ENERGY CONSUMERS AUSTRALIA

A Suite 2, Level 14, 1 Castlereagh Street
Sydney NSW 2000
T 02 9220 5500
W energyconsumersaustralia.com.au
@energyvoiceau
/energyconsumersaustralia
/energyconsumersaustralia

ABN 96 603 931 326

19 June 2019

Australian Energy Regulator

By email: aerinquiry@aer.gov.au

Ref 64872 - Forecasting for Retailer Reliability Obligation

Energy Consumers Australia appreciates the opportunity to comment on the Australian Energy Regulator's (AER) *Draft Interim Forecasting Best Practice Guideline: Retailer Reliability Obligation* (the Paper) of May 2019.

Energy Consumers Australia is the national voice for residential and small business energy consumers. Established by the Council of Australian Governments Energy Council in 2015, our objective is to promote the long-term interests of energy consumers with respect to price, quality, reliability, safety and security of supply.

The Paper is part of the process for implementing the Retailer Reliability Obligation (RRO). This is an important market-based reform, and we welcome the opportunity to participate in the design of the detailed implementation arrangements. As mentioned, we commissioned Finncorn to provide advice on aspects of this work, which we attach for your consideration. It is important to note that this work does not represent an Energy Consumers Australia position, but is used to identify areas for further engagement and thinking.

We make suggestions about how to strengthen the Draft Interim Forecasting Best Practice Guideline (the Guideline) under three broad headings in this submission: transprency; consultation; and ensuring good forecasts. In doing so, we also discuss the financial implications of the Australian Energy Market Operator's (AEMO) forecasts – both now and into the future.

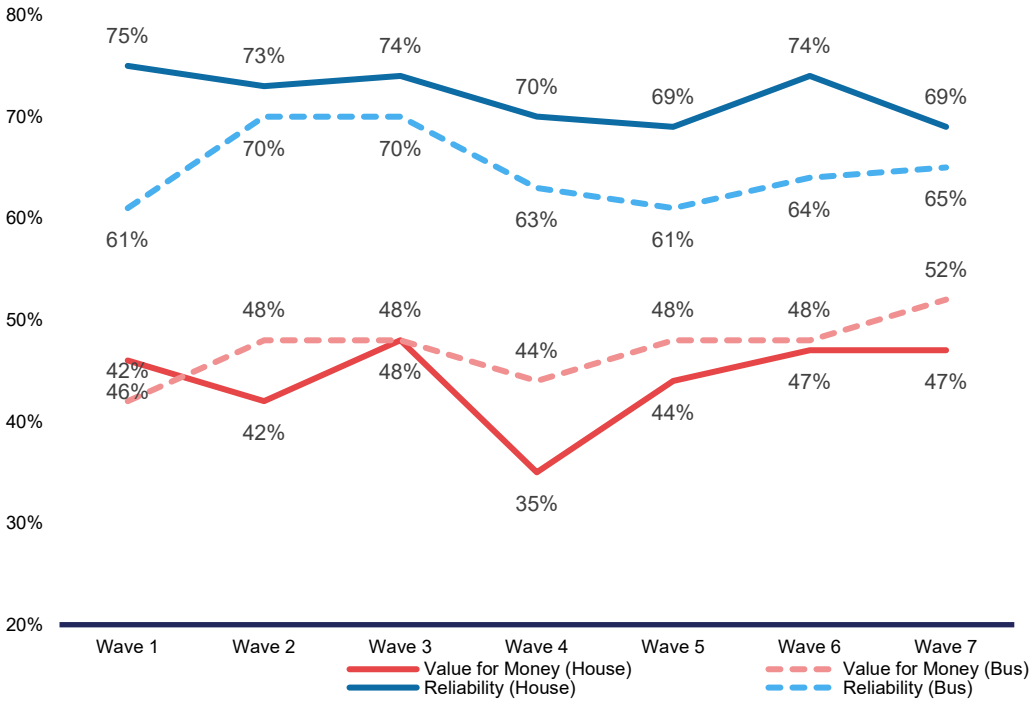
Consumer priorities

A key framing for this work is what consumers are telling us about energy services and priorities. The Energy Consumer Sentiment Survey indicates that affordability is their main priority. Figure 1 shows the response to questions about value for money and reliability over the seven waves of this research, for both households and businesses:[1]

- Value for Money
  "How would you rate the overall value for money of the products and services provided by your electricity company in the last 6 months?" (% 7 out of 10 or higher)"
- Reliability
  "Thinking about the reliability of your electricity supply, how satisfied are you with the number of times you've had loss of power, blackouts or other faults with your electricity supply in the past 6 months?" (% 7 out of 10 or higher)"

[1] https://energyconsumersaustralia.com.au/wp-content/uploads/Energy-Consumer-Sentiment-Survey-Report-June-2019.pdf

