Forecast Accuracy Report (FAR) Methodology

August 2020

Final Determination

1. Overview

AEMO develops a number of demand and supply forecasts for both operational purposes (short-term forecasts) and reliability/planning purposes (long-term forecasts) for the National Electricity Market (NEM).

This document focuses on longer-term electricity demand and supply forecasts and the methodology used by AEMO to assess the accuracy of these.

Within AEMO, these forecasts are used in a number of reliability and planning processes, including:

- Medium Term Projected Assessment of System Adequacy (MT PASA).
- The Electricity Statement of Opportunities (ESOO), and the associated reliability forecast used for the Retailer Reliability Obligation (RRO).
- Integrated System Plan (ISP).

Demand forecasts are also used by industry participants and governments for their own work.

To ensure the insights and advice derived from the forecasts are as accurate as can be expected, AEMO uses a continuous improvement process which includes the assessment of forecast accuracy, determining causes of forecasts deviating from actuals/observed values and identifying and implementing improvements to enhance the forecasts in future years.

The introduction of the reliability forecast[1] under the RRO rules in 2019 increased the importance of the forecast accuracy, and accordingly AEMO has increased the depth and breadth of the forecast accuracy reporting[2]. Similarly, AEMO formalised the reporting of the improvements to be made into a Forecast Improvement Plan that also includes a consultation step as outlined in AEMO's Interim Reliability Forecast Guidelines[3].

The Interim Reliability Forecast Guidelines outline methodology documents which explain various processes required to produce the Reliability Forecast. These methodology documents must be consulted on at least every four years using the Australian Energy Regulator's (AER's) Forecasting Best Practice Guidelines consultation procedure, to determine:

- The fundamental methodologies needed in the forecasting processes.
- The components on which the forecasts are to be based, and the way they are to be determined and used.
- The stakeholder engagement process for determining the forecasting methodologies, inputs, and assumptions.

Consultation on these methodologies is being staggered across the four years in recognition of the time commitment required by stakeholders to provide considered responses to these technical documents.

The FAR Methodology consultation is one of the methodology consultations required from AEMO under the Forecasting Best Practice Guidelines.

[1] Since 2019, AEMO has been required to develop and publish a Reliability Forecast as part of its ESOO, in accordance with the Retailer Reliability Obligation. For more information, see https://www.aer.gov.au/retail-markets/retailer-reliability-obligation.

[2] Previous years' forecast accuracy reports are at https://www.aemo.com.au/energy-systems/electricity/national-electricitymarket-nem/nem-forecasting-and-planning/forecasting-and-reliability/forecasting-accuracy-reporting.

[3] See https://aemo.com.au/en/consultations/current-and-closed-consultations/interim-reliability-forecast-guidelines.

1.1 The Forecast Accuracy Report

The FAR reports on the long-term demand and supply forecasts AEMO develops for the National Electricity Market (NEM), including Medium Term Projected Assessment of System Adequacy (MT PASA), the Electricity Statement of Opportunities (ESOO) and the associated reliability forecast[1] used for the Retailer Reliability Obligation (RRO), and the Integrated System Plan (ISP). Demand forecasts are also used by industry participants and governments for their own work.

To ensure the insights and advice derived from the forecasts are as accurate as can be expected, AEMO uses a continuous improvement process which includes the assessment of forecast accuracy, determining causes of forecasts deviating from actuals/observed values and identifying and implementing improvements to enhance the forecasts in future years. The FAR Methodology document describes the process by which AEMO assesses the forecast accuracy of the above forecasts.

1.2 Consultation process

This consultation was conducted in accordance with the AER's Interim Forecasting Best Practice Guidelines.

On 29 April 2020, AEMO initiated the first stage of the consultation with the publication of its draft FAR Methodology document and consultation notice[4]. Through this consultation, AEMO sought feedback on the draft FAR Methodology document to inform any changes to be applied in 2020 and beyond.

AEMO received feedback from the Forecasting Reference Group forum and one-on-one discussions, and an extensive submission from ERM Power, available on AEMO's consultation webpage. AEMO would like to thank ERM Power and others who participated in this process.

AEMO's approach in responding to submissions was to consider them with respect to the principles underlying AEMO's forecasting, which are outlined in Section 1.3.

On July 14th, 2020, AEMO published the draft determination[4] outlining AEMO's responses to issues raised by stakeholders. There were no further submissions following the draft determination, so AEMO now publishes the final FAR Methodology document without further changes. The results of this consultation are detailed in Section 2 of this document:

The publication of this Final Determination marks the conclusion of the consultation.

