Evaluating options to improve the system operator load forecast

Technical Advisory Service – SOW 073

Authority Project Code: WM.16.022.01

Transpower New Zealand Limited

May 2018

Keeping the energy flowing

Version	Date	Change
0.1	27/04/2018	Draft for Electricity Authority Review
1.0	11/05/2018	Final Version

IMPORTANT

Disclaimer

Copyright

Contact Details

Address:	Transpower New Zealand Ltd Waikoukou - 22 Boulcott St PO Box 1021 Wellington New Zealand
Telephone:	+64 4 495 7000
Fax:	+64 4 498 2671
Email:	system.operator@transpower.co.nz
Website:	http://www.transpower.co.nz

EXECUTIVE SUMMARY

The Electricity Authority engaged Transpower as system operator to define and assess options for improving medium-term load forecasting, providing sensitivity schedules, and how the two relate.

Our investigation into load forecast improvements sought input from a number of sources, including Transpower's 2012 load forecast trial and international consultation.

We identified a number of options for improving medium-term load forecasting, as variations of the four options specified by the Authority in its statement of work, and assessed these options according to accuracy and cost criteria agreed with the Authority.

On the basis of these criteria, our assessment concluded:

- The current load forecast model is unable to be future-proofed and will need to be replaced at some point. A full replacement of the load forecast system may have a lead-time of a number of years.
- In the intervening period, it is possible to make improvements to the current load forecast model that will produce benefits in the short-term. These include:
    - a number of quick wins that could produce immediate benefits, for example increasing the quality of current weather inputs.
    - updating the load forecasting model by recalibrating the current version
    - targeting ongoing intervention by Transpower staff
    - improving the input data to the current forecast, including increasing frequency of current weather inputs and getting new inputs
    - 'plugging in' the current TESLA model, accessed by ems via a web service, to the existing interface. Security and reliability concerns would be reduced by using the current tool as a backup.
- Making improvements to the current load forecast system now will enable assessment of the aspects of the model that create the most benefit. This will feed into the full system replacement specification.
- In the longer term, a replacement system should include more sophisticated models, full functionality for Grid Exit Point (GXP) level forecasts, and incorporate price responsiveness. This will allow accuracy to be maintained in the future, following significant uptake of new technologies such as roof-top solar, batteries, and smart home energy management systems. In the near term, these changes would improve accuracy during stressed market conditions and when transmission constraints bind.
- The choice between whether the full-replacement system is open and primarily controlled by Transpower, or closed and primarily controlled by an external vendor, is a trade-off between cost, accuracy and adaptability; the closed model has greater accuracy and is potentially as adaptable but comes at a greater cost.

Our investigation into sensitivity schedules and their relation to load forecast improvements concluded:

- Sensitivity schedules will help participants assess their risks and opportunities by signalling how much forecast prices would change for a certain change in forecast load, or other inputs. Both sensitivity schedules and load forecast improvements enhance information used in participants' decisions in the forecast horizon, and can be developed independently of, or alongside, each other.

- Sensitivity schedules will be of particular value during stressed market conditions, when prices are highly sensitive[1], especially at longer forecast horizons when other inputs are highly variable. During these times, even the best load forecast improvement option can lead to large price forecast errors.
- Providing information on load forecast variability in addition to sensitivity schedules will greatly assist participant decision making, by increasing visibility of the expected variation between forecast and final prices.
- Improving the load forecast to reduce outliers will limit unexpected variation between forecast and final prices. This will allow participants to act on price sensitivity information with greater confidence.

[1] Sensitive prices are where a small change in schedule inputs, e.g. load, lead to a large change in prices.

CONTENTS

1 Introduction

1.1 Purpose

This Technical Advisory Service (TAS) report responds to the Electricity Authority's (Authority) Statement of Work (SOW) 073 to define and describe options for improving medium-term load forecasting, providing sensitivity schedules, and how the two relate.

1.2 Audience

There are two audiences for this piece of work.

- The detailed investigation contained in the appendices, is written for analytical staff at the Authority whose work relates to the wholesale electricity market.
- The body of the report, in particular the executive summary, is written for a more general audience at the Authority.

1.3 Information Sources

The need to improve the accuracy of load forecasting is not specific to the New Zealand market. This is a challenge that all electricity markets are facing as markets shift to a distributed energy model that places the importance of the demand side on a par with that of the supply side.

Recognising there has already been research in this area, we carried out an information gathering exercise to seek input from a number of varied sources. These sources include:

- Load forecasting experts from overseas system operators, including the Australian Energy Market Operator (AEMO), and system operators from the United States of America (PJM and MISO).
- A range of subject matter experts within Transpower, with diverse skill sets including market operations, system coordination, market development, grid planning, grid economics, IST systems architecture and design, and analysts with statistical knowledge.
- Electricity Authority expertise in load forecasting.

We also:

- examined in detail how the current tool works
- reviewed the previous Transpower load forecasting work.

We then sourced a variety of data for the purposes of the quantitative analysis, including:

- Recent Transpower data, specifically
    - forecasts primarily from 2017, at forecast horizons of 2.5 and 24 hours
    - actual loads from 2015 – 2017.
- Historic data used in the Transpower load forecasting trial in 2012.
- TESLA – Energy Industry Forecasting Solutions (TESLA) forecasts[2] from 2017.
- Weather data over and above[3] that already provided by the Meteorological Service of New Zealand (MetService), namely:

[2] Energy Market Services (ems) has partnered with TESLA to deliver a highly accurate GXP level load forecast, out to 14 days in advance. For more details see http://ems.co.nz/portfolio/energy-information/loadforecast

[3] MetService already makes its weather data available to Transpower at 6 hourly intervals through the day.

- more precise (data with greater decimal place accuracy), more frequent (3-hourly forecasting) and more temporally granular (hourly data points) forecast data covering March – April 2017.
- observations of a range of weather parameters for 2017, including humidity, solar radiation, wind direction, rainfall and cloud cover.

1.4 Scope Mapped to Report Contents

The contents of this TAS report follow the scope outlined in section 3 of the Authority's Statement of Work as follows:

a) A summary of Transpower's work done on load forecasting to date:

Section 3

b) An outcome strategy map that will outline problem definition, business change capabilities required and potential benefits sought:

Section 4

c) A summary of options for delivering business change capabilities:

Section 6

including (but not limited to):

- Option 1: Any rapid deployment adjustments ("quick wins") to the current medium-term load forecast system to improve the accuracy of the load forecasts.
- Option 2: A refresh of the current medium-term load forecast model inputs using data, keeping model inputs and outputs in the current form (weather forecast input structure and regional load forecast outputs)
- Option 3: A refresh of the current medium-term load forecast model using recent load data and a modernised weather forecast (more granular temporal, geographical forecast inputs) and other inputs to load forecast which are not currently factored in, but preserving the current regional outputs.
- Option 4: A full replacement of the of the current medium-term load forecast system with more granular inputs and outputs e.g. nodal weather forecast inputs and load forecast outputs, and other inputs to load forecast which are not currently factored in. This option will include:
  - A model where access and rights to the intellectual property is open, transparent and adaptions enabled.
  - A model where access and rights to the intellectual property is closed.

d) An assessment of the above options against specified criteria, which will be defined with Authority during problem definition phase, but will include accuracy, operation and maintenance, and future proofing.

Section 5 contains the criteria for assessing options

Section 6.1 provides initial assessment of options

Section 7 provides the assessment of final options

e) A summary of the Sensitivity Schedules project and how this fits with improvement of medium term load forecasting.

Section 3.1 provides a summary of the previous Sensitivity Schedules project

Section 8 discusses how sensitivity schedules fit with load forecast improvements

Out of scope items are:

- Price estimates for the options identified.
- Consideration of improvement to system operator forecasting non-conforming load. (Note: the system operator does not currently forecast non-conforming load)
- Improvements to forecasting wind generation. (Note: the system operator does not currently forecast wind)
- Any consultation with industry on the options paper.
- Assessment of options according to price forecasting accuracy.

2 Context and Background

2.1 Electricity Authority Consultation on Making Hours-Ahead Price Forecasts More Accurate

This TAS investigation is an outcome of the Authority's project to improve information leading into real-time. On 9 February 2017, the Authority released the document "Making hours-ahead price forecasts more accurate"[4]. This consultation paper included four options for consideration:

Option A: (the preferred option) improve inputs into price forecasts under existing incentive arrangements (administrative/beneficiaries pay)

Option B: improve inputs into price forecasts and improve incentives (beneficiaries/ exacerbators pay arrangements)

Option C: encourage a voluntary hours-ahead market (market-like arrangements)

Option D: pursue a formal hours-ahead market (market-like arrangements)

On 15 August 2017, following a review of participant submissions, the Authority released a decision paper[5] outlining its decision to progress with Option A.

The original consultation contained analysis identifying conforming load forecast error was the single largest contributor to hours-ahead forecast price error. Therefore, following submissions from market participants who generally agreed that this should be investigated further, the next step was for the Authority to request that the system operator evaluate options to improve the medium-term load forecast. The medium-term load forecast provides the load information for conforming GXPs into the forward market schedules from 30 minutes to 7 days ahead of real-time.

2.2 Overview of Medium-Term Load Forecasting Systems

2.2.1 Load forecast systems

A typical load forecast system comprises:

Load forecast inputs Information that can influence load, including weather, historic load, and other inputs.

Load forecast model A model to convert inputs into forecasts of load.

Load forecast tool A tool for users of the model to access and edit the input data, configuring parameters, performing overrides, monitoring performance, etc.

IST infrastructure Infrastructure to send and receive data, house the tool, house the model, run the model, store the data, provide security, interface with the market system, etc.

People People to monitor and adjust the model and inputs, and provide IST support.

Together these components produce the load forecast.

[4] https://www.ea.govt.nz/dmsdocument/21777

[5] https://www.ea.govt.nz/dmsdocument/22436

2.2.2 Transpower's current load forecast

Transpower's current medium-term load forecast tool[6] models the forecast load for 10 geographical areas, every half hour, for all trading periods up to 15 days in the future. This is apportioned to the GXP level using each GXP's portion of area load for the same trading period the week before.

The current forecast is broken into four components:

Long-term component	Load that occurs at the same time of the day in prior weeks, on related days.
Short-term component	Any changes to the magnitude of the above, or any other load, that occurs at the same time of the day in the few days prior.
Weather component	Any variation in magnitude of either of the above that is due to temperature or wind speed. (Wind speed is only used for three areas.)
Refined component	This component only applies for periods during the same day in which the forecast is run. Forecast load is scaled to account for differences between recent forecasts and actual load.

2.2.3 History of Transpower's load forecast

The earliest version of Transpower's current medium-term load forecast tool was implemented for the purposes of assessing and maintaining system security. Between 2004 and 2012, forecast market schedule results based on the load forecast were only available within a 4-hour horizon of real-time. In 2007, a longer horizon forecast schedule was made available to market participants as a 'special winter schedule'. This change was not made through regulation; it was a response to the issues the industry faced during the winter of 2006 as a result of the inaccuracies of the then long-horizon forecast market schedule (derived from participant load bids).

In 2012, as a result of the Demand-Side Bidding and Forecasting (DSBF) initiative, the load forecast was used as an input to the NRS and PRS schedules (forecasting up to 36 hours ahead) and was included in the Code for the first time. This represented a significant expansion of the requirements of the load forecast tool.

[6] Further details about the current load forecast system can be found at *https://www.transpower.co.nz/sites/default/files/bulk-upload/documents/GL-SD-204%20Load%20Forecast%20Methodology%20and%20Processes.pdf*

3 TRANSPOWER'S PREVIOUS WORK

3.1 SENSITIVITY SCHEDULES

Sensitivity schedules provide a mechanism for indicating the sensitivity of prices to changes in inputs such as load. They enhance participants' ability to understand the variability of the price forecast, and hence assess their risk and opportunity.

In 2017, Transpower carried out a project to investigate the benefit of sensitivity schedules. A short summary of this project is given here.

3.1.1 What are sensitivity schedules?

Currently, market participants can see only the current price for energy based on forecast load and are not aware of the sensitivity of price to fluctuations in generation and/or demand. Sensitivity schedules provide a mechanism for indicating such price sensitivity.

The project examined implementation of sensitivity schedules for both the NRS and PRS long and short schedules. It recommended two additional schedules providing variations to each NRS and PRS, one with load increased from that forecast, the other with load decreased.

3.1.2 Requirements for the implementation of sensitivity schedules

The project investigated the modification of the SO market tools and associated systems and processes required to accommodate sensitivity schedules. Current-state capability to execute sensitivity analysis is hampered by inadequate tools and manual manipulation of input data. There is also no business process capability to carry out sensitivity analysis.

The proposed solution was based upon Transpower stakeholder requirements developed through workshops with subject matter experts and from a survey of industry participants. Implementation was expected to take approximately 10 months.

3.1.3 The value of sensitivity schedules

Published sensitivity schedule information would be of significant interest to market participants. This additional information would enable generators, reserve providers and price-sensitive load providers to revise their scheduling based on enhanced knowledge of the variability potentially impacting their decisions.

3.1.3.1 System security

Sensitivity analysis functionality would enable increased system security through heightened situational awareness. It would provide the ability to pre-empt security issues and understand their economic impacts, which would be of benefit to all customers and consumers.

3.1.3.2 Demand-side

Sensitivity schedules would provide useful additional information for customers with high sensitivity to price step-changes. Also, demand-side participants who on-sell currently cannot incentivise the price response of their customers, as they do not have information on the sensitivity of price. Although their customer base may not be significant enough to be able to influence price outcomes, provision of a price-risk profile would still be useful.

3.1.3.3 Supply-side

Supply-side participants currently find extrapolation of scenarios based on the available price stack information difficult. Inclusion of the results of energy-reserve co-optimisation and the impact on transmission constraints as part of sensitivity information would be especially beneficial for them. Such co-optimisation modelling can only be carried out in the market system.

3.1.3.4 Extreme pricing situations

Sensitivity information would be useful in extreme pricing situations, when small changes in load make a significant difference in price. According to one surveyed market participant, "the consequence of a load forecast being out by 1 MW can, at times, have a similar impact to being out by 200 MW at other times". This is illustrated by the situation on June 2, 2016, when prices were particularly sensitive. In this case, 12 MW less load for the 17:30 trading period would have caused prices to fall from around $4000 /MWh to around $300 /MWh. This is illustrated in Figure 1 with the vertical demand curve (blue), intersecting a steep supply curve (black).

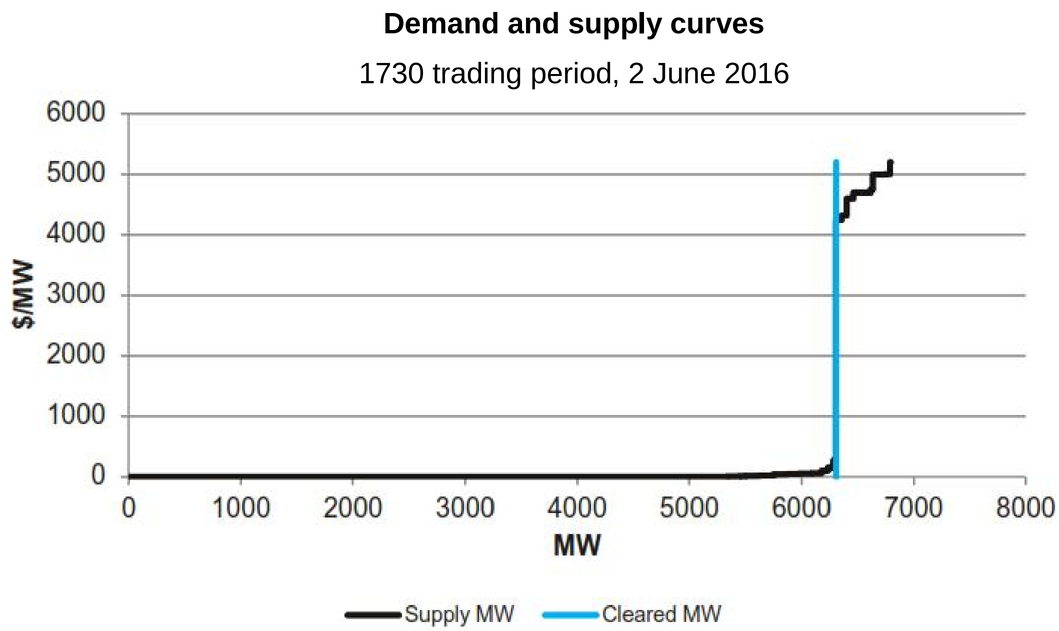


Figure 1 Extreme price sensitivity on 2 June 2016. An increase of 12 MW of load changes prices from $300 /MW to $4000 /MW.

3.2 LOAD FORECAST TRIAL

In 2010, Transpower decided to consider whether it could economically improve the quality of its load forecasts. In deciding how to approach this, we noted that:

- Electricity load forecasting is a specialist area
- Many companies specialise in it, with a variety of different forecasting approaches and business models
- Many of the specialist forecasting companies are based overseas
- Vendor's claims of the high accuracy that they could achieve for our situation could not necessarily be relied upon
- Trials against historical data had the potential to be gamed

For these reasons we decided that the only reliable way of finding the most cost-effective load forecast was to hold a live trial, agnostic to the forecasting approach, and internet-based to give equal access to New Zealand and overseas vendors.

The trial did not test all the functionality that we sought in the operational load forecast, as some of that functionality (e.g. taking account of demand response calls, knowledge of outages and distribution network reconfigurations etc.) would be too expensive for potential providers to implement prior to selection. Non-functional requirements such as reliance are also critical to an operational system and cannot be tested in such a trial. This left the risk that the winner of the trial could mark-up its prices for such additional functionality and requirements. To prevent this we required trial participants to quote a price pre-trial for the full functional, operational implementation.

We provided participants with actual load in real-time, historical load, weather, and historical weather observations, so they all had access to common basic data. They were free to include any other inputs that they thought useful. Every half-hour, they would upload their load forecasts for each region and every GXP. The trial was run from 1 March to 31 August 2012.

Before embarking on the trial, Transpower needed to determine a method of assessing the relative performance of trial participants, relative to each other and to Transpower's existing medium term load forecast (MTLF) model, especially important as such a trial generates a huge amount of data. In order to compare overall performance, Transpower developed a general performance index by weighting different spatial resolutions and horizons, based on their perceived utility to users of the forecasts and resultant schedules, such as the co-ordination centres, outage planners and market participants. We assessed forecasted load only – no attempt was made to measure the price impact that could have resulted.

A Request for Proposals was sent to 14 potential vendors, including all 11 respondents to an initial registration of interest. Five vendors (for the trial and potentially the operational system) were selected. Transpower's existing MTLF was treated as the 'sixth vendor' for comparison purposes.

The trial result was that two participants were well inside the cost-performance envelope, but four lay on the boundary, including the existing system as least performing but least cost. Cost-benefit analysis was conducted (and externally peer-reviewed), concluding that the net market benefits were highest for the most accurate but most costly vendor, TESLA.

The project was well advanced in terms of the functional and non-functional requirements and implementation details of the operational system and contractual arrangements with TESLA when, in late 2014, Transpower put the project on hold due to a change in budget priorities.

The functional requirements did not include incorporating price-responsiveness other than through Dispatchable Demand (DD) and Demand-Side Bidding and Forecasting (DSBF) difference bids. This would likely be an important addition should the project be re-started, given the real-time pricing project, and current expectations on the penetration of batteries, and energy and home management systems.

TESLA's approach is to develop individual forecast algorithms for each GXP (i.e. bottom-up rather than top-down) and monitor these continuously. If actuals diverge from the forecasts, then that forecast algorithm is adjusted. If this model, as was trialed in 2012, was to be used in Transpower's environment, the model would be run from inside the control rooms, TESLA would do its monitoring remotely, and any adjusted models would be sent to Transpower who would inject them into the control room forecasting tool.

TESLA's approach demonstrated the best performance, roughly halving the error of the existing forecast, and performed notably well during special events such as public holidays, daylight savings, a major snow storm, and during grid outages. Subsequently, Transpower through its Energy Market Services team (ems) has partnered with TESLA to deliver GXP level load forecasts, every trading period out to 14 days in advance. It should be noted that these do not include all the functionality that we sought in the operational load forecast, and thus the accuracy of these forecasts would be improved if operationalised in Transpower's control centres.

3.2.1 Changes to Problem Definition Since the Trial

There are four key differences between the problem as defined prior to the 2012 trial and the current problem:

- The Authority project that led to this current investigation focuses on the impact of load forecast improvements to forecast price accuracy. Transpower's trial concentrated on MW accuracy primarily, noting that MW accuracy will increase price accuracy, disregarding all other factors influencing price accuracy.
- During the six years since the trial, there have been changes in the load profiles of different regions, and increased network load control.
- There are several new technologies already in use and many on the horizon. The uptake of solar PV[7], electric vehicles (EVs), and batteries around the world, for example, has increased over the last six years. We are therefore in a better position to assess the uptake of these technologies and the type of modelling required for the load forecast.
- The trial was for limited functionality based on the inputs provided. An operational system should include additional inputs (e.g. taking account of demand response calls, knowledge of outages and distribution network reconfigurations etc.).

3.3 Load Forecast Improvements

A number of other changes have been made to the current load forecast outside of the load forecast trial. The key changes made to the model are outlined below.

- In 2011, the data used as historic load in the forecast model to form the historic profiles in the long and short-term components was changed from MV90 metering data[8] to SCADA data. This corrected issues with offered embedded wind farms in the MV90 data – as by using the MV90 data, wind generation was double counted, once as generation, and once as negative load.
- Following the 2012 trial:
    - SCADA keys[9] have been updated to align with the load they are forecasting.
    - Following this project, system performance improvements have allowed the timing of when the forecast updates in the market system to be brought forward so that each Price Responsive Schedule (PRS) and Non-Response (NRS) schedule uses a forecast that has run more recently, by half an hour compared to what had been the case. Note: This would not have affected the accuracy recorded against the forecast in the trial as it was based on the raw forecast rather than the load used in the schedule.
    - The model used by MetService for the weather forecasts provided was updated in April 2017.
- Since the trial, the current load forecast tool appears to have improved by approximately 20% as seen in Table 1, which compares 2012 and 2017 accuracies. TESLA data was also available for both periods, so is provided as a benchmark to illustrate how difficult each month was to forecast.

[7] Principally roof-top solar in New Zealand but there are large quantities of commercial-scale solar overseas.

[8] MV90 is revenue metering used in final pricing. Representing offered embedded wind generation as negative load (negative metering) was introduced on 1/03/2004.

[9] A mechanism to uniquely identify data

Table 1 Comparison of 2012 and 2017 load forecast accuracy, in terms of mean absolute percentage error (MAPE).

Mean Absolute Percentage Error (MAPE[10])	System Operator				TESLA		
	Short-horizon[11]		Long-horizon		Short-horizon		Long-horizon [12]
	2012 (2hr)	2017 (2.5hrs)	2012 (24hrs)	2017 (24hrs)	2012 (2hr)	2017 (2.5hr)	2012 (24hrs)
Apr	2.2%	1.9%	3.4%	3.3%	1.2%	1.3%	1.3%
May	2.7%	1.8%	4.1%	2.9%	1.2%	0.9%	-
June	2.6%	1.7%	4.7%	2.2%	1.0%	0.9%	1.2%
July	2.4%	2.3%	3.6%	3.5%	0.8%	1.3%	1.2%
Aug	1.7%	1.7%	2.3%	2.7%	0.8%	1.0%	1.3%

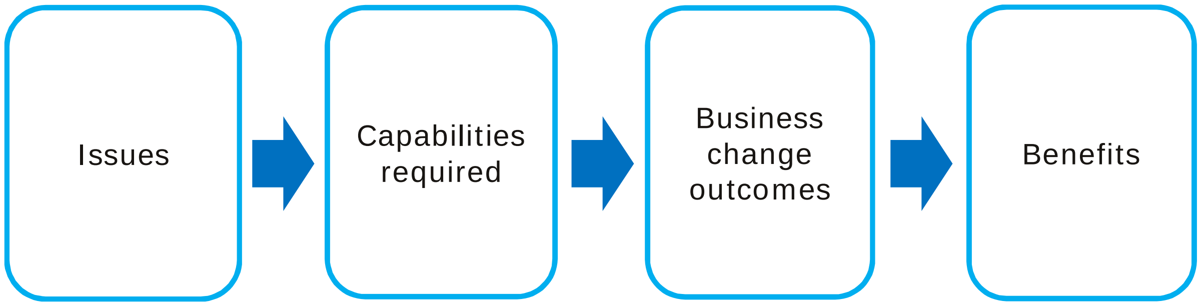
[10] See glossary for MAPE definition.

[11] The short-horizon varies between the 2012 data and the 2017 data, as the sample data was collected for different lead-times. However, both of these differentiate from the long-horizon which is the timeframe used by slower starting generators in order to make their decisions to fire up their plant.

[12] TESLA 24hr horizon forecast data for 2017 is not available (data is also unavailable for May 2012).

4 OUTCOME STRATEGY MAP

Transpower and the Authority agreed on an outcome strategy map to help determine the format of this investigation. An outcome strategy map helps a project team be specific about the areas it intends to target; the changes it hopes to see and the strategies appropriate to achieve these. Our map identified issues to address, capabilities required, business change outcomes, and benefits. The map has been used as a basis for this section.



This section explains how we derived the content of each of these areas, starting at the benefits and working back to the issues that need to be addressed with the current system.

As part of the option analysis, later in this report, we have identified how each of the proposed options deals with the issues we identified; this comparison is contained in a summary table in Section 7.1

4.1 BENEFITS

Improved load forecasts increase ***Competition, Reliability, and Efficiency***[13].

This investigation focusses on the benefits from increased forecast price accuracy. However, we also recognize that the load forecast is a critical piece of information that enables Transpower to operate a reliable and efficient electricity market. Other benefits are considered for completeness.

Competition benefits accrue as increased confidence in forecast prices encourages market participation.

Reliability benefits are likely to accrue from introducing a new forecast if, in the future, weather becomes a greater factor in determining load. Predicting solar and wind generation levels, in particular, can be very difficult, requiring sophisticated forecasting. The chaotic nature of weather means forecast variability information may also be critical to the system operator to enable modelling of various scenarios.

Reliability and ***Efficiency*** are an outcome of the system operator making better security assessments from improved load and price forecasts. The system operator can improve the accuracy of when transmission constraints are forecast and when scarcity situations are signaled; both increase allocative efficiency. Reliability is also increased by improving transmission constraint accuracy, along with employing better measures to alleviate voltage issues.

Efficiency is an outcome of participants making better decisions. If participants have a better gauge of load and spot price at the time they make a decision, they are able to make better decisions; this

[13] These are the elements of the Electricity Authority's statutory objective - to promote competition in, reliable supply by, and the efficient operation of, the New Zealand electricity industry for the long-term benefit of consumers.

increases allocative efficiency[14]. These efficiencies occur where participants require lead time to action a decision, or make multi-period decisions. This includes:

- Thermal generators, making commitment decisions
- Hydro generators, managing flows across a river chain
- Gentailers, assessing their portfolio position
- Industrial load, deciding whether to start or stop processes
- Participants considering hedge arrangements
- Participants charging batteries, including electric vehicles
- Network companies, or consumers exposed to spot prices, controlling heating of hot water cylinders
- Price responsive demand or unoffered embedded generation
- Peaking plant

Efficiency benefits accrue as Transpower, in its role as the grid owner, will be able to make better demand response calls, i.e. less unrequired calls and less missed opportunities to call. This increases dynamic efficiency[15] by reducing the need for transmission investment. Demand response calls typically require some lead time to organise and action the response.

4.2 Business Change Outcomes

All the benefits above accrue as a result of better decision making. Better decisions are enabled by business change outcomes arising from improved load forecasts, and improved price forecasts.

It is important to note that improved load forecasts will not necessarily result in improved price forecasts, due to the non-linear relationship between load and price. Price forecast errors will remain large at very high prices, due to high sensitivity, no matter how accurate the load forecast.

The following business change outcomes were identified:

- Improved information resulting from more accurate load forecasts
- Improved information resulting from more accurate price forecasts
- Improved signals of pre-response load and prices[16] as well as post-response load and prices[17].

These business change outcomes apply to decision makers in the following ways:

- Load and price forecasts are more accurate according to the dimensions relevant to the decision; these dimensions include
    - the geographical level, e.g. GXP or island: A participant needs to know prices at the GXP where they consume load, the system operator needs to know instantaneous reserve requirements at an island level.
    - time of day, e.g. peak period or any time: The system operator needs to ensure reserve requirements are met at peak times, a thermal generator needs to know prices at all times of day as they tend to run for several days at a time.
    - price/load range, e.g. prices around a generator's marginal cost, or high load times: Thermal plant owners need to make decisions by assessing whether prices are sufficient

[14] Allocative efficiency is the employment of resources for the best use from a society's perspective.

[15] Dynamic efficiency is the process of innovation and investment that allows the economy over time to produce a better quality or volume of goods and services demanded by society

[16] Pre-response load and prices are the load and price prior to participants reacting to information provided in a forecast

[17] Post-response load and prices are the load and price after participants have reacted to information provided in a forecast

    to meet their marginal costs, Transpower needs to assess peak loads to determine whether to request demand response.
  - time horizon relevant to the decision, e.g. one day out, a few hours out: Participants entering hedge arrangements need to make decisions at least one day out, whereas hydro generators may be able to wait until gate closure for some of their decisions.

- Load and price forecasts and pre- and post-response signals are sufficiently accurate to enhance market participants' decision making
- Load forecasts are sufficiently accurate to enhance the system operator's decisions when applying transmission constraints and assessing voltage stability
- Price forecasts are sufficiently accurate to enhance the system operator's decisions when assessing scarcity
- Pre- and post-response signals ensure that the system operator can account for the full range of demand response outcomes and so maintain appropriate security
- Load forecasts are sufficiently accurate for determining if interventions are required when transmission operators make a decision to instruct demand response or network load control
- Load forecasts remain sufficiently accurate in the future, for all users

4.3 Capabilities Required

The capabilities required of an improved load forecast will produce business change outcomes to deliver the required benefits. The key capabilities required can be summarised as

- ***Better input data***
  - Appropriate inputs
  - Data cleaned and validated
  - Data at sufficient granularity and provided with sufficient frequency
- ***Better ability to convert inputs into predictors of load.*** This depends on
  - The sophistication of the model(s)
  - Human intervention, in the form of
    - updating models, pre-forecast
    - adjusting forecast values, post-forecast
- ***Adaptability to the future.*** This depends on
  - Infrastructure
  - Contract arrangements
  - Tools/vendor
  - In-house expertise
- ***Secure and reliable system.*** This depends on
  - Infrastructure
  - Reputation of model vendor
- ***Publication of appropriate load forecast information to market***
  - Forecast load and price variability including sensitivity schedules (see section 8)
  - Pre- and post-response load and price signals
  - GXP level forecast load
- ***Incorporation of price response in the market system***
- ***Incorporation of GXP level forecasts in the market system***

4.4 ISSUES FOR LOAD FORECASTING

There are a number of issues that a load forecasting system needs to address to improve its predictive ability and meet the capabilities required. These are outlined below, in terms of their applicability to the current system[18]. The way in which these issues may apply in the future is outlined in section 5.2.2.

