Further Analysis of Within-island Basis Risk

Prepared by Energy Link

for

The Electricity Authority

November 2011

Quality Assurance Information

Name	EA WIBR scope Nov-11 FINAL.docx
File reference	E-EA-950
Issue Status	Issue 1, Final
Issue Date	16 November 2011
Client	Energy Link Ltd

Definitions

The following abbreviations and acronyms are used in this report.

Authority	Electricity Authority
ARI	Arapuni
BEN	Benmore
BPE	Bunneythorpe
Code	Electricity Industry Participation Code
em6	The market information system operated by the Energy Market Services division of Transpower
FTR	Financial transmission right
HAY	Haywards
HTI	Hataitai
IIBR	Inter-island basis risk (a.k.a. inter-island LPR)
INPD	Inter-nodal price difference
IRPD	Inter-regional price difference
ISL	Islington
LCE	Losses and constraints excess
LPR	Locational price risk
LPRTG	Locational Price Risk Technical Group
MRJD	Mean reversion jump diffusion
NI	North Island
OTA	Otahuhu
PAP	Papanui
SI	South Island
SPD	Scheduling, Pricing and Dispatch model
STK	Stoke
TKU	Tokaanu
WIBR	Within-island basis risk (a.k.a. within-island LPR or intra-island LPR)
WRK	Wairakei

Contents

1 Introduction

The Electricity Authority (the Authority) amended the Electricity Industry Participation Code (the Code) in October 2011 to allow for the introduction of financial transmission rights (FTRs), and an RFP was issued in August 2011 for the role of FTR Manager. The FTRs will initially hedge the price difference between two key nodes, one in each of the North Island (NI) and South Island (SI), and the FTRs will be funded from the losses and constraints excess (LCE). The FTRs will facilitate financial management of the 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) and has engaged Energy Link to advise the Locational Price Risk Technical Group (LPRTG) of any further analysis that is required to assess the need for LPR hedge instruments on a within-island basis.

The work undertaken and outlined in this report includes:

1. defining WIBR including:
    a. the scope of WIBR;
    b. the overlap between IIBR and WIBR;
    c. the factors that influence materiality of WIBR (e.g. when load, generation or hedge cash flows are influenced by WIBR);
    d. the drivers of WIBR (e.g. demand variations, generation outages, transmission outages, grid loading, offer strategy, grid topology, hydrology and storage);
2. discuss options (including any methodology) for assessing the level of WIBR on an on-going basis;
3. recommend an approach for assessing WIBR on an on-going basis, including an overview of the methodology;
4. detail how the analysis would be carried out and its output.

2 Summary

Basis risk in spot prices arises when a hedging strategy includes two prices that do not correlate perfectly, for example when a generator has retail load at two distinct nodes, or a large consumer hedges at a node which is distant from its load. Nodal prices may correlate well most of the time, but then this correlation can break down particularly when lines or equation constraints bind and price separations occur in simple fashion between two nodes, or in more complex ways due to the spring washer effect that occurs in loops in the grid.

WIBR is created by uncertainty in inter-nodal price differences (INPDs) within an island. INPDs exhibit relatively predictable seasonal changes, but with random noise superimposed, along with infrequent spikes or jumps up or down. It is the jumps, in

[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.

particular, that create INPDs that give the greatest degree of uncertainty and hence basis risk.

The spring washer effect is well known to produce jumps in INPDs that can extend over large areas of the grid and reach very high values, sometimes in the tens of thousands of dollars per MWh.

There are regions of the grid, however, in which WIBR is insignificant, which we will call WIBR regions. For example, the prices at nodes within urban Christchurch are all closely correlated to each other, and hence form a WIBR region. By identifying WIBR regions we can reduce the scope of the WIBR problem to decisions on what instruments, if any, may be required to assist in managing the uncertainty in inter-regional price differences (IRPDs). In the context of WIBR over New Zealand as a whole, some regions are more material than others, primarily those that contain large concentrations of generation, demand or both.

The recommended approach to assessing WIBR on an on-going basis includes a detailed methodology for identifying WIBR regions based on how nodal prices correlate over the period of a month, but going back for at least 60 months. Prices at nodes within a region must correlate each month to at least a defined thresholds value, likely to be at least 0.9 (90% correlation). The actual threshold value will be determined by trial and error during the analysis, taking into account the trade-offs between the number of regions, the computational effort involved in processing more regions, and the materiality of the regions identified.

Any approach to predicting WIBR into the future is likely to refer to the underlying drivers of WIBR, as they change over time. These drivers include line outages, generation outages, demand, generator offers, hydrology (wet and dry years), grid capacity and grid topology (how the grid is connected).

Looking forward, there are a number of approaches that could be taken to predicting WIBR, given the WIBR regions that are finally identified. Single or multiple scenario modelling is potentially very accurate, but would require excessive computational effort. Monte Carlo models appeal from a conceptual point of view and are computationally efficient, but require an effort to predict the values of key parameters as the grid and market change over time which is comparable to scenario modelling.

