spread events.

When the frequency of high price spread events is plotted against trading period, then a distinct pattern is evident in both islands: the frequency increases sharply in periods 15 and 16 (7:00 am to 8:00 am) to peak in a period shortly after, then falls off through the rest of day to period 43 when it falls sharply again to its overnight level. This effect was investigated in some detail, resulting in the development of a simple method of forecasting demand by trading period using the demand from the same period a week earlier. Figure 1 shows the frequency of high price spread events in the NI by trading period, along with the error in the simple demand forecast[2].

[2] Labeled 'NI RMS Error in Demand Forecast'.

Figure 1 – Frequency of Price Spreads and Demand Uncertainty by Period



It was postulated that leading into the morning demand peak, a number of factors create uncertainty for market participants including rapidly increasing and uncertain demand, along with a rapidly narrowing capacity margin, commencement of the new day's transmission outages, and in the SI, generators beginning planned outages. As the day progresses, it is possible that generators are able to fine tune their offers to reduce the occurrence of any constraints that appear leading up to and during the morning peak which, when combined with lower uncertainty over demand, would explain why the frequency of high price spread events consistently falls off during the day.

Of the sixteen WIBR drivers initially analysed, only two drivers failed to show some degree of positive association with either the frequency or magnitude of high price spread events. The distinct pattern of frequency by trading period also led us to add demand uncertainty as a key WIBR driver.

Projecting forward, the only drivers that might change significantly in future are grid capacity (via upgrades of the AC grid), reserves prices and HVDC transfers (once Pole 3 is commissioned). However, HVDC transfers are also strongly correlated with the frequency high price spread events across the day, so in respect of these three drivers we can only conclude that grid upgrades are likely to reduce WIBR (or at least until demand growth or the building of new generation 'uses up' the new capacity) but to a relatively small extent.

As to the other drivers that are positively associated with high price spread events, some will increase (demand, for example, is forecast to continue growing) and the rest are unlikely to move in a direction that will reduce WIBR (generation outages, circuit outages, capacity margin, wet and dry years).

Taken overall, there are some drivers that will tend to reduce WIBR in the foreseeable future, but on balance there are a greater number of stronger drivers that will tend to keep WIBR the same or to increase it in future. We conclude that WIBR will continue at similar levels into the foreseeable future, although the associations are likely to change over time as grid upgrades are completed.

However, there is a distinct pattern of frequency of high price spread events across the day, which appears to be linked to uncertainty in demand combined with the time of day at which circuit and generation outages commence, and potentially aggravated by the growing quantity of wind generation. It is therefore possible that WIBR will reduce 'at source' with the introduction of new forecast schedules at the end of this month, by providing the market with better forecast information, and particularly concerning the potential impacts of demand: if this occurs then the market will have more time to take action to reduce the occurrence of constraints. The impact of the new schedules can be assessed, in part, by any changes observed in the frequency of high price spread events across the day.

The Authority is currently pursuing work on improving the price formation process, and it may be that certain improvements could also reduce WIBR at source: for example, changing the way that constraints are modelled in SPD. We recommend the scope of this work be reviewed to determine whether or not it should include a measureable reduction of WIBR as an explicit goal.

3 Hedging and Basis Risk

This section is devoted to developing the context for WIBR, hedging strategy, and risk management in the electricity market. The simple examples used in this section illustrate how basis risk arises, why it is a key issue for market participants, and why certain choices were made in the course of the study.

The Authority defines basis risk (LPR) "the risk associated with unpredictable variations in the difference between spot prices for electricity at two nodes"[3] and states that "hedge contracts may not provide effective protection against this type of [basis] risk."[4]

It could be argued that a completely unhedged party is exposed to basis risk at their physical node[5] if a large INPD occurs (perhaps as a result of SWE) across an island, resulting in a very high or low[6] price at their physical node. However, such a party has presumably made a decision in the past not to be hedged at all, and so at least in terms of outcome, the risk of SWE is indistinguishable from the risk that prices will be high at all nodes in an island[7]. For this reason, basis risk is only relevant in the context of hedging strategy and, in particular, in the consideration of how the effectiveness of a hedging strategy is influenced by basis risk.

There are five generic hedging situations which include some element of WIBR:

1. generator sells at spot prices, and hedges at a distant node within the island;

[3] *Summary of Locational Price Risk Proposal, Explanatory Paper, Interim Report*, 19 April 2011.
[4] See http://www.ea.govt.nz/our-work/programmes/priority-projects/locational-hedges/
[5] The node at which they buy or sell electricity at spot prices.
[6] High price is of concern to a retailer or large consumer, low price is of concern to a generator.
[7] A SWE could produce much higher prices than a dry year, for example, but a dry year tends to last much longer, so the total impact of both could be similar.

2. retailer purchases spot prices to supply customers at a physical node, and hedges at a distant node within the island;
3. gentailer[8] sells at spot price at one physical node and purchases at spot price at a distant physical node within the island, and sells to customers at that node;
4. large consumer purchases at spot prices at a physical node, and hedges at a distant node with the island;
5. financial intermediary (such as a bank) buys or sells OTC hedges in the market at one node, creating an exposure to spot prices, and hedges the spot price risk at a distant node within the island (e.g. at one of the nodes at which futures contracts are available).

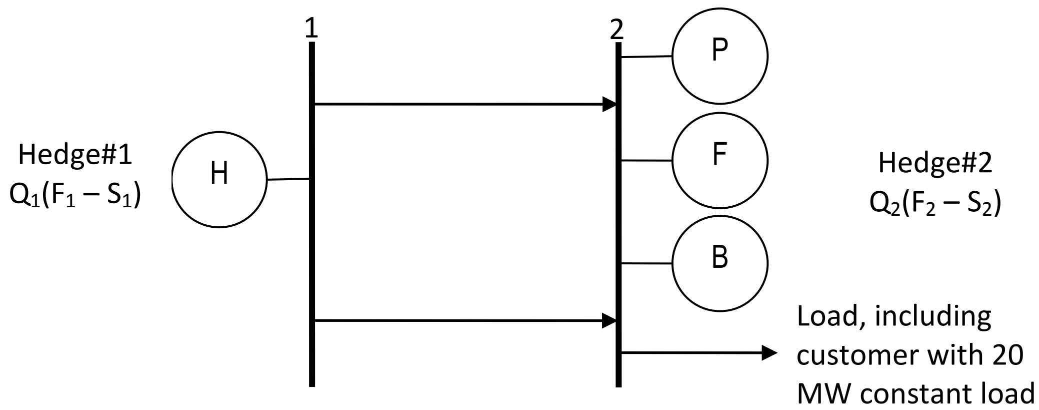
These five generic situations are described more fully in our report on the application of inter-island FTRs (refer to *Application of FTRs to Hedging Strategy, Part 1: Summary Report*, Energy Link, September 2010.[9]).

3.1 How Basis Risk is Created

In the following discussion, we refer to the simple grid and electricity market shown in Figure 2 and use this to illustrate how basis risk arises when parties attempt to hedge their spot price exposure with one or more hedges at one or more nodes located at a distance across the grid, but within the same island. The nodes are indicated in the figure by thick vertical lines.

Figure 2 shows a generator H at node 1, supplying energy across two identical transmission lines to node 2 at which there is a large customer bases and three more generators: generator B is a base-load generator, generator F is a firming[10] generator and generator P is a peaking generator. Under normal circumstances, all load is supplied by generators B and H.

Figure 2 – Simple Grid and Electricity Market



Losses are incurred on the transmission lines, and as a result the price at node 2 is higher than the price at node 1. In simple terms, and as long as the two lines run below

---

[8] Generator-retailer.

[9] http://www.ea.govt.nz/document/11910/download/our-work/programmes/priority-projects/locational-hedges/ftr-development/

[10] A firming generator fills the gap between base-load and peaking capacity.

their rated capacity, the price S_2 at node 2 is equal to the price S_1 at node 1 multiplied by one plus the marginal losses on the lines:

$$S_2 = S_1 \times (1 + 2Rp)$$

where R is the combined resistance[11] of the two lines and p is the total power flowing in the lines[12]. The price difference across the lines is proportional to the power flowing between nodes 1 and 2. The location factor of node 2 relative to node 1 can be defined as:

$$\ell_{21} = \frac{S_2}{S_1} = 1 + 2Rp$$

and the location factor of node 1 relative to node 2 is:

$$\ell_{12} = \frac{S_1}{S_2} = \frac{1}{1 + 2Rp}$$

However, when the two lines reach their rated capacity, then the simple relationships above no longer hold, and the price difference across the lines is determined not by marginal losses, but by the difference in the prices of offers that are dispatched at the two nodes. For example, if one of the lines has an outage and H's output is reduced, then B may not be able to supply all of the load at node 2, resulting in the dispatch of F and possibly also P. In this case the price at node 1 would be the offer price of H and the price at node 2 would be the offer price of node F or, if P was also dispatched, the offer price of P.

Let us suppose that a customer at node 2 has a constant load of 20 MW and they hedge 100% of the associated spot price risk hedge with generator H. There are two ways this can achieved:

- Hedge#1: hedge for a little more than 20 MW at node 1; or
- Hedge#2: hedge for 20 MW at node 2.

We will assume that the hedges are industry-standard contracts for differences (CFD).

Hedge#2

Hedge#2 is the natural choice, other things being equal, for the customer at node 2 simply because it eliminates basis risk. The total cost for the consumer with Hedge#2 in a trading period is given by

$$Cost = q_2 S_2 + Q_2 (F_2 - S_2)$$

where q_2 is the customer's load (20 MW in this example), Q_2 is the hedge quantity for Hedge#2 and F_2 is the strike (hedge) price for Hedge#2.

[11] Expressed as the per unit resistance divided by 100.
[12] This formula does not hold true for lines in a loop in a grid.

This can be rearranged to give

$$Cost = Q_2F_2 + q_2S_2 - Q_2S_2$$

The second formula shows that when 100% hedged so that $Q_2 = q_2$, then the residual risk associated with Hedge#2 is nil[13]. However the equivalent revenue formula for generator H (ignoring costs of generation) is given by

$$Revenue = q_1S_1 + Q_2(F_2 - S_2)$$

which rearranges to give

$$Revenue = Q_2F_2 + q_1S_1 - Q_2S_2$$

If we focus only on that amount of H's generation that matches Hedge#2 (which is less than the total output of generator H in this example) then we can see that even when 100% hedged, generator H retains residual exposure to the difference between S_1 and S_2.

Hedge#1
Under Hedge#1 the customer has trading period cost of

$$Cost = q_2S_2 + Q_1(F_1 - S_1)$$

which rearranges to give

$$Cost = Q_1F_1 + q_2S_2 - Q_1S_1$$

Under this hedge the customer now has residual exposure to the difference between S_1 and S_2 which is to say that there hedging strategy has basis risk, in addition to residual volume risk.

To ensure the customer has 100% hedge cover with Hedge#1, it is essential to adjust the hedge quantity by the location factor ℓ_{21} defined above[14]. In other words, under Hedge#1 the customer would select a hedge quantity equal to the expected value of ℓ_{21} over the term of the hedge[15]. This adjustment can be thought of as a way of 'amplifying' the payouts on the hedge at node 1 to match the larger movements in the spot price at node 2 where the customer has the spot exposure.

Under Hedge#1, of course, generator H does not have any exposure to the difference between S_1 and S_2, but the hedge quantity is not for 20 MW, but for 20 MW multiplied by the expected location factor ℓ_{21}.

---

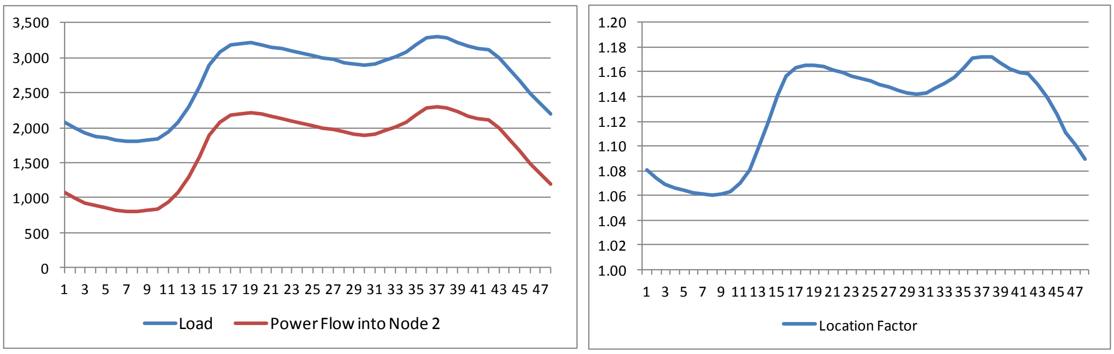
13 In practice, no hedge ever consistently achieves 100% hedge, so there is always some residual "volume risk" arising from an exposure to the spot price at the margin.
14 This is easy to show with a little bit of algebra.
15 A real hedge might be made up of a number of time zones, each with their own price and quantity, so the location factor adjustment would be made separately for each zone.

The above is perhaps better illustrated by a worked example, which covers one day, and in which the total demand at node 2 varies between 1,800 MW and 3,300 MW. The offers of generator H vary between $40/MWh and $130/MWh, generator B offers 1,000 MW at $0.01/MWh, generator F offers 500 MW at $300/MWh, and generator P offers 200 MW at $5,000/MWh.

The capacity of the lines connecting the two nodes is 1,750 MW each, so generators B and H can normally supply all the load at node 2 between them.

Figure 3 shows the total load at node 1, the power flowing into node 2 and the location factor $\boldsymbol{\ell}_{21}$, the latter varying from 1.06 to 1.17 over the day (which is deliberately set up to be quite a large diurnal variation).

Figure 3 – Load, Power Flow and Location Factor



An observation at this point is that the correlation between the prices at the two nodes is almost exactly equal to one, the importance of which will become apparent in section 4 when we define the cluster sets.

With the appropriate location factor adjustment on Hedge#1, Table 1 shows the two hedge quantities and strike prices, along with the load and generation actually hedged and the average price expected by the customer at node 2 (taking into account both the spot purchase and hedge settlements). The table shows, that with the location factor adjustment correctly applied in Hedge#1, and assuming the factor turns out to be equal to the expected actual location factor, the two hedging strategies achieve the same outcome.

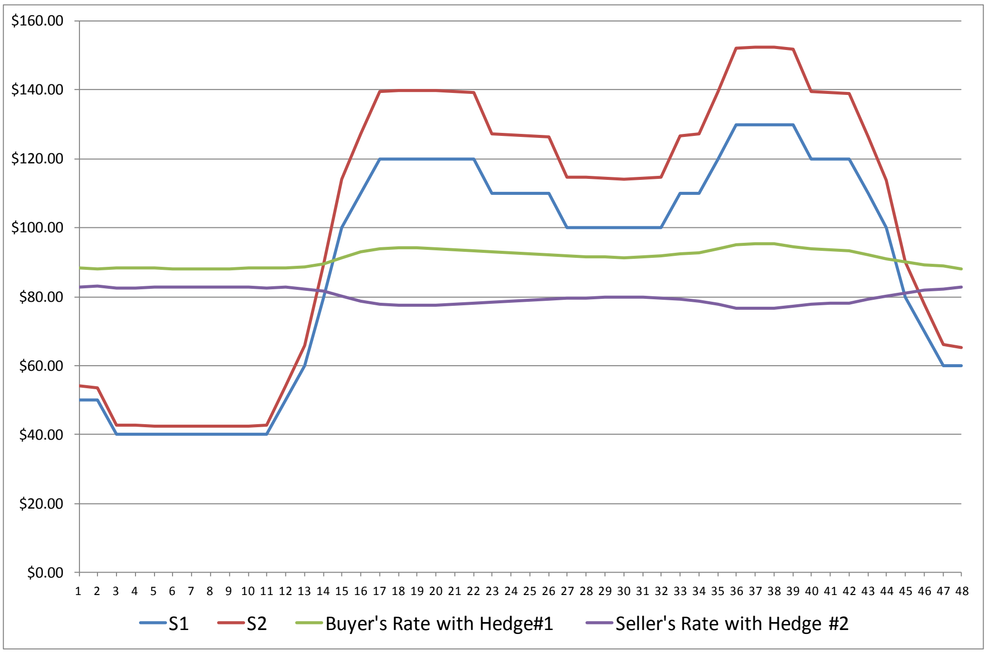
Table 1 – Hedge Parameters

Hedge	Node	Hedge Quantity (MW)	Strike Price ($/MWh)	Generation Hedged	Load Hedged	Price Expected at Node 2
Hedge#1	1	22.84	80.00	22.84	20.00	91.34
Hedge#2	2	20.00	91.34	22.84	20.00	91.34

The choice of location factor (1.142) for this example was simple: it is just the average of S_2 for the day by the average of S_1 for the day. In a real hedging situation the location factor adjustment is always made using an expected (forecast) location factor, which should allow for changes in power flows that are expected to occur with changes in demand, new generation build and transmission upgrades.

Nevertheless, the point here is that by making the appropriate location factor adjustment for Hedge#1, and as long as the actual location factors turn out to be approximately equal to the expected (forecast) location factors, a perfectly adequate hedging strategy can be formulated where the physical and hedge nodes are not the same (but within the same island). Figure 4 shows that the average buy and sell rates achieved by the customer and generator H, respectively, vary much less than the spot prices, despite large changes in price over the day and significant movements in location factors.

Figure 4 – Spot Prices and Average Rates



Now we look at what happens when larger price separations occur within an island, which can happen in one of two ways:

1. the prices at all nodes increase, as they do in a dry year for example, in which case the INPDs increase in proportion; or
2. a transmission constraint binds[16], often resulting in SWE.

In the first case, we can test the impact of this using our simple model, by increasing the offer prices of generators H and F by a factor of five, for example, which gives a range of prices at node 2 from $212/MWh to $514/MWh, but with the average achieved rate including hedges remaining at $91.34/MWh.

In a trading period when the location factor is equal to 1.142, the average purchase rate under Hedge#1 is still $91.34/MWh. With the location factor at its maximum value of 1.172 the average rate is higher at $111/MWh, and at the minimum location factor of 1.061 the average rate is lower at $75.11/MWh. The higher spot prices in this example

[16] There could be a line that constrains, or it could be an equation constraint. An equation constraint limits the combined flow on two or more lines.

amplify the small deviations from the expected average rate. That said, the deviation from $91.34/MWh is not great (+22%, -18%) and as long as the location factor does not move significantly from forecast, then Hedge#1 still achieves the expected costs for the customer over the day. Given the large movement in location factor across the day in this example, the customer could improve their hedge by having a different strike price and hedge quantity for day and night, for example.

SWE and other transmission constraints are associated with the much larger INPDs that can and sometimes do occur, and consequently much greater movements of location factor away from expected values. The simple model does not allow us to capture a full SWE because it does not include a suitable loop, but it does allow one line to be taken out of service. For example, if one line is out of service when demand is at its maximum value of just under 3,300 MW, then generators F and P are dispatched and the price at node 2 reaches $5,000/MWh and remains at $130/MWh at node 1. The resulting location factor ℓ_{21} is 38.5 and the additional cost to the customer in this one trading period is just under $49,000 which is equivalent to an additional 1.1 days of electricity supply at the hedged rate of $91.34/MWh.

Even with both lines in, if generator B has an outage then generator F is dispatched in this period, resulting in a price of $300/MWh at node 2 in this example, or an additional cost of just under $1,500.

Movements in location factor of this magnitude are difficult to predict, both in terms of their frequency and magnitude, and are therefore difficult to manage using a simple location factor adjustment.

In the above examples, outages created basis risk, but by implication, the choice of offers is also potentially a key WIBR driver: for example, if generator P chose to offer at $10,000/MWh then the price separation with one line out would be correspondingly greater than in our example.

Our examples also used a trading period when demand was at its highest, at which time the loading on the grid is correspondingly higher and the loss of transmission or generation capacity is therefore more likely to create basis risk. Thus we can see that transmission and generation outages, offer prices and demand are all implicated and may turn out to be significant WIBR drivers.

4 Cluster Analysis

The simple model used in section 3.1 above features only two nodes, whereas the real grid has around 250 GIPs and GXPs at which spot prices are published. For most nodes on the grid, there is a set of nodes whose prices correlate so well with that of node A that basis risk is immaterial if a spot price exposure at A is hedged at any node in the set. We have called these sets of closely correlated nodes clusters (or WIBR clusters)[17].

[17] The use of the term 'region' can be confusing because it conjures up images of nodes in a geographical region, whereas the nodes in a cluster are clustered in the sense of being electrically well connected, which could see a cluster spread out across large tracts of an island along one or more transmission long lines.

Once WIBR clusters are known, then anyone should be able to hedge with confidence at any node in the same cluster as their spot price exposure. This also simplifies the issue of assessing and managing WIBR because WIBR can be isolated between a relatively small number of clusters (instead of between a much larger number of nodes). For phase two of our study, the smaller number of clusters greatly simplified the analysis of WIBR drivers.

The first phase of the study first required that a test be developed for membership of any given node within a WIBR cluster. In principle, a cluster is made of nodes whose prices correlate above a given threshold with the prices at all other nodes in the cluster. However, this relatively simple definition of a cluster is complicated by two factors: firstly, one must define what is meant by 'correlation'; secondly, one must define the degree of correlation that is considered material in the context of hedging.

4.1 Correlation Defined

A key consideration in defining correlation was to use a correlation metric that is both well known, well understood and easy to apply, the obvious candidate being the correlation used in the CORREL function in the Excel spreadsheet program. CORREL uses the Pearson's product-moment coefficient in its sample form[18], which estimates the correlation coefficient of an entire population.

Pearson's coefficient is a widely used statistic that calculates the degree of correlation between two variables, and is equal to one when the two variables are linearly related. This explains why the correlation (calculated in Excel) between S_1 and S_2 in section equals one when the two transmission lines are unconstrained: the two prices are related by a simple formula which is linear in the power flowing between the nodes. While this high level of correlation persists, even when prices rise in a dry year, then the hedging strategy produces a stable average price across each day.

But when large price separations occur then the correlation falls away rapidly. For example, with node 2 at $300/MWh and node 1 at $130/MWh for one trading period in the day, the correlation is 0.90. But with node 2 at $5,000/MWh and node 1 at $130/MWh for one trading period in the day, the correlation falls to 0.23 for the day.

This all suggests that the Pearson's coefficient may perform well as a test for nodes that should be within the same cluster, for example we might define clusters by the rule that "all nodes should correlate to at least 0.9 with each and every node in the cluster at all times".

However, most nodes in the grid are connected to more than one other node, and the simple price relationships given in section 3.1 do not hold in general when there are loops in a grid. In the absence of binding transmission constraints, the price at any given node is a linear function of the flow in a number of lines (which in turn is influenced by generation and demand), so the price difference between two distant nodes is no longer a linear function of one variable. A potentially complicating factor is that SPD models losses in each line using three linear loss segments, resulting in stepped price differences between nodes (the HVDC link has six loss segments).

[18] For a concise description see http://en.wikipedia.org/wiki/Pearson_product-moment_correlation_coefficient.

While the issue of SPD's loss modelling is second order and likely to be immaterial, the issue of how abnormal power flows change correlations between prices at distant nodes is significant. However, a number of tests of the price differences between various pairs of nodes does suggest that the Pearson's coefficient (which is calculated in the CORREL function in Excel) is a reasonably good measure of correlation for defining clusters.

4.2 Materiality and the Cluster Threshold

Ideally, the criterion for a node to be a member of a cluster would be that it correlates with all other nodes in the cluster to at least some minimum value, so that any party subsequently hedging within the cluster could be confident that basis risk would be immaterial. However, the single value could be too high for some hedgers (more clusters, limiting nodes for hedging within a cluster) or too small (exposing the hedger to more risk than desired even when hedging within a cluster).

Each hedger has their own propensity to take risk and their own particular set of circumstances, so what amounts to material basis risk for one may be immaterial for another. For example, some larger consumers are 100% exposed to the spot price, while others prefer to be as close to 100% hedged as possible.

For the purposes of the study, and as a guide to readers, one month was used as the correlation period. For a node to be included in a cluster, it must correlate with all other nodes in the cluster to at least the chosen correlation test value, month by month from Jan-07 to Dec-11. A month was chosen for the correlation test period (as opposed to, for example, one year or five years) because one month is the length of the billing cycle.

Table 2 shows the impact of one inter-nodal price spike (between the physical and hedge nodes) out of the 1,440 trading periods in a month of 30 days, with each day having a typical daily profile[19], for a load of 100 MW and a hedge with a strike price of $74/MWh. The rightmost column shows the percentage increase in the monthly cost resulting from the spike. This load has a monthly electricity bill of just over $5.3 million.

Table 2 – Impact of One Price Spike on Correlation for a 100 MW Load

Correlation	Spike	Monthly Increase	Monthly Increase
0.99	$100/MWh	$5,000	0.09%
0.95	$230/MWh	$11,500	0.22%
0.90	$338/MWh	$16,900	0.32%
0.85	$431/MWh	$21,550	0.40%
0.80	$520/MWh	$26,000	0.49%
0.75	$610/MWh	$30,500	0.57%
0.70	$704/MWh	$35,200	0.66%
0.60	$914/MWh	$45,700	0.86%
0.50	$1,175/MWh	$58,750	1.10%
0.40	$1,535/MWh	$76,750	1.44%
0.30	$2,090/MWh	$104,500	1.96%

[19] Taken from the Haywards node.

Correlation	Spike	Monthly Increase	Monthly Increase
0.20	$3,110/MWh	$155,500	2.92%
0.10	$5,695/MWh	$284,750	5.34%

When transmission constraints occur, they can create large INPDs for as little as just one trading period, but they can also last for several hours in a day, and sometimes on more than one day in a month. For example, the table shows that a $100/MWh price spike occurring once in a month maintains a correlation of 0.99 while a spike about 12 times as large ($1,175/MW) reduces the correlation to 0.5. However, if the same $100/MWh spike occurred twelve times during the month, then the monthly increase in cost would be twelve times as large, or 1.1%, but the correlation would only fall to 0.89 (this is not shown in the table). Furthermore, if spikes can occur in one month then they can occur in another, which leads to the conclusion that the relationship between correlation and materiality is influenced by the number of large spikes in INPD as well as by the magnitude of each spike.

An increase in cost of 1.1% in one month in our example above is 0.09% over a year, which is a tiny percentage of the annual electricity bill. However, the absolute increase of almost $60,000 is still a substantial sum. So as well as considering the relative increase in costs due to spikes in INPDs, risk averse hedgers may also consider the absolute amounts potentially at risk.

In terms of this study, two conclusions were reached early on:

- there can be no choice clusters that will satisfy all potential hedgers for use in their respective hedging applications; and
- the possibility of multiple spikes means that risk averse hedgers are likely to prefer hedge nodes which are highly correlated with their physical nodes.

Another key consideration when it comes to forming clusters is whether nodes that are major centres of load and generation are included within one or more clusters: to be consistent with the work on FTRs to date, we refer to such nodes as hubs. If the prices at two hubs correlate to a high degree anyway, then they will be in the same cluster unless a very high threshold correlation value is chosen. But if they correlate to a lesser degree then it may be advantageous to use a threshold correlation value which allows the two hubs to be in different clusters for the purpose of assessing and designing instruments which hedge WIBR.

