The NEM Reliability Framework

November 2018

Additional information from AEMO to support its Enhanced RERT rule change proposal

Executive summary

AEMO has prepared this paper to support its Enhanced RERT rule change proposal. The paper considers the appropriateness of the existing NEM reliability framework in the context of the observed trends in the drivers of unserved energy. AEMO's findings are that:

Reliability is about ensuring that there is sufficient generation, demand response and interconnector capacity in the system to generate and transport electricity to meet consumer demand.

If there are insufficient resources the consequence will be **unserved energy (USE)** or load-shedding.

The risk of load shedding in the NEM is increasing due to:

- A tightening of the supply-demand balance following significant retirements of thermal generation, in the absence of sufficient new dispatchable capacity.
- The trend of increasing maximum temperatures leading to higher demands and lower supply due to de-rating of generation and transmission.
- The variability of renewable resources and the observed recent increase of forced outages at thermal plant.

Unserved energy can be characterised as a **tail risk** with a low probability of occurrence but potentially a high consequence. Whilst historically the NEM has experienced very little unserved energy (USE), a lack of USE in the past is not a good guide to future USE outcomes. This is because tail events are rarely experienced and because the supply-demand position has fundamentally changed.

The NEM reliability framework is not suited to increasing risk and uncertainty

The current NEM reliability framework must be able to deal with increasing tail risk but it is not well suited to this task. Instead it revolves around the comparison of an *average USE* measure of reliability to a single 0.002% reliability standard as the means of triggering a response.

The average USE measure suffers from the following issues:

- It does not adequately describe the shape and severity of the tail risk - other metrics such as conditional tail risk and "USE at risk" can provide better insights;
- It is only related to the cost of load shedding if all USE events are equally weighted. If there is a higher cost, for example, of longer USE events then this metric underestimates the cost of load shedding;
- It ignores risk aversion which is counter to most evidence of human behaviour. The usual risk averse approach to dealing with tail risk is to procure **insurance** and the reliability framework should focus on the appropriate level of insurance that RERT and other mechanisms can provide.

The reliability framework should incentivise the optimal resource mix with the lowest cost

A range of resources including demand response and distributed energy resources can play an important role in preventing and mitigating USE. The reliability framework should aim to encourage the deployment of the optimal resource mix with the lowest overall cost through both market and non-market means.

However, the infrequent and uncertain nature of tail events means that there is a lack of certainty of whether their costs can be recovered from the market, particularly for resources that have relatively high fixed purchase and establishment costs. For non-market revenue there is also investment uncertainty as to whether the reliability standard will be met, and whether RERT will be required again in subsequent years. Both of these factors can hamper investment and lead to overall higher costs than otherwise would be the case.

Proposed changes to the reliability framework

To address these issues, it is recommended that the procurement of RERT should be delinked from the reliability standard and that a **standing reserve** be created to provide the insurance function in the overall reliability framework.

The reliability framework should set the *level* of the required standing reserve over a defined horizon (akin to determining the sum to be insured) by taking account of:

- the nature of the tail risk - using a range of supplementary metrics;
- the risk appetite for different levels of load shedding expressed both in cost and limits terms;
- the cost structure and optimal mx of resources that can prevent or mitigate load shedding.

This approach would lead to a more stable investment environment that provides greater certainty to developers of resources as well as reducing the risk of load shedding.

Contents

Tables

Figures

1. Purpose

The Australian Energy Market Operator (AEMO) provides the following information on the appropriateness of the reliability framework in the National Electricity Market (NEM), as Supplementary Material to AEMO's Electricity Rule Change Proposal: Enhancement to the Reliability and Emergency Reserve Trader.

2. The NEM Reliability Framework

2.1 Elements of a Reliability Framework

A reliability framework is an approach to ensuring that there are sufficient resources in the electricity system to meet the required level of reliability.

