RELIABILITY PANEL AEMC

REVIEW

Reliability Panel AEMC

Draft report

Review of the form of the reliability standard and administered price cap

18 April 2024

Inquiries

Reliability Panel
Australian Energy Market Commission
Level 15, 60 Castlereagh Street
Sydney NSW 2000

E aemc@aemc.gov.au
T (02) 8296 7800

Reference: REL0086

About the Reliability Panel

The Panel is a specialist body within the Australian Energy Market Commission (AEMC) and comprises industry and consumer representatives. It is responsible for monitoring, reviewing and reporting on reliability, security and safety on the national electricity system, and advising the AEMC in respect of such matters. The Panel's responsibilities are specified in section 38 of the National Electricity Law.

Acknowledgement of Country

The AEMC acknowledges and shows respect for the traditional custodians of the many different lands across Australia on which we all live and work. We pay respect to all Elders past and present and the continuing connection of Aboriginal and Torres Strait Islander peoples to Country. The AEMC office is located on the land traditionally owned by the Gadigal people of the Eora nation.

Copyright

Citation

To cite this document, please use the following:
AEMC, Review of the form of the reliability standard and administered price cap, Draft report, 18 April 2024

Reliability Panel members

Executive summary

1 The national electricity market (NEM) is undergoing a significant transition. The characteristics and behaviours of the new system need to be understood so that system reliability can be maintained at a level that consumers value.

2 During the 2022 Reliability Standard and Settings Review (RSSR), the Reliability Panel (Panel) found that the nature and characteristics of reliability risks may be changing towards the end of 2028. Therefore, it recommended a more detailed review to determine if the form of the reliability standard is fit-for-purpose for consumers and the power system as the NEM transitions.

3 The Panel commenced this Review of the form of the reliability standard and administered price cap (APC) (Review) in March 2023. The purpose is to:

- better understand the characteristics of unserved energy (USE) and the nature of the risks as the power system transitions
- determine if the current form adequately reflects the changing reliability risk profile, or whether alternatives need to be considered
- review whether the existing form of the APC remains fit-for-purpose.

4 In November 2023, the Panel published a Directions Paper to seek feedback on the key findings and insights gained from the Panel's simulation modelling of the potential future changes to USE characteristics. Based on feedback, the Panel has carried out further modelling and analysis to assess whether the existing forms of the reliability standard and APC remain fit-for-purpose.

5 This Draft Report outlines the Panel's draft recommendations on the form of the reliability standard and APC. The Panel seeks stakeholder feedback on the draft recommendations to inform the Final Report, which will be published in June 2024.

Our draft recommendation is to maintain the current form of the standard

6 The current form of the reliability standard is based on expected USE value, which is a combination of the likelihood and size of events. Under this form, all events generated through reliability modelling influence the overall calculated expected value.

7 Based on the extensive modelling completed for this Review, the Panel considers that the existing form of the standard continues to be fit-for-purpose and can adequately capture the changing risk profile as the NEM transitions.

While the changing risk profile brings new challenges, the current form remains fit for purpose

8 There was a concern prior to this Review that the changing reliability risk in a high-VRE system may lead to very extreme reliability shortfalls and the current form of the standard may not adequately address such risk. This has been a critical focus for the Panel in its modelling work.

9 However, our latest modelling results do not suggest there is any significant risk of such extreme events. While there is a small risk of large USE events well into the future, these remain a small part of the overall reliability risk in the NEM (noting that no system can be perfectly reliable). Based on this, there is no clear need to change the form of the standard.

10 Our modelling has demonstrated that, while the characteristics of the reliability risk are expected to change in a number of ways, the current form of the reliability standard still adequately captures the vast majority of USE events that will likely arise in the NEM. Therefore, it remains an effective way to measure reliability risk and weigh it against the costs of increased reliability.

Preventing extremely rare USE events would likely cost more than consumers are willing to pay

11 Our simulations of the future NEM produced a small number of large but rare USE events which could still occur under the existing reliability standard. However, their probability and likelihood of occurrence is very small, approaching 1 in 100 or more years.

12 The Panel considers that the risk of such large, low-probability USE events cannot be adequately addressed by any form of the reliability standard, and would need to be addressed in other ways. This is because the reliability standard, as a tool, is not intended to achieve absolute reliability. Instead, it is designed to enable a trade-off between reliability and affordability such that it achieves a level of reliability based on consumers' willingness to pay. Using the reliability standard to address a small proportion of very rare USE events will likely result in an excessive cost burden on consumers, regardless of which form it takes.

13 The Panel has also considered that changing the form of the reliability standard would be unlikely to address how the market price settings could, in isolation, drive investment to manage the risk of severe, low-probability events. The market price settings are set at a level that will drive sufficient investment in generation and/or storage to meet the reliability standard. However, the commercial business case for generation and storage projects becomes increasingly difficult if they can only earn revenue on very rare occasions, regardless of the level of the market price settings and the form of the reliability standard.

We agree with AEMO that communicating the standard in tangible ways is important

14 The Panel considers that the current form of the reliability standard can be expressed in various ways to help illustrate the changing characteristics of reliability risk. In their submissions to the Issues Paper and Directions Paper, AEMO raised a concern that the current form of the standard is difficult for stakeholders and consumers to interpret in terms of real-world outcomes. AEMO considers that a measure related to depth, duration and frequency would be more suitable.

15 The Panel agrees that it is important for stakeholders to be able to understand reliability in a tangible way. However, this can be achieved without changing the form of the standard. The form of the standard, while defined as the expected value of unserved energy, can also be expressed using depth, duration and frequency. For example, in the 2023 ESOO AEMO presents the depth and duration of forecast USE from their simulations for each region. The Panel supports AEMO communicating the existing reliability standard in alternative ways in its materials.

We recommend process improvements to enhance the operation and implementation of the standard

16 The Panel has identified three possible process improvements to enhance the operation and implementation of the reliability standard. These improvements are either already underway or could be implemented by the Panel or AEMO.

17 Firstly, the Panel supports AEMO's continued work on enhancing its reliability modelling to take into account a larger range of weather conditions. AEMO's reliability modelling currently uses a set of 13 weather reference years based on the previous 13 years of weather data. Whilst these years do show a diversity of weather conditions, including some extreme events, it is still a limited set compared to the wider range of historical outcomes and potential future weather outcomes. We understand that AEMO is developing synthetic weather years for its future modelling and the Panel supports AEMO's continued work in this regard.

18 Secondly, the Panel considers there is scope to improve the modelling of future demand traces for reliability forecasting. AEMO constructs future demand traces to meet their forecasts for total annual energy and annual and seasonal peak demand. The Panel's modelling suggests that peak demand may not be a key driver of USE events in the future. The Panel understands that AEMO has been updating its forecasting methodology to better reflect other characteristics of demand, such as cumulative demand and other seasonal weather factors.

19 Finally, the Panel has identified an opportunity to improve the way in which it applies the VCR in its RSSR process. The 2022 RSSR used the AER's load-weighted jurisdictional VCR values as the base case, with re-weighted high and low case sensitivities. Consistent with feedback from stakeholders, the insights gained from this review create opportunities for alternative ways to apply or re-weight the VCR results for the RSSR. The weightings used to derive the AER's main state-based VCR values are based on historical customer outages from all causes. The Panel proposes that the VCR values used for the RSSR should be weighted according to the characteristics of future customer outages caused by reliability shortfalls, where feasible.

There are other ways to support reliability outcomes during the transition

20 The Panel recognises that as the market transitions to net zero, factors outside of the market frameworks can increasingly impact reliability in the short term. The Panel's modelling results suggest that the extent of extreme USE events is likely to remain limited. However, there are several other pathways available to support reliability as the NEM transitions if policymakers wish to mitigate these rare events.

21 While not forming part of the Panel's recommendations for this Review, the Panel has identified some possible pathways and factors outside the market frameworks to support reliability outcomes for consideration by other policymakers. These include:

- Reforms currently being considered or implemented to support reliability in the NEM, including the Capacity Investment Scheme and other jurisdictional investment schemes, progressive increases to the MPC and CPT, the Retailer Reliability Obligation and the Interim Reliability Reserve.
- The 2024 Draft ISP and additional tools outside the reliability framework to deliver the ISP and support AEMO's ability to operate the system.
- Addressing non-market barriers to the delivery of the ISP, such as supply chain, workforce and transmission constraints, planning and environmental approvals, and social licence issues.
- The uncertainty of thermal plant exits which is impacting market signals.

22 There are other support mechanisms that policymakers can consider exploring further to increase reliability at varying costs, including, for example, strategic reserves, capital grants to lower the funding needs for the project, swaption style arrangements, cap contracts or contracts for difference, reserve payments and government build-to-own or joint venture projects.

We have carried out rigorous modelling and evidence-based analysis to inform the draft recommendations

23 Since the publication of the Directions Paper, the Panel has undertaken further modelling work based on stakeholder feedback and as foreshadowed in the Directions Paper. The purpose of this further work was to improve an understanding of the changing reliability risk profile and to enable assessment on whether the current form of the standard remains fit for purpose.

24 Consistent with the approach taken for the Directions Paper, the Panel has undertaken further

modelling work based on a simulation of a virtual future power system that is deliberately constructed to create insights about its unserved energy profile. However, a number of changes and additions have been made based on stakeholder feedback to further improve confidence in the Panel's modelling results. These include implementing an alternative approach to removing capacity, validating insights using a full capacity model, using a much larger set of weather reference years and generating synthetic weather data to understand the likelihood of dark doldrums.

25 The Panel's further modelling reinforces the Directions Paper's four key insights about the changing reliability risk in the NEM. The Panel has re-explored these insights to gather additional evidence, improve confidence in the modelling results, and better understand some of the nuances. The results from the Panel's further modelling work reinforce the following four key insights as presented in the Directions Paper:

- given a constant level of reliability, expected USE events are likely to become deeper and/or longer but less frequent
- expected USE is likely to shift from mainly being in summer to winter
- USE events are likely to be driven increasingly by weather
- events are likely to occur across the day rather than just appearing in the evening peak.

26 The Panel has quantitatively evaluated the modelling results to understand whether the USE events that appeared in the simulation modelling are large enough or likely enough to be material to reliability forecasts such that they are adequately captured by the existing form of the reliability standard that is based on the expected USE value.

27 The results from the Panel's analysis indicate that the vast majority of USE outcomes adequately influence the level of reliability calculated under the current standard and therefore are captured by the current form of the standard:

- Most USE results seen in the modelling are either large enough, or likely enough to be material to reliability forecasts under the current standard.
- There are a small proportion of rare USE events, which may not be material to reliability forecasts under the current form. However, their probability and likelihood of occurrence approaches 1 in 100 or more years, indicating these events are too rare to be appropriate for any reliability standard as it is not intended to achieve absolute reliability.

28 To further test and improve confidence in these results, the Panel used various statistical resampling methods, including using a Markov chain Monte Carlo model to generate 100,000 synthetic years. This model resampled the modelling results from the virtual future power system constructed by using 94 reference years. This produced a more robust distribution of USE outcomes than the 94-year sample noting that 94 data points is relatively few from which to infer probabilities.

29 In addition to the insights that the current form adequately captures the majority of USE events, the Panel's modelling work has also identified that the changing reliability risk is unlikely to result in any extremely severe events. This analysis was undertaken in response to some stakeholders' concerns that the changing reliability risk may lead to very extreme tail USE events that could potentially lead to multiple day-long outages. The Panel's modelling results indicate that even under a very large range of weather conditions, the modelling did not produce any multi-day events (i.e. any events of more than one full day of unserved energy).

30 Further, the Panel's modelling has also demonstrated that there are very few events that would result in customers experiencing multiple rotational outages. The Panel's modelling results were

translated into a customer perspective by analysing the number of USE events in which customers would likely experience more than one round of rotational load shedding. Overall, multiple rotational load shedding events were very uncommon, with less than 2 per cent of total USE events resulting in customers experiencing more than one round of load shedding. For the avoidance of doubt, the modelling shows that USE events are rare, and these results indicated that only a very small fraction of those USE events lead to multiple load shedding.

Our draft recommendation is to maintain the current form of the APC

31 The Panel's recommendation on the form of the APC is to maintain its current form and regularly review its level through the RSSR process.

32 The Panel considers that the current form of the APC continues to be fit for its intended purpose of protecting market participants from extended periods of high prices. It is important that the APC be sufficiently high to minimise reliance on compensation, for generators to recover costs, and be able to contribute to reliability during administered pricing periods. To this end, the AEMC recently increased the level of the APC to $600/MWh, which the Panel sees as sufficiently high to serve this purpose. The level of the APC will be regularly reviewed to ensure this remains true.

33 In making its recommendation, the Panel has considered indexing the APC to CPI, but found that this was not necessary, provided the level of the APC is high enough to cover the short-run marginal costs (SRMC) of the marginal generator with minimal reliance on compensation.

34 The Panel acknowledges the importance of carefully monitoring the level of the APC given the events in June 2022 and the misalignment between the electricity APC and fuel prices.

35 That is why the Panel intends to review the level of the APC in each RSSR, and make any necessary changes to it as and when required. This will ensure the level continues to efficiently minimise reliance on compensation to recover the SRMC of the marginal generator or battery energy storage system (BESS).

Stakeholder submissions are due by 17 May 2024

36 Stakeholder feedback on this Draft Report will be instrumental in informing the Panel's final recommendations on the form of the reliability standard and APC.

37 Specifically, this Draft Report seeks stakeholder feedback on the following draft recommendations:

- **The form of the reliability standard |** Maintain the current form of the reliability standard while making a range of process improvements on how the current standard is communicated, operationalised and implemented.
- **The form of the APC |** Maintain the current form of the APC and undertake a regular review of the level of the APC with every RSSR.

38 After the Draft Report, the Panel will publish a Final Report in June 2024. The Final Report will set out the Panel's final recommendations on this Review based on stakeholder feedback on the Panel's draft recommendations, supporting rationales and evidence outlined in this Draft Report.

39 Following the Panel's Final Report on this review, the 2026 RSSR will then consider the level of the standard and the implementation of various process improvements recommended from this review, where appropriate.

Contents

Appendices

Tables

Figures

1 Introduction

1.1 The reliability standard and its role in the national electricity market

A reliable power system has an adequate amount of capacity (generation, demand response and interconnector capacity) to meet customer needs. This requires adequate investment in capacity, including sufficient investment to cover generator retirements, as well as an appropriate operational framework so that supply and demand can be maintained in balance at any particular point in time.

The reliability standard is a core element of the national electricity market's (NEM's) framework for delivering reliability. It provides a clear, actionable expression of the economically efficient level of generation and transmission capacity sought for the NEM.

The standard is an ex-ante standard to indicate to the market the required level of supply to meet demand on a regional basis. It is not a regulatory or performance standard that is 'enforced'. Rather, it is used to indicate the efficient level of reliability for the purposes of informing the market under the NEM reliability framework (see Appendix A for more information).

The standard is based on an economic trade-off made on behalf of consumers as to the appropriate level of reliability. It is a key input to the various market settings, that is, the market price cap (MPC), market floor price (MFP), cumulative price threshold (CPT), and administered price cap (APC) that define the price envelope that applies to spot market outcomes.

The Australian Energy Market Operator (AEMO) is responsible for operationalising the standard through its forecasting processes, by modelling and projecting whether the market is going to meet the standard. It does this across a number of time frames, from years ahead of real-time, up until real-time, through the various Electricity Statement of Opportunities (ESOO), projected assessment of system adequacy (PASA) and pre-dispatch processes.

1.1.1 Reliability events occur when supply does not meet demand

Understanding unserved energy (USE) and reliability and security events is essential to understanding how the reliability standard acts as an investment trigger supporting reliability.

The National Electricity Rules (NER) state that for the purposes of the reliability standard, USE in megawatt hours (MWh) includes energy demanded but not supplied due to power system reliability incidents resulting from:

- a single credible contingency event on a generating unit or an inter-regional transmission element that may occur concurrently with generating unit or inter-regional transmission element outages; or
- delays to the construction or commissioning of new generating units or inter-regional transmission elements, including delays due to industrial action or acts of God (such as extreme weather events).[1]

The NER specifies that a 'power system reliability incident' is an incident that AEMO considers would have been avoided only if additional active energy had been available to the relevant region or regions from generation, demand response or inter-regional transmission elements.

1 NER clause 3.9.3C(b)(1).

USE excludes energy demanded but not supplied due to power system security incidents resulting from:

- multiple credible contingency events[2], protected events or non-credible contingency events on a generating unit or an inter-regional transmission element, that may occur concurrently with generating unit or inter-regional transmission element outages;
- outages of transmission network or distribution network elements that do not significantly impact the ability to transfer power into the region where the USE occurred; or
- industrial action or acts of God at existing generating facilities or inter-regional transmission facilities.[3]

1.2 The AEMC has asked the Reliability Panel to review the form of the reliability standard and APC

In March 2023, the Australian Energy Market Commission (AEMC) issued terms of reference that requested that the Reliability Panel (the Panel) provide advice on the form of the reliability standard.[4] This followed the Panel's 2022 Reliability Standard and Settings Review (RSSR) which recommended a review to consider changing the form of the reliability standard.[5]

The Panel considered it necessary to review how the reliability risk profile is changing as the NEM transitions, and whether an expected value reliability standard remains appropriate in the transitioning NEM.

The terms of reference also requested that the Panel review the form of the APC. In the 2022 RSSR, the Panel identified the potential need to change the APC following the June 2022 market suspension and the administered pricing period. The Panel considered a range of options in the March 2023 Issues Paper and shortlisted two options for further consideration in the November 2023 Directions Paper.[6] The Panel has outlined its draft recommendation in this report.

1.3 The Panel consulted on its key findings and insights

The Panel published a Directions Paper for this Review in November 2023. The Directions Paper set out the key findings and insights from the Panel's modelling work on the potential changes to the characteristics of unserved energy as the NEM transitions.

The Panel's approach was to take a model of the NEM based on AEMO's ESOO and Integrated System Plan (ISP) Step Change scenario and remove capacity so that the model demonstrates unserved energy events.[7]

This deliberately under-resourced system model has materially less generation than is forecast in planning documents such as the ISP. This was done to create a larger data set from which to study the possible characteristics of USE in the future. It should be noted that the model is an extreme scenario that was used to provide a greater range and number of USE outcomes to study.

The Panel's simulation modelling has generated four key insights regarding the changing characteristics of reliability as the system transitions. Note that these trends will gradually emerge

2 NER clause 4.2.3.
3 NER clause 3.9.3C(b)(2).
4 Terms of reference found here.
5 Reliability Panel, 2022 RSSR found here.
6 Reliability Panel, 'Review of the form of the reliability standard and APC', Issues Paper and Directions Paper, found here.
7 AEMO 2022 ESOO, found here, AEMO 2022 ISP, found here.

as the NEM transitions to a higher variable renewable energy (VRE) penetration system in the 2030s and 2040s, and are not likely to emerge suddenly in the next 10 years. These trends include:

- given a constant level of reliability, expected USE events are likely to become deeper and/or longer but less frequent
- expected unserved energy is likely to shift from mainly being in summer to winter
- USE events is likely to be driven increasingly by weather
- events are likely to occur across the day rather than just appearing in the evening peak.[8]

The Directions Paper sought stakeholder feedback on the Panel's modelling approach and key findings to inform the Panel's draft recommendations, which are set out in this report.

The Panel also sought stakeholder feedback on the two options shortlisted for the form of the APC. The two options were retaining the current form of the APC and indexing the APC to the Consumer Price Index (CPI). The Panel has considered stakeholder feedback on the Directions Paper and outlined its draft recommendation in this report.

1.4 This Draft Report seeks stakeholder feedback on the Panel's draft recommendations

Building on stakeholder feedback to the Directions Paper, the Panel has carried out further work to gather additional information on the changing characteristics and to provide evidence on whether a change to the form of the reliability standard and APC is warranted. The draft recommendations for the Review outlined in this report have been informed by this work.

The key objective of this Draft Report is to present the Panel's draft recommendations and supporting rationale based on the Panel's analysis.

Specifically, the Draft Report seeks stakeholder feedback on the following draft recommendations:

- **The form of the reliability standard |** Maintain the current form of the reliability standard while making a range of process improvements on how the current standard is communicated, operationalised and implemented.
- **The form of the APC |** Maintain the current form of the APC and undertake a regular review of the level of the APC with every RSSR.

Stakeholder feedback on this Draft Report will be instrumental in informing the Panel's final recommendations on the form of the reliability standard and APC. The Panel will publish a Final Report for the Review in **June 2024**.

Following the Panel's Final Report on this Review, the 2026 RSSR will then consider the level of the standard and the implementation of various process improvements recommended from this Review, where appropriate.

8 Reliability Panel, 'Review of the form of the reliability standard and APC', Directions Paper, found here.

1.5 How the Panel has identified and assessed options

In identifying and assessing a range of options for this Review, the Panel has applied the assessment framework as outlined in the Issues Paper published in March 2023.

The Issues Paper outlined the Panel's assessment principles and approach to the Review and terms of reference. The Panel noted that it would apply to this Review an assessment framework that is consistent with the 2022 RSSR, which includes:

- the general assessment principles in the Panel's 2021 Final Guidelines Review of the Reliability Standard and Settings Guidelines (2021 Guidelines) to contribute to the National Electricity Objective (NEO)[9]
- the overarching assessment criteria and considerations set out in the terms of reference
- specific objectives and criteria for assessing different candidate metrics.

1.6 How to make a submission

Due date: The Panel invites feedback from stakeholders in response to this Draft Report. Submissions must be lodged with Panel by **17 May 2024**.

How to make a submission: Go to the AEMC's website, www.aemc.gov.au, find the 'lodge a submission' function under the 'Contact Us' tab, and select the project reference code REL0086.[10]

Tips for making submissions are available on our website.[11]

Publication: The AEMC publishes submissions on its website. However, we will not publish parts of a submission that we agree are confidential or that we consider inappropriate (for example offensive or defamatory content or content that is likely to infringe intellectual property rights).[12]

9 It should be noted that the NEO now includes the new emissions reduction component, which was implemented after the publication of the Issues Paper in March 2023.

10 If you are not able to lodge a submission online, please contact us and we will provide instructions for alternative methods to lodge the submission.

11 See: https://www.aemc.gov.au/our-work/changing-energy-rules-unique-process/making-rule-change-request/submission-tips.

12 Further information is available here: https://www.aemc.gov.au/contact-us/lodge-submission.

2 Stakeholder feedback on the Directions Paper

The Panel received 17 submissions to the Directions Paper that provided feedback on the modelling results and approach, views on whether this equates to a need to change the form of the standard, and the options on the APC.

This section provides a summary of stakeholder feedback which has informed the Panel as it developed its draft recommendations.

2.1 Implications of the modelling

2.1.1 The issue

The Directions Paper outlined the Panel's key modelling insights on the changing characteristics of reliability as the system transitions.[13] The Panel sought stakeholders' feedback on the implications of these insights in assessing whether the existing form of the reliability standard remains appropriate.

2.1.2 Stakeholder views

There was general agreement from stakeholders that the nature of reliability risk is changing. However, among stakeholders who expressed a clear view on the implications of the modelling, most of them considered the modelling results did not necessarily indicate there is a need to change the form of the reliability standard. Several of those considered that the existing form of the standard would remain suitable for the future and should not be changed. Some stakeholders did not provide a clear view in their submissions, noting further work is required to better understand the implications of the modelling.

For example, Shell Energy noted that:[14]

> Based on the modelling results presented, Shell Energy sees little evidence to clearly demonstrate that a change to the reliability standard is needed. The profile of USE presented does not appear to be markedly different from the existing market. Changing the resource mix to one more dominated by wind and solar generation only highlights that the real driver of USE is the volume of firm, dispatchable capacity in the market.

However, AEMO, Hydro Tasmania and the Strategic Policy and Delivery Division of the Department of Energy and Mining (SA Government) viewed the modelling findings as implying a new form of the standard may be required.[15]

AEMO considered that the Panel's modelling shows a shift in reliability risk and that "[t]he form of the reliability standard should be updated to clearly describe the changing type of risk in the system."[16] AEMO noted that:[17]

> [It] is supportive of the way the Directions Paper analysed and presented the findings, that is, describing the depth, duration and frequency of large events that were seen in the forecast. This is a clear and tangible way to describe the nature of reliability risk in the system – by describing the large events that are increasingly presenting the biggest

13 Reliability Panel, 'Review of the form of the reliability standard and APC', Directions Paper, 2023, p. 21-33.

14 Shell Energy submission to Directions Paper, p. 1-2.

15 Submissions to Directions Paper: AEMO, p. 1, 7; Hydro Tasmania, p. 1-2; SA Government, p. 1-2.

16 AEMO submission to Directions Paper, p. 1.

17 AEMO submission to Directions Paper, p. 5.

> reliability risk to consumers. It is therefore important that the form of the standard likewise communicate this risk in a way that tangible characterises the potential for large events in the system.

Further, AEMO outlined that in its view, "the priority should be to ensure that the form of the reliability standard explains the nature of large events that may occur, allowing governments and stakeholders to then make decisions and best manage their expectations of possible USE outcomes".[18]

The SA Government considers the modelling results provide "indicative evidence" for increased tail risk. The SA Government noted two shortcomings of the current reliability framework, including:[19]

1. the current form of the standard does not adequately capture risk owing to the changing dispersion of USE events throughout the year and the day and the influence of VRE on generation output
2. that it relies on the characteristics of historical USE events.

The SA Government's view is that "[t]he Panel's findings reinforce the Divisions' position that, particularly in its role as a market signal for investment decisions, the form of the reliability standard should better reflect how electricity is increasing being delivered and demanded".[20]

Further detail on the Panel's recommendation on the form of the standard are found in Chapter 3.

2.2 Feedback on the modelling approach

2.2.1 The issue

To assess the changing reliability risk as the NEM transitions, the Panel created a simulation model to explore the changing conditions under which there is unserved energy. The Directions Paper outlined the Panel's modelling approach, and sought stakeholders' feedback on the approach.[21]

2.2.2 Stakeholder views

Most stakeholders recognised the importance of undertaking the work to better understand the changing reliability risk, and acknowledged there were various challenges in accurately modelling the potential reliability risks as the NEM transitions.[22] Stakeholder views were varied on the modelling approach adopted by the Panel.

AEMO supported the Review's modelling approach and emphasised the importance of the role demand plays in weather-driven energy shortages.[23] However, AGL suggested that the choice to under-resource the model was the key driver of the findings of the exercise.[24] Shell Energy also noted that little USE occurs if the model is not under-resourced and thus determined it was unclear if the base case is realistic.[25]

18 AEMO submission to Directions Paper, p. 5-6.
19 SA Government submission to Directions Paper, p. 1-2.
20 SA Government submission to Directions Paper, p. 2.
21 Reliability Panel, 'Review of the form of the reliability standard and APC', Directions Paper, 2023, p. 15-20.
22 Submissions to Directions Paper: AEC, p. 1; Shell Energy, p. 1; AEMO, p. 2.
23 AEMO submission to Directions Paper, p. 3, 4.
24 AGL submission to Directions Paper, p. 1, 2.
25 Shell Energy submission to Directions Paper, p. 1.

Stakeholders, including the Australian Energy Council (AEC), CS Energy, Energy Users Association of Australia (EUAA) and Shell Energy also noted the differences between the generation mix used for the modelling exercise, which was based on the 2022 ISP, and the Draft 2024 ISP, with particular regard to the amount of gas-fired generation capacity.[26]

The AEC made some suggestions to further improve the Panel's modelling. These suggestions include that the modelling should include demand response, should test the ability of gas infrastructure to operate for extended periods, and that the Griffith University weather years be matched with expected demand.[27]

Some stakeholders made other suggestions to improve the modelling. These include:[28]

- the consideration of customer energy resources (CER) uptake as a factor in future reliability modelling (Ergon Energy Queensland)
- clarity regarding perfect foresight of batteries (Origin Energy)
- consideration of: (Origin Energy)
  - the declining reliability of thermal generation
  - the performance of older solar and wind farms over time
  - weather-related derating of transmission
- whether the exclusion of non-VPP demand response led to an overstatement of reliability risk (Shell Energy).