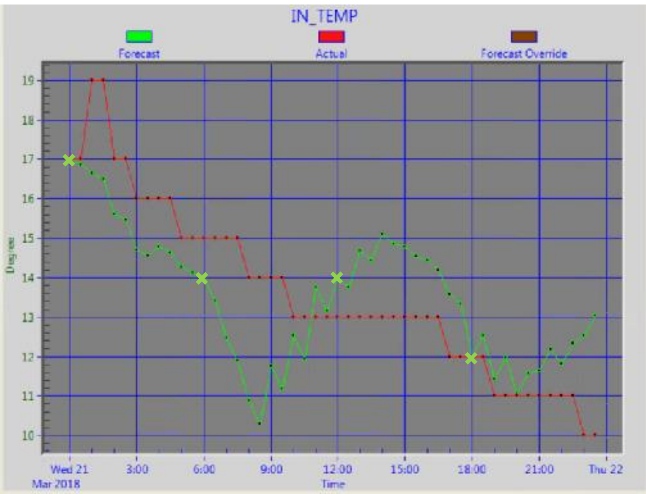
4.4.1 Input data quality

The current forecast has poor granularity and precision for normal weather patterns. The population weightings for regional temperatures within a load forecast area are also out of date.

Figure 2 indicates the granularity and precision issues, by comparing the current situation (LHS) and a possible future scenario (RHS). The LHS graph is an example of the current forecast and actual temperature inputs as used by the load forecast tool. The RHS graph is an example of what can be achieved with improved temperature inputs, and shows a better correlation between the forecast and actual temperature inputs.

At present, the forecast temperatures cover a period of 36 hours and are provided at 6 hourly intervals, to a precision of 0 decimal places (dp). This can be observed in the first graph; only temperatures for 00:00, 06:00, 12:00 and 18:00 are forecast (shown with green crosses). The remaining temperatures are interpolated by the load forecast tool using a profile based on historical temperature observations.

Hourly observed temperatures for the previous 24 hours are provided once a day, also to 0 dp. The software does not interpolate these, instead it makes the value for the intermediate half-hourly trading periods the same as the previous value. This has the effect of producing the saw-tooth pattern seen in the profiled forecast temperatures in the LHS graph.



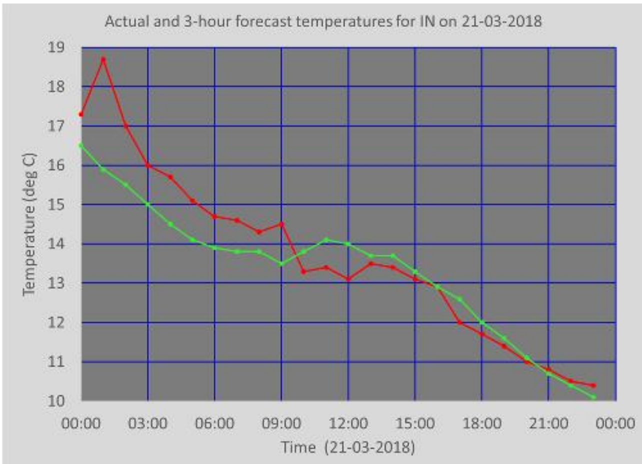


Figure 2 Comparison of currently used forecast and actual temperatures (LHS) with more precise and granular data available (RHS) Data is for 21 March 2018.

---

[18] Summary tables and graphs showing the accuracy of the data for the current load forecast in 2017 at 2.5-hour and 24-hour horizons are provided in Appendix 1. For comparison purposes, TESLA data is also provided. TESLA was chosen for this purpose as their forecast data is readily available, and because, as the most accurate trial participant, comparison indicates the accuracy possible.

4.4.2 Appropriate inputs

Data inputs to the current forecast are limited. Only temperature and, in three of ten areas, wind strength, are used. It is important that the appropriate data inputs are included in the forecast.

4.4.3 Forecasting during extreme weather

During extreme weather periods, load is poorly predicted by the model, which tends to underestimate the importance of extreme values of weather variables. Load forecast accuracy for an extreme weather event, comparing the current load forecast model (blue line) and the forecasts of the other trial participants in the 2012 Transpower trial is shown in Figure 3.

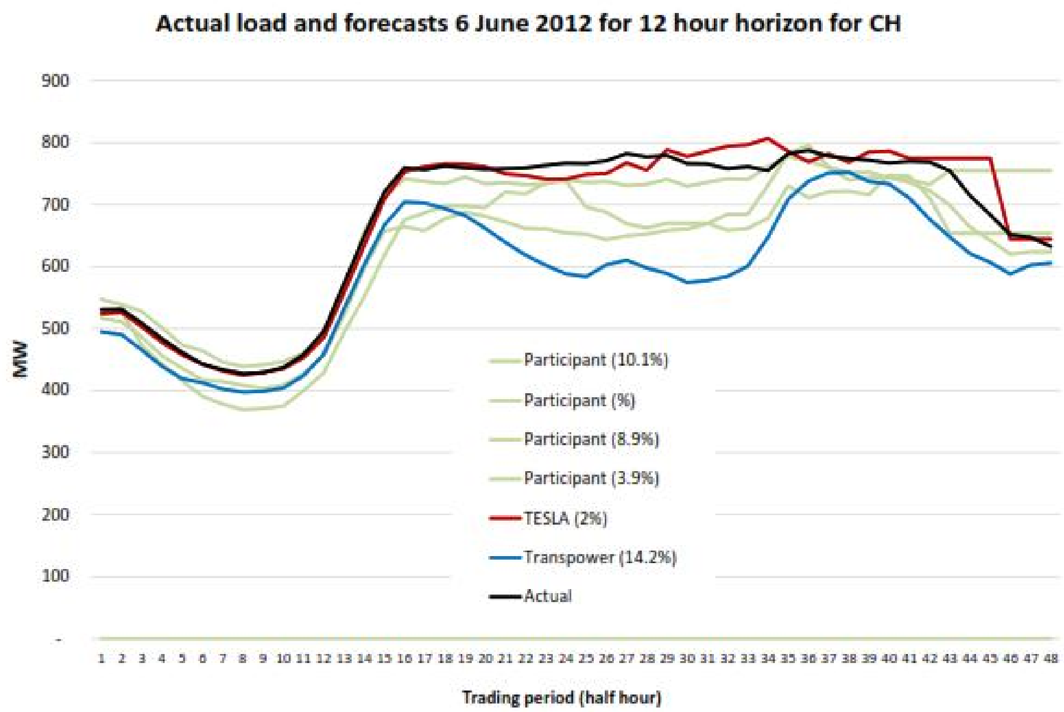


Figure 3 Comparison of trial participants' load forecast accuracy for an extreme weather event, for the Christchurch area

4.4.1 Changing load behaviour over time – years

The model has not been updated to take account of changes in the relationships between load and its determinants since 1996 (i.e. in terms of adjustable parameters relating to the long term, short term, and weather components). Load behaviour has changed significantly in that time, for instance the uptake of heat pumps. Annual updates of the model are advisable[19].

4.4.2 Changing load behaviour over time – seasons

The current forecast struggles with the existence of discrete 'seasons', such as school holidays, network load control periods, irrigation periods, and public holidays. The periods covered by such seasons in combination can be significant, as indicated in Figure 4. Regular updates are advised to capture discrete seasons.

[19] We have only recently discovered we have the ability to update the model ourselves and understood the importance of doing so.

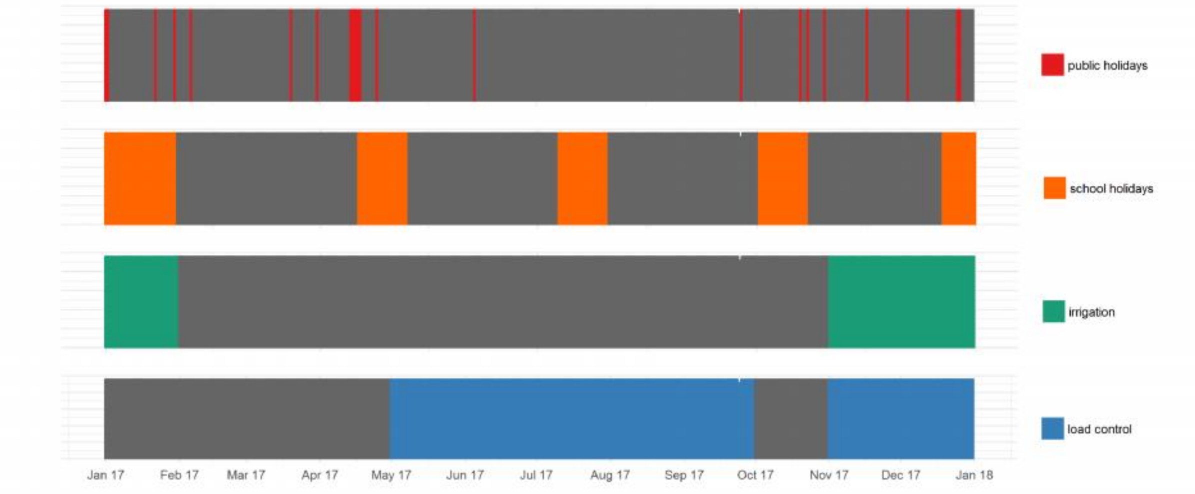


Figure 4 'Seasons' affecting load forecast accuracy

Forecasts are at their least accurate during the first week of these seasons and during the first week after. An example of this behaviour, for school holiday periods, is shown in Table 2. Further data for these seasonal effects is provided in Appendix 2.

Table 2 Mean absolute percentage errors (MAPE) related to school holiday seasons

2017, across NZ	MAPE for 2.5hr horizon
Comparison period	1.6%
School Holidays week 1	2.1%
School Holidays week 2	1.7%
Week after School Holidays	1.8%

The presence of poorly modelled seasons in the historical load data also has a knock-on effect to the forecast accuracy during 'normal' times of the year. The impact of discrete seasons and changing load behaviour over time has been confirmed by a trial re-calibration of the model. This trial put more emphasis on recent rather than long-term historic load as an input. This removed data from less relevant past seasons, improving the overall accuracy (see the Human Intervention discussion in Option 2a assessment within Appendix 3 for more detail).

4.4.1 Changing load behaviour over time – days/weeks

The forecast model does not deal well with large and frequent changes to the shape of the daily load profiles, such as during periods of fluctuating network load control (Figure 5).

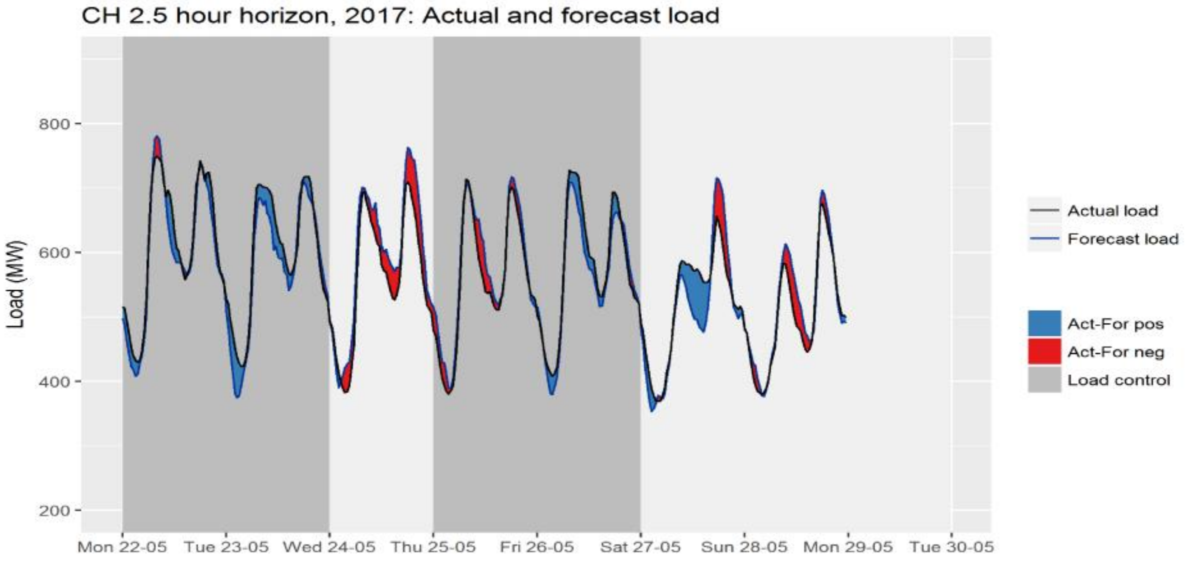


Figure 5 Load forecast response to fluctuating network load control in the Christchurch area, May 2017. Both the actual and forecast load are shown. Differences between these are shaded in blue (where the forecast is too low) and red (where the forecast is too high). Grey background shading shows days where the network company has indicated that network load control occurred.

4.4.2 Temporal load shifts

Redistribution of load during a day is not able to be modelled by the current forecast. For example, network load control, where load is reduced during peaks and brought back on following the peak period. The effect of this is seen in Figure 6, where the network load control (not predicted by the load forecast tool) has led to the load forecast overestimating the load in the two peak periods (red shading) and underestimating the load during the rest of the day (blue shading).

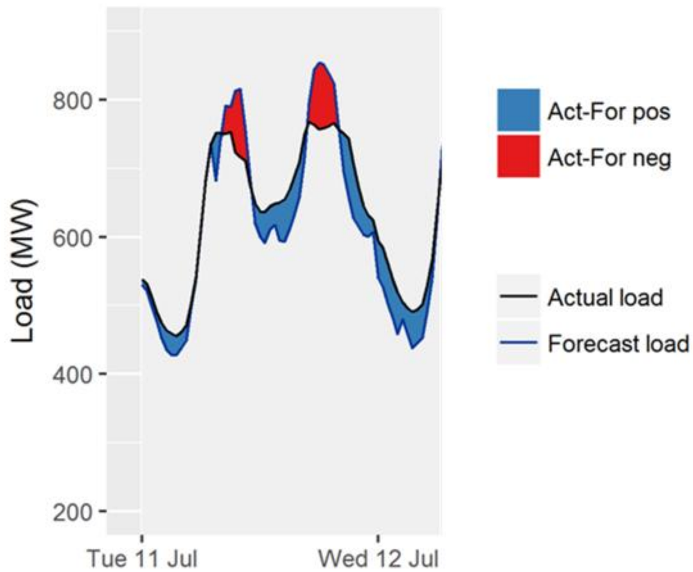


Figure 6 Forecast inaccuracy with a temporal shift in the expected load shape. Data for 11 July 2017, Christchurch area, 24-hour horizon.

4.4.3 Complex relationships between load and its determinants

The current load forecast does not model complex relationships between load and its determinants, leading to a reduced accuracy. This includes limited account of non-linearities and interactions between variables.

4.4.4 Price response

Price response is an issue because price is a determinant of load; participants may at times change their load in response to price. However, unlike other determinants of load, the issue is more complicated because load is also a determinant of price; cost of supply increases with increasing demand. In other words, the relationship between price and load is circular.

The current forecast does not explicitly consider price response.

The forecast load does not account for load response to price, except to the extent that similar response occurred in the historical load data used as an input. Significant price response tends to occur only at very high prices, however. The relative infrequency of such prices means that it is unlikely that similar response occurred in the historical data, therefore the forecast load is not a true post price response load.

The current load forecast does not provide a true pre-response signal to the market either (see section 6.1.1), as historical load data used in the forecast includes participants' load response to price at the time it occurred. Normalising the data to remove participants' response would be necessary for a true pre-response signal.

The same logic applies to network load control and DSI response.

4.4.5 Differences between GXPs

Different load characteristics at GXPs are not modelled independently. These characteristics include residential and commercial behaviour with a variety of daily profiles and different relationships between weather and load. If these were included they would increase the load forecast accuracy.

In the current system, load is forecast at an area level, and distributed to the GXP level using factors based on the previous week's actual load distribution. This approach was chosen historically to allow smoothing out of any irregularities at the GXP level.

4.4.6 Accounting for non-conforming load at a conforming GXP

Non-conforming loads, such as embedded generation or small industrial processes, are one of the types of load behavior which contribute to differences between GXPs, as mentioned above. They may require special consideration and techniques to detect and predict, and can be a relatively large source of forecasting inaccuracy.

4.4.7 Issues specific to the current model – refined component

The load forecast is particularly inaccurate when forecasting future days, as opposed to future periods within the same day. This is because the refined component of the forecast model generally improves forecast accuracy. When the refined component does not apply, forecast accuracy is usually worse.

The refined component can at times significantly reduce load forecast accuracy. It works best if a day's load profile shape is accurately predicted, i.e. a difference between the forecast and actuals now is likely to mean a similar difference later in the day. This is sometimes not the case, particularly during network load control seasons. Further details are given in Appendix 3A.

The refined component also resets to zero at midnight each day. This means the forecast produced at midnight is particularly inaccurate.

4.4.8 Human intervention error

The current forecast tool has poor usability, and is relatively difficult for an operator to understand and interpret. This increases the likelihood of human error.

4.4.9 Forecasting when many factors contribute

The load forecast is particularly poor in times of market stress, due to several of the above factors occurring concurrently. Accurate forecasts are particularly important during these times for a number of reasons; prices are highly sensitive to inaccuracies, slow start generation may be required, and load reduction is being considered for security.

4.4.10 Additional influence

In general, the above issues also corrupt the explanatory power and comparative importance of weather and historic load, thereby reducing forecast accuracy for other times, on top of when they directly apply.

5 Criteria for Option Comparison

We agreed a set of criteria with the Authority to compare the relative performance of the options derived as part of this assessment. These criteria can be split into two sub-categories, Accuracy and Cost, as follows:

- Accuracy of the current model in the near-term*
  - Average accuracy
  - Prevalence of outliers
  - Accuracy during market stress
  - Accuracy of transmission constraints
- Whole of project cost
  - Cost in the near-term (including operation and maintenance as referred to in the SOW)
  - Cost in the future**, required to maintain desired accuracy (this is "future-proof" as mentioned in the SOW)

*near-term describes the period prior to significant changes occurring in the future.

**future refers to the time when new technologies significantly change forecasting issues and solutions, as defined by Future State A and Future State B in section 5.2.2.

Reliability is also a critical requirement to ensure system security. All options must be attain the same high level of reliability. For this reason all options compare equally in term of reliability.

This section explains why each of these criteria are important aspects of the system to address.

5.1 Accuracy of the Current Model

5.1.1 Average accuracy

The more accurate the load forecast is, on average, the better forecast prices will be for the majority of the time.

Forecasts being accurate, on average, is most important at times of high prices, and over all forecast horizons.

The load forecast must be suitably accurate to ensure adequate security assessment by the system operator.

In quantitative assessments, we have generally described average accuracy using mean absolute percentage error (MAPE). Where positive/negative load forecast errors are referred to, these have been defined as actual load – forecast load. For analyses of the 2017 data, forecasting horizons of 2.5 hours and 24 hours have been used.

5.1.2 Prevalence of outliers

Outliers in the load forecast error are abnormally large differences between actual load and forecast load. The fewer outliers a load forecast has, the more likely price forecasts will be more accurate, on average, and the greater confidence participants have in forecast prices.

The relationship between load and price is often complex and non-linear, and is dependent on the steepness of the supply curve. Outliers in the load forecast error can lead to particularly high price forecast errors.

Example: An outlier load forecast value might result when the actual load was in the normal price range, where prices were not particularly sensitive to load change; whereas the load forecast was in the high price range, where prices are very sensitive to load change. This would result in an especially high price forecast error, out of proportion to the load forecast error.

Outliers in top-down forecast approaches which forecast by region can be expected to be more significant in their effect on price than outliers at the GXP level which occur in bottom-up forecast approaches.

It is important to have few outliers in the time horizon when participants need to make a final decision, e.g. the start-up horizon of a slow starting plant.

Outliers may also cause issues with security assessment. The system operator will always account for uncertainty when assessing security, but this is more difficult when there is an outlier as this is beyond the expected uncertainty level. For the purposes of operations, a forecast that has reasonable accuracy on average but achieves few outliers is more important than a forecast that is very accurate on average but has many outliers. This is most important within the 12-hour time horizon, and at the GXP level.

5.1.3 Accuracy in times of market stress

The market is under stress when the ability to meet demand for energy and reserves is low.

In times of market stress, prices are particularly sensitive, with a real possibility of extreme prices occurring which will heavily impact some participants.

Accurate load forecasts are highly important during these times to ensure participants have sufficient information to increase generation or decrease load, enter hedge arrangements and adequately structure their offers.

For several types of participants, the forecasts need to be accurate well ahead of time, e.g. slow start generation, DSI, price responsive demand, IL offers that depend on the availability of price responsive demand, and load reduction being considered for security.

The current load forecast is often particularly poor in times of market stress, due to several factors occurring concurrently. The high quantity and diversity of participant decisions, along with the prevalence of network load control, transmission constraints, and extreme weather conditions, make forecasting load more difficult.

5.1.4 Accuracy of transmission constraints

Transmission constraints can result in price separation between different GXPs. At times this leads to very high positive and negative prices, due to the high spring-washer effect[20].

Gentailers can be adversely impacted when they are short of generation in locations where prices are high but long in generation where prices are low, or vice versa. Any demand-side participant on a high price side could be heavily impacted if they do not receive accurate forecast prices to respond to. Transmission constraints incentivise increases in generation and reductions in load on the high price side, and vice versa on the low-price side. To react to these incentives, however, requires that these are adequately signaled in forecast schedules. This requires accurate load forecasts, particularly at GXP level.

[20] See glossary for definition of high spring-washer effect

5.2 Whole of Project Cost

We have broken down the cost into two timeframes; the near-term and the future.

5.2.1 Cost in the near-term

The best solution is the most cost-effective solution to deliver the required benefits. In order to assess these criteria we look at solution costs attributed to:

- The cost of the forecast tool
- The choice of platform
- Infrastructural issues
- The inputs to the model
- Operation and maintenance costs
- Staff resources

5.2.2 Cost in the future

One of the criteria requirements identified in the SOW was to consider future proofing of the load forecast. Although the requirement is that the load forecast must achieve a required level of accuracy in the future, this can always be achieved, but to do this has an associated cost. We therefore consider future proofing as a cost.

The need to adapt to the future depends largely on what the changes are and when they occur. We have investigated the effect of possible future events by defining two future states:

- Future State A - types of load and generation change, making load forecasting more difficult (costs increase for a given accuracy level).
- Future State B - improvements in data and computing power make forecasting cheaper or better (costs decrease for a given accuracy level, or accuracy increases for a given cost).

We note these are not mutually exclusive but they identify the drivers behind some of the future changes we may expect.

5.2.2.1 Future State A: New load types - solar PV, EVs, smart technology

The cost of load forecasting increases in this state.

Changes and issues characterising Future State A

Solar PV

Solar PV uptake is likely to increase.

Solar PV is already becoming more affordable, this is expected to continue. Solar PV can be highly difficult to forecast, often requiring specialist forecasters. Solar PV creates forecasting issues in accounting for accurate weather data, discrete seasons, changing load behaviour day to day. Uptake is likely to differ significantly between GXPs, depending on sunshine levels and relative affluence.

Uptake of solar PV will likely be at its greatest when the solar PV + battery combination becomes more economical. This creates forecasting issues in accounting for temporal load shifts, and changing behaviour month to month.

A combination of solar PV + hot water heating is potentially already economical, so uptake of this may increase in the nearer future. This also leads to forecasting issues relating to temporal load shift, as with solar PV + batteries (Hot water cylinders are similar to batteries, as both store energy).

Electric vehicles

Electric vehicle usage is likely to increase.

Although the majority of early uptake has been by individuals, the greatest impact on the load forecast is likely to be when there is a notable commercial fleet uptake, potentially creating discrete changes in load behaviour. This could create temporary challenges for the current load forecast as recent history would become less relevant (changing load behaviour over time). When recent load includes the new behaviour, it will be highly predictable by the forecast. Private uptake would not likely cause the same problem as it would be more gradual.

Changes in EV numbers could be large in the one-year timescale used to create the forecast model (necessary to determine relative importance of various inputs). This would limit the forecast accuracy if the model was not regularly updated.

Smart technology

Smart technologies enabling better load control are likely to increase. The greatest effect of this is likely to be increased price-responsive demand. The effect will be greater when used alongside solar PV and EVs.

Household load control to reduce electricity prices is likely to increase in the future to make use of smart technologies. Uptake of these technologies is likely to become more prominent due to increased ease of use, cultural shifts as the idea spreads and products are advertised, cost of products decrease, aggregators offer services either directly to consumers or via retailers who can provide a lower rate to their customers, etc.

Behaviour is likely to change month to month when uptake is at its greatest. But changes in behaviour will continue to occur as people develop new ways of using the systems following changes in technology, new apps, new aggregators, etc.

Relevant forecasting issues include price responsive demand, temporal load shift, changing load behaviour over time.

Increased load control

Incentives for energy consumers to reduce network costs may increase in the future. Future market regulations or arrangements may provide greater mechanisms for network companies to incentivise energy consumers to reduce peak loads, e.g. with time-of-use pricing.

Ahead market

An ahead market may occur in the future. This may change forecast accuracy requirements in the longer forecast horizons, when forecasting is inherently more difficult.

Organisational change

In a world of continual change it is likely that organisational change will occur. This would impact the ability of a forecaster (Transpower or an external vendor) to maintain their practice. There may be a cost associated with this, in providing redundancy of expertise.

Capabilities required in Future State A

This state has greater price responsive demand, more weather-sensitive demand, frequently changing load behaviour, and greater load differences between geographic regions.

This requires

- the ability to accommodate separate solar forecasts
- more than one weather forecast, and the ability to compare them

- variability forecasts
- price response prediction and provision of price elasticity to market system
- learning algorithms
- GXP level forecasts
- human intervention by several staff (in-house, by the vendor or both), including people who understand the load, people who understand weather forecasting, people who understanding forecasting models, 24/7 security operation

We believe that it will be important to incorporate price response in the market system in Future State A (see section 6.1.1). This will be the best mechanism to ensure that price forecast accuracy is maintained in the presence of large amounts of price responsive demand.

5.2.2.2 Future State B: New forecasting abilities - Data, data science, computer power

The cost of forecasting to a given accuracy level decreases in this state relative to that prior (either near-term or Future State A).

Changes and issues characterising Future State B

Improved forecasting techniques/expertise

Data science use in load forecasting is likely to increase in the future, either reducing cost of load forecasting by increasing competition among providers, or increasing ability to forecast accurately as technology improves, or both. Data science technology to improve load forecasts is already here; its use in load forecasting is likely to increase over time as more people gain expertise, and as data science use increases in industry in general.

Increased availability of data

Data relevant to load forecasting may become more readily available in the future. This includes information on types of load at a given meter, retail customers' plan type (e.g. time of use), and more data about embedded solar generation. Weather information may also be more granular, both geographically and temporally. This availability and greater detail will address the issue of input data quality. The data availability will also improve the effectiveness of data science techniques.

As an example, AEMO collects, and is looking to increase the collection of, large amounts of data on solar PV use to inform its forecasts. This includes feedback from solar PV via websites where people can log usage information, more weather metering stations, geostationary satellites to measure cloud cover, information from grid solar farms, etc.

Current data science techniques can be used to identify different types of loads, but require smart meters at the load site to do so. New Zealand does not currently have smart meters at every load point; load forecasting ability will increase when more smart meters are in use.

Regulations protecting private and confidential information may be a constraint on the availability of some types of data to load forecasting providers.

Batteries used for energy arbitrage

Battery usage for energy arbitrage is likely to occur at some point in the future. This is unlikely to occur soon, as it is not currently economical, and is unlikely to occur before battery usage in conjunction with solar PV is more common. Forecasting will be easier if batteries are used extensively for energy arbitrage as the shape of the daily load profile will be relatively flat, hence easy to predict.

Computing power + weather forecasting

Weather forecasting is likely to improve in the future as computing power improves, or cost less for a given accuracy improvement. Computer power is the current limiting factor to increasing the geographical granularity of weather forecasts at a reasonable cost. Aggregating this more highly granular weather forecasts increases the overall accuracy. Increasing the granularity of weather forecasts increases both the load forecasting accuracy of individual loads and at a regional level.

Capabilities required in Future State B

In addition to the implicit data science requirements, this state requires the forecaster to have additional data science expertise to monitor and update the models and clean the data.

6 Option Identification and Initial Assessment

6.1 Factors for Finalising Options

Options differ according to their levels of the various capabilities (section 4.3). The statement of work specified four options for investigation.

- Option 1: Any rapid deployment adjustments ("quick wins") to the current medium-term load forecast system to improve the accuracy of the load forecast.
- Option 2: A refresh of the current medium-term load forecast model inputs using data, keeping model inputs and outputs in the current form (weather forecast input structure and regional load forecast outputs)
- Option 3: A refresh of the current medium-term load forecast model using recent load data and an modernised weather forecast (more granular temporal, geographical forecast inputs) and other inputs to load forecast which are not currently factored in, but preserving the current regional outputs.
- Option 4: A full replacement of the current medium-term load forecast system with more granular inputs and outputs e.g. nodal weather forecast inputs and load forecast outputs, and other inputs to load forecast which are not currently factored in. This option will include:
    - A model where access and rights to the intellectual property is open, transparent and adaptions enabled.
    - A model where access and rights to the intellectual property is closed.

The options above differ according to input and model capabilities, and potentially adaptability.

We have added variations to these options to account for the capability of human intervention.

Sub-options have also been added to accommodate the capabilities of incorporating price response and GXP level forecasts in the market system. These have been included as sub-options as they can be implemented independently and at any time prior to the Future States. However, to simplify the comparison of options in the near term, we will consider these sub-options as being implemented at the same time as any other option.

In determining options, we have also considered the required capability of a secure and reliable system. Specifically, this entails refining the above options to ensure adequate reputation of the model vendor.

In summary, the factors we used to further define the SOW options were:

- Whether price responsiveness is incorporated in the market system
- Whether GXP level forecasts are incorporated in the market system
- Whether human intervention is utilised by the system operator
- Possible strategies to enable adaptability
- Reputation of the forecast model provider

Note: The remaining capability (publication of information) is covered in section 8 and Appendix 3.

In this section we describe the importance of the various factors and assess potential options considering these factors. This is then used to create a shortlist of options for final option comparison.

6.1.1 Price response incorporated in the market system

6.1.1.1 Why is this important and how can it be achieved?

We have already described how price responsiveness is an issue with the current system and how incorporation of price response in the market system is required in Future State A. In this section we focus on price responsiveness in the near-term, and how it relates to the options.

Price responsiveness could either be considered exclusively in the load forecast system, or incorporated in the market system. To explain the difference, we need to consider a few concepts and explain some potential issues.

Pre-response load forecast (LF_{PRE}) – in this example the pre-response load forecast is based on two input factors, weather (W) and historic actuals (H);

$$LF_{PRE} = aW + bH$$

NOTE: the historic actuals have been normalised, i.e. the effect of price response has been taken out of the historic data

Post-response load forecast (LF_{POST}) incorporates the predicted price (P) and the known price-elasticity (e);

$$LF_{POST} = aW + bH + eP$$

If a participant acts on a forecast price from a schedule that used a post-response forecast load, they may invoke a known paradox.

The prediction-signal paradox for price response

Incorporating price response into any load forecast can have perverse impacts. Making a change within the load forecast to take account of price response can mean the resulting forecast price may be insufficient to trigger the predicted response. This can be best illustrated by using an example,

An industrial process is uneconomic once the spot price is above $500. So, if the spot price is $510 it chooses not to operate. This is the “known price response”. Using this knowledge, their load is automatically removed from the load forecast and results in the spot price dropping to $480. As the industrial load requires a few hours to shut down, this is the forecast price which it uses to make its decisions on and therefore as the price is lower than $500 it continues to operate. What this means is that the information used due to a “known price response”, i.e. that the industrial load would choose not to operate, does not occur in reality, leaving the participant exposed to prices it was not willing to pay.

If a participant wants to understand when their inaction may lead to prices they should have acted on, they can compare the post-response load forecast to the pre-response load forecast. i.e. LF_{POST} to LF_{PRE}.