Figure 1: Consumer satisfaction with price and reliability

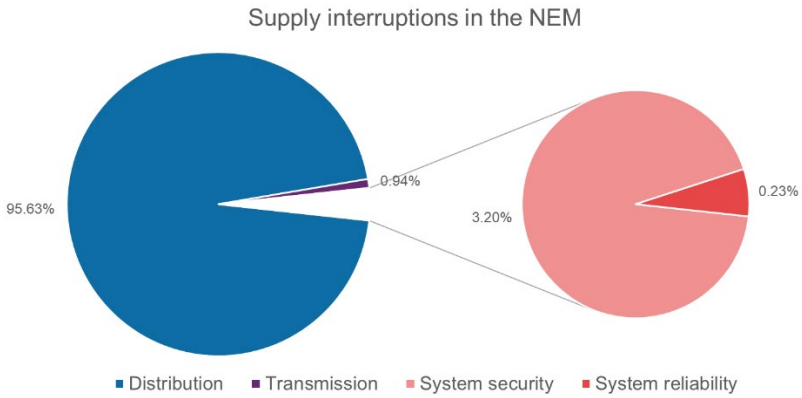


These survey responses demonstrate that consumers have significantly higher satisfaction with the reliability of electricity services than they do with price. The survey has also consistently shown that consumers rate the 'value for money' for electricity below that for all the other services (water, mobile phone, insurance and banking) included in the survey. We do not interpret this data as suggesting that consumers necessarily want a trade-off of lower prices for less reliability. We do, however, believe that it provides strong evidence for being cautious in instituting mechanisms to increase (or maintain) reliability, given the cost implications for consumers.

The need to ensure the RRO is appropriately calibrated is further reinforced by the observation that less than 4% of electricity supply interruptions are caused by system security or system reliability factors (see Figure 2).[2]

[2] Australian Energy Market Commission (AEMC) identified in its *Reliability Frameworks Review: Final Report 2018*

Figure 2: Sources of supply interruptions in the NEM: 2007-08 to 2016-17
(Source: AEMC Reliability Frameworks Review 2018



We believe this is significantly important in the context of the consultation on the RRO. In simple terms, the RRO operates by identifying periods where demand is likely to exceed supply and then requiring retailers to contract enough supply to cover the forecast demand. The initial consequence is too much demand for contracts chasing too little supply, and hence higher prices. The intention is that the demand for contracts will then underwrite the provision of additional supply.

Consequently, it is critical that the forecasting that will inform the decision about whether to trigger the RRO, is as robust as possible and not unduly conservative, given the cost implications for consumers. The Finncorn paper (Attachment A) identifies risks associated with overly conservative RRO settings and identifies ways to mitigate these risks.

In our submission to AEMO in response to their consultation on RRO forecasting, we supported two key recommendations; consistency of forecasting methodology for the T-3 and T-1 Reliability Gap; and the treatment of 'in progress' generation projects. The latter is a forecasting issue which is also relevant to the development of the Guideline.

New entrant generation

The Paper describes AEMO's proposed (new) approach to factoring-in new generation in its forecasts for the RRO as:

> *Projects that have commenced construction but do not yet meet all five of the criteria [in Table 3 of the Paper] are also published, because they have clearly made a formal commitment to construct (referred to as 'Com*' projects in AEMO's Generation Information page).*
>
> *In the 2018 ESOO, all new entrant generation and storage projects that were classified as committed, or had commenced construction, were included in the forecast. However, even though construction may have commenced, Com* may be less certain to proceed, particularly if connection approvals have not been provided. For the 2019 ESOO and future reliability forecasts, AEMO is proposing not to include Com* projects.*

AEMO considers that this more appropriately takes into consideration the cost of misclassification. At T-3, if more capacity is assumed committed than actually proceeds, there is a risk that a reliability instrument may not be requested, and it is not possible in subsequent years to reverse that decision based on new information. If, on the other hand, more capacity is built than assumed at T-3, this would serve to close any reliability gap and may result in no T-1 reliability instrument needing to be issued.

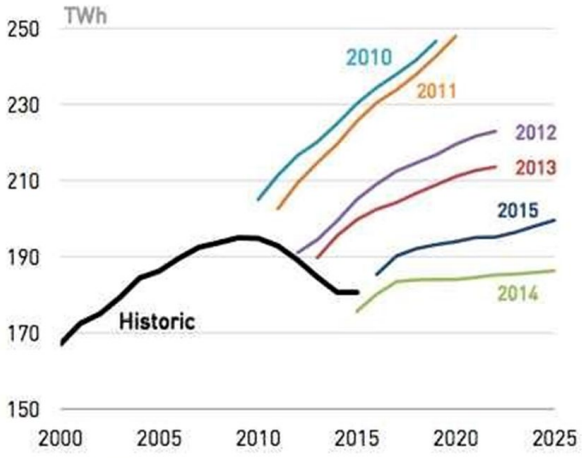
As observed above, there are costs associated with being conservative at T-3. While we accept that not all projects with AEMO's 'Com*' will come online at their scheduled dates, it is equally not true that none of them will. Good forecasting practice would be to use history as a guide to what the right 'risk-weighted' proportion of new generation should be (that is a proportion between the 100% previously used and the 0% now being proposed). The proportional likelihood can either be included through processes of weighted likelihood through multiple iterations or, more simply, by applying the weighting to the rated capacity of Com* generators. As noted in the Finncorn paper it is also appropriate to include a forecast 'lag' for average project delays.