4.3 Preparation of Data

It was originally planned to average prices at substations by half hour in order to reduce the total number of nodes to 108 in the NI and 67 in the SI, where we defined a substation as a collection of adjacent nodes which all start with the same three letters, for example the Stratford substation, SFD, includes SFD2201, SFD1101 and SFD0331. However, this approach soon fell over due to constraints that occur within substations, Roxburgh being a prime example: the transformer ROX_T10 links the 110 kV and 220 kV buses but has capacity of only 50 MW, which caused it to constrain during the study period and cause price separations between the two buses.

So early on in the study the decision was made to process the node data by voltage, to give “voltage nodes”. In most cases this meant processing nodal data, but in a few cases prices were averaged across two nodes, for example at OTA there are two nodes at 220 kV, so the voltage node OTA220 is the average of OTA2201 and OTA2202. The full list of voltage nodes is included in Appendix 1.

There are also a significant number of disconnected node situations in the study period, where the price at the disconnected node is set to zero. For example, the substation HWB consists of HWB2201, HWB0331, HWB0332 and HWB1101, and the first three of these nodes have prices published in the market. The prices at these three nodes normally correlate to a very high degree, but on 8-Apr-11 from trading period 14 to 35 inclusive, HWB2201 was disconnected from the grid and, as a result, had a price of zero. Should HWB2201 be bumped out of the same cluster as the other HWB nodes just because of this one day? Do disconnected nodes even constitute basis risk? Well, the short answer is yes, but in reality a disconnected node constitutes much less risk than the potential for an INPD of many hundreds or even thousands of dollars, so it was decided to ignore disconnected nodes. Thus clusters can include nodes that were disconnected during the study period.

Similarly, there were a handful of days when all prices were zero in an island, and so these days were also eliminated along with disconnected nodes.

4.4 Cluster Analysis

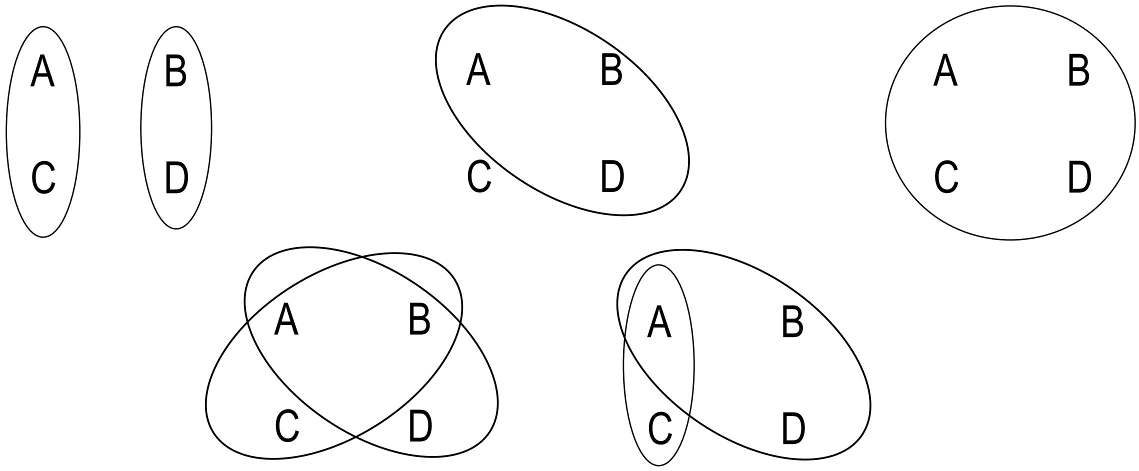
In an island with N nodes, there are N(N – 1)/2 node pairs, and there are 146 voltage nodes in the NI and 87 voltage nodes in the SI. For each of the 60 months in the study period, the correlation between the half hourly prices at each pair of nodes was calculated, resulting in 10,585 correlation values in the NI and 3,741 in the SI for each month.

The monthly correlation values were then tested for each node against all other nodes in a potential cluster: if the minimum correlation over the 60 months is greater than or equal to the chosen threshold correlation value, then the pair of nodes is in the same cluster[20]. This process is repeated until all nodes are allocated to one or more clusters, bearing in mind that it is possible for a node to be in a cluster all on its own.

Intuitively, one might expect clusters to always be distinct from each other, which is to say that any particular node can only be in one cluster at a time. But in reality clusters can and do overlap. For example, four nodes A, B, C and D are shown in Figure 5 in five different configurations of clusters (a cluster is indicated by an ellipse), all of which are possible: the lower two cluster configurations feature overlapping clusters.

[20] So it takes only one month to be below the threshold correlation value for the nodes to be in different clusters.

Figure 5 – Clustering of Four Nodes



The method developed to form clusters first required that all nodes in an island be sorted alphabetically: A, B, C, D in the figure above. The second node B in the list is tested for membership of the same cluster as the first node A, then the third node C is tested for membership of the same node as the first node A, and so on to the end of the list. This gives a cluster which contains the first node in the alphabetic list of all nodes in the island.

On the second pass the list is modified by placing the second node at the top of the list so the order is B, A, C, D and the cluster testing process is repeated as above. Then the third node is placed at the top of the sorted list to give C, A, B, D and the cluster testing processed repeated, and so on.

This continues until each and every node has been tested at the top of the sorted list of nodes. The process produces many duplicate clusters, and clusters which are subsets of larger clusters, leaving the set of unique clusters for the island and the threshold correlation value used in the clustering process.

4.5 Cluster Results

Arguably, some hedgers may be content with quite low correlation thresholds, but in this study only values down to 0.7 were used to produce clusters. Clusters were in fact produced for correlation thresholds of 0.7, 0.8, 0.9, 0.95 and 0.99.

The full list of clusters is included in Appendix 2, and the following maps show the larger clusters for monthly correlation thresholds of 0.8, 0.9 and 0.99.

A word of warning: the maps are included for illustrative purposes only, do not show the detail of each node in a cluster, and therefore can be misleading as a guide to hedging. Readers should refer to the cluster lists in the Appendix.

Figure 6 – Larger Clusters Obtained with Correlation = 0.8

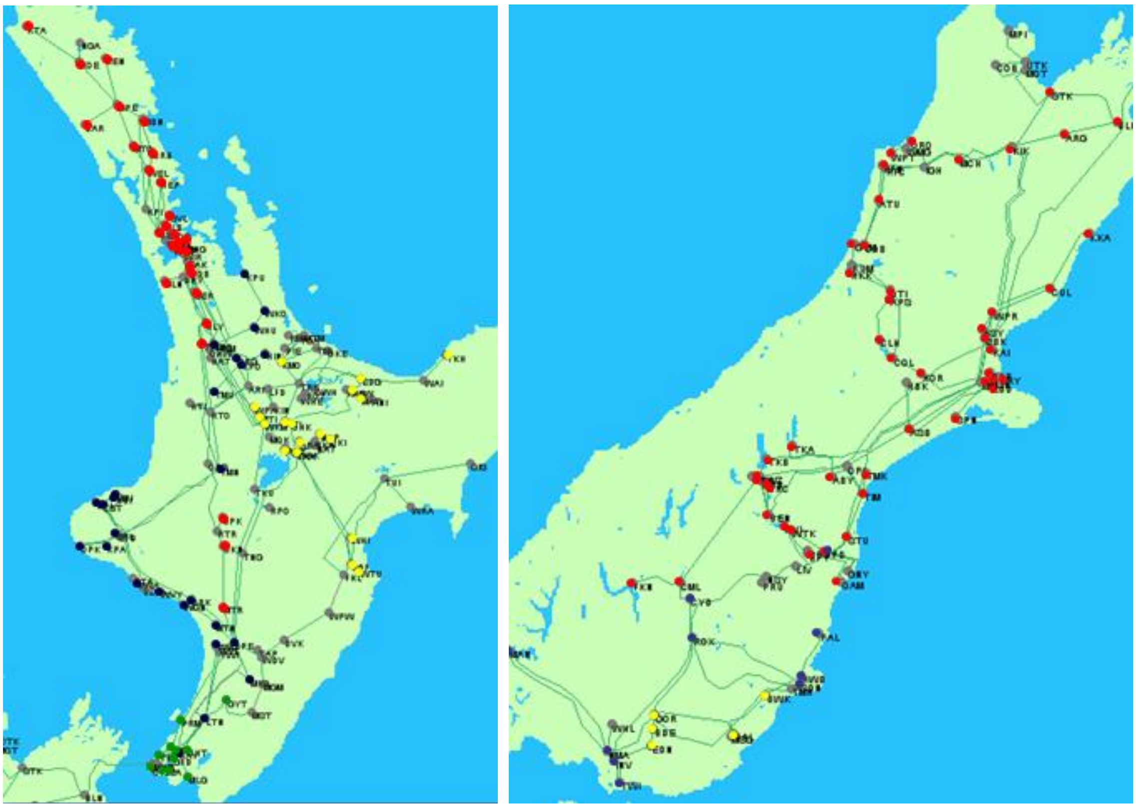


Figure 7 – Larger Clusters Obtained with Correlation = 0.9

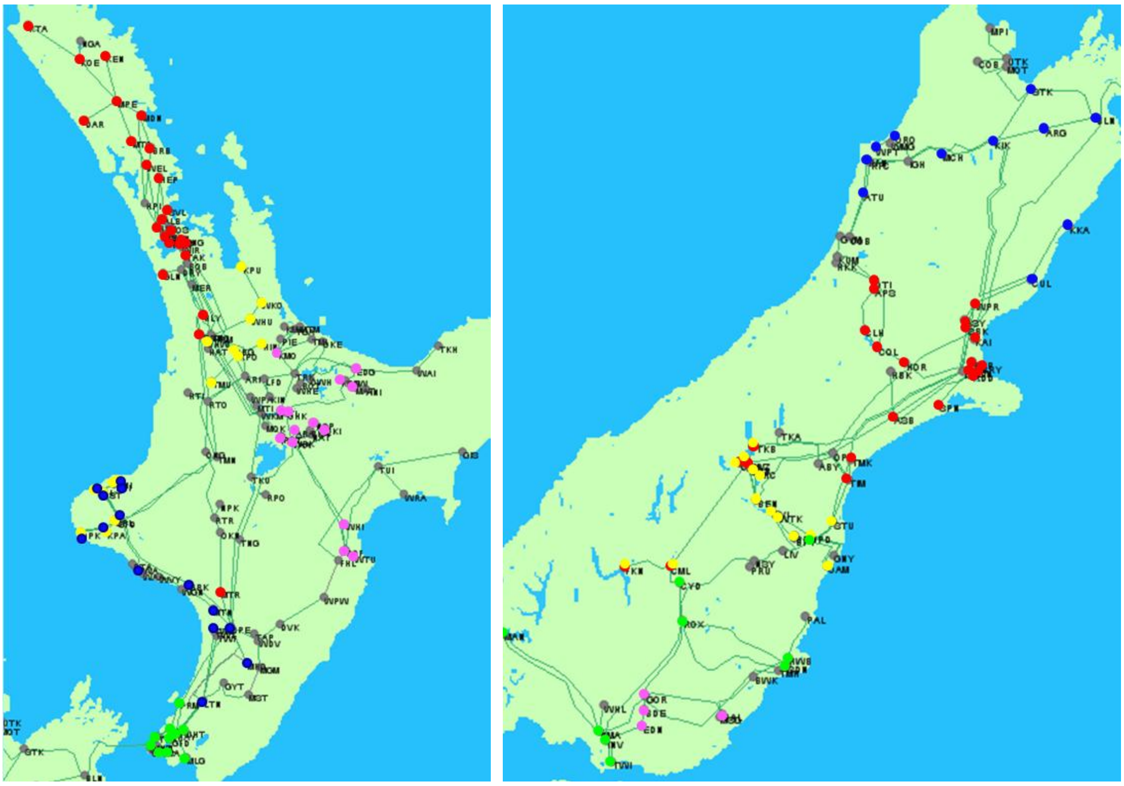
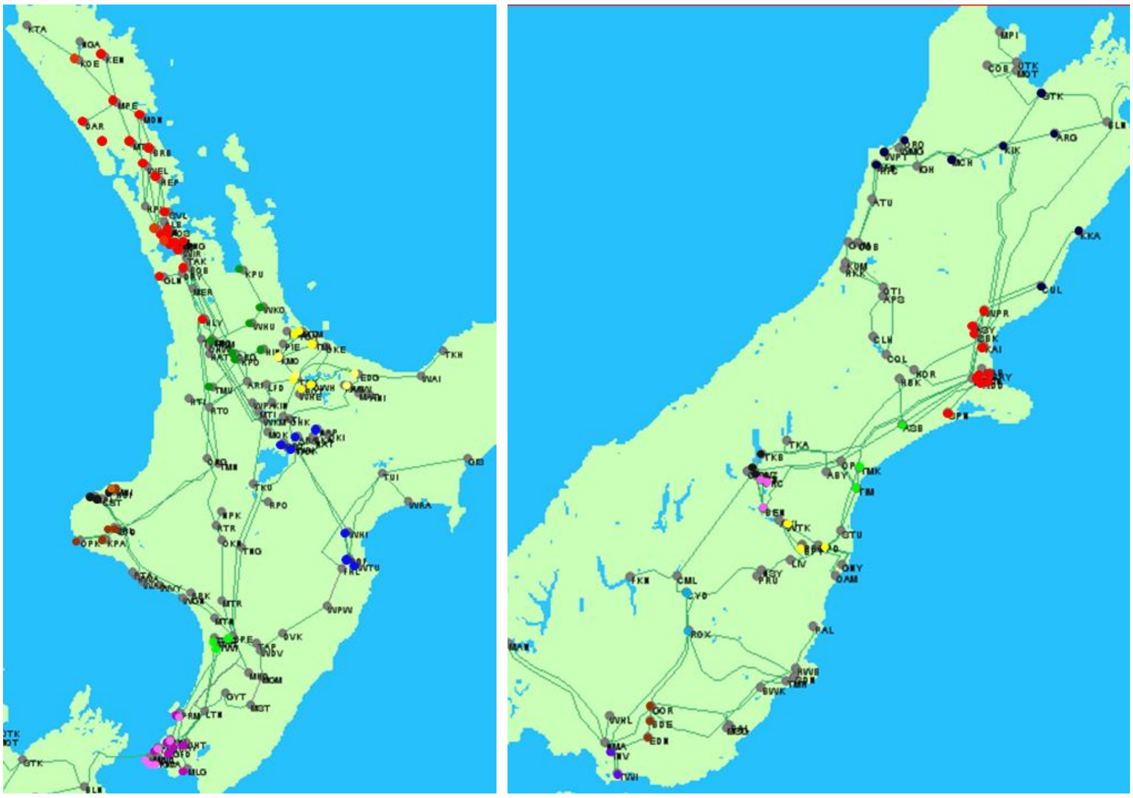


Figure 8 – Larger Clusters Obtained with Correlation = 0.99



If a hedger can define the materiality of basis risk in terms of the cluster correlation threshold, then they need only refer to the relevant clusters above and can then be confident that hedging at nodes within the same cluster as their physical node(s) will result in a hedging strategy that has immaterial basis risk. A word of caution, however, must be added to this statement: this assumes that WIBR remains the same as it was over the study period. In reality, it could change so that clusters either expand or contract over time.

For the purposes of phase 2 of the study, 0.9 was used as the correlation threshold, not because of any consideration of materiality, but in order to get enough clusters to give meaningful results without adding an excessive amount of data. This choice also allowed most key hubs to be in clusters of their own except where the correlation between major nodes is very high:

- WKM220, SFD220, HAY220, and BPE220 are in separate clusters;
- OTA110, OTA220, HLY220 and HLY033 are in one cluster together;
- HLY033 is also in another cluster without other major nodes;
- BEN, ISL and STK are in separate clusters;
- ROX110 is in a cluster without other major nodes;
- TWI220, CYD220, CYD033 and ROX220 are in a cluster together.

5 Analysis of WIBR Drivers

Having completed the cluster analysis, phase 2 of the study first required calculating the average price at each cluster (defined using 0.9 as the correlation threshold) for all half hours in the study period from 1-Jan-07 to 31-Dec-11.

All references to prices from here on, unless otherwise stated, are references to cluster prices calculated with a threshold correlation value of 0.9, and not to prices at individual nodes or voltage nodes.

5.1 WIBR Drivers

A measure of the degree of WIBR was calculated for each half hour. Three methods of calculation were tested:

1. the standard deviation of the cluster prices in the half hour;
2. the absolute spread of prices, equal to the maximum cluster price minus the minimum cluster price divided by a constant; and
3. the relative spread of prices, equal to the maximum cluster price minus the minimum cluster price, divided by the average cluster price in the half hour.

The standard deviation tends to hide more extreme events, for example when the price at only one cluster is much different to the others. The absolute spread of prices is tempting to use because it highlights extreme events such as 26-Mar-11 when provisional prices[21] reached $20,000/MWh. However, this measure tends to over-emphasise events with extreme prices, which causes the results of the analysis to be more difficult to interpret. The relative spread of prices was therefore used in phase 2.

The spread of prices was compared against a variety of candidates for being key drivers of basis risk within each island, including the following parameters recommended in our report in late 2011:

- total island demand;
- total capacity of circuits in outage;
- total loading on the grid in the island;
- total MW of generation having an outage within the island;
- total power transfer on the HVDC link, with northward flow defined as positive;
- change in offers during the half hour.

The capacity of equation constraints[22] was also considered as a candidate WIBR driver, but it was rejected on two counts. Firstly, it is actually quite difficult to track the capacity of equation constraints because they are not all permanent, so for example many only apply during outages. Secondly, the impact of equation constraints to limit power flows during outages should already be captured by analysing transmission and generation outages as WIBR drivers.

A significant effort was required to process the appropriate data for each half hour of the study period. Total island demand was only available back to 1-Jun-07, which effectively limited the study period to Jun-07 to Dec-11, although this still covered a total of 80,398 trading periods.

[21] During the study and at the time of writing, 26-Mar-11 provisional prices were the only prices available. It is anticipated that in the near future final prices will be published that are closer to $3,000/MWh.

[22] The capacity is equal to the constant value on the right hand side of the equation constraint.

Longer records of half hourly demand are available, but are net of embedded generation. The dataset used is known as the "GXP Demand" available from the Market Data menu in WITS, and is the raw demand data used in SPD to calculate final prices. This dataset was published starting in Jun-07 and is gross demand except at nodes where wind farms inject, so this demand series represents the net demand that must be supplied by all other generation.

Circuit outages were extracted from data available from em6 through to 15-Jun-09, and then from the SPD daily case data files which are now available from the Authority. As these outages were processed, however, it quickly became apparent that the total capacity in outage is a poor measure of the impact of the outages on the market, primarily because many lines and transformers, in particular, have listed ratings far in excess of their actual loading.

For example, many transformers have ratings of 2,000 MW despite being loaded at just a few percent of this value[23]. The loss of a transformer rated at 2,000 MW does not represent the loss of 2,000 MW of capacity if the transformer normally never exceeds a loading of 100 MW, for example.

Outages and arc flows were processed to determine the likely impact on power flows, based on the running average of the daily maximum power flowing in the outaged line. For example, if a line had listed capacity of 500 MW but normally peaked at 100 MW, then the loss of capacity due to the outage was taken as the lower figure. Transformer outages were excluded from the outage calculations, as were HVDC outages, the latter effectively being included in the HVDC flows data.

We also calculated the total transmission outage due to "transmission groups" being in outage, where we defined a transmission group as the line or lines connecting two nodes. In a transmission group outage, two nodes entirely lose their direct connection, which means that a group outage potentially has a greater impact on dispatch and pricing outcomes than an outage which leaves two nodes directly connected. For example, two lines connect directly between Whakamaru and Otahuhu. If one line is out then it is counted as an outage but not a group outage. If both lines are out then the outage is counted as a group outage.

A similar issue was faced with the total loading on the grid which, in principle, is equal to the total power flowing on the grid divided by the total listed capacity. With so many grid elements with capacity listed at 2,000 MW this approach would always indicate a lightly loaded grid. Instead, the arc flow on all lines was increased by 10%, 20% and 30% which sometimes resulted in one or more lines running over their capacity. The total overload in MW was then recorded against the relevant 10%, 20% or 30% flow increase. In effect, this approach provides a measure of how close, in physical terms[24], lines are running to their respective limits. For the purposes of phase 2, only the 30% overload figure was used, as it provided the clearest signal of grid loading.

[23] The reason for this is unknown (and was not checked), but may have to do with the transformers being highly unlikely to ever constrain.
[24] In market terms, generation is dispatched to ensure that lines are not overloaded.

Generation outages are more difficult to determine from actual generation data because it cannot be determined if the loss of generation is due to an outage or due to some other cause such as purely commercial factors, or simply just lack of demand.

The POCP web site has voluntarily disclosed data for the major thermal generators, so this data was used to estimate the total MW of generation outage in each half hour.

The change in offers in each half hour is the square root of the mean of the squares (RMS) of the change in offer quantity across the following price bands since the same half hour in the preceding week: up to $10/MWh, $10 to $30, $30 to $50, $50 to $100, $100 to $300, $300 to $500, $500 to $1,000, $1,000 to $3,000, $3,000 to $5,000 and over $5000.

In addition, we also looked at a range of other parameters that potentially play a causative role in creating basis risk:

- rates of change of the parameters described above;
- time of day, represented by trading period number;
- time of year;
- capacity margin, equal to the total quantity offered for the half hour less the total demand;
- maximum offer price;
- the total of the SIR and FIR price in the relevant island.

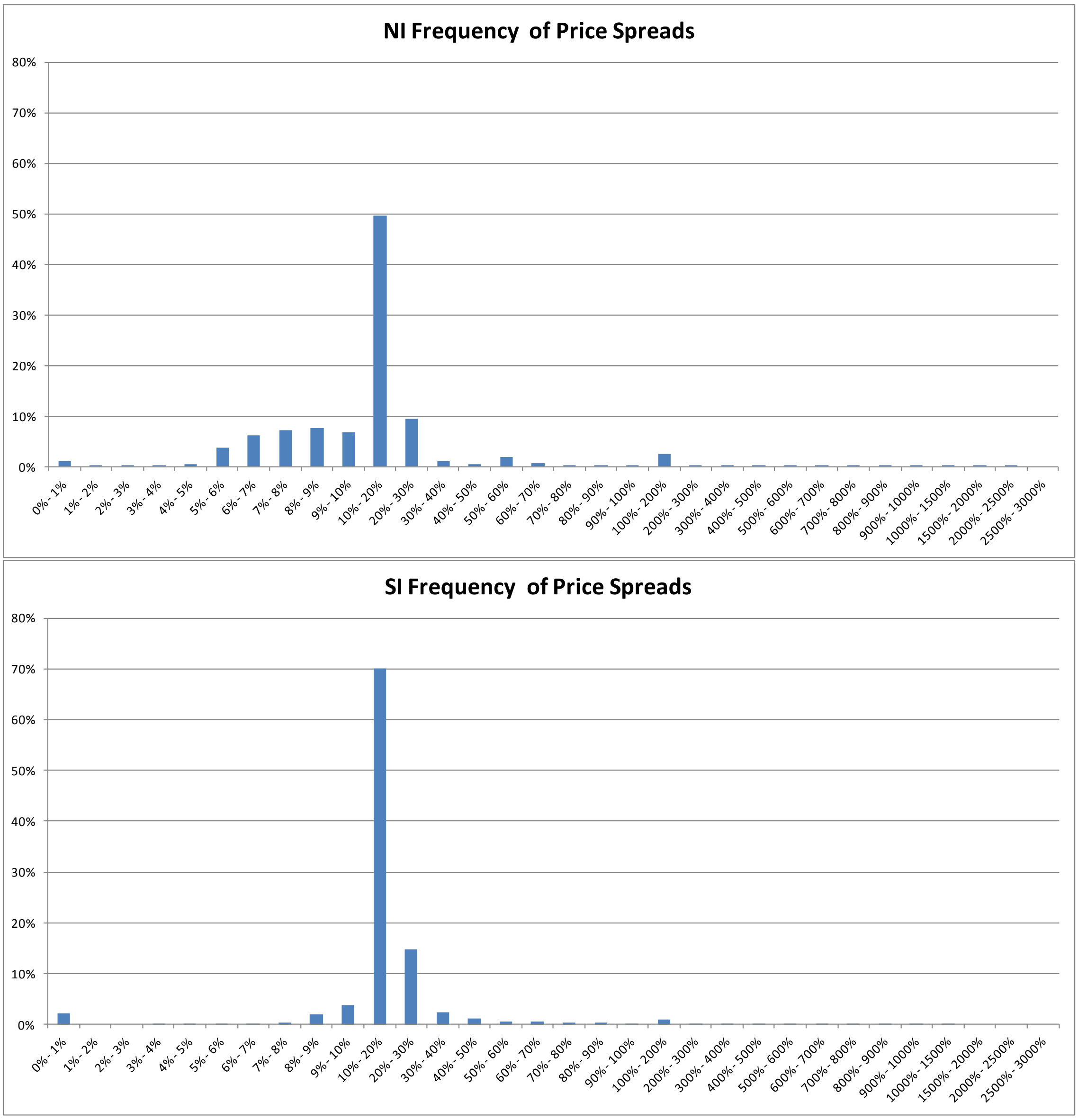
5.2 Results

This section contains a large number of charts with brief comments and observations, which are summarised in section 5.3.

The following two histograms show the frequency of price spreads[25] over all half hours of the study period, in bands from zero up to the highest spreads of just under 3,000%.

[25] Recall that the price spread is equal to the difference between the highest and lowest cluster price in the trading period, divided by the average cluster price in the period.

Figure 9 – Overall Frequency of Price Spreads



Price spreads in the NI appear in significant numbers from quite low values around 5% through to 30%, whereas in the SI they are more tightly grouped in the bands between 10% and 30%. Spreads well in excess of 100% are visible on the NI histogram but not so on the SI histogram, indicating the relative frequency of the extreme spreads is much higher in the NI.

If a simple correlation is performed between the parameters listed in section 5.1 and the spreads, over the study period, then the results are truly disappointing: the highest correlation between price spread and WIBR driver is only 0.09. However, using linear regression analysis produced a somewhat stronger result, indicating that circuit outages, generation outages and demand contributed about equally to creating the price spreads each half hour.

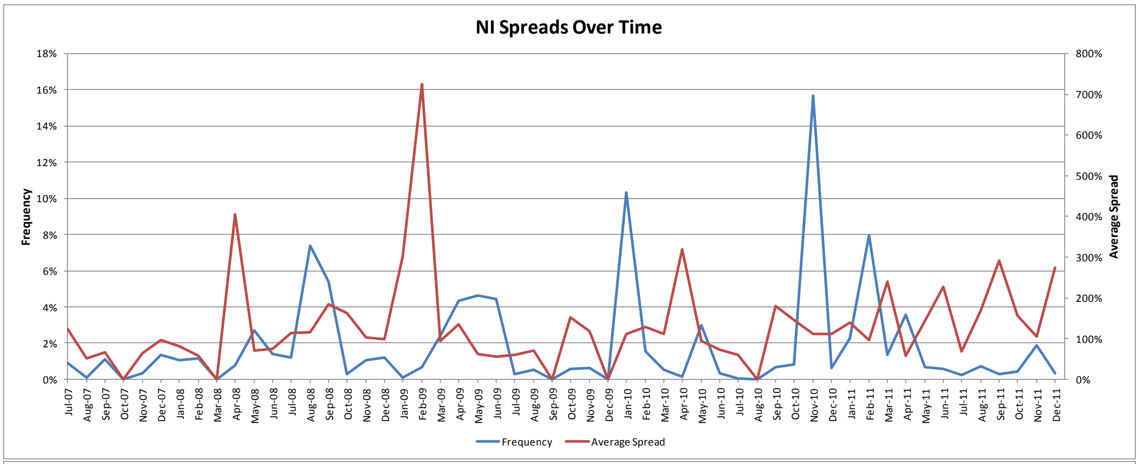
However, only a relatively small proportion of all trading periods have price spreads which are above the threshold to be considered as contributing to basis risk, so after initial testing, it was decided to focus on those half hours with the highest price spreads.