Option X – incorporate price response in the market system - this involves entering the post-response load forecast (LF_{POST}) into the market system as a number of separate pieces of information - pre-response load forecast (LF_{PRE}) and the price-elasticity (e). As the pre-response load forecast is calculated as a separate value, the participants can be provided with two price signals, e.g. the NRS could send the pre-response signal, while the PRS sent the post-response signal. Option X estimates the post-response load in the market system.

Option Y – price response is not incorporated in the market system, only the post-response load forecast (LF_{POST}) is entered in the market system. Option Y estimates the post-response load in the load forecast tool using an estimate of price (P) based on the price in the previous schedule.

The differences between these two options are:

- Only Option X allows for pre-response and post-response signals. This enables the NRS schedules to remain conservative for security assessment, as well as providing potentially important information to market participants.
- Option Y costs less than Option X
- Option X is more accurate than Option Y because
  - SPD, used in Option X, is better placed to estimate the effect of demand-price elasticity compared to a forecast system
  - For Option Y, price response updates due to changes in other inputs such as generation offers would not be provided until the following schedule was run
  - If the pre-response signal is not provided, as in Option Y, the price response may not occur. It is worth noting, however, that there is potential for de-stabilising action if pre- and post-response signals are sent; if all participants attempted to 'free-ride' on other participants' predicted response, for example, that response wouldn't occur.

6.1.1.2 What does it mean for the options?

Improvements to the current forecasting tool – Options 1-3: These are not able to incorporate price response in the market system. Price can be used as an input, but there is no mechanism to discover and export price response predictions from the model's inputs.

Options that require a new load forecast tool – Option 4 variations: Price response in the market system can be considered as an add-on to a base-level forecast system. Changes to the load forecasting system to incorporate this functionality are largely independent of other changes required to improve the load forecast system, therefore if the new system is in place the benefits and costs of incorporating price-responsiveness can be treated as self-contained.

6.1.2 GXP level forecasts incorporated in the market system

6.1.2.1 Why is this important and how can it be achieved?

The importance of GXP level forecasts at the individual GXP level and at an area level have been discussed in Section 4.4.5.

Ideally, GXP level forecasts could be incorporated in the market system directly as GXP loads in ahead schedules. Alternatively, the forecast system could provide GXP level forecasts, which would then be aggregated up to area level, and then apportioned to the GXP level using the current mechanism.

Incorporating GXP level forecasts in the market system allows for greater accuracy at the GXP level, leading to more accurate forecast prices, particularly when there is some likelihood that transmission constraints will bind.

6.1.2.2 What does it mean for the options?

Improvements to the current forecasting system – Options 1-3: These options cannot incorporate GXP level forecasts in the market system.

The cost of GXP level forecasting in the current system will outweigh any benefits regardless of the level of improvements in the current tool. Most of the benefit from GXP level forecasting comes from having unique models to match the characteristics of a GXP. The current tool is limited in its flexibility to test different models due to the time it takes to set up new models and limited visibility of model performance. GXP level forecasting within the current tool is particularly impracticable when the models need to be updated regularly to account for discrete seasonal effects and network load control, as is currently recommended at the area level. Time taken for adjustment of historic data along with these other limitations will be multiplied by the number of updates required.

Options that require a new load forecast system – Option 4 variations: Incorporating GXP level forecasting in the market system can be considered separately, as an add-on to a base-level forecast system. Changes to the load forecasting system to incorporate this functionality are largely independent of other changes required to improve the load forecast system, therefore if the new system is in place the benefits and costs of including GXP level forecasting can be treated as self-contained.

A variation to Option 4, Option 4c (see Section 6.2.2.3), considers the current TESLA model residing at ems being plugged into the current system, with the current model used as a backup (should the link to the TESLA functionality be unavailable). This provides an aggregated area level forecast made up from individual GXP level forecasts. Cost of new infrastructure would be minimised by piggy-backing off the existing infrastructure, and set-up and testing would have largely already been done. Note: this plug-in option does not provide forecasting to the market system at the GXP level, therefore feeding the ems TESLA forecast into a separate instance of our current tool can be used only while an area level forecast is workable.

6.1.3 Human intervention in-house

There are various levels of human intervention that can be applied to a forecast. The degree of intervention is categorised as “some” (low to medium effort), “high” and “significant”.

6.1.3.1 Why is this important and how can it be achieved?

Any accurate and reliable forecast requires at least some human intervention, as we learned in the information gathering stage of this assessment. It is widely used by other system operators and recommended by leading load forecast providers.

The cost is related to the amount of intervention required and will depend on how much improvement is desired. But there will be an amount of intervention after which the marginal gain will rapidly decline.

Where the intervention requires some expertise beyond procedural work, there would be a risk that the knowledge may be retained by a small number of staff, making a company vulnerable to staff departures. This in turn could lead to loss of accuracy in the forecast. Additional staff in the short term and a succession planning policy are important to ensure accuracy is maintained.

The tools and interfaces available to the intervener and a good understanding of the models minimise human error. A tool that is highly adaptable to an expert intervener reduces the likelihood of errors and makes errors easier to identify. For instance, an adjusted forecast could be easily tested with comparisons to forecasts from other and previous models by creating a visual output of the data.

Human intervention can be performed in-house, to varying degrees, or by the load forecast vendor.

6.1.3.2 What does it mean for the options?

Most options require human intervention to some degree.

Improvements to the current forecasting system – Options 1-3: Some level of intervention in the current system would be worth pursuing. However, intervention in the current tool is not highly efficient as the tool has limited adaptability and interveners have a limited view of the underlying algorithms. Achievable accuracy is limited and additional staff are required to mitigate the risk of human error.

Significant levels of intervention in the current system would take on increased complexity, requiring specialist expertise, beyond procedural work. To mitigate the risk of losing this specialist IP, 'over-recruitment' of staff may be necessary to provide redundancy. Therefore, the option of the current forecast with new inputs, model update, and significant human intervention has been excluded as an option because a new, open tool, option is likely both cheaper and more accurate.

Options that require a new load forecast system – Option 4 variations: Some, or high, intervention for a new open system would be worth pursuing. High intervention would be more workable than with the current tool, as open options are more flexible. The open model options with significant intervention are not workable. Significant levels of intervention in the open options, to the point where costs are equal to that of the closed options, are not likely to achieve significant accuracy above the closed option.

Closed options incorporate intervention by the vendor. For this reason additional intervention does not need to be considered.

6.1.4 System adaptability

6.1.4.1 Why is this important and how can it be achieved?

Change is a constant in modern life and systems need to adapt to the changes. It is therefore important that this is a feature of any of the options for the future.

The SOW refers to models where access and rights to the intellectual property are open and transparent and adaptions are enabled, and compares this to models where access and rights to the intellectual property is closed. It is important to distinguish between closed models and closed IP. The implication is that a closed model is rigid, with little ability to adapt. This could be the case of some closed-IP models, e.g. a 'black box' that is installed and then left to run without intervention, however, closed models can also be highly adaptive and adapted by load forecast professionals as the TESLA business model has demonstrated.

The risk of the closed system used in the near term is that it may not adapt to changes in the future. Additional costs will be required if another new tool is required in the future, and costs sunk into the original system may have been wasted. This risk depends on more than just the closed nature of the model; IST infrastructure options affect the costs of accommodating another new tool, contractual options and vendor choice affect the amount of sunk costs and the likelihood the vendor will adapt to the future.

6.1.4.2 What does it mean for the options?

Improvements to the current forecasting system – Options 1-3: These are not able to adapt as required, as their models are insufficiently sophisticated, and tools are too inflexible, to incorporate price response and GXP level forecasting in the market system.

Options that require a new load forecast system – Option 4 variations: Both the open and closed models can achieve system adaptability. Open systems are adaptable by definition. For closed models, contract specifications with the vendor ensuring accuracy is maintained to a certain level, with associated penalty or termination clauses, could significantly decrease any rigidity. This would enable near-term accuracy to be maintained without sacrificing future accuracy. It is likely that the vendor would have sufficient incentive to maintain their forecasting ability in order to maintain their practice. A closed system will be cost-effective in the future if no cheaper or better forecast systems become available before its payback period. Short renewal periods for the contract would minimise sunk costs, ensuring a closed system is cost-effective in the future.

6.1.5 Reputation of Provider

6.1.5.1 Why is it important?

The reputation of a load forecast provider informs us about the reliability of their model (a critical requirement for system operations).

6.1.5.2 What does it mean for the options?

All options need to be highly reliable.

Improvements to the current forecasting system – Options 1-3: The current system is already highly reliable.

Options that require a new load forecast system – Option 4 variations: We are not aware of, and do not believe it is likely that, there are any models with open-IP in current use by a transmission system operator. There would therefore be no reputation by which to gauge the reliability of such a model. The open models we have considered are therefore limited to options which are relatively open to the system operator only, such that they can be easily adapted and understood by in-house interveners. These open models, e.g. as provided by the vendor iTron, or the vendor used by PJM, have a good reputation.

TESLA is the example of a closed-IP model we have considered and they have a very good reputation.

6.2 Selected Options for Analysis

We selected nine final options for comparison, as well as sub-options for incorporating price response and GXP level forecasts in the market system. Table 3 below summarises how each factor was applied to the four SOW options.

For options that require a new load forecast system – Option 4 variations – the tools differ significantly depending on whether or not GXP level forecasts or price responsiveness are incorporated in the market system. Incorporation of GXP level forecasts or price response in the market system are therefore considered additions to a ‘base level’ forecast system.

Options for improving the current load forecast are only workable if forecasts remain at the area level and price response is not incorporated in the market system in the near-term. Incorporation of GXP level forecasts and price response in the market system would be expected to occur when a new system is implemented to accommodate Future State A.

Table 3. Selected options compared by factor.

Options		Price Response	GXP level forecast	Human intervention	System adaptability	Reputation of provider
Improve the current model	**1** Quick wins			√		√
	2 Refresh current model					√
	2a Refresh current model + some intervention			√		√
	3 Refresh current model + better inputs					√
	3a Refresh current model + better inputs + some intervention			√		√
New system required	**4a** New system, open + some intervention	√	√	√	√	√***
	4b New system, open + high intervention	√	√	√	√	√***
	4c New system, via new model, closed (TESLA plug-in with current tool as backup)	√	√	√*	√**	√
	4d New system, closed (e.g. TESLA in-house)	√	√	√*	√**	√

Key:

* intervention at vendor

** possible via regular contract renewal and accuracy clauses

*** provided the option is open to the system operator only (not open-source)

Excluded options:

- Improve current model + high or significant intervention
- New system, open + significant intervention

6.2.1 Improve the current model

6.2.1.1 Option 1 (SOW Option 1): Quick wins

Option 1 in the SOW specified "Any rapid deployment adjustments to the current medium-term load forecast system to improve the accuracy of the load forecasts". Several rapid deployment adjustments (quick wins) have been identified and included in this option:

- Increase the precision of the temperature inputs to 1 dp.
- Get up to date population weightings of different temperature measurements within a forecast area.
- Provide greater temporal granularity of weather forecasts.
- Perform current intervention more diligently.

6.2.1.2 Option 2 (SOW Option 2): Refresh current model

Option 2 in the SOW specified "A refresh of the current medium-term load forecast model inputs using data, keeping model inputs and outputs in the current form (weather forecast input structure and regional load forecast outputs)"

For this option, the quick wins in Option 1 are assumed to have been applied.

The model is refreshed by changing the current input parameters and re-running. Several parameters can be changed in the model, affecting how the various components are treated. There are potentially limitations on our ability to disclose these parameters due to copyright restrictions.

6.2.1.3 Option 2a (variation of SOW Option 2): Refresh current model + some intervention

This is a variation on SOW Option 2.

The major change in Option 2a is that this option includes some human intervention in addition to the aspects defined in Option 2. This may include updating the model regularly for different seasons, adjusting historic data, modifications for days adjacent to public holidays and extreme weather events, and estimation and adjustment for load control.

6.2.1.4 Option 3 (SOW Option 3): Refresh current model + better data inputs

Option 3 in the SOW specified “A refresh of the current medium-term load forecast model using recent load data and a modernised weather forecast (more granular temporal, geographical forecast inputs) and other inputs to load forecast which are not currently factored in, but preserving the current regional outputs”

For this option, the quick wins in Option 1 are assumed to have been applied. The model recalibration of Option 2 is also assumed.

New weather data can be acquired and added to the system, including cloud cover, humidity, wind direction, sunset time, and rainfall. These can then be configured to achieve the appropriate relationship with load for the model. Other inputs can also be added, to accommodate school holidays, or price, for instance.

6.2.1.5 Option 3a (variation of SOW Option 3): Refresh current model + better data inputs + some intervention

This is a variation on SOW Option 3.

The major change for Option 3a is that this option includes some human intervention in addition to the aspects defined in Option 3. Intervention would be similar to that in Option 2a.

6.2.2 New system required

6.2.2.1 Option 4a (variation on the first of the SOW Option 4 alternatives): New system, open + some intervention

This is a variation on the first of the two SOW Option 4 alternatives.

This option specified “A full replacement of the of the current medium-term load forecast system with more granular inputs and outputs e.g. nodal weather forecast inputs and load forecast outputs, and other inputs to load forecast which are not currently factored in. This option will include: A model where access and rights to the Intellectual property is open, transparent and adaptions enabled”.

The open model we consider does not have fully open IP. It is, however, relatively open to the system operator, such that it can be easily adapted and understood by in-house interveners. Some intervention is required.

Sub-options include base level, or GXP level forecasting and/or price response incorporated in the market system.

6.2.2.2 Option 4b (variation on the first of the SOW Option 4 alternatives): New system, open + high intervention

This is a variation on the first of the two SOW Option 4 alternatives.

The major change in Option 4b compared to Option 4a is that this option also includes greater (high) human intervention.

Sub-options include base level, or GXP level forecasting and/or price response incorporated in the market system.

6.2.2.3 Option 4c (variation on the second of the SOW Option 4 alternatives): New system, via new model, closed (e.g. TESLA plug-in with current model as backup)

This is a variation on the second of the two SOW Option 4 alternatives.

This option specified “A full replacement of the of the current medium-term load forecast system with more granular inputs and outputs e.g. nodal weather forecast inputs and load forecast outputs, and other inputs to load forecast which are not currently factored in. This option will include: A model where access and rights to the Intellectual property is closed”.

As mentioned earlier in the section, Option 4c utilises the current TESLA model residing at ems could plugged into the current tool, with the current tool used as a backup. Cost of new infrastructure would be minimised by piggy-backing off the existing infrastructure, and set-up and testing would have largely been already done. Feeding the ems TESLA forecast into a separate instance of our current tool while using the current tool as a backup is only workable if incorporating GXP level forecasting in the market system is not required.

This option variant is adaptable, as appropriate contract specifications are assumed possible.

In addition to new weather data and other inputs applied in Option 1 and 3, this option includes the ability to utilise separate models using different weather providers, more sophisticated models and better human intervention, and forecasting to the GXP level.

Sub-options include base level, or price response incorporated in the market system.

6.2.2.4 Option 4d (variation on the second of the SOW Option 4 alternatives): New system, closed (e.g. TESLA in house)

This is a variation on the second of the two SOW Option 4 alternatives.

This Option 4d involves an external provider, e.g. TESLA

The major change to Option 4c is that this option includes an external provider’s model housed inside IST infrastructure within Transpower as opposed to a model residing at ems which could be plugged into the current tool.

We have only carried out analysis on the sub-options GXP level forecasting, and GXP level forecasting + price response. We have not carried out analysis for the base or price response sub-options, as for these, Option 4d will deliver only similar accuracy to Option 4c, but at a higher cost.

7 Option Analysis

This section summarises the option assessment for the nine final options.

Section 7.1 compares the accuracy of each of the options according to how well they address the issues raised in Section 4.4.

Section 7.2 compares the whole of life cost criteria for each of the options against the cost criteria in Section 5.2.

Sections 7.3 compares the options against the full set of criteria in Section 5, including a summary for each option.

7.1 Assessment of Accuracy in Terms of Issues

Accuracy in the near-term differs between options depending on how well each option addresses the current issues with forecasting load. These differences are illustrated in Table 4 where options are ranked by each issue (1 is highest). Key points to note are:

- New systems address all issues to some extent. We expect the new in-house closed system option (Option 4d) with price responsiveness and GXP level forecasting incorporated in the market system, to address all issues better than, or at least as well as, any other option.
- Input data quality would be improved in all options, though more-so in new system options that can access a range of weather sources. The best use of multiple weather sources could be achieved by systems with high intervention, either in-house or at vendor, i.e. the closed options and open options with high intervention. This is because staff would be available to compare separate models - using separate weather sources - for each GXP.
- Appropriate inputs would be used in the two ‘better inputs’ current model options as well as all new system options.
- Annual changes in load behaviour will be addressed by all but the quick wins option.
- Systems incorporating price response in the market system best deal with the issue of price response. Non-price response Option 4 variants, as well as Options 3 and 3a, can also deal with price response by using prices from previous forecast schedules as inputs. This would be less accurate, as discussed in section 6.1.1.
- New systems can forecast at the GXP level, addressing the issue of differences between GXPs, and are able to better account for non-conforming load at a conforming GXP. The TESLA model has proven particularly competent in these areas. The issue of differences between GXPs is best addressed by options that incorporate GXP level forecasting in the market system; this is because forecast prices are more accurate, particularly when transmission constraints are binding.
- New systems address the issue of human intervention error; closed models by using trained and experienced staff at the vendor, open options by using models that are understandable to the intervener (not a ‘black box’) and tools that facilitate investigation and resolution of high forecast errors.
- The following issues are addressed by options with human intervention in the current forecast, but to a greater extent by the new model system options, due to a combination of more sophisticated models and more effective human intervention:
    - Forecasting during extreme weather. Some improvement is expected from intervention in the current system, particularly if better inputs are also used.
    - Seasonal changes in load behaviour. This is also partly addressed by the current model options with better inputs.

- Daily/weekly changes in load behaviour. Options with new models can employ artificial intelligence to achieve significant improvement over current model options.
- Temporal load shifts.

Table 4 Assessment of options according to accuracy achieved by addressing issues. Accuracy rankings are given, where 1 is the highest.

Options		**Accuracy by Issues**												
		input data quality	appropriate inputs	forecasting during extreme weather	changing load behaviour over time - years	changing load behaviour over time – seasons	changing load behaviour over time – days/weeks	temporal load shifts	complex relationships between load and its determinants	price response	differences between GXPs	account for non-conforming load at a conforming GXP	issues specific to current model – refined component	prevent human intervention error
Current model														
1 Quick wins		3												
2 Refresh current model		3			1									
2a Refresh current model + some intervention		3		3	1	3	2	2					1	
3 Refresh current model + better inputs		3	1	2	1	4				2				
3a Refresh current model + better inputs + some intervention		3	1	2	1	2	2	2		2			1	
4a New system, open + some intervention	base	2	1	1	1	1	1	1	1	2	4	2	1	1
	+ Price response	2	1	1	1	1	1	1	1	1	4	2	1	1
	+ GXP	2	1	1	1	1	1	1	1	2	2	2	1	1
	+ GXP + price response	2	1	1	1	1	1	1	1	1	2	2	1	1
4b New system, open + high intervention	base	1	1	1	1	1	1	1	1	2	3	2	1	1
	+ Price response	1	1	1	1	1	1	1	1	1	3	2	1	1
	+ GXP	1	1	1	1	1	1	1	1	2	2	2	1	1
	+ GXP + price response	1	1	1	1	1	1	1	1	1	2	2	1	1
4c New system, via new model, closed (TESLA plug-in with current model as backup)	base	1	1	1	1	1	1	1	1	2	2	1	1	1
	+ Price response	1	1	1	1	1	1	1	1	1	2	1	1	1
4d New system, closed (e.g. TESLA in-house)	+ GXP	1	1	1	1	1	1	1	1	2	1	1	1	1
	+ GXP + price response	1	1	1	1	1	1	1	1	1	1	1	1	1

7.2 Assessment of Whole of Life Cost

Costs to implement an improved load forecast differ between options depending on up-front capital costs, ongoing costs in the near-term, and future capital costs for infrastructure that enables desired accuracy to be met in the Future States. These costs are illustrated in Table 5. Key points to note are:

- Costs that are the same between options are not considered here. These include any ongoing costs in the future after the initial capital outlay to set up the required Future State infrastructure, and cost of a new model to accommodate Future State B.

- For all options, there are infrastructural costs that will be required either upon implementation (in the near term) or in the future. These have three key components:
    - Base level infrastructure to house a new model
    - Infrastructure to incorporate GXP level forecasting
    - Infrastructure to incorporate price response
- Options differ depending on when each component of the future infrastructure outlay occurs, either upon implementation of the option (in the near term), or in the future. All options that improve the current forecast, as well as options with TESLA plugged into the existing infrastructure, incur costs for all three components in the future. Options with full replacement systems will incur an initial (near term) cost for the base level infrastructure. Other cost components will be incurred either in the near-term or in the future depending on the option.

- For some options, there are additional capital costs.
    - For current model options, there are initial outlays including the (minimal) cost of quick wins and, except for the quick wins option, refreshing the model. Future outlay is also required for a new model and tools with the current model options.
    - For new system open options, initial outlay is required for a new model and tools.
- For most options, there are ongoing costs.
    - Depending on option, these may include costs for procuring better input data, or operational costs associated with intervention. For new system closed options, there are ongoing costs for vendor fees. Note, in Table 5 “new model and tools” does not apply to closed options as costs for a new closed model are categorised as “ongoing costs”.

Table 5 Assessment of options according to whole of life cost

Options		Costs														
		capital costs specific to some options	quick wins capital	refresh model capital	new model & tools	**capital costs for all options**	new in-house infrastructure	incorporation of price response in MS	incorporation of GXP forecasts in MS	**ongoing costs**	inputs	low operational costs	medium operational costs	high operational costs	new closed-IP model
Current model					▦		▦	▦	▦						
1 Quick wins			■		▦		▦	▦	▦						
2 Refresh current model			■	■	▦		▦	▦	▦						
2a Refresh current model + some intervention			■	■	▦		▦	▦	▦			■			
3 Refresh current model + better inputs			■	■	▦		▦	▦	▦		■				
3a Refresh current model + better inputs + some intervention			■	■	▦		▦	▦	▦		■	■			
4a New system, open + some intervention	base				■		■	▦	▦		■		■		
	+ Price response				■		■	■	▦		■		■		
	+GXP				■		■	▦	■		■		■		
	+ GXP + price response				■		■	■	■		■		■		
4b New system, open + high intervention	base				■		■	▦	▦		■			■	
	+ Price response				■		■	■	▦		■			■	
	+GXP				■		■	▦	■		■			■	
	+ GXP + price response				■		■	■	■		■			■	
4c New system, via new model, closed (TESLA plug-in with current model as backup)	base						▦	▦	▦		■	■			■
	+ Price response						▦	■	▦		■	■			■
4d New system, closed (e.g. TESLA in-house)	+GXP						■	▦	■		■	■			■
	+ GXP + price response						■	■	■		■	■			■

Key: Coloured squares indicate costs which are applicable to each option.
Checkered squares represent capital spend items required in the future.

7.3 Assessment Against Criteria

This section is a summary of the assessment of each option against the criteria. A full option assessment is in Appendix 3. An overview of the assessment is provided in Table 6 followed by an evaluation for each option.

The table is to be interpreted as follows:

- Lower values are better for all criteria. Values represent estimates of relativities by subject matter experts in Transpower, and should only be considered indicative. All values have been normalised such that the worst value for each criterion is 20 (lowest accuracy, highest cost) and the best is 0.
- A low accuracy number is better than a high accuracy number as accuracy in this sense represents “degree of error”.
- Accuracy (degree of error) differences were estimated using a mixture of qualitative and quantitative reasoning, according to the differences between options discussed throughout this section and in the appendices.
- For average accuracy (degree of error),
    - differences between the current forecast and the TESLA forecast are assumed to represent the range of improvements possible (Appendix 1 demonstrates some of these differences). This is a reasonable assumption given the clear difference between TESLA and other participants in the 2012 Transpower trial.
    - differences between the other options are less clear, as evidence was often not definitive. Some evidence suggested improved weather data would marginally improve average accuracy, other evidence suggested the potential was for large improvements. Trial model runs targeting specific seasons resulted in significantly lower errors for the targeted months, but the error for a full year was similar to the current model, and the errors seen on the targeted months were not inconsistent with typical variation month to month.
- We assume short contract renewal periods for the closed system options, therefore there are no sunk costs unnecessarily incurred when Future State B occurs. The criteria ‘cost in the future’ therefore represents additional costs required in Future State A only.
- Area-level outliers will be less prevalent for the bottom-up GXP level approach (available in new model options) when compared to a top-down approach (such as the current model). This is because of cancelling out effects which occur when aggregating positive and negative errors.
- For the purposes of costing, it is assumed Future State A or B will occur in 10 years[21]. The costs in the table represent discounted costs for the near-term period (over 10 years). The costs shown for Future State A or B represent the one-off cost at that time, discounted for the elapsed 10 years to provide compatibility with the other costs.
- All costs are discounted at 5% cost of capital.

[21] This period is only used for the purpose of calculating costs, and is not our estimation of actual timing of the Future States. Prediction of such timing does not form part of this assessment.

Table 6 Overall assessment of options against the accuracy (degree of error) and cost criteria. Values assigned range from 0 (best) to 20 (worst).

Options		accuracy (degree of error)				cost		
		average	outliers	stressed	constraints	in near-term	in future	discounted total
Current model		20	20	20	20	0	4	4
1 Quick wins		19	19	20	20	0	4	4
2 Refresh current model		19	19	20	20	0	4	4
2a Refresh current model + some intervention		16	10	18	19	2	4	5
3 Refresh current model + better inputs		18	12	19	19	3	4	7
3a Refresh current model + better inputs + some intervention		13	6	16	18	7	4	10
4a New system, open + some intervention	base	7	3	14	13	9	2	11
	+ Price response	7	3	5	13	14	1	15
	+ GXP	7	3	8	3	11	1	12
	+ GXP + price response	7	3	5	3	14	0	14
4b New system, open + high intervention	base	4	2	13	13	9	2	11
	+ Price response	4	2	3	13	14	1	15
	+ GXP	4	2	6	3	11	1	12
	+ GXP + price response	3	2	3	3	16	0	16
4c New system, via new model, closed (eg TESLA plug-in with current tool as backup)	base	0	2	12	10	15	3	18
	+ Price response	0	2	4	10	17	2	19
4d New system, closed (e..g TESLA in-house)	+ GXP	0	1	5	0	18	1	19
	+ GXP + price response	0	0	0	0	20	0	20

Option 1: Quick wins

Applying the quick wins is likely to bring some improvement in average accuracy and number of outliers, by helping address the issue of input data quality.

This option is likely to incur minimal costs in the near-term.

Costs for a new system will be required in the Future States, including infrastructure for the base level system, and incorporating price response and GXP level forecasting in the market system.

Option 2: Refresh current model

For this option, the Option 1 changes are assumed to have been applied.

Refreshing and re-calibrating the model is very likely to improve average accuracy and prevalence of outliers, by addressing the issue of changing load behaviour over several years.

This option incurs the costs of the quick wins option, both near-term and Future State costs.

In the near-term, small additional costs will be required for staff to test several models prior to implementing a model improvement.

Option 2a: Refresh current model + some intervention

This option assumes changes in Option 1 and 2 have been implemented.

Human intervention is expected to provide large improvements in reducing the prevalence of outliers and reasonable improvement in average accuracy, with some improvement during stressed conditions[22].

Human intervention is expected to improve forecasts by better handling changing load behavior over time, on both a seasonal and a weekly/daily basis, as well as temporal load shifts. This is largely due to intervention for school holidays and network load control.

This option incurs the same costs as Option 2, both near-term and Future State costs. In the near-term, this option incurs additional ongoing costs for human intervention.

Option 3: Refresh current model + better data inputs

This option assumes the changes in Option 1 have been made and the model has been re-calibrated.

Better inputs include new types of inputs and more frequent receipt of current weather forecast inputs. Both of these are likely to improve forecast accuracy, in terms of average accuracy and prevalence of outliers, by helping address the issues of input data quality and appropriate inputs.

This option incurs the costs of Option 2, both near-term and Future State costs. In the near-term, some additional ongoing costs will be required to receive new inputs.

Option 3a: Refresh current model + better data inputs + some intervention

This is a combination of Options 2a and 3, but in terms of accuracy Option 3a has accuracy greater than simply summing the accuracy of the individual options.

Some of the new inputs are more applicable to different seasons, for example irrigation season, therefore regularly changing the model will allow for greater improvements with more inputs.

Any intervention that requires estimation, for example load control, is less likely to be in error with a more fully specified model.

In the near-term, this option incurs the costs of Option 2a plus the additional cost of Option 3. Future State costs will be the same as for all previous options.

Option 4a: New system, open + some intervention

Base level

This option is expected to deliver the same improvements achieved with the above options, but to a greater extent. Any improvements due to intervention in the current model would instead be largely achieved by more sophisticated models, e.g. artificial intelligence.

This option has more sophisticated models and better human intervention than the above options. Human intervention would be more effective in this case because the tool is more flexible and understandable to the intervener.

Accuracy in every respect would be increased, especially during stressed market conditions. All the identified issues are addressed by a new system to some degree.

In the near-term, this option incurs the cost of new base level infrastructure, plus the same cost of new inputs as Options 3 and 3a, some extra human intervention resources compared to Option 3a, and the cost of new models and tools.

---

[22] Price accuracy during stressed conditions requires the accuracy of inputs to be very high.

Costs for infrastructure to incorporate GXP level forecasting and price response in the market system will be required in the Future States.

Price response incorporated in market system

Accuracy during stressed market conditions will be further improved by incorporating price response in the market system in addition to the base level system.

Costs for this option will be the same as for the base level option, noting that costs for infrastructure to incorporate price response in the market system will be incurred at the time the option is implemented (in the near term), not in the Future States.

GXP level forecasts incorporated in market system

Price forecast accuracy, compared to the base level system, will be improved by incorporating GXP level forecasting in the market system. In particular, it will improve the accuracy of constraints and high spring-washer effects.

Costs for this option will be the same as for the base level option, noting that costs for infrastructure to incorporate GXP level forecasts in the market system will be incurred at the time the option is implemented (in the near term), not in the Future States.

Price response and GXP level forecasts incorporated in the market system

This option provides accuracy improvements from both price response and GXP level forecasting being incorporated in the market system.

Costs for this option will be the same as for the base level option except costs for infrastructure to incorporate GXP level forecasts and price response in the market system will be incurred at the time this option is implemented (in the near term), not in the Future States.

Option 4b: New system, open + high intervention

This option is the same as Option 4a except with increased intervention, providing greater accuracy at each level (base, price, GXP incorporated), particularly during stressed market conditions.

The cost for this option is the same as Option 4a with additional ongoing costs for extra intervention in the near-term.

Option 4c: New system, via new model, closed (TESLA plug-in with current model as backup)

This option is only appropriate for base level forecasts and incorporating price response in the market system, not for incorporating GXP level forecasts.

Base level

We expect improved accuracy compared to the other base level open options (Options 4a and 4b), mainly due to better addressing of the issues of:

- differences between GXPs
- identifying and predicting non-conforming type loads and embedded generation at the conforming GXP, using a combination of sophisticated models and human intervention
- input data quality, by using multiple weather sources

In the near-term, this option incurs small set-up costs, plus the same cost of new inputs as Options 3 and 3a, similar cost for human intervention resources as Option 3a, plus high ongoing vendor fees.

In the Future States, costs for new base level infrastructure, plus incorporation of GXP level forecasting and price response in the market system will be required.

Price response incorporated in market system

Accuracy during stressed market conditions will be further improved by incorporating price response in the market system in addition to the base level system.