Detail on how the Panel has refined its modelling approach in line with stakeholder feedback is provided in Chapter 5 and Appendix B.

2.3 Relationship between this Review and the Value of Customer Reliability

2.3.1 The issue

Both the Issues Paper and the Directions Paper identified the interaction between this Review and the Value of Customer Reliability (VCR) process. The VCR is a key input into the reliability framework as the determinant of the cost of unserved energy. The interaction with the AER's upcoming 2024 VCR Survey has also been identified by the AER and the Panel has engaged closely with them to understand the implications of customer preferences to the form of the standard and identify opportunities to improve the way the Panel utilises the VCR in the reliability framework.

The Directions Paper also outlined that how customers value the changing nature of reliability would be an important consideration for the Panel to balance the affordability and reliability challenges that may emerge as a result of the changing reliability risk.

While the primary focus of the current Review is the form of the reliability standard, the Panel considered that VCR is more closely related to setting the level of the standard. The findings from the 2024 VCR will inform what the level of the standard needs to be as part of the 2026 RSSR process.

The Directions Paper also included a desktop review of the existing literature and international case studies that are relevant to the consideration of customer preferences towards reliability outcomes.

26 Submissions to Directions Paper: AEC, p. 4; CS Energy, p. 2; EUAA, p. 4; Shell Energy, p. 1.

27 AEC submission to Directions Paper, p. 4.

28 Submissions to Directions Paper: Origin Energy, p. 1; Ergon Energy Queensland, p. 1; Shell Energy, p. 1.

Stakeholders were asked for feedback on the relevance of the VCR to the form of the reliability standard as well as the potential synergies between AER's VCR work, this Review and future Reliability Standard and Settings Reviews.

2.3.2 Stakeholder views

Five stakeholders considered that the VCR is a key determinant of the *level* of the reliability standard but is not relevant to considerations of the *form* of the reliability standard (AEC, AEMO, AGL, EUAA, EnergyAustralia).[29] EnergyAustralia noted that even if the VCR methodology revealed the presence of risk aversion among consumers, this may not justify a change to the form of the standard.[30] Three of these stakeholders (AEC, AEMO, AGL), however, supported further examination of the VCR methodology and Panel engagement with the AER to improve these processes to better reflect future USE events.[31]

Four other stakeholders (Clean Energy Council (CEC), SA Government, Hydro Tasmania, Public Interest Advocacy Centre (PIAC)) also commented on the utility of further work on the VCR and/or collaboration with the AER to improve the methodology.[32] The CEC and SA Government expressed this point in the context of the impact of CER on customer preferences, with the SA Government encouraging cooperation to reflect "high levels of DER, 'tail' events, and the longer deeper events".[33]

Additional issues raised by stakeholders included:

- The ANU submission that the experiences of disadvantaged rural communities is not adequately captured by the current VCR methodology.[34]
- Ergon Energy Queensland's submission that the Panel should consider differences amongst customer groups, for example, those with and without rooftop solar or those in positions of hardship.[35]
- EnergyAustralia commented that recent reliability measures suggest a disparity in risk appetite between policymakers and consumers, and added further that the VCR methodology could inform these measures.[36]

2.4 Other issues for the form of the standard

Several stakeholders took the opportunity to provide feedback on issues they considered to be relevant to the Review but were not identified in questions to stakeholders provided in the Directions Paper. These are summarised below.

2.4.1 Stakeholder feedback

The CEC and Ergon Energy Queensland identified the interaction of the reliability outlook and the Capacity Investment Scheme (CIS), with the CEC commenting that the Panel should "consider the implications of changing the USE method in relation to other market signals such as the Capacity Investment Scheme or the Reliability and Emergency Reserve Trader."[37]

29 Submissions to Directions Paper: AEC, p. 3; AEMO, p. 7-8; AGL, p. 1; EUAA, p. 3-4; EnergyAustralia, p. 1, 2.
30 EnergyAustralia submission to Directions Paper, p. 1,2.
31 Submissions to Directions Paper: AEC, p. 3; AEMO, p. 7-8; AGL, p. 1.
32 Submissions to Directions Paper: CEC, p. 1; SA Government, p. 2; Hydro Tasmania, p. 1; PIAC.
33 SA Government submission to Directions Paper, p. 2.
34 ANU submission to Directions Paper, p. 1-2.
35 Ergon Energy Queensland submission to Directions Paper, p. 2.
36 EnergyAustralia submission to Directions Paper, p. 3.
37 Submissions to Directions Paper: CEC, p. 2; Ergon Energy Queensland, p. 1.

CS Energy and EUAA commented on the interaction between the reliability framework and government and jurisdictional measures more generally, noting that a new reliability standard could result in a duplicative cost to consumers for reliability in lieu of government measures to curb reliability risk.[38] Baringa, however, considered that government schemes can complement the reliability framework to ensure sufficient investment.[39]

Further detail on the Panel's views supporting reliability as the NEM transitions is provided in section 3.4.

2.5 The form of the administered price cap

2.5.1 The issue

The Directions Paper narrowed the suite of options for the form of the APC to two options; maintain the current form of the APC or index the APC to CPI. Stakeholders were asked to provide reasons for their preferred option.

2.5.2 Stakeholder views

Twelve stakeholders commented on the form of the APC. Most of these stated a preference towards indexing the APC to CPI, though many also considered the current form of the APC to be suitable.

Many of these submissions quoted consistency with the MPC and CPT process among the principal reasons for supporting an APC that is indexed to CPI. Three stakeholders did not state a preference between the provided options. However, one of these (Shell Energy) provide an alternative option for the APC to operate as a discounted MPC, which Shell Energy notes is indexed to CPI.[40]

EnergyAustralia explicitly preferred the current form to an indexed APC, commenting that indexation would create an additional administrative burden and, further, that the RSSR is a sufficiently frequent opportunity to adjust the level of the APC as required.[41] CEC, however, cited reduced future administrative burden as a reason to index the APC.[42]

The Panel's draft recommendation on the form of the APC is outlined in Chapter 8.

38 Submissions to Directions Paper: CS Energy, p. 3; EUAA, p. 1-2, 5.

39 Baringa submission to Directions Paper, p. 2, 5.

40 Shell Energy submission to Directions Paper, p. 3-4.

41 EnergyAustralia submission to Directions Paper, p. 3-4.

42 CEC submission to Directions Paper, p. 2.

3 The Panel's draft recommendation is to maintain the current form of the reliability standard

The Panel's draft recommendation on the form of the reliability standard is that:

- the form of the reliability standard should remain as expected USE because this form continues to be fit-for-purpose in the future NEM
- the Panel works closely with other market bodies to consider implementing improvements to the modelling and decision-making processes around setting and implementing the reliability standard, as well as how the standard is communicated.

This chapter outlines the Panel key reasons for its draft recommendation.

In addition, the Panel has also identified, for consideration by other policymakers, other tools and mechanisms that could support the NEM's future reliability needs based on this Review's key findings and insights. The Panel has also outlined its consideration of how to support AEMO in continuing to communicate the reliability standard in different ways, including using depth, duration and frequency.

3.1 Reasons for the Panel's draft recommendation

3.1.1 While the changing risk profile brings new challenges, the current form of the standard remains fit for purpose

Based on the extensive modelling completed for this Review, the Panel considers that the existing form of the standard continues to be fit for purpose and can adequately capture the changing risk profile as the NEM transitions. In the future, while the nature of reliability risk may change in a number of ways, the expected value of USE remains an effective way to measure that risk and weigh it against the costs of increased reliability. However, the NEM's transition will bring challenges in communication, system operation, and delivery of new transmission and supply.

Furthermore, there may be a need for complementary mechanisms alongside the standard if policymakers want a level of reliability above what consumers are willing to pay (see section 3.3).

Modelling results show that the current form reflects USE risk effectively for a range of risk profiles

The Panel's latest modelling reinforces the Directions Paper findings about the changes in reliability risk as the NEM transitions (see Chapter 6). The modelling showed that the shape of the USE distribution (the characteristics of events and how likely they are) would likely change. The key findings included the following:

- reliability risk is likely to shift from summer to winter over time
- as may be expected in a VRE-dominated system, weather is likely to have an increasingly strong influence on USE, likely stronger than that of forced outages
- while the overall level of risk will not necessarily increase, the USE events that do occur will likely become longer and deeper.

While reliability risk is changing, the Panel's view is that this does not result in a need to change how we assess reliability. Despite the changing characteristics, our analysis has shown that the existing form of the standard is capable of reflecting a wide range of risks. Therefore, the Panel considers the current form of the standard will remain fit for purpose for the future NEM.

As noted previously in the Review, the current form does place a low weighting on extremely rare events (those expected approximately once in 100 years).[43] However, the modelling shows that even these rare USE events would not be as extreme or extended as was originally suspected. Notwithstanding that, the Panel considers that the risk of extremely large, low-probability events cannot be adequately addressed by any form of standard and would need to be addressed in other ways, as discussed in section 3.1.3 and section 3.3.

The risk of rare, extreme USE events remains a small part of the overall reliability risk

The risk of severe or extended reliability shortfalls related to unfavourable weather conditions in a high-VRE system has been a critical concern for the Panel. However, the Panel's latest modelling results do not suggest there is any significant risk of such extreme events. While there is a small risk of large USE events in the future, these do not form a significant part of the overall reliability risk in the NEM (noting that no system can be perfectly reliable). Based on this, there is no clear need to change the form of the standard. In section 3.3, we explore a range of other ways to potentially manage the risk of large, low-probability reliability shortfalls.

3.1.2 The Panel agrees with AEMO regarding the communication of the standard

The Panel's modelling shows that the depth, duration and frequency of potential USE events may change as the NEM transitions, even if the total expected USE remains the same. In their submission to the Issues Paper, AEMO raised a concern that the current form of the standard is difficult for stakeholders and consumers to interpret in terms of real-world outcomes. AEMO considers that a measure related to depth, duration and frequency would be more suitable:[44]

> AEMO considers reliability should be communicated in terms of tangible measures for consumers such as depth, duration, and frequency.

The Panel agrees that it is important for stakeholders to be able to understand reliability in a tangible way, but considers that this can be achieved without changing the form of the standard. The Panel also notes that, while acknowledging the importance of ease of communication, the reliability standard has other purposes besides communication.

The form of the standard, while defined as the expected value of unserved energy, can also be expressed using depth, duration and frequency. For example, in the 2023 ESOO AEMO presents a bubble plot of depth and duration of forecast USE from their simulations for each region.[45] The plot shows a summary of all the simulation outcomes, highlighting the vast percentage of yearly outcomes in which there is no USE. It also shows outcomes that exceed the current 0.002 per cent USE reliability standard. AEMO stated in their submission to the Issues Paper that they are increasingly expressing USE in this way:[46]

> AEMO is of the view that expressing annual average USE as depth, duration, and frequency better describes the customer experience, which then can be set in a standard to reflect societal tolerance to tail events. AEMO considers this a priority and is increasingly characterising USE in this way as part of its ESOO and other forecasting functions.

Therefore, the current form of the standard remains fit for purpose because it can be expressed in various ways which may help to illustrate the characteristics of the risk. However, the Panel, notes

43 Reliability Panel, 'Review of the form of the reliability standard and APC', Issues Paper, p. 10, found here.

44 AEMO submission to Issues Paper, p. 1.

45 AEMO, 2023 ESOO, p. 60.

46 AEMO submission to Issues Paper, p. 14.

that communicating the reliability risk in terms of depth, duration and frequency may still raise some challenges as it still requires articulation of a probability outcome. The Panel supports AEMO communicating the existing reliability standard in alternative ways.

3.1.3 The market price settings in isolation are not designed to manage large, rare events

The reliability standard takes effect primarily through the market price settings, noting it is also used in other contexts such as AEMO's ESOO. The market price settings are set at a level that will drive sufficient investment in generation and/or storage to meet the reliability standard, as determined by RSSR modelling.

As outlined above, the reliability standard is an expected USE value which takes into account both the size and likelihood of potential USE events and the level of reliability customers value. The design of the wholesale market creates a price signal to incentivise investment in base load and peak generators. It can incentivise technologies such as gas-fired generation that are dispatched relatively rarely, as long as the price cap is high enough to provide revenue sufficiency from operating at those limited peak times. However, the commercial business case for generation and storage projects becomes increasingly difficult if they can only earn revenue on very rare and unpredictable occasions, regardless of the market price cap.

The Panel considers the market price settings alone are not a suitable tool for managing the risk of severe, low-probability events. In these circumstances, changing the form of the standard would still be unlikely to provide the incentives for the required investment. Therefore, the Panel would not be justified in recommending a change to the form (or even the level) of the standard based solely on a small risk of such events. This is supported by the modelling results which show limited, if any, risk of the extreme or extended USE events that were of concern at the outset of this Review.

This Review's modelling is based on the 2022 ISP, and much of our analysis focuses on the characteristics of USE, while assuming that expected USE will remain at or below 0.002 per cent of annual regional demand.[47] The timely delivery of the ISP will be important to maintain a high level of reliability in the NEM. The Panel notes that the ISP includes some types of supply with particularly challenging investment cases, such as long-duration storage. Additional mechanisms, including government support, could be needed in combination with the market price settings to manage the risk of low-probability reliability events. Section 3.3 explores potential ways to support reliability in the transition in more detail. However, this Review has not explored whether the consumer benefits of such support mechanisms would outweigh the costs.

3.1.4 Preventing extremely rare USE events would likely cost more than consumers are willing to pay

Any increase in reliability comes at a cost associated with building, maintaining and operating additional supply or transmission. The cost of that increased reliability needs to be paid by either the market (that is, consumers), taxpayers, or a combination of both. If consumers are to pay, the level of reliability must take into account the value they place on that level. Under clause 3.9.3A of the NER, the RSSR determines the appropriate level of the standard by balancing system cost against the value of customer reliability (VCR) derived by the AER.[48]

Since the reliability standard is designed to balance reliability and affordability, it is not suited to capture and respond to extremely rare events. The market price settings could not sufficiently

47 AEMO, 2022 ISP, found here.

48 NER clause 3.9.3A.

incentivise viable projects that address the risk of rare events without an excessive burden on consumers. This would be the case regardless of the form the reliability standard.

Building assets that target the risk of large, rare (1-in-100-year) events would be prohibitively expensive and could be considered an overbuild if these events do not come to pass in the asset's lifetime.

The Panel considers these events are too rare to be appropriate for any reliability standard to address, noting, once again, that the reliability standard is not intended to achieve absolute reliability and instead seeks to balance reliability and affordability.

3.1.5 Alternative forms of the standard do not have significant benefits over the current form

As outlined above, the Panel's draft recommendation is that the current form of the standard remains fit-for-purpose for the future NEM. However, for completeness, the Panel also assessed a short list of candidate forms to understand if there would be any benefit from changing the form. Table 3.1 below outlines the options and the Panel's assessment of each. Appendix F provides a more detailed explanation of this assessment.

The current form of the reliability standard outperforms the other candidate forms in that each of them either is more difficult to implement or communicate. The Panel's assessment of the ease of implementation included consideration of RSSR market modelling, translation to market price settings, and any need for additional consumer preference data. While some forms are better than the current form in capturing large, low-probability events, the Panel put a lower weight on these criteria given that extremely large events have not been found in the modelling as noted in Chapter 7.

The Panel further notes that the current form does detect large, low probability events, and weights them proportionally to their likelihood of occurrence. This can be bolstered through process improvements such as better communication (including discussing depth, duration and frequency independently) and updated modelling, discussed in more detail in section 3.2.

Given these findings and the modelling results, the Panel considers that there would be very limited, if any, benefit to using a different form of the standard. These limited benefits are not sufficient to justify the costs and complexity of changing the form.

3.1.6 Changes can be made to improve how the current form of the standard operates

The implementation and operation of the reliability standard draws on a number of different processes including the Panel's RSSR, the AER's VCR survey, AEMO's reliability modelling, and the operation of the wholesale market. These processes must work together effectively to ensure that reliability is maintained at the desired level that consumers are willing to pay.

While this Review's modelling has found that the current form of the standard adequately captures the changing risk profile, the Panel recommends exploring opportunities to enhance the operation and implementation of the standard. The Panel has identified three potential process improvements in section 3.2 below.

Further, as discussed in section 3.1.3, the reliability standard and market price settings may not be effective in mitigating the risk of rare, severe USE events. However, if governments wish to address these events, they could explore other mechanisms to do so (section 3.3).

The Panel's draft view is that these opportunities for process improvements and support mechanisms are a more proportionate response to address the changing future reliability risk than changing the form of the standard.

Table 3.1: Summary of the Panel's assessment of options

Option	Benefits	Issues	Panel's assessment
1. Expected USE / Annual demand (Current form)	• Calculation is straightforward • Captures depth, duration and frequency in a single metric. • Potential to communicate these dimensions separately	• Detects large, low probability events reasonably well, but not as strong as other options in this regard.	• Given there is very limited evidence of extremely large events in the modelling, the costs and complexity of changing the form would outweigh the benefits
2. Re-weight probability *Inflate* the probability of low-probability, large-USE events, then calculate the *average* USE	• Relatively easy to implement and directly targets 'tail risk' events. • Maintains the positive attributes of the current form.	• Needs a definition for 'large events', and it is unclear how much to inflate the probability of large events. • Will need additional data collection • The form is sensitive to which specific events to re-weight	• While stronger than the current form in capturing large, low-probability events, the implementation issue makes this a more complex option. • The Panel does not consider that there are significant enough benefits to this option to warrant a change.
3. Conditional value at risk (CVaR) approach Calculate the average USE overall and the average USE in the tail and take a weighted average of the two	• This approach has been suggested in the 2022 RSSR study as a method for dealing with tail risk that complements the existing standard. • Depending on the weightings, can better capture large events.	• Will need additional data collection (such as consumer preferences) • It is difficult to clearly communicate the approach to stakeholders • CVaR is less stable than the current form as more weight is put onto the tail component, even in the absence of large events.	• This is a complex standard that is difficult to communicate. • CVaR can be less stable than the current form. • Given there is very limited evidence of large, low-probability events in the modelling, there will be limited benefits of moving to this option.

Option	Benefits	Issues	Panel's assessment
4. Probability of exceedance Add an additional constraint to the current form. If the probability of a pre-defined event is too high, then the constraint is violated and needs to be addressed	• Easy to understand as it directly identifies 'bad' events and says we do not want these to happen (outside of very low probability events). • In the absence of large events, the probability of exceedance measure is equivalent to the current form.	• The constraint level may or may not be binding, creating difficulty in implementation with respect to market price settings • Will need to be informed by additional data collection	• While simple to communicate, the implementation issues associated with it mean that the Panel considers there are limited benefits of moving to this approach.
5. N-1 or N-2 redundancy Assess whether USE is below a threshold during the largest credible contingency (or two largest)	• Used for the New South Wales Energy Security Target, and the Panel understands it is being considered by other jurisdictions.	• Requires intensive modelling, which makes calculating the form difficult. • The constraint level may or may not be binding, creating difficulty in implementation with respect to market price settings. • Difficult to clearly communicate compared to the current form, particularly at an operational level	• Because this is difficult to implement and challenging to model for the RSSR, the Panel does not consider this a practical option.

Source: Panel analysis; Panel's 2022 RSSR final report, p. 41-44, found here; AEMO's Energy Security Target Monitor Report October 2023, p. 3, found here.

3.2 The proposed process improvements

The Panel has identified three possible process improvements that are either already underway or could be implemented by the Panel or AEMO. Some of these process improvements are complex and will take time for the Panel and AEMO to action.

3.2.1 Reliability modelling could take into account a larger range of weather conditions

AEMO's reliability modelling currently uses a set of 13 weather reference years based on the previous 13 years of weather data.[49] Whilst these years do show diversity of weather conditions, including some extreme events, it is still a limited set compared to the wider range of historical outcomes and potential future weather outcomes. We understand that AEMO is developing synthetic weather years for its future modelling and the Panel supports AEMO's continued work in this regard.

Using synthetic or historic data such as that from Griffith University, with corresponding demand traces, could allow the ESOO, ISP and RSSR modelling to incorporate the effect of variable weather conditions on USE.[50] This could be implemented by significantly increasing the sample of weather reference years modelled, or alternatively a smaller set of weather reference years with specific conditions could be modelled and weighted appropriately according to their likelihood. Note that further climate science work would be needed to inform these weightings.

Our modelling has shown that reliability in the future power system is more sensitive to weather than forced outages. Using additional reference years in reliability modelling would allow a wider range of weather conditions, a more diverse range of USE outcomes and provides an improved basis for averaging and weighting of events.

The current methodology for the ESOO and ISP uses a 13-year weather reference year set (updated from 11 to 13 reference years in the 2023 ESOO) and around 100 randomly generated forced outage patterns. It also considers 10 per cent PoE, 50 per cent PoE and 90 per cent PoE demand conditions by growing the same historical demand to match future peak demand and energy forecast to calculate a weighted average USE outcome.

While the current set of 11-13 reference years (FY2011 to FY2023 inclusive, depending on publication) includes some dark doldrum periods, using a larger sample allows for more diverse weather conditions, including dark doldrums, which could be weighted according to the likelihood of those conditions.

To date, the key limitation preventing the use of longer weather datasets has been the lack of coincident demand data. For this Review's modelling, we used a sequence of machine learning models to link historical temperature, solar power, time of day and a range of other parameters to customer demand based on the 83-year Griffith dataset as outlined in Appendix B.

Increasing the number of weather reference years in reliability modelling could create computational challenges related to run-time and computing resources. These challenges could be managed by using more reference years with fewer forced outage patterns if forced outage patterns are still considered important as VRE penetration increases. Note that the reduced impact of forced outages will occur after most thermal plant has exited, which is likely outside the

49 AEMO 2023 ESOO, found here.

50 The Panel's modelling has used a dataset of 83 years of nominal VRE generator output in the NEM, developed by Griffith University. J Gilmore, T Nelson and T Nolan, 'Quantifying the risk of renewable energy droughts in Australia's National Electricity Market (NEM) using MERRA-2 weather data', Griffith University Centre for Applied Energy Economics & Policy Research, 2022, found here.

10-year horizon of the next ESOO, so reliability modelling may need to continue using forced outage samples for the near future. Another alternative could be selectively modelling the more severe weather conditions and developing a computational adjustment to weight these results appropriately in the expected USE.

3.2.2 There is scope to improve the modelling of future demand traces for reliability forecasting

AEMO constructs future demand traces to meet their forecasts for total annual energy and annual and seasonal peak demand.[51]

Our modelling suggests that peak demand may not be a key driver of USE events in the future. We understand that AEMO has been updating its forecasting methodology to better reflect other characteristics of demand, such as cumulative demand and other seasonal weather factors.

To generate the demand traces for the 83 Griffith years, the project team used a sequence of machine learning models that were trained on the AEMO forecast data. The Panel understands that AEMO has been looking into similar methods to operate on synthetic weather.

The analysis for this review identified that in the future weather will become a more important driver of reliability than the failure of major equipment. This is driven by the distributed nature of the VRE facilities, the variability of the resource on which they rely, and that these facilities are an aggregation of many smaller production units where the failure of one component is unlikely to affect the output more than the potential variability of the resource.

Currently the probability of exceedance levels for demand are driven primarily by temperature considerations centred around summer peak conditions. While future forecasts predict that peak demand in most states is shifting to winter driven by factors such as electrification, the USE analysis showed links to dark doldrum conditions that may not be considered in the current forecasting practices.

The current calculation of reliability in the ESOO weights USE outcomes from the 10 per cent PoE and 50 per cent PoE based on weightings determined previously.[52] In the future both the determination of the mechanism for ranking demand outcomes and the weightings by which USE outcomes are combined into the overall reliability standard may need to change to incorporate this increasing link between weather variability and USE. It is that variability and uncertainty, particularly as it relates to days preceding or comprising dark doldrums, that may need to be the focus of future classifications of the probability of exceedance levels.

3.2.3 The Panel can improve the way in which the VCR is applied in the RSSR

The 2022 RSSR used the AER's load-weighted jurisdictional VCR values as the base case, with re-weighted high and low case sensitivities.[53] Further information about how the VCR was applied in the 2022 RSSR is available in Appendix G.

Consistent with feedback from stakeholders, the insights gained from this Review's modelling create opportunities for alternative ways to apply or re-weight the VCR results for the RSSR.

The weightings used to derive the AER's main state-based VCR values are based on historical customer outages from all causes. The Panel proposes that the VCR values used for the RSSR should be weighted according to the characteristics of future customer outages caused by reliability shortfalls, to the extent this is feasible.

51 AEMO 2023 ESOO, found here; AEMO Draft 2024 ISP, found here.

52 AEMO 2023 ESOO, found here.

53 Reliability Panel, 2022 RSSR, found here; AER, Values of Customer Reliability 2019, final report, found here.

There are two sources of information that the Panel could use for this.

- **Model of USE in the NEM |** The Panel could model USE events in the NEM for the RSSR's review period (e.g. 2028-32 for the 2026 RSSR). The model would be broadly similar to this Review's model but with a closer time horizon. The modelling results could be used to estimate characteristics such as the fraction of USE events that occur in winter as opposed to summer, and the fraction of USE events that occur at peak times of day. This Review has noted a shift towards winter reliability risk so we expect the weighting of the winter VCR would increase with time.
- **Rotational load shedding methods |** As noted in section 7.2.2, individual customers should only experience a short outage as a result of a reliability-driven USE event due to rotational load shedding. The Panel could use information on how different jurisdictions carry out load shedding to better understand the types of outages that customers would experience due to a reliability shortfall. The duration of an interruption during rotational load shedding and information about any exempt loads would be of particular interest and could guide the selection of sectoral VCR values.

The Panel could consider accounting for repeated outages in a short time frame (i.e. the case of multiple rotational load shedding), but given that such large USE events are rare, this is likely to be unnecessary and impractical.

In making this process improvement recommendation, the Panel reiterates its view outlined in the Directions Paper that there are two limbs to reliability metrics, including:

1. types of possible reliability scenarios, and
2. how much consumers are willing to pay to avoid those types of events.

The Panel considers that the question that is directly relevant to the intended scope of the current Review is the first limb referring to the types of possible reliability scenarios as the NEM transitions.

The findings from the VCR will inform, as part of the future RSSR, a more specific question on how much consumers value reliability in the context of these events. The Panel notes that extreme but rare outages (including repeated outages) are outside the scope of the VCR. As part of this Review, the Panel has worked with the AER to include two additional questions in the 2024 VCR pilot survey, addressing some characteristics of outages that may be associated with reliability shortfalls. The 2024 VCR Methodology Draft Determination also notes that the AER will carry out a separate workstream to explore the impact of prolonged and/or widespread outages on energy consumers.[54]Future work may be needed to further explore how customers value severe, widespread, very rare, or repeated outages.

3.3 Further supporting reliability during the transition

The Panel recognises that as the market transitions to net zero, factors outside of the market frameworks can increasingly impact reliability in the short term. The Panel's modelling results suggest that the extent of extreme USE events is likely to remain limited. However, there are a several other pathways available to support reliability as the NEM transitions if policymakers wish to mitigate these rare events.

54 AER, Values of Customer Reliability Methodology Draft determination, p. 2-3, found here.

3.3.1 There are reforms underway to support reliability in the NEM

Over the past several years, there have been different reforms introduced in the NEM to manage the increasing pressure on reliability:

- The Commonwealth's CIS and jurisdictional schemes will support investment in various generation and demand side projects.
- The AEMC has recently made a rule to progressively increase the MPC and CPT from 1 July 2025 to 30 June 2028 to support investment in the generation, demand response and storage projects needed to maintain reliability as the NEM transitions. This change will work to complement the CIS and jurisdictional schemes.
- In 2023, the AEMC recommended changes to the regulatory framework, promoting timely and efficient investment in and delivery of significant transmission. It is now working on several rule changes to implement those recommendations.
- The Retailer Reliability Obligation, introduced in 2019, is a measure that places obligations on liable entities to contract with firmed generation. This provides an incentive for investment in these assets and on-market generators to improve market liquidity in periods where reliability gaps are identified in regions. The RRO is currently triggered by the Interim Reliability Measure (IRM).[55]
- The Interim Reliability Reserve, an out-of-market capacity reserve, allows AEMO to enter contracts to keep unserved energy below the IRM.
- The Energy Security Board recommended that jurisdictions concerned about reliability in the short term could leverage AEMO's short-notice RERT panel. This mechanism provides AEMO access to out-of-market backup resources that can be called upon if reliability issues arise within the short notice timeframe (defined as between three hours and seven days' notice of a projected reserve shortfall).
- Energy Ministers are developing an orderly exit mechanism to allow thermal plants to act as a strategic reserve if there are reliability risks that will result from their retirement.