A typical framework comprises a number of elements including:

- **Reliability Measure** – a metric to quantify the reliability of the system.
- **Reliability Standard** – an articulation of the required level of the reliability measure.
- **Reliability Response** – the range of actions that are triggered by monitoring the reliability measure against the reliability standard e.g. procurement of resources.
- **Governance** – the approach to governing, monitoring and changing the reliability framework.

This framework is illustrated in Figure 1 below:

Figure 1 Reliability framework

2.2 The NEM reliability framework

The Reliability Panel defines reliability in terms of having sufficient generation, demand response and interconnector capacity in the system to generate and transport electricity to meet consumer demand[1].

The NEM reliability framework consists of the following elements:

- **The Reliability Measure** – expected unserved energy[2] (USE).
- **The Reliability Standard** - 0.002% of the total energy demand in that region for a given financial year.

[1] Definition from AEMC Reliability Panel – Annual Market Performance Review 2017 Final Report – March 2018

[2] In practice, "expected USE" is interpreted and implemented as "probability weighted average USE" according to its statistical definition. For example, see

Reliability Panel, *Reliability standard and settings review 2018 Final Report*, p2

- The reliability standard is related to the **reliability settings**, which are a set of parameters that limit the range of market prices:
  - The Market Price Cap (MPC) sets an upper limit on spot prices in the wholesale market. The current level is $14,500/MWh in financial year 2018-19.
  - The Cumulative Price Threshold (CPT) imposes a limit on sustained high prices in the wholesale market. The current level is $216,900 for 336 trading intervals in financial year 2018-19.[3]
  - The Administered Price Cap is the 'default' price cap that applies when the CPT is exceeded, the current level is $300/MWh.
  - The Market Floor Price sets a lower limit on spot prices in the wholesale market, the current level is -$1000/MWh.
- **Reliability Response** – AEMO forecasts USE through its Electricity Statement of Opportunities (ESOO) and Medium Term Projected Assessment System Adequacy (MTPASA) processes and these can lead to the following responses:
  - within market, including additional generator capability being made available to the market or new generation capacity investment (long term).
  - AEMO rescheduling transmission outages, or intervention including procuring capacity through Reliability and Emergency Reserve Trader (RERT) contracts, direction or instruction (eg to generate or to shed load).
- **Governance** – The reliability framework is governed by the AEMC's Reliability Panel (the "Panel") operating under the National Electricity Rules.

The reliability standard and settings are reviewed by the Panel every four years[4]. In reviewing the reliability standard, the Panel is guided by the following general principles[5] to meet the National Electricity Objective (NEO):

- 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 stable, predictable and flexible regulatory framework.

[3] Both the MPC and CPT are indexed by Consumer Price Index

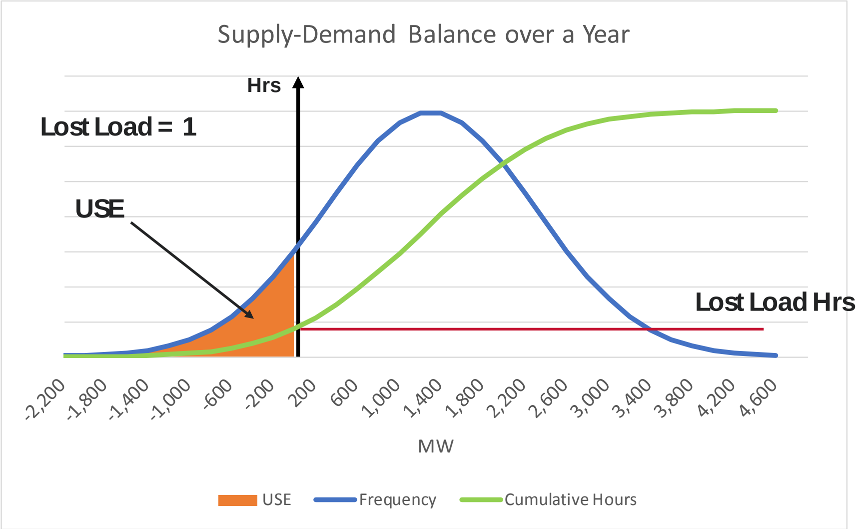
[4] Clause 3.9.3A (d)