Reliance on historical patterns of WIBR simply fails to anticipate the impact of changes to the grid and market over time, but does at least set context. Our recommend approach is a form of fundamental analysis which aims to quantify enduring relationships between WIBR drivers and IRPDs.

Price separations over the last five years would be identified and for each trading period concerned the price spread between WIBR regions would be quantified as the standard deviation of the average prices in the regions: the great the spread, the greater the IRPDs. WIBR drivers would be quantified in ach trading period concerned, and the resulting data processed to determine the relationships between the IRPD spreads and the WIBR drivers.

These relationships would then be used to look forward and predict how WIBR would change at the aggregate level, or down to the regional level in some cases, thus informing the LPRTG.

3 The Scope of WIBR

Basis risk (LPR) is defined as unpredictable movements in the price differences between nodes (Electricity_Commission 2010). Basis risk can be broken down in a number of ways, for example by classifying the cause of basis risk as being losses (the impact of marginal losses on nodal prices), or line constraints that bind (either single line constraints or equation constraints that limit the flow on a group of lines), or the impact of reserves when the HVDC link sets the reserve risk in an island (in this case the HVDC link is said to be 'economically constrained by reserves'). When price differences between nodes are greater or less than the price differences expected due to losses, then we say that a 'price separation' has occurred between the nodes concerned.

Basis risk can also be classified as WIBR or IIBR, where IIBR refers to the uncertainty associated with the price differences between nodes that are in different islands, and WIBR refers to the uncertainty associated with price differences between nodes in the same island. The subject of this paper is WIBR and how it can be assessed and how the growth or decline of WIBR can be predicted years into the future.

For our purposes, we are primarily concerned with unpredictable movements in the price difference between nodes in the same island, and particularly price separations, since price separations can and do produce the largest price differences between nodes.

However, WIBR and IIBR overlap in the sense that unpredictable movements in the price differences between certain nodes in an island add to uncertainty around price differences between nodes in different islands. For example, during a dry year there can be price separations on the HVDC link without any price separations occurring within either island. However, the HVDC link might remain unconstrained while a price separation occurs between BPE and HAY: in this case the price separation on the BPE-HAY circuits contributes to both IIBR and WIBR.

Just about any price separation within an island creates uncertainty in the price differences between nodes in the other island, so we can conclude that just about all WIBR contributes to IIBR, whereas there is some IIBR (i.e. price separations on the HVDC link) which do not contribute to WIBR. However, work undertaken earlier this year (Energy_Link 2011) shows that the impact of WIBR on IIBR is largely due to price separations between a limited number of key nodes within the NI.

3.1 The Nature of WIBR

WIBR is created by uncertainty in INPDs within an island. It is instructive to think of INPDs as exhibiting a combination of seasonal patterns with mean reversion, random noise which is normally distributed, and a range of spikes (or jumps) of varying sign (positive or negative), duration, amplitude and probability (Clelow, Strickland et al. 2001).

The seasonal impact on INPDs is often associated with the demand profile within each day and each week, plus changes in demand between climate seasons, and in the ebb and flow of storage in hydro lakes (since hydro inflows exhibit strong seasonal patterns). Random noise can be driven by factors such as noise in demand or inflows, outages, and offer strategies. Seasonal jumps often occur when inflows depart from normal, especially in very wet or very dry years. Short term jumps often occur when one or more line or equation constraints bind, often with aggravating factors such as changes in offers and concurrent outages.

Seasonal changes in INPDs are usually predictable because patterns of demand and inflows are well understood, for example winter tends to have high demand and low inflows in the SI. Noise in INPDs may be less predictable, but also of a magnitude that is not always material given the many other uncertainties involved in buying and selling electricity. But on the other hand, jumps tend to be highly unpredictable, often create large INPDs, and are what most would perceive as WIBR.

A key driver of jumps in INPDs is the spring washer effect which is unique to nodal electricity markets on AC networks, in which a binding line or equation constraint creates price separations over a region of the grid which can extend well beyond the constrained line or lines. For example, a constraint on BPE-TKU in trading period 22 on 21 August 2004 caused a spring washer which gave prices of -$1,311/MWh at TKU, $11,682/MWh at BPE (and higher down toward HAY and back up to HTI), $986/MWh at WRK and around $1,000/MWh at nodes on the East Coast, $0.4/MWh at ARI and $60 - $80/MWh most other places in the NI. This one constraint produced price separations between at least five regions of the grid in the NI. During this trading period the underlying causes of the spring washer were an HVDC link bipole outage and a line outage in the lower NI.

3.2 Regions

There are nodes between which WIBR is immaterial and can be ignored, for example Islington and Papanui in Christchurch are adjacent each other on the grid and exhibit a highly stable and predictable INPD[2]. There are a number of other nodes around ISL which exhibit similarly stable and predictable INPDs with respect to ISL, to PAP and to each other. Assuming that we can identify 'clusters' of nodes within which INPDs are highly stable and predictable, then we can label each of these clusters as a 'WIBR region': then the issue of WIBR becomes one of IRPDs as opposed to INPDs.