3.3.2 The ISP outlines substantial investment to maintain a reliable system

As the NEM undergoes its significant transformation, the market needs substantial new investment in generation, storage and transmission infrastructure, alongside the retirement of aging thermal generation. The 2024 Draft ISP outlines that to maintain a secure and reliable system, by 2050, the NEM will need:

- A seven-fold increase in grid-scale wind and solar projects
- 57 GW of storage capacity
- 10,000 km of new transmission lines and upgrades to existing networks.

3.3.3 We need tools outside of the reliability framework to deliver the ISP

Market incentives are the foundation of the current NEM reliability framework. The reliability standard is a central feature of the NEM because it establishes the market price settings that support the effective operation of and investment in the NEM. Historically, price signals have provided enough information for generation and demand-side resources to be built and dispatched. However, increasingly non-price and out-of-market drivers are creating barriers to the timely delivery of generation and transmission investment and infrastructure, even if the price signals are sufficient, influencing AEMO's ability to operate the system.

55 The Interim Reliability Measure is set at 0.0006 per cent USE in any region in any year.

3.3.4 New entry is being challenged by supply chain, workforce and transmission constraints

The Panel recognises that managing high-impact, low-probability events is becoming increasingly difficult as conventional reliability tools, such as directing scheduled generation and major load and customer demand management to facilitate resource exchange, are likely less effective in a high VRE world. In particular, project delays that result in deviations from the anticipated market development forecast in the ISP increase the challenges and difficulties for AEMO.

There are several programs underway to support timely project delivery. Governments have a role to play in working to address the non-market barriers to the delivery of the ISP, including:

- Supply chain constraints
- Workforce training and upskilling
- Planning and environmental approvals
- Social licence.

3.3.5 The uncertainty of thermal plant exits is impacting market signals

The unpriced cost of carbon emissions in the electricity sector means there is no strong in-market signal for generator exit to support emissions objectives.

Further, the exit of thermal capacity will likely result in periods of high and volatile prices between thermal plant retirements and new capacity entering the market. Governments and industry have identified that a critical requirement in the transition is to ensure new assets are in place before old assets retire. To achieve this, governments may need to introduce mechanisms to support asset entry and the predictable exit of aging thermal generation.

3.3.6 Some technologies may not be commercially viable without government support

Some events may be too rare to materially impact USE and too rare to be addressed via the market price settings (around 1 in 100 years). Or, even if the price settings are sufficient, other factors impact the timing, scale or ability to deliver the investment.

New entrants may need additional help beyond market frameworks and signals to build a business case due to variable weather-dependent revenues and reliance on extreme pricing events. Some assets outlined in the ISP, such as long-duration storage or pumped hydro, may only be commercially viable within ISP timeframes with some level of government support.

3.3.7 Support mechanisms can increase reliability, but these come at a cost

No power system can be 100 per cent reliable. Rare or unforeseen and unplanned-for reliability events can occur. Building a system with sufficient capacity to avoid any outages is prohibitively expensive, as it would involve significant over-capitalisation in power system assets, leading to much higher power prices than consumers would be willing to pay.

Governments should continue to consider what further tools may be needed outside the market price settings to provide the capacity to manage these rare events, mainly if these are above customer willingness to pay. This is particularly the case if Government policies and/or programs are not aligned with the direction or timing of the ISP. Any increase in reliability comes at a cost, which would need to be funded by the market (that is consumers), taxpayers or a combination of both.

A range of support mechanisms have already been introduced in the NEM or discussed in the academic literature that governments could consider. Each is a blunt instrument and comes with trade-offs. The scope of this Review does not cover recommendations for support mechanisms outside the reliability standard, but the Panel notes that examples of such instruments could include:

- Strategic reserve (in or out of market)
- Capital grants or concessional debt financing to lower the funding needs or cost of capital for the project
- Swaption style arrangements like the NSW LTESA or CIS for long-duration storage or pumped hydro (including CIS auctions for long-duration storage or pumped hydro)
- Cap contracts or contracts for difference
- Reserve payments like the Hornsdale Power Reserve
- Government build-to-own or joint venture.

4 The draft recommendation will contribute to the energy objectives

As outlined in the Panel's Issues Paper published in March 2023, the Panel has been guided by the NEO in making recommendations in this Review.

The National Electricity Objective (NEO) as stated in the National Electricity Law (NEL) is:[56]

> to promote efficient investment in, and efficient operation and use of, electricity services for the long term interests of consumers of electricity with respect to -
>
> a. price, quality, safety, reliability and security of supply of electricity; and
>
> b. the reliability, safety and security of the national electricity system; and
>
> c. the achievement of targets set by a participating jurisdiction -
>
> i. for reducing Australia's greenhouse gas emissions; or
>
> ii. that are likely to contribute to reducing Australia's greenhouse gas emissions.

4.1 How the Panel applied the NEO to its draft recommendation

In the Issues Paper, the Panel outlined that it would consider its recommendations against the general assessment principles in the Review of the Reliability Standard and Settings Guidelines of:[57]

- allowing efficient price signals while managing price risk
- delivering a level of reliability consistent with the value placed on that reliability by customers
- providing a predictable and flexible regulatory framework.

Further, it outlined that a change to the form should deliver a standard that:

- better reflects customer expectations and the value they place on system reliability
- is clear, simple and implementable
- captures increasing diversity and complexity in the characterisation of reliability risk
- achieves reliability at low cost
- is agnostic to future market design and technology changes
- delivers a common reliability experience across all NEM regions.

4.2 Maintaining the form of the reliability standard contributes to the NEO

Having regard to issues raised in the terms of reference, the Panel is satisfied that the draft recommendation will, or is likely to, contribute to the achievement of the NEO. The Panel has determined that its draft recommendation to maintain the current form with various process improvements will likely deliver the greatest net benefit for the long-term interests of consumers.

The draft recommendation is consistent with the assessment criteria and maintains the reliability standard that strikes a desired balance between affordability and reliability in the power system. The Panel's draft recommendation recognises that the existing reliability standard remains fit for purpose and appropriate in light of the changing nature of reliability risk in the evolving NEM.

56 National Electricity (South Australia) Act 1996, found here.

57 Reliability Panel, 'Review of the Reliability Standard and Settings Guidelines, Final Guidelines, 2021, found here.

The Panel's modelling work has identified that the existing form can adequately address the vast majority of USE outcomes seen in the Panel's simulation modelling. While there is a small proportion of rare USE events, which may not be material to reliability forecasts under the existing form, their probability or likelihood of occurrence approaches 1 in 100 years or more. This means that these events are too rare to be appropriate for any reliability standard, noting that the reliability standard is not intended to achieve absolute reliability and instead seeks to balance reliability and affordability.

The market price settings to which the reliability standard is linked could not effectively provide sufficient incentives for viable projects to prevent the occurrence of those rare events without an excessive burden on consumers, regardless of the form of the reliability standard.

Overall, the Panel has formed a view that maintaining the current form of reliability standard is a recommendation that will likely deliver the greatest net benefit for the long-term interests of consumers consistent with the NEO.

4.2.1 The Panel has considered the new emissions component of the NEO

Since the publication of the Issues Paper, the list of matters forming part of the long-term interests of energy consumers in the NEO was updated to include the achievement of targets set by a participating jurisdiction for reducing or that are likely to contribute to reducing Australia's greenhouse gas emissions. Accordingly, the Panel has considered whether the new emissions component of the NEO is a relevant consideration for this Review.

The Panel recognises that the NEM is going through unprecedented change as ageing coal-fired generation retires and is being replaced by renewable energy and storage during the transition towards net zero. In its 2024 Draft ISP, AEMO's forecast that about 90 per cent of coal-fired generation will exit the market by 2034-35, while grid-scale wind and solar will almost triple by 2030, and storage capacity will increase by over seven-fold by 2050.

Ageing coal-fired generation is already exiting the market faster than anticipated, which is placing increasing pressure on the electricity system. Significant reliability risks arose in June 2022 when the market was suspended, due in part to the increasing unreliability of ageing coal-fired generation. Further reliability risks following the closure of coal-fired generation have also been forecast by AEMO in its 2023 ESOO. The ESOO finds that reliability risks are increasing with unserved energy in NSW and Victoria now forecast to be above the reliability standard during the period relevant to this review.

These trends demonstrate that emissions policies are part of the broader context for this review. However, the Panel has determined that there are no direct or indirect emissions impacts arising from the Panel's recommendations for the Review. While this Review is intended to consider the form of the reliability standard and APC, the reliability standard and APC are linked to market price settings that are designed to drive efficient and timely investments in the context of the transition to a net zero energy system.

The Panel is recommending retaining the current form of the reliability standard and APC. Given this, it is unlikely that the Panel's recommendations for this Review will have any direct or indirect emissions impacts that are assessable under the NEO.

While the Panel considers there are no emissions impacts arising from this review, it will re-consider them as part of future RSSR processes. The future RSSR may identify and assess the impact of emissions and emission reduction activities arising from setting the level of the reliability standard and market price settings, subject to the Panel's recommendations.

4.3 If necessary, the Panel can reassess whether the form remains suitable in the future

The Panel's draft recommendation is to retain the existing form of the reliability standard because it is still fit for purpose in the changing energy system. This conclusion is based on our analysis of the reliability framework, modelling results, stakeholder feedback, and available information on how customers value reliability.

However, the Panel also acknowledges that the outlook for the future NEM could change significantly and new information is constantly becoming available. That said, the Panel intends to monitor the changing risk profile and any changes to market design or consumer preferences in future RSSRs, and consider whether the form may need to be reviewed again.

It is worth noting some of the new information that is expected to be available in the future includes:

- Regular ESOO and ISP updates which include new costs and constraints
- The 2024 VCR review and subsequent VCR reviews
- Changing weather and climate data
- Performance of VRE resources in a changing climate
- Evolving customer demand and new behind-the-meter technologies
- Generator entry and exit
- Transmission developments
- Potential future changes to the market design and State and Federal incentives
- More information about the CIS.

5 The Panel has undertaken further modelling work based on stakeholder feedback

Since the publication of the Directions Paper, the Panel has undertaken further modelling work based on stakeholder feedback and as foreshadowed in the Directions Paper.

This further work aimed to improve understanding of the changing reliability risk profile and assess whether the current form of the standard remains appropriate. A short summary of the modelling methodology is provided here; however, a much larger section in Appendix B describes the methodology in greater detail.

5.1 The Panel has modelled the changing nature of reliability risks in the evolving NEM based on a simulation, but this is not a forecast

The Panel has undertaken the modelling based on a simulation of a virtual future power system that is deliberately constructed to create insights about its unserved energy profile.

These results, however, are not a projection or forecast of whether the future power system will meet the reliability standard. Instead, the simulations examine the composition or distribution of different types of reliability events in a possible future NEM to better understand the nature of the reliability events the system could face under stress.

The base model used for the Directions Paper modelling was the AEMC USE Simulation Model (AUSM) which was designed to generate a sufficient number and range of USE events to study the changing nature of reliability. The AUSM uses the generator and network construction schedule and demand forecasts from the 2022 final ISP for the period FY2028 to FY2043. Like the ISP, AUSM assumes that the power system is constructed to meet the emissions trajectory and does not consider delays to critical infrastructure or changes to actual project location or timing.

A key purpose of this model was to analyse the distribution of USE as the NEM transitioned from a lower VRE-penetration system to a higher VRE-penetration system. This modelling produced a number of insights described in the Directions Paper, including that if USE events occur they may be deeper and longer but less frequent, and also that weather will become increasingly more important in USE outcomes.

To confirm and extend these insights, changes and additions have been made to the modelling approach which are described below and in more detail in Appendix B. These changes and additions have been made on the basis of stakeholder feedback to make the findings more robust, and also to suit the purpose of this second phase of modelling which is to understand whether the current form of the reliability standard is fit for purpose given the changing reliability profile.

5.2 Changes and additions have been made to improve confidence in the modelling results

Modelling for this report involved making changes and additions to the base AUSM model used in the Directions Paper. These are described briefly below, and in more detail in Appendix A.

- **An alternative approach to removing capacity |** A slightly different approach for generating enough USE to study was made in this revision. The focus in this stage of modelling was to target the removal of capacity which is more directly equivalent to USE. As such, the supply reductions to the base model comprised firstly removing the forecast demand side participation units, some new or existing Open Cycle Gas Turbine (OCGT) units and, finally, a small number of VRE units.
- **Validating insights using a full capacity model |** Stakeholders raised concerns that the AUSM produced potentially unrealistic results due to the removal of capacity to generate USE. In response to this, the Panel has included a model that uses the full capacity forecast in the 2022 ISP to validate the findings and provide an alternative picture of reliability risk based on a more reliable system.
- **A much larger set of weather reference years |** The Directions Paper found that weather was a stronger driver of USE outcomes than simulated forced outages. To capture a greater range of weather conditions in the model, the 83 weather years provided by Griffith University was fully incorporated into the model. Consequently, the modelling considered each of these reference years rather than any forced outage samples. Demand traces were also generated to match each of the 83 historical weather reference years using machine learning models based on historical temperature, weather data and a variety of other inputs.
- **Applying a univariate approach to understand the changing risk |** An additional approach for studying the impacts of a changing VRE mix was studied in this phase of modelling. Instead of analysing the changing USE outcomes across financial years, we fixed the system at a single financial year and only changed the VRE penetration. This was done so that the impact of VRE penetration could be isolated from other variables such as the changing demand profile.
- **Resampling results using statistical methods |** To confirm whether the current form of the standard is fit for purpose given the changing nature of reliability, the probability of USE also warranted investigation. To generate a more robust and accurate picture of the likelihood of USE events, the USE results from a 94 reference year set (83 Griffith reference years and 11 AEMO reference years) were resampled using a Markov chain Monte Carlo approach.
- **Generating synthetic weather data to understand likelihood of dark doldrums |** Professor John Boland from the University of South Australia was commissioned to apply statistical methods to the 83 weather reference years to understand the probability of dark doldrums. This analysis was used to understand in more detail what the absolute risks of dark doldrum periods are, and to better understand the relationship between dark doldrums and USE outcomes.

Note that the AUSM is reconfigured to generate specific models to target different findings. Wherever analysis or figures are presented, we have specified the model from which the results were drawn. Details of each of these models and any other assumptions or methodology considerations can be found in Appendix B.

6 Additional evidence to confirm the changing characteristics of the reliability risk

The Panel's modelling completed for the Directions Paper generated four key insights about changing reliability risk in the NEM. The Panel has explored some of those findings further to gather additional evidence, improve confidence in the key insights, and better understand some of the nuances. This chapter outlines additional evidence to demonstrate the changing characteristics of the reliability risk as outlined in the Panel's Directions Paper. Note that these findings relate to the period of FY2028 - FY2043, with a particular focus on the later half of this horizon. As such, the Panel expects the trends described below to appear gradually as the NEM transitions to a higher VRE-penetration system, rather than appearing suddenly in the near future.

6.1 Given a constant level of reliability, USE events may become longer and deeper but less frequent

Consistent with the findings described in the Directions Paper, the modelling suggests that as the penetration of VRE increases during the transition, if USE events occur, they may be longer and deeper but less frequent.

The analysis in the Directions Paper showed that for a given level of reliability, both the typical and largest USE events that might be expected were longer and deeper as the VRE penetration increased. This does not mean that reliability risk or the overall amount of USE would increase in the future NEM, but rather that the same amount of USE might be concentrated on a smaller number of events.

Another finding in the Directions Paper was that USE events may be spreading out across the day rather than appearing only in the evening peak. This finding is not discussed further in this paper as it is a direct result of USE events becoming longer and being driven more by weather, and may be less relevant to the form of the reliability standard.

The modelling has confirmed that as the NEM transitions, USE events will still primarily begin after sun-down during the evening peak. USE in these events is correlated with low wind output and may extend into the following daylight hours until solar resources become available. Typically, even under low solar conditions, as long as there is sufficient wind resource, the grid and CER battery storages are well charged (note that CER behaviour is based on ISP assumptions).

However, when a low solar day is followed by low wind in the evening/night, the battery storage deplete quickly without wind support. The outage will continue until either demand drops to a point where the supply-demand is restored or adequate solar or wind resources are available. The spreading out of USE events across the day, therefore, is primarily a result of the limited number of VRE resources which in turn extends the length of the events.

Additional details on how USE events may become longer and deeper can be found in the Directions Paper and in Appendix section C.1.

6.1.1 Additional findings

In subsequent modelling, the Panel used an alternative approach to study the impact of increasing VRE-penetration on USE outcomes. This approach involved keeping all other variables like demand fixed and modifying VRE-penetration only.[58]

58 See Appendix B for details on how the model for this phase was developed.

The results of this analysis support the conclusion that the depth and duration of USE events may increase in the transition to a higher VRE penetration system and impact depth more than duration.

This finding is outlined in more detail in Appendix section C.1.

6.2 As the NEM transitions, reliability risks may shift from mainly being in summer to winter

The modelling results suggest that winter may become a period of more significant reliability risk than summer in the future NEM. Several factors including changing demand patterns combined with increased VRE penetration may be driving this shift.

The modelling for the Directions Paper, which studied USE longitudinally between FY2028 and FY2043, showed a clear shift of reliability risk from summer to winter.

The analysis for the Directions Paper showed that while the penetration of VRE in each region differed, the broad trend was consistent across all regions, that is, as the NEM transitions, the proportion of USE events in winter compared to other periods also increases.

Additional details on how reliability risk may be moving to winter periods can be found in the Directions Paper and in Appendix section C.2.

6.2.1 Additional findings

The finding that as the NEM transitions, USE will move from being mainly in summer to mainly in winter was also clear in the modelling for this Draft Report. However, additional analysis was done using an alternative approach which revealed that the driver of this change may not primarily be VRE-penetration, but by other factors such as demand.

To isolate the potential impact of varying VRE penetration on the seasons in which reliability shortfalls occur, the latest analysis examined the variation of VRE penetration against a constant operational demand for in a given future financial year (FY2040). Results showed that varying utility scale VRE penetration did not materially impact the season in which USE events occurred. Instead, the analysis has shown that the shift from summer to winter is closely linked to winter operational demand exceeding summer but that other factors such as the evolution of both native demand and rooftop photovoltaics (PV), contribute to this outcome.

This finding is outlined in more detail in Appendix section C.2.

6.3 USE events may be driven increasingly by weather

The modelling performed to date shows that, as the NEM transitions, USE events may be increasingly driven by weather patterns compared to conventional plant failures.

Furthermore, in a future high VRE power system, weather-driven USE events may be more extreme. Further modelling incorporating a wider range of weather reference years has also produced a wider range of USE outcomes and revealed that whilst dark doldrums are a significant driver of USE outcomes, not all dark doldrums lead to USE.

The Panel's modelling work undertaken for the Directions Paper found:

- as the NEM transitions, weather patterns may have a far greater impact on the mean time between events, depth and duration of USE events than thermal plant outages (modelled by forced outage samples) or different PoE demand levels
- 'dark doldrums' or periods of very low wind and solar availability may significantly impact the depth and duration of USE events
- as the proportion of VRE in the system increases, the impact of these dark doldrums or periods of low wind and solar availability on the depth and duration of USE events may also increase.

Additional details on how reliability risk may be driven increasingly by weather can be found in the Directions Paper and in Appendix section C.3.

6.3.1 Additional findings

Stakeholders broadly agreed with this changing driver of USE outcomes. Further, modelling for the Draft Report extended this analysis to understand in more detail the impact of weather conditions on USE outcomes. This involved including more weather reference years in the model and examining in more detail the relationship between dark doldrums of different sizes and USE outcomes. The results of this analysis indicate that:

- including more weather reference years leads to a greater range of USE outcomes
- dark doldrums are a significant contributor to USE, although not all dark doldrum periods produce USE.

Further, the Panel commissioned Professor John Boland from the University of South Australia to conduct statistical analysis on the likelihood of dark doldrums of different sizes and lengths. The additional findings related to weather are outlined in more detail in Appendix section C.3, and the results from John Boland's study are outlined in Appendix E.

7 Despite the changing reliability risk, the current form of the reliability standard remains fit for purpose

7.1 The current form of the reliability standard can adequately address the vast majority of USE outcomes seen in the modelling

The current form is based on expected unserved energy value, which is a combination of the likelihood and size of events. Under this form, all events generated through reliability modelling influence the overall calculated expected value.

However, small or rare events may have such a small value that they will not make a material difference to the calculation of the overall level of reliability. There was a concern prior to this Review that future USE events seen in a NEM with higher VRE penetration may not be adequately addressed by the current expected value form of the reliability standard.

We have quantitatively evaluated the modelling results to understand whether the USE events we see are large enough or likely enough to be material to reliability forecasts under the existing form of the standard based on the expected value.

The results from this analysis indicate that the vast majority of USE outcomes adequately influence the level of reliability calculated under the current standard:

- Most USE results seen in the modelling are either large enough, or likely enough to be material to reliability forecasts under the current standard.
- There are a small proportion of rare USE events which may not be material to reliability forecasts under the current form. However, their probability and likelihood of occurrence approaches 1 in 100 or more years. Therefore, these events are too rare to be appropriate for any reliability standard, noting that the reliability standard is not intended to achieve absolute reliability and instead seeks to balance reliability and affordability.

To test whether the modelled USE results are adequately addressed by the current expected value form, the Panel used statistical re-sampling to extend the modelled results and overlaid them with several assumptions on materiality.

Results from the calibrated reduced-capacity model using 94 reference years were re-sampled using a Markov chain Monte Carlo model to generate 100,000 synthetic years.[59] This produced a more robust distribution of USE outcomes than the 94 year sample noting that 94 data points is relatively few from which to infer probabilities. More information about this methodology can be found in Appendix B.

These results were then transformed into a cumulative probability curve (utilising probability and size of the USE event) that could be compared to an expected value curve that represents 5 per cent of the current level of the standard. The level of 5 per cent was considered appropriate for defining material outcomes consistent with a typical cutoff for statistical significance.

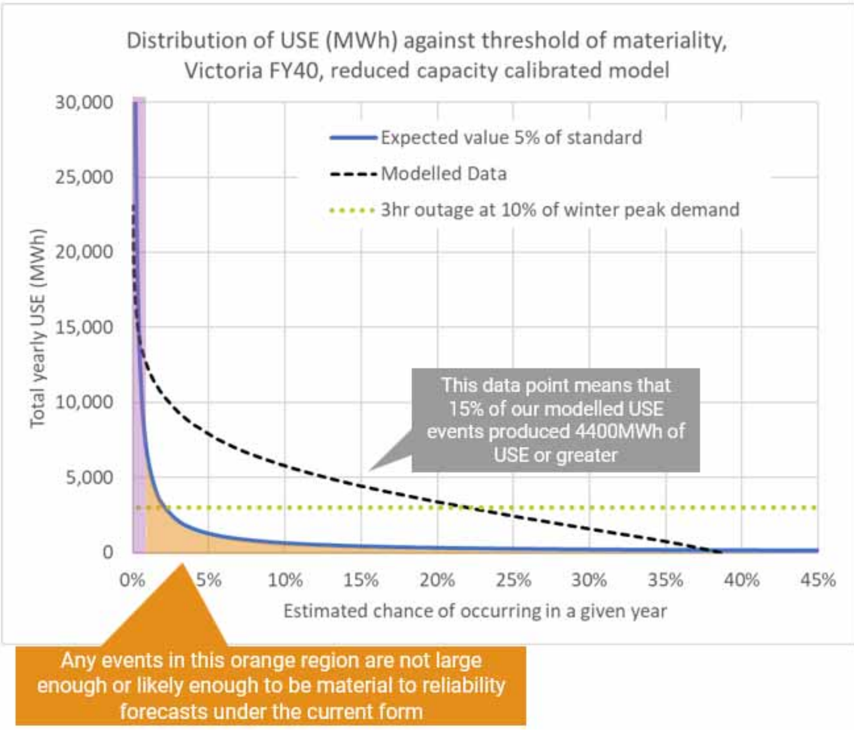
59 94 years comprising 83 reference years from the Griffith data and 11 reference years from AEMO.

The vast majority of USE events in the distribution were found to sit above this 5 per cent line, meaning that the vast majority of USE outcomes seen in the modelling could have a probability and consequence large enough to change the value determined by the current average expected value form. The Panel notes that this conclusion broadly holds under a range of conditions, including:

- using the full capacity ISP modelling results
- in all regions
- for all years modelled, FY2035 to FY2040
- when the 5 per cent line is increased to a more conservative 10 per cent of the current standard
- when using the non-resampled 94 data points.

These results are best described graphically, for example as in Figure 7.1 below.

Figure 7.1: The vast majority of events produced in the modelling are material to reliability forecasts under the current standard



This figure demonstrates whether the modelled USE events would likely make a difference in breaching the reliability standard under its current form and level.

The x-axis is the likelihood of a USE outcome in the modelled year, and the y-axis is the total MWh of USE in a modelled year. These results are for Victoria in FY2040 using the calibrated reduced-capacity model for 100,000 Monte Carlo simulated reference years.

- The solid blue line represents a USE outcome which has an expected value of 5 per cent of the current standard, which we consider material to reliability forecasts under the current form.

- The orange shaded area represents the region in which USE outcomes would have such a small, expected value that they would not be material to reliability forecasts under the current standard.
- The purple area represents USE outcomes that occur at a frequency of 1 in every 100 years or greater – these are considered too rare to be effectively addressed by the pricing mechanisms that drive the investment needed to meet the standard, and so are not considered relevant to determining the form of the standard. Section 3.2.3 provides more detail on this concept.
- The black dashed line is a cumulative density curve representing USE outcomes for 100,000 Monte Carlo simulations based on 94 modelled reference years.
- The green dotted line is just a reference line that represents a 3-hour outage taking out 10 per cent of load for the duration of the outage, assuming that demand is the average winter load at 6pm.

Wherever the black dotted line crosses below the blue line indicates a set of USE outcomes produced by the modelling which may not make a material difference to the calculation of reliability under the current form.

USE events that occur in the orange section are small enough or of low enough probability that they do not make a material contribution to the average USE on which the standard is calculated. These are, to some degree, similar to the expected large number of simulations in which there is no USE. The difference being that when they occur in reality they warrant operator intervention and have the potential to result in interruptions to consumers.

Whilst the entire distribution is shown on this chart, it is also only the left hand side of the chart which is of concern, as these represent the larger and rarer events that constitute the 'tail risk' that stakeholders are concerned about.

The black dashed line does not lie within the orange area, except for a very small section on the right side representing very small USE events. This means that the vast majority of USE events seen in the modelling are large enough or likely enough to make a material difference to the calculation of reliability, and so the current form of the standard is adequate to address these events.

There is a small region above the black dashed line and below the blue line, which lies in the purple region. As can be seen from the figure, the probability of these events is small, approximately 1 per cent. As such, they would be expected to occur, at most, once every 100 years.

As noted earlier in this section, the reliability standard is not intended to achieve absolute reliability and thus is not expected to capture such extremely rare events. The market price settings to which the reliability standard is linked could not effectively provide sufficient incentives for viable projects to their occurrence without an excessive burden on consumers, regardless of the form of the reliability standard.[60]

The broad conclusion holds across all the modelling results. Consistent with all reliability assessments, the analysis also revealed a small number of modelled results that were likely but small. These events are captured by the standard and do not on their own warrant a change in the form.