The cost for this option will be the same as for the base level option, except that costs for infrastructure to incorporate price response in the market system will be incurred at the time the option is implemented, not in the Future States.

Option 4d: New system, closed (e.g. TESLA in-house)

Base level*

This option is similar to Option 4c at the base level, with slightly greater accuracy as in-house models can receive feedback from the market system for demand response instructions, outages, and load transfers between GXPs.

Average accuracy and prevalence of outliers improve compared to any level of the open options (Options 4a and 4b), but:

- open systems with price response incorporated in the market system are more accurate during stressed market conditions than base-level closed systems
- open systems with GXP level forecasting incorporated in the market system are more accurate in forecasting transmission constraints than base-level closed systems.

Costs for this option are similar to Option 4c except that base level infrastructure costs are incurred at the time of implementation instead of in the Future States.

*Note, this option has not been included for final options assessment as it is assumed the benefits due to increased accuracy over option 4c at base level, in terms of our criteria, will not exceed the increased cost. It has been included here for completeness as it may be viable for Transpower to use this option for DSI, which would require a highly accurate in-house model.

Price response incorporated in the market system**

Incorporating price response in the market system in addition to the base level system will improve accuracy during stressed conditions compared to all other options.

Costs for this option will be the same as for the base level option, noting that costs for infrastructure to incorporate price response in the market system will be incurred at the time the option is implemented (in the near term), not in the Future States.

** Note, this option has not been included for final options assessment as it is assumed the benefits due to increased accuracy over option 4c with price response, in terms of our criteria, will not exceed the increased cost. It has been included here for completeness as it may be viable for Transpower to use this option for DSI, which would require a highly accurate, in-house model.

GXP level forecasts incorporated in market system

Incorporating GXP level forecasts in the market system in addition to the base level system will improve accuracy of constraints compared to all other options. This option can make especially good use of GXP level forecasts incorporated in the market system, demonstrated by TESLA's outstanding accuracy at the GXP level in the trial.

Costs for this option will be the same as for the base level option, noting that costs for infrastructure to incorporate GXP level forecasts in the market system will be incurred at the time the option is implemented (in the near term), not in the Future States.

Price Response and GXP level forecasts incorporated in the market system

This option provides accuracy improvements from incorporating both price response and GXP level forecasting in the market system.

Costs for this option will be the same as for the base level option except costs for infrastructure to incorporate GXP level forecasts and price response in the market system will be incurred at time the option is implemented (in the near term), not in the Future States.

8 How Sensitivity Schedules Fit with Load Forecast Improvements

8.1 Sensitivity Schedules and Load Forecast Improvements Both Increase Efficiency and Competition

Sensitivity schedules enhance participants' ability to understand the price forecast distribution, hence assess their risk and opportunity. Sensitivity schedules and load forecast improvements both improve the value of information used by participants ahead of real-time.

Efficiency can be increased by an understandable forecast price distribution. This occurs when participants reduce load instead of paying more than they were willing. An improved load forecast also increases efficiency as less substantial errors will occur from acting on forecast prices in error.

Competition can be increased by an understandable forecast price distribution. Market participation will increase if risks and opportunities are clearer. An improved load forecast also increases competition, particularly if it is reliable, with few outliers. Market participation increases when participants have confidence in the load forecast and other inputs affecting forecast prices.

8.2 Load Forecast Improvement vs Sensitivity Schedules

There may be times when sensitivity schedules will provide significantly more benefit to participants than load forecast improvements or vice versa, and times when the most benefit will come from a combination of the two.

8.2.1 Sensitivity schedules provide more benefit when prices are high

Sensitivity schedules are likely to be more valuable than load forecast improvements to load participants when forecast prices are high. Load forecast accuracy improvements will also enable more efficient decisions but the benefit will decrease relative to sensitivity schedules as price sensitivity to input error increases. Prices will be most sensitive to input error when prices and input errors are both high, for example at longer load forecast horizons.

Figure 7 demonstrates how extreme prices may be seen ahead of time without sensitivity information. This case on 2 June 2012 was highly sensitive to load changes, with 12 MW less load causing a reduction from approximately $5000/MWh to $300/MWh. The extreme spot prices were not predicted in the longer horizons when input errors are high. It is relevant to note that 12 MW was approximately 0.2% of load at the time, far less than the MAPE achievable with the very best forecast, at any horizon. This demonstrates that extreme price forecast errors can occur even with a very high level of load forecast improvement.

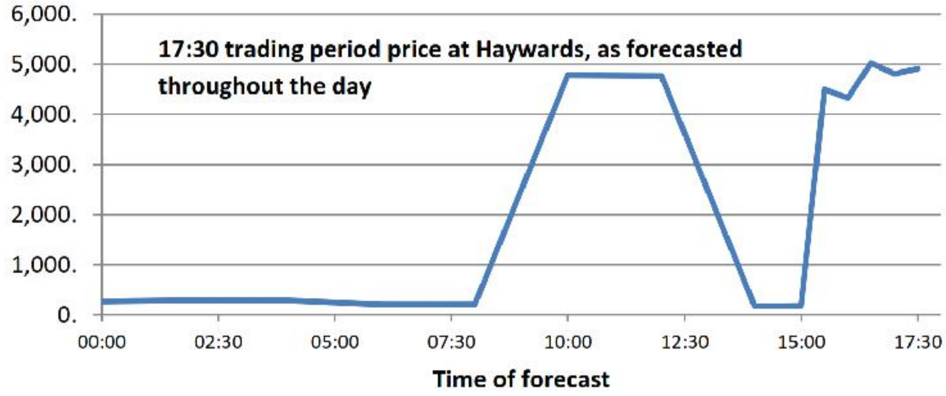


Figure 7. Variation in price forecast with forecast horizon on 2 June 2016, when price sensitivity was extreme.

8.2.2 Sensitivity schedules provide more benefit when input errors are high

Sensitivity schedules are valuable when input errors are high. Even with maximum improvement in load forecast error, total input error will still be high at longer horizons due to the influence of other inputs, particularly wind generation.

As an example, Figure 8 gives the total forecast error from conforming load, non-conforming load, and wind generation for June 2017. Using TESLA data to approximate the best forecast possible, minimum total error achievable by improving the load forecast at the 24-hour horizon is 1.8%, compared to 2.3% with the current forecast.

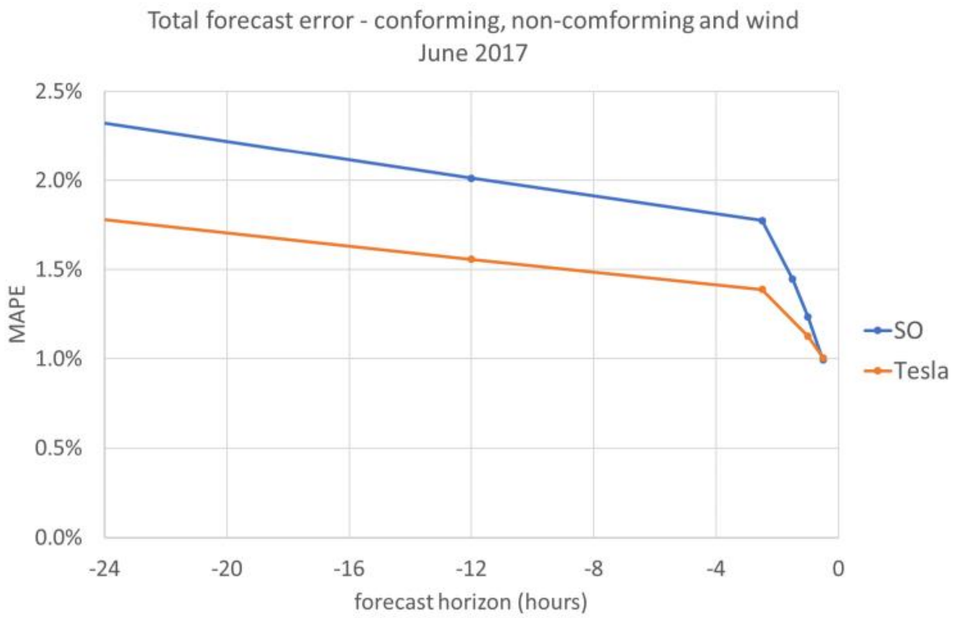


Figure 8. Total forecast error, including conforming load, non-conforming load, and wind generation. Comparison of current forecast and TESLA for June 2017.

8.2.3 Load forecast improvements provide more benefit when prices are normal and stable

Load forecast improvements would be more beneficial than sensitivity schedules when prices are in a normal range and relatively stable (i.e. insensitive).

8.3 Using Load Forecast Improvements and Sensitivity Schedules Together

8.3.1 Sensitivity schedules and load forecast variability

Load forecast variability information could be used alongside sensitivity schedules to enhance participants' understanding of the forecast price distribution. For example, confidence limits could be published at different time horizons. Sensitivity schedules are two schedules that adjust the load of the forecast schedules, one up, one down. If these were fixed adjustments, e.g. 100 MW, participants could see whether the 100 MW deviation was inside or outside the load forecast confidence limits, to gauge the probability of the adjusted prices. Alternatively, the load forecast variability could inform the adjustments, e.g. the two schedules could represent the load values with 10% and 90% probabilities of being exceeded.

Prices do not have to be extremely sensitive (as in the 2 June example above) for the 1.8% minimum error shown in Figure 8 to have a high impact on forecast prices. With prices only in the medium-high range, load forecast variability information combined with sensitivity schedules can usefully expose the risk. Participants may find this information useful in conjunction with sensitivity schedules.

8.3.2 Sensitivity schedules and total input variability information

Variability information would ideally include other varying components of price forecasts. Variability information that included wind, and potentially non-conforming load, would be much more useful than conforming load information alone.

8.3.3 Sensitivity schedules and modelled total input variability information

A forecast of variability/uncertainty that modelled various factors would be better than a simple representation. More information could be provided on the variability by forecasting how much it might change at different times, depending on different factors. For example, load forecasts vary more with higher prices, higher load, and more extreme temperatures. Both load and wind forecasts vary more when weather is changing quickly. For example, when wind is changing quickly, a slight inaccuracy in predicting the timing of an increase or decrease in wind energy can lead to a very large error for a given trading period.

AEMO have in the past published variability information for their load forecast, consisting of simple scaling factors based on historic forecast variability to give confidence limits at different horizons. However, they have recently developed a more complex Bayesian Belief Network model with information on a number of sources of variability, to cope with increases in solar generation and consequent dependence of load forecasts on weather[23]. At least one load forecasting vendor also provides uncertainty forecasting.

8.3.4 Sensitivity schedules, variability information, and reduction in load forecast outliers

Sensitivity schedules used alongside variability information will provide most benefit if load forecast improvements are made to reduce the prevalence of outliers. The benefit from implementing sensitivity schedules alongside variability information will be derived from participants' increased understanding of the forecast price distribution. Outliers in the load forecast error are unexpectedly large variations caused by factors not explained within the load forecast system. The variability model would provide

[23] https://www.aemo.com.au/-/media/Files/Electricity/NEM/Security_and_Reliability/Power_System_Ops/LOR-Reserve-Level-Declaration-Guidelines-Final-V10.docx

information on the size of known variations but outliers would remain unexplained. Outliers would therefore reduce the value of the variability information by limiting participants' understanding of the forecast price distribution. This would in turn limit participants' confidence in making decisions, hence competition benefits.

8.3.5 Sensitivity schedules used to indicate capacity requirements

Sensitivity schedules can also benefit system coordination in their assessment of energy and reserve requirements during peaks. System coordinators monitor the market for its ability to meet capacity requirements, to provide sufficient signals of the need for slow starting generation. This may require increasing the load forecast to meet the intra-period peak, which can sometimes be significantly higher than the average for the period. Sensitivity schedules could instead provide the peak signal of the instantaneous peak within a period, allowing load forecasts to be more indicative of spot prices.

8.3.6 Sensitivity schedules used to indicate price response

Sensitivity schedules could indicate variation from price response, which could be used by participants instead of signals provided from incorporating price response in the market system. If price elasticity information was published from the load forecast tool, participants could deduce pre- and post- response signals using this information alongside sensitivity schedules.

9 Key Findings

The following are the key findings from this investigation, including important points required to understand the problem and compare the options. Table 7 provides comparison of a sensitivity schedules option with a range of load forecast improvement options, to indicate some key differences and similarities for benefits of improved information to participants, future requirements, and costs.

Table 7. A summary of selected options along with sensitivity schedules and variability information.

	Value of information to participants		**Future-proof?**	**10-year costs***
Options	**normal conditions**	**stressed conditions**		
Improve current	medium	low	No	low
TESLA plug-in with current tool as backup	medium-high	low-medium	No	high
New system, closed, (TESLA in-house) + price response + GXP	very high	medium-high	Yes	very high
New system, open, + price response + GXP	high	medium	Yes	medium
Sensitivity schedules + variability information	medium	very high	Yes	medium

* Costs for options that are not future proof include capital costs to accommodate a new tool at the end of 10 years, discounted to take account of cost of capital.

9.1 Overall

- Significant accuracy improvements can be made to the medium-term load forecast, especially at longer forecast horizons, as demonstrated by the TESLA model in the load forecast trial.
- Price forecast errors will be large when prices are high, even with significant load forecast accuracy improvements, especially at longer forecast horizons.
- Sensitivity schedules will provide more benefit in these circumstances, in terms of price forecast accuracy.
- Variability information for schedule inputs, including load and wind generation, combined with sensitivity schedules, will be highly valuable to participants decision making. The better the variability information, the more value to participants.
- Reducing the prevalence of outliers will be the most important load forecast improvement if sensitivity schedules and variability information are implemented. This will ensure competition benefits as participants will have more confidence in decision making if they do not need to consider the possibility of unexplained forecast price variability due to outliers.
- There are many options for load forecast improvements in the near-term phase before new models are required in the future. Greater improvement can be made with a new model compared

to any current forecast improvement options. This model could be closed or open depending on acceptable cost and required accuracy. The greater the accuracy, the greater the costs.

- Key issues for the current forecast are
    - temporal shifts of energy such as network load control,
    - changes in load behaviour over time, including
        - over several years, e.g. uptake of heat pumps,
        - within a year, e.g. holidays, irrigation load, network load control,
        - day to day, e.g. irrigation load reduction when it rains, load control
    - poor and insufficient weather data, including
        - insufficient inputs, e.g. no account for humidity, wind direction, wind speed in most forecast areas.
        - low quality of current inputs

9.2 System Requirements for the Future

- Any future load forecast solution will require a new model and base infrastructure, to incorporate GXP level forecasts in the market system, and to incorporate price response in the market system. This will enable appropriate account of solar PV, batteries, sophisticated models making use of increases in available data, and home energy management systems.
- Incorporation of price response or GXP level forecasting in the market system need not be implemented at the same time a new system is acquired. These can be added separately and independently at a later date.
- All options for improving the load forecast can adapt to the future by investing in appropriate infrastructure and models at any time prior to significant future changes occurring. The less infrastructure invested now, the more required in the future.
- No option will have significant sunk costs when adaptation to the future is required. This includes improvement options within the current system (as costs incurred are mostly annual and required in any future forecast system) and closed options (if these costs are annualised with contracts that specify accuracy criteria and have short renewal periods).
- Human intervention is a key component of any accurate load forecast system, whether performed in-house or by an external vendor.

9.3 Improvements to the Current Forecast in the Near-Term

- The current forecast is much less of a 'black box' than initially believed, as Transpower has the ability to adjust the forecast model.
- Quick wins can be achieved such as:
    - Increasing the precision of the temperature inputs from 0 to 1 decimal places
    - Updating population weightings of different temperature measurements within a forecast area
    - Increasing temporal granularity of weather forecasts, from 6-hourly to 1-hourly data points each time a forecast is received
    - Increasing diligence in current intervention
- Human intervention in the current forecast, including regularly adjusting the forecast model, has potential to provide reasonable improvements in accuracy, particularly by reducing the prevalence of outliers.

- There are potentially conflicting analyses regarding the impact of weather; some analysis suggests improved weather information in the current forecast will not provide much improvement, though different analysis suggests otherwise. Any improvements are likely to mostly be in average accuracy.

9.4 Quick Deployment of a New Model

- Plugging in the current TESLA model, provided to ems via a web service, into the existing interfaces can be done relatively quickly and cheaply in terms of infrastructure. To ensure security and reliability the current model could be used as a backup. This option would be inadequate for the Future States, when GXP level forecasts and price response need to be incorporated in the market system, since the current system would be an inadequate backup. However, if full functionality with an in-house solution was desired in the near-term then the TESLA model can provide greater accuracy by feeding back information for outages, inter-GXP load shifts, and demand side initiative calls.
- Incorporating GXP level forecasts in the near term will increase accuracy of forecasting transmission constraints, and therefore spring washer effects.
- Incorporating price response in the near term will mainly increase accuracy during times of market stress.

10 Next Steps

We suggest the next steps to be:

Short term

- Consolidate all the quick wins, to capture benefits at little cost.
- Increase the amount of human intervention (outside of real time) in the current system.
- Incorporate improved input data, including more frequent receipt of current weather inputs as well as new inputs such as solar radiation, humidity, wind direction, and rainfall. Trialing intervention and new inputs in our current model would shed light on the potential improvements in each option and provide us with greater understanding of the opportunities and challenges of in-house vs at-vendor, specialist intervention.
- Implement sensitivity schedules alongside the load forecasting work to support market decision-making during times of market stress, particularly for those participants making decisions in longer time horizons such as thermal generators and gentailers entering hedge arrangements. If this is a major driver for improved load forecasts, sensitivity schedules should be given priority.

Medium term

- Investigate if a new system is appropriate in the near-term. If high accuracy in the near-term is desired, in parallel to or prior to sensitivity schedules, we suggest investigation into plugging in the current TESLA model to the existing interface, to provide area level forecasting. To ensure reliability and security the current model could be used as a backup.

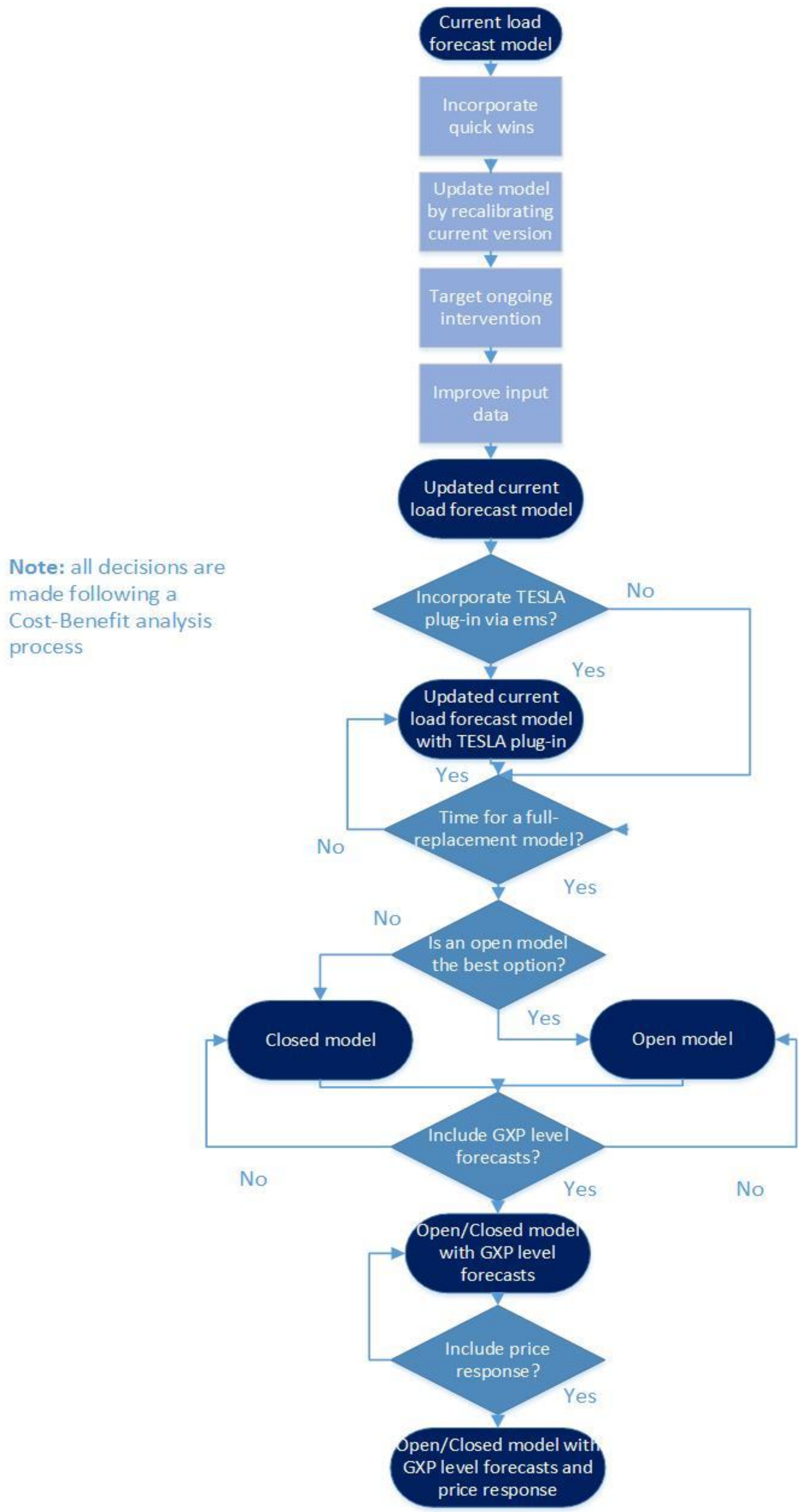
Long term

- Implement a new forecast system, including new base infrastructure to accommodate new models, along with infrastructure to accommodate GXP level forecasting and price responsiveness in the Market System[24]. In the long term the current load forecast model even with all improvements outlined above will not be fit for purpose. It currently does not have the required adaptability to account for the expected new load behaviour outlined in Future State A, or the capability to capture the benefits that may come from data science outlined in Future State B (see Section 5.2.2 for more details on Future States).
- Determine if the new system should be closed or open. An open system model cannot be fully open to the public. A closed system will provide greater accuracy but at greater cost. Adaptability concerns of a closed model, such as reduced and costly adaptability, are material but can be mitigated through robust contract arrangements. These arrangements would likely take shape as performance incentives ensuring the vendor has economic reason to maintain and update their model, and shortened service agreements allowing the system operator opportunities to re-negotiate or change providers as we transition through Future States A and B.
- Provide input variability information in the new system to support sensitivity schedules. This could be improved incrementally from a simple high-level model to a more complicated and detailed model incorporating wind generation, and incorporating several factors that affect the extent of variability. Such factors may include temperature, wind, and price. We believe variability information will be more relevant in the future, particularly due to difficulties in forecasting solar, and increased price elasticity of demand.

---

[24] GXP level forecasting and price response can be added immediately or at a later date.

The flow chart below outlines the next steps for load forecast improvements.



APPENDIX 1 SUMMARY OF LOAD FORECAST ACCURACY BY MONTH

Summary data for the current load forecast at the 2.5-hour forecast horizon in 2017

Values are given per month, for New Zealand and the 10 load forecast areas, namely Christchurch (CH), West Coast (WC), Invercargill (IN), Northland (NL), Auckland (AK), Bay of Plenty (BP) Hamilton (HM), Napier (NR), Palmerston North (PN) and Wellington (WN).

The variables shown are Mean Absolute Error (MAE) (MW), Mean Average Percentage Error (MAPE), mean error (actual load – forecast load) (MW), and standard deviation (MW).

SO load forecast, 2.5 hour-horizon

type	month	NZ	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
MAE	Jan	55.5	12.5	6.0	9.5	6.9	17.3	3.6	7.9	4.5	6.4	6.6
MAE	Feb	39.0	8.3	5.1	7.7	5.5	13.4	3.3	7.0	4.3	4.8	5.3
MAE	Mar	40.0	11.4	5.3	8.6	4.9	13.3	3.0	6.2	3.8	4.6	5.7
MAE	Apr	65.1	15.2	6.1	9.0	7.1	16.9	4.6	9.0	5.9	6.2	9.4
MAE	May	69.5	20.1	7.0	9.7	8.2	22.3	5.9	11.6	6.1	6.0	10.4
MAE	Jun	67.8	23.7	7.3	12.3	8.7	21.1	5.4	10.6	6.0	6.4	10.2
MAE	Jul	96.9	27.6	9.4	16.5	10.7	26.8	7.0	14.7	8.8	7.8	13.7
MAE	Aug	67.6	23.6	7.3	13.0	8.6	20.6	6.0	11.0	5.5	5.9	9.9
MAE	Sep	65.4	19.0	6.2	9.2	8.8	19.3	5.3	9.4	5.4	6.2	9.4
MAE	Oct	49.5	14.8	6.0	8.2	6.5	15.2	4.3	7.6	5.2	5.2	7.8
MAE	Nov	34.8	9.3	5.3	7.0	4.8	10.4	3.1	6.6	3.9	4.0	4.8
MAE	Dec	47.1	10.8	5.7	8.0	5.2	13.6	3.5	7.4	4.5	4.4	5.5
MAPE	Jan	1.8%	2.8%	3.4%	3.1%	2.4%	2.5%	2.2%	2.3%	2.4%	2.8%	0.9%
MAPE	Feb	1.2%	1.7%	2.5%	2.4%	1.8%	1.7%	1.9%	2.0%	2.2%	1.9%	0.7%
MAPE	Mar	1.2%	2.4%	2.9%	2.6%	1.6%	1.6%	1.7%	1.8%	1.8%	1.9%	0.7%
MAPE	Apr	1.9%	3.2%	3.3%	2.8%	2.2%	2.2%	2.4%	2.6%	3.1%	2.5%	1.2%
MAPE	May	1.8%	3.7%	3.5%	2.7%	2.3%	2.5%	2.7%	3.1%	2.8%	2.3%	1.2%
MAPE	Jun	1.7%	4.1%	3.7%	3.4%	2.4%	2.2%	2.4%	2.8%	2.7%	2.4%	1.1%
MAPE	Jul	2.3%	4.6%	5.0%	4.3%	2.7%	2.6%	3.0%	3.7%	3.7%	2.9%	1.3%
MAPE	Aug	1.7%	4.3%	3.9%	3.7%	2.3%	2.2%	2.7%	2.7%	2.5%	2.2%	1.0%
MAPE	Sep	1.8%	3.8%	3.3%	2.7%	2.5%	2.1%	2.5%	2.3%	2.7%	2.3%	1.0%
MAPE	Oct	1.4%	3.1%	3.4%	2.5%	1.9%	1.8%	2.3%	2.1%	2.8%	2.1%	0.9%
MAPE	Nov	1.0%	1.9%	2.9%	2.0%	1.5%	1.3%	1.7%	1.8%	2.0%	1.6%	0.6%
MAPE	Dec	1.5%	2.3%	2.9%	2.4%	1.7%	1.9%	2.1%	2.1%	2.5%	1.8%	0.8%
mean error	Jan	12.9	1.9	1.0	-0.7	-0.2	3.0	-0.1	4.6	0.8	1.6	1.0
mean error	Feb	10.6	0.4	0.3	0.8	0.5	5.1	0.1	4.6	0.2	-1.0	-0.4
mean error	Mar	4.3	-1.6	-1.5	-0.6	0.2	3.1	0.4	3.4	-0.1	0.3	0.6
mean error	Apr	4.4	0.0	0.3	-0.2	0.0	1.1	0.1	3.9	-0.5	0.3	-0.7
mean error	May	5.4	0.0	0.0	0.6	-0.1	2.6	-0.2	3.9	-0.4	-0.1	-0.8
mean error	Jun	14.3	1.0	0.2	0.4	0.4	5.7	0.5	4.5	0.6	0.6	0.4
mean error	Jul	10.0	-0.3	0.3	0.4	0.9	3.2	0.3	4.8	0.2	0.4	0.0
mean error	Aug	-11.9	-5.1	-0.4	-3.3	-1.1	-1.2	-0.5	3.0	-0.8	-0.3	-2.1
mean error	Sep	-6.9	-4.6	-0.2	-0.7	-0.4	-1.1	-0.6	3.4	-0.7	-0.4	-1.6
mean error	Oct	7.1	2.8	0.6	1.0	-0.8	0.5	-0.3	3.7	-0.1	0.0	-0.3
mean error	Nov	12.0	3.8	0.5	1.4	0.2	2.1	-0.3	3.5	0.8	0.1	-0.3
mean error	Dec	-1.5	-2.5	-2.0	0.0	-0.5	0.6	-0.4	4.4	-0.6	-0.2	-0.3
sd	Jan	77.1	16.8	8.1	12.6	10.3	25.1	5.2	9.2	5.8	8.8	9.1
sd	Feb	53.2	11.2	6.6	9.8	8.4	17.7	4.4	7.9	5.4	6.6	7.4
sd	Mar	54.1	15.8	6.9	11.5	6.7	18.4	4.2	7.2	4.9	6.1	8.5
sd	Apr	91.2	20.4	8.0	12.1	10.1	24.0	6.5	16.0	8.2	8.4	12.6
sd	May	93.5	26.1	9.5	13.0	11.5	30.7	8.0	15.4	8.0	8.1	13.4
sd	Jun	90.7	32.4	9.8	16.5	13.7	28.0	7.3	13.4	8.1	8.2	18.5
sd	Jul	132.8	36.2	12.6	21.4	14.6	38.1	10.2	19.8	12.1	10.8	18.6
sd	Aug	87.7	30.4	9.8	16.3	11.5	28.0	7.9	14.1	7.3	7.7	13.1
sd	Sep	88.5	24.5	7.8	12.2	12.1	26.9	7.1	12.3	7.2	8.2	12.8
sd	Oct	66.3	23.6	7.6	10.6	9.0	20.9	5.7	9.3	6.8	6.9	10.5
sd	Nov	42.8	11.7	8.2	8.9	6.4	13.6	4.2	7.7	4.9	5.3	6.2
sd	Dec	68.6	14.9	7.5	10.5	7.2	19.3	4.9	8.1	6.0	5.8	7.4

Summary data for the current load forecast at the 24-hour forecast horizon in 2017

Values are given per month, for New Zealand and the 10 load forecast areas, namely Christchurch (CH), West Coast (WC), Invercargill (IN), Northland (NL), Auckland (AK), Bay of Plenty (BP) Hamilton (HM), Napier (NR), Palmerston North (PN) and Wellington (WN).

The variables shown are Mean Absolute Error (MAE) (MW), Mean Average Percentage Error (MAPE), mean error (actual load – forecast load) (MW) and standard deviation (MW).