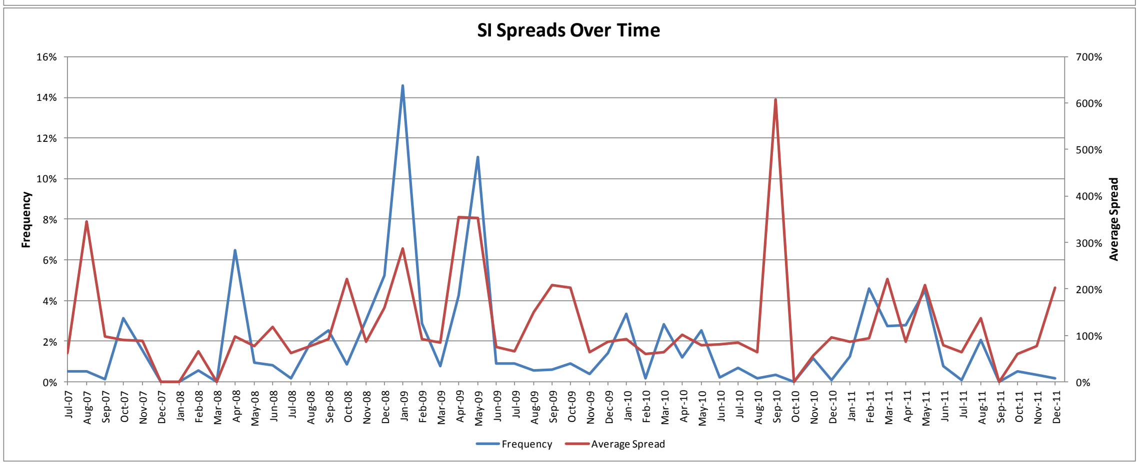
The average spread over the study period was 21.5% in the SI and 19.5% in the NI, so the threshold spread for the analysis was chosen to be 50%, which is high enough to indicate a spread significantly above average (and almost certainly associated with a binding constraint as opposed to losses), but low enough to give a large sample to work with. The sample size reduced from 80,399 periods down to 4,514 in the NI (5.6% of the total periods) and to 2,793 in the SI (3.5% of the total periods), which reinforces the initial observation that there are more extreme events in the NI than in the SI.

In the following, we refer to these two sets of events as 'high price spread events', and the graphs below typically show the frequency of these events and the average magnitude of the price spread.

Figure 10 shows how the frequency and magnitude of spreads has changed by month over the course of the study period. The charts show periods where high spreads have occurred more or less frequently, but the only trend that is evident late in the study period is that the average spread during the high spread events is increasing in the NI.

Figure 10 – Spreads over Time





How to Interpret the Following Graphs

Most of the following figures show graphs of high price spread events against a range of values of a WBIR driver. In each graph, the frequency is shown in two ways: in blue and in green.

The blue curve shows frequency relative to the total number of high price spread events, which is much less than the total number of periods in the study.

The green curve shows frequency relative to the total number of periods from 1-Jun-07 to 31-Dec-11. For example, in Figure 12 (which shows high spread events by HVDC flow) the blue curve on the NI graph shows that more high price spread events occur when the flow is 400 – 450 MW north than for any other value of flow: about 9% of all high price spread events occurred in this flow range.

But the green curve shows that flow of 650 – 700 MW north is the most likely range in which high price spread events might occur: about 27% of all periods in the study with flow in this range had events.

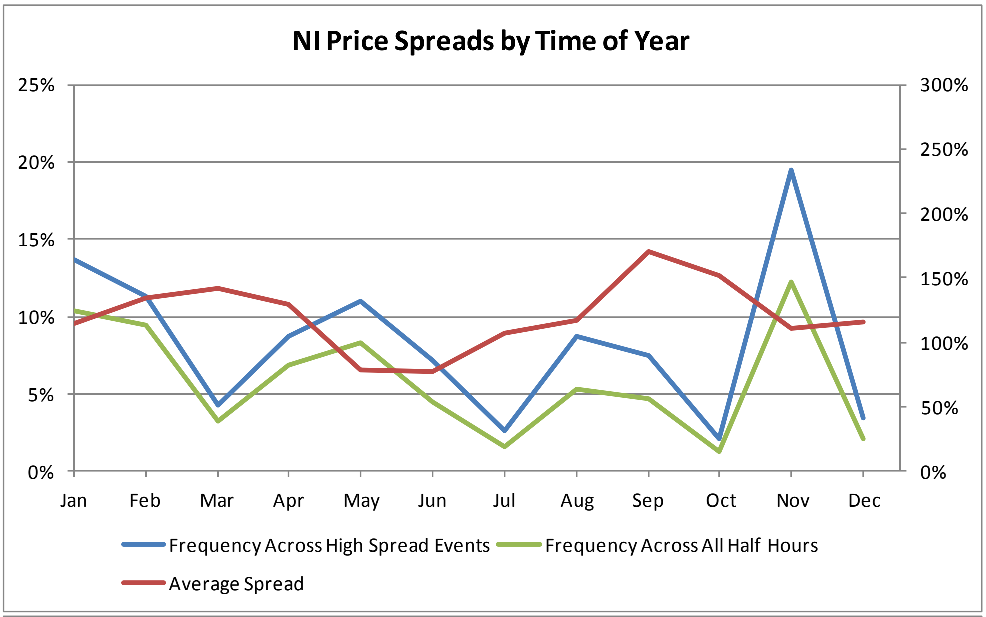
The blue curve is shown for reference, but it is the green curve that is the most important measure of the frequency of high price spread events.

The red curve shows the average magnitude of high price spread events, i.e. the average price spread (relative to the trading period average price) during these events.

On all graphs, the left hand vertical axis is for frequency (blue and green curves) and the right hand vertical axis is for the average magnitude of price spread (amongst high price spread events).

Figure 11 shows the high price spreads by month over the study period. The SI has two distinct peaks in January and in April-May which are due to the events of the 2008 dry year. The NI has peaks in January, May, August-September and November, arising from Jan-10, May-09, August-September 2008 and Nov-10, respectively.

Figure 11 – Spreads by Time of Year



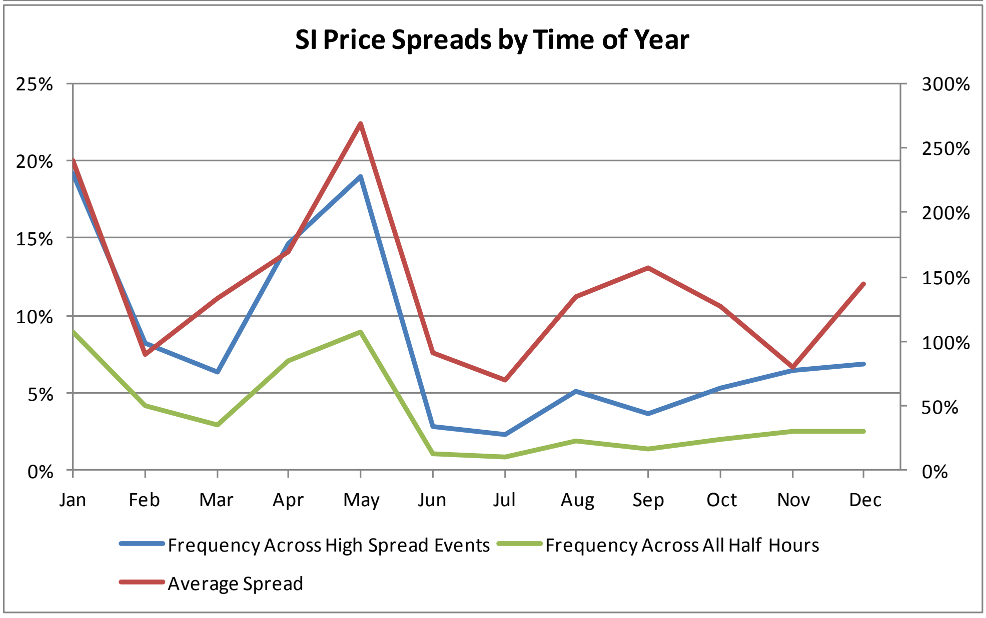


Figure 12 shows high price spread events against the flow on the HVDC link, where positive flow is in the northward direction. In the NI, high price spread events, when they occur, increase in frequency with northward flow to a peak in the 400 – 450 MW band. But taken over all periods they peak close to the extreme ends of the spectrum of flow, i.e. the highest flow values had a greater likelihood of a high price spread event. The average price spread is more or less constant across the range of flows.

In the SI, high price spread events are more common with high flows northward, and the average spread is also significantly higher with northward flows.

Figure 12 – Price Spreads by HVDC Flow

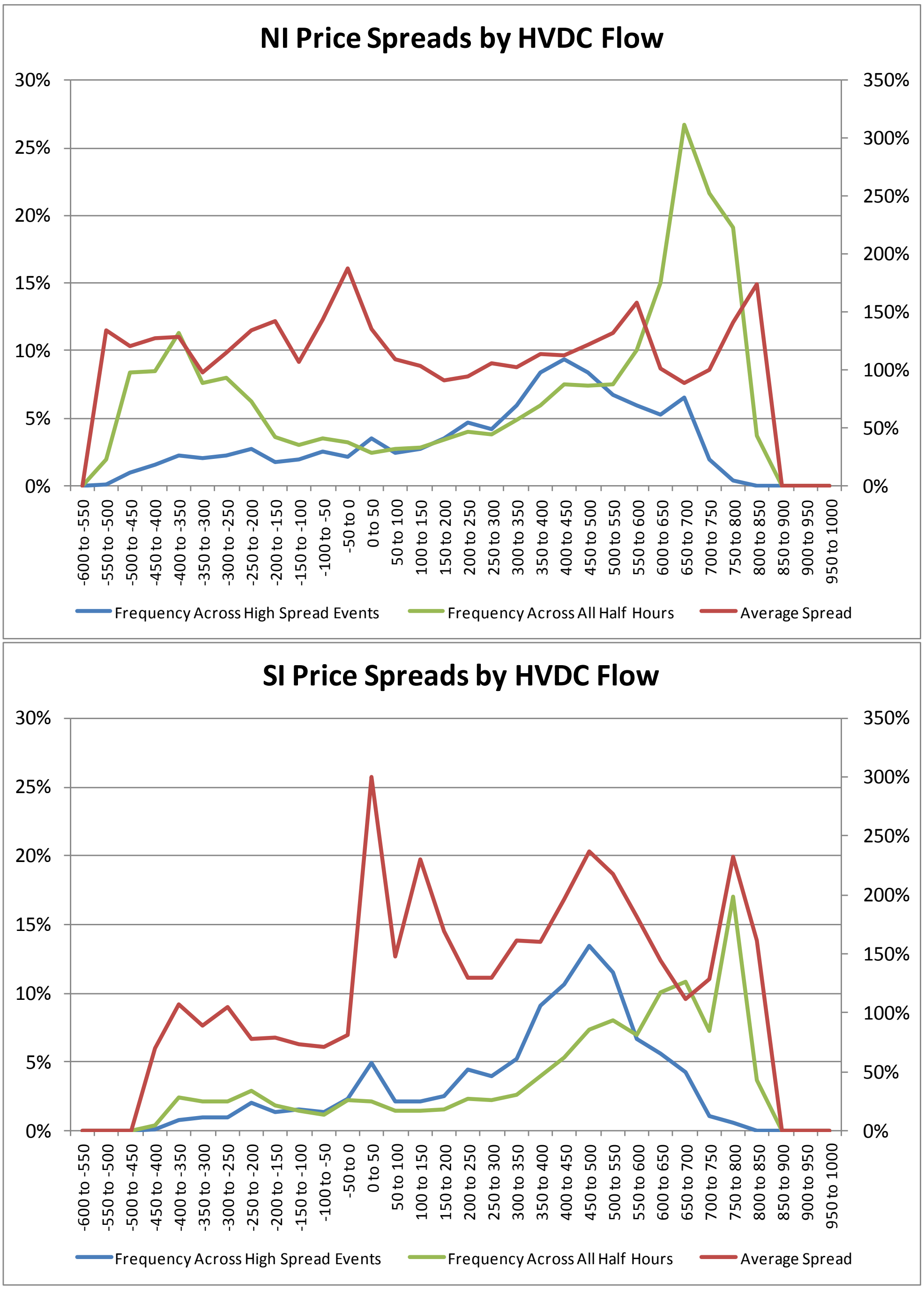


Figure 13 shows high price spread events against total demand for the relevant island. In the NI, amongst these events only, they are most likely to occur when demand is between 2,900 MW and 3,400 MW but taken over all periods they are much more likely to occur with demand in excess of 4,300 MW than with lower demand. The average spread, however, is more or less constant over the entire range of NI demand.

In the SI, these events are most likely to when demand is between 1,500 and 1,800 MW but with average spread peaking at lower levels of demand up to 1,200 MW.

Figure 13 – Price Spreads by Demand

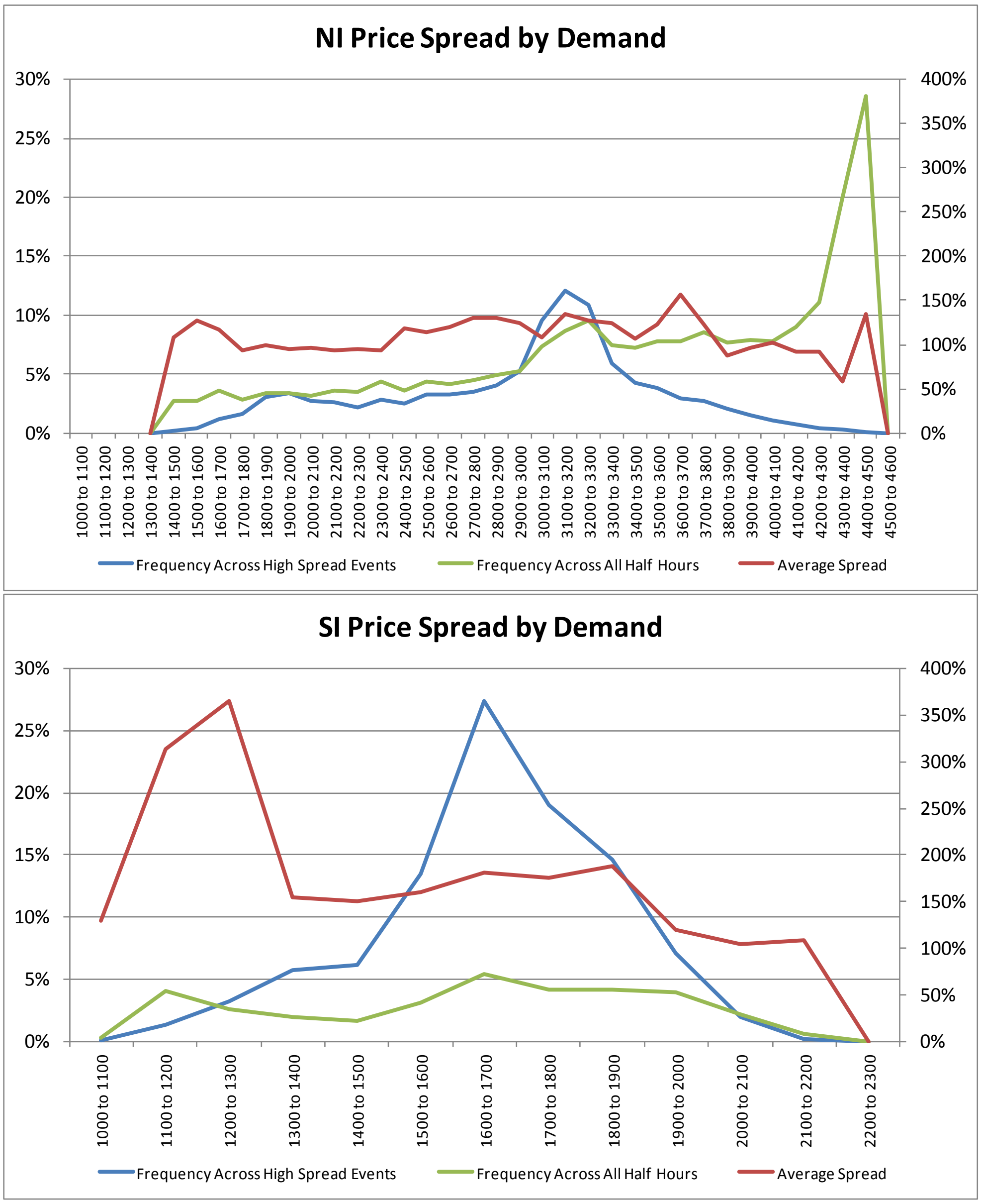
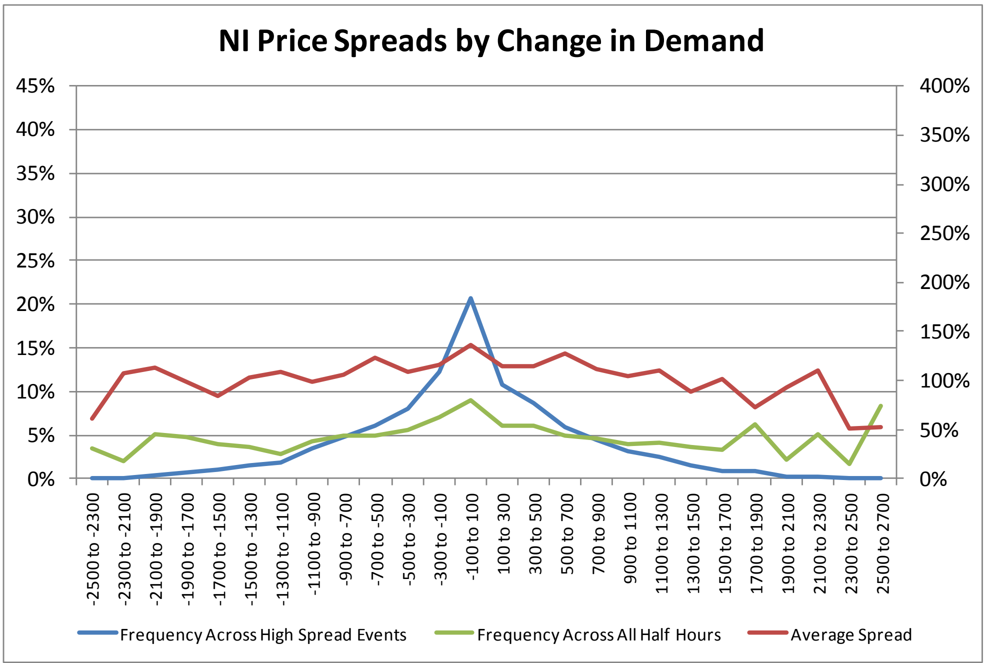


Figure 14 shows high price spread events by change in demand from one trading period to the next. Rather surprisingly, the frequency has a peak at around zero change in demand in both islands. In the NI the average spread has a weak peak at around zero change whereas the SI exhibits strong peaks at the extremes of high and low change in demand, with a smaller peak at around zero change.

Figure 14 – Price Spreads by Change in Demand



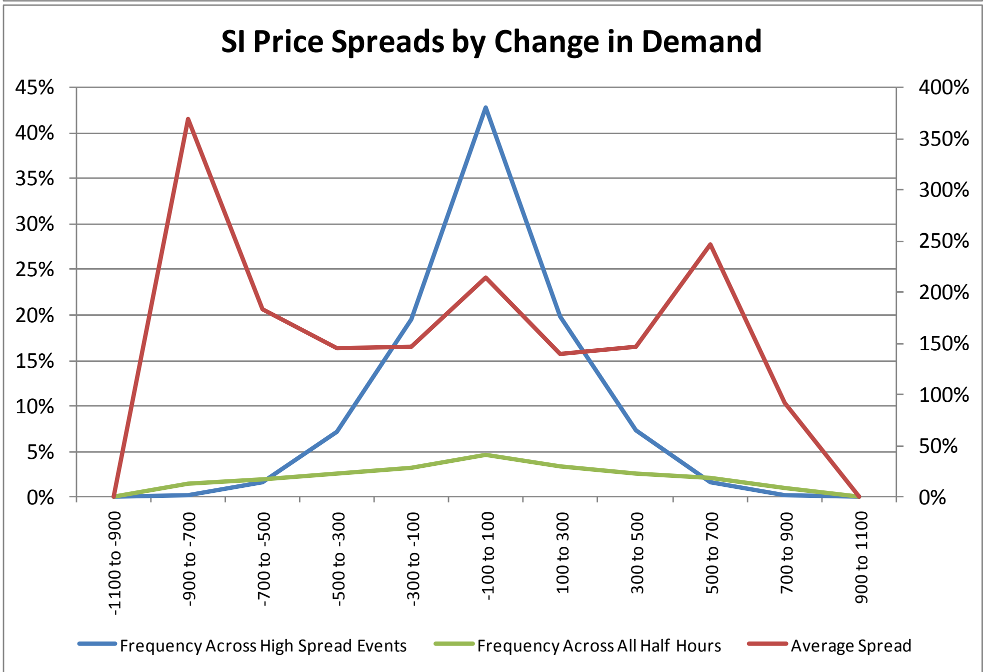
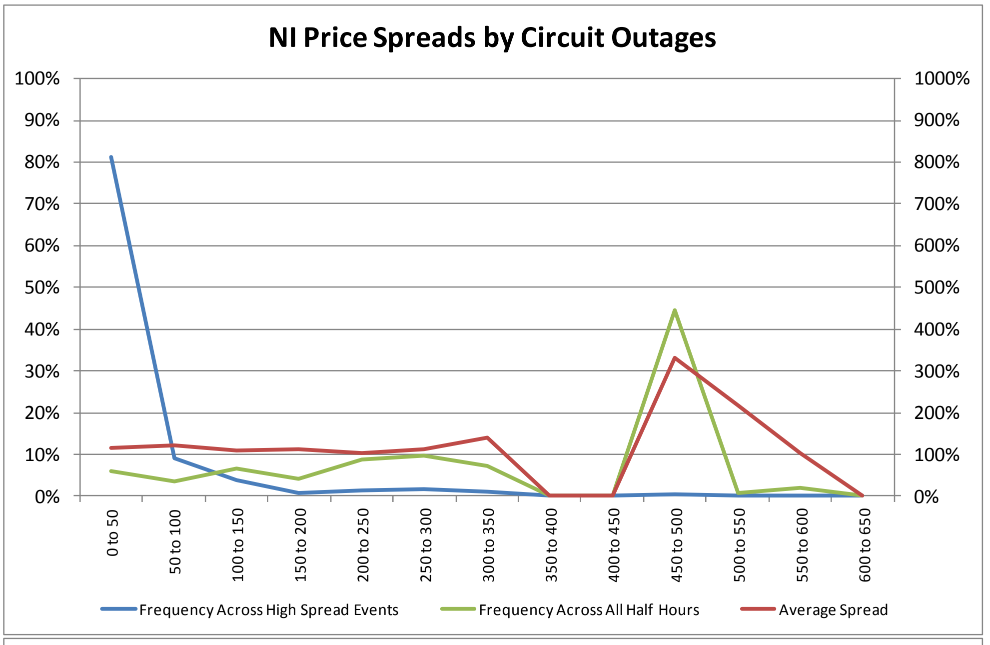


Figure 15 shows high price spread events by the total MW effectively lost through circuit outages. In the NI the green curve shows that events are much more likely at higher values of circuit outage, with a strong peak between 450 and 500 MW. The average spread also exhibits a peak at higher circuit outage values.

In the SI, the frequency and magnitude patterns are stronger than in the NI.

Figure 15 – Price Spreads by Circuit Outages in MW



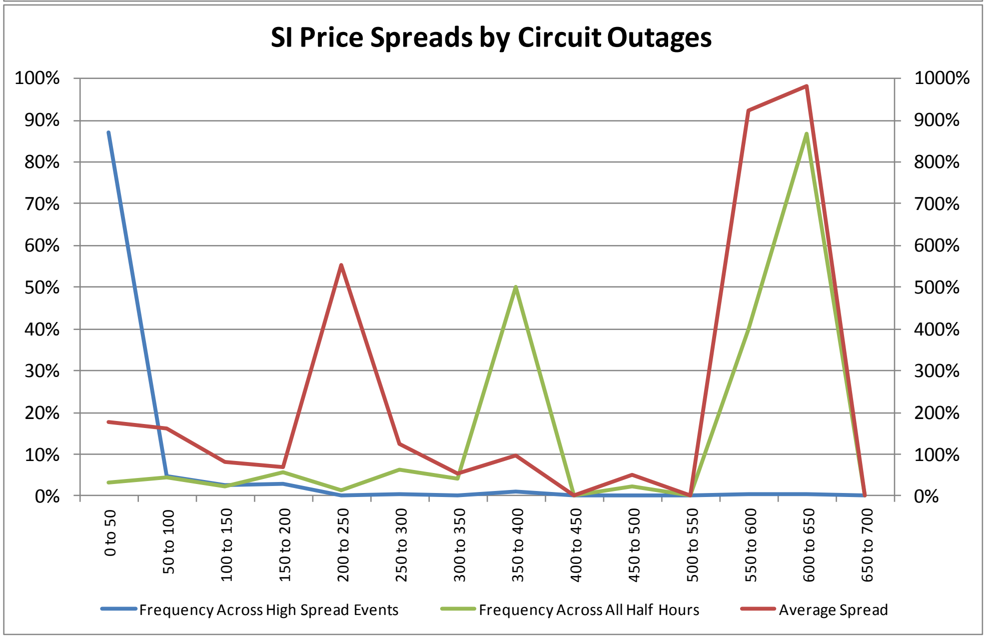


Figure 16 shows high price spread events by the change in circuit outages between trading periods. Both islands show peaks in frequency in the 0 – 100 MW of circuit outage change, but the green curves show that events are more likely when the rate of change in circuit outage is highly positive or negative, and particularly so in the SI. Both islands also show peaks in magnitude at the extremes, particularly in the SI.