60 The modelling results for all mainland regions of the NEM, for financial years FY2035-2040 are detailed in Appendix D, and include a number of sensitivities using the full capacity model and using less conservative assumptions regarding materiality.

7.2 Even though the risk profile is changing, the modelling has not produced any extremely severe events

7.2.1 There are no extreme multi-day events

In response to the Issues Paper, some stakeholders expressed concerns that the changing reliability risk might lead to very extreme tail USE events that could potentially lead to multiple day-long outages. Our results indicate that even under a very large range of weather conditions, the modelling did not produce any multi-day events.

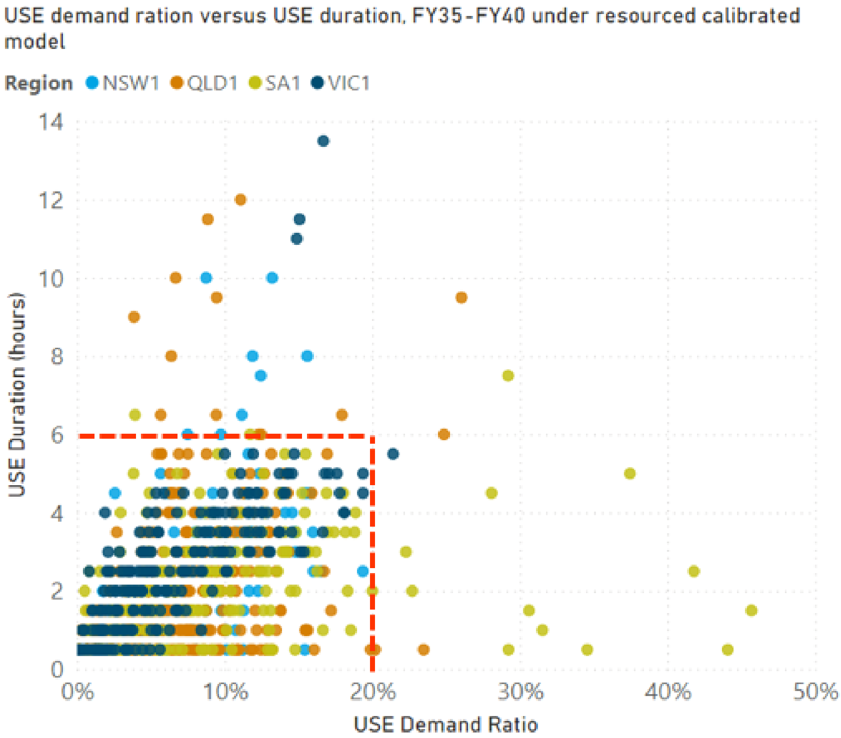
While the modelling results indicate a shift to potentially longer and deeper events, the Panel's modelling has not found any events of more than one full day of unserved energy even under a very large range of weather conditions.

The depth, duration and frequency of USE events were studied using a full capacity model based on the 2022 final ISP and a calibrated reduced-capacity model derived from the 2022 final ISP. Both models tested FY2035-FY2040, and the longest event produced is less than 24 hours in total USE duration, and the overall distribution did not produce a significant number of outlier events.

In these results we 'clustered' USE events separated by a period less than 16 hours – see Appendix B for more details on this methodology. Whilst the results hold true without this clustering, the Panel considers that it is more reasonable to group events that are separated by a short period. Consequently, a single USE event may stretch out across several days, and comprise 2 or more discrete, short, events. Considering the standard refers to annual expected USE we focused this analysis on the total duration of USE (termed USE duration in this report) within this period rather than the absolute chronological length of the clustered event.

Figure 7.2 shows the USE duration and the USE demand ratio for the calibrated reduced-capacity model for financial years 2035-2040 inclusive. Each dot represents a USE event, where an event is defined as a group of intervals with USE that are 16 hours apart or less. (See Appendix B for more information on the 16 hour clustering methodology.)

Figure 7.2: The longest events in the calibrated reduced-capacity model are 14 hours in USE duration or less, and whilst there are some deep events they are short in duration



This figure shows that 96 per cent of events lie in the lower left quadrant, less than 6 hours of USE and less than 20 per cent of demand not being met during the period. While there are a handful of very deep events, these are all isolated in South Australia and are generally very short in duration (the majority are three hours or less in USE duration).

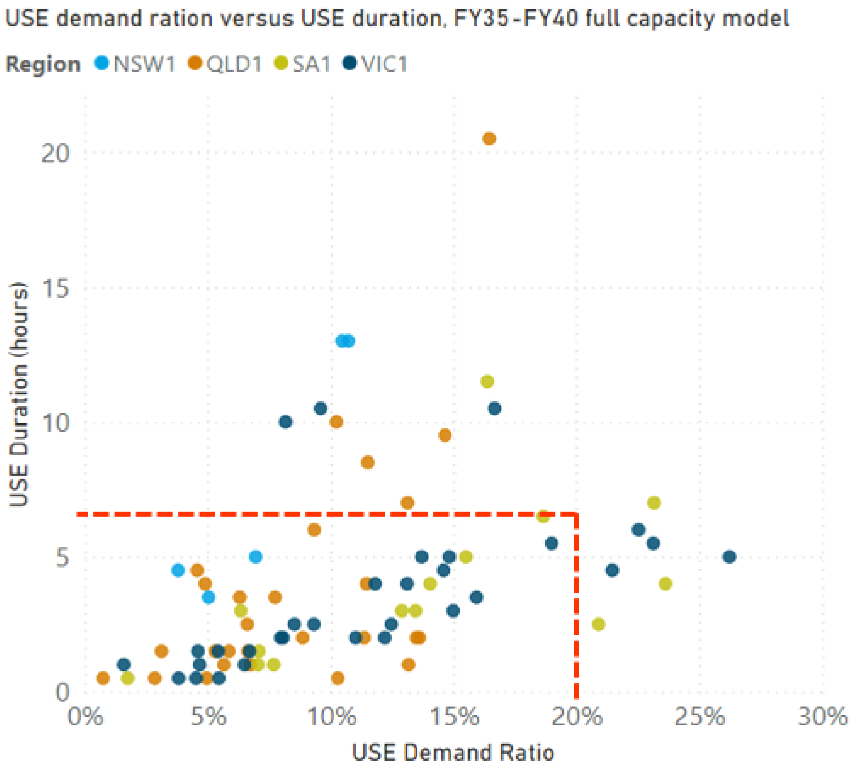
The finding that South Australia produces particularly deep USE events compared to other regions is also corroborated in AEMO's 2023 ESOO. This document shows that the reliability outcome in South Australia for FY 2024 includes many USE outcomes between 40 per cent to 50 per cent of average regional demand, compared to other regions which only show a single USE outcome above 25 per cent of average regional demand.[61] It should be noted that AEMO's assessment for FY 2024 includes retirements and delays to PEC which will have increased the USE outcomes.

As also shown in Figure 7.2, the longest event seen in this model had a USE duration less than 14 hours. This addresses stakeholder concerns regarding a high VRE system resulting in an extremely long tail of events lasting more than one day.

61 AEMO, 2023 ESOO, p. 143, Figure 81, found here.

Figure 7.3 below is the equivalent to Figure 7.2 but shows the results from the full capacity model.

Figure 7.3: In the full capacity model the longest event is 21 hours of USE duration, and there are fewer deep events



This model in Figure 7.3 produces regional reliability outcomes well below the current reliability standard (i.e. a more reliable outcome), unlike the calibrated reduced-capacity model that produces outcomes at the reliability standard. Consequently, there are significantly less dots on this chart.

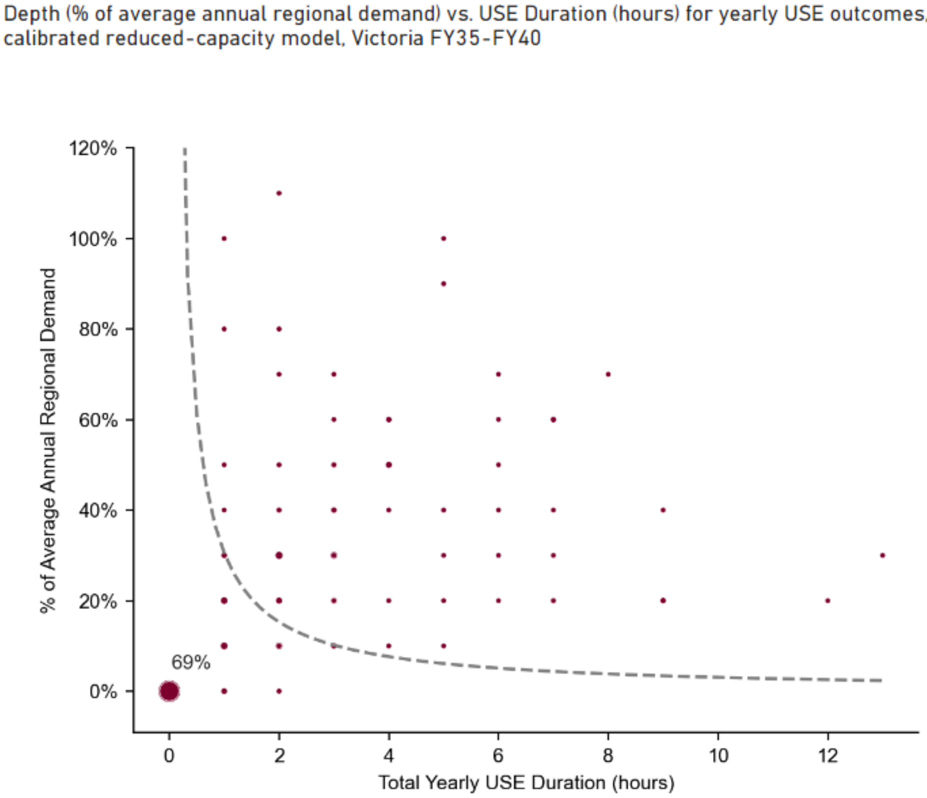
Broadly the conclusions remain the same. Most events generated have a USE duration of less than 6 hours long and remove less than 20 per cent of demand during the event. The full ISP model produced fewer deep events in South Australia, however it did produce a longer 21 hour USE event in Queensland.

These results lend further evidence to our finding that, while the reliability risk may be changing to longer and deeper events, we have not found evidence that such change reflects an extremely long tail of events lasting more than one day.

The Panel has also used a similar visualisation tool to that used by AEMO in the 2023 ESOO to provide a comparable view to their FY2024 forecasts. Figure 7.4 below shows the results of the calibrated reduced-capacity model for Victoria for financial years FY2035 to FY2040. The results in this figure are the combination of the calibrated reduced-capacity model using inputs generated by the AEMC and those for the same period using AEMO ISP demand and VRE resources – as such it covers 83 Griffith years and 11 AEMO ISP reference years.

In Figure 7.4, each dot represents a yearly USE outcome from each of the 94 reference years modelled but do not represent a strict likelihood, as they are based on only 94 reference year outcomes rather than a larger sample of more than 2,000 as produced by AEMO. The x-axis shows the total USE duration of the modelled year, and the y-axis shows the depth of the USE outcome as a percentage of average regional demand. The Panel notes that the AEMO chart and this chart are similar but are based on significantly different financial years. Our modelled results are for a future NEM where the average operational demand is significantly lower because of the significant forecast offset from rooftop PV during the day and CER.

Figure 7.4: Total yearly USE outcomes range for 0 to 14 hours in duration and up to 110 per cent of average annual regional demand



While AEMO's simulated results are based on a larger number of samples in the order of 4000 (made up of 100 outage samples, with three PoE demand profiles and 13 weather reference years), the modelling results presented in this analysis are based on a much larger set of 94 different weather patterns which, in the future, are a stronger driver of USE than stochastic forced outages. We conclude therefore that while our modelling does not include a large number of simulations, the results are still robust and a good representation of USE outcomes.

The dotted line on this chart represents an outcome that is above the level of the reliability standard (0.002 per cent of regional annual load). Events above this line are considered 'large' by the definition that AEMO uses in the 2023 ESOO, which refers to an individual USE event in a single year large enough to breach the standard in that year. These large events can occur under the reliability standard as the system will be reliable for a majority of the time, however some events may be large enough to move the average significantly.

This figure also demonstrates that, while the predominant outcome continues to be no USE, when events do occur they are likely to be large. This chart also demonstrates that there are no particularly large outliers, either in terms of very large depth or duration. The total duration of yearly USE duration is still below one day, and while there are some quite deep events they do not last for a long duration. Note that the y-axis presents depth as the percentage of average regional demand, rather than percentage of demand during the event. As USE events are most likely to occur in the evening peak during periods of high demand, they are likely to be a much smaller percentage of event demand – for example the dot in this chart at two hours and 110 per cent of average annual regional demand is made up of 8 modelled USE events, none of which took out more than 10 per cent of load at the time.

7.2.2 There are very few events that would result in customers experiencing multiple rotational outages

The modelling results were translated into a customer perspective by analysing the number of USE events in which customers would likely experience more than one round of rotational load shedding.

The Panel notes that the customer experience of most USE events is likely to be similar, regardless of the depth and duration of the region-wide event. If a USE event occurs (and emergency measures such as RERT are exhausted), the required operational capacity margin will be achieved by rotational load shedding.

Initiated by a direction from AEMO through the Transmission Network Service Provider (TNSP), Distribution Network Service Providers (DNSPs) turn off some segments of customer load to match the level of curtailment in the AEMO direction. Each customer segment is only disconnected for a limited time before the load shedding 'moves on', to spread the impact of USE across customers.

Although the modelling indicates some risk of long or deep events, individual customers would only experience an outage for this limited time. This finding is consistent with some feedback from stakeholders, including Endgame Economics' submission to the Issues Paper, which noted:[62]

> Consequently, there is not a one-to-one relationship between region-wide USE and the length of individual customer outages. A large 'tail-risk' USE event would generally be rotated around different groups of customers so that the duration of the outage to any individual customer will be a fraction of the duration of the total event across the region.

The Panel acknowledges that very long events might lead to two or more rounds of rotational load shedding. This means that some customers experience repeated short outages within 1-2 days, and that deeper events would mean rotational load shedding moves through customers more quickly.

However, our modelling results suggest that this scenario would be extremely rare, as outlined below. These results were derived from applying reasonable but simplified load shedding assumptions to the PLEXOS modelling results. These are discussed in more detail in the Appendix section B.6.

Table 7.1 below shows the total number of events that occurred in the calibrated reduced-capacity model simulations based on the Griffith data for the six years FY2035-FY2040, and the number of events that led to more than one round of rotational load shedding (termed 'multiple load shedding events').

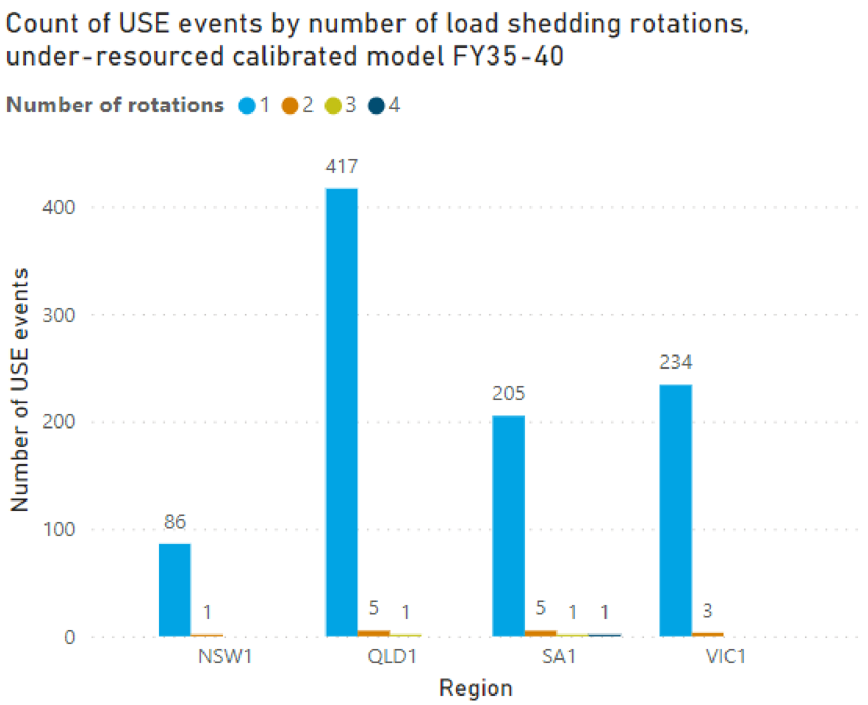
62 AEC submission to Issues Paper, Endgame Economics report, found here.

Table 7.1: Number of USE events in which more than one round of rotational load shedding occurred

Region	NSW	QLD	SA	VIC	NEM
Total events	87	423	212	237	959
Events with more than one round of rotational load shedding	1	6	7	3	17

This is also shown visually in Figure 7.5 below.

Figure 7.5: The vast majority of USE events only lead to one instance of rotational load shedding



Overall, multiple rotational load shedding events were very uncommon, with less than 2 per cent of total USE events leading customers to experience more than one round of load shedding. For the avoidance of doubt, the modelling shows that USE events are rare, and these results indicated that only a very small fraction of those USE events lead to multiple load shedding.

All events with multiple load shedding were in June, which is consistent with our findings that reliability risk is highest in winter. Even though Queensland experiences significantly more summer USE events than other regions, the risk of an event leading to multiple load shedding is still highest in winter. Most of these events only reached the second round of load shedding (i.e. no customer experienced more than two outages). Three events reached the third round of load shedding, and one event reached the fourth round.

We note there are a range of operational challenges with rotational load shedding. In particular, its efficacy in future events that occur during daylight hours may be reduced as a result of rooftop PV allowing some parts of the distribution network to generate more energy than they consume. Our simplified rotational load shedding analysis does not account for such situations.

However, the larger USE events in our simulations typically occur in the evening, night, and other times with a lower solar contribution. In these conditions, rotational load shedding would likely be effective as few, if any, distribution feeders would be net supply.

These results were generated by simulating how load shedding would move through blocks of customer load given the shape of the USE event and load at the time. A few simplifying assumptions were needed which are outlined in Appendix section B.6.

8 The form of the administered price cap

The Panel's draft recommendation is to maintain the current form of the APC and regularly review its level through the Review of Reliability Standard and Settings process.

8.1 The Panel's draft recommendation is to maintain the current form of the APC

The Panel's draft determination is that the current form of the APC should be maintained.

However, the Panel acknowledges the importance of carefully monitoring the level of the APC given the events in June 2022 and the misalignment between the electricity APC and fuel prices.

That is why the Panel intends to review the level of the APC in each RSSR, and make any necessary changes to it as and when required. This will ensure the level continues to efficiently minimise reliance on compensation to recover the short-run marginal cost (SRMC) of the marginal generator or Battery Energy Storage System (BESS).

This Review has considered several options for the form of the APC that might continue to protect participants from extended periods of high prices, while also providing sufficient prices to cover the short-run marginal costs of the generation needed to maintain reliability.

While a total of six options were initially considered in the Issues Paper for the form of the APC, the Directions Paper determined that only the current form of the APC and indexing the APC to CPI would progress for further consideration.

This shortlist was created because of stakeholder feedback, which provided that certainty in the level of the APC was required for market participants to successfully manage risk and provide liquidity in contract markets.

The following sections outline the Panel's reasons for its draft recommendation on the form of the APC.

8.2 Reasons for the Panel's draft recommendation

8.2.1 Stakeholders value a stable APC

No stakeholder submissions to the Issues Paper supported linking the APC to dynamic fuel prices or having a trigger mechanism to increase the APC under certain market conditions. Stakeholders considered that stability in the level of the APC was important for market participants to hedge effectively and for the liquidity of contract markets. Furthermore, stakeholders generally considered that the events of June 2022 would have been prevented by a sufficiently high level of the APC.

As a result of this feedback, the Directions Paper shortlisted only the current form of the APC and indexing the APC to CPI as options and sought feedback from stakeholders on their preferences between these two. Many stakeholders considered that the current form of the APC would be suitable for the future. Further commentary on stakeholder feedback to the form of the APC is provided in section 2.5.

8.2.2 The current form of the APC remains fit for purpose

The current form of the APC continues to be fit for its intended purpose of protecting market participants from extended periods of high prices. It is important that the APC be sufficiently high for generators to recover costs, and thus be able to contribute to reliability during administered pricing periods. To this end, the AEMC recently increased the level of the APC to $600/MWh, which the Panel sees as sufficiently high to serve this purpose.[63] The level of the APC will be regularly reviewed to ensure this remains true.

The Panel could see some merit in indexing the APC to CPI, but found that this was not necessary, provided the level of the APC is high enough to cover the short-run marginal costs of the marginal generator. While the MPC and CPT are indexed, the Panel does not consider this is necessary for the APC. This is because the MPC and CPT are an investment signal in a way that the APC is not. The purpose of the APC is to protect market participants from exposure to extended periods of high prices. To this end, the APC is a meaningful determinant of the hedging strategy of market participants and of contract prices, but it does not meaningfully influence investment in generation assets.

The Panel also notes that while fuel costs do increase over time, they are volatile and not necessarily well represented by the CPI. This has been evident over the last few years, where fuel price inflation has, at times, significantly exceeded CPI.

The Panel does, however, intend to update the level of the APC as and when required during the regular Reliability Standard and Settings Review, to ensure that the short-run costs of the marginal generator or BESS can continue to be recovered under administered pricing conditions. The Panel considers that the four-yearly RSSR is sufficiently frequent to provide any necessary changes to the level of the APC while maintaining a stable environment for market participants.

8.3 The Panel's draft recommendation contributes to the energy objectives

The Panel is satisfied that the draft recommendation to maintain the current form of the APC will contribute to the NEO.

The current form of the APC supports the long-term interests of consumers, particularly with regard to the price component of the NEO. This is because:

- the current form of the APC achieves its purpose by protecting consumers from sustained periods of high prices while retaining the incentive for generators and storage assets to contribute to reliability during administered pricing periods
- the level of the APC can be changed as and when required every four years during the RSSR to ensure the APC continues to contribute to the NEO in the future.

The form of the APC has been considered in the context of the newly included emissions reduction component of the NEO. The Panel has determined that the form of the APC does not have a significant impact on emissions reduction. For further discussion on emissions impacts of the review, see Chapter 4.

63 In the 2022 RSSR process the Panel recommended an APC of $500/MWh (Final Report, found here). In the Rule change to implement the RSSR recommendation the AEMC considered that a higher APC would more effectively cover the cost of investment based on the latest cost projections from AEMO's 2022 IASR and facilitate greater opportunities for grid scale battery developments ('Amendment of the Market Price Cap, Cumulative Price Threshold and Administered Price Cap', Final Determination, found here).

A The reliability framework in the NEM

The current reliability standard, is set at the maximum average forecast unmet energy, or USE for each financial year, as a proportion of the total energy supplied in a region.[64] This appendix provides background on how the reliability standard is operationalised in the NEM and key concepts and terms underpinning it.

A.1 The reliability framework in the NEM is designed to ensure reliability is delivered at a level consumers value

A reliable power system has enough capacity (generation, demand response, interconnector and energy storage capacity) to meet consumer needs. To maintain reliability, a power system needs investment in enough new capacity to meet changing demand patterns and to cover generators as they retire. A reliable supply also needs reserves that allow demand and supply to balance when unexpected changes occur to demand and supply.

No power system can be 100 per cent reliable. Rare or unforeseen events can always occur. Building a system with sufficient capacity to meet all rare or unforeseen events is prohibitively expensive, as it would involve significant over capitalisation in power system assets leading to much higher power prices than consumers would be willing to pay.

Reliability can be increased by encouraging investment in more capacity (either generation or demand side) in the system. However, this comes with additional costs. Since the NEM is a one-sided market, customers do not have direct input on what they are willing to pay for a unit of electricity. Without the market price settings, prices in rare circumstances could rise far beyond what customers are willing to pay. The reliability standard seeks to address this by balancing between the consumer value gained from increasing reliability, versus the costs that this may entail. These trade-offs are implemented through the market price-settings, based on what consumers value in relation to the reliability sought.

The NEM operationalises reliability through a broad reliability framework to deliver investment in reliable and affordable electricity to customers. The framework comprises a mix of reliability settings, market prices and incentives, and operational decisions to keep supply and demand in balance.

Figure A.1 summarises the framework which includes the provision of reliability through spot market incentives supplemented by information and intervention by AEMO as a last resort.

64 NER clause 3.9.3C(a).

Figure A.1: The reliability framework in the NEM

A.2 The reliability standard underpins the market settings

The reliability standard is a central feature of the NEM because it establishes the key market price settings that support the effective operation of and investment in the NEM. It guides the decisions of market participants through the reliability settings, AEMO's forecasts and the RRO.

The reliability standard is an ex-ante standard that indicates the efficient level of supply required to meet demand on a regional basis. It is not a regulatory or performance standard that is 'enforced'. The standard is intended to provide a clear, expression of the economically efficient level of generation and inter-regional transmission capacity sought for the NEM.

The reliability standard also serves as a trigger for intervention in the market, should that be necessary to shore up reliability, through intervention measures such as AEMO's dispatch of emergency reserves. Crucially, the NEM's reliability framework seeks to deliver customers both reliable and affordable electricity. The reliability standard represents a trade-off between the dual objectives of reliability and affordability.

Market incentives are the foundation of the current NEM reliability framework. Prices in the spot and contract markets provide signals for generation and demand-side resources to be built and dispatched and indicate supply and demand balance in each region over time.

The reliability standard influences new entrant investment revenue potential through the level of the market price settings, being the MPC, CPT, APC and MFP.

The market price settings act to limit the prices that generators receive for supplying electricity and therefore the revenue to support investment decisions. The level of the market price settings determined through the RSSR is set at the revenue required by the marginal new generation entrant to maintain the expected level of USE in the NEM to 0.002 per cent.

A.3 The standard also guides AEMO in its role as the system operator

Another key role of the standard is to guide AEMO's decisions in its role as the system operator. AEMO is responsible for operationalising the reliability standard through its forecasting and operational processes. AEMO uses the reliability standard to:

- inform market participants, NSPs and potential investors, over ten-year, two-year and six-day outlooks by publishing forecasts of the adequacy of supply to meet the standard
- initiate intervention action to maintain power system reliability and security where practicable by monitoring demand and generation capacity.

In operational timescales, AEMO issues lack of reserve (LOR) notices to inform the market when supply scarcity conditions exist. AEMO declares LOR conditions when it determines there is a potential shortfall of supply to meet forecast demand. LOR notices are either LOR1, LOR2, and LOR3 in order of increasing supply scarcity.[65]

As effective as information processes and market incentives can be in delivering the desired reliability outcomes from commercial investment decisions, they do not always elicit the outcomes needed. If the market fails to respond to the market price settings and the information AEMO publishes, AEMO may need to intervene in the market. Under the NER, AEMO can intervene in the NEM to ensure a reliable and secure electricity supply.[66]

65 Further information can be obtained from AEMO's Reliability Standard Implementation Guidelines, found here.

66 NER clause 3.20, clause 4.8.9.

B Modelling methodology

B.1 Key definitions and assumptions

The results of the AEMC USE Simulation Model (AUSM) model have been characterised by exploring the key characteristics of USE events: duration, depth, USE (MWh) and mean time between events. Furthermore, a range of assumptions and definitions have been adopted to analyse the modelling results as summarised in this section.

B.1.1 Defining the characteristics of the USE events

The characteristics of the USE events have been defined as follows:

- **Duration** – measured in hours, refers to either:
  - Event duration defined as hours between the start and end of a cluster of events (see section below on clustering), or
  - USE duration defined as the total hours of USE within a cluster of events
- **Depth** – measured in MW, refers to the half-hourly values reported in the modelling.
- **USE** – measured in MWh, refers to the energy lost in the event. For comparison purposes, in most cases in this report USE is reported as one of two ratios:
  - USE event demand ratio is the total USE that occurred in an event divided by the total demand in the event, or
  - USE annual demand ratio which is defined as the total USE that occurred in an event divided by the annual average regional demand.