The financial implications of AEMO forecasting

Market participants and market bodies all rely on AEMO's short term and long-term forecasting. Daily decisions about market participation are informed by AEMO forecasts, while decisions by AEMO itself on whether it needs to intervene (for example to order a generator on) are also based on forecasts.

The load-shedding that occurred in South Australia in February 2017 was due to AEMO underestimating demand in its forecasts that in turn was based on a poor choice of weather forecast data provider.[3]

Figure 3: AEMO electricity consumption forecasts and outcomes 2010-2015



A consequence of errors with longer term forecasting is reflected in the outcome colloquially known as the 'hedgehog' diagram (see Figure 3). From 2010 on AEMO forecast a continuation of what had been a steadily increasing total consumption despite a downturn in demand apparent from 2008.

[3] https://www.aemo.com.au/-/media/Files/Electricity/NEM/Market_Notices_and_Events/Power_System_Incident_Reports/2017/System-Event-Report-South-Australia-8-February-2017.pdf

Hugh Saddler has estimated that more than half this reduction was due to major energy efficiency measures that AEMO could have observed and included in the forecast.[4] The extent of the savings due to efficiency were demonstrated for AEMO in a report by Pitt&Sherry in 2016.[5]

This forecasting error is one of the factors that resulted in over-investment in distribution assets.[6]

We note in addition that the Energy Security Board (ESB) in its *Converting the Integrated System Plan into Action Consultation Paper* of May 2019 has identified a potential role for the Guideline (AER) in making the Integrated System Plan (ISP) actionable. Given this proposition, and the significance of AEMO forecasting generally, we think it is appropriate that the AER uses a wider context than just the precription of the RRO rules that requires the development of the Guideline.

For these reasons we think that the AER Guideline should be prepared to not only meet the requirements of the RRO rules (which requires that the AER make the Guideline) but as a tool that could be applied more widely.

The Draft Interim Forecasting Best Practice Guideline

We admit to some confusion as to what in the Paper constitutes the Guideline. This is the heading given to section 3 of the Paper but this section also includes reasoning in support of the positions described in the Paper.

The Guideline otherwise is simply modelled on the three limbs identified in the draft Rules as needing to be included in the Guideline. These can be paraphrased as the forecasts need to be good, transparent and consulted on. These should not in practice be treated equally – the value of transparency and consultation is as much about making sure the forecasts are good as it is about building trust.

As best we understand it a core part of the Guideline splits the consulting process into two parts. In the first part AEMO would once every four years use the 'Forecasting Best Practice Consultation Procedures' to determine fundamental methodologies to be used in forecasting, the components of the forecast and process for determining them, and the stakeholder engagement process.

We do not agree with the proposition that methodologies and components should be unchanged for four year periods. We have more comments on consultation below.

We discuss the three elements in the order in which we think they shjould appear in the guideline – transparency, consultation and ensuring forecasts are good.

Transparency

To state that the basic inputs, assumptions and methodology that underpin the forecasts be disclosed is a weak form of expressing transparency. It isn't sufficient to suggest that they 'should be disclosed' – they must be disclosed.

---

[4] http://www.australian-options.org.au/wp-content/uploads/2015/08/5-Hugh-Saddler-electricity.pdf and https://theconversation.com/why-is-electricity-consumption-decreasing-in-australia-20998

[5] http://www.aemo.com.au/Electricity/National-Electricity-Market-NEM/Planning-and-forecasting/-/media/0667A56BE96141F7B0314794E33CA12F.ashx

[6] https://www.accc.gov.au/system/files/Retail+Electricity+Pricing+Inquiry%E2%80%94Final+Report+June+2018_0.pdf Page 159

It is not clear what the AER's intent is in relaiton to transparency. Indeed , we think the Paper is conflating 'disclosure' (in our terms transparency) with stakeholder engagement. In Section 3.2 'Forecasting methodologies, inputs, and assumptions to be disclosed' the Paper actually describes a consiultation process.