Figure 16 – Price Spreads by Change in Circuit Outages

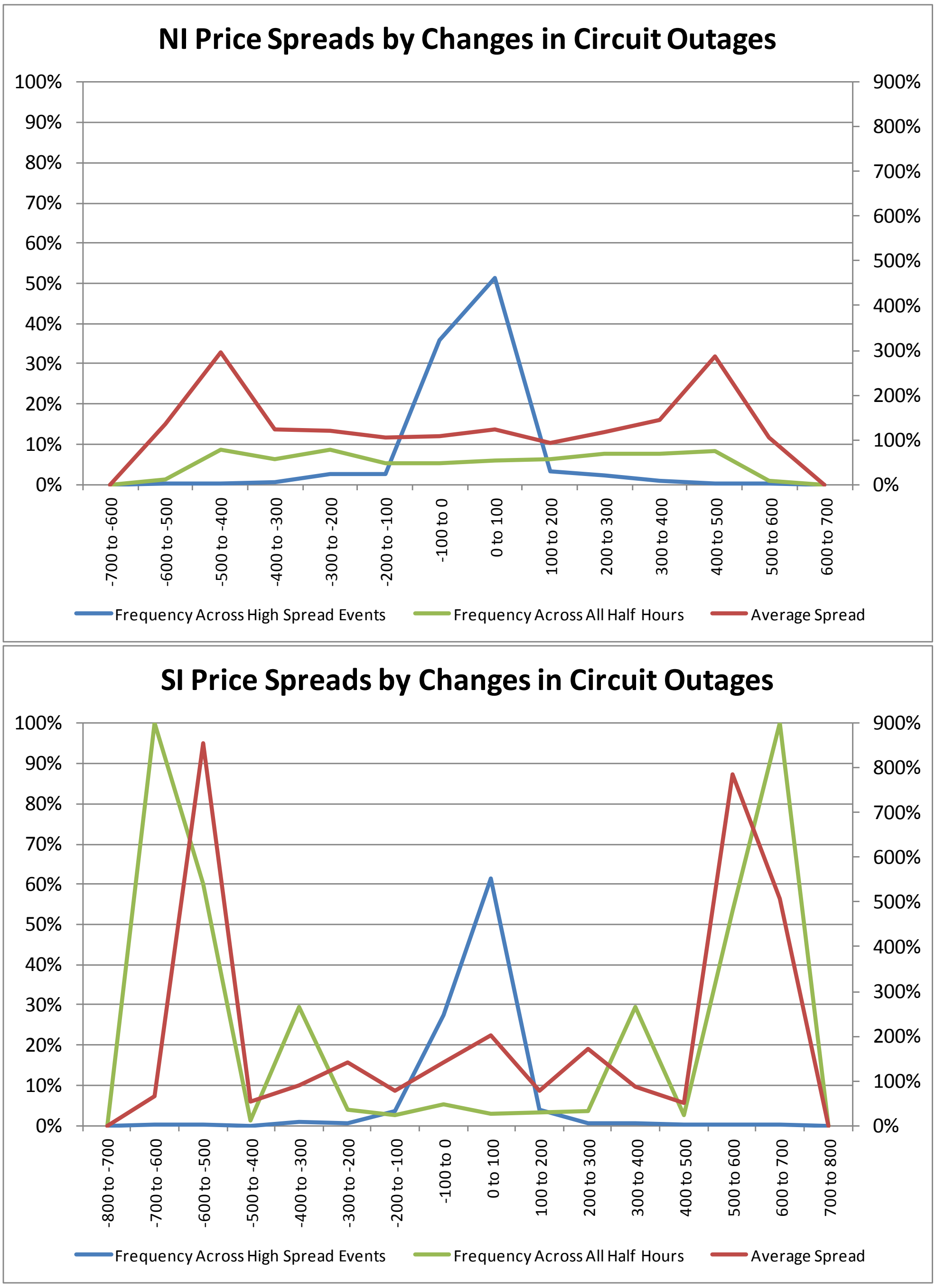
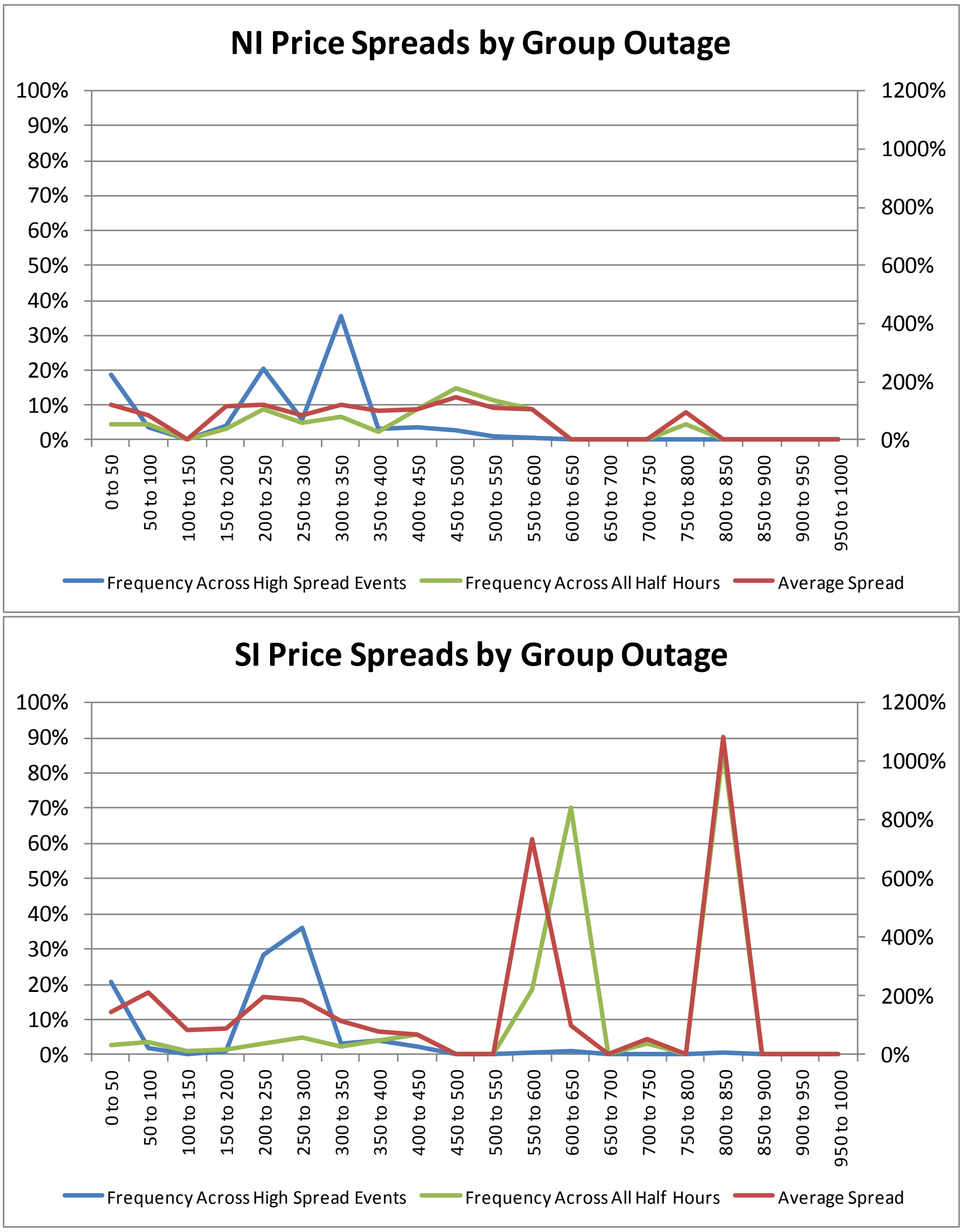


Figure 17 shows high price spread events by group outage[26]. The relationship between price spreads and group outages in the NI is weak, but in the SI frequency and average spread peak at higher group outage values.

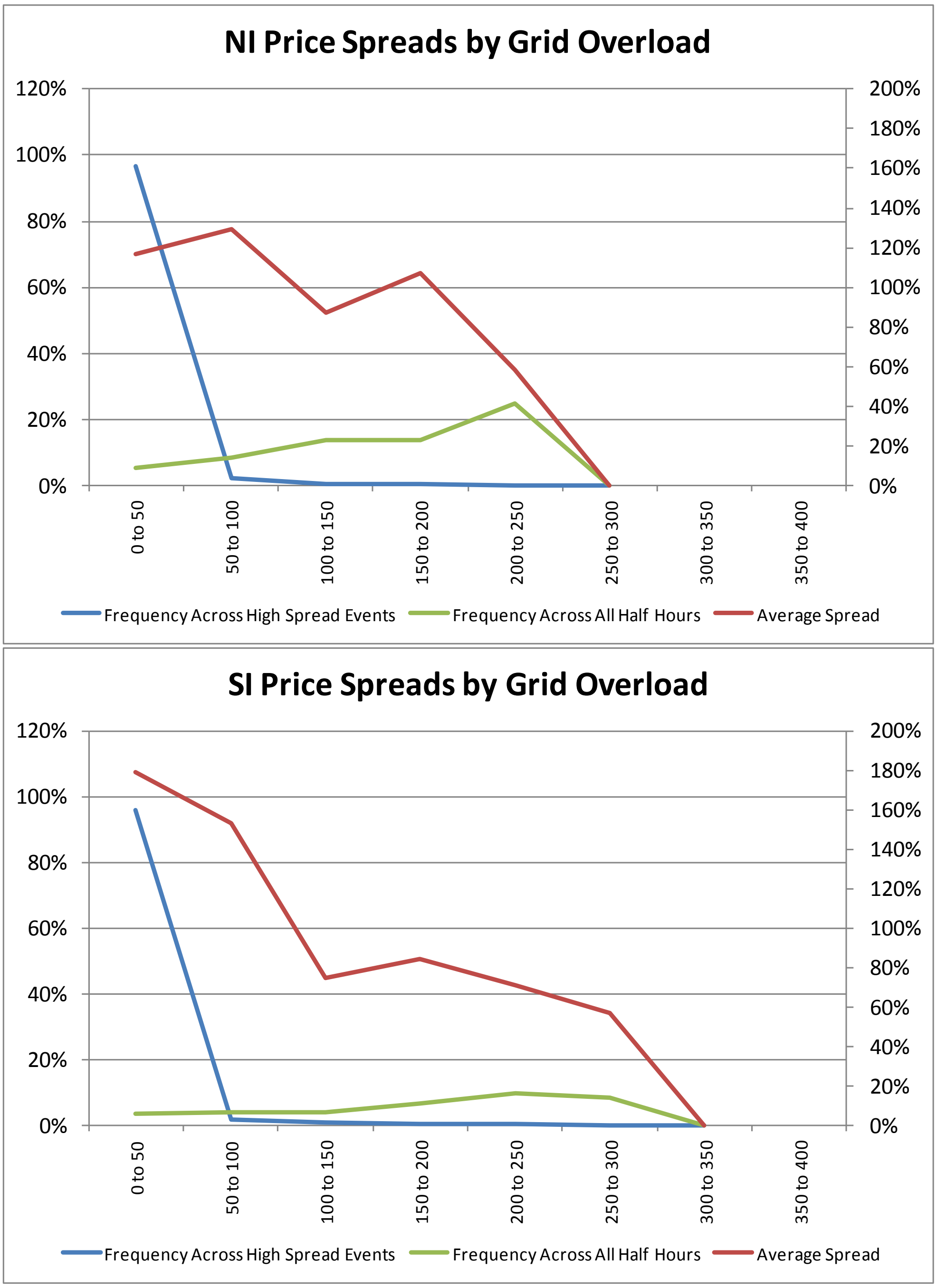
Figure 17 – High Price Spread Events by Group Outage in MW



[26] A group outage occurs when all of the lines directly connecting two nodes are taken out of service at the same time.

Figure 18 shows high price spread events by grid overload[27]. During periods with events, high price spreads occur more often at zero or low values of overload, although events are more likely generally when overload is higher. Somewhat surprisingly, average spreads reduce as overload increases.

Figure 18 – High Price Spread Events by Grid Overload



[27] The total MW over capacity that would result in increasing all arc flow in a trading period by 30%.

Figure 19 shows high price spread events by total generation planned outages in MW. In both islands the frequency amongst high price spread events peaks at lower outage values, but in the SI events are much more likely when the total generation outage is higher. In both islands the average price spread increases with the total MW in outage.

Figure 19 – High Price Events by Generation Outage in MW

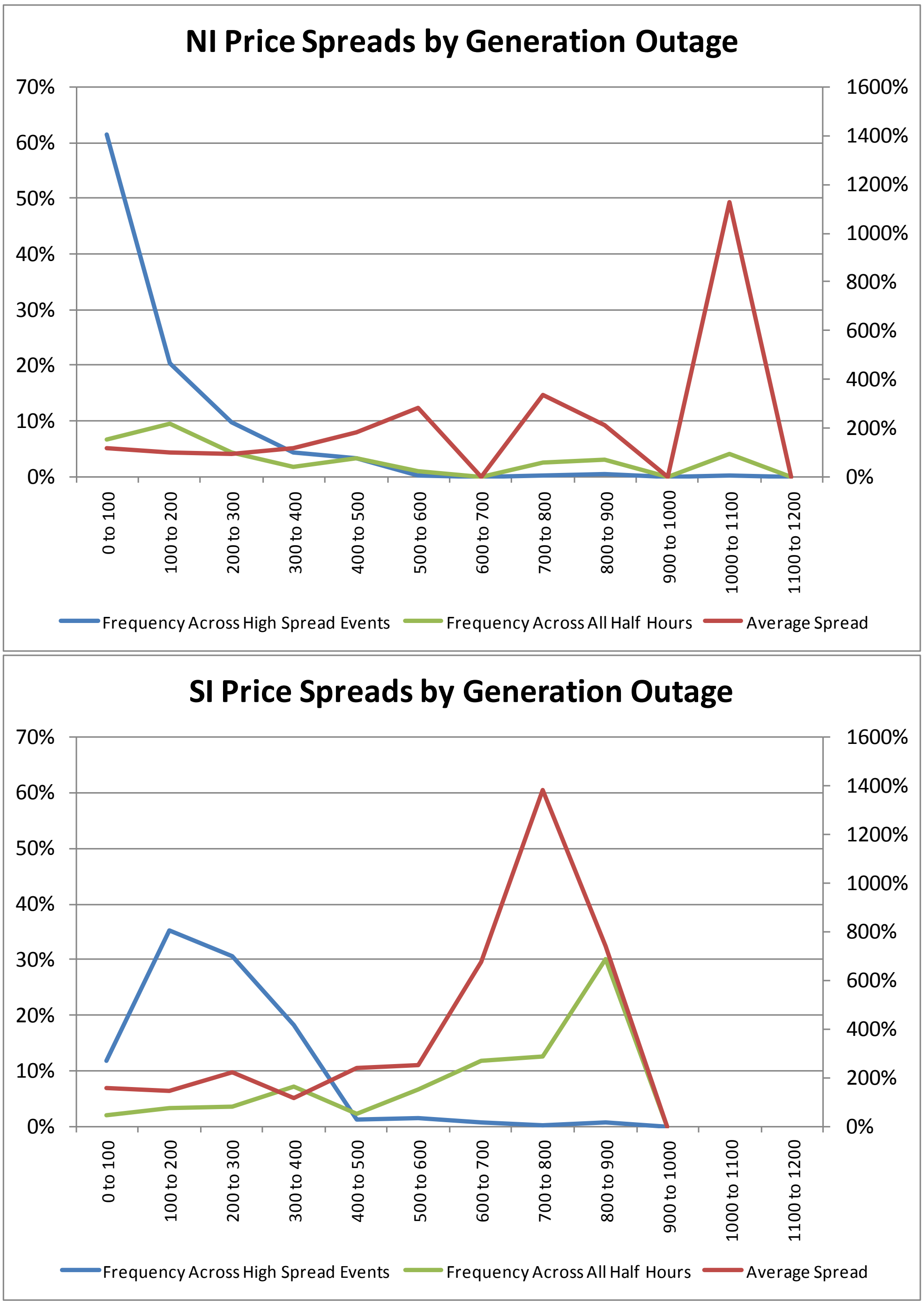
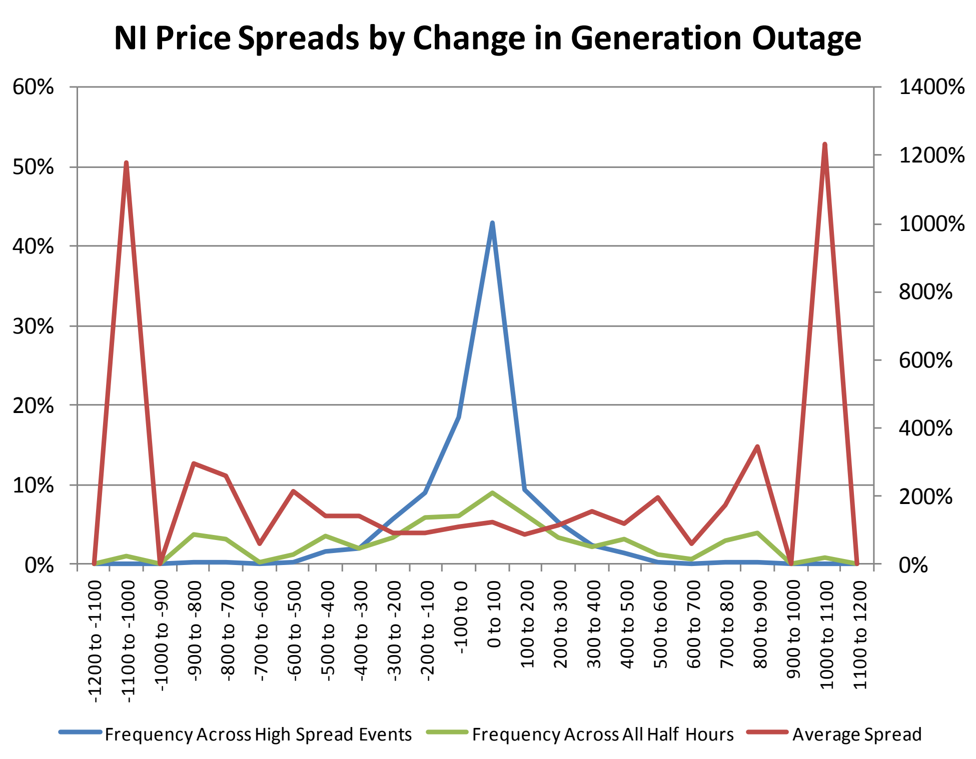


Figure 20 shows high price spread events by change in total generation outage. In the NI the frequency of events peaks strongly around zero change, whereas in the SI events are more likely to occur when the rate of change is high, either positive or negative. In the SI a high rate of increase in generation outage is associated with high average spread, and in the NI a high rate of decrease in outage is also associated with higher average spreads.

Figure 20 – High Price Spread Events by Change in Generation Outage in MW



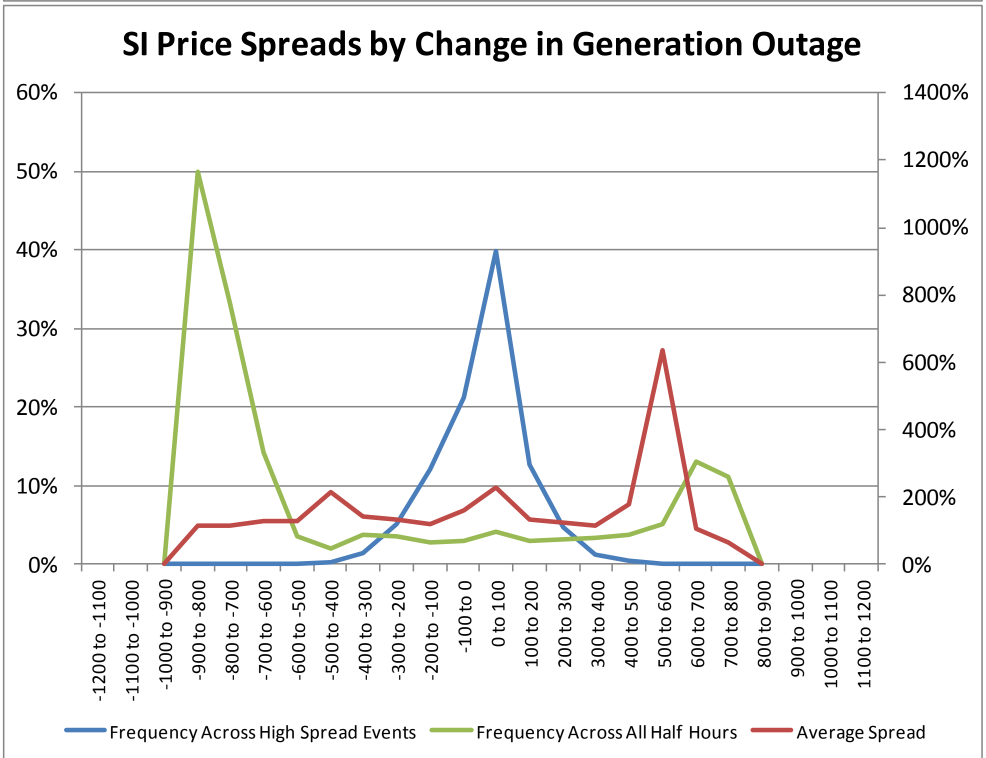
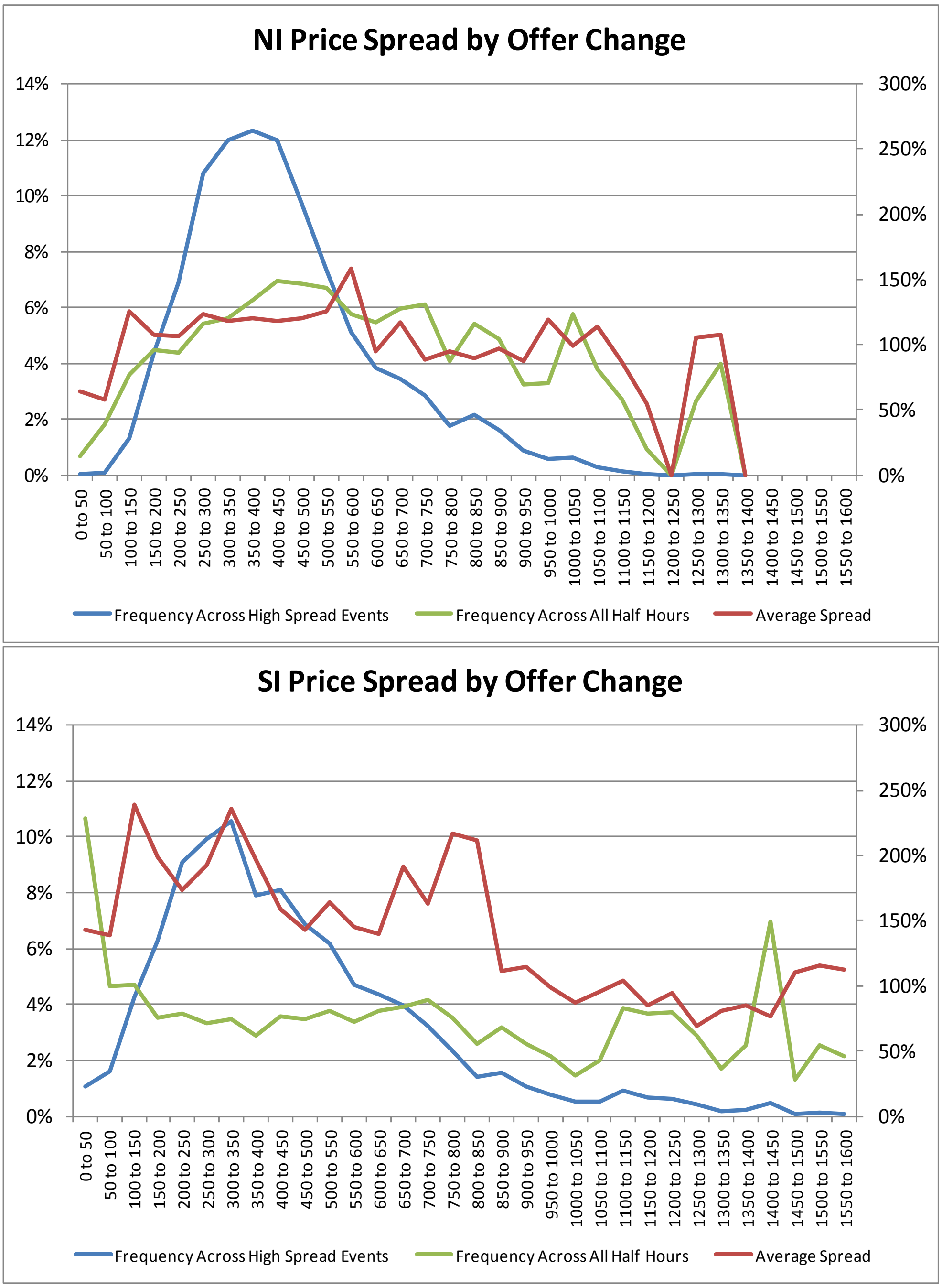


Figure 21 shows price spreads by the RMS change in offers[28] between trading periods. Amongst high price spread events, the frequency peaks at low to moderate offer change values (300 – 400 MW) in both islands. In both islands, frequency across all periods and average spread vary weakly across the range of offer changes. If anything, higher frequencies and spreads are associated with lower values of offer change.

Figure 21 – High Price Spread Events by Change in Offers



[28] The change in offers in each half hour is the square root of the mean of the squares (RMS) of the change in offer quantity across the following price bands since the same half hour in the preceding week: up to $10/MWh, $10 to $30, $30 to $50, $50 to $100, $100 to $300, $300 to $500, $500 to $1,000, $1,000 to $3,000, $3,000 to $5,000 and over $5000

Figure 22 shows high price spread events by the maximum offer price, which is calculated as the average offer price in each of the price bands used for the offer change analysis (so the maximum offer shown in the graphs may be less than the highest offer in the trading period). The NI shows a moderate peak in frequency over all periods at the high end of the maximum offer price range.