  It should be noted that these ratios do not have units.
- **Mean time between events** – measured in days, refers to the average number of days between USE events. While this has not been reported to date, it may form part of future analysis on the form of the standard and in the subsequent Reliability Settings and Standards Review.

B.1.2 'Clustering' of USE events

The modelling simulations were based on 30-minute dispatch intervals and therefore a single USE event could span one or many consecutive 30-minute intervals.

As many of the simulated events occur because of weather conditions over many days and, in particular, may relate to solar cycles, consecutive USE events with less than 16 hours between the end of one event and the start of the next have been 'clustered' together. For example, two 6-hour USE events separated by a period of 6 hours will be clustered, giving an event duration of 18 hours but a USE duration of only 12 hours.

The 16-hour buffer window was chosen for the following reasons:

- it captures events that impact morning and evening peaks over consecutive days
- isolated short events are unaffected
- long events remain and a longer buffer window does not extend the longest USE events
- other characteristics of USE do not change materially with a longer buffer window.

B.2 Directions Paper methodology

Table B.1 below gives a high level overview of the methodology used in the Directions Paper.

Table B.1: Overview of the methodology used in the Directions Paper

Modelling step	Description
Development of the AUSM	A simulation model based upon the 2022 ISP was developed to study the nature of USE from FY2028 to FY2043. The scope of this model was to understand the changing nature of unserved energy rather than the absolute reliability risk. This model was built using the 2022 ISP with updates based on the Draft 2023 IASR. Additional changes to this model including doubling of the southerly VIC-NSW transmission limit and adding additional capacity are described in Appendix A to the Directions Paper.
Inclusion of 11 AEMO reference years	As the AUSM model was built using the 2022 ISP (Step Change scenario), 11 different AEMO reference years (2010-2021) were utilised for modelling. These included solar and wind traces at the REZ level which is required for modelling generation capacity. They also included future demand traces as per the 2022 ISP. The Panel notes that the 2023 ESOO published by AEMO includes an additional two reference years.
Removal of capacity to generate USE	As the purpose of the modelling was to study the changing nature of USE, capacity was taken out of the model to ensure there was enough USE generated with which to study. The process for deciding how much capacity to remove was iterative, and focused on ensuring there was enough USE generated in each region. This process is described in Appendix A to the Directions Paper.
Apply calibration process to stress testing model	USE outcomes from this reduced capacity model were calibrated to ensure that the average level of USE across all reference years is at or below the standard. This was done in a post-hoc fashion on the USE output data, where a fixed amount of firm capacity was added back in to each USE event across all reference years such that the overall average was at or below the standard. This process was repeated in the modelling for the Draft Report.

B.3 Latest methodology

This section outlines the latest modelling methodology steps as per the Draft Report, see Table B.2 below. More information on new changes can be found in Section A.4.

Table B.2: Modelling methodology for the Draft Report

Modelling step	Description	New addition?
Inclusion of 11 AEMO reference years	11 different AEMO reference years were utilised (2010-2021) for modelling. These included solar and wind traces at the REZ level which is required for modelling generation capacity. They also included future demand traces as per the 2022 ISP.	No
Inclusion of 83 years of Griffith generation data	Griffith University has provided 83 years (1940-2022) of solar and wind trace generation data. This has been mapped to the nearest REZ for modelling purposes.	Yes
Generation of machine learning demand estimation model	A machine learning model was developed by leveraging the 11 AEMO reference years as training data. The model takes a variety of variables as inputs including time variables, solar/wind power and temperature variables. The model outputs future demand traces (out to 2050) and was applied to the 83 years of Griffith generation data allowing 83 reference years to be available for modelling.	Yes
Synthetic weather generation to understand dark doldrum likelihood	1,000 years of synthetic weather traces were created from a model trained on the 83-year Griffith dataset. The model was built by Professor John Boland from the University of South Australia. The 1,000 synthetic years provide enough data allowing calculations of probabilities of dark doldrums (e.g. X days of low wind/solar).	Yes
Develop reduced capacity PLEXOS model	Here, alterations are made to the PLEXOS models seen in the Directions Paper including the removal of DSP and some OCGT gas capacity to increase USE in an unbiased fashion.	Yes
Univariate approach to changing risk profile	An alternative approach for studying the impacts of a changing VRE mix was studied in this phase of modelling. Instead of analysing the changing USE outcomes across financial years, we fixed the system at a single financial year and modified the VRE capacity in the system. This was done so that the impact of VRE penetration could be isolated rather than being mixed in with a changing demand profile, amongst other things.	Yes
Develop full capacity model	Following stakeholder feedback that the reduced capacity model may be unrealistic in its presentation of USE, the base 2022 ISP model was used with minimal changes. The changes made relate to changing the VRE generation source files for existing VRE generators to the closest REZ trace. This facilitates the inclusion of the additional 83 reference years into the model.	Yes

Modelling step	Description	New addition?
Apply calibration process to stress testing model	USE outcomes from the reduced capacity model have been calibrated to ensure that the average level of USE across all reference years is at or below the standard. This was done in a post-hoc fashion on the USE output data, where a fixed amount of firm capacity was added back in to each USE event across all reference years such that the overall average was at or below the standard.	No
Develop Markov chain Monte Carlo model	To test the current form of the standard, annualized USE outcomes are required. 94 (83+11) reference years / data points are available per financial year and whilst this is an improvement on the 11 reference years used in the Directions Paper, this is a small number of datapoints from which to infer probabilities and hence calculate the current form of the standard. There are however, over 1.5 million data points at the 30-minute level (from the 94 years) that were then leveraged in a Markov chain Monte Carlo process to develop 100,000 synthetic USE year outcomes.	Yes

Note: It should be noted that the machine learning model would not be possible without the AEMO demand traces and the extensive modelling work undertaken by AEMO to produce those traces.

B.4 Changes and additions for the Draft Report

The following section goes into detail describing the new modelling methodology additions post Directions Paper.

B.4.1 Inclusion of 83 years Griffith data for generation

Griffith University provided the AEMC with confidential hourly calculated nominal wind and solar power data for 83 years from 1 January 1940 to 31 December 2022 for around 54 mainland NEM locations.[67] In most cases, at least one and, on some occasions, up to five reference locations were provided for each onshore renewable energy zone (REZ) as defined in the 2022 ISP. For those REZs where hourly traces were available for more than one location, the hourly nominal information was averaged.

A simple average was used because a more sophisticated weighted local capacity average could not be developed given that detailed local development information is not available for the period FY2035-FY2040. As changes to the installed capacity in each REZ for each financial year are applied during the simulations, PLEXOS only requires the nominal resource performance for each of the 83 reference year REZ traces. As is the norm with creating future chronological models from historical data some adjustments were necessary to ensure chronological and calendar consistency such that Mondays from history were still Mondays in the future.

B.4.2 Machine learning demand estimation model

In order for the 83 years of Griffith generation data to be useful for simulating USE events, it is vital that demand traces be estimated that align with the generation data. For example, as solar

67 The data was developed in a 2022 study: J Gilmore, T Nelson and T Nolan, 'Quantifying the risk of renewable energy droughts in Australia's National Electricity Market (NEM) using MERRA-2 weather data', Griffith University Centre for Applied Energy Economics & Policy Research, 2022, found here.

generation decreases, operational demand is expected to increase due to a reduction in the rooftop PV component. Thus it is important that the relationship between generation and demand be maintained throughout the modelling.

It should be noted that AEMO already provides this relationship through their extensive suite of long-term demand modelling where demand traces are estimated out to 2050 and leverage a variety of inputs including but not limited to temperature related variables including CDD (cooling degree days), HDD (heating degree days) and time related variables etc. Eleven reference years were provided by AEMO (published in the 2022 ISP) that include generation input traces (wind/solar power) and corresponding demand traces (operational demand, native demand). The 11 reference years have been fundamental to the modelling in this project and it was observed that there is a great deal of variability in USE outcomes depending on the different reference years. As a result, a collection of machine learning models were built to estimate or approximate demand traces that correspond to the 83 reference years from the Griffith dataset. Details of this model are outlined below.

Model training

The model was trained on the 11 AEMO reference years of data which was supplemented with additional temperature data. Inputs and outputs from the training data are characterised as follows:

- **Inputs:**
    - **Time related variables:** These include time of day, day of week and week of year. Time of day variable was found to perform strongly as it was able to identify peak demand periods.
    - **VRE variables:** These included AEMO solar and wind power traces.
    - **Temperature variables:** These included HDD and CDD for each state. BOM weather stations near to Australian capital cities were leveraged to source this data.
- **Outputs:**
    - operational demand
    - native demand.

As part of the model testing, results indicated that there was value in producing separate models to predict each calendar year of demand (2023-2042 were used for this project), for each state and for each type of demand (operational or native). In total, this means a suite of 80 models were produced.

A random forest regression machine learning model structure was utilised for each of the 80 models. This model structure was chosen due to its ability to model complex non-linear relationships and for its strong validation performance. Furthermore, random forest models leverage ensemble methods which tend to provide stable results that perform well out of sample.

Model validation

Out of sample testing was conducted where 20 per cent of the available data was excluded from the training process. The models were then trained on the 80 per cent of available data before testing on the remaining 20 per cent. When utilising AEMO ISP demand traces as the goal for regression, the estimated demand traces were around 94-96 per cent accurate depending on state and estimation year. Furthermore, more than 90 per cent of the variability observed was explained by the fitted model. These are strong results however it should be noted that this sort of exercise is expected to have much lower error rates than say predicting observed demand. Observed demand, is real observed data that is expected to include a great deal of noise that models will struggle to estimate. Instead, our model is trying to estimate AEMO ISP demand traces which could include less noise than real observed data.

Despite the high overall accuracy rates it is important to test the accuracy under certain homogenous groups to ensure that the model works for different populations, e.g. does the model still work well for certain months of the year or for periods of low wind/solar. In total the model was tested over the following groups:

- Per month of the year
- Per time of day
- For different quantiles of wind power
- For different quantiles of solar power
- For different quantiles of demand.

In general, the model had similar error rates for most of these splits. This is particularly important for groups of low wind or low solar power (dark doldrums) as these are expected drivers of USE events. However, the model was found to have diverging error rates for the different quantiles of demand. Error rates are defined as the percentage difference between actual (AEMO ISP demand traces) and estimated (machine learning) demand traces. The validation data was split into 100 groups of 1 per cent quantiles based on operational demand. The bottom 99 per cent of the data was found to have low error rates of below 6 per cent, but the top 1 per cent was found to have an error rate of 12 per cent.

Whilst not ideal, this reduced performance at the extremes is not uncommon in modelling. Generally, model fitting processes aim to perform well over the full range of possible outcomes and can struggle to estimate edge cases. The model fitting process could be adjusted to prioritise fitting for high demand periods but this would likely reduce the overall accuracy of the model. As such, the model was retained as is and a post model adjustment was applied to address the demand spikes.

It was observed that the underestimation of demand spikes occurred primarily in winter and summer where operational demand spikes are relatively high. As a result, the top 5 demand spikes from the ISP 2013 reference year demand traces were taken for both winter and summer seasons in each calendar year (10 demand spikes per year). The results of the machine learning model were then amended such that the corresponding top five demand spikes for both winter and summer were aligned between the machine learning output and the ISP 2013 reference year.

The machine learning results were stretched so that the demand peaks are aligned and all other demand results are linearly scaled such that the lowest demand from that given day remains unchanged. This method ensures that the demand traces are relatively smooth and that the model is capturing the top 10 demand spikes per calendar year.

B.4.3 Synthetic weather trace model

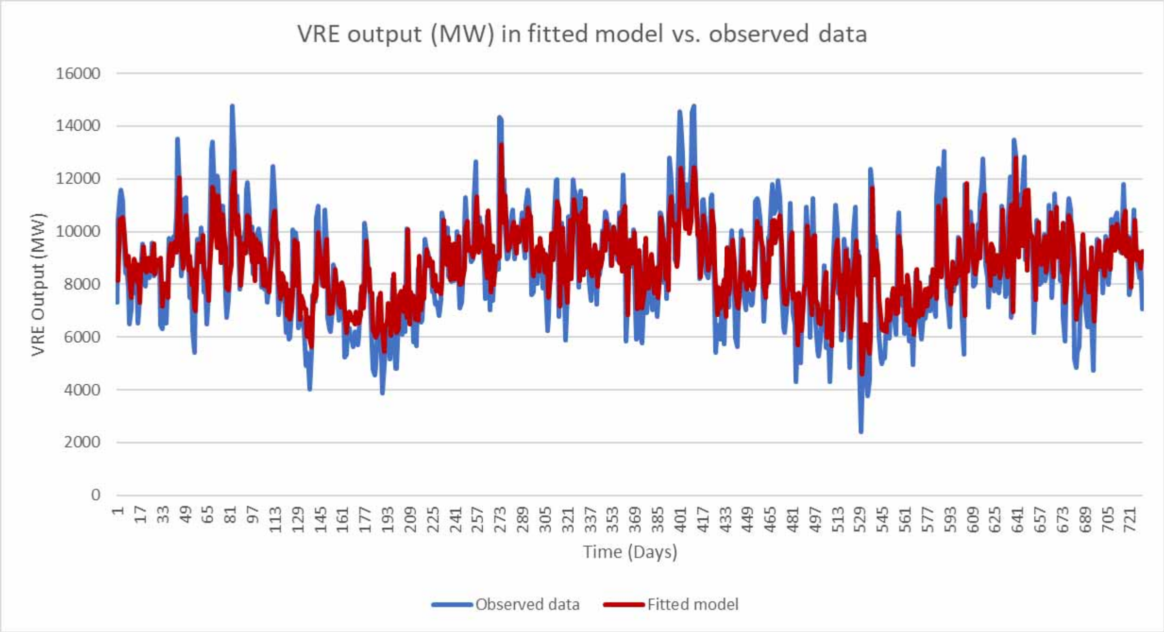
As found in the Directions Paper, dark doldrums play a key role in driving USE events. They may not be the only driver of USE events but, as explored in Appendix section C.3.2, they are likely one of the largest drivers of USE. As a result, it is important that we develop an understanding of their likelihood. For example, if we find that high depth / high duration USE events are largely caused by 1 in 50 year dark doldrum events then there may be evidence to suggest that more weather reference years should be used when modelling USE. It is unlikely for a 1 in 50 year dark doldrum to occur in an 11 year reference year sample but it is likely to occur in an 83 year reference year sample.

Professor John Boland from the University of South Australia was commissioned to perform statistical analysis on the 83 Griffith weather years to develop a robust and precise understanding of the likelihood of dark doldrums of various definitions. The methodology used to perform this analysis incorporates elements of statistical bootstrapping, where the patterns in the input set were decomposed into different components and then resampled to generate a much larger set of data with which to base statistically significant conclusions on. The detailed methodology is described below:

1. **Daily average VRE output per state**. 83 years of solar and wind power traces from the Griffith data set were utilized here. First, the data was transformed from standardized units to power units (MW) by applying the forecasted FY2040 capacities from each REZ. Next, results were aggregated to the regional level and solar/wind power were combined to produce VRE power output. Finally, the daily average was taken.
2. **Seasonality component**. The daily VRE output was found to have significant seasonality components. The first seasonal relationship occurs with 1 cycle per year and indicates that VRE output is higher during summer than winter. This is expected noting that daily solar output is generally higher in summer than winter. Furthermore, a second but less powerful seasonal relationship was found with 2 cycles per year. Seasonality was identified using the power spectrum where frequencies that contribute most significantly to explaining the variance of the data are identified. Once seasonality is identified, the seasonality is modelled via a Fourier series representation.
3. **Autoregressive component**. Next, the seasonality component is subtracted from the original data and the residual data is tested for ARIMA type features. The Box-Jenkins methodology is used to determine which ARIMA model best fits the residual data. It was found that the data follows an AR(3) process. In other words, the outcomes of today depend on the 3 previous days. This follows intuition noting that weather patterns tend to last for more than 1 day. Furthermore, the day before was found to matter most and the 3 day prior result found to matter the least.
4. **Noise component**. Next, the autoregressive component was removed and it was found that the remaining data exhibits no remaining autocorrelation. The result of the fitted model (including seasonality and autoregressive components) is included in figure X below. We can observe that there is significant remaining variability attributed to noise.
5. **Climate risk adjustment**. It was observed that the variance in the noise term slightly increased through time from 1940 to 2022. This could potentially be caused by climate change but it is difficult to prove. It should be noted that we are interested in estimating USE events into the future, as such, the variance of the noise term was linearly extrapolated out to 2040. Results for the synthetic solar/wind power trace model were produced both with and without the climate risk adjustment. The increased variance of the climate risk adjustment led to more volatile weather patterns and a higher likelihood of dark doldrums.

Figure B.1 shows the results of the fitted model against the observed data. The fitted model (red) follows the observed data (blue) well but there is remaining variability attributed to noise. Large noise variance is expected when modelling weather variability which is known to include a large degree of randomness.

Figure B.1: The fitted model matches observations well, although there is additional variability related to noise



Once the model is created, synthetic years can be simulated by estimating seasonal and autoregressive components from the model and then simulating noise terms randomly with replacement as observed in the 83 year dataset. This process was repeated 1,000 times to produce 1,000 synthetic years. There was little difference in results when comparing 1,000 and 2,000 synthetic years and so 1,000 years was used (convergence of simulation).

Once the 1,000 synthetic years are created, probabilities of certain dark doldrums can be quickly interpreted. Dark doldrums are defined as a certain number of consecutive days with VRE output beneath a certain quantile. Probabilities were produced for quantiles 5 per cent, 10 per cent and 25 per cent and for dark doldrums ranging from two to seven days in length.

Results from this analysis are outlined in Appendix E.

B.4.4 New calibrated stress testing PLEXOS model

A number of changes were made to the AEMC Unserved Energy Simulation Model (AUSM) built in the first phase of work for the Directions Paper. These changes were decided upon based on stakeholder feedback and to take a more targeted approach to understanding the changing nature of reliability risk and the impact of a wider range of weather conditions.

The new phase 2 model again used the 2022 ISP as a basis. A number of stakeholders suggested using the updated 2024 draft ISP as a basis, however the project team considered that there would be little benefit to updating the model considering the purpose of the modelling. Firstly, the modelling findings relate to a potential future high VRE penetration system and are not critically dependent on the most up-to-date buildout assumptions. Secondly, the project team considers that the 2024 ISP is likely to generate less USE than the 2022 ISP due to a significant increase in firm gas and storage in the newer model. Given the intention to use an reduced capacity model to study the characteristics of USE, we do not believe that updating the buildout assumptions will benefit the study materially.

Similar to the approach used in the model for the Directions Paper, effort was made to ensure that enough USE was generated in each region such that it could be studied. In this phase of modelling it was decided that demand side participation (DSP) generation would be taken out first as this is a much better proxy for USE and as DSP will only be activated in the model when there would otherwise be USE. Following this, a moderate capacity of OCGT generation in each region was removed, and finally it was necessary to still remove 20% of VRE capacity in SA and NSW to generate sufficient levels of VRE. Table B.3 roughly the volume of capacity removed in each region, noting that where capacity was removed across multiple financial years only FY40 has been given.

Table B.3: Capacity reductions were made in each region, focusing on DSP, OCGT and finally VRE

Region	DSP Capacity Removed (FY2040)	OCGT Capacity Removed	VRE Reduction
New South Wales	~960MW	664MW	20% reduction across all years
Queensland	~610MW	346MW	
South Australia	~205MW	105MW	20% reduction across all years
Victoria	~760MW	292MW	
Tasmania	~60MW	208MW	

An additional 83 reference years were also included in this model, which is described at the start of this Appendix. As the weather traces related to these 83 years were only available in REZs, each VRE generator in the model was associated with the weather traces from its closest REZ. This association was applied even to the original 11 reference years so that there is consistency between the AEMO and Griffith reference years.

As the modelling is focused on the impact of varying weather conditions in a higher VRE penetration environment, generator forced outages were all turned off and no simulations of forced outages were run. Results from the first phase of modelling indicated that the variation in outcomes from forced outage samples was far less than the variation produced by differing weather years, and so the project team believed that focusing time and computing resources on hundreds of forced outage simulations would not be beneficial.

B.4.5 Univariate analysis approach

The modelling in the Directions Paper studied the changing nature of USE as the NEM transitioned from a lower VRE system to a higher VRE system. This transition occurred across financial years where other changes were also present such as increasing demand, the profile of demand and rooftop PV was changing, and the capacity/mix of other technologies were changing such as storage and thermal units.

To ensure the findings from this approach were robust, the modelling team undertook a univariate approach to modelling in this phase, where the system was fixed at a single financial year and only utility VRE capacity was varied. FY2040 was chosen for this analysis, and VRE was increased and decreased by blocks of 10 per cent so that there were scenarios of up to 30 per cent less VRE and up to 30 per cent more VRE. Note that each scenario with an increased or decreased VRE capacity was put through the calibration process such that the overall level of reliability was still at or below the standard – this ensures that each scenario can be compared as they are all at a similar level of reliability.

B.4.6 Full capacity PLEXOS model

Following stakeholder feedback that the reduced capacity model may be unrealistic in its presentation of USE, the base 2022 ISP was used with minimal changes. The changes made relate to changing the VRE generation source files for existing VRE generators to the closest REZ trace. This facilitates the inclusion of the additional 83 reference years into the model.

B.4.7 Markov chain Monte Carlo model

As explored in chapter 7, PLEXOS modelling outputs have been used to investigate if the current form of the standard adequately captures the risk profile seen in modelling results. In order to do this, the current form of the standard (Expected value of yearly USE) must be replicated which requires calculation of USE outcomes at the yearly level and determining associated probabilities. 94 reference years have been used in modelling (83 Griffith + 11 AEMO) which provide 94 data points per financial year from which to evaluate the form of the standard.

Whilst 94 reference years is an improvement on the 11 reference years available from the Directions Paper modelling, 94 data points is a relatively small number from which to evaluate probabilities robustly. It is important to remember that the 94 years can be decomposed into more granular 30-minute intervals which provides approximately 1.5 million data points at the 30 minute level.

As a result, a Monte Carlo model was built to simulate synthetic yearly USE outcomes from the 30 minute granular data. Specifically, a Markov chain Monte Carlo model was used to take into account the fact that USE events tend to occur in clumps of multiple 30 minute periods. In other words, the probability of USE in a given 30 minute period depends on if the previous 30 minute period had USE or no USE.

The simulation model can be summarised as follows:

1. Determine if a given 30-minute period contains USE. This is determined by random simulation and its probability is dependent on if the previous time step contained USE or no USE. If it contains USE go to step 2 else go to step 3.
2. If it contains USE, we sample with replacement from the observed 30 minute USE values from the total 94 years (empirical distribution).
3. Repeat steps 1 and 2 for all 30 minute intervals in each simulation year. Add up all USE values for the given simulation year to determine a single yearly USE outcome.
4. Repeat steps 1 to 3 for the required number of simulation years.

The model was run for 100,000 simulation years. This is a sufficient number of simulations runs considering that key statistics such the average yearly USE values don't change much when comparing 100,000 and 1 million simulations (convergence of simulations). It should also be noted that the results of the 100,000 years largely follow that of the 94 years except that the inferred probabilities are less volatile due to the greater sample size.

For more information regarding how the 100,000 simulated USE years were applied to evaluate the current form of the standard please see section 7.1.

B.5 Key limitations of the modelling

As PLEXOS was the modelling software used to perform this study, it comes with several limitations, including but not limited to:

- **Perfect foresight**
  - This is a known issue fundamental to modelling the energy market using tools like PLEXOS. In the modelling for the Directions Paper, this was mitigated somewhat by reducing the total energy storage available by half. In the modelling for the Draft Report, battery durations were returned to their original values. This was done as many stakeholders felt that the capacity in the model should be as close to the ISP build as possible, and the issue of perfect foresight is not fully resolved by halving battery life.
- **Restricting storage to predetermined state of charge targets**
  - To mitigate this, we have shifted the solve window for modelling away from midnight to 10 am, when resource availability should be at its average daily peak.
- **Minimum interval lengths of 30 minutes**
  - No mitigation methods were attempted, but acknowledgement should be made that events are highly unlikely to fit neatly into 30-minute intervals.
- **Programmatic responses to events**
  - No mitigation methods were attempted, but acknowledgement should be made that the actual running of the network and participants still has a human component that will react differently to software.
- **Historical only weather patterns, which do not consider future climate conditions**
  - A significant range of reference years was used to capture the scope of known climactic conditions going back to 1940, which is a significant improvement on the 11-13 reference years used in typical reliability modelling, including by AEMO in the ESOO.
  - However, it must be noted that these reference years are based on historical weather conditions, and may not capture the variability and changing weather conditions that may be seen in the future.
  - When undertaking the analysis of dark doldrum frequency, Professor John Boland did take into account increased variability in generating the synthetic weather years, however, these were not strictly used in the USE simulation model.

Additionally, as this model is not a forecast, all observed results must be considered acknowledging they were synthesised using unrealistic versions of the NEM. This also means that they are likely more severe than anything that would occur, and while important, do not precipitate specific actions to mitigate them. Specifically, USE events have been generated in order to understand the characteristics of USE events when they do occur. This report seeks to highlight the characteristics of USE events when they occur rather than provide an indication of how often USE events are likely to occur in the future.

As detailed in appendix B.4.2, the machine learning model underestimates demand by about 12 per cent when estimating demand spikes in the top 1 per cent. Models often struggle to estimate the edge cases or extreme cases and in order to mitigate this limitation an adjustment has been applied (see appendix B.4.2 for more detail).

For the Markov chain Monte Carlo model, it should be noted that 100,000 simulated years is never as valuable as 100,000 real years which is not currently available and so this is a relatively unavoidable limitation. Furthermore, one could explore fitting a distribution to the size of the USE outcomes in a 30 minute period and then sampling from the fitted distribution rather than the empirical distribution. This would then allow 30 minute USE outcomes that were not seen in the observed data. Whilst this could be useful, care would need to be taken when sampling from the tails of the fitted distribution for which there is low data available for fitting.

B.6 Load shedding assumptions

Section 7.2.2 shows analysis relating to the impact of USE events on customers through rotational load shedding. This analysis was based on a number of simplifying assumptions detailed below:

- Equally sized blocks are used for load shedding, with the size chosen so that each block's maximum load is approximately 100 MW. This led 180 blocks for New South Wales, 140 each for Queensland and Victoria, and 50 for South Australia.
- Customers can be shed for up to two hours at a time. This is based on indicative information provided publicly by DNSPs[68], and news reports of the 2019 rotational load shedding event in Victoria.[69] The Panel understands that shedding periods may be shorter in some instances or in jurisdictions such as South Australia.
- Where the depth of USE decreases during an event, some load is restored. Those customers are considered to have finished their turn and will not be disconnected again until the next rotation.
- All loads are subject to load shedding. We note that certain loads may be exempt in reality.
- As for all of our analysis, the 16-hour clustering window is applied. This means that events occurring within a few days of each other, but not within 16 hours, are not counted as multiple load shedding.

However, the Panel notes that very deep events may have broader impacts if large segments of customer load need to be disconnected at the same time.

68 'What is load shedding?', Endeavour Energy website, found here; 'Load shedding', Citipower & Powercor website, found here; 'Electricity load shedding', Jemena website, found here; 'Load shedding', United Energy website, found here; 'Load shedding', AusNet website, found here; 'Load shedding', SAPN website, found here.

69 Harrison, D, 'What caused the blackouts in Melbourne, and do Victorians need to get used to power cuts?', ABC News, 26 January 2019, found here.