[5] Reliability Panel, *Review of reliability standard and settings guidelines,* 1 December 2016

2.3 Measuring reliability

Reliability is a quantification of the *risk* of load shedding occurring and is calculated using forecasts of the supply-demand balance. Figure 2 illustrates the relationship between the most common metrics:

Figure 2 Common metrics for reliability



The supply-demand balance is forecast for each period of the year for a range of scenarios that take into account different weather conditions and the varying availability of supply (due to outages and de-rating). For each scenario the following metrics are calculated:

- USE = Unserved Energy in MWh = orange shared area under the supply-demand curve
- Lost Load Hours = number of hours in the year where there is any amount of unserved energy = intersect of the y-axis with the green cumulative number of hours curve.
- Lost Load Outcome = 1 or 0, depending on whether there has been any unserved energy in the scenario.

The reliability metrics are calculated across the full range of scenarios:

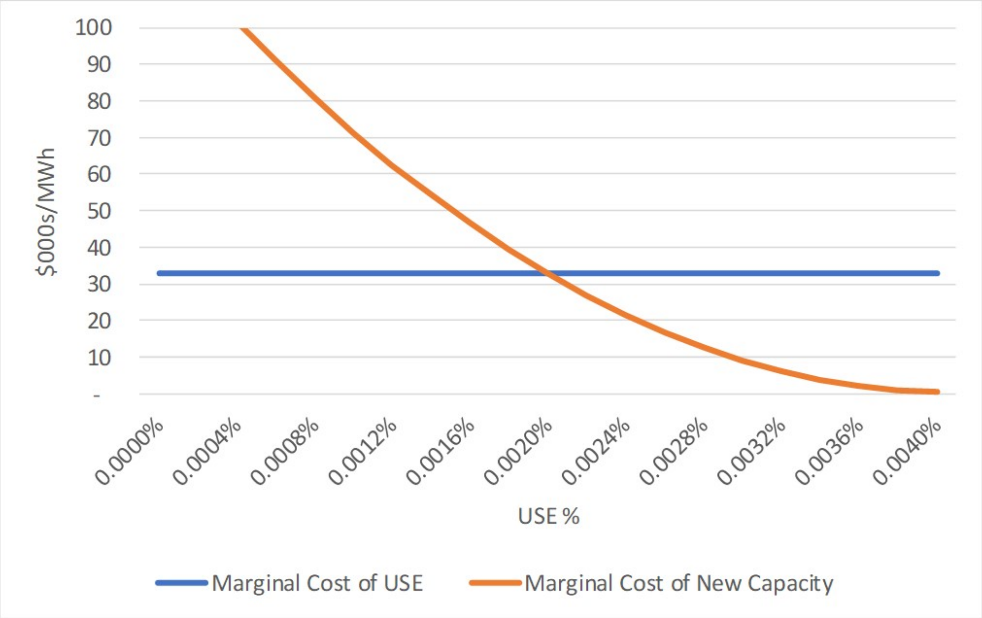
- **Average USE** = average of USE across all scenarios.
- **LOLE** = Loss of Load Expectation = Average of Lost Load Hours across all scenarios.
- **LOLP** = Loss of Load Probability = Average of Lost Load Outcomes across all scenarios.

Variations on this theme include metrics like 1 in 10 LOLE which is the average of lost load hours over a subset of scenarios for weather that is only expected to occur once every ten years (“P10”).