- In the section on 'Methodologies' the Paper says "The AER considers it best practice for AEMO to follow the process set out in section 3 in determining and detailing the approach to each of these streams."
- In the section on 'Assumptions, scenarios and sensitivities' the Paper says "The AER considers best practice development of the scenarios and sensitivities, where appropriate, and the supporting narratives, will follow the approach set out in section 3."
- In the section 'Inputs' the Paper says "To facilitate stakeholders' understanding of the fundamental drivers of energy industry development, the AER considers it best practice for AEMO to use the process set out in section 3 to determine the components that will feed into the models used to determine the demand forecasts and the supply forecasts on which *reliability forecasts* will be produced."[7]

We have a different view about the requirements of a guideline on disclosure and transparency. The first is that the availability of the informatioon needs to be clearly advertised and promoted. Transparency must go beyond informing the members of a forecasting reference group.

The second is that the information needs to be accessible. That means it needs to be available in easy to access, downloadable machine readable form. The best proactive principle is that if data comes from a data base tools should be available for users to choose what part of the datset they wish to extract. [8]

The third is that it should be comprehendable. Where industry specific terms are used these need to be included in a glossary and links to the glossary provided in the location from which the data is accessed.

Finally the guideline needs to provide clarity about the attribution of 'confidentiality' to data. The path of least resistance for a body such as AEMO to acquire data is by agreeing to treat it as confidential. The guideline needs to be clear about the circumstances under which it should accept claims of confidentiality.

In accordance with process that bodies such as the AER and the Australian Competition and Consumer Commission (ACCC) use where confidential information is used the Guideline must specify the requirement for a regime that allows confidential information to be accessed by appointed experts who exercise appropriate confidentiality deeds. That is, stakeholders need the ability to appoint an expert who can look at the data and provide advice to their client about whether the client should challenge the use of the data.

[7] The references here to 'section 3' are quite confusing because they appear in section 3. From the context it is clear the Paper is merely referring to the stem of section 3 (for want of a better term section 3.0) this throws us back to the Consultation Forecasting Best Practice Procedures.

[8] As an example, AEMO's historic aggregate price and demand files are available for one time period for one region at a time. Given the output is a simple CSV file the parameter choice should permit the selection of multiple time periods and multiple regions.

Consultation

The Guideline has introduced Forecasting Best Practice Consultation Procedures in Appendix A that are "modelled on the *Rules consultation procedures* in clause 8.9 of the NER". It is not clear to us that these procedures are the right model for the task.

The suggested procedure is based on what we describe as the 'hub and spoke' model of consultation wherein all stakeholders are positioned as being in a bilateral conversation with the central agency (in this case AEMO). Best practice consultation seeks to benefit from collective knowledge and skills, improvements are developed through the interaction of parties who have a stake in the outcome but possibly different, even conflicting, principles and priorities.

The first step is that in forecasting AEMO should be actively developing a community of interest with whom AEMO invests in a joint development of skills and experience in forecasting. In developing this community AEMO must avoide filtering participation on the basis of AEMO's perception of skill (sometimes described as people who are 'qualified representatives'[9]) and instead put the emphasis on stakeholders who understand the impact of the forecast.

Workshops, seminars, small group meetings with these stakeholders are an effective way of consultation. Importantly such meetings shouldn't be conducted as a means to enable formal submissions to be written, they should be documented and that documentation used itself as an input. We note that AEMO has used a number of such forums very effectively already in the ISP and RRO processes.

We consider that meetings and seminars can profitably be the first stage of a process before issues papers, proposed methodologies and initial inputs are prepared. The first meeting in each cycle should consider the effectiveness of the last forecasting round (that can be assessed against outcomes.

The second point is that AEMO should be engaging at every step of the process. This might only be by publishing an outcome or result and inviting comment. The important part is being open to comment, inviting comment and considering comment. It should be relatively less important to make a requirement to be able to demonstrate how comments have been addressed (as outlined in Part (e) of the proposed procedure). Stakeholders have a greater interest in AEMO's resources being spent in considering content rather than detailing how they did so.

Finally, it is important that consultation is iterative. As we will discuss the process of building a forecast isn't just a linerar process from inputs, scenarios and sensitivies and methodologies to outcomes. The outcomes of initial modelling can impact on the choices made on inputs, scenarios and sensitivities.

Ensuring forecasts are good

We have chosen to use the word 'good' to cover the range of variables that are inherent in the assessment of forecasting.

Forecasting is the process of describing future states of nature with respect to a number of variables of interest. They take many forms. A single point estimate is usually an attempt to measure the expected value (statistical mean) of a variable. More complex forecasts attempt to estimate the distribution of values across the whole universe of possible states of nature.