On the assumption that the number of WIBR regions in each island is much less than the number of nodes, working with WIBR can be greatly simplified by working with regions rather than nodes. There are currently 160 nodes with prices in the NI and 93 in the SI, but these can be grouped into substations (a substation contains one or more nodes at a range of voltages, e.g. PAP0661 and PAP0111) on the assumption that nodes at the same substation are in the same WIBR region, of which there are 108 in the NI and 67 in the SI.

We can speculate that there may be five to ten WIBR regions in the SI and 20 to 30 in the NI, but the number of regions depends to a certain extent on how much WIBR is allowed within each region when they are defined. Allowing more WIBR within regions will give less regions, and vice versa.

[2] Over the period 1/1/06 to 31/10/11 the INPD exceeded 20% in 20 trading periods, or 0.02% of the time.

3.3 Materiality

WIBR is of concern to anyone buying or selling at spot prices only to the extent that the unpredictability of INPDs creates uncertainty in net cash flows, where two or more cash flows are driven by different nodal prices, e.g. a cash flow at OTA and a cash flow at BEN.

Cash flows may be impacted by INPDs when price differences occur between nodes where an entity has generation and load (gentailer), generation and hedge (generator), load and hedge (consumer), customers and hedge (retailer), or hedge and hedge (intermediary). While nodes or regions with a high concentration of load are readily identifiable, nodes with high levels of hedging are not so easy to identify. Furthermore, hedging strategies can change over time and therefore materiality associated with nodes or regions can also change.

However, while a change in location of hedging may have the effect of shifting WIBR from one party to another, the nodes or regions involved do not change. For example, a generator at node A and consumer at node B may enter into a hedge referencing node A one year, and then node B the next year. In the first year, the consumer has the WIBR associated with the hedge, but this shifts to the generator in year two.

If materiality is defined in the context of New Zealand as a whole[3], then we conclude that WIBR is of most interest between nodes which have a high concentration of generation, or demand, or both. Thus, once WIBR regions are identified, then further analysis should focus on those regions with the greatest concentration of generation and demand. WIBR regions with little or no generation or demand can either be ignored or be the subject of secondary analysis.

3.4 WIBR Drivers

Of particular concern to most parties facing some degree of WIBR are the occurrence of price separations when one or more line or equation constraints bind, and which often create large IRPDs. Assuming that the grid is designed to cater for “normal loadings” without constraining, then in principle, price separations are caused by abnormal or atypical loadings. These in turn could be caused by a number of factors either on their own or together in some way:

- line outages: causes other lines to be loaded in an abnormal or atypical fashion, e.g. 21 August 2004;
- generation outages: causes changes in power flows which causes lines to be loaded in an abnormal or atypical fashion;
- demand: record or unanticipated peaks in demand causes lines to be loaded in an abnormal or atypical fashion;
- offers: changes in offer strategy by one or more generators causes lines to be loaded in an abnormal or atypical fashion;
- hydrology[4] (inflows and storage): spill, or low inflows combined with low storage, cause lines to reach their limit, e.g. low generation at Manapouri and constraints in the lower SI;

[3] Which is appropriate in the regulatory context.
[4] This cause has elements of both generation outages and offer changes.

- grid capacity: as demand grows, or if generation is built far from load centres, the potential for constraints increases due to generally higher loading on certain areas of the grid;
- grid topology (how the grid connects): loops in the grid create the potential for spring washer, e.g. the permanent split between WPW and FHL on the East Coast of the NI reduces the potential for spring washer in the 220/110 kV loop in the NI by breaking one loop.

Analysis of spring washer events shows that a combination of the above causes is often at play, for example a line outage combined with a generation outage, or two line outages at once.

4 Options for Assessing WIBR

Given the number of nodes on the grid in each island, it makes sense to divide each island into WIBR regions within which WIBR is below some threshold value. The task of assessing WIBR on an on-going basis then comes down to assessing WIBR between a much smaller number of regions.

The threshold value should be chosen based on the trade-off between the number of regions, and the materiality of each region. It may be that many regions are not material and can be amalgamated into one or more neighbouring regions, or alternatively immaterial regions may simply be ignored while the focus is on regions that are material.

It may also be that regions change over time as WIBR drivers change, e.g. due to demand growth or to grid upgrades. A larger number of regions would facilitate monitoring of changes in regions, e.g. do two regions tend to become one over time, or does one larger region become two over time?

4.1 Regions: Clustering of Nodal Prices

Nodes may be clustered into regions within which WIBR is low: in other words, the price of each node within a region must correlate highly with the prices at all other nodes within the same region.

It is reasonable to assume that nodes that are distant on the grid will not correlate as highly as nodes that are closer. By distance, we refer to the number of lines that are connected between two nodes. So nodes within a region will tend to be located in close proximity on the grid.