2.4 Setting the reliability standard

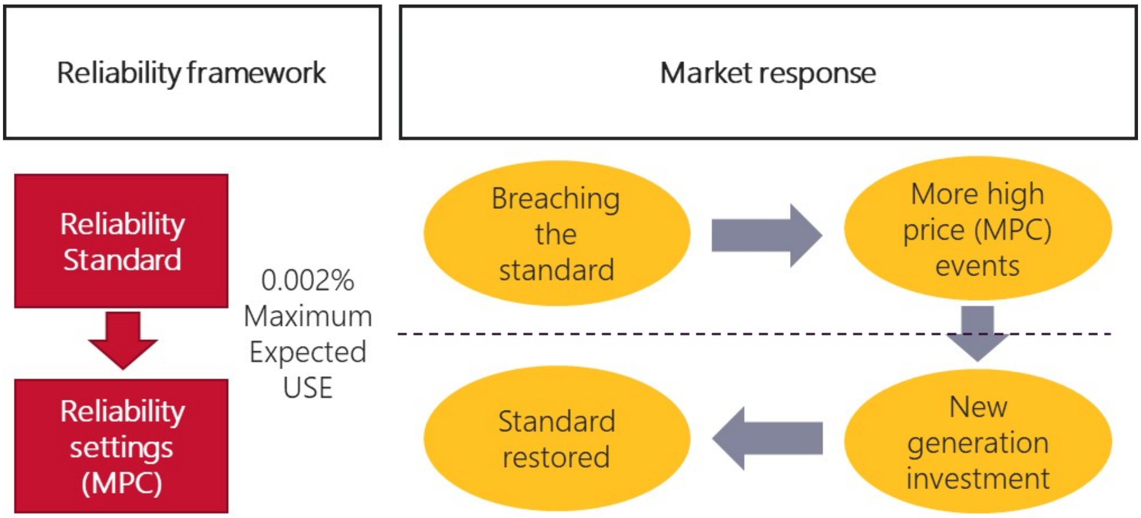
The reliability standard aims to reflect the economically efficient level of reliability to be provided by the market by balancing the benefit of reliability of electricity supply to customers and the cost of generation associated with its provision. This is shown in Figure 3: At higher levels of reliability (represented here by USE decreasing along the x-axis from right to left), the marginal cost of generation increases, until it exceeds the marginal benefit of increased reliability. The point where the marginal benefit and cost curves intersect is the theoretically efficient level of reliability.

Figure 3 Efficient level of reliability



In an energy only market, new investment is made on the basis of recovering capital costs from expected spot market prices[6] in the future. The reliability settings and in particular, the MPC, are set so that they provide sufficient investment incentive for new entrant generators to ensure the reliability standard can be delivered.[7] In theory, in an energy only market, a breach of the reliability standard causes MPC events to be frequent enough for new generation capacity to profitably enter the market. Figure 4 shows the linkage discussed above.

Figure 4 Reliability framework and generation investment



The current form of the NEM reliability standard and the 0.002% threshold were set in 1998[8] and have remained virtually unchanged[9]. While the Reliability Panel acknowledged at the time that the 0.002% level

---

[6] Or selling contracts on the basis of expected future spot prices.

[7] The reliability settings also limit the financial risk of market participants from excessive price volatility on the market. Reliability Panel, *Review of reliability standard and settings guidelines, Final Determination,* 1 December 2016, p 27.

[8] Reliability Panel, *Determination on reserve trader and direction guidelines,* June 1998, pp 5-9,

[9] The only exception was the expected USE has changed from "over ten years" to "a given financial year".

is an “on balanced” decision “to introduce a common approach across the National Market”[10], subsequent reviews have not revised this number as the Panel considered that the value is broadly consistent with AEMO’s value of customer reliability (VCR) estimate and the marginal generator being a gas turbine[11].

2.5 International comparison

This section presents a comparison of reliability standards across jurisdictions internationally. Whilst standards are set in terms of common reliability measures, implementation can vary depending on a range of factors including input assumptions, modelling methods, choice of sensitivity analyses, and other factors.

Table 1 summarises the different metrics that are used in reliability frameworks across international jurisdictions. Reliability standards help determine market price caps and other settings in energy only markets, and capacity requirements in capacity markets.