C Reliability risk is changing

C.1 If USE events occur they may be longer and deeper

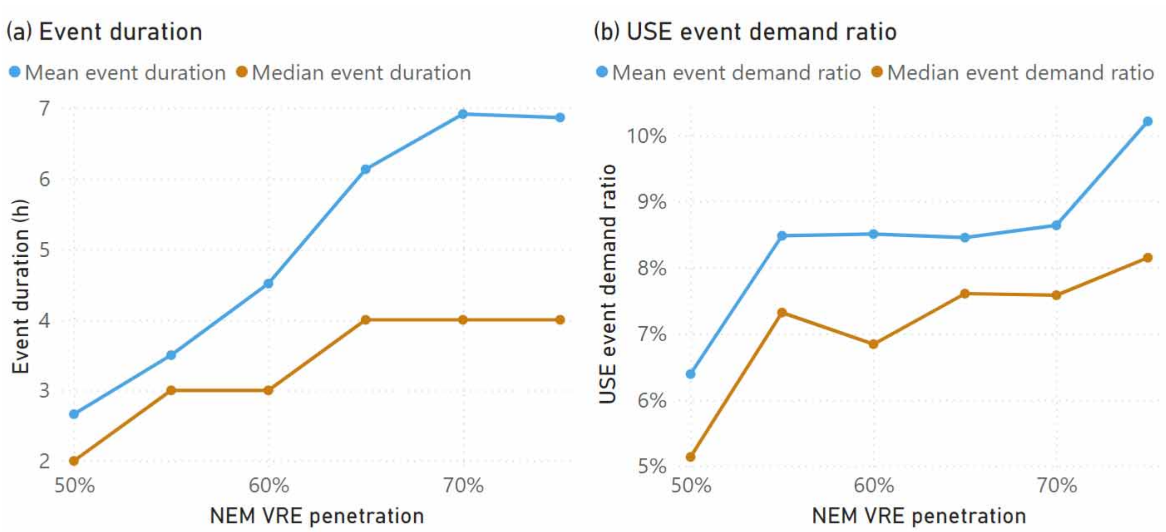
Consistent with the findings described in the Directions Paper, the modelling suggests that as the penetration of VRE increases during the transition, if USE events occur they may be longer and deeper but less frequent. The analysis showed that both the typical and largest USE events that might be expected were longer and deeper as the VRE penetration increased. This does not mean that reliability risk or the overall amount of USE would increase in the future NEM, but rather the same amount of USE might be concentrated in a smaller number of events.

C.1.1 Findings from the Directions Paper

Analysis for the Directions Paper indicated that while the depth, and duration of USE events were increasing as the NEM transitioned from FY2028 to FY2043 the events were becoming less frequent. While these results were potentially confounded by other factors, such as load growth and regional VRE levels, the latest modelling of a single year (FY2040) has confirmed the findings. The raw modelling results for the Directions Paper delivered levels of unserved energy that were at times well above the reliability standard. To make the results more comparable these raw results were calibrated back to the reliability standard by applying a capacity offset such that the overall level of reliability was set to the current reliability standard or below. Note that this calibration process is described in more detail in Appendix A of the Directions Paper.

Chart (a) in Figure C.1 shows the mean and median event duration plotted against the NEM-wide VRE penetration. The event duration has a clear increasing trend up to at least 70 per cent VRE installed capacity.[70] Chart (b) shows that the mean and median USE event demand ratio also increases with NEM-wide VRE penetration.[71]

Figure C.1: Event duration and USE event demand ratio increase as VRE penetration increases – results from the Directions Paper



70 The event duration represents the total time between the beginning and end of an event. The clustering process applied to the USE data 'joins' USE events where there is less than 16 hours between the end of one period of USE and the start of the next. As such there will be a difference between the number of hours of USE in an event and the event duration.

71 The USE Event Demand ratio represents the ratio of the amount of USE compared to the total customer demand during the event. Note that this chart and all others in this section are based on results that have been calibrated using the first calibration approach described in the Directions Paper.

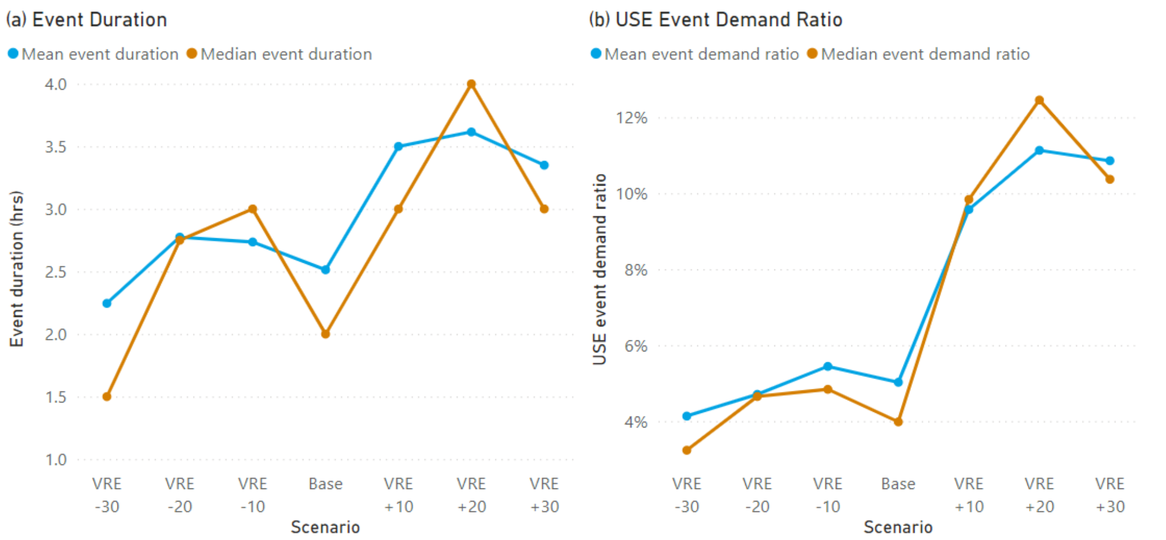
Trends beyond 70 per cent are based on materially fewer simulated outcomes and may be affected by significant variation in the technology mix across the regions, and so may not be statistically significant. Additional details on these findings can be found in the Directions Paper.[72]

C.1.2 Additional findings

To confirm this finding for the Draft Report, the Panel created a similarly reliable system. However, exploring outcomes from a single financial year in which the level of VRE penetration was incrementally varied. The results of this analysis support the conclusion that the depth and duration of USE events may increase in the transition to a higher VRE penetration system. The new analysis also suggests the increase may be more significant in terms of depth rather than duration.

Figure C.2 below similarly demonstrates the changing depth and duration of USE events as the VRE penetration in the NEM varies around the base case from the updated modelling. The scenarios to the left and right of the base case represent systems with an increase or decrease in the total VRE capacity by a percentage of the base case – for example 'VRE +10' represents the scenario in which the total VRE capacity was increased by 10 per cent.

Figure C.2: The univariate analysis on VRE penetration supports the finding that depth and duration of events increase as VRE penetration does

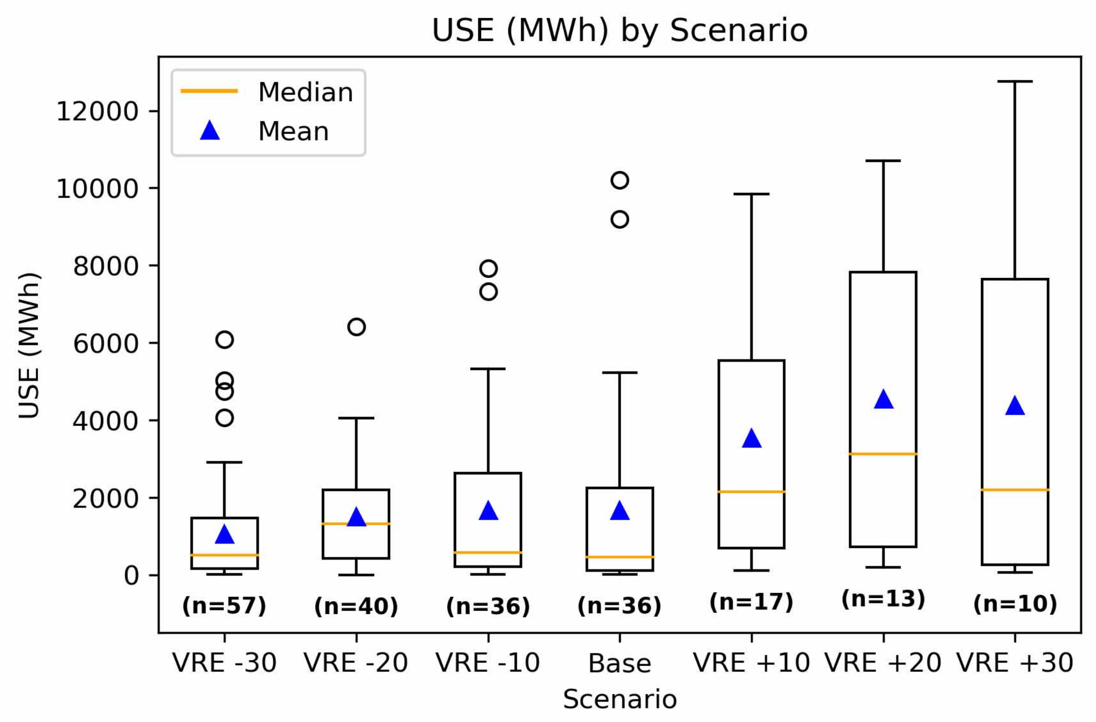


These results are slightly different to the results presented in the Direction Paper and Figure C.1 above. This is due to the adjustments to the model described in Appendix B.

Modelling for the Draft Report also explored the trends in the size of USE events in terms of total energy (MWh) at different levels of VRE penetration. This analysis was conducted on the reduced capacity model using AEMO VRE and demand input files, re-calibrated to the reliability standard. VRE volumes were adjusted in 10 per cent increments at the same time an equivalent level of firm capacity was also made. Figure C.3 below demonstrates the increasing size of USE events on a MWh basis.

72 Found here.

Figure C.3: Under a system held at a constant level of reliability, the mean, median and variability of USE events increases as VRE penetration increases, confirmed by univariate analysis



Note: **Box and whisker plots:** The ends of the box represent the 25th and 75th percentiles, that is, half of events fall within the box. The median and mean are marked as shown. In this case, the upper whiskers (short horizontal lines) indicate the 99th percentile and the lower whiskers indicate the minimum value. These values were chosen to emphasise the top 1 per cent of events in terms of USE duration. Events in the top 1 per cent are shown as small circles. The number of events in each scenario is notated below each box in parentheses.

Both the mean and the median USE event increase as the VRE penetration in the system increases. The mean is consistently higher than the median, suggesting a skewed distribution in which there is a larger tail of large USE events compared to small USE events.

The increasing length of the boxes demonstrates the increasing trend in the variability of USE – i.e. in a higher VRE system there will be a wider distribution of USE events rather than outcomes being clustered tightly around the mean. It is also important to note that the upper whiskers representing the 99th percentile of outcomes show a consistent increase with VRE penetration, again highlighting that USE outcomes if they do occur have a greater likelihood of being longer or deeper.

This chart also shows the number of USE events by scenario, which decreases as VRE penetration increases. This result demonstrates that a system at the same level of reliability may experience fewer USE events, but these events may be deeper and longer as the VRE penetration increases.

C.2 As the NEM transitions, reliability risks may shift from mainly being in summer to winter

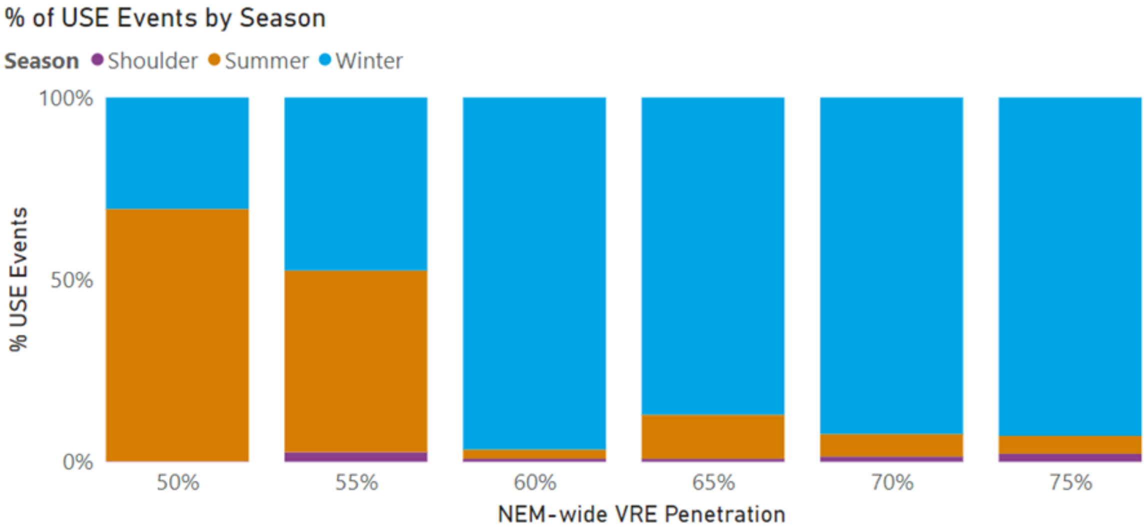
The modelling results suggest that winter may become a period of more significant reliability risk than summer in the future NEM. Several factors including changing demand patterns combined with increased VRE penetration may be driving this shift.

C.2.1 Findings from the Directions Paper

The modelling for the Directions Paper, which studied USE longitudinally between FY2028 and FY2043, showed a clear shift of reliability risk from summer to winter. The trend of USE events moving predominantly to winter months as the NEM transitions is clear across all the modelled results in the Directions Paper.

Figure C.4 below shows the percentage of unserved energy events occurring in different periods of the year as the VRE penetration in the region increases

Figure C.4: USE events move from primarily in summer to winter, as shown in the Directions Paper



Note: Summer represents months from December to March (inclusive) and Winter represents months from May to August (inclusive).
Note: This chart uses calibrated data.

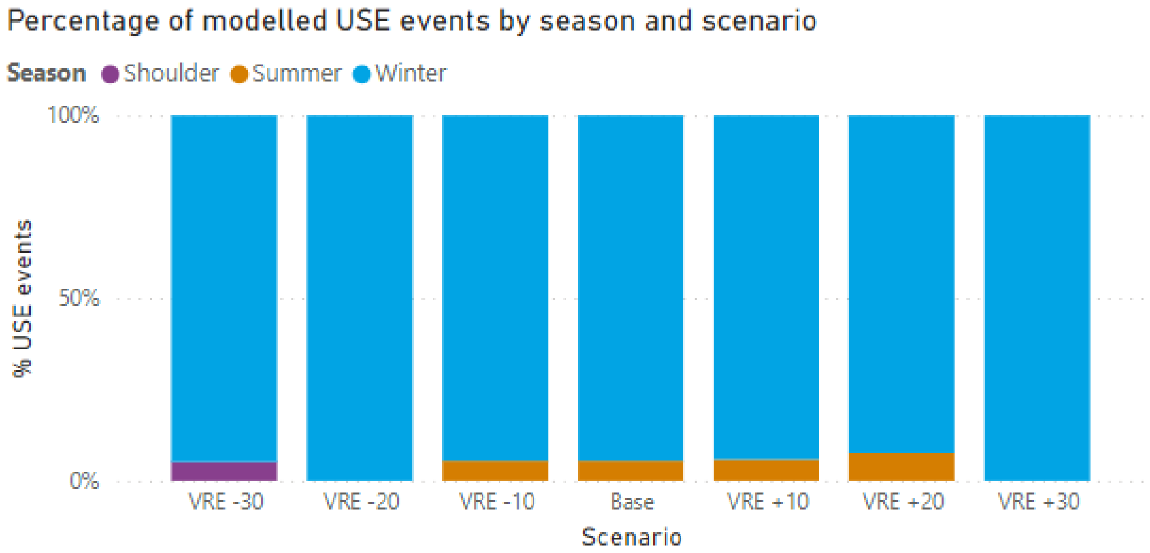
The analysis for the Directions Paper showed that while the penetration of VRE in each region differed, the broad trend was consistent across all regions – i.e. as the NEM transitions, the proportion of USE events in winter compared to other periods also increases.

It is worth noting that Queensland is the only region that continues with a more significant proportion of summer events for VRE penetration greater than 65 per cent. More information on this can be found in Section 5.2 of the Directions Paper.

C.2.2 Additional findings

Upon further analysis taking a fixed year (FY2040) and varying utility solar and wind capacity only, this trend was less evident. Operational demand was kept constant (as of FY2040) and so effectively rooftop solar was also kept constant. However, these results do show that the vast majority of USE events around FY2040 are likely to occur in winter as demonstrated in Figure C.5 below.

Figure C.5: Percentage of total USE events by season for incremental changes in NEM-wide VRE penetration for FY2040

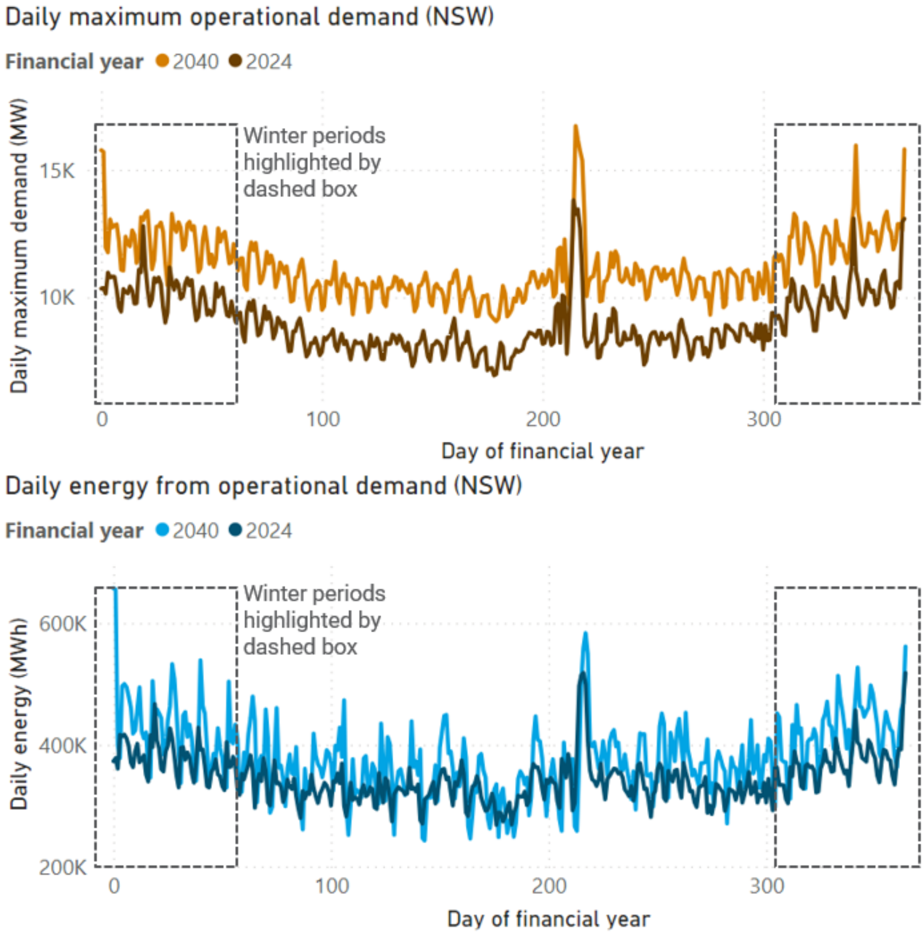


This suggests that the shift from summer to winter USE events observed in the Directions Paper results may be driven by multiple factors and not simply by the increase in utility solar and wind penetration. Note that our modelling for the Draft Report used the same operational demand traces (FY2040) for the lower and higher VRE scenarios, so distributed solar capacity was effectively kept constant. This means that Figure 5 only captures the effects of differences in large-scale VRE, and not distributed solar. Evolving demand patterns are likely to be another key factor. Customer demand is changing for several reasons including electrification, growth in CER including EVs and batteries, demographic changes, and changing weather patterns.

Some of the trends in demand between now and 2040 are shown in Figure C.6 below which highlights the forecast demand for NSW for FY2024 and FY2040, based on AEMO's 2011 reference year.[73] The winter periods from 1 May to 1 September are highlighted with a dotted box.

73 These traces are AEMO's demand forecasts for FY2040 based on reference year 2011. Our modelling used these traces as a direct input to the PLEXOS model, and also to generate demand traces for the Griffith years.

Figure C.6: There is significantly higher daily energy in winter from FY2024 to FY2040



Note: Data shown is based on 2011 in AEMO's set of reference years, as an illustrative example.

The key differences between the FY2024 and FY2040 based on the 2011 reference year traces include:

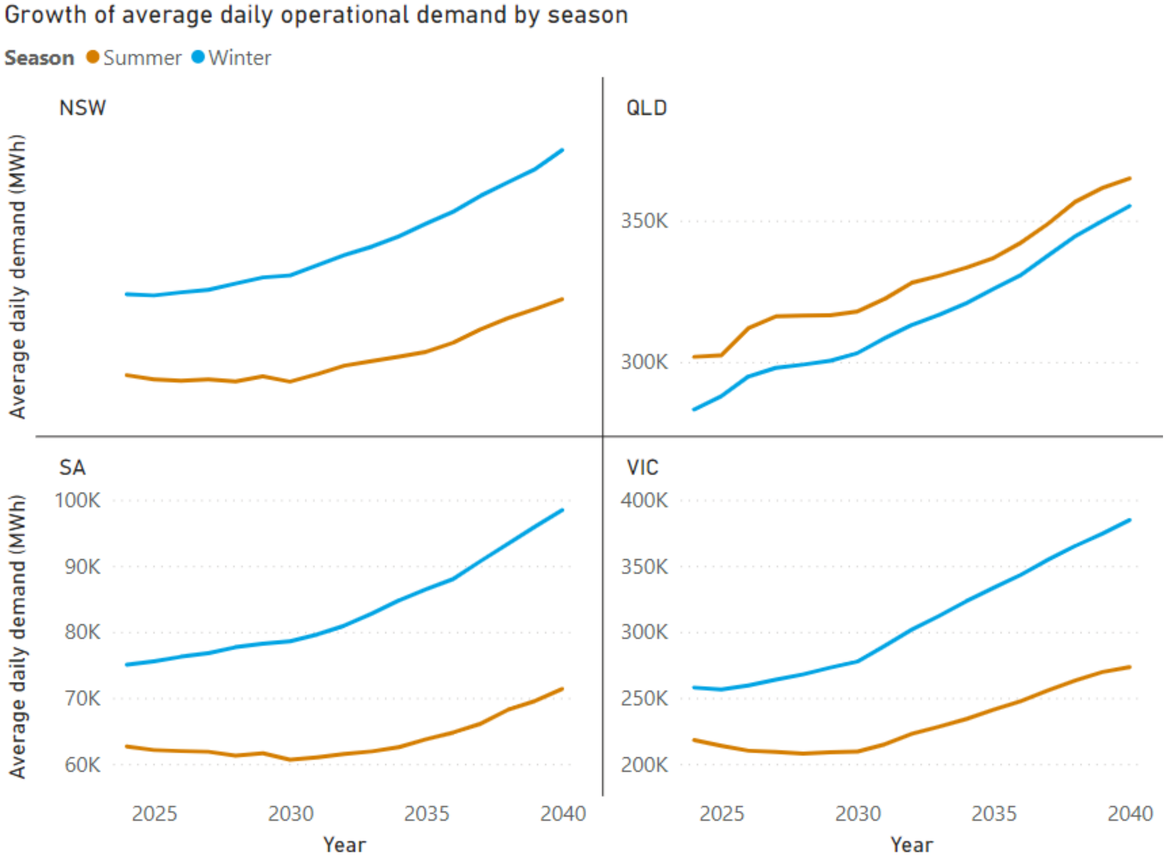
- Overall higher demand and greater daily energy variability,
- Higher peak demand, with summer peaks remaining slightly higher than winter peaks, and
- Materially higher daily energy in winter and slightly higher daily energy in summer, meaning the difference between average winter demand and average summer demand is larger in 2040.

These features are generally consistent across different reference years for New South Wales, Victoria, and South Australia, although there considerable variability between weather reference years. Queensland has a significantly different demand pattern as noted below.[74]

Figure C.7 is derived by averaging all 11 reference years and shows that the average daily energy in winter grows much more quickly than summer over the period FY2024-FY2040. For New South Wales in 2024, daily energy is 10 per cent higher in winter compared to summer, and by 2040 this difference is 17 per cent. There is an even larger shift for Victoria and South Australia as shown in Table C.1 below. These changes over time may reflect increasing demand for electrified heating in winter, and a greater contribution from distributed solar in summer.

74 Tasmania also has different demand characteristics and is largely excluded from our analysis. It experiences little to no USE in the model due to its very high VRE construction targets and hydro generation resources.

Figure C.7: Average daily energy (MWh) grows faster in winter than summer



Note: Summer is defined as December to March inclusive, and winter is May to August inclusive.

Table C.1: The difference between winter daily energy and summer daily energy for each region

Financial year	NSW	QLD	SA	VIC
2024	10%	-6%	20%	18%
2040	17%	-3%	38%	41%

The shift towards winter reliability risk may be caused by this rapid growth in winter demand in combination with the higher VRE penetration.

In Queensland, daily energy is higher in summer than in winter from 2024 to beyond 2040. Winter demand nevertheless grows slightly more quickly than summer demand as shown in Figure C.7. This is consistent with the finding from the Directions Paper that Queensland retains more summer reliability risk than other regions even as the NEM transitions.

The increased risk of winter USE events appears to be related to sustained high demand over a day or more, rather than afternoon peaks. Higher demand days coinciding with lower solar resources could stretch the supply-demand balance, deplete storage, and create the potential for USE events. There may be an opportunity for future modelling to further investigate the link between future demand characteristics and reliability risk.

C.3 USE events may be driven increasingly by weather

All of the modelling performed to date shows that as the NEM transitions, USE events may be increasingly driven by weather patterns compared to conventional plant failures, furthermore in a future high VRE power system weather driven USE events may be more extreme. Further modelling incorporating a wider range of weather reference years has also produced a wider range of USE outcomes, and revealed that whilst dark doldrums are a significant driver of USE outcomes, not all dark doldrums lead to USE.

Modelling for the Directions Paper revealed:

- as the NEM transitions, reference years (driven primarily by temperature and weather) may have a far greater impact on the mean time between events, depth and duration of USE events than forced outage samples or different PoE demand traces'
- 'dark doldrums' or periods of very low wind and solar availability may significantly impact the depth and duration of USE events
- as the proportion of VRE in the system increases, the impact of these dark doldrums or periods of low wind and solar availability on the depth and duration of USE events may also increase.

Stakeholders broadly agreed with this changing driver of USE outcomes. Modelling for the Draft Report extended this analysis to understand in more detail the impact of weather conditions on USE outcomes. This involved including more weather reference years and examining in more detail the relationship between dark doldrums of different sizes and USE outcomes. The results of this analysis indicate that:

- including more weather reference years leads to a greater range of USE outcomes
- Dark doldrums are a significant contributor to USE, although not all doldrum periods produce USE.

C.3.1 Findings from the Directions Paper

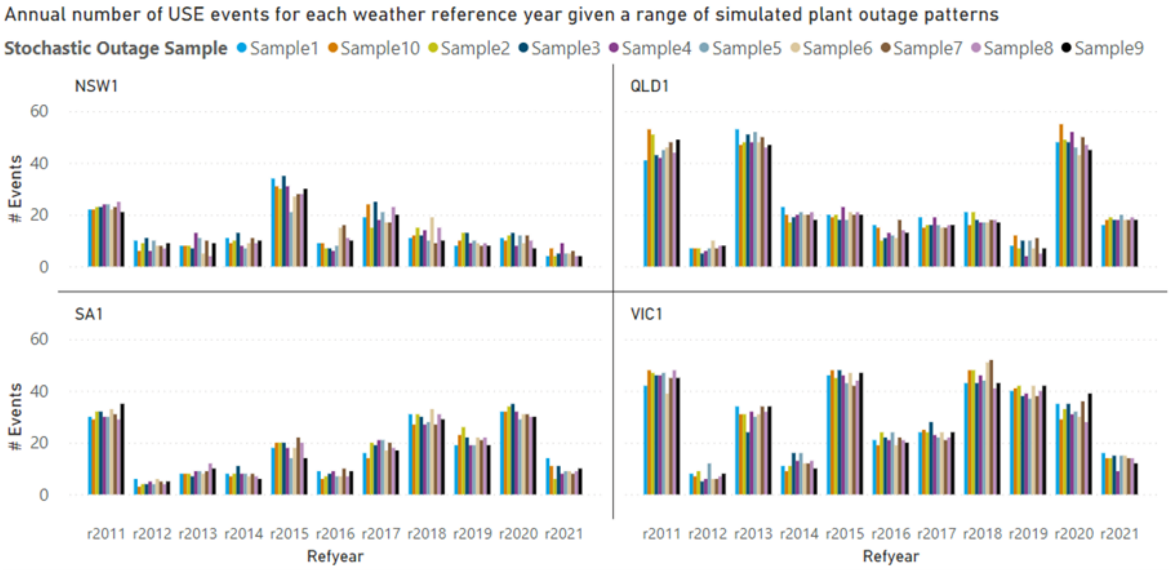
In the Directions Paper phase of modelling, it was demonstrated that forced outage samples do not drive USE outcomes as much as weather and demand traces.