There may arise a situation in which the price at node A correlates highly with the price at node B, and the price at node B correlates highly with the price at node C, but the price at node A does not correlate so highly with the price at node C. To be within a region, then the price at a node must correlate highly with prices at all other nodes in the same region. The potential for nodes to be within two regions, and the number of regions, will be determined to an extent by the threshold that is chosen as the minimum level of correlation for membership in a region. This threshold will be chosen carefully, and the choice may involve a trade-off between the number of regions, the degree of cross-over between regions, and the strength of the correlations that determine region membership.

To illustrate the approach to clustering nodes, six nodes were sampled using prices from 70 months from Jan-06 through to Oct-11. Correlations between prices were calculated in three different ways. First, the correlation over the whole 70 months was calculated. Second, the correlation was calculated by month. Third, prices were averaged into six hourly blocks and correlations calculated by month. The results are shown in the three figures below for Benmore, Islington, Papanui (connected directly to Islington), Stoke at the top of the SI, Haywards and Otahuhu. The data is also tabulated in Appendix 1 - Sample Correlations.

The three figures show the correlation between pairs of nodes. The closer the correlation is to one (data point near the outside of the circle), the more the prices at the two nodes tend to move together.

Taking the correlation period to be the entire 70 months (Figure 1), BEN, ISL, PAP and STK all correlate highly with each other, with the minimum correlation BEN:STK sitting at 0.995. Using this approach, these four nodes appear to form a region. In reality, constraints can and do occur between BEN and ISL and between ISL and STK (the latter are less common since ISL-KIK was upgraded), while ISL and PAP are clearly within a region with correlation equal to 1.000. Using this approach, the correlation threshold for defining regions would have to be set very close to one, since ISL and STK correlate to the tune of 0.998 over this period.

Figure 1 - Nodal Correlations at the Half Hourly Level Over 70 Months

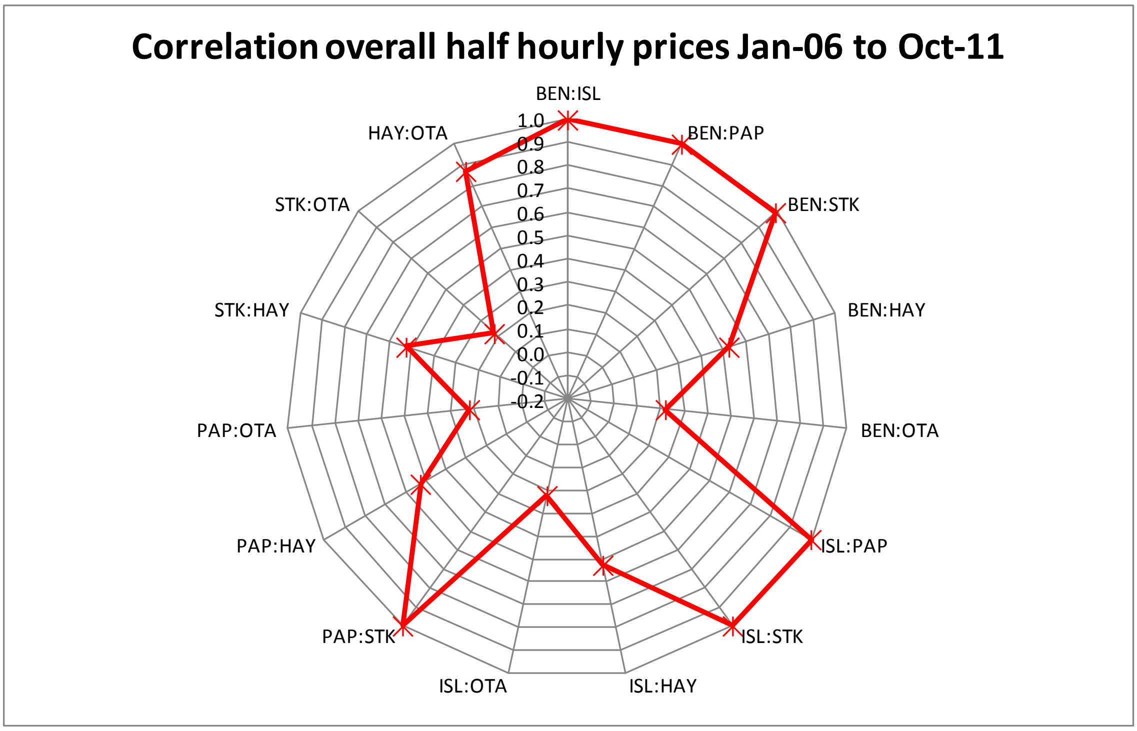


Figure 2 shows a different approach to Figure 1, but still using half hourly prices. The correlations behind this figure are calculated for each month, and the figure shows the minimum monthly correlation from the 70 month period. Using this approach, a threshold correlation of 0.9 would identify ISL and PAP as being nodes within the same region, and all other nodes being in other regions. Reducing the threshold to 0.89 would add BEN to the ISL, PAP region.