SO load forecast, 24 hour-horizon

type	month	NZ	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
MAE	Jan	120.4	25.0	12.8	18.3	11.9	34.7	6.3	11.7	8.2	13.9	10.8
MAE	Feb	73.6	11.0	7.9	13.1	9.8	25.8	6.1	11.8	6.5	8.3	10.3
MAE	Mar	67.0	21.3	8.5	13.6	8.7	21.4	5.2	9.4	5.5	6.0	8.5
MAE	Apr	114.4	29.2	9.4	14.9	10.5	27.1	7.5	11.4	9.6	9.6	16.6
MAE	May	108.4	28.4	9.2	15.6	12.1	31.6	8.7	17.1	9.2	8.9	16.8
MAE	Jun	86.3	29.1	8.8	16.8	12.4	26.5	6.8	13.7	7.8	9.5	12.1
MAE	Jul	141.8	32.9	12.3	20.9	14.9	35.2	9.5	20.3	11.3	12.4	19.7
MAE	Aug	105.6	33.7	9.4	18.4	12.8	32.6	9.1	15.2	8.2	8.8	15.1
MAE	Sep	99.0	28.2	8.9	14.3	12.8	29.3	7.9	14.3	7.5	9.0	13.0
MAE	Oct	71.1	23.6	9.1	11.9	10.0	22.3	6.7	10.2	7.6	7.7	12.0
MAE	Nov	60.5	15.0	6.8	9.7	7.9	17.1	5.2	9.3	5.5	6.1	9.1
MAE	Dec	82.8	15.5	8.4	14.0	7.8	25.1	5.7	10.9	8.1	6.6	8.4
MAPE	Jan	3.8%	5.5%	6.9%	5.8%	4.0%	4.9%	3.8%	3.4%	4.4%	5.7%	1.5%
MAPE	Feb	2.1%	2.2%	3.9%	4.0%	3.1%	3.1%	3.4%	3.2%	3.4%	3.3%	1.2%
MAPE	Mar	1.9%	4.5%	4.7%	4.1%	2.8%	2.5%	2.8%	2.6%	2.7%	2.4%	1.0%
MAPE	Apr	3.3%	6.2%	5.0%	4.7%	3.2%	3.4%	3.9%	3.3%	4.9%	3.9%	2.0%
MAPE	May	2.9%	5.2%	4.7%	4.3%	3.4%	3.5%	3.9%	4.5%	4.2%	3.3%	1.9%
MAPE	Jun	2.2%	5.0%	4.6%	4.7%	3.4%	2.8%	3.1%	3.6%	3.5%	3.6%	1.3%
MAPE	Jul	3.5%	5.6%	6.7%	5.6%	3.7%	3.4%	4.1%	5.1%	4.9%	4.6%	1.9%
MAPE	Aug	2.7%	6.2%	5.0%	5.4%	3.4%	3.5%	4.0%	3.7%	3.8%	3.2%	1.6%
MAPE	Sep	2.7%	5.6%	4.9%	4.2%	3.6%	3.2%	3.7%	3.5%	3.7%	3.3%	1.4%
MAPE	Oct	2.1%	5.0%	5.1%	3.6%	3.0%	2.7%	3.5%	2.7%	4.1%	3.0%	1.4%
MAPE	Nov	1.7%	3.1%	3.6%	2.8%	2.4%	2.1%	2.8%	2.5%	2.8%	2.4%	1.1%
MAPE	Dec	2.6%	3.3%	4.2%	4.2%	2.6%	3.5%	3.3%	3.1%	4.4%	2.7%	1.1%
mean error	Jan	15.1	4.3	3.7	-3.3	-1.2	-0.7	-1.0	4.0	1.9	5.0	2.4
mean error	Feb	21.5	0.6	0.8	2.4	2.0	14.0	0.4	5.6	0.5	-4.3	-0.4
mean error	Mar	2.7	-4.4	-4.0	-1.7	0.8	4.4	1.4	3.9	0.2	0.8	1.4
mean error	Apr	-8.1	-1.9	0.7	-1.4	-0.3	-2.4	0.4	1.6	-1.8	-0.4	-2.8
mean error	May	4.9	-0.7	-0.1	1.3	-0.8	5.9	-0.6	3.1	-1.2	-0.5	-1.4
mean error	Jun	19.6	0.0	0.9	2.7	0.5	5.9	0.2	6.8	0.9	1.5	0.4
mean error	Jul	24.0	-1.6	-0.9	2.7	4.0	9.2	1.2	8.0	-0.1	0.7	0.9
mean error	Aug	-33.5	-8.7	-0.9	-7.7	-2.5	-7.3	-1.1	3.0	-2.2	-0.5	-5.5
mean error	Sep	-25.9	-9.4	-1.0	-2.0	-1.6	-6.6	-1.4	1.9	-1.7	-0.8	-3.4
mean error	Oct	9.6	6.7	1.9	2.3	-1.5	-0.2	-0.7	3.2	-0.2	-0.2	-1.7
mean error	Nov	19.0	8.7	1.7	3.4	0.5	1.6	-0.8	3.5	2.0	0.1	-1.8
mean error	Dec	-13.3	-5.2	-3.5	-0.7	-1.0	-3.0	-0.8	4.1	-1.6	-0.4	-1.2
sd	Jan	163.5	31.1	15.9	22.5	16.4	50.9	9.5	14.7	10.5	16.8	14.6
sd	Feb	102.1	20.1	11.2	18.0	13.4	34.5	8.0	14.5	7.8	10.8	13.9
sd	Mar	91.4	28.6	9.8	17.6	11.5	28.0	7.1	11.1	6.9	8.1	13.1
sd	Apr	153.5	39.6	12.4	19.5	14.9	36.2	9.8	14.7	12.6	13.0	21.9
sd	May	145.4	36.9	12.0	21.2	16.3	44.1	11.7	23.1	12.1	11.6	21.5
sd	Jun	114.0	53.9	11.1	21.2	18.5	35.0	9.2	17.3	10.1	11.7	16.7
sd	Jul	190.6	43.2	15.5	26.8	20.7	52.0	13.5	26.5	14.9	17.1	26.6
sd	Aug	133.7	42.8	12.0	20.9	16.5	41.7	11.9	19.4	10.5	11.7	18.9
sd	Sep	129.5	34.4	11.1	19.2	17.4	38.1	10.6	19.0	9.7	12.3	17.6
sd	Oct	97.0	31.7	10.9	15.2	13.4	29.1	8.8	12.9	9.8	10.1	16.4
sd	Nov	76.8	17.0	9.2	11.6	10.3	22.5	7.2	11.6	6.9	8.4	11.8
sd	Dec	114.9	20.6	10.4	17.5	10.7	35.1	7.6	13.8	10.7	8.7	11.1

Summary data for the TESLA load forecast at the 2.5-hour forecast horizon in 2017

Values are given per month, for New Zealand and the 10 load forecast areas, namely Christchurch (CH), West Coast (WC), Invercargill (IN), Northland (NL), Auckland (AK), Bay of Plenty (BP) Hamilton (HM), Napier (NR), Palmerston North (PN) and Wellington (WN).

The variables shown are Mean Absolute Error (MAE) (MW), Mean Average Percentage Error (MAPE), mean error (actual load – forecast load) (MW) and standard deviation (MW).

TESLA load forecast, 2.5-hour horizon

type	month	NZ	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
MAE	Jan	32.4	11.3	8.4	8.7	4.3	9.5	2.8	4.8	4.4	5.1	4.5
MAE	Feb	25.2	13.2	7.6	6.5	6.2	8.7	2.2	4.8	4.5	5.3	4.0
MAE	Mar	28.2	10.0	5.0	7.0	6.1	10.2	2.7	5.0	3.6	4.4	4.1
MAE	Apr	40.6	15.2	7.9	9.1	6.2	14.8	4.1	8.0	4.3	5.1	6.6
MAE	May	34.5	13.3	5.6	7.0	6.2	19.4	3.8	8.0	4.1	5.8	6.0
MAE	Jun	34.9	17.2	6.0	7.5	6.4	13.4	5.0	7.6	4.7	5.6	6.5
MAE	Jul	56.0	19.9	6.5	8.8	7.8	17.7	4.9	9.5	6.0	7.5	8.9
MAE	Aug	40.4	14.5	6.1	8.1	7.2	16.6	3.8	7.5	4.5	4.5	6.6
MAE	Sep	42.6	14.0	6.3	7.8	7.9	13.6	4.6	7.5	4.5	5.6	7.4
MAE	Oct	27.0	10.9	5.5	7.3	5.8	9.9	3.2	5.2	4.1	4.6	5.4
MAE	Nov	24.5	8.0	6.4	10.6	4.4	7.4	2.6	4.5	3.3	4.0	3.0
MAE	Dec	27.1	9.2	6.9	9.2	5.5	8.7	2.4	5.8	4.7	3.9	3.8
MAPE	Jan	1.0%	2.4%	4.3%	2.8%	1.4%	1.3%	1.8%	1.4%	2.4%	2.2%	0.6%
MAPE	Feb	0.7%	2.7%	3.6%	2.0%	2.0%	1.1%	1.3%	1.4%	2.3%	2.2%	0.5%
MAPE	Mar	0.8%	2.1%	2.7%	2.1%	2.0%	1.2%	1.5%	1.4%	1.8%	1.9%	0.5%
MAPE	Apr	1.2%	3.3%	4.1%	2.8%	2.0%	1.8%	2.2%	2.3%	2.2%	2.1%	0.8%
MAPE	May	0.9%	2.4%	2.8%	1.9%	1.7%	2.2%	1.7%	2.1%	1.8%	2.3%	0.7%
MAPE	Jun	0.9%	2.9%	3.0%	2.1%	1.7%	1.4%	2.3%	2.0%	2.1%	2.2%	0.7%
MAPE	Jul	1.4%	3.3%	3.4%	2.3%	2.0%	1.7%	2.2%	2.4%	2.5%	3.0%	0.8%
MAPE	Aug	1.0%	2.6%	3.3%	2.2%	1.9%	1.7%	1.7%	1.8%	2.1%	1.7%	0.7%
MAPE	Sep	1.1%	2.7%	3.5%	2.3%	2.2%	1.4%	2.1%	1.9%	2.2%	2.0%	0.8%
MAPE	Oct	0.8%	2.3%	3.1%	2.2%	1.8%	1.2%	1.6%	1.4%	2.1%	1.8%	0.6%
MAPE	Nov	0.7%	1.6%	3.3%	3.0%	1.3%	0.9%	1.4%	1.2%	1.7%	1.6%	0.4%
MAPE	Dec	0.8%	1.9%	3.3%	2.7%	1.8%	1.2%	1.4%	1.6%	2.5%	1.7%	0.5%
mean error	Jan	20.5	7.0	7.2	1.0	1.3	2.3	-1.7	-1.1	2.1	0.7	1.5
mean error	Feb	10.8	6.9	6.2	1.0	-2.5	3.5	0.1	0.3	-2.8	-0.6	-1.3
mean error	Mar	5.7	1.4	1.3	2.8	-2.0	-0.2	1.4	1.1	-0.2	-0.9	1.0
mean error	Apr	-2.4	-8.2	5.7	-3.5	3.2	0.6	0.6	-2.9	-0.2	-0.5	2.6
mean error	May	0.2	2.7	0.0	-1.3	-1.3	0.7	-0.2	-0.4	1.4	-3.8	2.5
mean error	Jun	-1.5	10.8	-0.5	0.1	-2.8	-3.9	-3.9	-0.7	-0.5	-2.9	2.8
mean error	Jul	5.1	11.4	-1.4	0.8	-0.2	0.0	-1.3	-2.0	0.9	-3.6	0.7
mean error	Aug	-10.2	1.5	0.2	-2.1	-3.4	-5.8	-0.6	-0.6	-0.7	0.6	0.6
mean error	Sep	1.0	-2.6	1.4	-0.4	0.1	-1.0	2.1	-1.3	0.6	0.2	1.8
mean error	Oct	-0.4	2.6	-0.7	1.7	1.6	-0.7	0.4	-2.6	-0.9	-0.5	-1.3
mean error	Nov	12.6	4.5	1.7	9.4	-1.4	-0.9	-0.5	-0.5	0.7	0.0	-0.2
mean error	Dec	-5.3	-1.1	-3.5	5.4	-2.1	-1.9	-0.3	0.3	-3.1	-0.4	1.3
sd	Jan	35.4	13.2	7.4	10.7	5.6	12.1	3.3	5.8	5.3	6.4	6.2
sd	Feb	30.1	15.9	6.2	8.2	9.2	10.8	2.8	6.2	5.0	6.7	5.3
sd	Mar	36.7	13.7	6.4	8.4	9.2	14.1	3.3	6.3	4.6	5.4	6.0
sd	Apr	54.6	18.6	8.6	11.5	7.5	20.4	5.6	10.2	6.3	6.6	9.5
sd	May	46.1	18.2	7.5	9.2	7.8	26.5	5.0	10.5	5.4	6.2	8.4
sd	Jun	47.4	21.7	8.0	9.7	8.3	17.6	5.1	9.9	6.4	6.4	9.3
sd	Jul	79.2	24.8	8.4	11.6	11.0	25.7	6.9	12.5	8.4	8.5	12.3
sd	Aug	51.9	20.7	7.8	10.6	9.3	24.0	5.2	9.9	6.1	6.0	9.8
sd	Sep	61.3	18.9	7.8	10.3	10.8	19.7	6.1	9.9	6.2	7.7	11.4
sd	Oct	36.7	14.9	7.1	9.4	7.5	13.7	4.3	6.4	5.4	6.0	7.6
sd	Nov	29.2	9.6	9.0	8.1	6.1	10.6	3.2	5.9	4.2	5.2	4.1
sd	Dec	35.3	12.0	8.6	10.1	7.8	11.0	3.1	7.3	5.2	5.0	5.0

Difference summary data showing the effect of manual intervention by coordinators, current load forecast, 2.5-hour horizon, 2017

The system operator data in the summary tables on the previous pages are based on load forecasts which are the direct outputs of the load forecast tool. Upon occasion, manual overrides are applied to these forecasts by the coordinators, prior to the data being used as the load input to market schedules such as the NRS. Changes to the summary data due to these overrides are given in the table below. Negative (red) values indicate improvements in errors due to manual overrides, and blue (positive) values indicate a worsening in errors due to manual overrides.

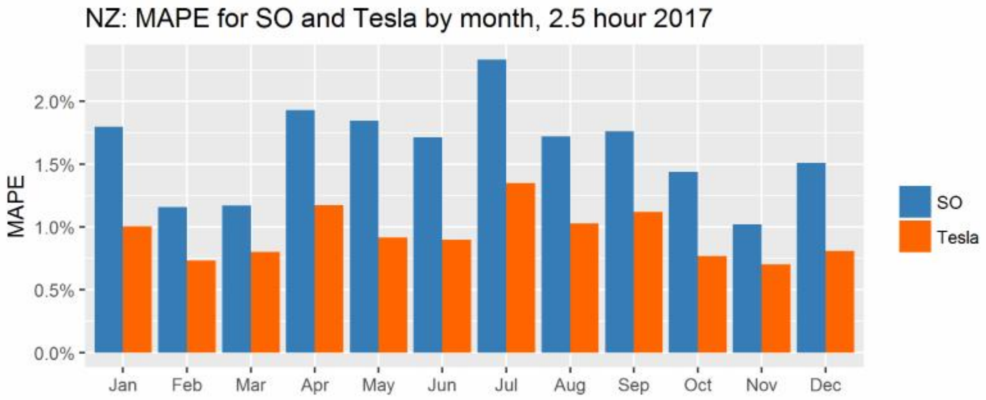
type	month	NZ	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
MAE	Jan	-1.8	-0.3	0.3	-0.2	-0.1	0.2	-0.1	0.4	-0.1	0.1	0.0
MAE	Feb	-0.8	0.3	0.3	0.4	0.0	0.0	0.0	-0.1	0.1	0.1	0.2
MAE	Mar	-0.2	0.1	-0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0
MAE	Apr	0.9	0.2	0.1	-0.2	0.2	0.2	0.1	-0.4	0.1	0.1	0.5
MAE	May	-1.8	-0.3	0.2	0.4	-0.1	-0.2	-0.1	0.1	0.1	0.0	-0.2
MAE	Jun	-0.9	-0.7	0.0	-0.1	0.0	0.1	0.0	0.1	0.0	0.0	-0.3
MAE	Jul	0.5	-0.4	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.1
MAE	Aug	-0.2	-0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MAE	Sep	0.5	0.0	0.1	0.1	0.1	0.3	0.0	0.1	0.0	0.0	0.0
MAE	Oct	-0.5	-0.4	0.4	0.0	0.4	-0.2	0.0	0.0	0.0	0.1	0.1
MAE	Nov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MAE	Dec	-0.2	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0
MAPE	Jan	-0.06%	-0.07%	0.15%	-0.06%	-0.05%	0.03%	-0.07%	0.12%	-0.09%	0.03%	0.00%
MAPE	Feb	-0.03%	0.05%	0.14%	0.10%	0.00%	0.00%	-0.01%	-0.02%	0.02%	0.05%	0.01%
MAPE	Mar	-0.01%	0.01%	-0.03%	0.01%	0.00%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%
MAPE	Apr	0.05%	0.03%	0.07%	-0.05%	0.09%	0.06%	0.08%	-0.06%	0.07%	0.07%	0.06%
MAPE	May	-0.04%	-0.05%	0.14%	0.14%	-0.02%	-0.02%	-0.02%	0.04%	0.06%	0.02%	-0.01%
MAPE	Jun	-0.03%	-0.12%	-0.01%	0.00%	0.00%	0.01%	0.01%	0.02%	0.01%	0.01%	-0.03%
MAPE	Jul	0.01%	-0.08%	-0.01%	0.02%	0.01%	0.00%	0.01%	0.00%	0.01%	0.01%	0.00%
MAPE	Aug	-0.01%	-0.07%	0.07%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
MAPE	Sep	0.01%	0.01%	0.03%	0.03%	0.04%	0.04%	0.01%	0.02%	0.01%	0.01%	0.00%
MAPE	Oct	-0.02%	-0.08%	0.19%	-0.01%	0.13%	-0.03%	-0.02%	-0.01%	0.03%	0.03%	0.01%
MAPE	Nov	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
MAPE	Dec	0.00%	0.01%	0.04%	-0.01%	0.00%	0.00%	0.00%	0.00%	-0.02%	0.09%	0.00%
mean error	Jan	2.9	0.2	0.9	0.4	0.2	-0.2	0.2	0.6	0.0	0.4	0.3
mean error	Feb	-1.9	-0.4	0.3	-0.6	-0.3	-0.1	0.0	-0.1	-0.2	-0.3	-0.3
mean error	Mar	-0.1	0.1	0.1	0.0	0.0	-0.2	0.0	0.0	0.0	0.0	0.0
mean error	Apr	-0.2	0.3	0.1	-0.2	0.0	-0.3	0.0	-0.8	0.0	0.1	0.4
mean error	May	-5.0	-0.5	-0.4	-0.7	-0.4	-1.1	-0.3	-0.5	-0.4	-0.4	-0.3
mean error	Jun	-1.2	-0.4	-0.2	0.0	0.0	-0.1	0.0	-0.1	0.0	0.0	-0.4
mean error	Jul	1.3	1.2	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0
mean error	Aug	0.6	0.4	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
mean error	Sep	-0.1	0.2	0.1	0.1	-0.1	-0.4	0.0	0.0	0.0	0.0	-0.1
mean error	Oct	1.6	0.4	0.5	0.2	0.1	0.2	0.0	0.1	0.1	0.1	-0.1
mean error	Nov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
mean error	Dec	0.7	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0
sd	Jan	-3.5	-0.5	0.2	-0.2	-0.3	0.2	-0.1	0.5	-0.2	0.2	-0.1
sd	Feb	-1.6	0.7	0.4	0.8	-0.1	-0.2	-0.1	-0.2	0.1	0.3	0.3
sd	Mar	-0.1	0.1	-0.1	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0
sd	Apr	0.0	0.3	0.2	-0.4	0.1	0.1	0.1	-2.5	0.2	0.3	0.7
sd	May	-3.9	-0.5	0.3	0.5	-0.3	-0.5	-0.1	0.3	0.3	0.1	-0.3
sd	Jun	-1.3	-1.5	-0.1	-0.6	0.0	0.1	0.1	0.1	0.0	0.0	-2.5
sd	Jul	-0.6	-0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1
sd	Aug	-0.1	-0.4	0.2	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
sd	Sep	0.6	0.1	0.1	0.1	0.2	0.4	0.0	0.1	0.1	0.0	0.0
sd	Oct	-1.3	-3.3	0.5	0.0	0.7	-0.4	-0.1	-0.1	0.0	0.1	0.1
sd	Nov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
sd	Dec	-0.3	0.0	0.2	0.0	-0.1	0.0	0.0	0.0	0.0	0.6	0.0

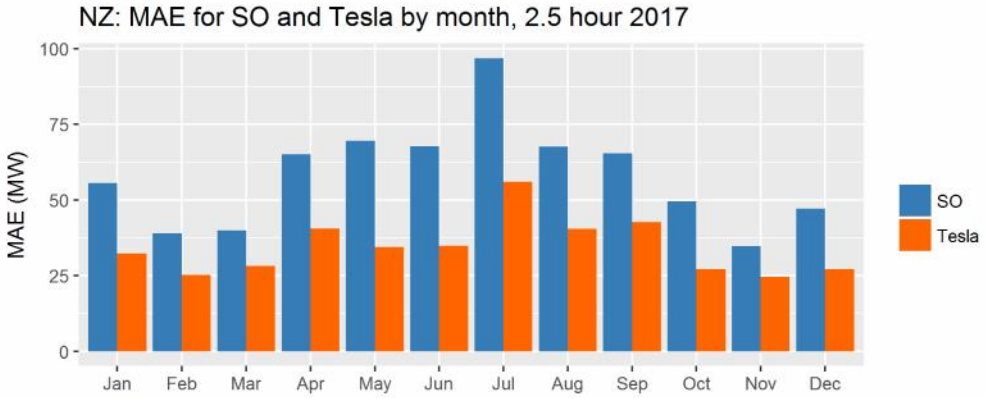
Difference summary data showing the effect of manual intervention by coordinators, current load forecast, 24-hour horizon, 2017

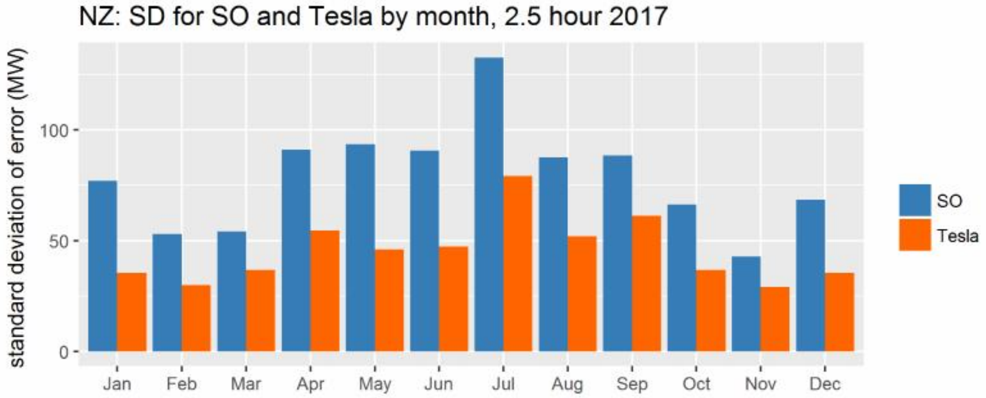
The system operator data in the summary tables on the previous pages are based on load forecasts which are the direct outputs of the load forecast tool. Upon occasion, manual overrides are applied to these forecasts by the coordinators, prior to the data being used as the load input to market schedules such as the NRS. Changes to the summary data due to these overrides are given in the table below. Negative (red) values indicate improvements in errors due to manual overrides, and blue (positive) values indicate a worsening in errors due to manual overrides.

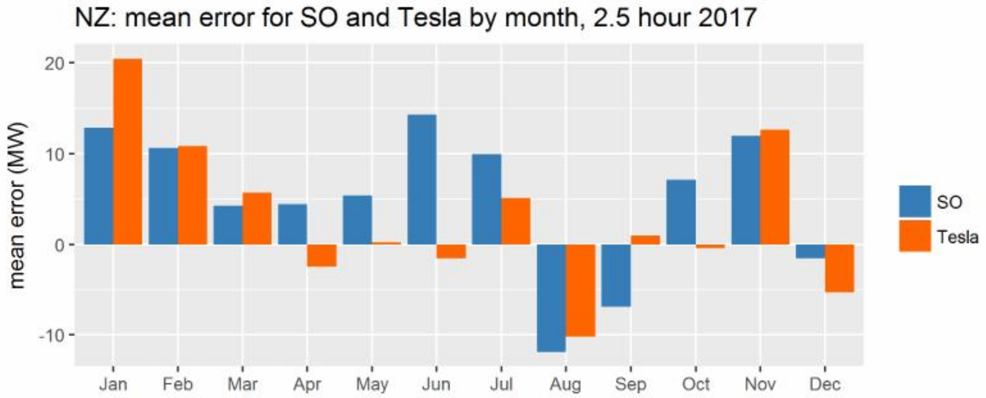
type	month	NZ	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
MAE	Jan	-0.3	0.1	0.1	0.1	0.0	-0.1	-0.1	0.1	-0.1	0.0	0.1
MAE	Feb	-4.3	-0.4	-0.2	-0.4	-0.2	-1.4	-0.2	-0.5	0.3	0.2	-0.1
MAE	Mar	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MAE	Apr	-0.8	-1.3	0.1	0.5	0.3	-0.5	0.1	-0.3	0.0	0.4	-0.3
MAE	May	-1.6	-0.1	0.0	-0.1	0.0	0.0	0.2	0.2	0.3	0.2	-0.4
MAE	Jun	-0.9	-0.5	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MAE	Jul	-0.7	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MAE	Aug	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MAE	Sep	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MAE	Oct	0.6	0.0	0.0	0.0	0.1	0.2	0.0	0.1	0.0	0.1	0.1
MAE	Nov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MAE	Dec	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MAPE	Jan	-0.01%	0.02%	0.04%	0.02%	0.00%	-0.02%	-0.05%	0.03%	-0.04%	-0.01%	0.02%
MAPE	Feb	-0.11%	-0.07%	-0.10%	-0.12%	-0.05%	-0.14%	-0.11%	-0.11%	0.13%	0.07%	-0.01%
MAPE	Mar	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
MAPE	Apr	0.01%	-0.24%	0.08%	0.14%	0.12%	-0.02%	0.09%	-0.06%	0.02%	0.15%	-0.02%
MAPE	May	-0.03%	-0.02%	0.02%	-0.02%	0.00%	0.00%	0.07%	0.05%	0.12%	0.08%	-0.03%
MAPE	Jun	-0.02%	-0.07%	0.00%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
MAPE	Jul	-0.02%	-0.03%	0.00%	-0.01%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
MAPE	Aug	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
MAPE	Sep	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
MAPE	Oct	0.02%	-0.01%	0.00%	0.00%	0.02%	0.02%	0.02%	0.02%	0.02%	0.02%	0.01%
MAPE	Nov	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
MAPE	Dec	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
mean error	Jan	1.6	0.2	0.1	0.2	0.2	0.4	0.1	0.2	0.1	0.1	0.1
mean error	Feb	-4.0	0.1	0.1	0.0	-0.6	-1.4	-0.3	-0.6	-0.4	-0.5	-0.5
mean error	Mar	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
mean error	Apr	9.1	1.2	0.7	0.9	0.8	2.2	0.5	0.7	0.5	0.7	0.9
mean error	May	-5.8	-0.1	0.0	-0.1	-0.8	-1.8	-0.5	-0.8	-0.5	-0.6	-0.7
mean error	Jun	1.5	1.4	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
mean error	Jul	2.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
mean error	Aug	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
mean error	Sep	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
mean error	Oct	0.6	0.0	0.0	0.0	0.1	0.2	0.0	0.1	0.0	0.1	0.1
mean error	Nov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
mean error	Dec	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
sd	Jan	-0.2	0.1	0.0	0.1	0.0	-0.1	-0.1	0.0	-0.1	0.0	0.1
sd	Feb	-7.1	-6.4	-1.7	-2.7	-0.2	-2.5	-0.3	-0.4	0.5	0.1	-0.2
sd	Mar	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
sd	Apr	-3.8	-2.2	0.0	0.9	0.1	-1.0	0.1	-0.6	0.0	0.5	-0.4
sd	May	-2.6	-0.1	0.0	0.0	0.0	0.0	0.3	0.3	0.3	0.3	-0.6
sd	Jun	-6.1	-14.7	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
sd	Jul	-2.0	-0.4	0.0	-0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
sd	Aug	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
sd	Sep	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
sd	Oct	1.1	0.0	0.0	0.0	0.3	0.4	0.0	0.2	0.1	0.0	0.1
sd	Nov	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
sd	Dec	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Graphs of summary data by month, comparing SO and TESLA forecast errors at the 2.5 hour forecast horizon. Data for New Zealand.

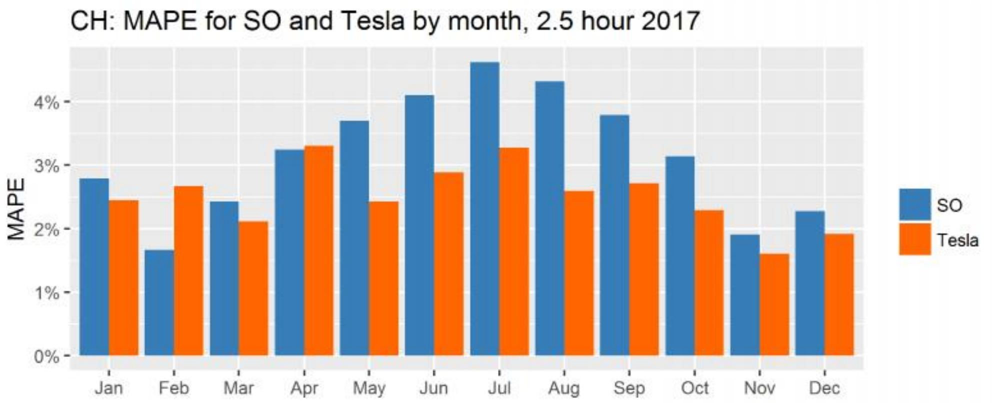


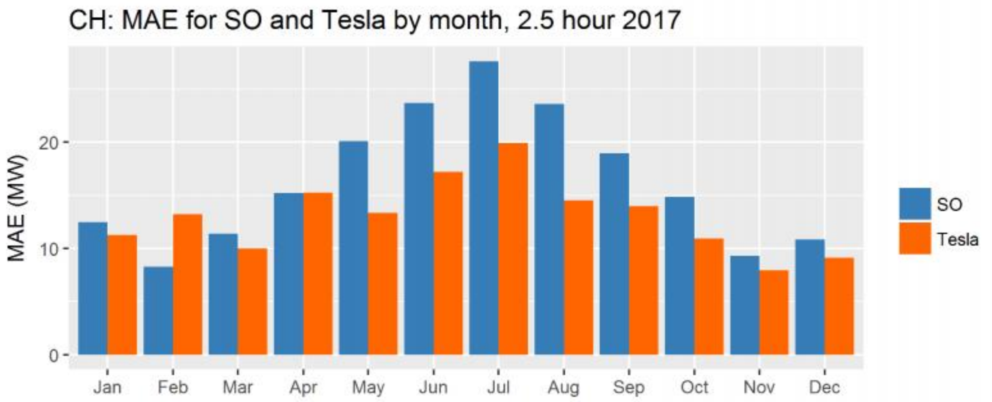


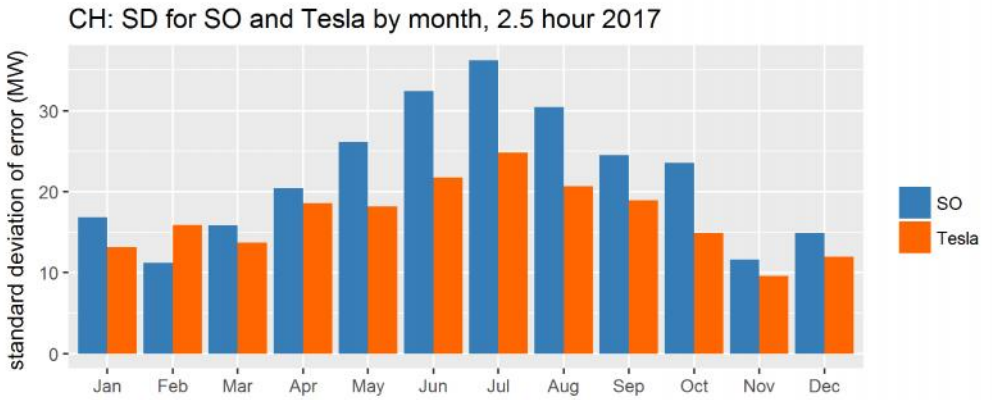


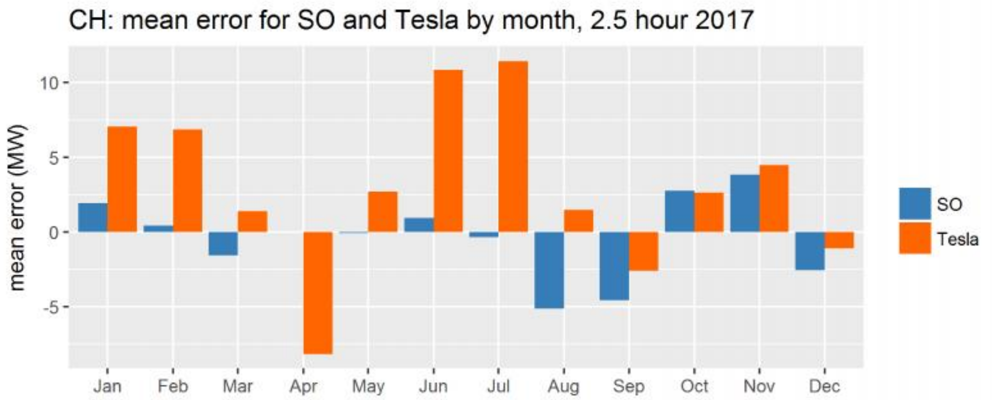


Graphs of summary data by month, comparing SO and TESLA forecast errors at the 2.5 hour forecast horizon. Data for Christchurch.