Figure C.8 below shows the number of USE events generated for different reference years and stochastic outage samples, where the reference year is grouped on the x-axis, and the colour of the bar represents the forced outage sample. These results were generated in a reduced-capacity model.

Figure C.8 shows that the variability between reference years is greater than that generated by different forced outage samples. AEMO supported these findings and are already working on using more and potentially synthetic weather reference years with associated demand traces in their modelling of the future NEM.

The Directions Paper provides additional details on how weather has been demonstrated to drive USE outcomes.

Figure C.8: There is far more variation in number of USE events between reference years than between outage samples, indicating that year-to-year weather variability is a more important driver of expected USE than forced plant outages – analysis from the Directions Paper



C.3.2 Additional findings

Based on this finding, the modelling team chose to integrate the 83-year Griffith weather data with estimated demand into the simulation, rather than continuing to run multiple forced outage samples.

Modelling for the Directions Paper observed that when dark doldrums periods are inserted into existing weather and demand traces the depth, duration and variability of USE outcomes all increases. To refine this observation we fully integrated the 83-year Griffith weather data set into the simulation model so that a more extensive set of weather conditions, including dark doldrums of various types, could be modelled. To fully integrate these weather traces into the model, a representative demand profile was also constructed to match each of the 83 reference years. This was achieved using a machine learning approach described in Appendix section B.4.2.

As expected, the additional weather reference years produced a far greater number of USE events, including a larger number of deeper and/or longer events. This is shown in Figure C.9 below, which shows the USE duration and the USE demand ratio for the calibrated reduced-capacity model for financial years 2035-2040 inclusive. Each dot represents a USE event, where an event is defined as a group of intervals with USE that are 16 hours apart or less. (See Appendix section B.1.2 for more details on this 16-hour clustering.) The colour of the dot represents the set of weather years it was drawn from, where the dark blue dots are events that were produced from the 11 reference years provided by AEMO, and the green dots represent events produced from the extended 83-year set.

Figure C.9: In the calibrated reduced-capacity model, the extended weather data set produced significantly more USE events

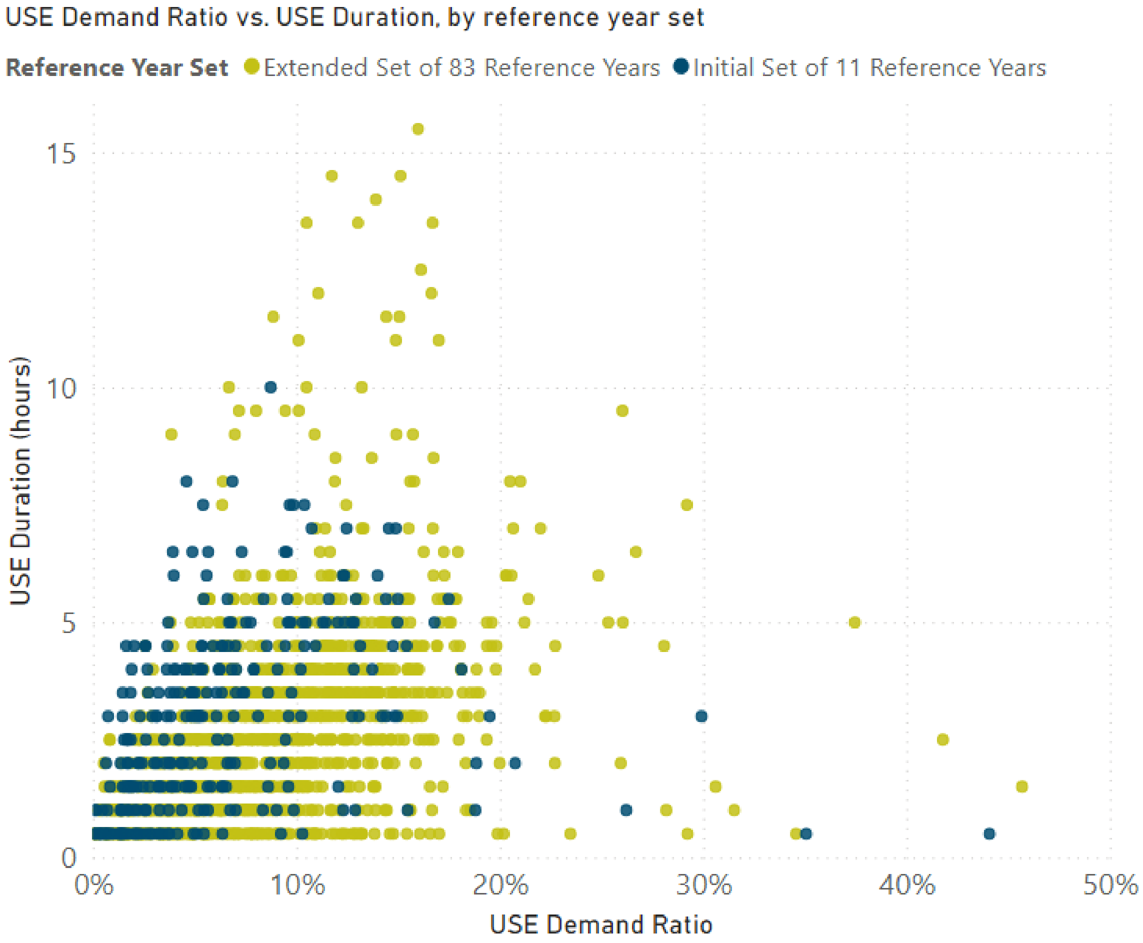


Figure C.9 demonstrates that the extended set of 83 years produced a large number of USE events, and in particular, produced a greater number of long and deep events. This indicates that modelling a greater number of weather patterns may lead to a more robust and perhaps more indicative result. Including additional weather reference years in reliability modelling is an improvement that has been flagged by AEMO already, as the Panel understands AEMO is undertaking work to increase the set of weather years used in their modelling. The Panel supports this work and has included it as a recommended process improvement in this report.

Another avenue of investigation in this stage of the modelling was to understand in more depth the relationship between doldrum periods and unserved energy. Working with Professor John Boland of the University of South Australia, the modelling team performed extensive statistical analysis of the 83 year Griffith data to define doldrums in each region as being below the 5th, 10th and 25th percentile of annual VRE generation and explored the correlation between doldrums and USE outcomes.

The likelihood of dark doldrums defined as daily VRE generation being below the 10th quantile of yearly VRE generation is shown in Table C.2 below. Likelihood is represented as 1 in X years, e.g. an entry of 4 refers to a likelihood of 1 in 4 years.

Table C.2: Likelihood of dark doldrums of differing lengths, where a dark doldrum is defined as a period in which daily VRE generation is below the 10th quantile of yearly VRE generation for FY40

Dark doldrum length (days)	NSW	QLD	VIC	SA	Whole NEM
2	5 in year	5 in year	6 in year	6 in year	5.6 in year
3	2 in year	2 in year	2 in year	1.7 in year	2.2 in year
4	1.3	1.2	1.5	1.9	1.3
5	3	3	4.5	6.3	2.8
6	7	6.6	10	21.2	7.3
7	9	6.5	19	41.5	7.7

This table shows that the absolute likelihood of doldrums is relatively rare, with 2 day doldrums defined by the 10th percentile output across the NEM likely to occur in 2040 at a frequency of 5-6 per year, and longer doldrums of 7 days or more occurring only once every 8 years or more. It should also be noted that results are region dependent (for example a 7 day dark doldrum occurs once in 42 years for SA) and likely depend heavily on the proportion of utility solar or wind in each region and/or the weather variability in each region.

The methodology describing how these probabilities were developed can be found in Appendix section B.4.3, and the full set of results is provided in Appendix E.

The correlation between doldrum periods were then compared with the daily conditions when USE occurred to look for a causal relationship between doldrums and USE outcomes. This showed that whilst doldrums are a significant contributor to USE outcomes, they are not the only driver of USE, and not every doldrum leads to USE.

The modelling team undertook statistical analysis on the VRE generation that was produced in the phase 2 model to determine whether each daily period was in a dark doldrum of differing definitions. In the following analysis we define a dark doldrum for this analysis as periods of two or more days in which the daily VRE generation was at or below the 10th percentile of annual VRE generation. We can then examine two alternative perspectives:

- Do days with USE events correspond with dark doldrum periods?
- Do dark doldrum periods correspond with days with USE?

Figure C.10 shows the correlation between days with USE and the likelihood of dark doldrums of different doldrum lengths. It shows that whilst almost 25 per cent of USE events occurred on days not related to a dark doldrum, around 70 per cent of USE events occurred during dark doldrum periods of 2 days or more.

Figure C.10: USE days may - but not always - coincide with dark doldrums

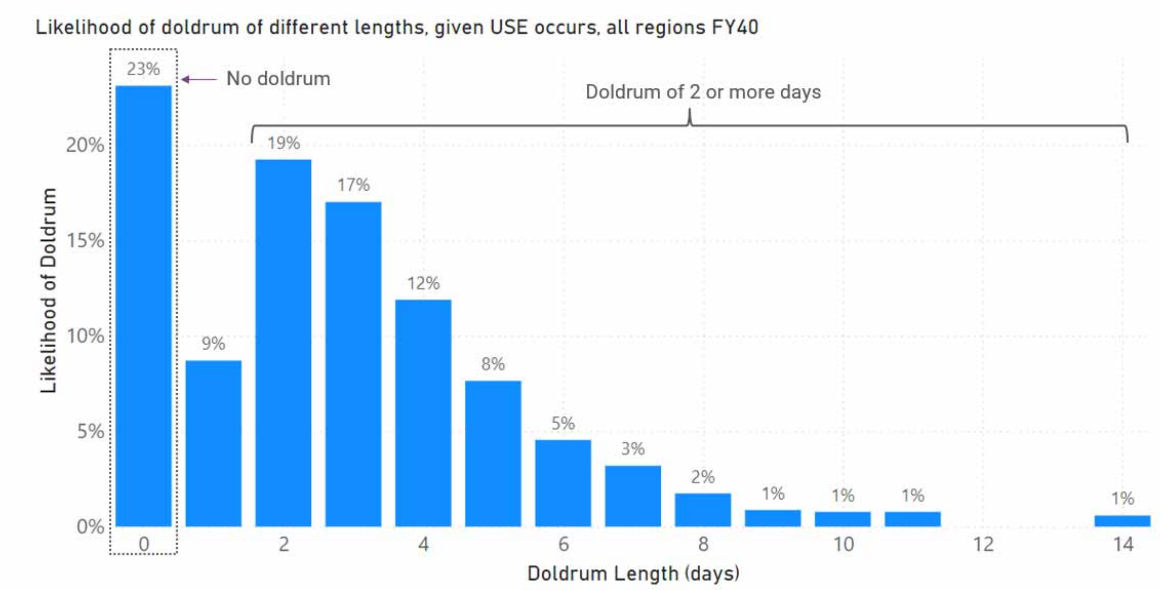
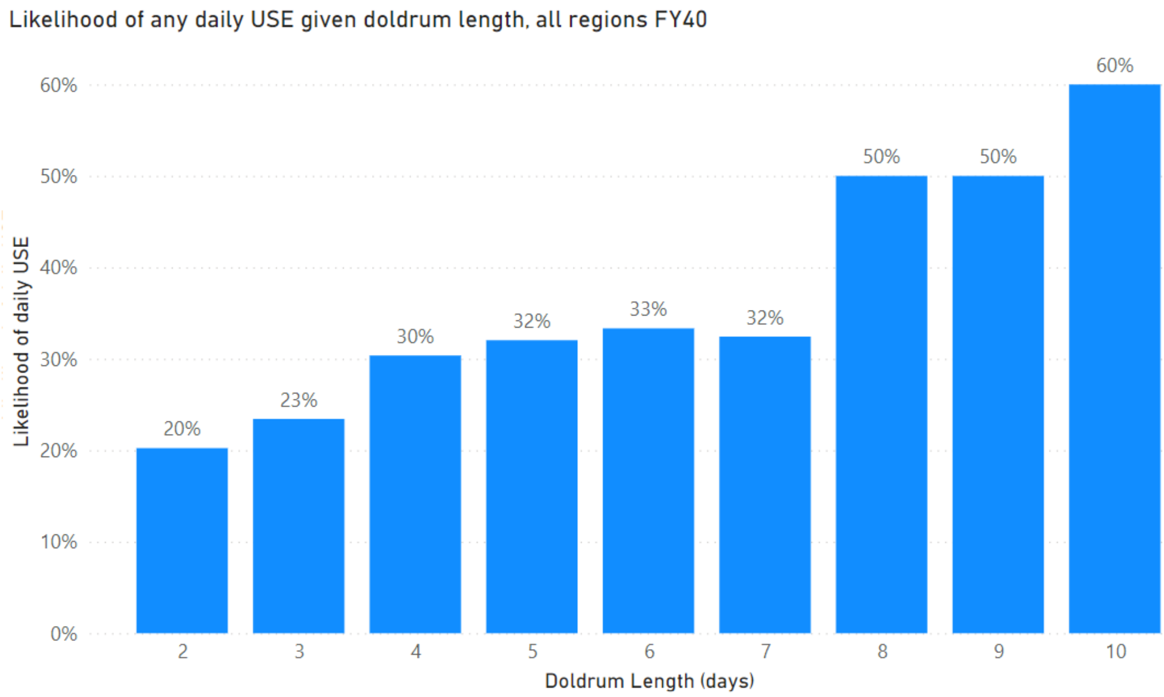


Figure C.11 shows the alternative approach in which each doldrum period is tested for its correlation to USE events. This figure demonstrates that dark doldrums lead to daily USE between 34 per cent to 60 per cent of the time, with doldrums of eight days or longer being increasingly more likely to lead to USE. Taking all doldrums of two days or greater, the likelihood of producing daily USE is approximately 43 per cent.

Figure C.11: Longer dark doldrums are more likely to cause USE



Note that sample size for each bar in these charts is not equivalent – as the length of doldrum increases, the number of events seen in the modelling dramatically decreases, for example there were 215 two-day doldrums in the modelling but only 5 ten-day doldrums.

Overall this analysis of correlation between doldrums and USE outcomes has revealed that when defining a doldrum of 2 or more days of VRE output below the 10th percentile:

- The likelihood of USE given the region is in a doldrum is approximately 43 per cent
- Given a USE event in a region, the likelihood that it occurred during a doldrum is approximately 68 per cent.

This indicates that while doldrums are highly correlated with USE (two-thirds of USE events occur during a doldrum), a doldrum is certainly not a guarantee that USE will occur, as more than half of doldrums in the modelling did not produce any USE at all. More work is needed to understand the most significant drivers of USE outcomes, which will be explored in more detail in the final report.

It must also be noted that one may define a doldrum using a more or less conservative threshold and defining it in different ways will change the strength of the correlation. For example, defining a doldrum less strictly at 2 or more days of VRE generation below the 25th percentile would increase the likelihood that a doldrum occurs alongside USE from 68 per cent to 89 per cent but would conversely decrease the likelihood of USE given the region is in a doldrum from 43 per cent to 38 per cent.

D Supplementary analysis for assessment of the current form

The figures below are an extension of section 7.1 in the main body of the report, and demonstrate that the current form is fit for purpose given the modelling results for Queensland, New South Wales and South Australia. Note that only results for FY2040 have been provided in this report, though the Panel notes that the years FY2035 - FY2039 were also analysed in the same way and produced results consistent with the overall finding. The way that this chart has been constructed and its implications are provided in more detail in section 7.1.

Figure D.1: The vast majority of events produced in the reduced capacity calibrated model are material to reliability forecasts under the current standard in QLD

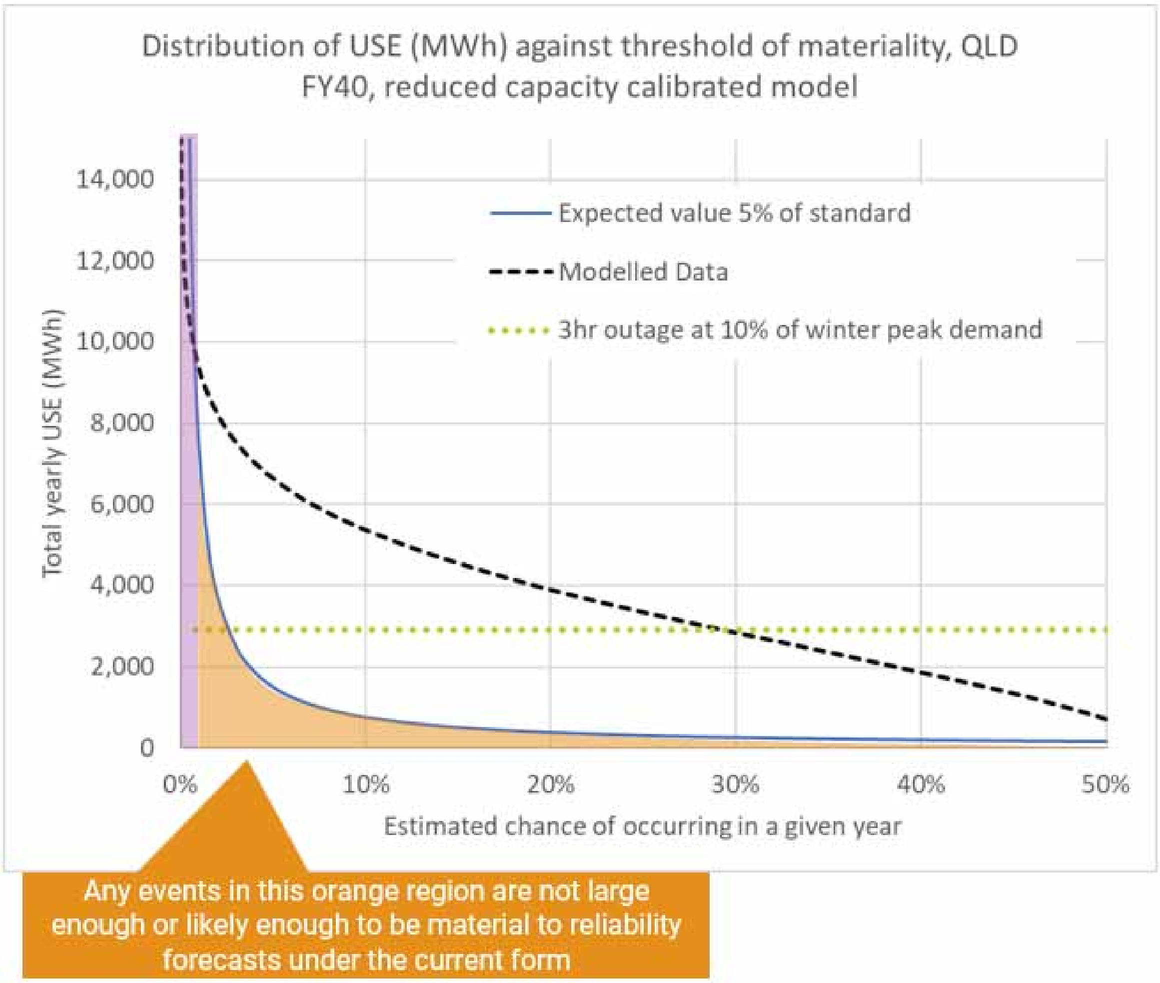


Figure D.2: The vast majority of events produced in the reduced capacity calibrated model are material to reliability forecasts under the current standard in NSW

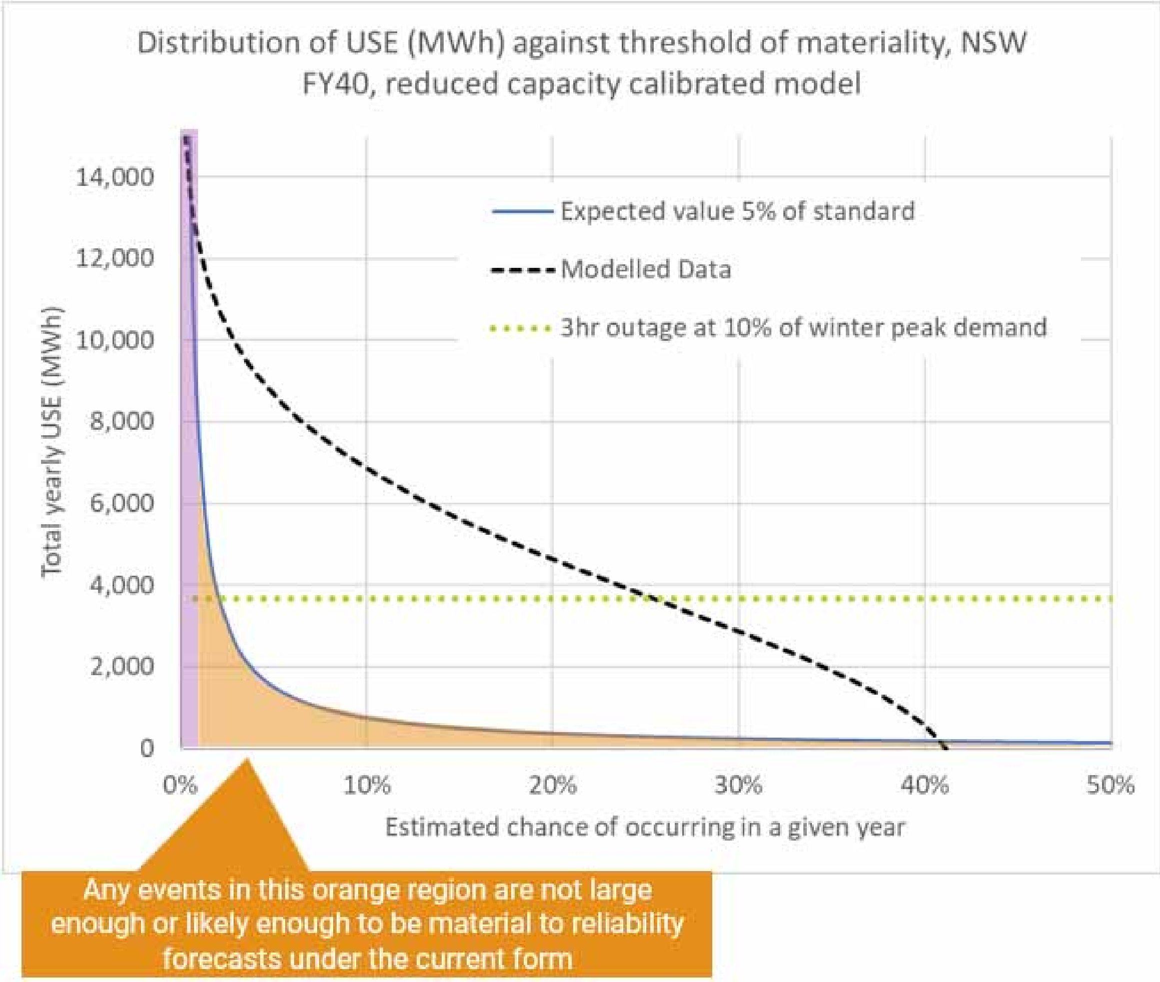


Figure D.3: The vast majority of events produced in the reduced capacity calibrated model are material to reliability forecasts under the current standard in SA

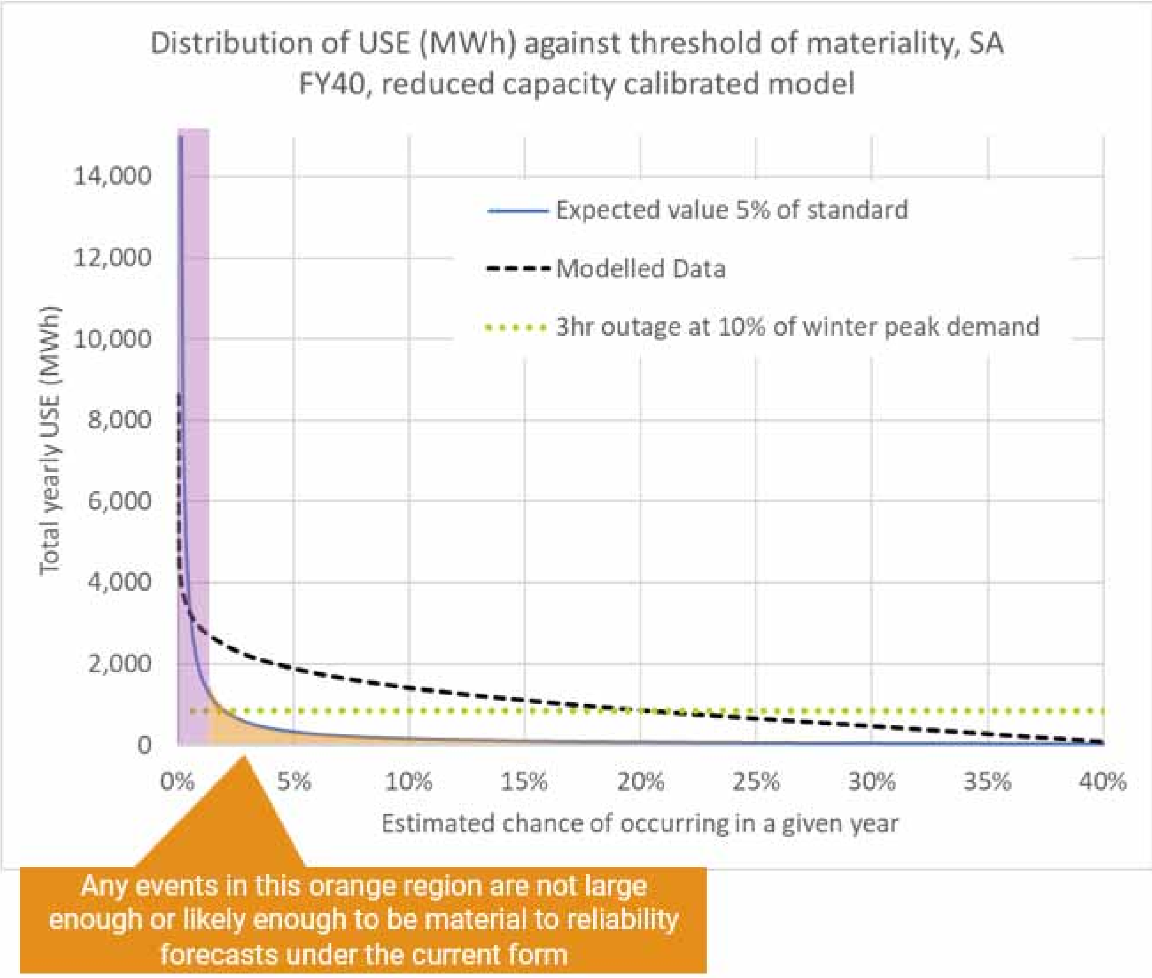


Figure D.4 below shows the result for Victoria in FY2040 using the full capacity model rather than the under-resourced calibrated model. The modelled events using this model show a consistent result, where the vast majority of USE events are materially captured by the current form of the standard, and those that are not materially captured are too rare to be appropriate for the reliability standard anyway.

These results are consistent with other regions in other financial years, though only Victoria in FY2040 has been included in this report.