Figure 2 - Monthly Nodal Correlations at the Half Hourly Level Over 70 Months

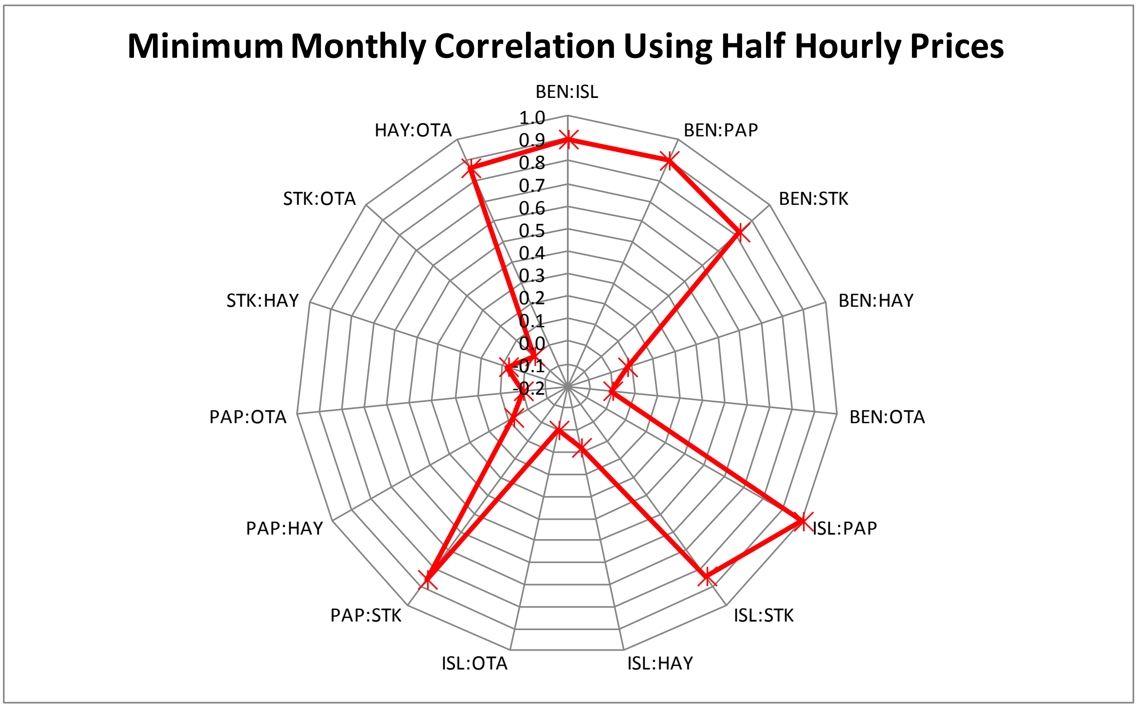
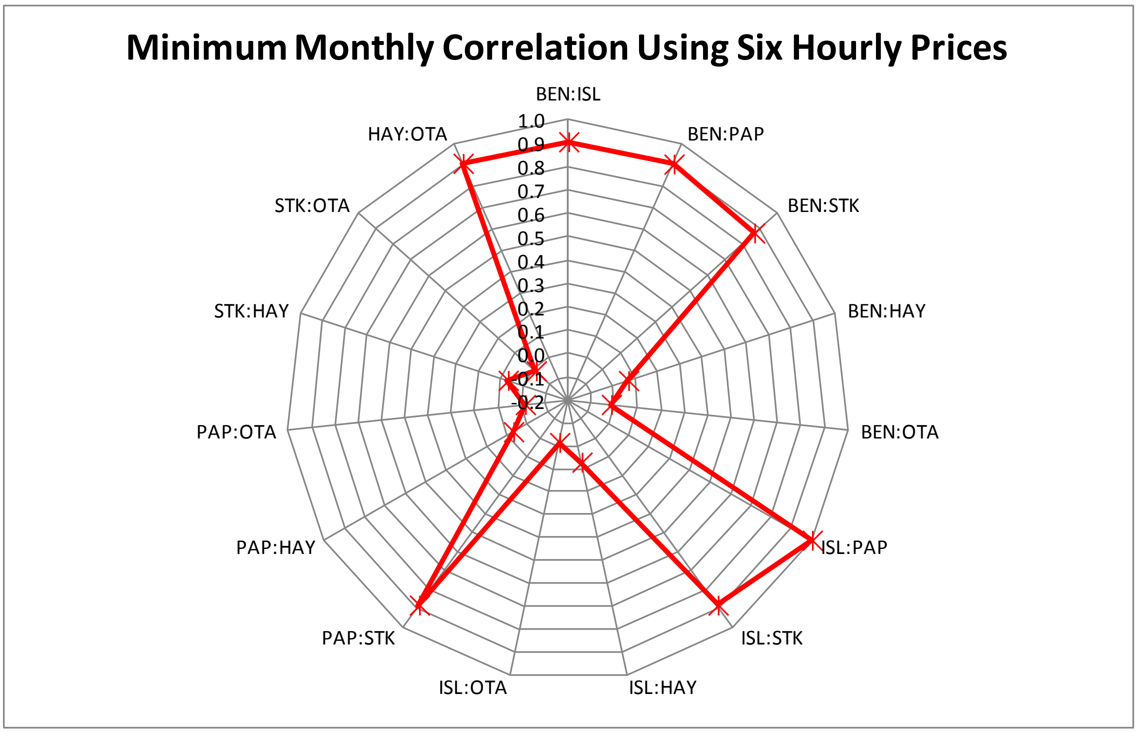


Figure 3 shows the effect of using prices averaged over six hours, but still using the minimum monthly correlation as the parameter for choosing nodes in a region. In this case a threshold of 0.91 identifies only ISL and PAP as being in the same region.