[9] For example, this phrase was used in AEMO's description of its proposed Consumer Engagement Panel in the *2019 Planning and Forecasting Consultation Paper: Scenarios, Inputs, Assumptions, Methodology, Timeline, and Consultation Process* of February 2019

Consequently it is very hard to ever describe a forecast as 'accurate' – the outturn state of nature will almost certainly be one that sits within the distribution, but that alone tells you nothing. Only through a repeated application that generates a distribution of outturns that can be compared to the distributions forecast can be used to assess if the estimation was good

The Guideline should therefore not be focussing on the *post hoc* evaluation of the forecasts but should instead be providing advice on how the forecast process can be conducted in a way that best meets the requirements of the forecast. This is where the Guideline should be providing advice on the choice of input sources, methodologies, scenarios and sensitivities.

Input sources should be accurate and up to date. For example using reference years for weather that don't reflect the increase in temerature anomolies is a bad choice of inputs. Expert assistance should also be sought on the input choices (Bureau of Meterology for weather, Reserve Bank of Australia or other sources for Gross Domestic Product growth, Australian Bureau of Statistics for population growth, experts who can model energy efficiency programs for per premise consumption).

Forecasting methodologies should be complete and consistent. Most importantly methodologies should attempt to make as many variables as possible internal to the model. For example, nothing in the current ISP model estimates the impact on price of the forecasts, yet wholesale price will affect both the timing of generator exit and the level of consumption demand (both through its impact on Districuted Energy Resources (DER) and energy efficiency investment). As a second example the current ISP uses a forecast of DER uptake as an input to the model rather than something that could be affected by future policies driven by the conclusions of the model.

We have some concern that the commercial package Plexos is a bit of a 'black box' and we are unclear on how much ongoing external validation the computational elements of the model are subject to.

On completeness we do not believe that modelling to 2040 is sufficient when the change in generating fleet will need to continue to 2050. We therefore support AEMO's decision to include 'step change' scenario in its current ISP process[10].

Scenarios are devices used to model vastly different trajectories. For example, energy transition studies by Geels et al use two parthways (plus a neutral pathway) to model energy transition in the United Kingdom, one based on technological substitution (enacted by incumbent actors), and one based on broader system transformation (enacted by new entrants).[11] Scenarios are therefore complete models using different pathways. These are typically constructed as extreme polar opposites to provide boundary conditions for the distribution.

The number of scenarios, however, needs to be determined by the correlation between different variables inherent in the possible pathways. We think that in forecasting for the ISP (and not so much the RRO) there are sufficient uncorrelated variables that two or four scenarios might be insufficient to map the boundaries of the set of future states. Examples include the role of hydrogen in transport and heating, whether storage and generation is centralised or decentralised, the extent to which emissions reduction objectives are achieved continuosly or through sudden policy changes.

---

[10] https://reneweconomy.com.au/aemo-to-model-step-change-in-energy-transition-and-major-emission-cuts-85604/

[11] Geels, FW, McMeekin, A & Pfluger, B 2018, 'Socio-technical scenarios as a methodological tool to explore social and political feasibility in low-carbon transitions: Bridging computer models and the multi-level perspective in UK electricity generation (2010–2050)', *Technological Forecasting and Social Change*.

In general sensitivities are used as small incremental changes to inputs from which distribution of outcomes can be determined. The extreme of this is a full Monte Carlo simulation where repeated runs with inputs drawn from distributions estimated for those inputs, from which a distribution of outputs can be presented.

However, sensitivities play an additional role when forecasting complex systems. The well known 'butterfly effect' describes the sensitivity that can apply to model outcomes for model inputs in certain boundary conditions.[12] A more significant example of its impact is the way that weather forecasting is now conducted. The forecasters have a model of a complex system into which they input actual weather observations. To test the forecasts they re-run them using small variations in the input parameters (that is, deviating from what was actually measured) and are able to get a measure of the 'stability' of the forecast.[13]

In the case of energy market forecasts, sensitivity analysis of this kind identifies where the model is sensitive to particular assumptions. This can then justify applying additional resources to the analysis of the method of developing the input. As an example in our comments of the RRO Forecast we identified a concern with the relatively arbitrary approach to deciding what weighting to apply to the likelihood of new generation being available – AEMO having decided to go from 100% to 0% and our suggestion that the percentage should be modelled. If this input is found to have a very significant impact more resources might be put in to firming that assessment. Alternatively it might be found that the choice between 100% and 0% isn't at all material.

These are the kinds of issues the Guideline should explore.. In practical terms it means that after transparency is applied to inputs and methodologies and these have been consulted on, AEMO needs to demonstrate how the choices on inputs, methodologies, scenarios and sensitivities provide the best unbiased estimates of the distribution of possible future states of nature.

Conclusion

The requirement for the AER to develop a Forecasting Best Practice Guideline is a substantial part of the RRO Rules designed to ensure that the RRO is aligned with the long-term interests of energy consumers. The ESB's proposal to use the Guideline as part of making the ISP actionable extends the potential significance of the Guideline to decisions that will have even greater implications for consumers than the RRO. In this submission we have suggested ways to strengthen the Guideline consistent with this significant role.