1.3 Principles applied in considering this feedback

In considering how to take this feedback into account, AEMO has applied the following principles that align with the AER's forecasting best practice guidelines[5]:

- Forecasts should be accurate, unbiased, and based on comprehensive information.
- Transparency is important to provide stakeholders with confidence in the forecasts.

In particular, in assessing the merit of any proposed changes to the methodology, AEMO has considered whether:

- The change will materially improve the accuracy of the reliability forecasts.
- The change will materially improve transparency in the way AEMO assesses its forecasting performance, and provide confidence that performance is being assessed against appropriate metrics.
- The expected benefits outweigh the implementation costs borne by AEMO and/or industry participants.

---

[4] Available at https://aemo.com.au/consultations/current-and-closed-consultations/forecast-accuracy-report-methodology.

[5] See https://www.aer.gov.au/retail-markets/retail-guidelines-reviews/retailer-reliability-obligation-interim-forecasting-best-practice-guideline.

2. Discussion of material issues raised

This section discusses the material issues raised by stakeholders in submissions to stage one of the consultation along with AEMO's considerations and conclusions. It is unchanged from the draft determination, and as there were no submissions received on the draft, the headings below now constitute AEMO's final determination.

2.1 Adjusting demand

The historical time series for operational sent-out demand (referred to in the methodology as metered demand) is useful for modelling actual demand, provided the following are taken into account: unusual activities such as AEMO directions, Reliability and Emergency Reserve Trader (RERT) activation, load shedding, distribution network outages, DSP, and voluntary load reduction. Adjusting demand refers to the process of performing those changes to allow operational sent-out energy to be used as a data source.

2.1.1 Adjustments for price-responsive Demand Side Participation

Issue summary and submissions

ERM Power suggested AEMO's adjustment methodology for price-responsive DSP is not suitable for highly variable load types below a size threshold, and recommended that a minimum load size and acceptable baseline accuracy threshold be defined.

Assessment and conclusion

AEMO's view is that it is rarely appropriate to utilise a size or accuracy threshold for determining inclusion within a forecast; it is often the case that including an imperfectly forecast component is better than missing out on the component altogether.

AEMO is currently focusing on larger sites to avoid the issue described by ERM Power, and will consider the addition of smaller, more varying loads if good estimates can be found for individual days through methodology improvements made to AEMO's DSP forecasting work.

AEMO will continue the current adjustment methodology for price-responsive DSP and have made it clear in the methodology document that the adjustment is only made for DSP resources where good estimates can be determined with confidence.

2.1.2 Network reliability programs estimated demand side participation

Issue summary and submissions

ERM Power submitted that the data verification process for NSPs' DSP is unclear, and recommended that the FAR methodology set out a verification process to be applied by NSPs and AEMO for the provision of estimated DSP.

Assessment and conclusion

AEMO have amended the methodology to describe the basic first pass review – calculating anticipated DSP performance with respect to the scale of the relevant load. Immaterial and infeasible values will be addressed appropriately.

Beyond that, AEMO notes committing to a detailed verification methodology is unlikely to serve the evolving nature of DSP solutions, and will continue the current practice of case by case manual review by AEMO's experts, with resource allocation in proportion to the expected contribution to DSP.

2.1.3 Distribution network outages

Issue summary and submissions

ERM Power submitted:

> *The methodology notes that "AEMO seeks to get an estimate of customers without power for the relevant period from the relevant NSP, or an estimate of the impact in MW directly if available." However, the methodology fails to set out the process by which the NSP is required to compile and verify their estimates.*

ERM Power further recommended the methodology sets out a verification process for provision of this data.

Assessment and conclusion

AEMO disagrees with the recommendation, on the basis that the benefits are unlikely to outweigh the implementation costs. The load profile created by adding the estimated energy lost due to the outage back onto operational sent-out energy is inspected for feasibility, and immaterial changes are ignored. This review process is appropriate given the overall scale of the distribution network outages.

2.1.4 Voluntary load reductions – general

Issue summary and submissions

ERM Power questioned the inclusion of pool pumps, washing machines, clothes dryers and dishwashers in the Demand Reduction Calculator (DRC) tool as appliances that would normally be expected to consume energy during the peak demand periods.