Figure 3 - Monthly Nodal Correlations at the Six Hourly Level Over 70 Months



The two conclusions to be drawn from the analysis of the small sample of six nodes are that:

1. the correlation analysis must be conducted at the monthly level to ensure differentiation of regions; and
2. prices could be averaged down below the half hourly level with only a minor impact on the threshold correlation value and the ability to identify clusters of nodes that make up WIBR regions.

The second conclusion is important because it would allow us to reduce the data handling requirements by approximately an order of magnitude, if this is required in the full analysis recommended in section 5. Although six hourly prices were used in the sample analysis, four hourly prices are more likely to be used in a full analysis as it is generally recognised that four hour blocks capture peaks well.

4.2 Assessing WIBR Between Regions

In this section we briefly consider five different approaches to assessing WIBR into the future:

1. multi-scenario modelling;
2. scenario analysis;
3. Monte Carlo simulation;
4. historical analysis;
5. fundamental analysis.

The emphasis in the discussion below is on finding a method which is suitable for use in the regulatory setting, where the costs and benefits to New Zealand as a whole are of prime importance. The time frame under consideration is relatively long at 5 - 10 years, but this is required so that anticipated grid upgrades and demand growth can be factored in to WIBR assessments, and also to be consistent with the Authority's approach to cost-benefit analysis.

4.2.1 Multi-scenario modelling

A model such as *EMarket*[5], which was used earlier this year to project the LCE for 15 years as part of the assessment of IIBR, models the grid in detail more than sufficient to assess WIBR between regions. For the 15 year projection of the LCE, however, no outages were modelled due to the increased modelling effort this would have required. As it was, the simulations were run at the day-night level, three demand growth scenarios were modelled, along with 80 years of inflows in each year modelled, and reserves were enabled: this exercise modelled 3,555 individual years and represented 15 days of uninterrupted modelling on one PC.

When the potential for any line on the grid to have an outage, along with any generator outage, along with changes in offer strategy, and all modelling needing to be to the four hourly level, the number of modelled years expands to the extent that the exercise would be excessively long and very expensive to undertake. Even limiting outages to key lines and generators would see the exercise expand by a factor in the hundreds. The quantity of output data would also be vastly greater and would require a major computational effort to process it into a useable form.

[5] The SDDP model is also able to model the grid to the level required.

Although multi-scenario modelling has the advantage of being able to relate cause and effect with certainty, this approach is best used to analyse a limited number of scenarios that may be of particular interest from time to time. For example, by limiting the modelling exercise to one year and to the investigation of WIBR between two specific regions.

4.2.2 Scenario analysis

An approach routinely used by market participants, grid owners and analysts is to model particular scenarios, for example the impact on the market of a known or likely planned outage or grid upgrade sometime in the near future, over a limited period typically anywhere from one half hour to a few weeks. In our case, we use our *EMarketOffer* model which models the grid and equation constraints in full detail and includes reserves and the ability to change offers in any way desired.

This approach is very accurate as long as key inputs are either known or can be predicted accurately, but it is also more time-consuming than multi-scenario analysis. While it is appealing to have accurate analysis on particular scenarios, focusing on a small number of scenarios could mean that many important future scenarios are missed.

In the regulatory context, this approach is probably only useful for illustrating particular scenarios or possibilities, for investigating a small number of historical events, or for supplementing comprehensive analysis.

4.2.3 Monte Carlo simulation

Section 3.1 discussed the conceptual model of WIBR as being a combination of seasonal patterns with mean reversion, random noise, and a range of jumps. Monte Carlo models in this context are times series simulations that use repeated random sampling as they move from one sample period to the next (Energy_Link 2010). Mean-reversion jump diffusion (MRJD) Monte Carlo models have been applied to electricity prices and can also be applied to INPDs, for example the INPD (or IRPD for that matter) between two nodes could be modelled as:

$$\frac{\Delta INPD_t}{INPD_t} = \alpha(\overline{INPD} - INPD_t)\Delta t + \sigma\varepsilon\sqrt{\Delta t} + \kappa dq$$

where $\overline{INPD}$ is the mean value of INPD, α is the rate at which the INPD returns to the mean, κ is the proportional jump size, and dq is a random variable that is mostly zero but is occasionally equal to one when a jump occurs.