Thank you for the opportunity to contribute to the development of the supporting arrangements for the RRO. Please do not hesitate to contact David Havyatt, Senior Economist, on 02 9220 5500 or david.havyatt@energyconsumersaustralia.com.au, if you would like to discuss this submission further.

Yours sincerely,

Lynne Gallagher
Acting CEO

[12] "Nearly 45 years ago, during the 139th meeting of the American Association for the Advancement of Science, Edward Lorenz posed a question: 'Does the flap of a butterfly's wings in Brazil set off a tornado in Texas?'" https://www.americanscientist.org/article/understanding-the-butterfly-effect

[13] See Lewis, M 2018, *The Fifth Risk*, WW Norton & Company.

Summary and recommendations

In this paper we briefly outline the process, the impact on market participants and consumers, particular areas of concern in relation to the forecasting of Reliability Gaps and develop suggestions to minimise negative cost impacts on consumers within an acceptable reliability outcome.

Our recommendations are:

- **Recommendation 1:** ESB and AEMO should (a) clarify what the ESOO USE forecast really represents (b) clarify what the RRO USE forecast trigger is intended to be and (c) adjust the RRO process to differ from the ESOO USE forecast if the two are not identical.
- **Recommendation 2:** AEMO's Forecasting Report should be supplemented with a lookback of their forecast accuracy on maximum demand over the life of their ESOO-style forecasting of the metric.
- **Recommendation 3:** AEMO should include "Com*" projects, using a risked basis of the capacity size and timing to account for uncertainty of delivery. These projects should be subject to a relatively high risk-weighting of the capacity being delivered within three years (80%?), with a reasonable delay assumed (6 months on top of project estimated COD?)
- **Recommendation 4:** AEMO should consider a similar pipeline-risking approach to projects which have met some of the five tests for being considered "Committed" (and thus included in the forecast) – particularly in relation to projects which have secured offtake arrangements and internal and external financing commitments.
- **Recommendation 5:** AEMO should simulate the P90 maximum demand case and include it explicitly, except in the circumstances that the P50 maximum demand case results in zero USE.
- **Recommendation 6:** AEMO's process should allow for a proportional buffer – for example, a 20% breach of the Reliability Standard (from 0.0020% to 0.0024%) before requesting an RI from the AER.
- **Recommendation 7:** AEMO should maintain the same process and threshold for LOLP in defining the RG at both T-3 and T-1.

Review of engagement and supporting material

As requested, we are assisting Energy Consumers Australia ("ECA") in engagement with the RRO design and implementation process. This involves reviewing material from both AER and AEMO in relation to the Retailer Reliability Obligation ("RRO") and participating in structured engagement. This has included:

- Attendance at two AER workshops to date (with Contracts and Firmness to follow on June 14th)
- Review of three of AER's four draft RRO Guidelines (on the **Market Liquidity Obligation**, **Reliability Instrument**, and **Forecasting Best Practice,** again with **Contracts and Firmness** to follow)
- Review of relevant AEMO material including:
    - The **Reliability Forecasting Methodology Issues Paper** (April 2019) and **Addendum** (May 2019)
    - The associated slide pack developed for their Reliability Forecasting Methodology workshop on May 9th (which we were not invited to attend);
    - The associated **Reliability Standard Implementation Guidelines**, draft 17th April 2019 (updated to account for the RRO-related aspects of reliability);
    - The **ESOO Methodology Document** April 2019 (for the 2019 ESOO)
    - The **Electricity Demand Forecasting Methodology Information Paper** April 2019 (for the 2019 ESOO); and

- The most recent **Forecast Accuracy Report 2018**.

- Public submissions – where we note some substantial concerns. For example, ERM Power's submission to AEMO on the Reliability Forecasting Methodology Issues Paper stated they believe ***"AEMO is intending to apply an excessively conservative approach to forecasting USE in the 2019 Electricity Statement of Opportunities"***.

Less frequent interventions would be better for consumers

It is our premise that the long-term interests of consumers would be better served by less-frequent occurrences of intervention under the Retailer Reliability Obligation ("RRO").

The interventions occur when AEMO requests (and is likely granted by the AER) a Reliability Instrument ("RI"), based on an ESOO forecast of a Reliability Gap ("RG") three years in the future. The RI will either:

- **have no effect** (in the case where all retailers collectively agree with AEMO that the reliability risk is present and so, needs to be managed in their businesses via contracting in the normal course); or
- **cause retailers to undertake additional contracting for compliance purposes**, in excess of their commercial judgement (when they judge that AEMO's forecast is overly conservative).

In the latter case, demand for contracts will rise.