Assessment and conclusion

AEMO's understanding of appliance use is informed by the Residential Baseline Study (RBS)[6], a periodically updated analysis produced by the Department of Industry and Science on behalf of the Trans-Tasman Equipment Energy Efficiency (E3) Program. AEMO anticipates an update to the RBS in 2021. AEMO does not assume that all pool pumps, washing machines, clothes dryers, and dishwashers are consuming at peak time, but that one or more of these common appliances could be used at peak times and therefore are included in considerations for voluntary load reduction potential. The methodology document has been clarified.

2.1.5 Voluntary load reductions – air-conditioning

Issue summary and submissions

ERM Power submitted that:

- Current electricity costs have already impacted users' air-conditioning usage behaviour, and behavioural change may be continuing to evolve.
- Consumers' practice of raising the air-conditioning set point may not change power usage if their device is already operating on full power.

ERM Power also asked whether AEMO assumes air-conditioners are running in unoccupied residences.

[6] See https://www.energyrating.gov.au/document/report-residential-baseline-study-australia-2000-2030.

Assessment and conclusion

AEMO agrees with ERM Power that consumers' use of air-conditioners, including set temperature, continues to change over time. Such changes, along with the potential use of air-conditioners in unoccupied premises, would require substantial resources to measure and track over time.

Rather than utilising an expensive bottom-up data or assumption driven methodology, AEMO's estimates of voluntary reductions in air-conditioning load are based on the high level approach outlined in the methodology. Given the scale of the demand reduction under discussion, AEMO views this high level approach as suitable, that is, the expected benefits to the forecast outweigh the implementation costs.

2.2 Demand forecasts

2.2.1 Calculating minimum and maximum demand

Issue summary and submissions

ERM Power recommended that:

> *AEMO extract from their modelling outputs the 1, 10, 50, 90 and 99% POE (Probability Of Exceedance) forecasts on at least a monthly basis and compare these to the actual monthly outcomes for both the day of working weekday maximum demand, cross-checked with the maximum daily demand outcome on the day of maximum or minimum working weekday daily maximum temperature outcomes. This would allow the accuracy of the model to be considered over a range of potential future scenarios, as opposed to the current summer and winter single day accuracy measurements.*

Assessment and conclusion

AEMO acknowledges the practical challenges in accurately assessing a probabilistic demand forecast where the actual figure of interest occurs at an annual frequency. This is supported by the University of Adelaide's statistical assessment[7]:

> *It is challenging to retrospectively assess the accuracy of a probabilistic forecast against a single point observation. Communicating forecast accuracy across a range of stakeholders provides a further challenge, given the highly technical nature of probabilistic forecast assessment.*

AEMO uses a multi-model ensemble to generate the annual/seasonal forecasts, and does not produce a monthly equivalent model. To support insights into monthly demand, AEMO has previously published actual monthly demand with a monthly maximum demand forecast derived from the demand traces. For the sake of transparency, AEMO agrees to continue publishing these values.

However, it is critical that readers of this information note it cannot be assumed that monthly demand values have the same distribution as annual demand values. The University of Adelaide's expert report states:

> *[the use of monthly datapoints to assess annual demand forecasts] may violate the distributional assumptions necessary to assess forecast accuracy. First, weekly or monthly forecasts in the same location are unlikely to be independent. For example, a heatwave at the end of one month may continue to the start of the subsequent month; more broadly, months within the same season will be subject to the same model conditions (e.g., El Niño or La Niña), giving the impression of systematic bias in forecasts.*

In summary, monthly demand forecasts and their respective actuals are unreliable for assessing annual demand forecasts. The first point in the above quote implies that, given the correlation between months for such factors as weather, comparing monthly demand forecasts and their actuals is only statistically valid on a month by month basis. For example, a large set of historical February monthly demand forecasts is needed to correctly assess February monthly demand forecast performance; it is inappropriate to supplement any shortage of February monthly demand data with adjacent months.

[7] See https://www.aemo.com.au/-/media/files/electricity/nem/planning_and_forecasting/accuracy-report/forecastmetricsassessment_uoa-aemo.pdf.

The report continues:

Second, the minimum/maximum demand over shorter time intervals (weeks or months) might not have the same distribution as the seasonal minimum/maximum demand. Inappropriate use of aggregated data could bias estimates of accuracy for the actual quantities of business need (i.e., seasonal minimum/maximum demand).

This implies that even if demand across adjacent months were uncorrelated, one should not expect a 10% POE monthly demand forecast to be exceeded once a year on average, but rather expect the 10% POE monthly demand forecast for a particular month to be exceeded in that month once every 10 years.