In principle, deployment of an MRJD model is straightforward and only requires the choosing of parameters which accurately reflect the underlying processes driving INPD. MRJD models are computationally efficient and, once set up and tuned, can be run many times, very quickly.

While conceptually simple, and useful in illustrating the observed behaviour of INPD, applying MRJD to the real world has a number of significant hurdles to cross. The parameters, including the mean INPD, are usually estimated from historical data, so are only predictive in a meaningful way as long as nothing changes in the underlying grid and market.

In addition, real world INPDs exhibit seasonality (requiring the mean to follow a seasonal pattern) and also more than one jump process, e.g. short term jumps caused by outages and longer term jumps reflecting changing patterns of generation as occurs in dry years. Obtaining accurate results even in the historical sense requires significant extension of the simple MRJD model above, and also requires other models or processes which predict how key parameters might change over long periods of time.

While combining MRJD with multi-scenario analysis provides some solutions to these problems, this approach also introduces much longer processing times as discussed in section 4.2.1 and is unsuitable for the purposes required by the Authority.

4.2.4 Historical analysis

Relying solely on historical INPD data suffers the same problem as those suffered by MRJD, as it is not predictive if anything changes. However, historical data is always useful in the sense that it indicates context. For example, if INPDs have been $20,000/MWh in the past, then one must assume they could be at least this large in the future (unless something changes which would prevent this).

Of course, historical analysis remains the first choice for assessments of WIBR that are entirely backward-looking.

4.2.5 Fundamental analysis

This approach is predictive in a wider sense, but not reliant on using large numbers of scenarios. In essence, it mines historical WIBR data in search of enduring relationships between WIBR and the drivers of WIBR listed in section 3.4. The relationships would not be specific to individual lines, for example, but would be of the form “WIBR increases as grid capacity rises”, or perhaps “the most significant driver of WIBR is the occurrence of line outages”.

The fundamental approach is to answer the question: given WIBR during month X in the past, what WIBR drivers were factors present during month X and to what degree? This analysis essentially builds up a multi-factor regression model[6] for which the inputs can be predicted well into the future.

We recommend that fundamental analysis be used for on-going forward-looking assessment of WIBR, and the methodology is outlined in section 5 below.

5 Recommended Approach

The methodology for a thorough analysis of WIBR proceeds in three phases, using half hourly prices, but noting that four hourly prices could also be used if the computational requirements of half hourly analysis turn out to be excessive (which seems unlikely).

[6] Or if not a regression model per se, at least a set of qualitative and quantitative relationships which can be used to predict WIBR in the future.

1. **Phase 1: Identifying regions**
   Determine a set of regions within which WIBR is below threshold, and between which WIBR is above threshold.

2. **Phase 2: Processing WIBR Drivers**
   Given that WIBR regions are defined, then some set of one or more factors must cause WIBR between regions: process the WIBR driver data and determine relationships between IRPDs and each driver.

3. **Phase 3: Projecting Forward**
   Using the WIBR regions, and the impact of each WIBR driver, discuss how each driver will change in future (out to 5 years, for example) and hence draw conclusions about how WIBR will evolve.

5.1 Identifying Regions

The grid will be split into two groups of substations (which we will refer to as nodes in the following) - 108 in the NI and 67 in the SI - and the average half hourly[7] price in each region will be calculated over at least a five year period (the study period).

The number of node pairs is equal to N(N-1)/2 where N is the number of nodes, giving 5,778 node pairs in the NI and 2,211 node pairs in the SI (see for example, Appendix 1 - Sample Correlations for the 15 pairings of six nodes). For each node pair in each island, correlations will be calculated for each month in the study period, giving a total of at least 346,680 correlations in the NI and 132,660 correlations in the SI. From this data, only the month with the lowest correlation over the study period will be used to determine membership of WIBR regions.

A correlation threshold will be chosen, initially 0.9. Correlation data will be processed for each island to identify clusters of nodes for which the minimum monthly correlation value exceeds the threshold for all nodes in each cluster. In other words, a cluster is formed from nodes which correlate in all months of the study period to at least the threshold correlation value. These clusters will form the WIBR regions in each island.

There will be an iterative process at this point, in which the threshold value will be tested against the number of regions. For example, if 0.9 produces 50 regions in the NI then a higher threshold is likely to be used, although this may also depend on how many regions are actually material. For example, if there are 50 regions but only 20 are material, then it may be decided that no further adjustment to the threshold is required as further analysis can focus on the 20 material regions.

The iteration will continue until we are satisfied with the trade-offs between the number of regions, materiality and the amount of data to be processed in the next phase.

5.2 Processing WIBR Drivers

Phase 2 attempts to establish relationships between IRPDs and the WIBR drivers listed in section 3.4, with the focus being on IRPDs created by price separations. The analysis

[7] This appears to be feasible, but if problems are encountered with the volume of data then four hourly prices will be used.

will use all half hours in the study period when price separations occurred, identified by either one of:

1. the Shadow Price field contained in the arc flow data available from em6 being greater than zero, indicating a binding line constraint; or
2. the UpperMarginalPrice field in the SPD final pricing case data files available from NZX Energy via the Authority[8] being greater than zero, indicating a binding line constraint.