Table 1 Comparison of reliability metrics used internationally

Metric	Annual Standard	Jurisdiction	Supplementary Requirement	Market Type
USE	**0.002 %**	WEM (Aus)	Reserve margin = greater of 7.6% or largest unit	Capacity
		NEM (Aus)		Energy only
	300 MWh (0.0005%)	AESO (Alberta, Canada)		Energy only
1 in 10 LOLE	**2.4 hours**	NY-ISO, PJM, ISO-NE (US)		Capacity
		ERCOT (Texas)	Non-binding 13.75% reserve margin[12]	Energy only
	3 hours	National Grid (GB)	Sufficient capacity for a 1 in 10 year winter peak	Capacity
	3 hours	RTE (France), Elia (Belgium)	< 20 h lost load 95% of the time	Capacity
	8 hours	EirGrid (Ireland), Portugal	Index of load served > threshold 95% of the time	Energy only
LOLP	**4 %**	NWPCC (US)		Capacity
	15 %	OCCTO (Japan)	Based on 0.3 days/month LOLP during peak periods	Energy only
No formal requirement		Germany, Nord Pool, CAISO (US)	Various bespoke metrics.	Capacity

[10] Reliability Panel, *Determination on reserve trader and direction guidelines,* June 1998, p 8.

[11] Reliability Panel, *Reliability standard and settings review 2018,* 30 April 2018, pp14-16

[12] ERCOT is moving towards economically optimum and market equilibrium reserve margins in lieu of static reserve margins.

In comparing the international approaches to the NEM it can be seen that the USE metric is relatively unusual and most jurisdictions have adopted simpler LOLE and LOLP metrics. Recognising that load shedding is most correlated with extreme cold or hot weather some jurisdictions focus their metrics on these conditions e.g. the 1 in 10 LOLE metric used in the US. This has the added benefit of reducing computational effort and targeting the discussion of reliability on the highest risk periods.

Whilst LOLE and LOLP can be easier to understand, unlike USE, they do not provide any information on the amount of unserved energy if an event is expected to occur. To overcome this draw-back they are often considered alongside supplementary metrics to present a fuller picture of the risk of unserved energy.

The jurisdictional comparison also reveals that the NEM standard is generally not as restrictive as standards in other countries. For comparison, the 2018 ESOO USE forecast in Victoria over 2018-19 is close to the NEM's 0.002% standard but translates into a 1 in 10-year LOLE of 7.2 days (compared to a standard of 2.4 days in the US) and an LOLP of 31%.