Figure 22 – High Price Spread Events by Maximum Offer Price

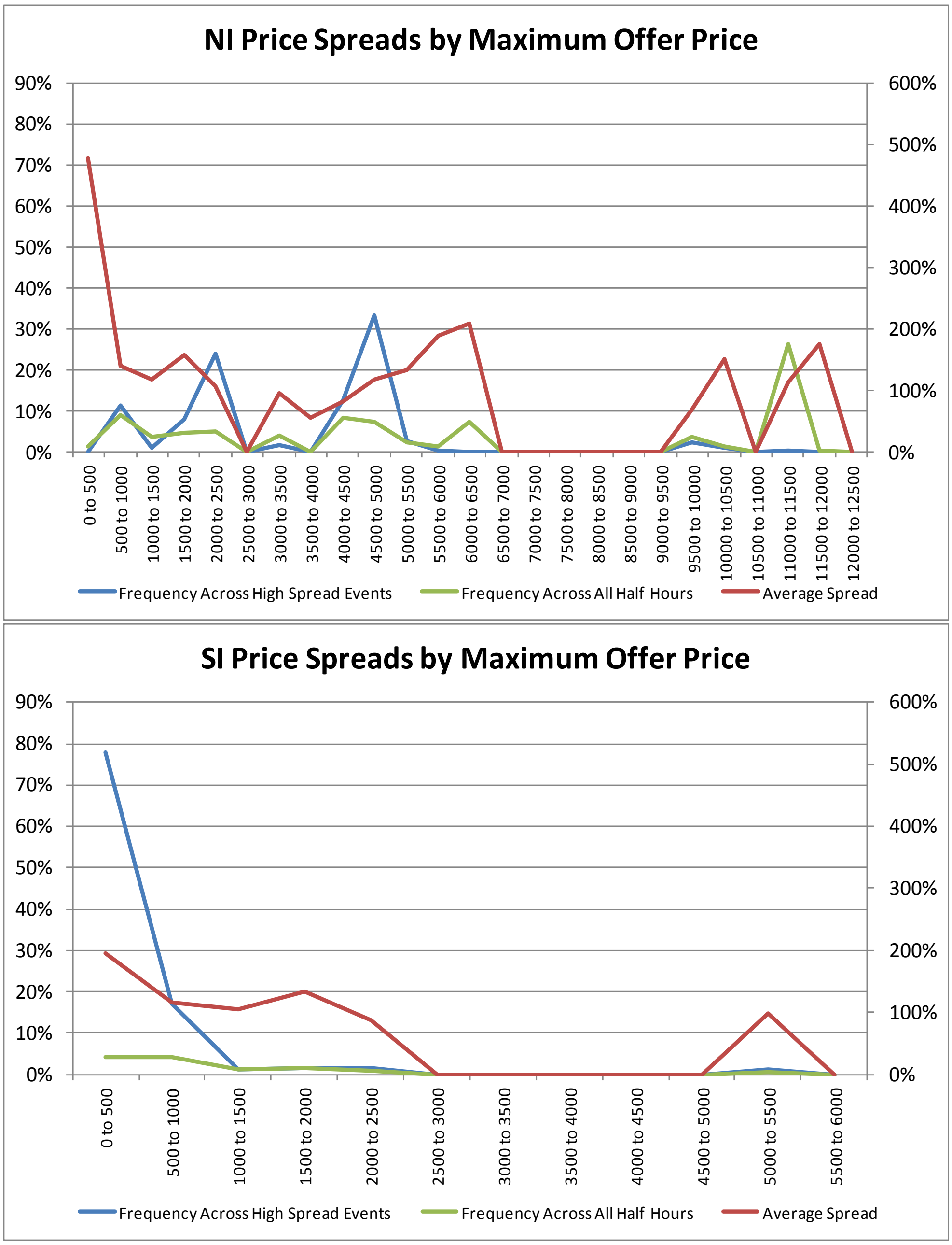
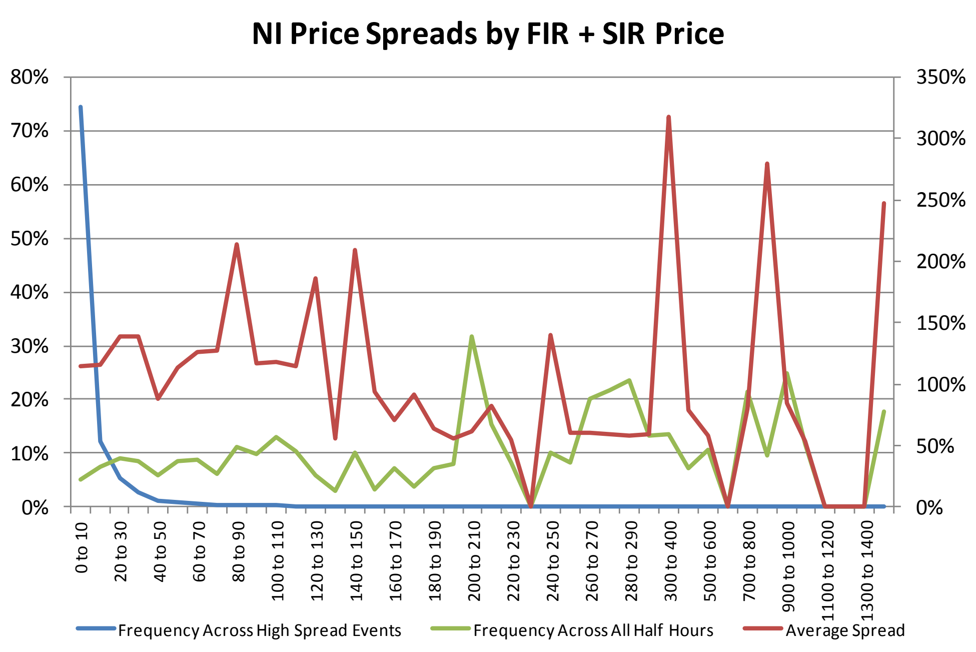


Figure 23 shows high price spread events by the sum of the FIR and SIR prices in the island, an indicator of the balance between the supply of and demand for reserves (i.e. higher price indicates tighter balance). About 95% of all events occur when the sum of the reserve prices is less than $40/MWh in the NI and $10/MWh in the SI. In the NI events are more likely to occur at higher values of the sum of the reserves prices, and average price spread is also greater at higher reserve prices.

Figure 23 – High Price Spread Events by FIR +SIR Price



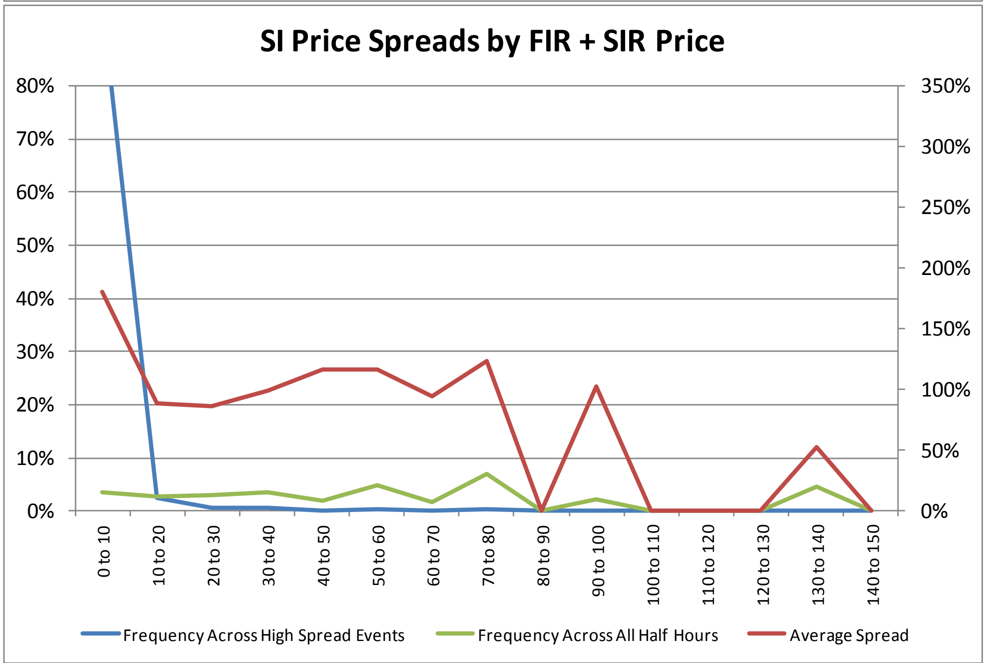


Figure 24 shows high price spread events by capacity margin, defined as the total quantity offered in the trading period less total island demand. The frequency of events when they occur peaks mid-range, but overall periods events are more likely to occur when the capacity margin is at the low end of the range, i.e. when the gap between supply and demand is small. The average price spread peaks at low to moderate values of capacity margin.

Figure 24 – High Price Spread Events by Capacity Margin

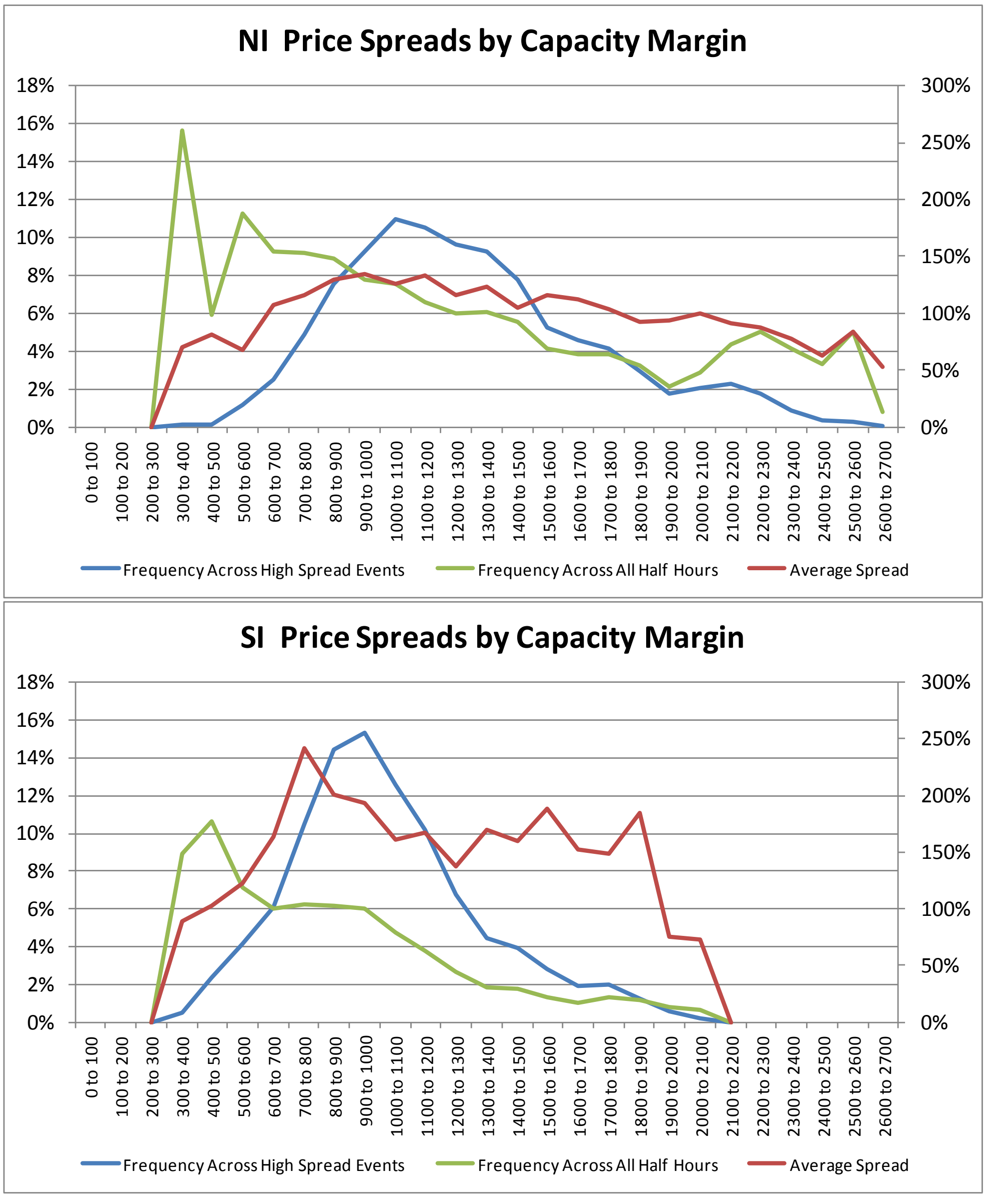


Figure 25 shows high price spread events by change in capacity margin. Amongst the events, the peak occurs at around zero change, and both islands exhibit a peak in average spread at this point, but with another peak in the SI at for larger decreases in capacity margin, and a small peak in the NI at high increases in capacity margin. Over all periods in both islands, events are most likely to occur around zero change and again at high rates of increase and decrease in capacity margin.

Figure 25 – High Price Spread Events by Change in Capacity Margin

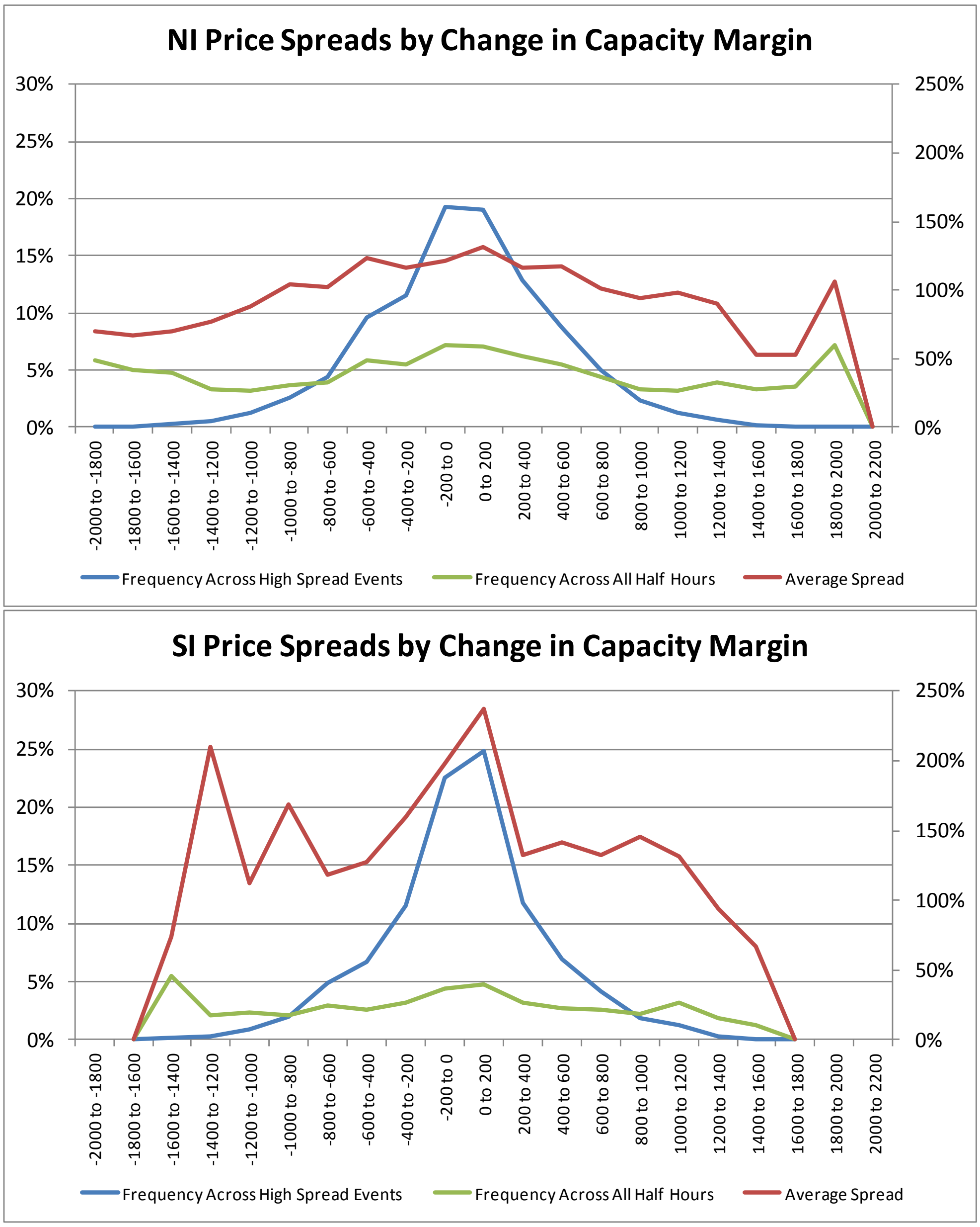
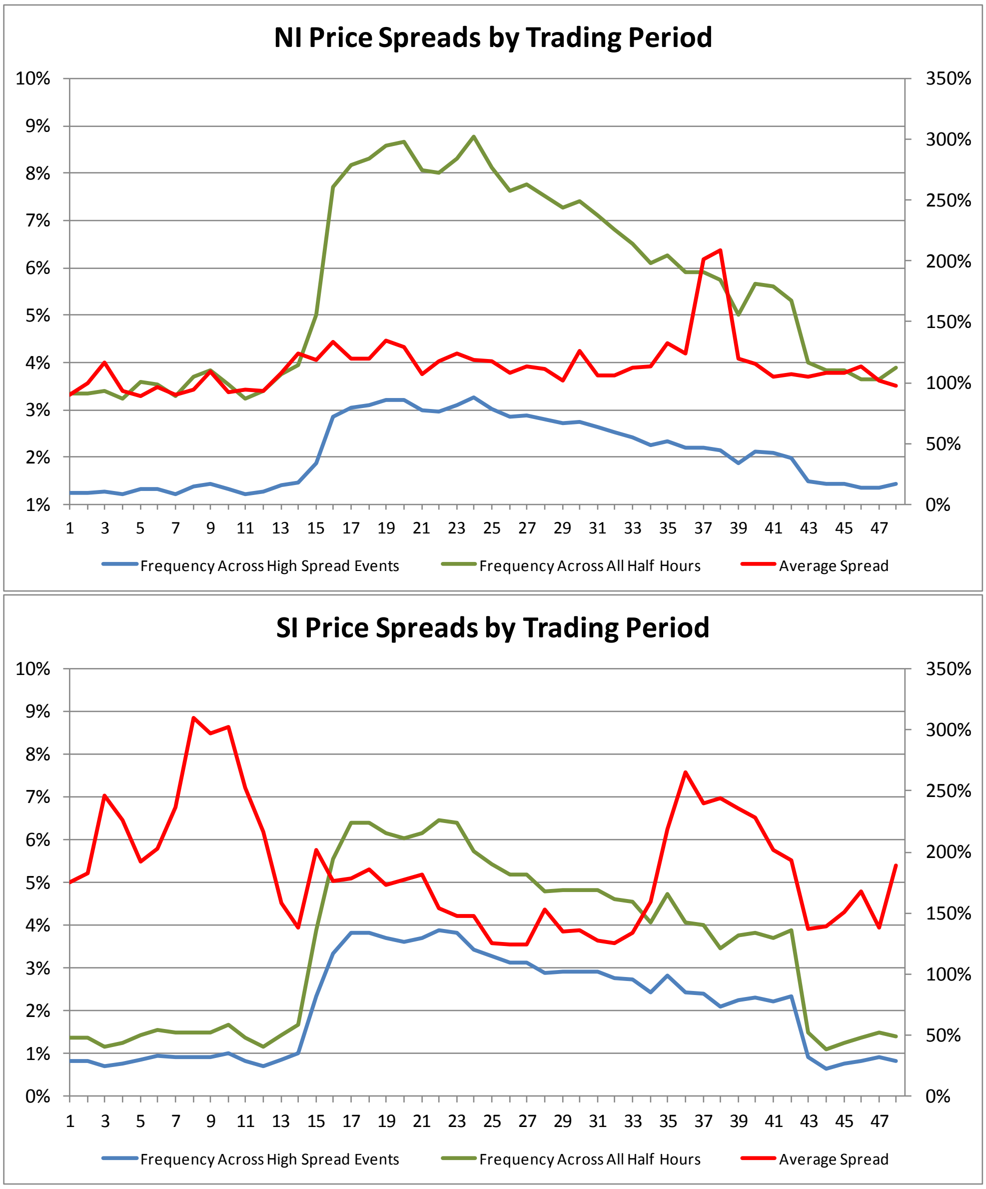


Figure 26 graphs the frequency and average value of high price spread events by trading period, and shows a distinct pattern: the frequency increases sharply in periods 15 and 16 (7:00 am to 8:00 am) to peak in a period shortly after, followed by a second peak two hours later, then falls through the rest of day to period 43 when it falls sharply again to its overnight level.

In both islands the average spread has an evening peak between periods 36 and 38, although in the SI the evening peak falls off slower than in the NI. The SI actually peaks overnight in periods 8 through 10.

Figure 26 – High Price Spreads by Trading Period



5.3 Summary of Observations

The observations from section 5.2 are summarised below.

1. High price spreads are more likely in the NI and are less tightly grouped around the average of 19.5% than they are in the SI (where they are tightly grouped around the average value of 21.5%);
2. The average spread during the high spread events trended slightly upwards in the NI in the last year of the study period;
3. **Time of Year**
   a. SI price spread frequency peaked in Jan, and April-May 2008 which are largely the result of the 2008 dry year;
4. **HVDC Flows**
   a. High price spread events in the NI are associated with higher flows on the HVDC link in both directions, but particularly with northward flow;
   b. High price spread events in the SI are associated with high northward flows on the HVDC link;
   c. Price spreads in the SI are higher when the HVDC link is flowing northward;
5. **Demand**
   a. In the NI, high price spreads are more likely with higher demand;
   b. In the SI, high price spreads are more likely when demand is between 1,500 and 1,800 MW but with average spread peaking at lower levels of demand up to 1,200 MW;
   c. The frequency of high price spread events peaks at around zero change in demand in both islands;
   d. The SI exhibits strong peaks at the extremes of high and low change in demand, with a smaller peak at around zero change;
6. **Circuit Outages**
   a. High price spread events are more likely and larger on average with higher circuit outages;
   b. High price spread events are more likely when the rate of change in circuit outage is high, but particularly in the SI;
   c. High price spread events are greater in magnitude at higher rates of change in circuit outages;
   d. In the SI, high price spread events are more likely and larger on average with higher group outage MW;
7. **Grid Overload**
   a. In both islands high price spread events are more likely with higher values of grid overload, but average price spreads reduce as overload increases;
   b. In the SI price spread events are much more likely when the total generation in outage is higher;
8. **Generation Outage**
   a. In both islands the average price spread increases with the total generation in outage;
   b. In the NI the frequency of high price spread events peaks strongly around zero change in generation outage;

c. In the SI events are more likely to occur when the rate of change in generation is high positive or negative;
d. A high rate of increase in generation outage is associated with high average spread in both islands;
e. In the NI a high rate of decrease in generation outage is also associated with high average spread;

9. Offer Change

a. Amongst high price spread events, the frequency peaks at low to moderate offer change values (300 – 400 MW) in both islands;
b. The NI shows a moderate peak in frequency over all periods at the high end of the maximum offer price range;

10. Reserves

a. The great majority of high price spread events occur when the sum of the reserve prices is less than $40/MWh in the NI and $10/MWh in the SI;
b. In the NI events are more likely to occur at higher values of the sum of the reserves prices, and the average price spread is also greater at higher reserve prices;

11. Capacity Margin

a. High price spread events are more likely to occur when the capacity margin is at the low end of the range;
b. Amongst high price spread events the peak in frequency occurs at around zero change in capacity margin;
c. Both islands exhibit a peak in average spread at zero change in capacity margin, with smaller peaks in the SI when capacity margin decrease is in the range from 1,400 - 1,600 MW, and in the NI when capacity margin increase is in the range from 1,800 – 2,000 MW;
d. In both islands the frequency of events peaks at around zero change in capacity margin and also at high rates of increase and decrease in capacity margin;

12. Time of Day

a. In both islands the frequency of high price spread events jumps in trading periods 15 and 16 to a peak shortly after, then falls progressively over the day, and finally falls sharply in period 43;
b. average magnitude of spread has an evening peak between periods 36 and 38;
c. in the SI the evening peak in magnitude falls off slower than it does in the NI;
d. in the SI the overall peak in the average magnitude of spread occurs overnight in periods 8 through 10.

5.4 Scoring the WIBR Drivers

Various attempts were made to quantify correlations between the WIBR drivers and the frequency and average magnitude of high price spread events, but these were frustrated by the highly non-linear character of many of the relationships evident in the graphs in section 5.2, plus the small number of samples at higher values of the WIBR drivers. In the end, a relatively simple and partly subjective method of scoring each WIBR driver was chosen and the results of this are shown in Table 3 and Table 4 below.

How the WIBR Drivers are Scored

The scoring system is not totally objective, but is intended to indicate the ranking of a WIBR driver in terms of the strength of the associations between the WIBR driver and either the frequency or magnitude of high price spread events. High accuracy is not required to achieve this ranking.

Scoring is done separately for the green curve (frequency over all half hours in the study) and the red curve (average magnitude of the spread) for each WIBR driver. The blue curve is not scored. A maximum and minimum value of the relevant variable is taken from the data behind the relevant graph over the range of values taken by the WIBR driver, and one is divided by the other. If an increase in the frequency or magnitude is associated with an increase in the WIBR driver, then division is done so that the score is greater than one, and vice versa, down to a minimum possible value of zero.

If the driver is a rate of change driver, for example the change in demand by period, it is was scored based on the highest of the scores at the negative and positive change ends of its graph.

Some drivers have peaks near the end of their range, but with a sharp fall at higher (or lower) values, so in these cases the peak value was taken.

Some drivers have peaks near the end of their range which could be the result of random effects in a small number of samples at the extreme values, but no attempt was made to correct for this.

A driver that appears to have no association with a WIBR driver would appear on its graph as a straight horizontal line and would score one.

Example: HVDC Flow and NI high price spread events (Figure 12)

The green frequency curve peaks at just under 27% in the 650 – 700 MW north range, and has a minimum of 2.4% in the 0 – 50 MW north range, so northward HVDC flow is obviously associated with high price spread events, and more so than southward flow. So HVDC flow frequency is scored as 27/2.4 = 11.3 which puts it at the top of the list for frequency of high price spread events in the NI.

The tables below are sorted first by frequency score and then by magnitude score. Table 5 shows the average of the scores across both islands.

Table 3 – NI WIBR Scores

NI Driver	Frequency Score	Magnitude Score
HVDC Transfer	11.3	1.9
Demand	10.6	1.0
Time of year (month)	7.5	2.2
Circuit Outage	7.5	2.9
Capacity Margin	7.0	0.5
FIR + SIR Price	5.0	2.2
Grid Overload	4.5	0.5
Group Circuit Outage	3.4	1.0
Maximum Offer Price	2.9	0.4
Trading Period	2.8	2.3
Circuit Outage Change	1.6	2.7
Capacity Margin Change	1.0	0.6
Demand Change	0.9	0.4
Offers Change	0.6	0.7
Generation Outage	0.4	9.6
Generation Outage Change	0.4	9.9

Table 4 – SI WIBR Scores

SI Driver	Frequency Score	Magnitude Score
Circuit Outage Change	33.3	3.9
Circuit Outage	25.6	5.6
Group Circuit Outage	19.4	7.4
Generation Outage	15.9	8.8
Capacity Margin	13.4	0.3
Time of year (month)	10.0	3.8
HVDC Transfer	7.1	2.6
Trading Period	6.5	2.5
Generation Outage Change	4.3	5.4
Grid Overload	2.9	0.3
FIR + SIR Price	1.3	0.3
Capacity Margin Change	1.0	0.9
Offers Change	0.2	0.5
Demand Change	0.2	1.2
Maximum Offer Price	0.2	0.5
Demand	0.1	0.3

Table 5 – Average Score across both Islands

Driver Both Islands	Average Frequency Score	Average Magnitude Score
Circuit Outage Change	17.5	3.3
Circuit Outage	16.5	4.2
Group Circuit Outage	11.4	4.2
Capacity Margin	10.2	0.4
HVDC Transfer	9.2	2.2
Time of year (month)	8.8	3.0
Generation Outage	8.2	9.2
Demand	5.4	0.7
Trading Period	4.6	2.4
Grid Overload	3.7	0.4
FIR + SIR Price	3.2	1.2
Generation Outage Change	2.4	7.7
Maximum Offer Price	1.5	0.4
Capacity Margin Change	1.0	0.7
Demand Change	0.6	0.8
Offers Change	0.4	0.6

A score above one indicates a degree of positive association (correlation) between the WIBR driver and the frequency or magnitude of high price spread events. In the NI, 11 drivers score above one for frequency, but only eight for magnitude. The strongest associations for magnitude are with generation outages, but there are several strong associations with frequency.

In the SI, 11 drivers score above one for frequency, but this set of drivers is different to the 11 in the NI. Circuit and group outage, and generation outage are particularly strong drivers for magnitude in the SI, and a total of nine drivers scoring above one for magnitude.

In both islands, magnitude is considerably less likely than frequency to be strongly associated with any particular WIBR driver. Across the two islands, no driver gains a score greater than 10 for magnitude versus eight for frequency, and there are 14 scores of five or greater for frequency against six for magnitude. The lower ratings for magnitude are particularly evident in the NI where high price spread events are twice as common as in the SI. This all suggests that the magnitude of high price spread events, when they occur, has a significant random component, which in turn suggests that the magnitude could be harder to predict accurately than frequency.