Confining the analysis to trading periods when price separations occur reduces the sample size to a small percentage of the half hours, probably no more than 5,000 over the entire study period, and focuses on the periods when IRPDs create the greatest WIBR.

For each half hour in the sample, the following data will be prepared:

- degree of price separation expressed as the standard deviation (i.e. the spread) of prices across all WIBR regions;
- weekly average island price, taken over all periods that do not have price separations;
- total loading on the grid in the island, as indicated by the average loading calculated using arc flow data available form em6;
- number of line outages, and total capacity in MW of lines in outage: outage data is available from SPD final pricing case data and from files on em6;
- whether or not the HVDC link is having an outage: total MW unavailable;
- total MW of generation having an outage within the island;
- island demand;
- total power transfer on the HVDC link, with northward flow defined as positive: positive values indicate a low shortage probability, low values indicate a high shortage probability (which is associated with high levels of southward transfer);
- change in offers during the half hour: this requires processing the island supply curve to estimate how much the supply curve changed during the price separation, relative to the same half hour in the week prior to the event, and this would be done by calculating the sum of the squared difference between supply curves[9].

Grid topology will not be assessed as part of the analysis because this would require a significant amount of additional work which would probably add little value, given that line outages alter grid topology, and that the default grid topology is fixed for long periods.

Once the WIBR drivers are processed, it will be a relatively straightforward exercise to search the sample data to find relationships between the spread of regional prices during price separation events and the WIBR drivers. The objective is to determine relationships such as, for example, “WIBR is driven by line outages” or “WIBR is driven by a combination of a line outage, high demand and a change in offer prices”.

[8] This field is definitely available in these files back to July 2009, but we have to confirm that it is available in earlier final pricing case data files once we have the data.

[9] Take the supply in the relevant half hour and at intervals of 10 MW, calculate the difference between the price offered during the price separation event and the same half hour in adjacent days. Square the differences, sum the squares, and then take the square root (root mean squared difference). The larger the final difference, the more the offers must have changed during the price separation event.

These relationships will be quantified where possible, for example using some form of multi-factor regression analysis.

5.3 Projecting Forward

Assuming that significant relationships can be found between IRPDs and the WIBR drivers, then it will be possible to draw conclusions about WIBR in future. For example, if the conclusion is that WIBR is driven by line outages then since line outages are likely to be a feature of the market forever, it would be reasonable to conclude that WIBR will continue ad nauseam.

Or if WIBR is driven by grid capacity then it would be reasonable to conclude that there will be long periods in future, after grid upgrades, when WIBR will fall until demand growth reduces grid capacity. In this case, conclusions may be able to be drawn about which regions will experience WIBR in future, i.e. high growth areas, and when WIBR might fall or rise.

The final part of the study, therefore, will be to use these relationships to draw conclusions about WIBR in the future, to the extent that this can be done without significant further analysis, or if further analysis is required then this could be commissioned as a separate exercise. But in any case, the overarching objective is to inform the LPRTG about how WIBR may evolve over time.

5.4 Study Outputs

There are three key outputs from the recommended analysis:

A. a list of the WIBR regions as they fell in the last five years;

B. a much improved understanding of the underlying drivers of WIBR, ranked in order of importance and with an understanding of how the various drivers interact to create WIBR;

C. analysis and commentary on how WIBR may evolve in future, including some indication of the WIBR regions that are most likely to experience changes in WIBR in the foreseeable future (the period over which there is some certainty about grid upgrades and demand growth, say five years).

6 Appendix 1 - Sample Correlations

	BEN:ISL	BEN:PAP	BEN:STK	BEN:HAY	BEN:OTA	ISL:PAP	ISL:STK	ISL:HAY	ISL:OTA	PAP:STK	PAP:HAY	PAP:OTA	STK:HAY	STK:OTA	HAY:OTA
Minimum Monthly Correlation Using Half Hourly Prices	0.896	0.896	0.821	0.075	-0.003	0.999	0.842	0.078	0.000	0.860	0.079	0.000	0.078	0.000	0.860
Minimum Monthly Correlation Using Six Hourly Prices	0.904	0.904	0.867	0.068	-0.017	1.000	0.884	0.072	-0.014	0.884	0.072	-0.014	0.072	-0.014	0.908
Correlation overall half hourly prices Jan-06 to Oct-11	0.996	0.996	0.995	0.524	0.220	1.000	0.998	0.525	0.221	0.998	0.525	0.221	0.525	0.221	0.869

7 Appendix 2 - References

Clelow, L., C. Strickland, et al. (2001). Extending mean-reversion jump diffusion. Energy Price Risk Management. United Kingdom.

Electricity_Commission (2010). Consultation Paper: Managing Locational Price Risk: options. Wellington.

Energy_Link (2010). Validation of Electricity Modelling for Energy Outlook. Dunedin.

Energy_Link (2011). Losses and Constraints Excess Projections. Dunedin.