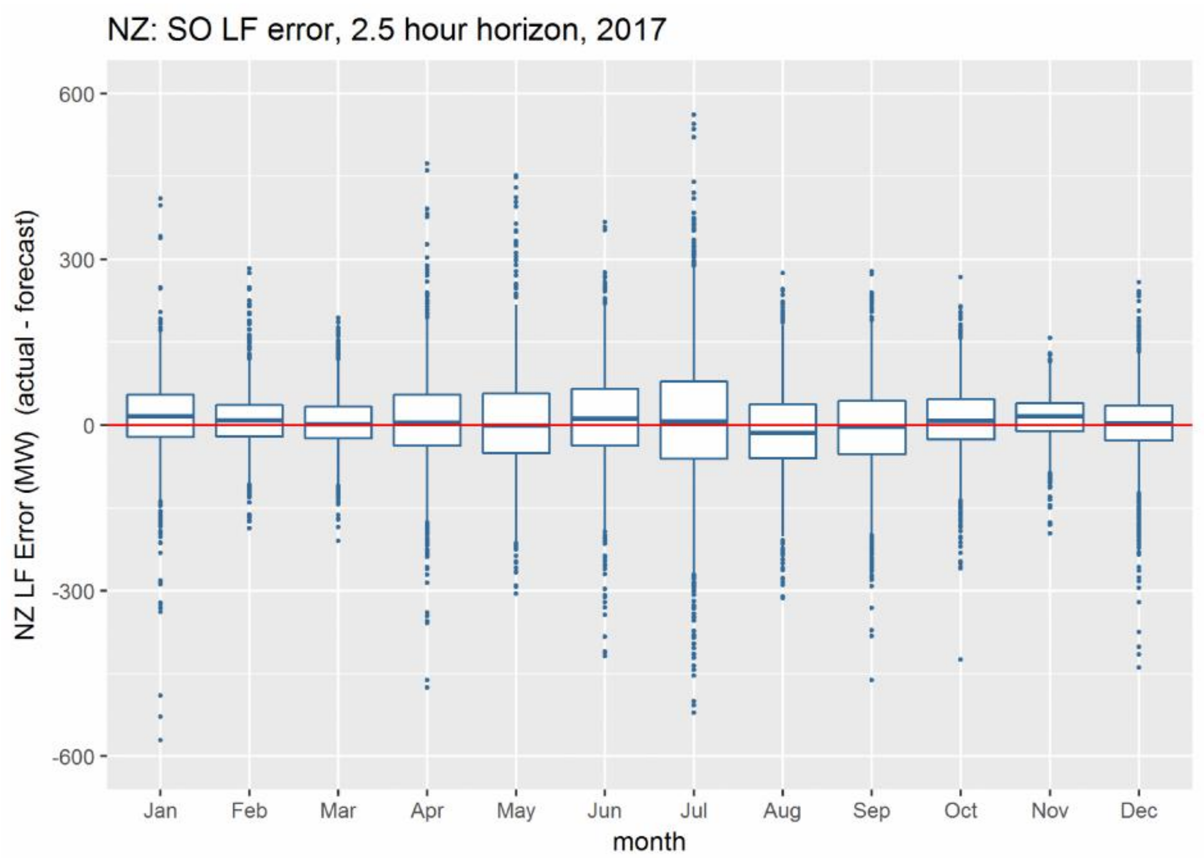


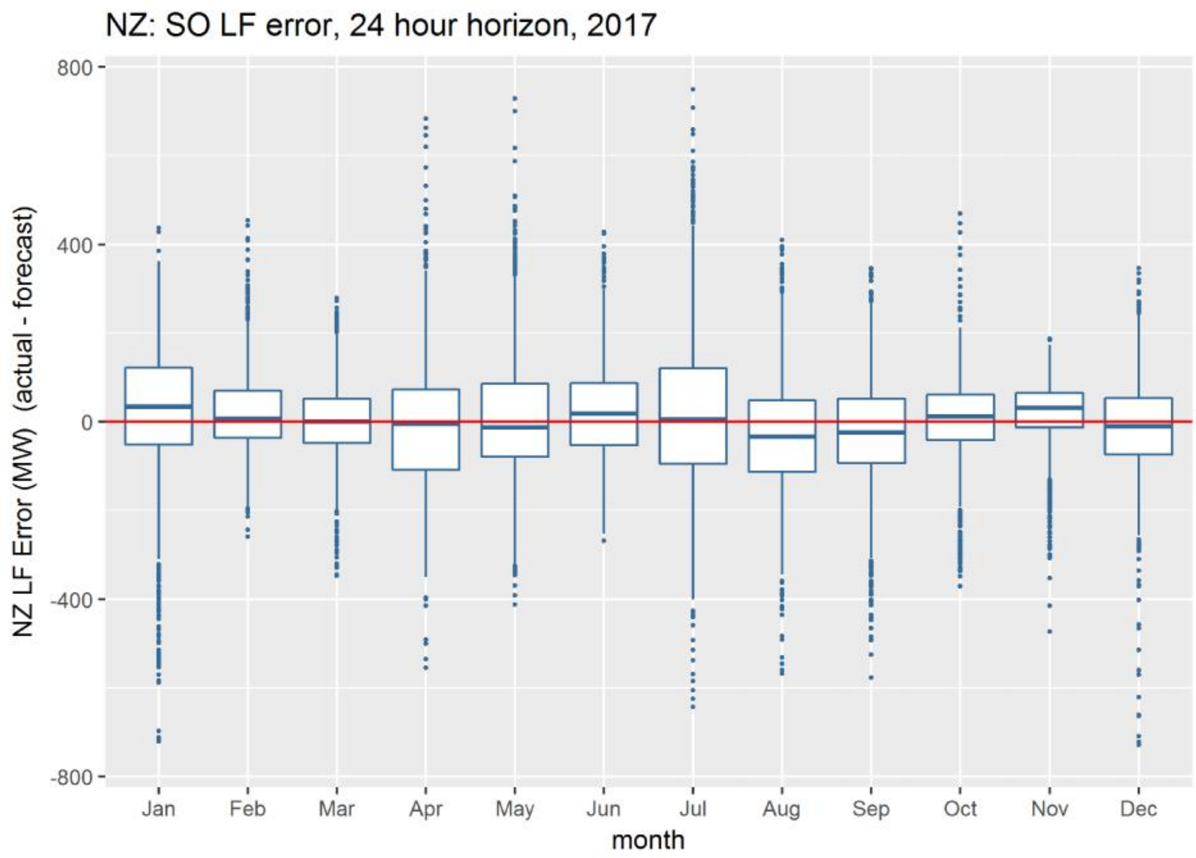


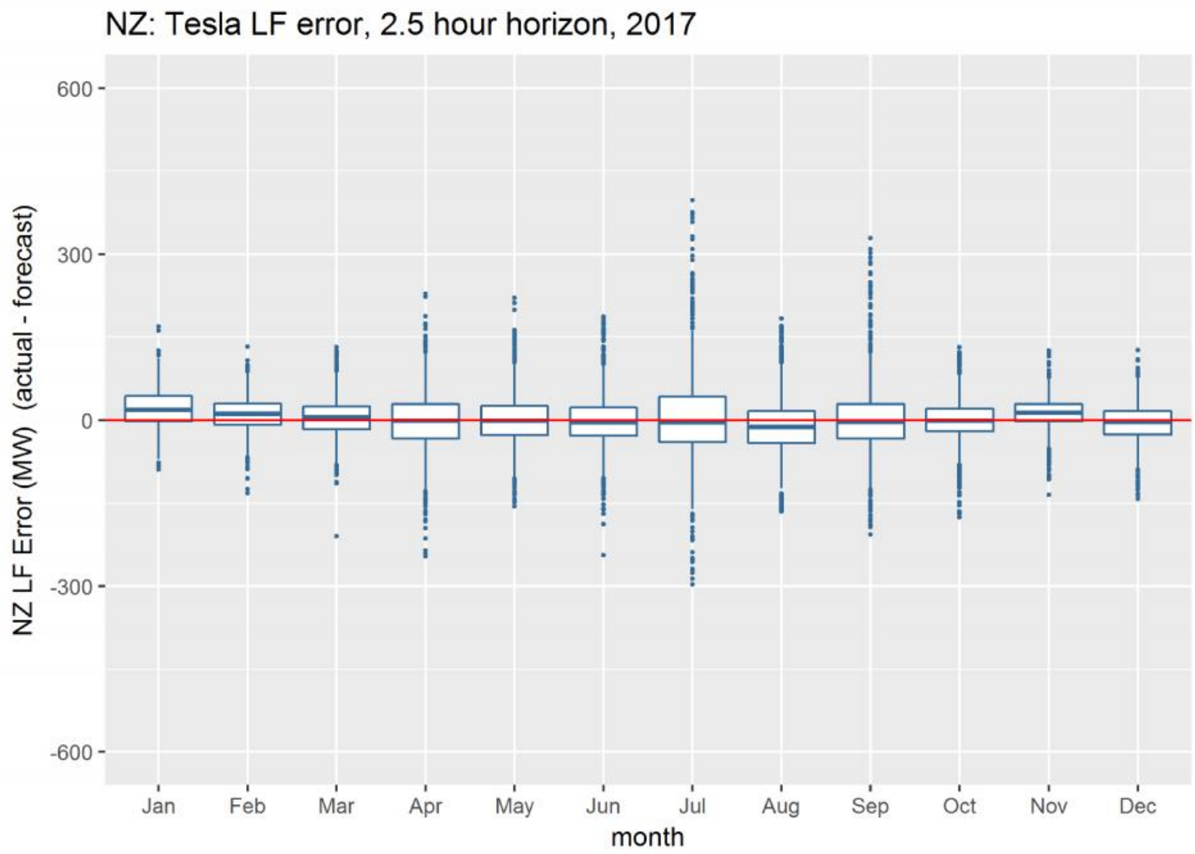


Box and whisker plots for New Zealand for the current (SO) load forecast at 2.5-hour and 24-hour horizons, and for the TESLA load forecast at a 2.5-hour horizon.

The edges and centre of the box indicate the 25th and 75th percentiles (quartiles) and the median respectively. The ends of each 'whisker' line are a maximum of 1.5 times the interquartile range. Data points beyond this are shown as dots and are considered outliers.




NZ: SO LF error, 24 hour horizon, 2017
NZ LF Error (MW) (actual - forecast)
800
400
0
-400
-800
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
month


NZ: Tesla LF error, 2.5 hour horizon, 2017
NZ LF Error (MW) (actual - forecast)
600
300
0
-300
-600
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
month

Error count and difference information, current load forecast, 2.5-hour horizon, NZ wide, 2017

The system operator data in the box and whisker plots on the previous pages are based on load forecasts which are the direct outputs of the load forecast tool. Upon occasion, manual overrides are applied to these forecasts by the coordinators, prior to the data being used as the load input to market schedules such as the NRS. Counts of errors, and differences in counts due to overrides, in particular MW tranches are given in the table below. Negative (red) values indicate improvements in error counts due to manual overrides, and blue (positive) values indicate a worsening in error counts due to manual overrides.

horizon	month	> +/- 400 MW	+/- 300 to 399 MW	+/- 200 to 299 MW	+/- 150 to 199 MW	+/- 100 to 149 MW
2.5 hour	Jan	4	9	12	39	157
2.5 hour	Feb	0	0	9	21	62
2.5 hour	Mar	0	0	1	24	84
2.5 hour	Apr	4	10	41	83	180
2.5 hour	May	5	10	47	78	201
2.5 hour	Jun	3	9	50	85	181
2.5 hour	Jul	18	43	131	117	227
2.5 hour	Aug	0	2	54	98	194
2.5 hour	Sep	1	4	55	81	177
2.5 hour	Oct	1	0	14	42	127
2.5 hour	Nov	0	0	0	5	28
2.5 hour	Dec	3	2	24	38	95

horizon	month	> +/- 400 MW	+/- 300 to 399 MW	+/- 200 to 299 MW	+/- 150 to 199 MW	+/- 100 to 149 MW
2.5 hour w overrides	Jan	4	6	8	38	154
2.5 hour w overrides	Feb	0	0	6	17	62
2.5 hour w overrides	Mar	0	0	2	23	84
2.5 hour w overrides	Apr	0	11	53	85	181
2.5 hour w overrides	May	1	7	50	74	198
2.5 hour w overrides	Jun	3	7	49	87	171
2.5 hour w overrides	Jul	15	43	133	118	235
2.5 hour w overrides	Aug	0	2	54	98	192
2.5 hour w overrides	Sep	1	4	57	83	177
2.5 hour w overrides	Oct	0	0	13	46	121
2.5 hour w overrides	Nov	0	0	0	5	28
2.5 hour w overrides	Dec	3	1	24	38	95

horizon	month	> +/- 400 MW	+/- 300 to 399 MW	+/- 200 to 299 MW	+/- 150 to 199 MW	+/- 100 to 149 MW
difference	Jan	0	-3	-4	-1	-3
difference	Feb	0	0	-3	-4	0
difference	Mar	0	0	1	-1	0
difference	Apr	-4	1	12	2	1
difference	May	-4	-3	3	-4	-3
difference	Jun	0	-2	-1	2	-10
difference	Jul	-3	0	2	1	8
difference	Aug	0	0	0	0	-2
difference	Sep	0	0	2	2	0
difference	Oct	-1	0	-1	4	-6
difference	Nov	0	0	0	0	0
difference	Dec	0	-1	0	0	0

Error count and difference information, current load forecast, 24-hour horizon, NZ wide, 2017

The system operator data in the box and whisker plots on the previous pages are based on load forecasts which are the direct outputs of the load forecast tool. Upon occasion, manual overrides are applied to these forecasts by the coordinators, prior to the data being used as the load input to market schedules such as the NRS. Counts of errors, and differences in counts due to overrides, in particular MW tranches are given in the table below. Negative (red) values indicate improvements in error counts due to manual overrides, and blue (positive) values indicate a worsening in error counts due to manual overrides.

horizon	month	> +/- 400 MW	+/- 300 to 399 MW	+/- 200 to 299 MW
24 hour	Jan	54	33	152
24 hour	Feb	5	12	66
24 hour	Mar	0	9	55
24 hour	Apr	23	47	188
24 hour	May	23	54	149
24 hour	Jun	3	14	93
24 hour	Jul	80	81	196
24 hour	Aug	13	46	150
24 hour	Sep	13	39	129
24 hour	Oct	4	18	49
24 hour	Nov	2	3	23
24 hour	Dec	12	9	78

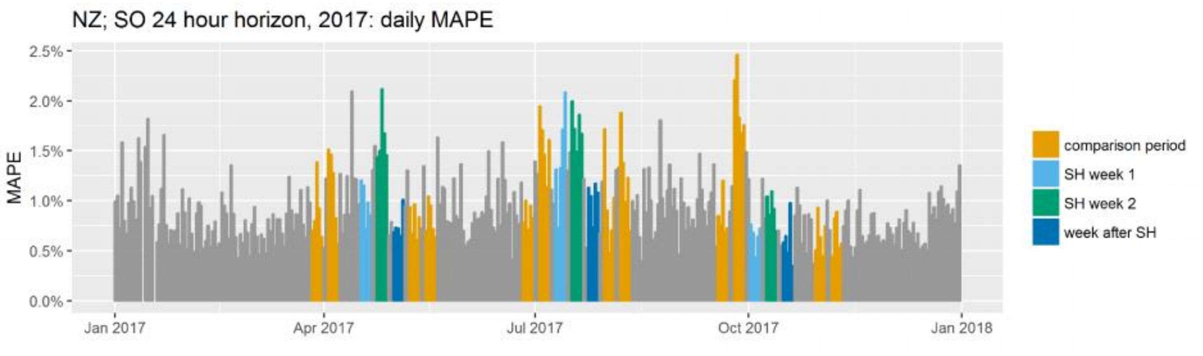
horizon	month	> +/- 400 MW	+/- 300 to 399 MW	+/- 200 to 299 MW
24 hour w overrides	Jan	54	33	155
24 hour w overrides	Feb	4	9	46
24 hour w overrides	Mar	0	9	55
24 hour w overrides	Apr	17	39	196
24 hour w overrides	May	19	48	154
24 hour w overrides	Jun	2	14	94
24 hour w overrides	Jul	79	77	194
24 hour w overrides	Aug	13	46	150
24 hour w overrides	Sep	13	39	129
24 hour w overrides	Oct	4	18	54
24 hour w overrides	Nov	2	3	23
24 hour w overrides	Dec	12	9	78

horizon	month	> +/- 400 MW	+/- 300 to 399 MW	+/- 200 to 299 MW
difference	Jan	0	0	3
difference	Feb	-1	-3	-20
difference	Mar	0	0	0
difference	Apr	-6	-8	8
difference	May	-4	-6	5
difference	Jun	-1	0	1
difference	Jul	-1	-4	-2
difference	Aug	0	0	0
difference	Sep	0	0	0
difference	Oct	0	0	5
difference	Nov	0	0	0
difference	Dec	0	0	0

APPENDIX 2 SEASONAL EFFECTS ON ACCURACY IN THE CURRENT LOAD FORECAST

2A: SCHOOL HOLIDAYS

The graph identifies the days comprising the school-holiday-related categories. The plot is of daily Mean Absolute Percentage Error (MAPE) for the current (SO) 2.5-hour horizon load forecast for all New Zealand. The comparison period consists of 2 weeks either side of each school holiday period, avoiding weeks containing public holidays. Week 1 and week 2 of the school holidays, and the week after the school holidays are considered separately. Weekends are excluded from all categories.



The table on the following page compares 2017 error statistics by school holiday category for the current (SO) load forecast at 2.5-hour and 24-hour horizons, and for the TESLA load forecast at a 2.5-hour horizon across New Zealand and all 10 forecast areas.

SO forecast, 2.5 hour horizon

type	category	NZ	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
MAPE	rest of year	1.6%	3.0%	3.2%	2.8%	2.1%	2.1%	2.3%	2.4%	2.5%	2.2%	0.9%
MAPE	comparison period	1.6%	3.5%	3.9%	3.0%	2.0%	1.8%	2.3%	2.5%	2.6%	2.3%	0.9%
MAPE	SH week 1	2.1%	4.0%	3.9%	3.6%	2.4%	2.4%	2.5%	3.0%	3.0%	2.7%	1.2%
MAPE	SH week 2	1.7%	3.4%	3.6%	3.2%	2.3%	2.0%	2.5%	2.4%	3.2%	2.6%	1.1%
MAPE	week after SH	1.8%	3.4%	3.6%	2.9%	2.2%	2.2%	2.1%	2.5%	2.8%	2.1%	1.1%
MAE	rest of year	55	15	6	9	7	17	4	9	5	5	8
MAE	comparison period	62	19	7	11	7	18	5	10	6	6	8
MAE	SH week 1	81	22	8	13	9	23	6	12	7	7	12
MAE	SH week 2	64	17	7	11	8	17	5	9	7	7	10
MAE	week after SH	71	19	7	11	8	22	5	10	6	6	10
mean error	rest of year	4.3	-0.6	-0.2	-0.2	-0.1	1.8	-0.1	4.0	0.0	0.1	-0.3
mean error	comparison period	4.9	-0.2	0.2	-0.5	-0.2	1.7	0.1	4.0	0.1	0.2	-0.5
mean error	SH week 1	16.9	3.0	1.8	1.7	1.6	4.1	-0.4	4.7	0.0	0.8	-0.4
mean error	SH week 2	-12.2	-3.0	-2.1	-0.5	-2.5	-0.9	-0.8	2.4	-1.9	-0.9	-2.0
mean error	week after SH	23.3	2.7	1.3	1.6	1.4	8.2	0.4	5.1	0.7	0.8	1.2
sd	rest of year	78	22	8	13	10	24	6	11	7	7	12
sd	comparison period	85	26	10	15	10	24	7	15	8	8	12
sd	SH week 1	115	35	11	17	13	35	9	20	10	10	18
sd	SH week 2	89	23	9	15	11	25	7	12	9	9	14
sd	week after SH	101	26	9	14	12	30	7	13	8	8	13

SO forecast, 24 hour horizon

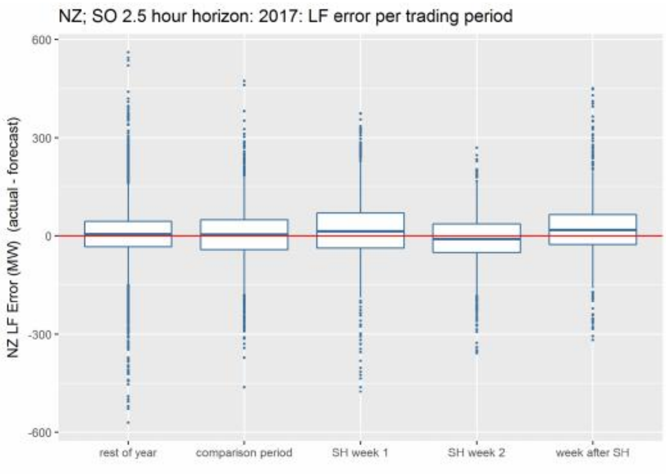
type	category	NZ	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
MAPE	rest of year	2.6%	4.6%	4.9%	4.4%	3.2%	3.4%	3.6%	3.4%	3.9%	3.4%	1.4%
MAPE	comparison period	2.5%	5.4%	5.5%	4.4%	2.9%	2.5%	3.3%	3.4%	3.7%	3.4%	1.4%
MAPE	SH week 1	3.2%	5.7%	4.8%	5.0%	3.8%	3.7%	4.1%	4.4%	4.4%	4.5%	2.0%
MAPE	SH week 2	2.8%	5.6%	5.2%	4.3%	3.7%	3.0%	4.1%	3.4%	4.6%	4.1%	1.5%
MAPE	week after SH	2.4%	3.9%	4.4%	4.4%	3.0%	2.9%	2.7%	3.0%	3.3%	3.1%	1.6%
MAE	rest of year	92	23	9	15	11	28	7	13	8	9	12
MAE	comparison period	96	29	11	16	11	24	7	14	8	9	14
MAE	SH week 1	129	31	10	18	15	37	9	18	10	12	20
MAE	SH week 2	104	29	10	15	13	27	8	13	10	10	14
MAE	week after SH	95	21	8	16	11	28	6	12	7	9	15
mean error	rest of year	2.1	-1.9	-0.3	-0.6	-0.1	1.8	-0.3	4.3	0.0	0.1	-0.7
mean error	comparison period	0.4	1.2	0.8	-0.1	-0.5	-2.0	0.5	3.0	-0.4	0.3	-2.3
mean error	SH week 1	49.2	5.6	4.5	5.2	6.1	14.1	0.3	8.7	0.4	3.2	1.1
mean error	SH week 2	-59.1	-11.3	-6.3	-3.0	-6.8	-8.9	-2.6	-2.4	-6.3	-5.1	-6.3
mean error	week after SH	41.0	9.9	1.9	4.3	2.4	12.8	-0.6	6.1	1.4	2.0	0.7
sd	rest of year	126	34	12	19	15	38	10	16	10	12	16
sd	comparison period	133	39	14	21	14	33	10	18	11	13	19
sd	SH week 1	188	44	11	23	22	57	13	26	14	18	29
sd	SH week 2	127	35	11	20	17	35	10	16	10	12	18
sd	week after SH	136	28	10	20	16	39	8	17	10	11	19

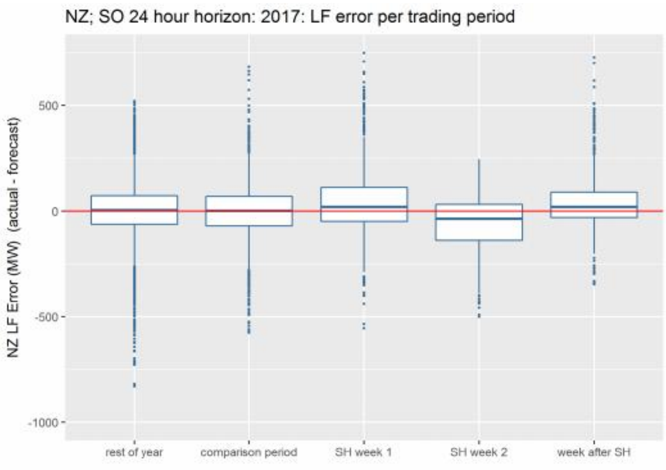
TESLA forecast, 2.5 hour horizon

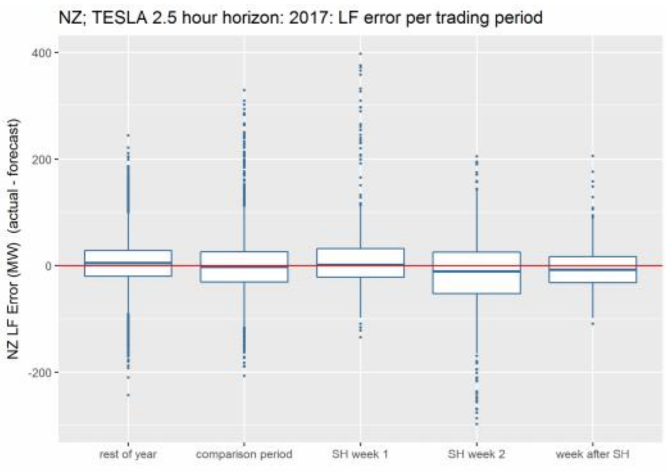
type	category	NZ	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
MAPE	rest of year	0.9%	2.5%	3.3%	2.4%	1.8%	1.4%	1.7%	1.7%	2.2%	2.0%	0.6%
MAPE	comparison period	1.0%	2.5%	3.3%	2.2%	1.7%	1.7%	2.0%	1.9%	2.0%	2.1%	0.6%
MAPE	SH week 1	1.0%	2.8%	4.2%	2.6%	2.0%	1.4%	1.5%	2.0%	2.0%	2.3%	0.6%
MAPE	SH week 2	1.4%	3.0%	3.5%	2.5%	2.2%	1.7%	2.3%	2.1%	2.6%	2.7%	0.9%
MAPE	week after SH	0.8%	2.3%	3.2%	2.1%	1.6%	1.4%	1.3%	2.0%	1.7%	1.9%	0.7%
MAE	rest of year	32	13	6	8	6	11	3	6	4	5	5
MAE	comparison period	40	14	6	8	6	16	4	7	4	5	6
MAE	SH week 1	42	16	9	9	7	14	3	8	4	6	6
MAE	SH week 2	56	16	7	9	8	16	5	8	6	7	9
MAE	week after SH	31	13	6	8	6	13	3	8	4	5	7
mean error	rest of year	4.4	3.5	1.7	2.0	-1.3	-0.5	-0.5	-0.3	-0.5	-0.8	1.0
mean error	comparison period	1.2	1.7	0.2	0.0	-0.9	-0.5	0.3	-1.5	0.9	-1.1	2.1
mean error	SH week 1	12.8	0.8	4.6	-2.3	3.3	4.9	0.4	-0.4	0.6	-0.5	1.4
mean error	SH week 2	-17.4	2.3	0.2	-2.3	1.2	-7.9	-0.1	-4.8	-0.6	-3.7	-1.7
mean error	week after SH	-6.5	3.3	-1.7	-0.2	2.6	-2.2	-1.1	-4.0	-0.5	-1.6	-1.0
sd	rest of year	42	17	8	10	8	17	5	8	6	6	8
sd	comparison period	57	20	9	10	9	23	6	10	6	7	9
sd	SH week 1	68	22	10	12	10	22	6	12	6	8	9
sd	SH week 2	75	22	9	12	10	21	7	9	8	8	12
sd	week after SH	39	17	7	10	8	17	4	9	5	6	10

Box and whisker plots for New Zealand for the current (SO) load forecast at 2.5-hour and 24-hour horizons, and for the TESLA load forecast at a 2.5-hour horizon. The school holiday categories are compared in each graph.

The edges and centre of the box indicate the 25th and 75th percentiles (quartiles) and the median respectively. The ends of each 'whisker' line are a maximum of 1.5 times the interquartile range. Data points beyond this are shown as dots and are considered outliers.

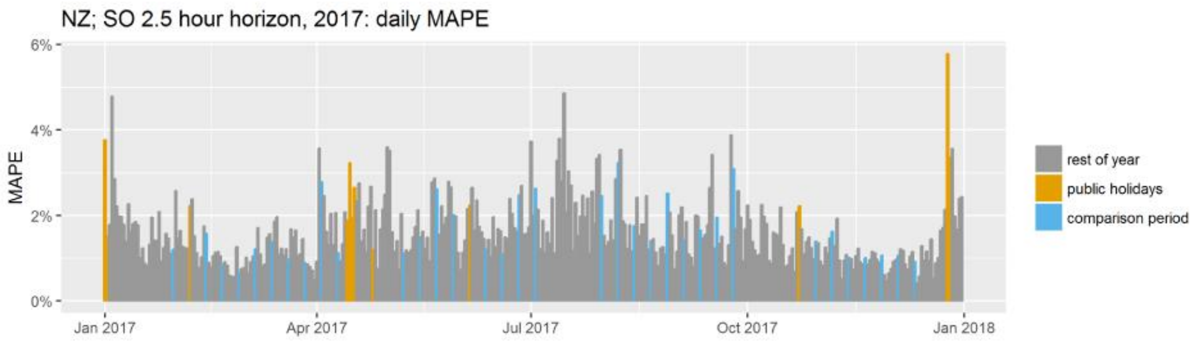






2B: Public Holidays

The graph identifies the days comprising the public-holiday-related categories. The plot is of daily Mean Absolute Percentage Error (MAPE) for the current (SO) 24-hour horizon load forecast for all New Zealand. The public holidays category includes all nationally-celebrated public holidays. The comparison period is made up of all Mondays excluding those that fall in the school holidays.



The table compares 2017 error statistics by public holiday category for the current (SO) load forecast at 2.5-hour and 24-hour horizons, and for the TESLA load forecast at a 2.5-hour horizon across New Zealand and all 10 forecast areas.