But why is it, for example, that capacity margin is strongly associated with frequency but not with magnitude? As Figure 24 shows, low capacity margin is much more likely to be associated with high price spreads than high capacity margin, which intuitively makes sense (the system may be under stress, and patterns of flow atypical, when the capacity margin is low), but the average spread falls off with capacity margin below 700 MW, which is counterintuitive.

Even more puzzling is the low ranking of change in demand: one could reasonably assume that rapid changes in demand might be associated with system stress and high price spreads. But Figure 14 shows that frequency peaks when the change in demand is around zero, and similarly the magnitude has a peak at around zero change in demand.

The answer may be found in the interaction of a number of factors, and the biggest clue is the graphs shown in Figure 26, which shows the frequency and magnitude of high price spread events by trading period. Although trading period on its own is not as strongly associated with high price spreads as a number of other WIBR drivers, the patterns exhibited in Figure 26 are so striking, and so similar in both islands for frequency, in particular, that they warranted further investigation.

The pattern of frequency across the day raised a number of questions. For example, one could easily imagine that the morning peak in demand could be associated with high price spread events, but then why is the evening peak less so? What other factors are at play across the day?

The following graphs show a number of the strongest WIBR drivers plotted across the day, starting with Figure 27 which has average HVDC flow by trading period, plotted alongside the frequency of high price spread events in the NI. As the following graphs show, the strongest WIBR drivers also tend to exhibit strong seasonality across the day with rapid change leading into the morning peak.

Figure 27 – Average HVDC Flow by Trading Period

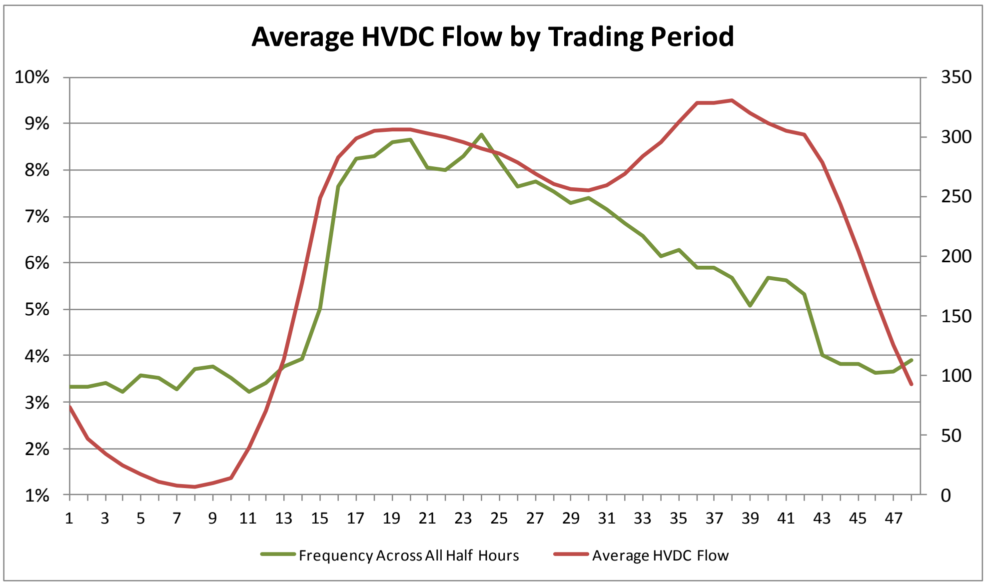


Figure 28 shows the frequency of high price events over all periods (green curve), along with average island demand and capacity margin by trading period.

Figure 28 – Average Demand and Capacity Margin by Trading Period

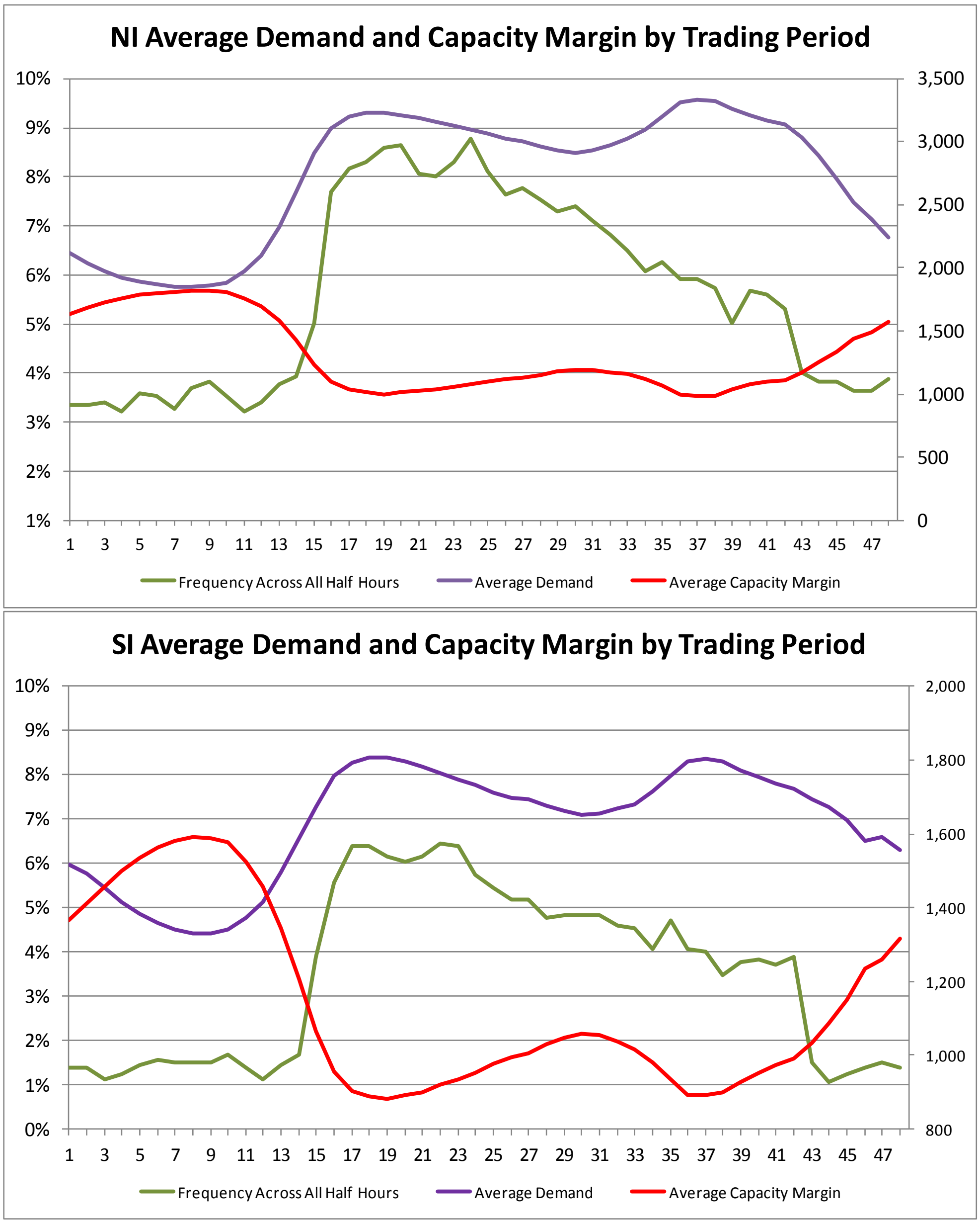
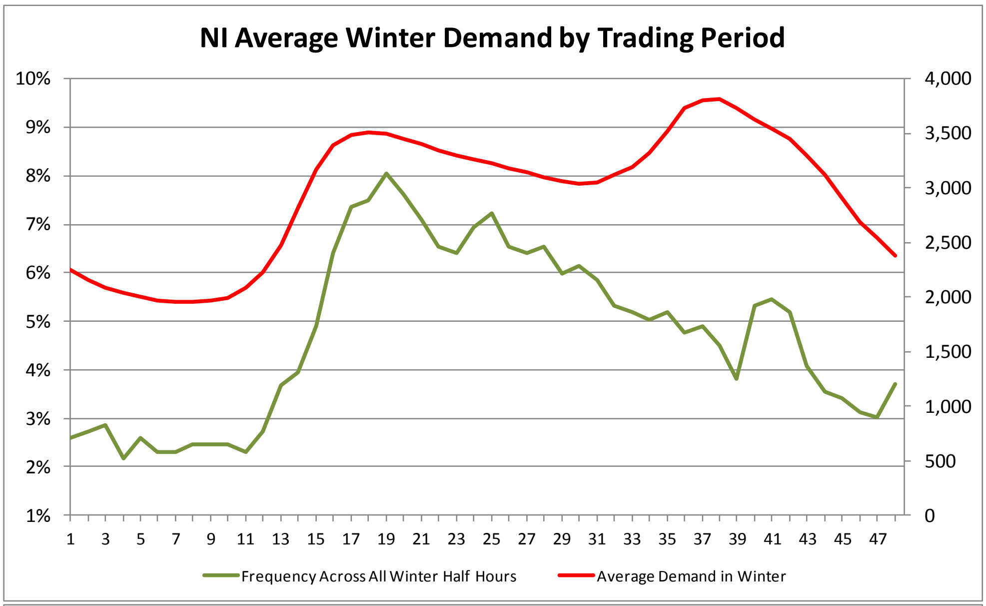


Figure 29 and Figure 30 shows winter and summer demand by trading period, alongside the frequency of high price spread events by trading period.

Figure 29 – Average Winter Demand and by Trading Period



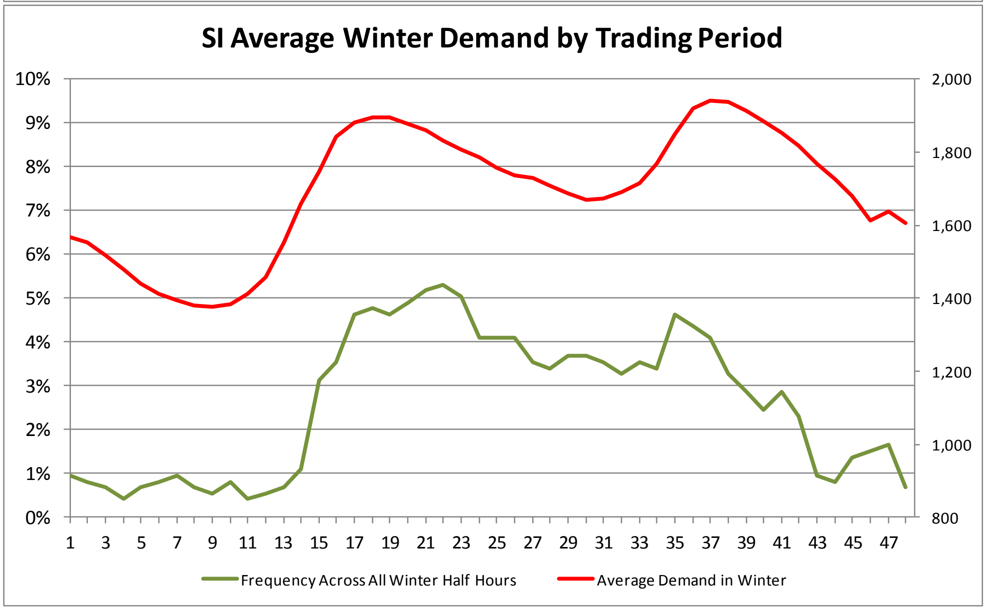
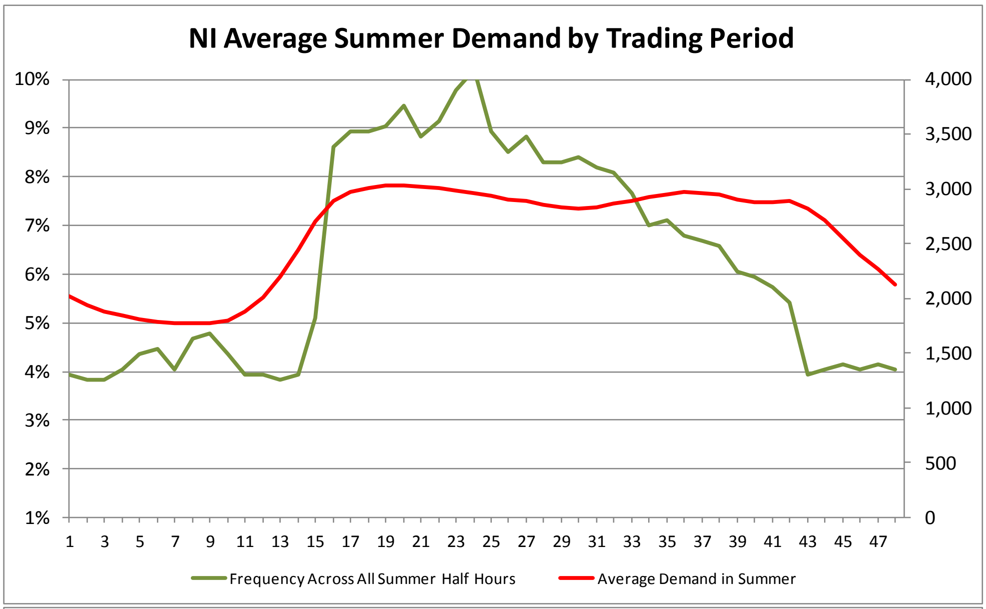


Figure 30 – Average Summer Demand by Trading Period



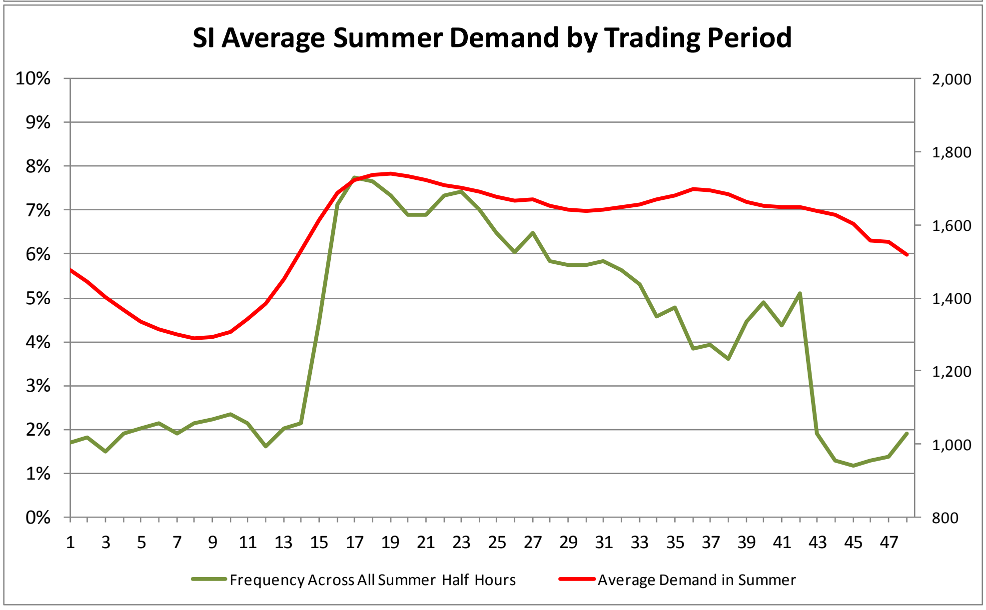


Figure 31 and Figure 32 show the average price spread by trading period alongside the frequency of high price spread events by trading period.

Figure 31 – Average Price Spread in Winter

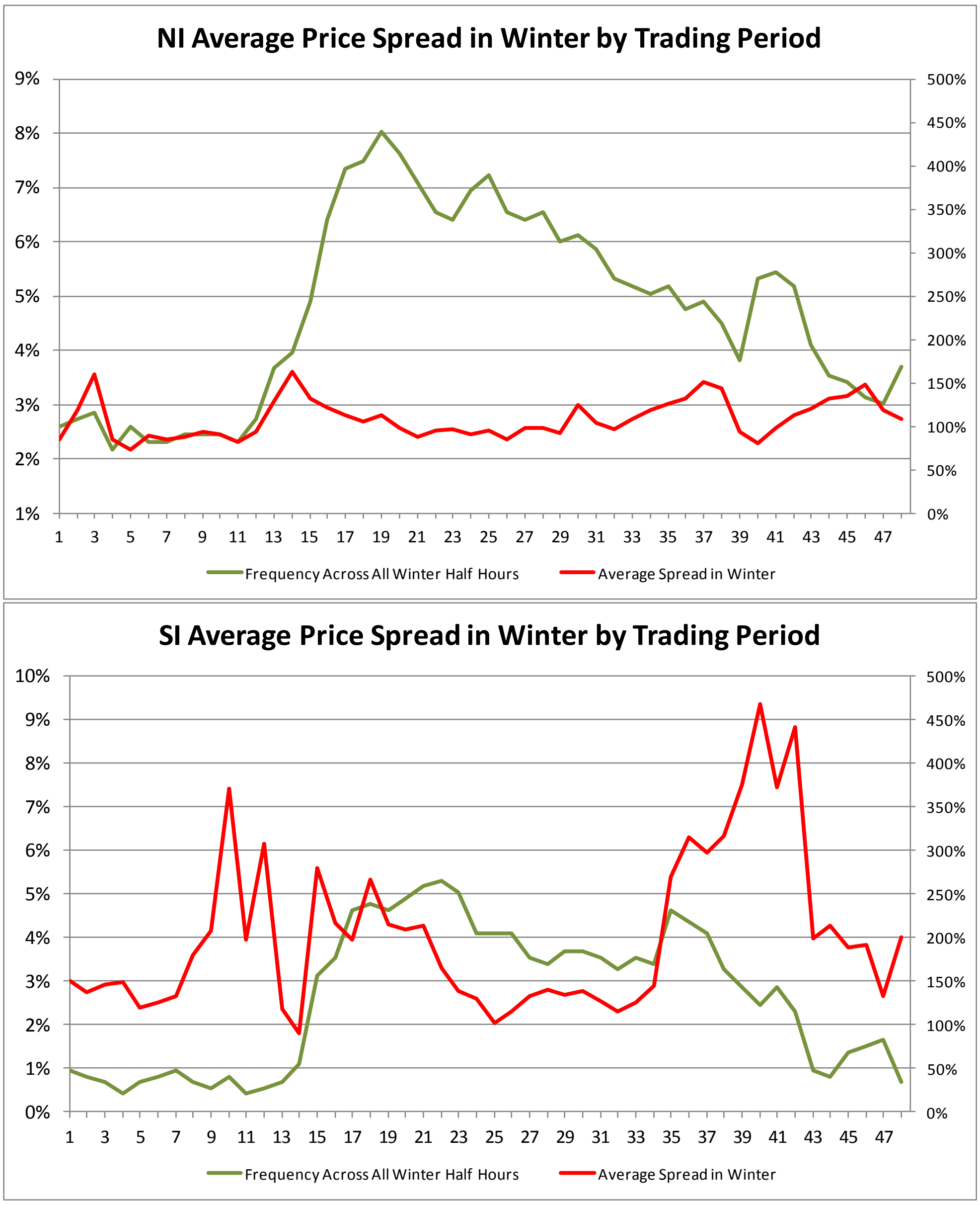


Figure 32 - Average Price Spread in Summer

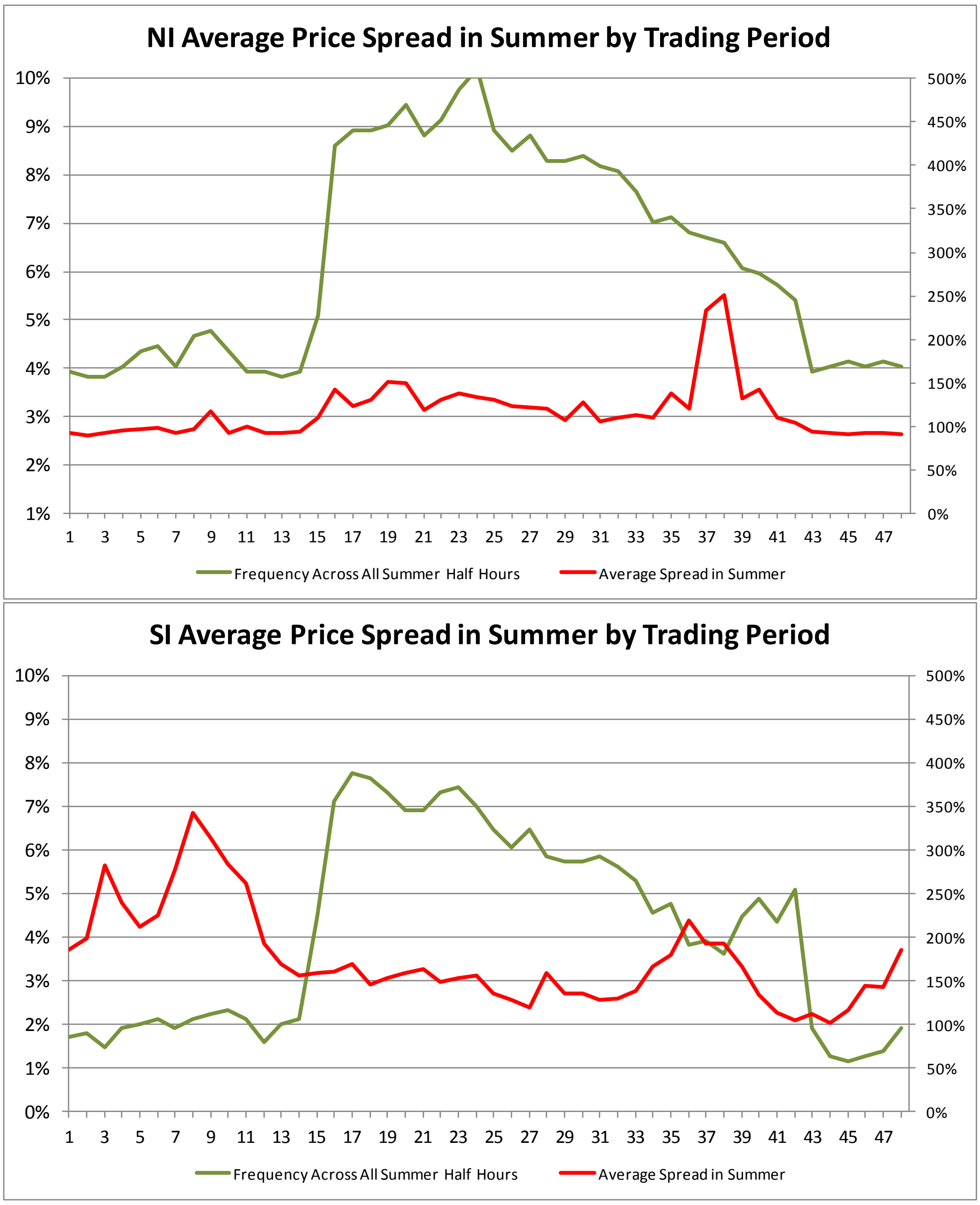
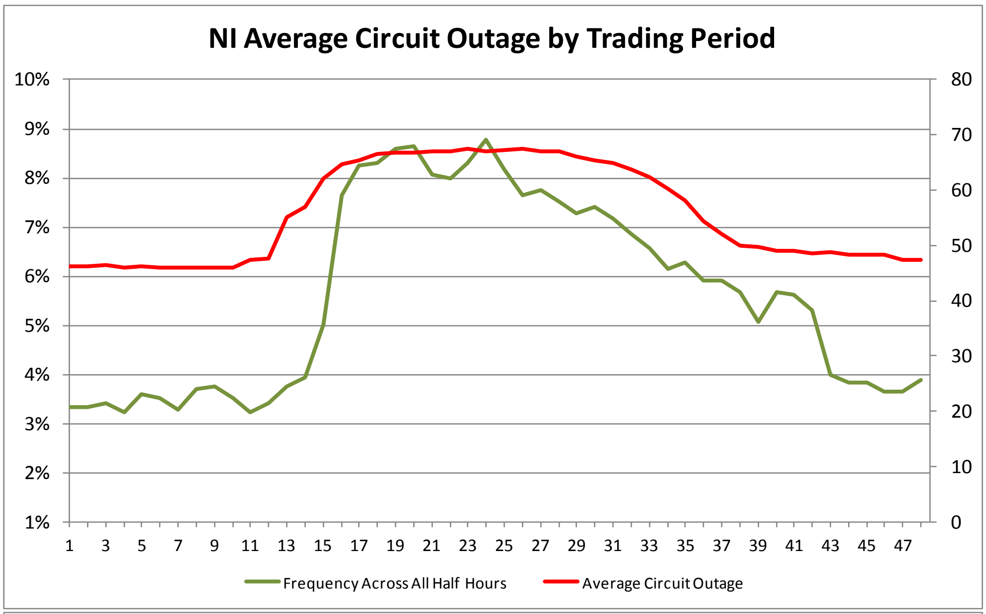


Figure 33 shows the frequency of high price events over all periods, along with average circuit outage.