To meet that demand, market participants will deliver a combination of:

1. **new sources of reliability** (such as additional demand response, batteries, pumped hydro, peaking generation); and
2. **additional contracting from existing sources** (such as higher levels of contracting from existing thermal plant units).

In the first case, total systems costs are increased by the new capacity. In the second case, additional risk is being taken by the suppliers of reliability contracts at the margin, where an unplanned outage of the supporting capacity could lead to large contractual losses.

The outcome is passed directly to consumers: the cost of such contracts entered into by retailers is the basis for their Cost of Goods Sold, essentially a hedged version of the wholesale pool price exposure.

The more contracting is compelled, the less relevant is the outcomes for the spot price – the assertions that the RRO would drive down spot prices are not really relevant. All that matters to consumers is the retailers' overall Cost of Goods Sold, driven by contracts prices, not spot prices.

In short, when AEMO proves to be too conservative in driving compliance activity under the RRO, consumer prices will be higher than otherwise.

AEMO's ESOO processes can tend to the conservative

Until the advent of the RRO, the ESOO was informational – in a sense, AEMO's ESOO was a means to "ring the alarm bells" to highlight the potential need for further capacity investment to meet the Reliability Standard.

Since it was market participants who were exposed to the risks and opportunities, they would use the information in the ESOO but form their own judgement as to whether AEMO's forecast was the appropriate basis for their investment decisions.

Clearly, the consequence for AEMO of being too relaxed in relation to future reliability were worse than being too conservative: it is much more embarrassing if the lights go out when you said they wouldn't, than if they stay on when you said they might not.

Furthermore, there was no real consequence for AEMO in being too conservative, because they are only one view – ultimately participants would face the consequences of their own decisions.

This changes when the ESOO forecast creates a specific compliance obligation via the RRO.

As a result, we should critically examine whether the ESOO process as it stands is an appropriately balanced method to compel participant activity, rather than suggest or recommend it.

"False Positive" risks to consider and potentially, challenge

In this section we present a few areas of concern to us, based on our review of AEMO and AER materials, and associated submissions.

Is the trigger for a RI the "expected" USE?

It is unclear whether the process AEMO undertakes for the ESOO really leads to an "expected" USE forecast, where "expected" is commonly defined and understood to mean a P50 (or 50% probability of exceedance) case in any given year.

Other terms such as "maximum expected USE" crop up (e.g. the introduction to the ESOO Methodology Document) – which could be something else entirely, such as the "USE expected in an extreme weather year" or some other far more conservative interpretation.

Whether or not the actual nature of AEMO's version of the term "expected USE" is acceptable from the point of view of the ESOO in general, or the purposes of the Reliability Panel and the Reliability Standard is outside our scope here.

The concern we have is that the RRO – based on our understanding from the policy design process – was only meant to be invoked in the case of very clear shortfalls in future capacity, implicitly due to a market failure (since in hindsight, the NEM has met the Reliability Standard in the past without the RRO).

We understand that was to mean forecast USE exceeding the Reliability Standard in the P50 case, or with a 1-in-2 year probability.

Recommendation 1: ESB and AEMO should (a) clarify what the ESOO USE forecast really represents (b) clarify what the RRO USE forecast trigger is intended to be and (c) adjust the RRO process to differ from the ESOO USE forecast if the two are not identical.

Have AEMO's forecasts proved to be skewed conservatively, or not?

In reviewing AEMO's 2018 Forecast Accuracy Report, we note there is no statistical back-testing of the forecast accuracy of AEMO's maximum demand forecasts over time – the ultimate top-down view of whether AEMO has been forecasting well or poorly (albeit, backwards-looking).

In our view, this would involve:

1. Gathering the set of historical maximum demand outcomes in each of the 5 NEM regions for a given year;
2. Comparing these outcomes with the immediately preceding AEMO ESOO forecasts for P90, P50 and P10 maximum demand in that region.

The first of these are shown clearly in the Report (e.g. Figure 10 for NSW, etc.), but there is no lookback to AEMO's prior forecasts, only those for (in this case) the preceding 2017 ESOO.

The data set may be a bit sparse, but over time we would expect to see the P90 forecasts exceeded in 90% of the years, P50 in half the years, P10 in 10% of the years. There would also be an opportunity to identify any apparent trend in forecast accuracy – getting better or worse over time?

Recommendation 2: AEMO's Forecasting Report should be supplemented with a lookback of their forecast accuracy on maximum demand over the life of their ESOO-style forecasting of the metric.

Should the generation project pipeline be better assessed?

We are particularly concerned by AEMO's very strict definition of when a generation project is "committed enough" to be included in the modelling of the gap, three years forward.

If there is a tight supply-demand balance looming (such that AEMO may feel obliged to call for an RI), then we would expect a substantial pipeline of capacity to be mobilising to address it three years out – exactly as we observe today. We know from recent history that solar PV, battery storage and gas reciprocating engine capacity projects can be conceived and delivered well within three years, as can the contracting of demand response.