SO forecast, 2.5 hour horizon

type	category	NZ	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
MAPE	rest of year	1.6%	3.2%	3.4%	2.8%	2.0%	2.0%	2.3%	2.4%	2.5%	2.2%	0.9%
MAPE	public holidays	2.7%	3.8%	3.6%	3.7%	3.6%	3.1%	3.3%	3.0%	4.1%	2.9%	1.4%
MAPE	comparison period	1.6%	3.0%	3.6%	3.0%	2.2%	2.2%	2.2%	2.5%	2.7%	2.2%	0.9%
MAE	rest of year	57	16	6	10	7	17	5	9	5	6	8
MAE	public holidays	78	16	6	11	11	20	5	10	6	6	10
MAE	comparison period	62	17	7	11	8	20	5	10	6	6	8
mean error	rest of year	6.1	-0.3	-0.1	0.1	0.1	2.1	0.0	4.1	0.2	0.3	-0.3
mean error	public holidays	-9.3	-2.2	-0.5	-2.0	-3.9	-0.9	-0.8	3.4	-1.4	-1.3	0.3
mean error	comparison period	-0.3	-0.1	-0.1	-1.3	-0.1	2.5	-0.2	3.0	-1.5	-1.3	-1.1
sd	rest of year	81	24	9	13	10	25	7	12	7	8	12
sd	public holidays	105	21	8	14	18	26	7	13	8	8	13
sd	comparison period	86	24	10	14	11	29	7	13	8	8	12

SO forecast, 24 hour horizon

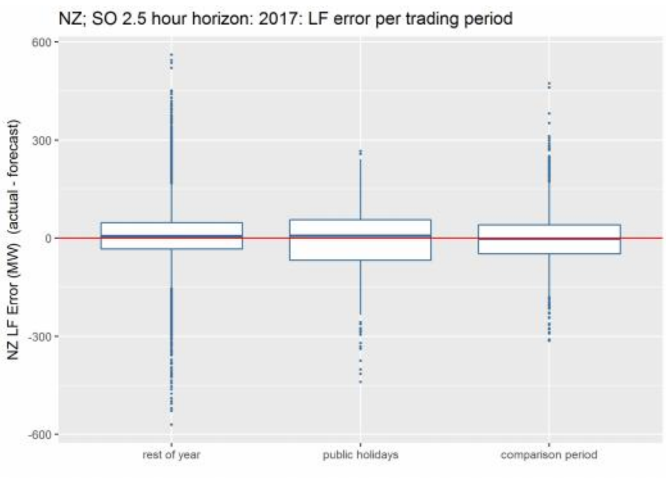
type	category	NZ	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
MAPE	rest of year	2.6%	4.8%	5.0%	4.4%	3.1%	3.1%	3.5%	3.4%	3.8%	3.4%	1.4%
MAPE	public holidays	4.8%	6.7%	5.8%	6.8%	5.7%	5.6%	5.5%	4.3%	7.5%	4.6%	2.1%
MAPE	comparison period	2.4%	4.6%	4.7%	4.4%	3.4%	3.0%	3.3%	3.7%	3.8%	3.3%	1.3%
MAE	rest of year	93	24	9	15	11	27	7	13	8	9	13
MAE	public holidays	141	28	10	20	17	37	9	14	12	10	14
MAE	comparison period	93	25	9	15	12	28	7	14	8	9	12
mean error	rest of year	7.2	-0.5	0.0	0.5	0.3	2.2	0.0	4.5	0.4	0.7	-0.8
mean error	public holidays	-90.8	-17.9	-2.8	-12.0	-12.6	-21.9	-4.6	0.3	-9.0	-6.2	-4.0
mean error	comparison period	-4.9	0.5	-0.2	-2.3	0.7	4.9	-0.8	0.8	-3.2	-2.9	-2.5
sd	rest of year	129	34	12	20	15	38	10	17	10	12	18
sd	public holidays	163	32	12	22	22	44	11	19	12	11	19
sd	comparison period	130	49	12	20	17	38	10	19	10	11	17

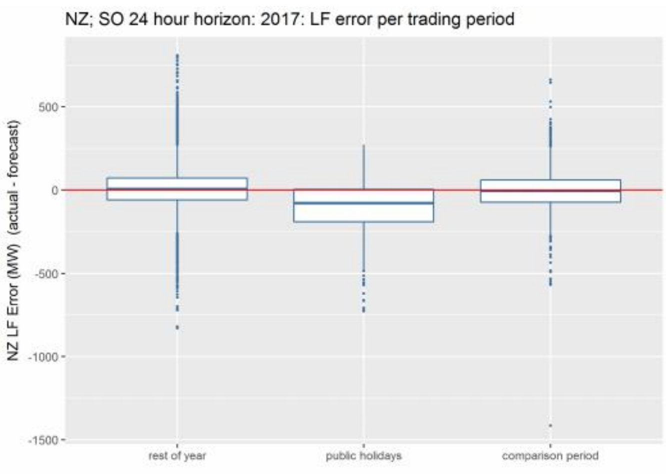
TESLA forecast, 2.5 hour horizon

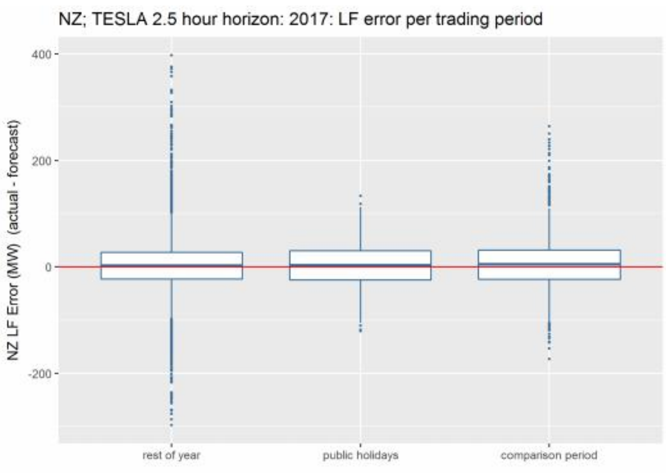
type	category	NZ	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
MAPE	rest of year	0.9%	2.5%	3.3%	2.3%	1.8%	1.4%	1.8%	1.8%	2.1%	2.0%	0.6%
MAPE	public holidays	1.1%	3.2%	3.8%	3.5%	1.9%	1.7%	2.2%	2.0%	3.0%	2.3%	0.9%
MAPE	comparison period	1.0%	2.3%	3.3%	2.3%	1.8%	1.6%	1.7%	1.7%	1.9%	2.0%	0.6%
MAE	rest of year	34	13	6	8	6	12	4	7	4	5	6
MAE	public holidays	33	14	7	10	6	12	4	6	5	5	6
MAE	comparison period	38	13	7	8	6	14	4	7	4	5	6
mean error	rest of year	2.6	3.3	1.3	1.2	-0.8	-1.0	-0.3	-0.9	-0.2	-1.0	0.9
mean error	public holidays	2.6	-2.7	1.9	2.4	1.9	0.6	-2.0	-0.7	-1.3	-0.5	2.9
mean error	comparison period	5.7	2.3	1.9	1.4	-1.4	2.2	0.0	-1.0	0.1	-1.3	1.6
sd	rest of year	48	18	8	10	9	18	5	9	6	7	8
sd	public holidays	42	19	8	13	7	16	4	8	7	7	8
sd	comparison period	52	17	9	11	9	22	5	9	5	7	9

Box and whisker plots for New Zealand for the current (SO) load forecast at 2.5-hour and 24-hour horizons, and for the TESLA load forecast at a 2.5-hour horizon. The three public holidays categories are compared in each graph.

The edges and centre of the box indicate the 25th and 75th percentiles (quartiles) and the median respectively. The ends of each 'whisker' line are a maximum of 1.5 times the interquartile range. Data points beyond this are shown as dots and are considered outliers.

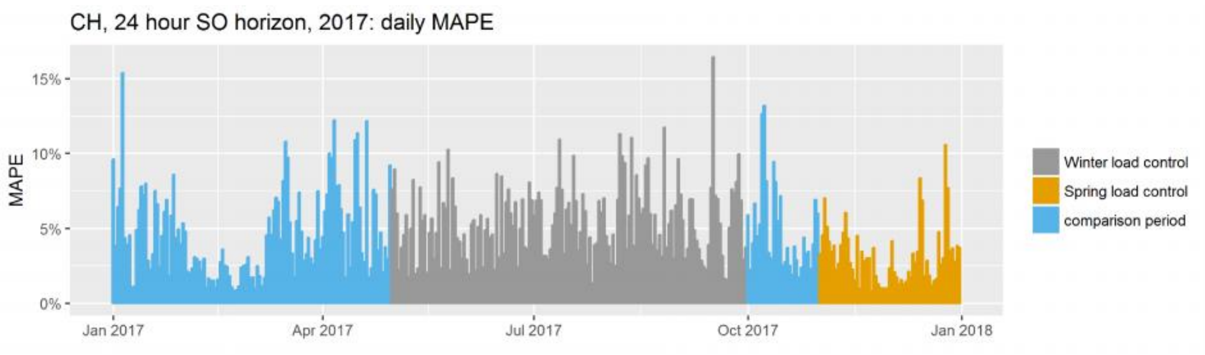






2C: Load Control

The graph identifies the days comprising the network load control related categories used. The plot is of daily Mean Absolute Percentage Error (MAPE) for the current (SO) 24-hour horizon load forecast for Christchurch forecast area. The load control periods are as indicated by the network operator, Orion. These affect the CH and, to a lesser extent, WC forecast zones, and in turn, the NZ-wide data. The load control periods are defined as from the first day of load control, to 2 weeks after the last day. Load control primarily occurred in Winter, with some instances in Spring. The comparison period is made up of the rest of the year. Our analysis focuses on the winter period.



The table compares 2017 error statistics by load control category for the current (SO) load forecast at 2.5-hour and 24-hour horizons, and for the TESLA load forecast at a 2.5-hour horizon across New Zealand and all 10 forecast areas. The ratio of MAPE for the load control period : comparison period is also given.

SO forecast, 2.5 hour horizon

type	category	NZ	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
MAPE	winter load control	1.9%	4.1%	3.9%	3.4%	2.4%	2.3%	2.6%	2.9%	2.9%	2.4%	1.1%
MAPE	comparison period	1.5%	2.7%	3.1%	2.7%	2.0%	2.0%	2.1%	2.1%	2.5%	2.2%	0.9%
MAE	winter load control	74	23	7	12	9	22	6	11	6	6	11
MAE	comparison period	50	13	6	9	6	15	4	8	5	5	7
mean error	winter load control	2.2	-1.8	0.0	-0.5	-0.1	1.8	-0.1	3.9	-0.2	0.0	-0.8
mean error	comparison period	7.8	0.7	0.1	0.1	-0.1	2.5	0.0	4.1	0.1	0.3	0.1
sd	winter load control	101	30	10	16	13	31	8	15	9	9	16
sd	comparison period	70	18	8	11	9	22	5	10	6	7	10
MAPE ratio	load control / comparison	1.25	1.54	1.25	1.26	1.22	1.19	1.26	1.35	1.17	1.07	1.29

SO forecast, 24 hour horizon

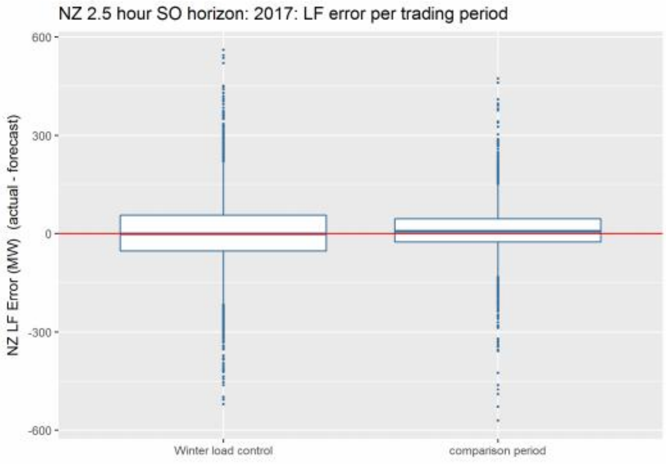
type	category	NZ	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
MAPE	winter load control	2.8%	5.5%	5.2%	4.8%	3.5%	3.3%	3.8%	4.1%	4.0%	3.6%	1.6%
MAPE	comparison period	2.7%	4.7%	5.1%	4.4%	3.2%	3.3%	3.5%	3.1%	3.9%	3.7%	1.4%
MAE	winter load control	108	30	10	17	13	31	8	16	9	10	15
MAE	comparison period	89	22	10	14	10	26	6	11	7	9	12
mean error	winter load control	-2.2	-4.0	-0.4	-0.6	-0.1	1.5	-0.3	4.5	-0.9	0.1	-1.8
mean error	comparison period	8.0	1.1	0.6	-0.4	-0.1	2.8	0.1	3.6	0.1	0.3	-0.2
sd	winter load control	147	43	12	22	18	43	12	21	12	13	21
sd	comparison period	126	31	12	19	14	37	9	14	10	13	16
MAPE ratio	load control / comparison	1.05	1.18	1.01	1.09	1.09	0.99	1.08	1.33	1.03	0.98	1.13

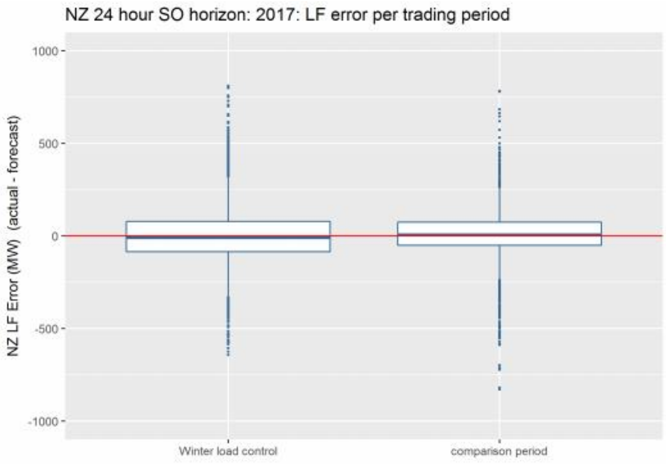
TESLA forecast, 2.5 hour horizon

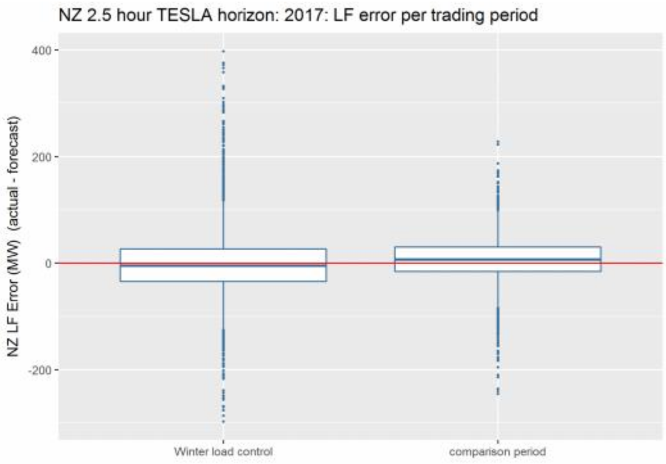
type	category	NZ	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
MAPE	winter load control	1.1%	2.8%	3.2%	2.2%	1.9%	1.7%	2.0%	2.0%	2.1%	2.2%	0.7%
MAPE	comparison period	0.9%	2.6%	3.5%	2.4%	1.8%	1.3%	1.7%	1.6%	2.2%	2.0%	0.6%
MAE	winter load control	42	16	6	8	7	16	4	8	5	6	7
MAE	comparison period	31	12	7	8	6	11	3	6	4	5	5
mean error	winter load control	-1.1	4.8	-0.1	-0.6	-1.5	-2.0	-0.8	-1.0	0.3	-1.9	1.6
mean error	comparison period	6.6	1.9	3.8	0.6	0.4	1.0	0.2	-1.0	-0.4	-0.3	0.5
sd	winter load control	59	22	8	10	10	23	6	11	7	7	10
sd	comparison period	41	16	8	10	8	15	4	7	6	6	7
MAPE ratio	load control / comparison	1.18	1.09	0.91	0.92	1.03	1.29	1.19	1.28	0.99	1.11	1.21

Box and whisker plots for New Zealand for the current (SO) load forecast at 2.5-hour and 24-hour horizons, and for the TESLA load forecast at a 2.5-hour horizon. The winter load control period is compared with the rest of the year. in each graph.

The edges and centre of the box indicate the 25th and 75th percentiles (quartiles) and the median respectively. The ends of each 'whisker' line are a maximum of 1.5 times the interquartile range. Data points beyond this are shown as dots and are considered outliers.

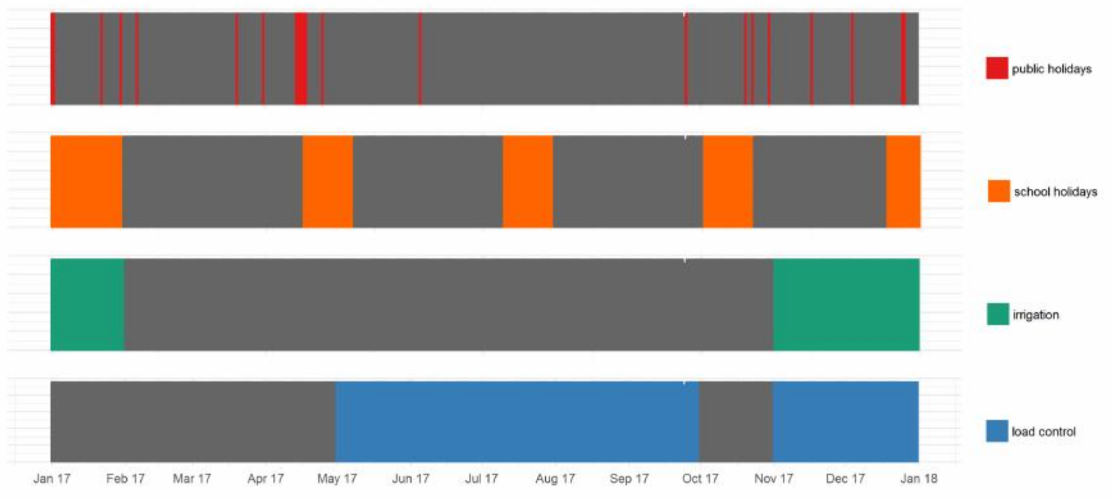






2D: All Seasons

Four seasonal effects which impact the load forecast accuracy are shown in the graph; public holidays, school holidays, load control and irrigation.



The table compares 2017 error statistics by all seasons category for the current (SO) load forecast at 2.5-hour and 24-hour horizons, and for the TESLA load forecast at a 2.5-hour horizon across New Zealand.

year1	horizon	category	MAE	MAPE	mean_error	sd
2017	2.5 hour SO	comparison period	45.6	1.3%	11.8	64.2
2017	2.5 hour SO	all seasons	61.4	1.7%	3.4	86.5

year1	horizon	category	MAE	MAPE	mean_error	sd
2017	24 hour SO	comparison period	79.6	2.2%	23.2	109.1
2017	24 hour SO	all seasons	98.2	2.7%	-2.2	136.1

year1	horizon	category	MAE	MAPE	mean_error	sd
2017	2.5 hour Tesla	comparison period	28.9	0.8%	10.7	37.3
2017	2.5 hour Tesla	all seasons	35.9	1.0%	0.9	50.4

APPENDIX 3 DETAILED OPTION ASSESSMENT AND COMPARISON

3A: IMPROVE CURRENT FORECAST OPTIONS

We will first consider the options which are based on improvements to the current forecast, namely Options 1 - 3.

Overview of criteria fulfillment

Accuracy

Some accuracy improvements are achievable within current forecast options, but accuracy is limited by the sophistication of the models, and inflexibility of the tool limiting the effectiveness of intervention.

Cost

Costs for current forecast improvement options range from minimal to medium.

Cost effectiveness in the future

The current forecast tool is unlikely to accommodate Future States A or B with any reasonable level of accuracy. The future cost for all current forecast improvement options is therefore assumed to be the cost of a brand new system, including base level infrastructure and new model, as well as infrastructure to incorporate GXP level forecasts and price response in the market system

None of the current forecast options will be able to maintain sufficient accuracy in Future State A, and better forecasts will be accommodated in Future State B.

Current forecast options have comparatively simplistic models and therefore require significant human intervention to achieve high levels of accuracy in any state. The complexities in Future State A are likely to render such an approach impracticable.

In Future State B, whether or not Future State A is concurrent, high value-for-money improvements will be possible by adopting new models. The cost of staff required to intervene in the system is not likely to change, though data science expertise may replace procedural requirements specific to the tool.

Improvements to our current load forecast are only workable if neither GXP level forecasting or price response incorporated in the market system are required.

Option 1 (SOW Option 1): Quick wins

Applying the quick wins is likely to bring some improvement in average accuracy and number of outliers, which will be equally important for the future, and can be achieved with minimal cost.

There are four rapid deployment adjustments considered:

- Increase the precision – number of decimal places - of the temperature inputs
- Get up to date population weightings of different temperature measurements within a forecast area
- Provide greater temporal granularity of weather forecasts
- More diligent human intervention

The first three of these improve weather data. All three improve weather forecasts; the first two also improve weather actuals.

Deployment of these adjustments is likely to improve average load forecast accuracy, particularly when weather is more extreme than usual. It is not expected to reduce the number of outliers, though is likely to reduce the size of some. It is likely to have little impact on accuracy at times of market stress. It will be equally important in the future. It is not expected to incur a cost. The impact of improved weather data will be mostly attributable to the more accurate weather forecasts, including profiling improvements, as opposed to the more accurate weather actuals.

Sensitivity of actual load to weather is expected to differ according to time of day, and be greater when temperature is particularly high or low. Forecasts of weather are also less reliable at the extremes, so the overall effect of weather forecast inaccuracies on load forecasts is expected to be greatest in the high and low ranges.

Weather forecast inaccuracies are considered a major source of error to international experts at PJM, AEMO, and MISO. Weather can be more extreme in their jurisdictions however, impacting load more compared to New Zealand.

Table 8 shows a comparison of available error statistics for the load forecast and the load forecast model. Since the load forecast is based on forecast weather inputs, and the load forecast model is based on actual weather inputs, the difference between these MAPEs gives some indication of the relative proportion of the errors which could be attributed to the quality of weather forecast data, and the maximum improvements which could be gained by improving the current forecast data.[25]

Table 8. MAPE statistics for Dec 2017 at a 24- hour forecast horizon. Forecast statistics (based on weather forecasts) are compared with model statistics (based on actual weather observations)

	average across areas	average excl WN	CH	WC	IN	NL	AK	BP	HM	NR	PN	WN
Load forecast MAPE	3.2%	3.5%	3.3%	4.2%	4.2%	2.6%	3.5%	3.3%	3.1%	4.4%	2.7%	1.1%
Load forecast model MAPE	3.0%	2.9%	2.8%	3.4%	3.5%	2.1%	2.9%	2.8%	2.5%	3.7%	2.5%	3.3%

The four rapid deployment adjustments are discussed in further detail below.

Temporal Granularity

Providing greater temporal granularity of weather forecasts will improve the forecasts.

Weather forecasts are currently provided for 6 hourly increments (midnight, midday, 6am and 6pm), with the times in between filled in based on profiles from historic weather actuals (see profiling effect in Figure 9). The weather forecast accuracy for these in-between times will be improved by using actual, rather than profiled, forecasts.

[25] Note, the load forecast MAPE should be greater than the model MAPE. The load forecast used for the load forecast MAPE above, however, had a slightly different underlying model than that used for the model MAPE. While we don't expect this to have a significant impact, it is notable that the Wellington region MAPE is larger for the model than the forecast, the opposite of what we expected. This is likely attributable to one model using the previous year's December data in the historic dataset used to create the model, and the substantial differences in weather between December 2016 and 2017 in Wellington.

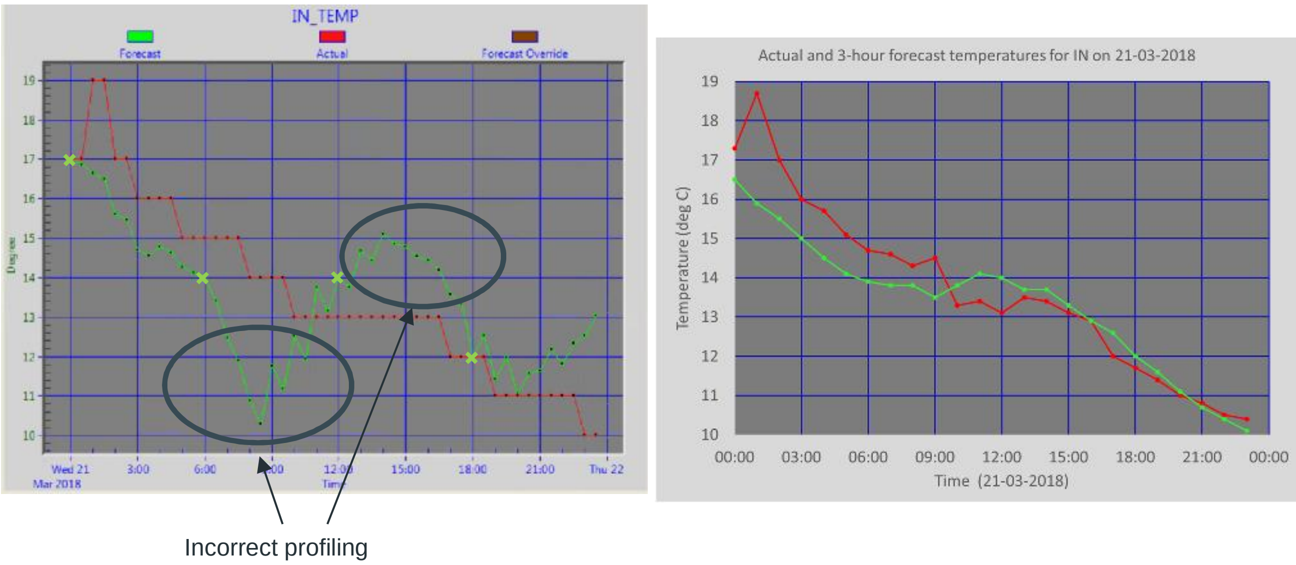


Figure 9 Comparison of currently used forecast and actual temperatures (LHS) with more precise and granular data available (RHS). The incorrect profiling determined by the current model is indicated.

Precision

Increasing the precision of temperature forecast data means increasing the accuracy of temperature forecasts. Temperature data is currently provided to 0 decimal places.

Increasing the precision by one decimal place could change the forecast value by up to 0.5 degrees.

Increasing the precision of actual temperature data will also provide some improvement, as this data is used in model estimation.

Population Weightings

Getting up to date population weightings is likely to increase the accuracy of the load forecast. This will improve both the weather forecast, and the weather actuals.

The temperature data, forecasts and actuals, used for a forecast area is the weighted average of a few different measurement points within the area. The weightings are based on population values at the time we updated forecast areas, in 2002. Populations have changed substantively during that time. Additional measurement points could also be used for Wellington (Lower Hutt and Porirua) and Auckland (Whenuapai), where few points are currently utilised.

Human Intervention

More diligent human intervention will provide improvements in accuracy mainly in number of outliers and accuracy of constraints. It will provide small improvement in average accuracy, negligible impact during stressed conditions, will be equally advantageous in the future, and will incur little cost. Such intervention could include

- Process improvements to ensure accurate update of moveable holidays such as Easter.
- Prompt update of factors that apportion load between GXPs within an area. This will lead to more accurate forecast constraints.

- Prompt update of data quality issues. This requires enabling of sufficient alerts, and staff development. This would reduce the number of outliers in the load forecast error, but due to its infrequent occurrence, isn't likely to have a large impact on average accuracy.

  An example of such data quality issues relates to SCADA quality flags. The actual load data for a given area currently becomes zero when the data quality of a single GXP within the area is flagged as suspect within SCADA. This affects the forecast via both the refined component and the historic load data used in subsequent days' forecasts. Specifically,

  - The impact on the refined component causes large decreases in forecast loads, toward a default minimum close to zero, for all periods of the same day. This occurs for all forecasts until the underlying issue is fixed, or a coordinator manually overrides the forecast.
  - The impact on historic load means decreased forecasts for future days, for the periods when the issue occurred. The impact is most prevalent for the next few days, but can still be noticeable for several weeks, for days of the same day type.

Option 2 (SOW Option 2): Refresh current model

For this option, the rapid deployment adjustments are assumed to have been applied.

Refreshing and re-calibrating the model is very likely to improve average accuracy and prevalence of outliers, but finding the best model may take some weeks.

The model is current in the sense that it always runs on the latest year's data, updated daily. However, the adjustable parameters within the model, related to long-term, short-term, refined, and weather components, have not been re-calibrated since the model was developed.

It is likely a much better model can be found, given the current model was developed a long time ago (1996). The market has changed since 1996, with increased network load control, increased irrigation load, massive changes in heating and air conditioning, particularly the uptake of heat pumps, increased insulation, etc. None of these determinants are a specified part of the model, but they are likely to have an impact on adjustable parameters. For example, greater use of heat pumps and air conditioning may have altered the relationship between weather variables and load from that currently specified. Increased load control may mean that better overall accuracy can now be achieved by configuring to alter the balance between recent and older load profiles.

We achieved greater accuracy with the first model we trialed, primarily as proof of concept. As minimal effort was applied to designing this model, we expect larger improvements are possible.

A few trial model runs demonstrated that different models were suitable to different areas. This suggests potential for improvement by customising models by area, as the current approach is fairlyuniform across areas.

Creating a new model is expected to take some time and require dedicated and trained staff and/or consultants. This is because suitable relationships between inputs and load need to be determined and the nature of the output provided by the load forecast tool is not conducive to rapid comparison of potential model updates.

The improvement achievable from a single model re-calibration is likely to be limited, due to the model's inability to account for discrete seasonal changes. This option is based on a one-off refresh of the model. Several discrete seasonal effects are known to exist in load behaviour currently, for example irrigation, school holidays and network load control. Different models will be appropriate for different seasons. For these reasons, a single model will need to compromise. This is especially true without any new inputs.

Option 2a (variation of SOW Option 2): Refresh current model + some intervention

This option assumes changes in the above options have been implemented.

Human intervention is expected to provide a reasonable improvement in average accuracy and in outliers, with some improvement during stressed conditions, but at some cost.

Human intervention enables the development of a number of different models, and provides the ability to change between these models on a seasonal or daily/weekly basis, as required.

Human intervention is expected to improve forecasts by better handling discrete seasonal effects, changing load behaviour over time, and temporal load shifts. Two of the biggest current applications of this are believed to be school holidays and network load control.

Recent analyses support the hypothesis that network load control and school holidays are major problems for the current load forecast. For periods where school holidays and network load control were not expected to have a major effect, the differences between TESLA and the current forecast were much smaller (see Appendix 2). This suggests significant improvement can be made to the forecast if these issues can be dealt with.

Addressing impact of seasonal effects

The greatest benefit from addressing seasonal effects is likely due to improved accuracy throughout the year as the model can make better use of the historic load and weather data. The current model has a high reliance on very recent data, as long-term data too often represents very different load behaviour, from a different season. Intervention to address seasonal effects would mean that long-term data could be appropriately utilised. The impact of weather also gets distorted as the model can't predict its relationship with load when colder weather leads to decreased load due to load control.

Trial model adjustments have shown that different models achieve better accuracy at different times of the year (and in different areas).