Figure 33 – Average Circuit Outage by Trading Period



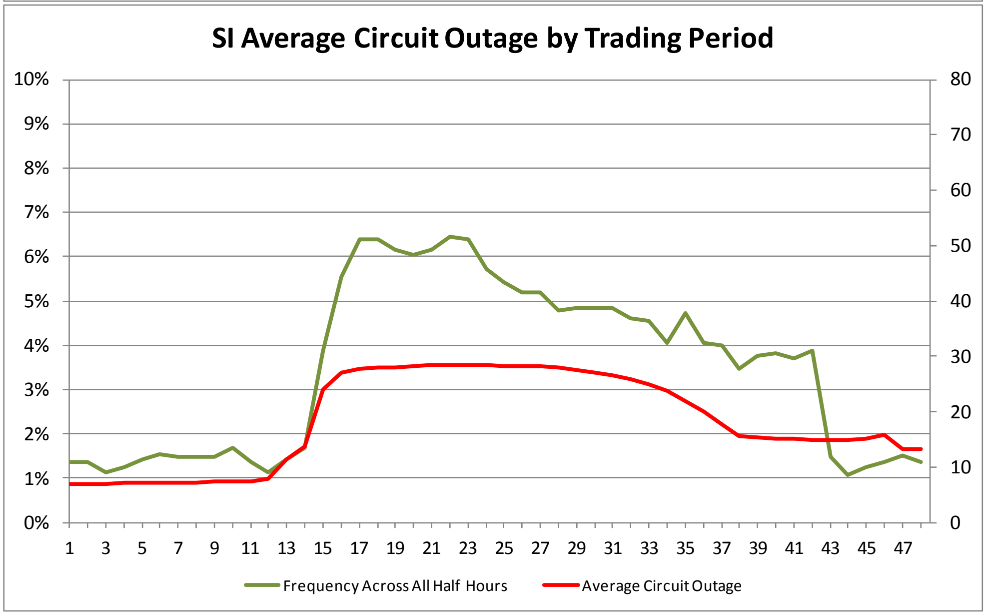
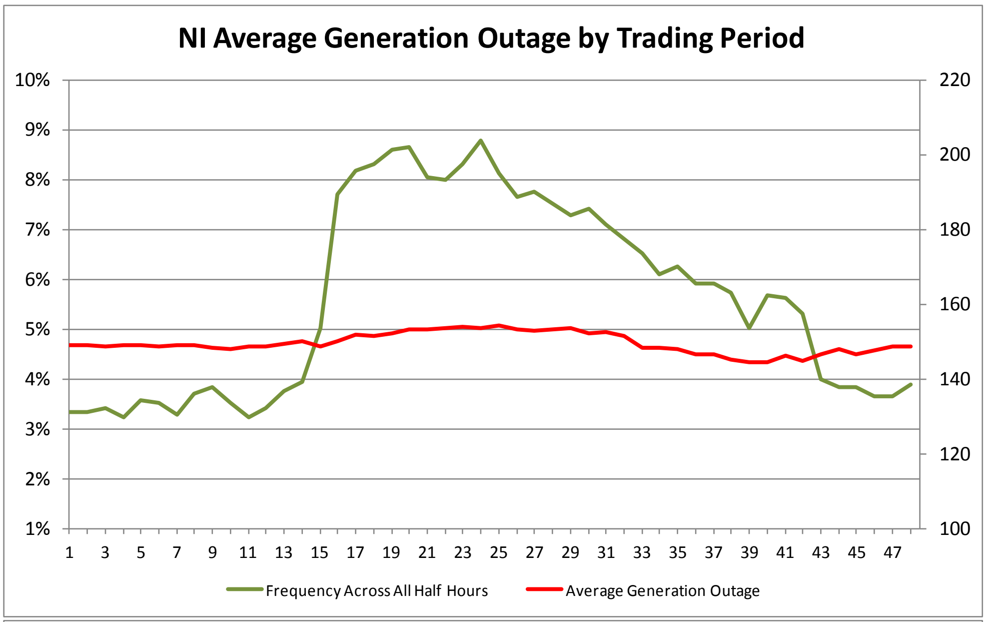
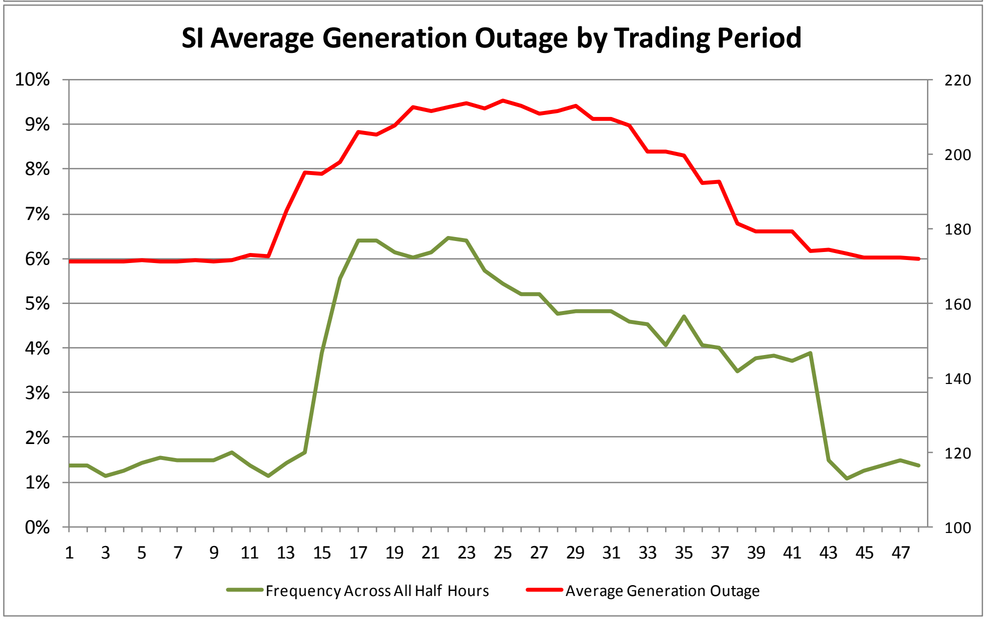


Figure 34 shows the frequency of high price events over all periods, along with average generation outage.

Figure 34 – Average Generation Outage by Trading Period





Correlations between the frequency by trading periods and the averages graphed above are shown in Table 6 below.

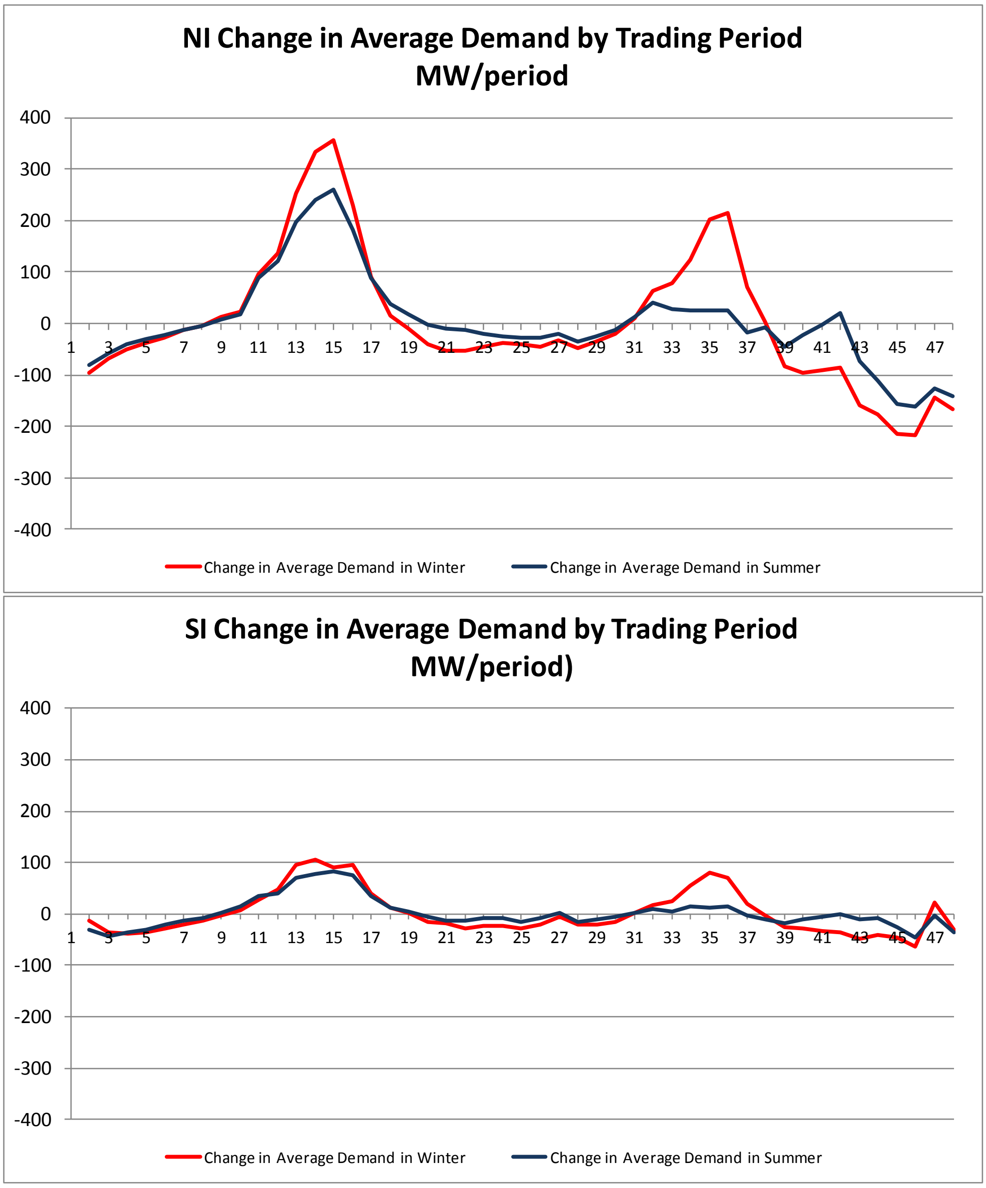
Table 6 – Correlations with Frequency by Trading Period

Island	Average HVDC Flow	Average Demand	Average Winter Demand	Average Summer Demand	Capacity Margin	Average Generation Outage	Average Circuit Outage
North Island	0.78	0.78	0.72	0.83	-0.82	0.66	0.93
South Island	0.79	0.78	0.80	0.75	-0.82	0.91	0.93

Given the relatively large of the sample size, the values in Table 6 indicate strong correlations with the frequency of high price spread events across the day.

Figure 35 shows the change in average demand by trading period for both winter and summer. The change in average demand is highest in the morning peak, particularly in winter in the NI, but despite rapid changes in demand leading into the evening peak, the frequency of high price spread events consistently falls across the day before falling off to the overnight level in period 43.

Figure 35 – Change in Average Demand by Trading Period



6 Discussion

Phase 3 of the study involved analysing the results of phase 2 for correlations, and causal relationships where possible, and then using this information to project into the foreseeable future, attempting to answer the question: will WIBR increase, decrease or stay the same in future? The answer will inform the LPRTG's work on the need for, and design of instruments that hedge WIBR.

The WIBR driver scores in the tables in section 5.4 show the associations between a number of drivers and the frequency and average magnitude of high price spread events over the study period. These associations and the outlook are summarised in Table 7 below.

Table 7 – Summary of WIBR Drivers

NI Driver	Impact Assessment	Outlook
Capacity Margin	Strongly associated with frequency of high price spreads in the SI, moderately so in the NI. Appears to play a key role in the morning peak, along with demand.	Capacity margin will fluctuate depending on a range of factors, but has reduced since the 1990s. Unlikely to significantly increase in future on average, but will fluctuate from year to year depending on increments in supply and demand.
Circuit Outage Group Circuit Outage	Strongly associated with frequency and magnitude of high price spreads in the SI, and with frequency in the NI. The increase of circuit outages seen leading into the morning peak probably adds to the risk of high price spread events.	Outages will continue in future. Unless the timing of outages can be changed, e.g. started earlier, they will continue to create WIBR.
Demand	Strongly associated with frequency of high price spreads in the NI, with a key role in the morning peak along with capacity margin.	Demand is forecast to grow in the longer term, and there is no evidence to suggest that the ramp up to the morning peak will slow in future.
FIR + SIR Price	Moderate association with frequency of high price spread events in the NI.	The addition of Pole 3 will reduce the propensity of the HVDC link to set the reserve risk in both islands, leading to a general fall in reserve prices and less WIBR, along with greater transfers across the HVDC link.
Generation Outage	Strongly associated with frequency of high price spread events in the SI, moderately associated with magnitude of these events in both islands.	There is nothing to indicate that generation outages will get any less or more common in future.
HVDC Transfer	Moderate to strong association with frequency of high price spread events in both islands.	The addition of Pole 3 will allow significantly greater transfers in both directions on the HVDC link relative to most of the study period, to a large extent because the HVDC link will typically not set the reserve risk until transfers of around 800 MW[29] in the NI and 600 MW in the SI. Based on the study period, higher transfers could lead to an increase in WIBR, although the results for reserves indicate the opposite.
Time of year (month)	Moderate to strong association with frequency of high price spread events in both islands.	During the study period, time of year is primarily associated with seasonality of inflows, i.e. wet and dry years. Inflows will continue to impact significantly on flows on the grid for decades to come.
Trading Period	Frequency of high price spread events show a very distinct pattern over the day, peaking with the morning peak in periods 15 and 16, falling across the day then falling sharply in period 43.	This pattern appears to be a function of demand, capacity margin (offers less demand), circuit outages and generation outages, plus possible dynamic effects associated with experience that traders gain during the day. The association with WIBR is unlikely

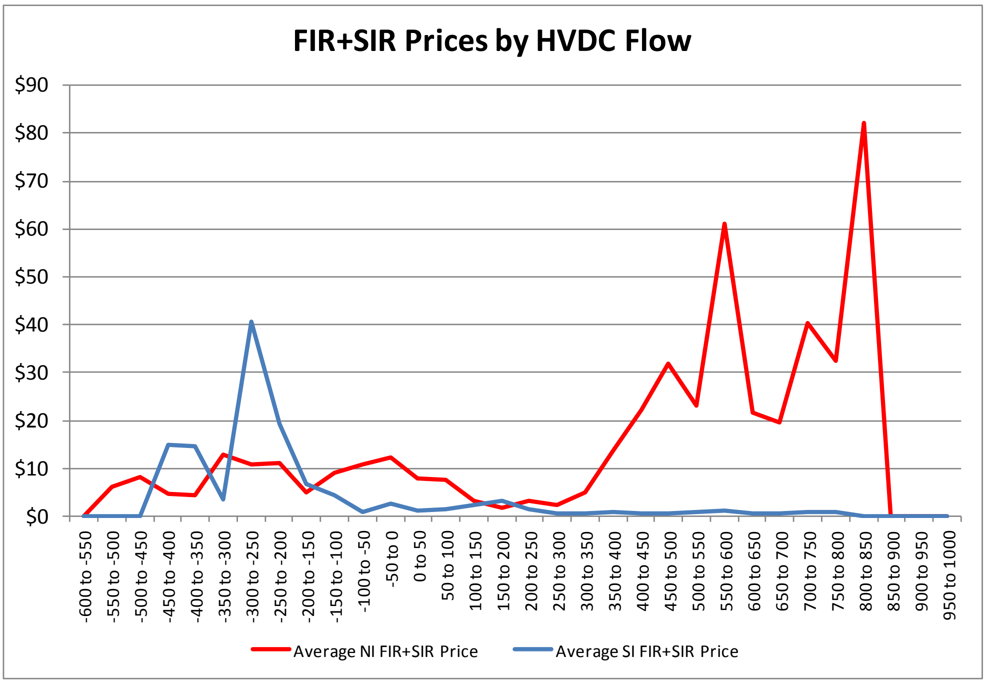
[29] We assume the ability of Pole 2 + Pole 3 to cover the reserve risk of the HVDC link is around 500 MW. In this case the HVDC link would need to transfer 500 MW plus the output of the largest generator in the receiving island before it would set the risk, so 800 – 900 MW in the NI and 620 MW in the SI.

NI Driver	Impact Assessment	Outlook
		to change in future unless steps are taken to, for example, increase the rate at which experience is gained across the day.
Grid Overload	Grid overload is associated with the frequency of high price spread events to a moderate extent in both islands.	Grid upgrades completed after 2011 are likely to reduce the frequency of high price spread events.
Offers Change Maximum Offer Price	Not associated with high price spreads	

The results that indicate that WIBR might reduce in future are those for reserves prices and grid overload. The latter is an intuitive result: if grid capacity is increased then one would expect less WIBR, other things being equal.

The outlook for reserves prices is more complex due to the interaction between reserves and HVDC transfers. These two drivers are related in that higher transfers on the HVDC link have required large amounts of reserves over most of the study period, leading to higher reserves prices, as shown in Figure 36, to cover the reserve risk while Pole 1 has operated at low capacity (nil for southward transfers capacity). On the one hand, Pole 3 will reduce the need for reserves, which is likely to reduce reserves prices. But on the other hand, it will allow higher HVDC transfers, which the results of the study show are associated with a higher frequency of high price spread events. It is not clear from the data whether high price spread events were purely a result of high HVDC transfers or whether they resulted from interactions in the provision of reserve and generation (to cover reserve risk associated with HVDC transfers) which modified dispatch and power flows in ways which led to constraints. Adding to the complexity of the overall picture for HVDC flows and reserves, is the strong correlation between HVDC flow and time of day (refer to Figure 27 and Table 6).

Figure 36 – Reserves Prices by HVDC Flow



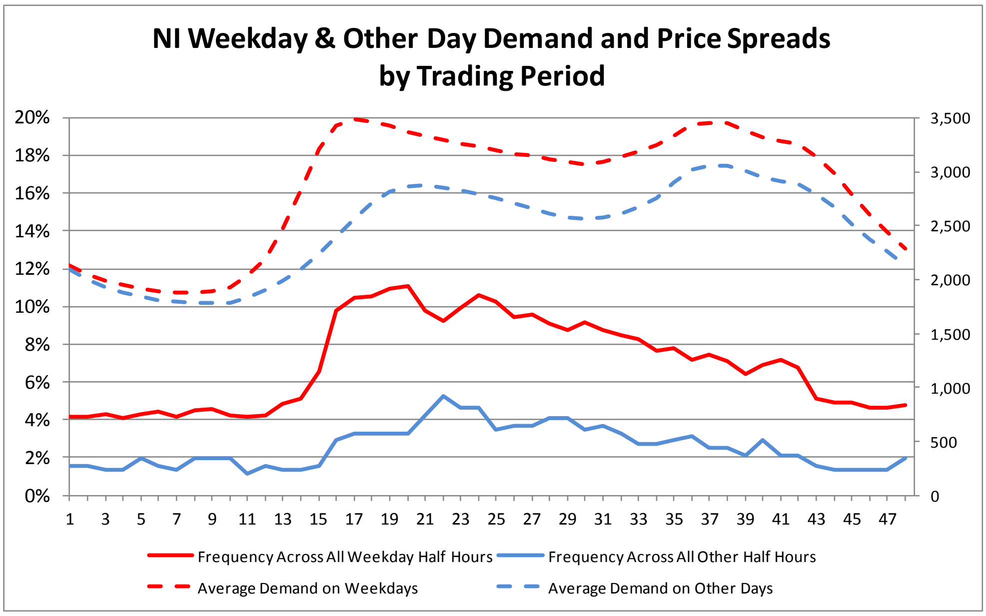
6.1 Dynamic Effects Associated with Time of Day

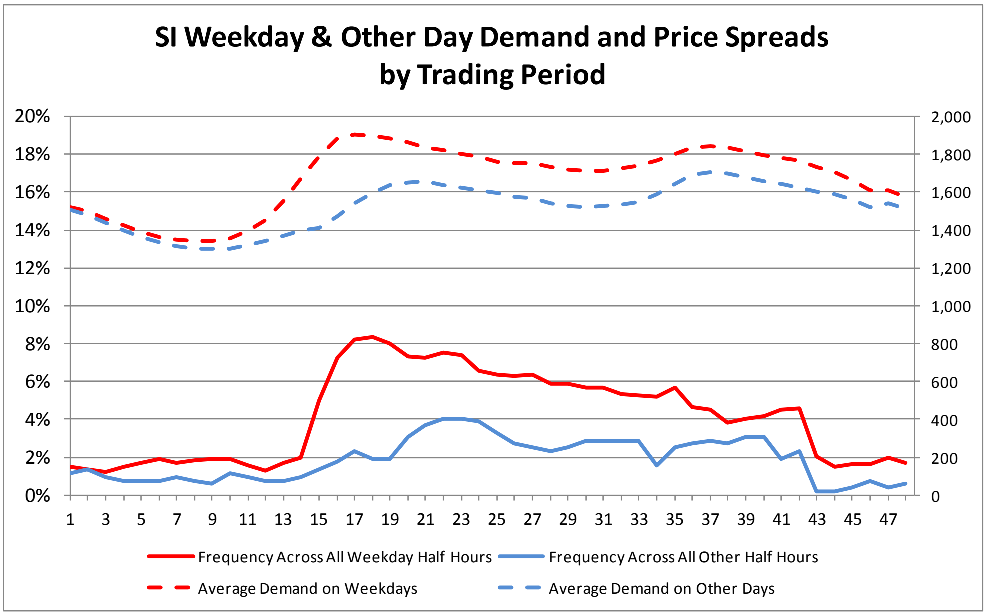
The graphs in sections 5.4, starting at Figure 26 and running through to Figure 35, help to explain why the frequency of high price spread events peaks when the change in demand is around zero: the rate of change of demand slows as demand reaches its morning peak, which is also when the frequency of high price spread events peaks.

The distinctive pattern of frequency across the day suggests a more dynamic process is involved in the creation of high prices spread events than is captured by looking at each WIBR driver on its own. We postulate that leading into the morning peak, a number of factors create uncertainty for market participants including rapidly increasing demand, along with a rapidly narrowing capacity margin, commencement of the new day's transmission outages, and in the SI, generators beginning planned outages. Early in the morning, there may be greater uncertainty as to where demand will peak than later when the demand trend may be clearer. As the day progresses, it is entirely possible that generators are able to fine tune their offers to reduce the occurrence of any constraints that appear leading up to and in the morning peak, which would explain why the frequency of high price spread events consistently falls off during the day.

Figure 37 shows the frequency of high price spread events over all trading periods in weekdays and other days, plus the average demand for weekdays and other days. The pattern observed in Figure 26 is again evident, although the graphs show that high price spread events are significantly more likely on weekdays. Furthermore, the peak in frequency moves to the right as the peak in demand moves to the right for other days, which is an indication that the characteristic morning peak in frequency is strongly associated with demand.

Figure 37 - Weekday & Other Day Demand and Price Spreads by Trading Period

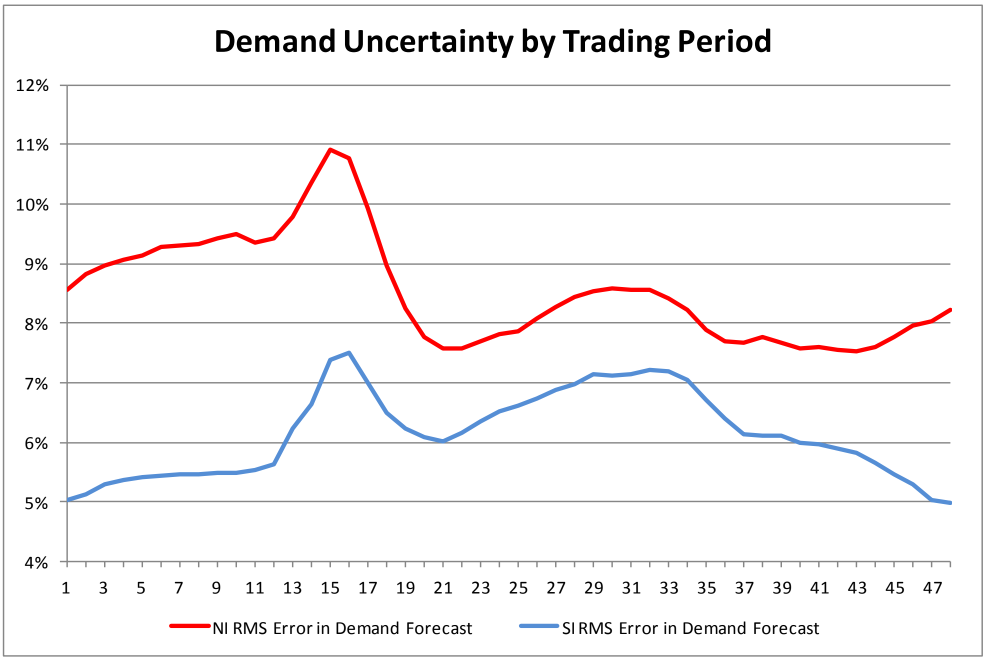




To investigate the role of uncertainty further, a simple forecast of demand was created for all periods from 8-Jun-07 to 31-Dec-11 in which, for each island, the demand from the same trading period in the previous week was used as the forecast for the period demand. The RMS error is shown in Figure 38 below, after removing a handful of outliers, and in both islands it can be seen to peak in periods 13 through 17, and the highest two values in both islands are in periods 15 and 16, which is also when the frequency of high price spread events rises sharply towards the peak for the day.

The RMS error in the simple demand forecasts is lower for the rest of the day, including during the evening peak in demand.

Figure 38 – Demand Uncertainty by Trading Period



The implication of this is quite clear: uncertainty around demand in the morning peak is a key causative factor for high price spread events. The corollary of this is also clear: reduce uncertainty in demand forecasts and you reduce WIBR.

7 Conclusions

Sixteen WIBR drivers were initially anlaysed across 7,307 periods when high price spread events occurred in one or other island. With the exception of only two drivers, all of these drivers showed some degree of positive association with either the frequency or magnitude of these events.

Projecting forward, the only drivers that might change significantly in future are grid capacity (via upgrades of the AC grid), reserves prices and HVDC transfers (once Pole 3 is commissioned). Of these three drivers, we can conclude:

- grid upgrades are likely to reduce WIBR (or at least until demand growth or the building of new generation 'uses up' the new capacity);
- the commissioning of Pole 3 may or may not change WIBR, depending on exactly how reserves and HVDC transfers interact to make constraints more or less likely.

The degree to which grid capacity (or lack of it) associates with WIBR is low to moderate compared to the other WIBR drivers, so the impact of AC grid upgrades on WIBR may in fact be difficult to measure against the general background of WIBR.

Over the majority of the study period, the HVDC link was the main bottleneck for inter-island transfers, but this will no longer be the case once Pole 3 is commissioned. To get large transfers across the link, power must first be gotten to one or other end of the link, creating the potential for bottlenecks in the AC grid to become more obvious, and to create WIBR as a result.

Where these bottlenecks create INPDs between Otahuhu and Benmore, inter-island FTRs and basis swaps[30] will be available to hedge this risk. But where they create INPDs between one of the two FTR hubs and nodes within the same island (for example due to the SWE) then increases in HVDC flow will increase WIBR. One possible outcome is that HVDC transfers will be less strongly associated with high price spread events up to a point, but then more strongly associated at higher flows. A complicating factor is that HVDC flows exhibit a strong pattern across the day, so that the apparent association between HVDC flows and frequency of high price spread events may actually be more to do with time of day than with the HVDC flow itself.

As to the other drivers that are positively associated with high price spread events, some will increase (demand, for example, is forecast to continue growing) and the rest are not likely to move in a direction that will reduce WIBR (generation outages, circuit outages, capacity margin, wet and dry years).

Taken overall, there are some drivers that will tend to reduce WIBR in the foreseeable future, but on balance there are a greater number of stronger drivers that will tend to

[30] A basis swap is a hedge against INPD formed by buying and selling two CFDs or futures contracts for the same period but at different nodes. For example, the payout on a basis swap consisting of a futures contract bought at Otahuhu and a futures contract sold at Benmore increases as difference between the spot price at Otahuhu and the spot price at Benmore increases.

keep WIBR the same or to increase it in future. We conclude that WIBR will continue at similar levels into the foreseeable future, although the associations are likely to change over time as grid upgrades are completed.

However, there is a distinct pattern of the frequency of high price spread events across the day, which is linked to uncertainty in demand combined with the time of day at which circuit and generation outages commence. A potentially aggravating factor, which has not been considered in this study, is the role of wind generation, which adds to the uncertainty around demand forecasts in future. The frequency of events peaks just after the morning demand peak, which is a function of uncertainty around demand at this time of day, in particular. It may be possible to reduce WIBR by reducing this uncertainty through improvements in the market's forecasting processes. Or even, if the benefits of reducing WIBR outweigh the costs of changing the pricing process, by changing the way that prices are formed.

7.1 Further Work

The dynamic effects associated with the morning peak raise a number of questions around the basic functioning of the market. For example:

- Are transmission and generation outages scheduled to start at the best times? Could start times be staggered?
- Do traders have all of the information they need with respect to the uncertainty surrounding demand, capacity margin, and the impact of outages?
- Could the various schedules of forecast prices be improved?
- Does the way that prices are formed give the best trade-off between transparency and the ability to manage basis risk?
- Is the way that SPD models transmission constraints the best trade-off between managing the grid within limits and signaling the likelihood of a SWE to the market?
- Does two hour gate closure limit the ability of traders to respond to situations in an optimal fashion?

We note the Authority is already pursuing work on improving price formation, which will likely extend to some of the items listed above. We recommend that the scope of this work be reviewed to determine whether or not it should include a measureable reduction of WIBR as an explicit goal.

A key advantage of making improvements in demand forecasting and the pricing process, as opposed to adding hedging instruments to the market, is that it should be possible to observe changes in the frequency of high price spread events across the day if the improvements make a significant difference to the ability of traders to anticipate and manage circumstances which could create WIBR. It may also prove to be more cost-effective to improve existing processes than to introduce new hedges, for example, by reducing the number of new hedges that are required.