3. Unserved Energy in the NEM

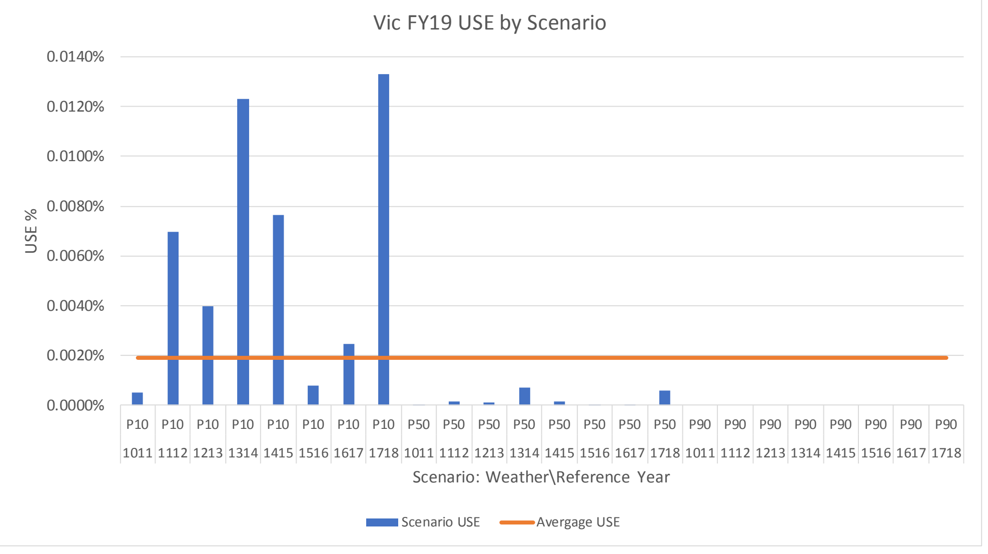
3.1 Measurement of USE

In ESOO and MTPASA, AEMO measures USE on an ex-ante basis by forecasting the distribution of the hourly supply-demand balance in each NEM region for a range of simulations across a multiple-year horizon. If supply is less than demand in any hour this creates USE. The ESOO reports the weighted average USE metric 100 simulations for each of 8 reference years weighted across the P10, P50 and P90[13] weather scenarios.

Figure 5 illustrates the components of USE in the 2018 Electricity Statement of Opportunities (ESOO) forecast for Victoria for financial year 2018-19. The weighted average USE is 0.0019% which is just below the 0.002% standard.[14]

The chart shows that USE is highly skewed to the P10 scenarios with limited USE in the P50 scenarios. All but two of the eight reference years in the P10 scenarios produced USE that was above the 0.002% threshold and some were considerably higher, being up to seven times more than the standard.

Figure 5 Distribution of Victorian USE for FY19 in ESOO 2018

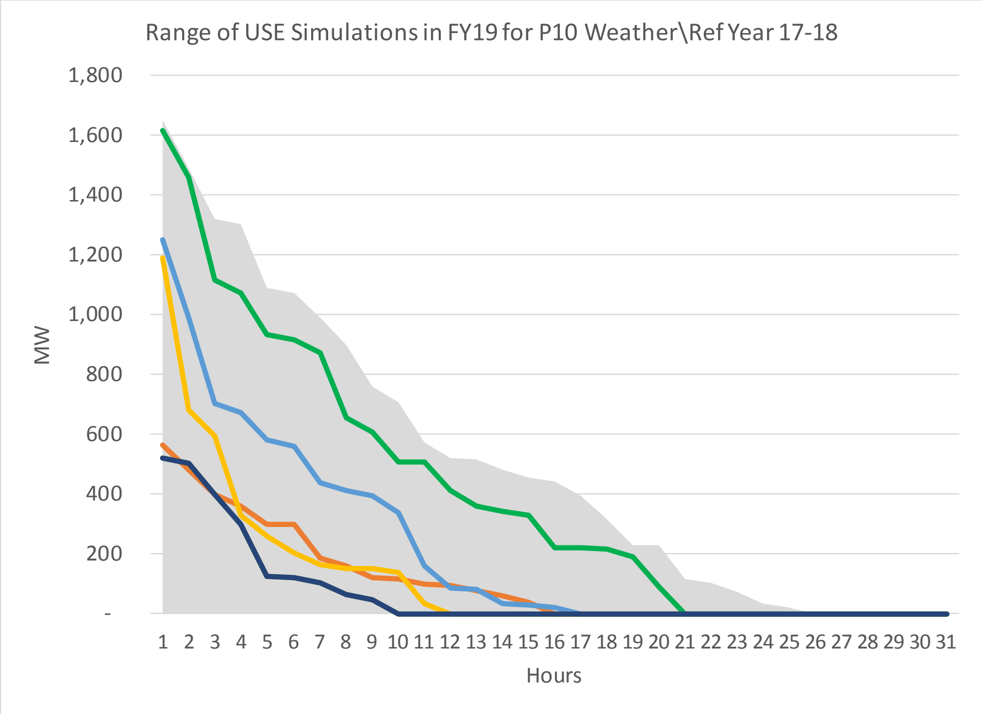


The USE metric for each of the 24 weather/reference year scenarios is itself an average compiled from the USE outcome for each of 100 simulations. The distribution of USE from these simulations is shown in Figure 6 for the scenario representing P10 weather and the 17-18 reference year. This scenario has an average USE of 5,913 MWh or 0.0133%.