Human intervention for school holidays

Human intervention is likely to improve how the current load forecast deals with school holidays. Creating new variables to handle school holidays has already been mentioned as a potential improvement. The amount of improvement achievable by this is limited by the model's inability to account for a separate load profile for school holidays. Different kinds of load are used in school holidays, at different times of the day. School children and some parents may get up later in the morning, and play video games rather than showering, for example.

Figure 10 compares average load for the July school holiday period (from 2015 – 2017) with average load for the week before and after. The morning peak is shifted to ~ 0.5 hours later and the morning load curve is broader after the peak. The same behaviour was seen across all the 2-week school holiday periods. In addition, the school holiday inter-peak load is a higher proportion relative to the morning peak load. This effect was strongest in the July holidays, and is consistent with additional residential home heating being used during the day.

A given model sets the way historic data is used to determine the load profile (i.e. long and short components used as inputs in the model), and the importance of the load profile compared to other variables. When there are discrete seasons, historic data from a different season is less relevant for the load profile. The current forecast is particularly inaccurate during the first week of school holidays and during the first week after, when the historic data from a different season is a significant component of the load profile (see Table 9).

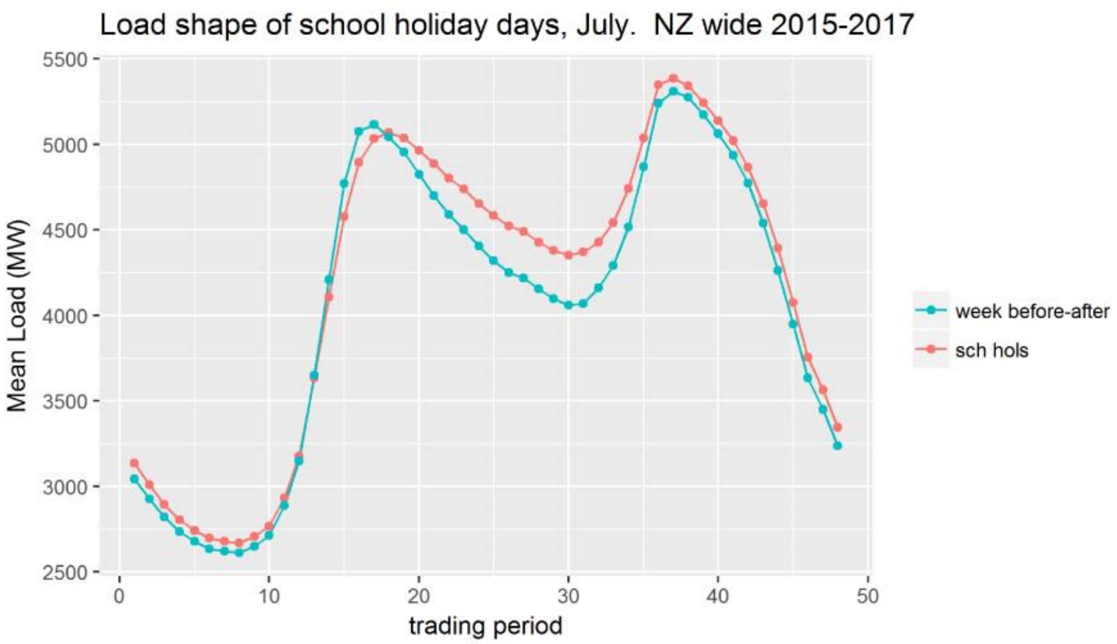


Figure 10 Average actual load for the July school holiday period, compared to the weeks before and after. Data from 2015 – 2017.

Table 9 Comparison of current forecast and TESLA errors for school holiday seasons

2017 NZ wide sch_hol	SO 2.5 hr MAPE	TESLA 2.5 hr MAPE	SO 24 hr MAPE
Comparison week	1.6%	1.0%	2.5%
SH week 1	2.1%	1.0%	3.2%
SH week 2	1.7%	1.4%	2.8%
Week after SH	1.8%	0.8%	2.4%

The model can be altered at the start and middle of new seasons, to change the way historic data is used to determine the load profile. This will provide some improvement, but the improvement will still be limited.

Altering recent historic data in addition to altering the model will bring the most improvement.

There will still be a transition period where the forecasts are poor. If school holiday days are forecast prior to any alterations, perhaps within longer forecast horizons, these forecasts will remain poor. If alterations are made before the school holidays start, forecasts for days before the school holidays start will likely be poor.

Human Intervention for network load control

Human intervention is likely to improve how the current load forecast deals with network load control.

Human intervention during every load control day could make significant improvements to the forecasts. Load control could be predicted and accommodated within the tool, if we knew the load control decision criteria of network companies and the GXPs belonging to their network.

Historic data can be normalised for the effect of load control, ensuring uncontrolled load is forecast. This would allow prediction of load control by assessing against the control criteria, if an area was set up representing the monitored network. Correction could then be made by reducing peak loads and shifting the reduced energy into the surrounding periods.

Figure 11 demonstrates how the current forecast fails to accommodate load control, the red areas showing over-forecasts, the blue under-forecasts, for two days in Christchurch during the 2017 winter.

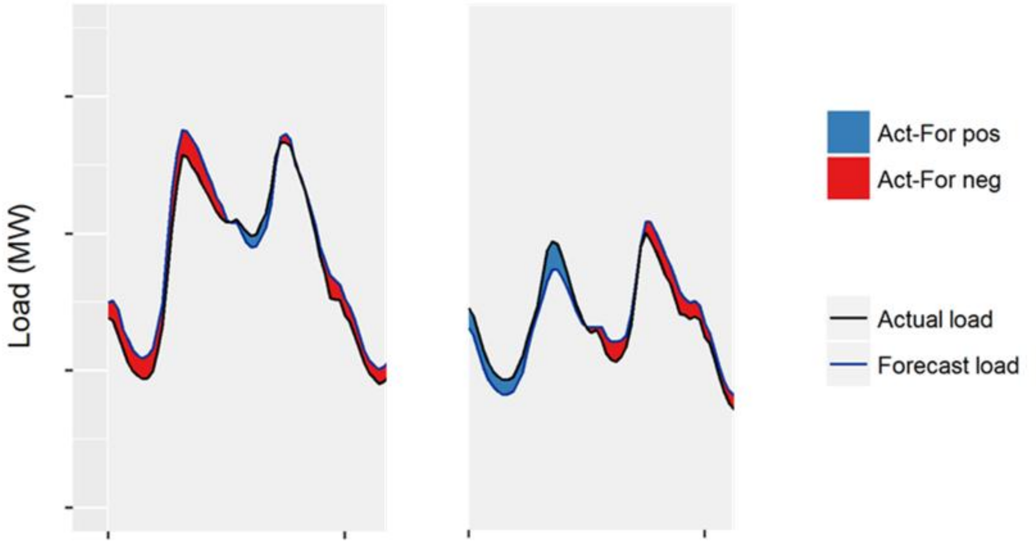
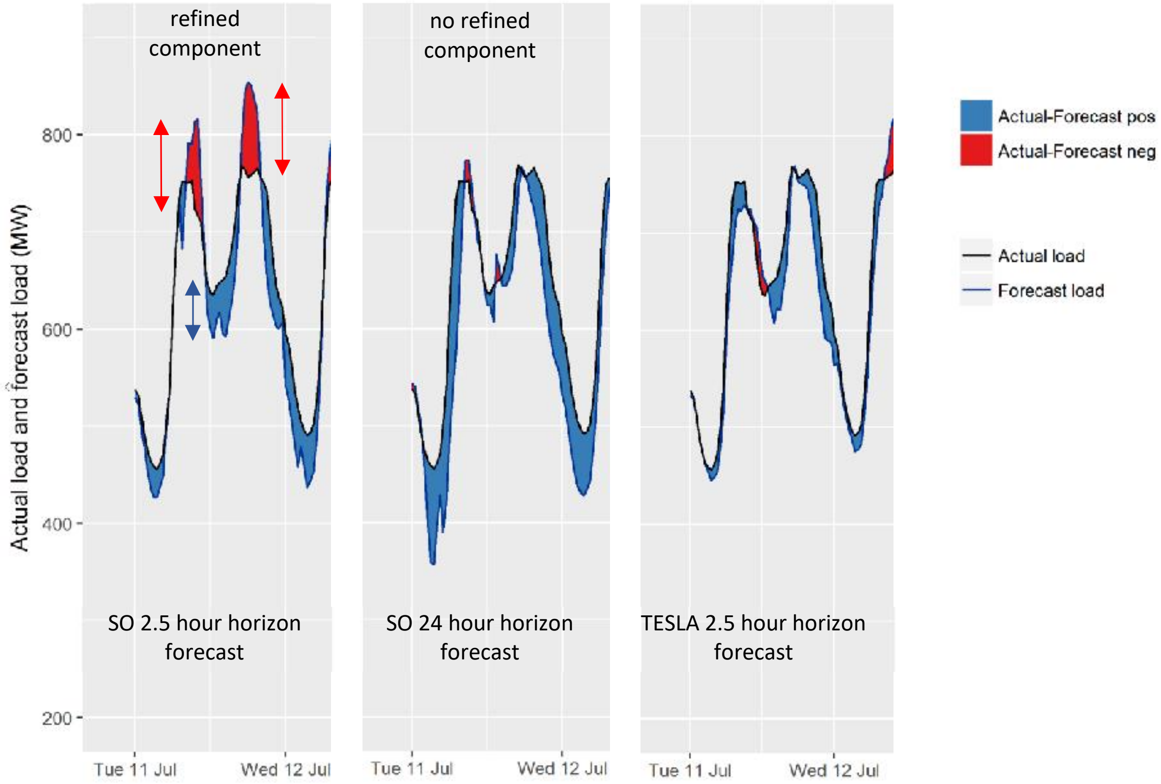


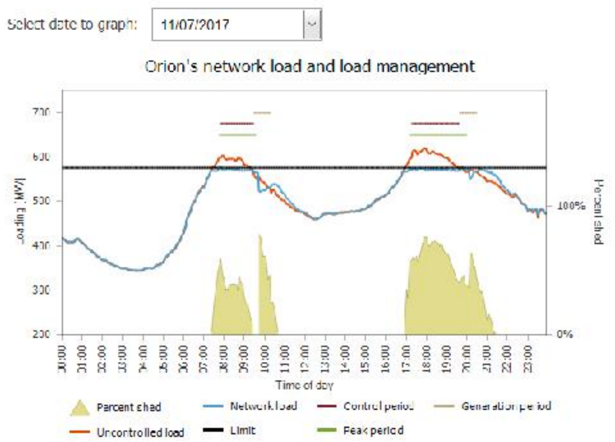
Figure 11 Examples of forecast inaccuracy with network load control. On the day on the left, the forecast is inaccurate because it has assumed a non-load-control load shape on a day with load control. On the day on the right, the forecast is inaccurate because it has assumed a load-control load shape on a day where there is no load control.

- The first day, on the left, had network load control that the forecast didn't predict because there was no load control during the period used for the data in the historical profile; this resulted in over-forecasting where load was reduced by the network company for the morning peak, and under-forecasting where load was shifted to the midday trough. Prior intervention to normalise the data would ensure the model provided forecasts like these, even if the historic profile used days that had load control, as the controlled load would be added back on when the data was normalised. An intervener could then use the network company's decision criteria to estimate the amount of load control according to the amount of over-forecast, and then manually shift that load into the midday trough. The intervention would also need to normalise the data to add it back the load control.
- The second day, on the right, had no network load control but the forecast predicted there was because there had been during the period used to create the historical profile; this resulted in under-forecasting where load in the historical profile had been reduced by the network company in the morning peak, and over-forecasting where load in the historical profile was shifted to the midday trough. Prior normalisation of the data would prevent this.

The more human effort applied to correct for load control, the greater the improvements.

It would be difficult to fully accommodate load control in our current system if we did not know which GXPs were relevant, or what control decision criteria were being used. Intervention to change the model frequently should help, to change the way history is used to create the profile, based on how representative different days in history are expected to be. A library of shapes based on days that had a similar temperature at the peak could also help. However, estimation would still be prone to error, particularly as the model is under-specified (not accounting for important variables). New inputs could help somewhat with this.





Orion graph taken from https://online.oriongroup.co.nz/LoadManagement/default.aspx

Figure 12 Comparison of forecasts for CH area on 11 July 2017, when load was controlled by Orion.The SO LF refined component (in the 2.5 hour forecast, but not the 24 hour forecast) scales based on current actual load and expected load shape, and so makes the LF errors worse.

Other human intervention

There are other opportunities for human intervention beyond load control and school holidays. Some of these are detailed with explanations below.

- If soil moisture inputs were unattainable, adjustments could be made after a rain storm to account for increased soil moisture, hence less irrigation load.
- Adjustments could be made to account for people taking leave on Mondays preceding a Tuesday public holiday, or Friday's following a Thursday public holiday. Adjustments could be made for known behaviour unique to certain days, such as ANZAC day when many shops that are shut in the morning re-open in the afternoon. These adjustments will reduce outliers.
- The refined component could be turned off in certain situations where the shape of the days load is expected to differ substantially from the profile predicted by historic data. The refined component scales the rest of the day's forecast based on observed differences between the forecast and recent actuals. This works best if the shape is accurately predicted, i.e. a difference between the forecast and actuals now is likely to mean a similar difference later in the day. This is often not the case. Network load control, for example, could mean an under-forecast in the middle of the day when load is brought on following morning load shed, or in preparation for evening load shed, and an over-forecast in the evening when load is shed (Figure 12). The refined component will increase the evening forecast to adjust for the higher actuals in the middle of the day, further increasing the magnitude of over-forecast in the evening.
- Intervention to use a different forecast method within the current tool could be utilized when sudden weather changes occur. This would reduce outliers, particularly in the longer forecast horizons, when the refined component is not acting. The model currently relies heavily on recent history, so if a day isn't easily explainable by recent history, the forecast error can be high.

Option 3 (SOW Option 3): Refresh current model + better data inputs

This option assumes quick wins have been utilised and the model has been re-calibrated. Better inputs include new types of inputs and more frequent receipt of current weather forecast inputs.

New types of inputs and more frequent receipt of current weather forecasts are both likely to improve forecast accuracy, in terms of average accuracy and prevalence of outliers.

New Inputs

Weather Inputs

The current forecast uses temperature as its sole weather variable for most areas, with wind speed additionally used in just three areas. Typically, other load forecast providers use a range of additional variables, including cloud cover, dew point, sunset time, soil moisture, rainfall, wind direction, etc. Temperature, wind speed, wind direction, cloud cover, and solar radiation were used in the 2012 trial.

Additional weather variables provide additional load information. Soil moisture determines whether irrigation is required, rainfall would count as an irrigation substitute, sunset time would account for the peak that occurs when people turn their lights on, cloud cover will also impact when people turn their lights on, humidity increases transfer of heat or cold to a body and makes air conditioning work harder, wind direction along with speed affects wind chill factor (which hastens the transition of a body to a change in ambient temperature), etc.

It is not known whether soil moisture information can be attained, but the other weather variables mentioned should be easily attainable, though at a cost.

It is likely significant improvement can be made by adding new weather inputs. Some time and effort would be required to accommodate the inputs appropriately. The inputs would need to be combined in meaningful ways and their relationships to load would need to be defined. This would require some research and studies.

For instance, irrigation load depends on soil moisture, but rainfall may be a substitute. It is unlikely that the relationship between soil moisture and the need to irrigate is linear; there may be a threshold soil moisture level before a given person decides to irrigate, and this threshold might change per person. There may also be a rainfall threshold before rainfall is thought of as a substitute. If soil moisture is particularly low people may decide to irrigate even if it rains.

Other Inputs

New, non-weather, inputs would also be beneficial, for example price, and school holiday information.

Significant improvement is expected from an appropriate account of school holidays. However, the best account for school holidays within the current tool would require some human intervention in addition to new input variables. Furthermore, the current tool is limited in its ability to account for discrete seasonal effects such as school holidays, even if input data is available.

School holiday inputs can be added, such as an indicator variable that tells the system whether or not a day is a school holiday. This will add or subtract a constant amount, determined by the model fit, to the load determined from the historic profile and weather information. Interaction variables between the school holiday indicator and the weather variables would allow for a different relationship between weather and load to apply to school holidays. More energy might be consumed on a rainy day, for instance, as opposed to a sunny day when school children are more likely to play outside.

Forecast prices from the most recent forward schedules could be used as an input to allow some account for price response, as discussed in section 6.1.1.

Controlled load, due to network load control, could be added as an input if we had historic data of controlled load and could forecast load control.

Quantitative evidence

Trial participants used better inputs and had better accuracy compared to the current forecast at the time. All trial participants had more inputs than the current forecast, including wind direction, solar radiation, rainfall, and school holiday information. Four out of five participants usually performed better than the current forecast, at least for horizons of 6 hours or more, with the current forecast particularly inaccurate at longer horizons. It is expected the improved input data were a factor in their forecast accuracy; this was the only factor known to be common to all participants except the current forecast.

TESLA may have had better inputs than other participants and had greater accuracy. Trial participants were not excluded from using additional weather information to that provided for the trial. We expect TESLA may have been the only participant to use additional information, which could suggest improved weather information is critical to forecasting, given how much more accurate they were than other participants, particularly in the longer horizons when weather forecasting is more difficult. However, it is more likely that the major reason for their greater accuracy was their effort and ability to account for the unique characteristics at each conforming GXP, especially non-conforming types of load and embedded generation.

Simple regression runs suggested additional weather variables could improve accuracy. We ran simple regression models of 2017 Auckland conforming GXP load against several weather variables to get an idea of usefulness of additional weather information. New weather inputs included humidity, rainfall, wind speed (not used in Auckland in our current model, though the data is available) and solar radiation. A rough approximation of the current forecast was used as a base model prior to adding the new inputs; this included load for the same period the day before, load from the same period the week before, trading period indicators, temperature, and squared temperature (as used in the current model). The MAPE of the model including all additional inputs along with interaction terms was approximately 0.1% less than the approximation of the current model, 4.2% compared to 4.3%. All additional variables used inputs in their raw form however, implying linear relationships to load; greater improvement would be expected if inputs were transformed to create variables with more appropriate relationships to load, as there is no reason to expect these relationships should be linear. Greater improvements may also be expected in other regions where some of the additional inputs are more relevant, for example in the South Island rainfall may help explain irrigation load.

More frequent receipt of weather forecast

More frequent receipt of weather forecasts is likely to increase accuracy, on average and in outliers.

This is likely a major contributor to overall weather forecast accuracy, which can improve load forecasts as explained above in the assessment of rapid deployment adjustments.

Infrequent receipt of weather data may be causing load forecast inaccuracies when weather changes substantially from recent history. System coordinators occasionally adjust the load forecast in the morning of days when weather changes substantially from recent history. They will often remove their override later in the day when the forecast has become more accurate; the forecast may have become more accurate due to new weather data coming in, or it may just be the effect of the refined component.

Option 3a (variation of SOW Option 3): Refresh current model + better data inputs + some intervention

This is a combination of options 2a and 3, but in terms of accuracy is greater than the sum of its parts.

Some of the newer inputs are more applicable to different seasons, for example irrigation season, therefore regularly changing the model will allow for greater improvements with more inputs.

Any intervention that requires estimation, for example load control, is less likely to be in error with a more fully specified model (i.e. one with better data inputs).

3B: NEW SYSTEM REQUIRED

We assume a system with a new model will need to be at least as accurate, secure, reliable, and future-proof as our current tool. A new tool could achieve additional accuracy, on average and with fewer outliers, beyond any improvement of the current tool. Cost effectiveness could be maintained in the future better than with the current tool.

New models enable GXP level forecasts and price response to be incorporated in the market system, as well as response to events such as DR calls, and all the other drivers discussed above in human intervention.

New models require the cost of better inputs (as with options 3 and 3a) and at least low operational costs.

Base level

The achievable accuracy of a new tool would depend on the system, with greater cost for greater accuracy. Any system would require at least 0.5 FTEs in-house.

Accuracy

A new system could achieve better accuracy than any of the current tool improvements.

A new system can achieve greater accuracy compared to any of the current tool improvements by

- utilising a more sophisticated model
- utilising an automatically, continuously updating model
- utilising a suite of models
- utilising better human intervention
    - better tools and interfaces to enable effective intervention
    - interveners who understand the models, to whom it is not a 'black box'
    - awareness improved by comparing several models
- providing bottom-up forecasting

This allows the system to account better for some of the issues identified in the outcome strategy map. Explanations are detailed below.

- Changing load behaviour over time - seasons

    A new tool can account better for discrete seasons. The following can be improved:

    - Identification of seasons using greater expertise, better tools and user interfaces, data science and learning techniques.
    - Account of seasonal effects; the current day's load profile can be better separated from other times, complex weather relationships and interactions are more likely to pick up different kinds of weather effects attributable to different seasons, learning models and data science techniques can be utilised, and effects can differ by GXP.

- Changing load behaviour over time – day/weeks

    A new system can better account for frequently changing load behaviour as models can automatically and continuously update, people can more easily intervene, data science techniques can find relationships where humans could not easily do so (particularly in the future when load may change more frequently but more data may be available). Changing patterns specific to GXPs can be more easily identified when freed from the noise of the aggregate.

- Temporal load shifts

  A new system can better account for temporal shifts of load; a total daily energy value can be forecast prior to distribution throughout the day, learning algorithms can pick up patterns better, and historic data can be more easily normalised for past load shifts.

- Relationships between weather variables and each other, weather variables and load

  Human expertise can better understand and model relationships between variables. A more sophisticated model, such as a neural network model, can discover important relationships that may be comparatively hidden from human eyes.

- Different scenarios, e.g. extreme weather

  A new system can better account for different scenarios (including extreme weather) by using the most applicable model for each.

  A new tool may have greater ability to adapt to, or choose the optimal model for, different scenarios. Neural network models, for instance, perform best when weather is in a relatively normal range. These could be used within this range, with other models used at other times. A model that can predict temporal load shifts could be used during the network load control season, but may not perform as well as another model in other seasons. Having a suite of models allows for reduced complexity, greater effective sample sizes, and less chance of over-fitting. It enables greater situational awareness to adapt to scenarios, due to the reduced complexity and ability to compare a range of forecasts.

  A new system can better account for extreme weather scenarios. Simple weather variables are unlikely to fit a relationship to load that is particularly good at the extremes; the lower sample size of extreme values means the fit will be more accurate in the normal range instead. The days profile is also atypical during extreme weather. A suite of models and more sophisticated models is likely to enable better account of extreme weather, for example load lost due to lines tripping may be detectable and appropriately accounted for. Several weather forecasts feeding into separate load forecasts would enable better assessment of weather effects and sensitivities when there are fronts.

- Human intervention error.

  The tools and interfaces available to the intervener and the intervener’s understanding of the models could minimise human error. A tool that is highly adaptable to an expert intervener reduces the likelihood of errors and makes errors easier to identify. For instance, an adjusted forecast could be tested, with comparisons to forecasts from other and previous models easily visualised.

Accuracy is likely to be improved during stressed market conditions. More sophisticated models and greater expertise will allow better account for stressed market conditions. Sophisticated models can better account for the relationships under stressed conditions, but the small history of data would make it difficult for any statistical model. Investigation and planning by expert staff is likely to make a big difference, with appropriate models then applied. Expertise in understanding the load as well as in statistical techniques would be highly valuable, especially if the tools are highly adaptable to enable modelling of unique situations. Accuracy will therefore differ between options depending on the level of expertise.

Cost in the near-term

The cost of the system differs per option.

Cost in the future

Cost effectiveness in the future differs depending on closed vs open options.

Price response incorporated in the market system

It is expected price response incorporated in the market system will mainly improve the accuracy during stressed market conditions. Accuracy on average will only be marginally improved, with a small reduction in size of outliers. The cost of incorporating price response in the market system will be reasonably significant.

There is also a risk that the system will mis-estimate price response, as increases in demand do not always lead to higher non-response prices, due to non-linearities related to reserve sharing optimisation, introduced with the National Market for Instantaneous Reserves. This would need to be carefully considered when the system was designed.

Price responses can be better accounted for, as price elasticity (the variation attributable to prices) can be provided to the market system from the load forecast model, potentially on an ongoing basis, with the degree of price elasticity continually learnt and updated.

GXP level forecasts incorporated in the market system

Accuracy increases with improved forecasts at GXP level, particularly for constraints and at times of market stress. During times of market stress prices are particularly high, as is the difference between loss tranches. A small proportionate error inaccuracy can therefore have a large magnitude impact on price.

GXP level forecasts enable greater price forecast accuracy when transmission constraints bind, such as with spring-washer effects.

Forecast price accuracy will increase if the market system is updated to accommodate GXP level forecasts. This is due to more accurate losses forecast by SPD and better signaling of spring-washer effects. The potential contribution of more accurate forecast losses to forecast prices is not negligible compared to the contributions of load forecast improvements herein considered. Using GXP level forecasts translates to a small price effect in general due to losses, and a large effect at times by reducing spring-washer effects, in addition to the effect of increased accuracy of total New Zealand load.

A new system is likely to cost substantially less without GXP level forecasts incorporated in the market system.

If GXP level forecasting is required in the near-term, the current tool is an inadequate backup, therefore greater security for the TESLA forecast will then be required. New interfaces will also need to be built with the market system to provide GXP level forecasts. If GXP level forecasting is not required in the near term, these costs can be deferred to the future.

Open Systems

An adaptable system will maintain accuracy in Future State A and continue to be cost effective in Future State B.

In comparison to a closed system, a system which is relatively open would be cheaper but with lower accuracy.

Option 4a (variation on the first of the SOW Option 4 alternatives): New system, open + some intervention

This is similar to the general comments above.

Option 4b (variation on the first of the SOW Option 4 alternatives): New system, open + high intervention

Increased intervention would enable greater accuracy, particularly in terms of prevalence of outliers and during stressed market conditions.

Accuracy could be improved using several in-house FTEs but probably couldn't reach the accuracy of the closed system. There would be a risk of maintaining expertise when staff move on. Further FTEs would be required if this risk were to be mitigated. The IST infrastructure could be slightly cheaper, the vendor fee could be substantially less.

Closed Models

Closed models are more accurate than open models but likely incur greater costs. Future accuracy and costs differ between closed options, as discussed in the following sections.

A closed model is expected to be more accurate than an open model. There were a number of comparatively open participants in the trial. Differences between TESLA and the other participants therefore provides some indication of accuracy improvements of closed vs open. These were likely attributable to different levels of

- Model sophistication, particularly account of network load control and non-conforming-type loads at conforming GXPs
- Human intervention, both in quantity and attempt to understand the uniqueness of different load locations
- Account for differences between GXPs
- Quality of weather data

Costs differ between closed-IP options although both require the same annual vendor fee.

There are also some accuracy differences between the different closed options, as discussed in the following sections.

Option 4c (variation on the second of the SOW Option 4 alternatives): New system, via new model, closed (TESLA plug-in with current model as backup)

This option is only appropriate for base level forecasts.

Accuracy cannot be maintained in the future when GXP level forecasting needs to be incorporated in the market system.

Future cost will include infrastructure to bring the tool in house, incorporate GXP level forecasting and price response in the market system.

Option 4d (variation on the second of the SOW Option 4 alternatives): New system, closed (e.g. TESLA in-house)

This option is appropriate for base forecasts or for forecasts incorporating GXP level forecasts or price response in the market system.

Accuracy is greater than Option 4c, even at base level, as models can be adjusted to incorporate feedback from the market system for GXP ties, outages, and demand response calls.

This option can make better use of GXP level forecasts incorporated in the market system, demonstrated by TESLA's particular accuracy at the GXP level in the trial.

Cost is greater than the above option as infrastructure to bring the tool in house is done initially, rather than in the future. If GXP level forecasting or price response are incorporated in the market system now, they will also incur capital costs that will then not be needed in the future.

This option can meet the required accuracy in the future by incorporating GXP level forecasting and/or price response in the market system if they were not initially incorporated.

APPENDIX 4 GLOSSARY

AEMO	Australian Electricity Market Operator
CE	Contingent Event. An event deemed likely enough to occur that instantaneous reserve is procured to maintain frequency above 48 Hz
Code	Electricity Industry Participation Code 2010 is a set of rules that govern New Zealand's electricity industry
Conforming GXP	A Grid Exit Point where the load is less than 250 MW per year and follows a "typical" two peak daily profile of morning peak and evening peak demand profile with periods of lower load during the rest of the day and at night
DD	Dispatchable Demand, an initiative to allow purchasers to bid price-responsive load bids into the spot market to provide certainty on the price they paid for their load
Demand	Load or embedded generation at a conforming GXP
DSBF	Demand-Side Bidding and Forecasting, an initiative to allow purchasers at conforming nodes to enter a different demand quantity to see the price effect of a different load level at that GXP
DSI	Demand Side Initiative. Demand reductions initiated by Transpower to reduce the requirement for transmission maintenance and investment
dp	Decimal places
ems	Energy Market Services, Transpower's commercial market group
Frequency Keeping	An ancillary service procured to control of system frequency to within normal band (50 ± 0.2 Hz)
Gentailers	Electricity companies that are vertically integrated generator - retailers
GXP	Grid Exit Point. A point of connection on the grid at which electricity predominantly flows out of the grid
High spring-washer effect	This occurs when there is a transmission constraint on the system means lower priced generation is unable to be used to satisfy load in the constrained area and therefore higher priced generation is required to meet this load instead. It results in a separation of prices

IL	Interruptible Load. A form of instantaneous reserve by which load is switched off automatically if frequency falls below 49.2 Hz
IR	Instantaneous Reserve. An ancillary service consisting of backup generation or interruptible load procured and dispatched in order to mitigate the risk of a credible under-frequency event
Island (electrical)	A separate power system consisting of generation and load, disconnected from the rest of the grid
IST	Information Services & Technology
MAPE	Mean absolute percentage error. A measure of prediction accuracy of a forecasting method in statistics $$\left(\frac{1}{n}\sum\frac{\lvert Actual - Forecast\rvert}{\lvert Actual\rvert}\right) * 100$$
MetService	Meteorological Service of New Zealand
MISO	The system operator for the Midcontinent of the United States of America
MS	Market System
MTLF	Medium Term Load Forecast
MV90	Revenue meter data mainly sourced from Transpower's ION meter.
NRS	Non-Response Schedule – takes two forms, a short form of the current trading period and the following 7 trading periods, and a longer form of the current trading period and the following 71 trading periods. These forecasts can be seen by market participants ahead of the trading period and provide an indication of price per trading period. As they do not include the same inputs as the final pricing schedule their results will vary from the final price. One of the key variances is that the NRS includes a forecast load instead of the metered load at a GXP
OFR	Over frequency Reserve. An ancillary service procured to arrest a rapid increase in frequency from a sudden drop in load
PJM	The system operator for a large part of the North West of the United States of America
PLSR	Partly-Loaded Spinning Reserve. Type of synchronous instantaneous reserve
PRS	Price Responsive Schedule – takes two forms, a short form of the current trading period and the following 7 trading periods, and a longer form of the current trading period and the following 71 trading

	periods. These forecasts can be seen by market participants ahead of the trading period and provide an indication of price per trading period. As they do not include the same inputs as the final pricing schedule their results will vary from the final price. One of the key variances is that the PRS includes a forecast load instead of the metered load at a GXP. One factor where the PRS differs from the NRS is that prices are included in the bids by purchasers at non-conforming nodes
Trading period	A period of 30 minutes ending on each hour or 30 minutes past each hour on any trading day
RTP	Real Time Pricing
SCADA data	Operational data collected from Supervisory control and data acquisition systems
SPD	Scheduling, Pricing and Dispatch model. The Linear Programming model that is used to optimise the variables in the market
TESLA	TESLA – Energy Industry Forecasting Solutions
Voltage support	An ancillary service procured by the injection or absorption of reactive power in order to maintain system voltage