Furthermore, even if it was valid to assume the monthly distributions were uncorrelated and representative of the annual demand distribution, 12 datapoints is insufficient to meet the statistical 'law of large numbers' required for a meaningful or fair assessment of AEMO's probabilistic annual demand forecast. In other words, 12 datapoints per year suffers from the same challenges as a single annual data point, but is more likely to instil a false sense of confidence in the ability to assess probabilistic annual demand forecasting performance.

For the avoidance of doubt, the reasons above mean the following erroneous hypothetical statements are not statistically supported:

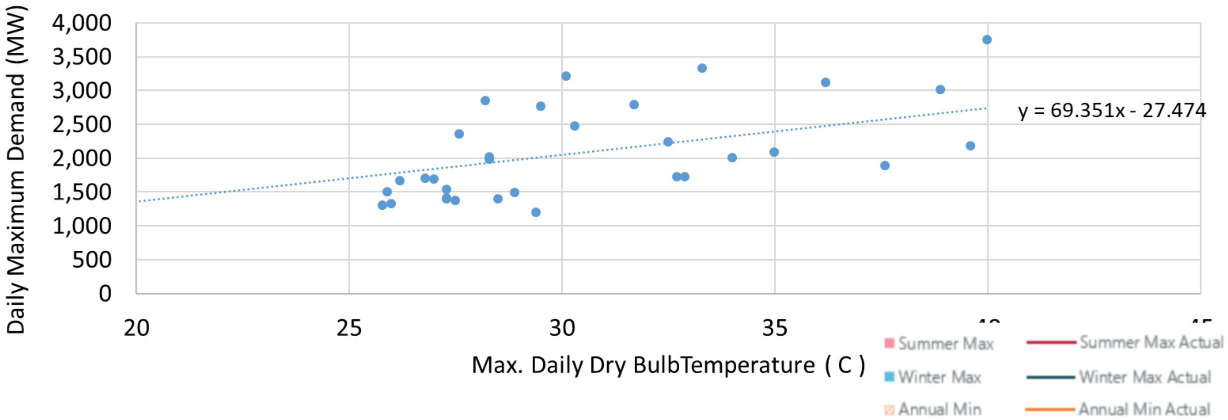
These four-in-a-row monthly actual figures are well above the monthly 50% POE forecasts, proving that AEMO's annual demand forecast is biased.

In the last twelve months, there's been no monthly actuals above AEMO's 10% POE monthly forecast, therefore AEMO is over-forecasting annual demand.

In the last twelve months, there's been three monthly actuals above AEMO's 10% POE monthly forecast, therefore AEMO is under-forecasting annual demand.

The suggestion of 'cross-checking' demand forecasts with temperature unfortunately suffers related statistical limitations. The single annual demand figure (or even a monthly demand figure) is not enough to meet the statistical law of large numbers and avoid erroneous conclusions being drawn from the stochastic nature of consumer temperature response.

Figure 1 Estimated maximum daily residential demand vs maximum daily temperature, New South Wales



The above figure (taken from Figure 5 in the FAR Methodology report) shows the large range of possible maximum demands at any given temperature, illustrating that cross checking demand against temperature is not a reliable means of assessing maximum demand forecasts.

AEMO acknowledges ERM Power's concern about the risks of inaccurate forecasts. In reference to the performance of the 2019 Queensland forecast, AEMO used a variety of metrics to determine that the forecast was too low and required upward revision. In the 2019 Forecast Accuracy Report[8], AEMO described how inaccuracy in input forecasts, energy consumption forecasts, summer maximum, and monthly maximum forecasts collectively built an evidence base for forecast revision. The upward revision was not simply determined due to an observed exceedance of the 10% POE threshold. In the case of South Australia, a similar exceedance was observed that was not associated with the same commentary as there was no evidence to suggest an upward revision was warranted.

In general, the extreme nature of 1% and 99% POE forecasts makes them inherently challenging to forecast accurately. Ideally, any desired level of accuracy can be achieved by running a very large number of simulations in order to capture enough instances of extreme outcomes and accurately reflect their probability. In practice, costs associated with computer simulation time limits the number of simulations and hence the accuracy of extreme POE forecasts. Thus, the FAR's assessment of probabilistic forecast accuracy is focused towards more central measures (10%, 50%, and 90%), which are not only easier to forecast, but also (relatively) easier to assess.

In conclusion, to support insights into monthly demand, AEMO will continue publishing monthly actual demand along with monthly maximum demand forecast derived from the forecast demand traces. However, for the reasons explained above, AEMO will not use this data to assess demand forecast accuracy. AEMO will continue to use the methods referred to in the University of Adelaide's expert report, as set out in the methodology document.