[13] P10, P50 and P90 refer to the weather conditions that are likely to be exceeded only once every 10 years (P10), once every 2 years (P50) and nine times every ten years (P90).

[14] AEMO's current MTPASA modelling shows the average USE will be more than 0.002% in Victoria for financial year 2018-19. This is because compared to 2018 ESOO, the MTPASA modelling has incorporated additional outage information that will likely arise in the 2018-19 summer.

Figure 6 USE Duration Curves for P10 Scenario/1718 Refence Year

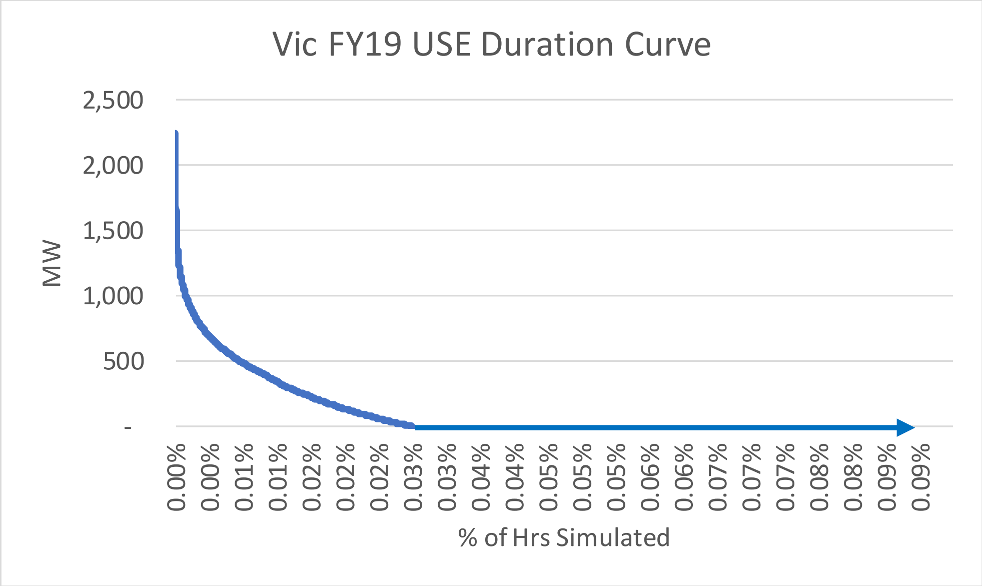


The chart shows several individual USE duration lines (i.e. USE outcomes for an individual simulation ordered from highest USE to lowest USE) and the grey shaded area shows the range of all outcomes from the 100 simulations.

Whilst the average USE is 5,913 MWh in this scenario, there are some simulations with very high amounts of USE (up to 1,600 MW) and some with very long durations (up to 25 hours). This wide range of outcomes is not apparent when considering the single average USE statistic.