Figure D.4: Even using the full capacity model, the majority of events produced are material to reliability forecasts under the current standard in VIC

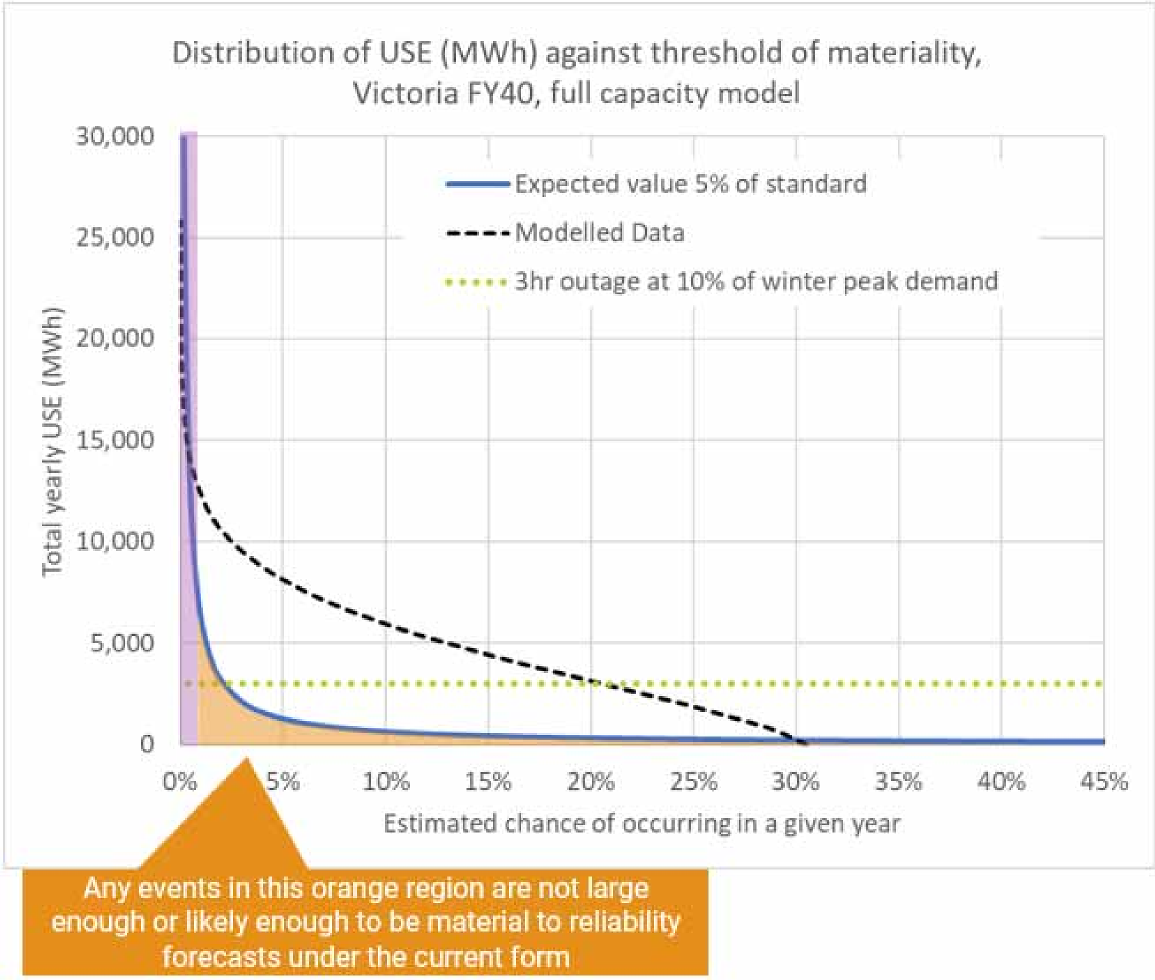
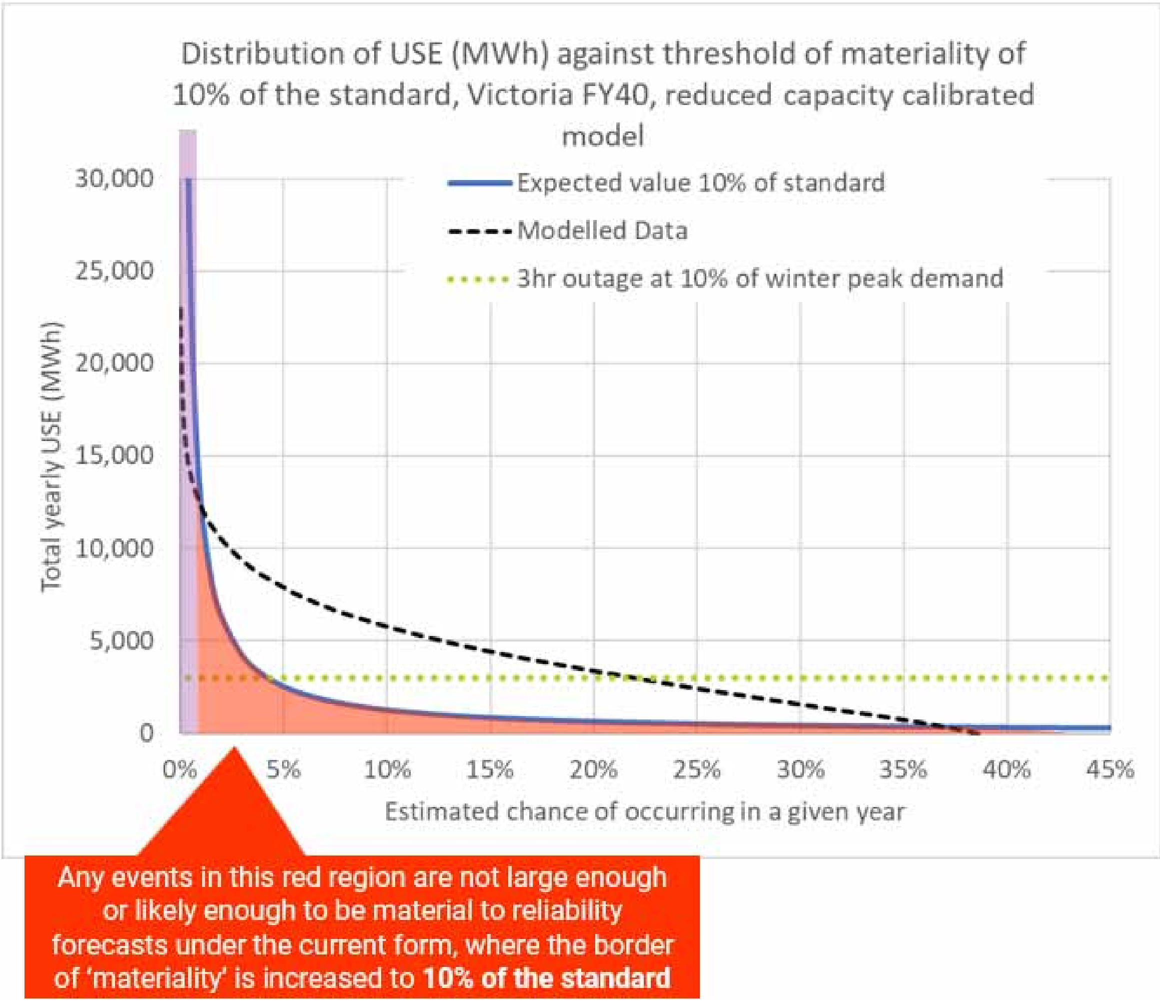


Figure D.5 below shows the results of the under-resourced calibrated model for Victoria in FY2040, but assuming a less conservative view of what is material to reliability forecasts. Instead of using 5 per cent as the threshold for materiality, this analysis assumes that any event with expected USE of up to 10 per cent of the reliability standard is not material.

Figure D.5: Even using a less strict definition of materiality, the majority of events produced are still material to reliability forecasts under the current standard in VIC



This figure shows that even if we define a threshold for materiality at an expected value of 10 per cent of the standard rather than 5 per cent, the vast majority of events produced by the model are still considered material, and those that are not material are either too rare to be addressed by the standard anyway or are extremely small in size.

This result broadly holds across all regions and financial years.

E Results of the analysis on dark doldrum likelihood

The tables below show the likelihood of dark doldrums of various lengths and definitions based on the statistical bootstrapping approach performed on the 83 years of weather data for FY40. The Panel engaged a consultant, Professor John Boland of the University of South Australia, for this work. The assumptions and methodology that sit behind these results are described in Appendix B section 4.3.

The likelihood of dark doldrums varies within states likely due to different solar/wind proportions and different weather patterns. Unsurprisingly, the likelihood of dark doldrums decreases as the length of the dark doldrum increases. Results for the modelling with climate risk adjustments generally increases the likelihood of dark doldrums indicating that there may be a greater risk of dark doldrums in 2040 than historically.

Note that in these tables, likelihood is represented as 1 in X years, e.g. an entry of 4 refers to a likelihood of 1 in 4 years. However, entries of the form ‘3 in year’ mean that this type of doldrum would be expected 3 times per year on average i.e. it is more frequent. There is a table for each definition of a doldrum, where a doldrum is described as a period in which the total daily VRE generation is below either the 5th, 10th, or 25th quantile of total VRE generation. The threshold in each region that defines this quantile is also provided at the bottom of the table.

Table E.1 shows dark doldrum likelihood vs region vs length without climate risk adjustment for the 5 per cent quantile level.

Table E.1: 5 per cent quantile analysis results of the likelihood of dark doldrum events of different lengths by region without additional climate volatility

Dark doldrum length (days)	NSW	QLD	VIC	SA	Whole of NEM
2	3 in year	2.7 in year	3.7 in year	3.3 in year	3 in year
3	1	1.2	1.2	1.4	1
4	4	3.5	3.8	5.4	4
5	11	10.6	11.6	22.5	10.7
6	33	25.7	43	90	29.2
7	43	45	120	154	45
Threshold (MW)	*3,620*	*5,500*	*2,230*	*1,110*	*16,000*

Table E.2 shows the likelihood of a dark doldrum vs region vs length with climate risk adjustment for the 5 per cent quantile level.

Table E.3 shows the dark doldrum likelihood vs region vs length with climate risk adjustment for the 10 per cent quantile level.

Table E.4 shows dark doldrum likelihood vs region vs length with climate risk adjustment for the 25 per cent quantile level.

Table E.2: 5 per cent quantile analysis results of the likelihood of dark doldrum events of different lengths by region with additional climate volatility

Dark doldrum length (days)	NSW	QLD	VIC	SA	Whole of NEM
2	3.5 in year	3.6 in year	5.6 in year	3.7 in year	4.4 in year
3	1	1.1 in year	1.7 in year	1.1	1.4 in year
4	2.9	2.6	2.1	4.4	2.4
5	8.5	6.5	5.7	18.4	6.1
6	21.8	15.4	14.3	48.8	13
7	46	18.9	26.8	92.9	22.4
Threshold (MW)	*3,620*	*5,500*	*2,230*	*1,110*	*16,000*

Table E.3: 10 per cent quantile analysis results of the likelihood of dark doldrum events of different lengths by region with additional climate volatility

Dark doldrum length (days)	NSW	QLD	VIC	SA	Whole of NEM
2	5 in year	5 in year	6 in year	6 in year	5.6 in year
3	2 in year	2 in year	2 in year	1.7 in year	2.2 in year
4	1.3	1.2	1.5	1.9	1.3
5	3	3	4.5	6.3	2.8
6	7	6.6	10	21.2	7.3
7	9	6.5	19	41.5	7.7
Threshold (MW)	*4,260*	*6,180*	*2,820*	*1,380*	*17,850*

Table E.4: 25 per cent quantile analysis results of the likelihood of dark doldrum events of different lengths by region with additional climate volatility

Dark doldrum length (days)	NSW	QLD	VIC	SA	Whole of NEM
2	11 in year	10 in year	11 in year	12.3 in year	10.9 in year
3	5 in year	5 in year	5 in year	5 in year	5.6 in year
4	3 in year	2.7 in year	2.25 in year	2.4 in year	2.7 in year
5	0.6	1.5 in year	1	1.5 in year	1.5 in year
6	1	1	2	2.3	1
7	0.7	1.6 in year	2	2.3	1.2
Threshold (MW)	*5,540*	*7,392*	*3,940*	*1,900*	*21,000*

F Details of alternative form assessment

F.1 Steps taken to assess alternative forms and findings

To assess the suitability of the current form of the standard against alternative candidate forms, the Panel:

1. **Developed a shortlist of alternative candidate forms.** An initial list of alternative forms was drawn from the Mathematics in Industry Study Group (MISG) report, stakeholder feedback, economic principles and other jurisdictions.[75]

   The Panel carried out work to understand how (if at all) each alternative form could be implemented and then categorised different types of forms into groupings by function. The Panel then drew up a shortlist list alternative forms that represented a more readily implementable example from each category.

2. **Developed a comparison framework.** To assess the alternative forms, the Panel operationalised the assessment criteria from the 2023 Issue Paper.[76] This multi criteria assessment framework includes both qualitative and quantitative elements.

   - The qualitative elements focused on how easily an alternative form could be implemented and whether it could be clearly communicated. The ease of implementation had three sub-categories for assessment:
     - How easily it could be implemented in the market modelling for the RSSR, where the level of the reliability standard is set;
     - How easily it could be implemented through the market price settings, which are used to realise the reliability standard; and
     - How much additional data would be required for implementation, including consumer preference data.
   - The Panel then set out to quantitatively assess the performance of any alternative standard, based on its the ability to 'capture' large, low probability events and its stability in the absence of large, low probability events.

3. **Assessed candidate forms.** To this end, the Panel built a quantitative testing model that took the USE events from the PLEXOS modelling as an input, and artificially constructed 'tail risk' to test the quantitative performance of each candidate form. The model did this by setting aside the largest observed events and then assigning each of the remaining events a probability of occurring consistent their modelled distribution, but with a scaling factor, so that each year and jurisdiction was at the reliability standard under the current form of the standard (expected USE). Large events were similarly assigned a probability, and subsequently added back to the other events.

   This process was repeated using different random samples of the non-large events. Alternative candidate forms were then implemented in the model.[77]

   To assess the stability of each candidate form, we kept the large events set aside (so each jurisdiction-year was at the current reliability standard), and calculated how consistent the alternative metrics were at evaluating this. To assess the ability to capture large low probability events, we added the large events back in and tested whether there was a material

75 MISG advice to the AEMC in 2023.

76 Found here.

77 Accurately parameterizing these alternative candidate forms was not possible given current available data, but can be updated as more data or information becomes available. For example, CVAR requires two parameters: 1. choosing which part of the tail of the distribution is at risk, and 2. The relative weighting between the entire distribution and the tail.

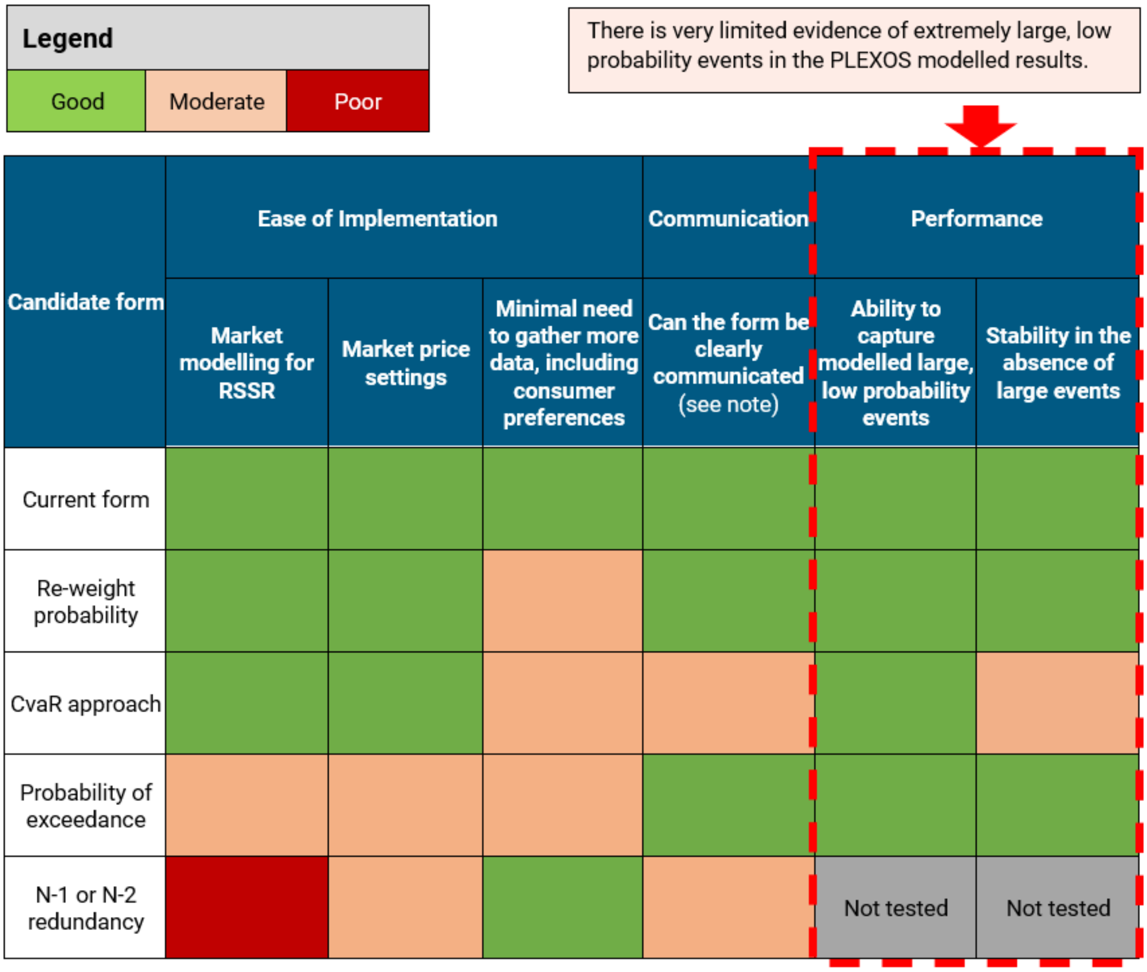
change in the reliability metric. More detail on the quantitative assessment model can be found in appendix F.2.

Figure F.1 below shows a summary of our assessment framework for the alternative candidate forms. Table F.1 provides more details on the assessment of each form. The current form of the reliability standard outperforms the other candidate forms in that each of them either is more difficult to implement or communicate.

We further note that the current form still does detect large, low probability events, but weights them proportionally to their likelihood of occurrence. This can be bolstered through the process improvements and updated modelling discussed in previous sections.

Given these findings and the modelling results, the Panel considers that the costs and complexity of changing the form would outweigh any limited benefits offered by the alternative form.

Figure F.1: Candidate forms assessment



Note: As noted in section 3.2, AEMO considers that the form of the reliability standard should explain the nature of large events that may occur. The Panel agrees that AEMO should continue to communicate the existing standard in different ways, which is consistent with the current expected USE standard.

Table F.1: Details on alternative candidate form assessment

Alternative form	Rationale for assessment
Current form	• Calculation of the current form is straightforward and easy to implement.
Re-weight probability	• The values to be re-weighted will need to be informed by additional data collection (such as consumer preferences). • The form is sensitive to which specific events are re-weighted.
CVaR approach	• The values of the CVaR weighting parameters will need to be informed by additional data collection (such as consumer preferences). • Difficult to clearly communicate due to inherent mathematical complexity. • There is a trade-off between CVaR's ability to capture large, low probability events and the form's stability. When the CVaR form is set to be more sensitive to large events, it can be less stable in years without tail events.
Probability of exceedance	• Calculation is more resource intensive than the current form. • The constraint level may or may not be binding. This introduces difficulty in implementation with respect to market price settings. • Appropriate setting of the constraint level will need to be informed by additional data collection. • The form captures large events to a greater extent than the current form. • If the constraint is not binding, this is equivalent to the current form
N-1 or N-2 redundancy	• Requires intensive modelling, which makes calculating the form difficult. • The constraint level may or may not be binding. This introduces difficulty in implementation with respect to market price settings. • Difficult to clearly communicate compared to the current form, particularly at an operational level.

F.2 Quantitative assessment model

To assess the performance of the shortlisted forms, the quantitative testing model constructed by the Panel takes the USE events from the PLEXOS modelling as an input, and artificially constructs 'tail risk'. The model achieves this by:

- Setting aside the largest observed events from the other 'non-large' events.
- Creating many random samples of these non-large events.
- Assigning a probability to each of the non-large events in accordance with their modelled probability distribution.
- Scaling these probabilities by a constant factor, such that each sample, year, and jurisdiction was at the reliability standard under the current form of the standard (expected USE).
- Assigning a probability to each of the large events in accordance with their modelled probability distribution.
- Adding these large events to each sample of non-large events.

Shortlisted alternative candidate forms were then implemented in the model, and results were evaluated against each of the two quantitative sub-categories under Performance:

- To assess the stability of each alternative candidate form, we considered how consistent each alternative form performed among the samples of non-large events.
- To assess the ability of each alternative candidate form to capture large, low probability events, we measured the difference in value between the samples that include large events and those that did not, for each form.

Note that empirically, the PLEXOS modelling did not find the types of large, low probability events that the quantitative model set out to capture, hence the need to artificially induce 'tail risk'.

The Panel did not assess the quantitative performance of the N-1 or N-2 redundancy form because it is qualitatively different to other forms and does not fit in the current modelling framework. Given the form's performance in the qualitative assessment criterion, the Panel did not consider that a quantitative assessment of the form would change the overall position.

F.3 Forms not shortlisted for assessment

Several alternative forms were considered by the Panel, but not shortlisted for detailed assessment. Rationale for their exclusion is included in Table F.2 below.

Table F.2: Alternative forms not shortlisted for detailed assessment

<table>
<tr><th>Alternative form</th><th>Rationale</th></tr>
<tr><td>Median event expected USE / annual demand</td><td>During preliminary quantitative testing, this form was outperformed by the current expected USE form in terms of both stability and ability to capture large events.</td></tr>
<tr><td>Mean outage duration or frequency</td><td>This was part of the MISG report. However, they do not assess depth. As we are concerned with both the depth and duration, these are insufficient.</td></tr>
<tr><td>Re-weighted demand traces</td><td>This form is the 're-weight probability' approach. Further, this approach would put more weight on all USE events, not just tail events.</td></tr>
<tr><td>The average annual outcomes that are at or above a one-in-10-year probability may not be greater than x per cent of average regional load shed for 4 hours, or equivalent</td><td>This form, initially proposed by the AEMO, is similar to the 'probability of exceedance' approach as covered in the body of the report.</td></tr>
<tr><td>Set limit on the value of x per cent quantile USE event / annual demand</td><td rowspan="4">• These are derived from the MISG report.
• They are similar to the shortlisted 'probability of exceedance' approach.
• The probability of exceedance approach is preferred as it explicitly considers the probability of an event. Specifying a 'limit' as in these approaches necessitates elimination, rather than reduction of risk and therefore may be prohibitively expensive.</td></tr>
<tr><td>Set limit on the 'most probable maximum risk'</td></tr>
<tr><td>Set limit on the expected value of the maximum USE</td></tr>
<tr><td>Set limit on the maximum probable duration or frequency</td></tr>
</table>

G The application of the VCR in the 2022 RSSR

Clause 3.9.3A of the NER and the RSSR Guidelines currently require the Panel to have regard to the value of customer reliability in determining the level of the standard and settings.[78] The value that consumers place on reliability is a key consideration for determining an efficient level of the standard.

Box 1: Values of Customer Reliability

The AER's VCR review, conducted every 5 years and most recently in 2019, surveys consumers to understand how much they are willing to pay for reliability (that is, to avoid outages). The VCR survey includes questions about a range of outage characteristics (such as season, duration, and time of day) and covers a wide range of customer segments (all NEM regions, several climate zones, urban and regional customers, residential, commercial and industrial customers).

The AER uses these discrete outage characteristics to create 32 outage scenarios, each with a different VCR. They then take a weighted average according to the probability of each scenario based on historical outage data from distribution network Regulatory Information Notices (RINs).

Most outages that consumers experience are due to localised distribution failures. Reliability events have only caused approximately 0.3 per cent of customer outages historically in the NEM.

Source: AER VCR final report 2019, found here; Reliability Panel's Annual market performance review 2018, p. xlvi, found here.

Previous RSSRs have used the AER's final state-based VCR values for the base case. These values were weighted by customer load and outage characteristics (based on RIN data). However, as mentioned above, the VCR survey actually provides much more detailed information. The 2022 RSSR used some of this information to conduct high VCR and low VCR sensitivities.

In the high VCR sensitivity, the Panel re-weighted the VCR results to include only one-hour outages, as this duration was most closely aligned with the length of rotational load-shedding interruptions.[79] This formed a high case as customers value short outages more highly on a per-kWh basis.[80] This higher sensitivity was much higher than the base case VCR.

For the low VCR sensitivity, the Panel excluded large commercial and industrial customers from the VCR results. This is because large customers tend to have a higher VCR due to the commercial consequences of an outage, and the Panel considered that residential customer segments, which have a lower willingness to pay, may be disconnected first in the case of rotational load shedding. This lower sensitivity was slightly lower than the base case VCR.[81]

The high VCR sensitivity significantly affected the cost curve and suggested a significantly tighter reliability standard of 0.001 per cent expected USE. The results of the low VCR sensitivity were very similar to the base case. The Panel used the base case for their final decision on the level of the standard, as they considered the value used in the high VCR sensitivity did not truly reflect customers' willingness to pay, and would have led to an excessively high MPC.[82]

78 NER clause 3.9.3A; Reliability Panel, 'Review of the reliability standard and settings guidelines', 2021 Final Guidelines, p. 2, found here.

79 The VCR surveys customers about discrete outage durations of one, three, six and 12 hours (AER VCR final report 2019, p. 8, found here). Rotational load shedding typically disconnects customers for between 30 minutes and 2 hours at a time depending on the jurisdiction (see Appendix section B.6).

80 That is, customers' average willingness to pay to avoid a three-hour outage is less than three times that for a one-hour outage, as correctly noted by Endgame Economics' submission to the Issues Paper for the AEC, found here.

81 Reliability Panel, 2022 RSSR Final Report, p. 53, 108-109, found here.

82 Reliability Panel, 2022 RSSR Final Report, p. 54-58, found here.

Abbreviations

2021 Guidelines	Panel's 2021 Final Guidelines Review of the Reliability Standard and Settings Guidelines
AEC	Australian Energy Council
AEMC	Australian Energy Market Commission
AEMO	Australian Energy Market Operator
AER	Australian Energy Regulator
APC	Administered price cap
AUSM	AEMC USE Simulation Model
BESS	Battery energy storage system
CDD	Cooling degree days
CEC	Clean Energy Council
CER	Customer energy resources
CIS	Capacity Investment Scheme
CPI	Consumer Price Index
CPT	Cumulative price threshold
CVaR	Conditional value at risk
DNSP	Distribution Network Service Provider
ESOO	Electricity Statement of Opportunities
EUAA	Energy Users Association of Australia
HDD	Heating degree days
IASR	Inputs, Assumptions and Scenarios Report
IRM	Interim reliability measure
IRR	Interim reliability reserve
ISP	Integrated System Plan
LOR	Lack of reserve
MFP	Market floor price
MISG	Mathematics in Industry Study Group
MPC	Market price cap
NEL	National Electricity Law
NEM	National electricity market
NEO	National Electricity Objective
NER	National Electricity Rules
PASA	Projected assessment of system adequacy
PIAC	Public Interest Advocacy Centre
PoE	Probability of exceedance
PV	Photovoltaics
SRMC	Short-run marginal cost
RERT	Reliability and Emergency Reserve Trader
REZ	Renewable energy zone
RIN	Regulatory Information Notice
RRO	Retailer Reliability Obligation
RSSR	Reliability Standard and Settings Review
TNSP	Transmission Network Service Provider
USE	Unserved energy
VCR	Value of Customer Reliability
VRE	Variable renewable energy