In addition, with the introduction of new pricing schedules from the end of June 2012, it is recommended that the impact of these new schedules on WIBR be assessed as to their impact on the frequency of high price spread events across the day. We understand that the accuracy of demand bids has not been monitored for a number of years, which is likely to have reduced the ability of the PDS to predict demand, especially leading into the morning peak, so the introduction of two new schedules, assuming they forecast demand better than the PDS, may have a measureable impact on WIBR across the day.

8 Appendix 1 – Nodes and Clusters

The following nodes were used in the study. Each node represents all the nodes of the same voltage at a substation.

NI Voltage Nodes

ALB033	HAY110	MER033	PEN033	TNG050
ALB110	HAY220	MGM033	PEN110	TRK011
ARA220	HEN033	MHO033	PEN220	TRK220
ARI110	HEN220	MLG011	PNI033	TUI011
ATI220	HEP033	MLG033	PPI220	TUI110
BOB033	HIN033	MNG033	PRM033	TWC220
BOB110	HLY033	MNG110	RDF033	TWH033
BPE033	HLY220	MNI011	RDF220	UHT033
BPE050	HTI033	MPE033	ROS110	WAI011
BPE220	HUI033	MST033	ROS220	WDV011
BRB033	HWA033	MTI220	ROT011	WDV110
BRK033	HWA110	MTM033	ROT033	WEL033
CBG011	KAW011	MTN033	ROT110	WGN033
CPK011	KAW110	MTO033	RPO220	WHI011
CPK033	KAW220	MTR033	SFD033	WHI220
CST033	KEN033	NAP220	SFD220	WHU033
DAR011	KIN011	NPK033	SVL033	WIL033
DVK011	KIN033	NPL033	SWN220	WIR033
EDG033	KMO033	NPL110	TAK033	WKM220
FHL033	KOE033	NPL220	TGA011	WKO033
GFD033	KPA110	OHK220	TGA033	WPA220
GIS050	KPO110	OKI011	TKH011	WRA011
GLN033	KPU066	OKI220	TKR033	WRK033
GYT033	KTA033	OKN011	TKU033	WRK220
HAM011	KWA011	ONG033	TKU220	WTU033
HAM033	LFD110	OPK033	TMI033	WVY011
HAM050	LTN033	OTA110	TMN050	WWD110
HAM220	MAT110	OTA220	TMU011	
HAY011	MDN110	OWH011	TMU110	
HAY033	MDN220	PAK033	TNG011	

SI Voltage Nodes

ABY011	EDN033	OTI011
ADD011	FKN033	PAL033
ADD066	GOR033	PAP011
APS011	GYM066	PAP066
ARG110	HKK066	RFN110
ASB033	HOR033	ROX110
ASB066	HOR066	ROX220
ASY011	HWB033	SBK033
ATU110	HWB220	SDN033
AVI220	INV033	SPN033
BAL033	INV220	SPN066
BDE011	ISL033	STK033
BEN110	ISL066	STK220
BEN220	ISL220	STU011
BLN033	KAI011	TIM011
BPD110	KIK011	TKA011
BPT110	KKA033	TKA033
BRY011	KUM066	TKB220
BRY066	MAN220	TMK033
BWK110	MCH011	TWI220
CLH011	MLN066	TWZ033
CML033	MOT011	WPR033
COB066	MPI066	WPR066
COL011	NMA033	WPT011
COL066	NSY033	WTK011
CUL033	OAM033	WTK033
CYD033	OHA220	WTK220
CYD220	OHB220	
DOB033	OHC220	
DOB066	ORO110	

8.1 Cluster Listings

In the cluster listings below, the following voltages codes are used:
0 = 011
1 = 110
2 = 220
3 = 033
5 = 050
6 = 066

For example, OTA2 refers to the two 220 kV nodes at Otahuhu, OTA2201 and OTA2202.

The clusters listed below are based on half hourly prices over the period 1-Jan-07 to 31-Dec-11 inclusive. If using these clusters when developing or reviewing hedging strategy, then note that cluster membership is likely to change over time. It is also important to work with the clusters that are relevant to your organisation's risk preferences.

8.2 Clusters Formed with Correlation = 0.7

NI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0.7	ALB1	BPE2	ARA2	ARA2	ALB1	CPK0	ARA2	ALB1	FHL3	ARA2	ARI1	DVK0	ARI1	MGM3	WPW3	WIL3	TNG5	RDF3	LFD1
	ALB3	BPE3	ATI2	ATI2	ALB3	CPK3	MTI2	ALB3	GIS5	ATI2	HTI3	WDV0	HTI3	MST3					
	BOB1	BPE5	EDG3	EDG3	BRK3	GFD3	NAP2	BRK3	TUI0	TKU2	KIN0	WDV1	TNG0						
	BOB3	BRK3	KAW0	KAW0	CST3	GYT3	OKI0	CST3	TUI1	TKU3	KIN3								
	BRB3	CBG0	KAW1	KAW1	HAM2	HAY0	OKI2	HAM2	WRA0										
	DAR0	CST3	KAW2	KAW2	HAM3	HAY1	PPI2	HAM3											
	GLN3	HAM0	KMO3	KMO3	HAM5	HAY3	RDF2	HAM5											
	HEN2	HAM2	MAT1	MAT1	HUI3	KWA0	RPO2	WGN3											
	HEN3	HAM3	MTI2	MTI2	HWA1	MLG0	TKU2												
	HEP3	HAM5	MTM3	MTM3	HWA3	MLG3	TKU3												
	HLY2	HAY2	NAP2	NAP2	KPA1	MST3													
	HLY3	HIN3	OHK2	OHK2	MNI0	PNI3													
	KEN3	HUI3	OKI0	OKI0	NPL1	PRM3													
	KOE3	HWA1	OKI2	OKI2	NPL2	TKR3													
	KTA3	HWA3	OWH0	OWH0	NPL3	UHT3													
	MDN1	KPA1	PPI2	PPI2	OPK3	WWD1													
	MDN2	KPO1	RDF2	RDF2	SFD2														
	MER3	KPU6	ROT0	ROT0	SFD3														
	MNG1	LTN3	ROT1	ROT1	TMN5														
	MNG3	MHO3	ROT3	WHI0	WVY0														
	MPE3	MNI0	TGA0	WHI2															
	MTO3	MTN3	TGA3	WKM2															
	MTR3	NPL1	TKH0	WPA2															
	NPK3	NPL2	TMI3	WRK2															
	OKN0	NPL3	TRK0	WRK3															
	ONG3	OPK3	TRK2	WTU3															
	OTA1	SFD2	WAI0																
	OTA2	SFD3	WHI0																
	PAK3	TMN5	WHI2																
	PEN1	TMU0	WKM2																
	PEN2	TMU1	WRK2																
	PEN3	TWC2	WRK3																
	ROS1	WGN3	WTU3																
	ROS2	WHU3																	
	SVL3	WKO3																	
	SWN2	WVY0																	
	TAK3																		
	TMN5																		
	TWH3																		
	WEL3																		
	WIR3																		

SI	1	2	3	4	5	6
0.7	ABY0	BPD1	HWB2	BAL3	BPD1	BPD1
	ADD0	CYD2	HWB3	BDE0	COB6	KUM6
	ADD6	CYD3	INV2	BWK1	MOT0	MLN6
	APS0	HWB2	INV3	EDN3	MPI6	SPN6
	ARG1	HWB3	MAN2	GOR3		
	ASB3	INV2	NMA3			
	ASB6	INV3	PAL3			
	ASY0	MAN2	ROX1			
	ATU1	NMA3	SDN3			
	AVI2	PAL3	TWI2			
	BEN1	ROX2				
	BEN2	SDN3				
	BLN3	TWI2				
	BPD1					
	BPT1					
	BRY0					
	BRY6					
	CLH0					
	CML3					
	COL0					
	COL6					
	CUL3					
	DOB3					
	DOB6					
	FKN3					
	GYM6					
	HKK6					
	HOR3					
	HOR6					
	ISL2					
	ISL3					
	ISL6					
	KAI0					
	KIK0					
	KKA3					
	MCH0					
	MLN6					
	NSY3					
	OAM3					
	OHA2					
	OHB2					
	OHC2					
	ORO1					
	OTI0					
	PAP0					
	PAP6					
	RFN1					
	SBK3					
	SPN3					
	SPN6					
	STK2					
	STK3					
	STU0					
	TIM0					
	TKA0					
	TKA3					
	TKB2					
	TMK3					
	TWZ3					
	WPR3					
	WPR6					
	WPT0					
	WTK0					
	WTK2					
	WTK3					

8.3 Clusters Formed with Correlation = 0.8

NI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
0.8	ALB1	BPE2	ARA2	ATI2	ALB1	ARA2	CPK0	ARA2	ARA2	GFD3	BOB1	FHL3	ALB1	DVK0	RPO2	HTI3	ARI1	RPO2	ARI1	WPW3	WIL3	TNG5	RDF3	MGM3	LFD1
	ALB3	BPE3	ATI2	EDG3	ALB3	ATI2	CPK3	ATI2	ATI2	GYT3	BOB3	GIS5	ALB3	WDV0	TKU2	NPK3	KIN0	TNG0	HTI3						
	BOB1	BPE5	EDG3	KAW0	CST3	EDG3	GFD3	EDG3	EDG3	HAY0	HLY3	TUI0	BOB1	WDV1	TKU3	ONG3	KIN3								
	BOB3	BRK3	KAW0	KAW1	HAM2	KAW0	GYT3	KAW0	KAW0	HAY1	MER3	TUI1	BOB3												
	BRB3	CBG0	KAW1	KAW2	HAM3	KAW1	HAY0	KAW1	KAW1	HAY3	MTR3	WRA0	TMN5												
	DAR0	CST3	KAW2	KMO3	HAM5	KAW2	HAY1	KAW2	KAW2	MLG0	NPK3														
	GLN3	HAM0	KMO3	MAT1	HUI3	KMO3	HAY3	KMO3	KMO3	MLG3	OKN0														
	HEN2	HAM2	MAT1	MTM3	HWA1	MAT1	KWA0	MAT1	MAT1	MST3	WIR3														
	HEN3	HAM3	MTI2	NAP2	HWA3	MTI2	MLG0	MTI2	NAP2	PNI3															
	HEP3	HAM5	NAP2	OHK2	KPA1	NAP2	MLG3	NAP2	OHK2	PRM3															
	HLY2	HAY2	OHK2	OWH0	MNI0	OHK2	PNI3	OHK2	OWH0	TKR3															
	HLY3	HIN3	OKI0	ROT0	NPL1	OKI0	PRM3	TKH0	TKH0	UHT3															
	KEN3	HUI3	OKI2	ROT1	NPL2	OKI2	TKR3	TRK2	TRK2																
	KOE3	HWA1	PPI2	ROT3	NPL3	PPI2	UHT3	WAI0	WAI0																
	KTA3	HWA3	RDF2	TGA0	OPK3	RDF2	WWD1																		
	MDN1	KPA1	TKH0	TGA3	SFD2	TKU2																			
	MDN2	KPO1	WHI0	TKH0	SFD3	TKU3																			
	MER3	KPU6	WHI2	TMI3	TMN5																				
	MNG1	LTN3	WKM2	TRK0																					
	MNG3	MHO3	WPA2	TRK2																					
	MPE3	MNI0	WRK2	WAI0																					
	MTO3	MTN3	WRK3																						
	MTR3	NPL1	WTU3																						
	NPK3	NPL2																							
	OKN0	NPL3																							
	OTA1	OPK3																							
	OTA2	SFD2																							
	PAK3	SFD3																							
	PEN1	TMN5																							
	PEN2	TMU0																							
	PEN3	TMU1																							
	ROS1	TWC2																							
	ROS2	WGN3																							
	SVL3	WHU3																							
	SWN2	WKO3																							
	TAK3	WVY0																							
	TWH3																								
	WEL3																								

SI	1	2	3	4	5	6	7
0.8	ABY0	ABY0	BPD1	HWB2	BAL3	BPD1	BPD1
	ADD0	ADD0	CYD2	HWB3	BDE0	COB6	KUM6
	ADD6	ADD6	CYD3	INV2	BWK1	MOT0	MLN6
	APS0	APS0	HWB2	INV3	EDN3	MPI6	SPN6
	ARG1	ARG1	HWB3	MAN2	GOR3		
	ASB3	ASB3	INV2	NMA3			
	ASB6	ASB6	INV3	PAL3			
	ASY0	ASY0	MAN2	ROX1			
	ATU1	ATU1	NMA3	SDN3			
	AVI2	BEN1	PAL3	TWI2			
	BEN1	BEN2	ROX2				
	BEN2	BLN3	SDN3				
	BLN3	BRY0	TWI2				
	BPD1	BRY6					
	BPT1	CLH0					
	BRY0	CML3					
	BRY6	COL0					
	CLH0	COL6					
	CML3	CUL3					
	COL0	DOB3					
	COL6	DOB6					
	CUL3	FKN3					
	DOB3	GYM6					
	DOB6	HKK6					
	FKN3	HOR3					
	GYM6	HOR6					
	HKK6	ISL2					
	HOR3	ISL3					
	HOR6	ISL6					
	ISL2	KAI0					
	ISL3	KIK0					
	ISL6	KKA3					
	KAI0	MCH0					
	KIK0	MLN6					
	KKA3	NSY3					
	MCH0	OHA2					
	MLN6	OHB2					
	OAM3	OHC2					
	OHA2	ORO1					
	OHB2	OTI0					
	OHC2	PAP0					
	ORO1	PAP6					
	OTI0	RFN1					
	PAP0	SBK3					
	PAP6	SPN3					
	RFN1	SPN6					
	SBK3	STK2					
	SPN3	STK3					
	SPN6	TIM0					
	STK2	TKB2					
	STK3	TMK3					
	STU0	TWZ3					
	TIM0	WPR3					
	TKA0	WPR6					
	TKA3	WPT0					
	TKB2						
	TMK3						
	TWZ3						
	WPR3						
	WPR6						
	WPT0						
	WTK0						
	WTK2						
	WTK3						

8.4 Clusters Formed with Correlation = 0.9

NI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.9	ALB1	CBG0	BPE2	ARA2	BPE2	EDG3	BPE2	CPK0	ARA2	ATI2	ARA2	BPE2	BOB1	FHL3	MTI2
	ALB3	CST3	BPE3	ATI2	BPE3	KAW0	BPE3	CPK3	ATI2	EDG3	ATI2	BPE3	BOB3	GIS5	TKU2
	BRB3	HAM0	BPE5	EDG3	BPE5	KAW1	BPE5	GFD3	KAW1	KAW0	EDG3	BPE5	HLY3	TUI0	TKU3
	DAR0	HAM2	BRK3	KAW0	BRK3	KAW2	BRK3	HAY0	KMO3	KAW1	KAW0	HAY2	MER3	TUI1	WKM2
	GLN3	HAM3	CST3	KAW1	CST3	KMO3	CST3	HAY1	MAT1	KAW2	KAW1	LTN3	WIR3	WRA0	WPA2
	HEN2	HAM5	HUI3	KAW2	HAY2	MAT1	HUI3	HAY3	MTI2	KMO3	KAW2	TWC2			
	HEN3	HIN3	HWA1	KMO3	HUI3	MTM3	KPA1	KWA0	NAP2	MAT1	KMO3	WVY0			
	HEP3	HUI3	HWA3	MAT1	HWA1	NAP2	LTN3	MLG0	OHK2	NAP2	MAT1				
	HLY2	KPA1	KPA1	NAP2	HWA3	OWH0	MHO3	MLG3	OKI0	TKH0	NAP2				
	HLY3	KPO1	LTN3	OHK2	KPA1	ROT0	MNI0	PNI3	OKI2	TRK2	WAI0				
	KEN3	KPU6	MHO3	OKI0	LTN3	ROT1	MTN3	PRM3	PPI2	WAI0					
	KOE3	MNI0	MNI0	OKI2	MNI0	ROT3	NPL1	TKR3	RDF2						
	KTA3	NPL1	MTN3	PPI2	NPL1	TGA0	NPL2	UHT3	WKM2						
	MDN1	NPL2	NPL1	RDF2	NPL2	TGA3	NPL3	WWD1	WPA2						
	MDN2	NPL3	NPL2	WHI0	NPL3	TKH0	OPK3								
	MNG1	OPK3	NPL3	WHI2	OPK3	TMI3	SFD2								
	MNG3	SFD2	OPK3	WRK2	SFD2	TRK0	SFD3								
	MPE3	SFD3	SFD2	WRK3	SFD3	TRK2	TWC2								
	MTO3	TMU0	SFD3	WTU3	TWC2	WAI0	WGN3								
	MTR3	TMU1	TWC2												
	OTA1	WHU3													
	OTA2	WKO3													
	PAK3														
	PEN1														
	PEN2														
	PEN3														
	ROS1														
	ROS2														
	SVL3														
	SWN2														
	TAK3														
	TWH3														
	WEL3														

NI	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
0.9	ALB1	DVK0	ARI1	MTR3	ARI1	GYT3	HTI3	LFD1	MGM3	RDF3	RPO2	TNG0	TNG5	WIL3	WPW3
	ALB3	WDV0	KIN0	NPK3	HTI3	MST3	ONG3								
	BRB3	WDV1	KIN3	OKN0											
	TMN5														

SI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0.9	ADD0	ADD0	AVI2	ADD0	ARG1	BPD1	ASB3	BAL3	DOB3	ABY0	COB6	HWB3	BAL3	TKA0	KUM6
	ADD6	ADD6	BEN1	ADD6	ATU1	CYD2	ASB6	BDE0	DOB6	MLN6	MOT0	PAL3	BWK1	TKA3	
	APS0	APS0	BEN2	ASB3	BLN3	CYD3	BEN1	EDN3	GYM6	SPN6	MPI6	ROX1			
	ASB3	ASY0	BPD1	ASB6	CUL3	HWB2	BEN2	GOR3	HKK6						
	ASB6	BRY0	BPT1	ASY0	KIK0	INV2	CML3								
	ASY0	BRY6	CML3	BEN1	KKA3	INV3	FKN3								
	BRY0	CLH0	FKN3	BEN2	MCH0	MAN2	OHA2								
	BRY6	HOR3	OAM3	CML3	ORO1	NMA3	OHB2								
	CLH0	HOR6	OHA2	FKN3	RFN1	ROX2	OHC2								
	CML3	ISL2	OHB2	OHA2	STK2	SDN3	TKB2								
	COL0	ISL3	OHC2	OHC2	STK3	TWI2	TWZ3								
	COL6	ISL6	STU0	TKB2	WPT0										
	FKN3	KAI0	TKB2	TWZ3											
	HOR3	MLN6	TWZ3												
	HOR6	NSY3	WTK0												
	ISL2	OTI0	WTK2												
	ISL3	PAP0	WTK3												
	ISL6	PAP6													
	KAI0	SBK3													
	MLN6	SPN3													
	OHA2	SPN6													
	OTI0	TIM0													
	PAP0	TMK3													
	PAP6	WPR3													
	SBK3	WPR6													
	SPN3														
	SPN6														
	TIM0														
	TKB2														
	TMK3														
	TWZ3														
	WPR3														
	WPR6														

8.5 Clusters Formed with Correlation = 0.95

NI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
0.95	ALB1	ALB1	ARA2	EDG3	BPE2	CPK0	BRK3	CBG0	ATI2	BPE2	BPE2	GIS5	ALB1	MHO3	BOB1	MTR3	WDV0	TKU2	HTI3	GYT3	KIN0	WPW3	WIL3	TNG5	TNG0	RPO2	RDF3	MGM3	LFD1	ARI1
	ALB3	ALB3	ATI2	KAW0	BPE3	CPK3	CST3	HAM0	EDG3	BPE3	BPE3	TUI0	ALB3	MTN3	BOB3	NPK3	WDV1	TKU3	ONG3	MST3	KIN3									
	BRB3	BRB3	KAW1	KAW1	BPE5	GFD3	HUI3	HAM2	KAW0	BPE5	BPE5	TUI1	BRB3	WGN3	MER3	OKN0														
	DAR0	DAR0	KMO3	KAW2	BRK3	HAY0	HWA1	HAM3	KAW1	LTN3	BRK3	WRA0	TMN5	WVY0	WIR3															
	GLN3	GLN3	MAT1	KMO3	CST3	HAY1	HWA3	HAM5	KAW2	WGN3	MTN3																			
	HEN2	HEN2	MTI2	MAT1	HUI3	HAY3	KPA1	HIN3	KMO3	WVY0																				
	HEN3	HEN3	NAP2	MTM3	KPA1	KWA0	MNI0	KPO1	MAT1																					
	HEP3	HEP3	OHK2	NAP2	LTN3	MLG0	NPL1	KPU6	NAP2																					
	HLY2	HLY3	OKI0	OWH0	MNI0	MLG3	NPL2	TMU0	TKH0																					
	HLY3	KEN3	OKI2	ROT0	NPL1	PNI3	NPL3	TMU1	WAI0																					
	KEN3	KOE3	PPI2	ROT1	NPL2	PRM3	OPK3	WHU3																						
	MDN1	KTA3	RDF2	ROT3	NPL3	TKR3	SFD2	WKO3																						
	MDN2	MDN1	WHI0	TGA0	OPK3	UHT3	SFD3																							
	MNG1	MDN2	WHI2	TGA3	SFD2	WWD1																								
	MNG3	MNG1	WKM2	TKH0	SFD3																									
	MPE3	MNG3	WPA2	TMI3	TWC2																									
	MTO3	MPE3	WRK2	TRK0																										
	OTA1	MTO3	WRK3	TRK2																										
	OTA2	OTA1	WTU3	WAI0																										
	PAK3	OTA2																												
	PEN1	PAK3																												
	PEN2	PEN1																												
	PEN3	PEN2																												
	ROS1	PEN3																												
	ROS2	ROS1																												
	SVL3	ROS2																												
	SWN2	SVL3																												
	TAK3	SWN2																												
	TWH3	TAK3																												
	WEL3	WEL3																												

SI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
0.95	ADD0	ARG1	CYD2	AVI2	HWB2	BPD1	BAL3	ASB6	DOB3	BPD1	ABY0	CLH0	BPD1	APS0	COB6	COL0	TKA0	BWK1	HWB3	KUM6	MOT0	NSY3	PAL3
	ADD6	ATU1	CYD3	BEN1	INV2	BPT1	BDE0	CML3	DOB6	OHA2	MLN6	HOR3	OAM3	OTI0	MPI6	COL6	TKA3						
	ASB3	BLN3	INV2	BEN2	INV3	STU0	EDN3	FKN3	GYM6	TKB2	SPN6	HOR6	STU0										
	ASB6	CUL3	INV3	BPD1	MAN2	WTK0	GOR3	TKB2	HKK6	TWZ3													
	ASY0	KIK0	MAN2	OHB2	NMA3	WTK2																	
	BRY0	KKA3	NMA3	OHC2	SDN3	WTK3																	
	BRY6	MCH0	ROX2	STU0	TWI2																		
	HOR3	ORO1	TWI2																				
	HOR6	RFN1																					
	ISL2	STK2																					
	ISL3	STK3																					
	ISL6	WPT0																					
	KAI0																						
	MLN6																						
	PAP0																						
	PAP6																						
	SBK3																						
	SPN3																						
	SPN6																						
	TIM0																						
	TMK3																						
	WPR3																						
	WPR6																						

8.6 Clusters Formed with Correlation = 0.99

NI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
0.99	ALB1	ALB1	ALB1	ARA2	GFD3	KMO3	ALB1	CBG0	CPK0	CST3	BPE2	HUI3	KPA1	KAW1	EDG3	TUI0	MTI2	HLY2	ATI2	BOB1	KAW0	HLY3	HAM2
	ALB3	ALB3	ALB3	NAP2	HAY0	MTM3	ALB3	HAM0	CPK3	HUI3	BPE3	MNI0	MNI0	MAT1	KAW2	TUI1	WKM2	HLY3	KMO3	BOB3	KAW1	KOE3	HAM3
	BRB3	BRB3	GLN3	OKI0	HAY1	OWH0	BRB3	HIN3	KWA0	MNI0	BPE5	OPK3	OPK3	TKH0	KMO3	WRA0	WPA2	TWH3	OHK2	MER3	MAT1	KTA3	HAM5
	DAR0	DAR0	HEN2	OKI2	HAY3	ROT0	DAR0	KPO1	PNI3	NPL1	LTN3	SFD2	SFD2	WAI0	TRK2								
	GLN3	GLN3	HEN3	PPI2	MLG0	ROT1	HEN2	KPU6	PRM3	NPL2	TWC2	SFD3	SFD3										
	HEN2	HEN2	HEP3	RDF2	MLG3	ROT3	HEN3	TMU0	TKR3	NPL3													
	HEN3	HEN3	HLY3	WHI0	PNI3	TGA0	HEP3	TMU1	WWD1														
	HEP3	HEP3	MNG1	WHI2	PRM3	TGA3	HLY3	WHU3															
	HLY3	HLY3	MNG3	WRK2	TKR3	TMI3	KEN3	WKO3															
	KEN3	KEN3	OTA1	WRK3	UHT3	TRK0	KOE3																
	MDN1	MDN1	OTA2	WTU3																			
	MDN2	MDN2	PAK3																				
	MNG1	MNG1	PEN1																				
	MNG3	MNG3	PEN2																				
	MPE3	MPE3	PEN3																				
	MTO3	MTO3	SVL3																				
	OTA1	OTA1	SWN2																				
	OTA2	OTA2	TAK3																				
	PEN2	PEN1																					
	PEN3	PEN2																					
	ROS1	PEN3																					
	ROS2	SVL3																					
	SVL3	SWN2																					
	SWN2	TAK3																					
	TAK3																						
	WEL3																						

24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
MTN3	WDV0	TKU2	MTR3	NPK3	MHO3	KIN0	HWA1	WVY0	WPW3	WIR3	WIL3	TNG5	TNG0	TMN5	RPO2	RDF3	ONG3	MST3	MGM3	LFD1	HTI3	HAY2	GYT3	GIS5
WGN3	WDV1	TKU3	OKN0	OKN0	MTN3	KIN3	HWA3																	

SI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
0.99	ADD0	ARG1	BPD1	ASB3	BEN1	BDE0	CYD2	INV2	OHA2	CML3	COL0	DOB3	HOR3	HWB2	MAN2	BPD1	BPD1	TKA0	ABY0	APS0	ATU1	AVI2	BAL3
	ADD6	CUL3	BPT1	ASB6	BEN2	EDN3	CYD3	INV3	TKB2	FKN3	COL6	DOB6	HOR6	SDN3	NMA3	OAM3	STU0	TKA3					
	ASY0	KIK0	WTK0	TIM0	OHB2	GOR3	ROX2	TWI2	TWZ3														
	BRY0	KKA3	WTK2	TMK3	OHC2																		
	BRY6	MCH0	WTK3																				
	ISL2	ORO1																					
	ISL3	RFN1																					
	ISL6	STK2																					
	KAI0	STK3																					
	MLN6	WPT0																					
	PAP0																						
	PAP6																						
	SBK3																						
	SPN3																						
	SPN6																						
	WPR3																						
	WPR6																						

24	25	26	27	28	29	30	31	32	33	34	35	36	37
BLN3	BWK1	CLH0	COB6	GYM6	HKK6	HWB3	KUM6	MOT0	MPI6	NSY3	OTI0	PAL3	ROX1