At T-1 this may all be apparent, but by then the damage may be done in terms of driving up contract demand and prices, if at T-3 an RI has been granted by the AER based on AEMO taking an overly pessimistic view on the pipeline as it sees it.

A key concern is the total exclusion of projects which have commenced construction (defined as "**Com***") prior to finalising their grid connection and generator performance standard requirements. These outstanding matters could well lead to delay in the project's stated Commercial Operations Date ("COD"), but it seems very likely indeed that almost all such projects will have sorted this out within the three years.

AEMO applies a high degree of science and rigour in forecasting demand under many more complex variables than this – yet takes a very crude approach indeed to supply, which we believe is far too conservative in light of the consequences of false positives.

In particular, AEMO would be aware that tight supply-demand conditions in three years will be correlated with identified but uncommitted generation capacity projects advancing to completion – in many cases, within the forecast period and so, helping to address any threat to reliability.

Recommendation 3: AEMO should include Com* projects, using a risked basis of the capacity size and timing to account for uncertainty of delivery. These projects should be subject to a relatively high risk-weighting of the capacity being delivered within three years (80%?), with a reasonable delay assumed (6 months on top of project estimated COD?)

Recommendation 4: AEMO should consider a similar pipeline-risking approach to projects which have met some of the five tests for being considered "Committed" (and thus included in the forecast) – particularly in relation to projects which have secured offtake arrangements and internal and external financing commitments.

Do the weightings of P90, P50 and P10 USE cases make sense for an expected USE case?

As a specific subset of the general concern above, we note that as a final step in deriving the "expected USE" forecast, AEMO weight the USE from P10 maximum demand simulations to 30.44% and the USE from P50 maximum demand simulations to 69.56%.[1]

The mathematics and statistics leading up to this outcome is probably sound, but the final step is a simplification of the "right" answer, which would be weighting BOTH the P90 and P10 cases to 30.44% and the P50 case to 39.12%.

The simplification is historically driven by avoiding further computational processes, which does not sound like a critical constraint to us in 2019. AMO's explanation for the validity of the simplification is confusing:

- On one hand AEMO state "The P50 and P90 outcomes are very close and so their weightings can be aggregated".
- Further on, AEMO state they "assume the USE is zero in the P90 case". We expect this is very likely to be the case at the moment – it would be extremely improbable that a simulation would throw up enough forced outages at the same time to cause USE in a P90 max demand case. However this may not always be the case, especially if (e.g.) a major thermal closure was being included in the scenario, ahead of committed replacement capacity.

These two statements can only both be true if both the P50 and P90 cases are in fact, zero USE.

Perhaps the P50 USE case is also zero (or very close to it) today, in which case the issue is immaterial.

[1] Refer to Section 5.2.2 of AEMO's "ESOO Methodology Document", April 2019

However, to the extent that EITHER OR BOTH the P50 case and the P90 case are non-zero, which might well be a plausible circumstance should one more coal plant announce an exit, the methodology used will over-estimate expected USE.

Recommendation 5: AEMO should simulate the P90 maximum demand case and include it explicitly, except in the circumstances that the P50 maximum demand case results in zero USE.

Should there be a materiality threshold?

The concept of a "material reliability gap" appears to have been weakened – it is now defined as any forecast breach of the Reliability Standard, no matter how small.

Since the existence of an RG and the request by AEMO for a RI is a "cliff edge" decision which drives every retailer participant into a compliance task which may increase consumer costs, we do not think this is appropriate.

A better approach would be to reinstate a materiality buffer, to ensure that modest forecasting errors do not prompt costly action based on a false positive.

Recommendation 6: AEMO's process should allow for a proportional buffer – for example, a 20% breach of the Reliability Standard (from 0.0020% to 0.0024%) before requesting an RI from the AER.

Should the T-3 RG be structurally more conservative than the T-1 RG?

AEMO have proposed to apply a broader definition of the RG at the T-3 point at which they request the RI. This is because at the T-1 point, they may only narrow the extent of the RG, not widen it, based on the better information and forecast at that time.[2]

While this is sensible from their perspective in guarding against errors by allowing them to be corrected at T-1, we do not think it is appropriate – because the T-3 RG will drive immediate contracting activity, on a basis that is clearly skewed to be more conservative than the ultimate outcome.

Recommendation 7: AEMO should maintain the same process and threshold for LOLP in defining the RG at both T-3 and T-1.

[2] E.g. if the T-3 RG is Dec-Jan weekdays 4pm-8pm, the T-1 RG can only be a smaller subset of this (e.g. Jan weekdays 6pm-8pm). AEMO do this by applying a looser threshold (Likelihood of Loss Probability of 2% at T-3, rather than 5% at T-1) for including a trading interval in the RG.