2.3 Supply forecasts

2.3.1 Demand side participation (DSP) demand forecasts

Issue summary and submissions

ERM Power expressed concern regarding the:

> *low values used by AEMO in their supply forecasts compared to the level of DSP observed by participants at times of high prices in the market.*

ERM Power further disagreed with the decision criteria as it:

> *fails to acknowledge that DSP not subject to an awarded RERT contract remains free to participate for in-market demand side response and assumes that awarding of RERT contracts will be explicitly required in the future. This in effect introduces a circular outcome that by excluding this price responsive load from the reliability assessment, a resulting reliability gap is artificially manufactured which then determines ongoing RERT procurement is required.*

Assessment and conclusion

In response to stakeholder submissions on both the FAR Methodology and DSP Methodology consultations, AEMO has reassessed options for including RERT panel responses to the extent they can be reliably assessed and validated based on evidence. This matter has been addressed in the DSP Methodology Consultation.

2.3.2 Generating resources supply availability

Issue summary and submissions

ERM Power recommended that actual to forecast comparison data be presented on a regional basis separately for both scheduled and semi-scheduled generating resources.

---

[8] At https://www.aemo.com.au/-/media/files/electricity/nem/planning_and_forecasting/accuracy-report/forecast_accuracy_report_2019.pdf?la=en&hash=DCD762A3035664F4F4F53430FABB0846.

Assessment and conclusion

AEMO agrees with ERM Power on the benefits in reporting on the availability of supply side resources at a regional level, separating scheduled generation from variable renewable energy (VRE). This analysis will provide a generalised view of the accuracy of supply resourced during hot days.

AEMO will perform the regional analysis, and where insightful, will continue to report on generator performance at a technology level.

2.3.3 Generating resources supply availability PASA availability

Issue summary and submissions

ERM Power noted the diversity of demand outcomes that can occur across the top 10 summer days, and consequently recommended AEMO substitute "PASA availability" for "maximum availability" in supply availability for gas-fired generation.

In the market rules, *PASA availability* is defined as "the physical plant capability (taking ambient weather conditions) … that can be made available during that period, on 24 hours' notice" whereas *maximum availability* is the "total MW capacity available for dispatch". For forecast assessment purposes, the key difference is that PASA availability could include capability that may be physically available but where a participant has chosen not to offer this into the market for dispatch.

Assessment and conclusion

AEMO agrees that PASA availability is often a good metric, particularly in the example cited by ERM, and has in the past used "PASA availability" rather than "maximum availability" for the forecast accuracy analysis.

However, PASA availability can be an unreliable metric. For example, when a unit trips and reduces maximum availability to zero, the plant operator may neglect to adjust PASA availability. Thus, PASA availability is subject to data errors and is therefore inappropriate for the forecast accuracy assessment.

AEMO will therefore continue using maximum availability as a more consistent metric.

2.3.4 Generating resources supply availability outlier treatment

Issue summary and submissions

ERM Power noted its understanding that

> *AEMO truncates the forecast supply side availability curves to "eliminate outliers that may occur with very low probabilities"… in displaying actual outcomes, AEMO continues to include these very low probability outcomes which in ERM Power's view results in a degree of assessment bias.*

ERM Power recommend the assessment and graphical representation include the full range of potential outcomes from AEMO supply availability assessment modelling.

Assessment and conclusion

Given enough simulations, the full range of simulated supply availability extends from zero to full capacity, providing little insight. As evidence of this, prior to the 2019 Forecast Accuracy Review, AEMO presented the supply availability curves with the full range of availability that occurred across the simulations. Stakeholders at the time noted that showing the full range was not a good representation of the simulated outcomes, because the minimum availability observed across all simulations is, by definition, an outlier event and would be a highly volatile measure.

AEMO agreed with this feedback and reverted to a more standard statistical approach, which is to show the equivalent of a 95% confidence interval. If the historical outcomes fall outside this range, it indicates that the input assumptions used (being capacities and outage rates) could be incorrect, but AEMO also understands there is also a chance that the historical performance itself could be an outlier event. AEMO remains of the view that the 95% approach is superior to an approach that shows the full range of simulated outcomes.

A1. Additional issues raised

Table 1 Minor changes

Organisation(s)	Comment	AEMO response
ERM Power	Amend example table 1 in the Forecast Accuracy Report methodology to show that <20,000 impact in half-hour interval has a zero-adjustment value.	AEMO has made this clarification.
ERM Power	Provide support/derivation of 1 customer = 2kW	AEMO has made this clarification.