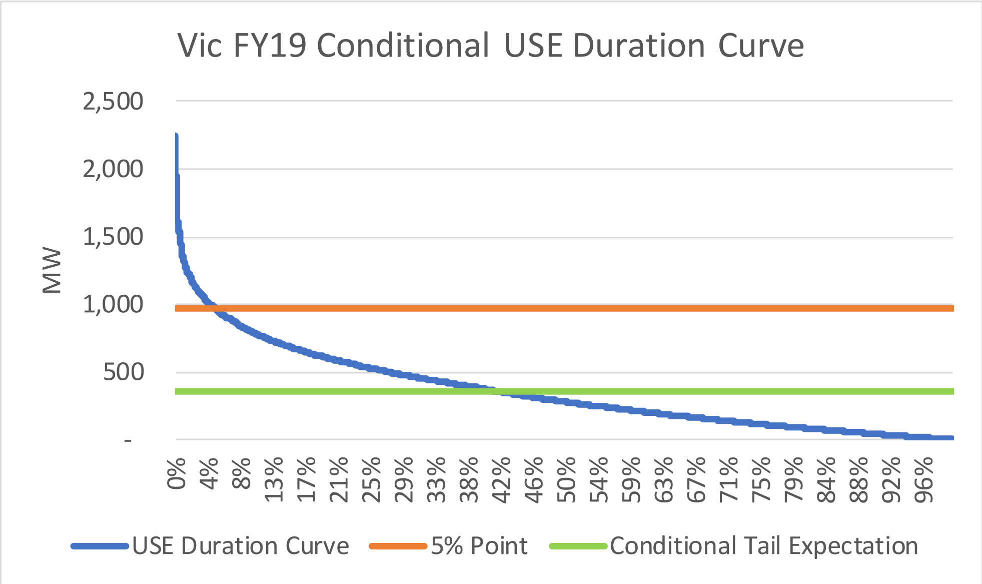
The full extent of the USE tail can be appreciated by combining all the simulations from each of the 24 weather/reference year scenarios in Figure 7. The chart shows that the full distribution of USE forms a very sharp tail with some simulated hours having more than 2000 MW of USE due to multiple, independent, generator outages coinciding with high demand driven by hot weather. However, the tail has a low probability of occurring with just 0.0269% of all simulated hours producing USE.

Figure 7 Full USE Duration Curve for Victoria FY 2018-19



Another way of describing a tail distribution is to focus on metrics that are *conditional* on the occurrence of an event in the tail. The Conditional Tail Expectation (or Tail Conditional Expectation) is the average MW of USE *providing that a USE event occurs*. In Figure 8, it is shown by the green line which is at 363 MW. The other metric is the 5% point of the distribution which is at 977 MW. The interpretation of this metric is that if 977 MW of resources were procured they would prevent USE in all but 5% of the cases when there is a USE event.

Figure 8 Alternative USE Statistics



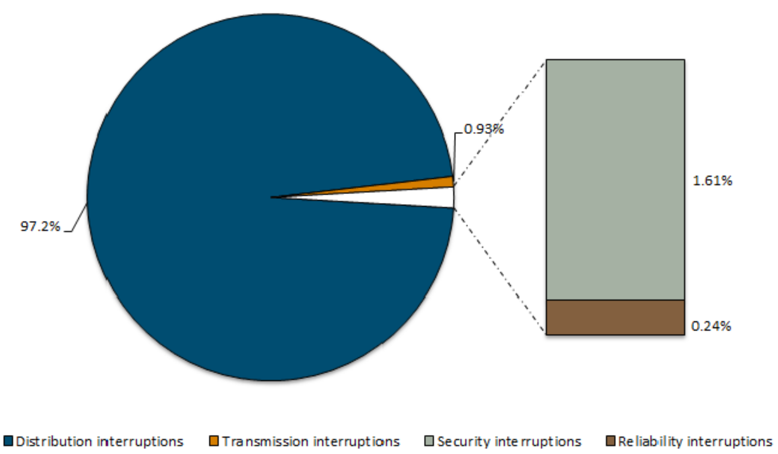
Metrics such as these can provide more detail on the nature of the USE risk. In fact, the average USE measured in MWh can be broken down into two metrics which describe the *size* and *likelihood* of USE.

Average USE (MWh) = Conditional Tail Expectation (MW) * LOLE (hrs)

3.2 Why we have not seen much USE in the past

Historically, the NEM has experienced very little USE that can be attributed to reliability causes (i.e. insufficient investment in supply to meet demand). The USE that has occurred has primarily been the result of distribution events as is shown in this extract from the AEMC's 2017 Annual Market Performance Report. This is not surprising. Over the last ten years the NEM supply-demand position has been generally favourable, driven by excess generation supply combined with flat to falling maximum demands.

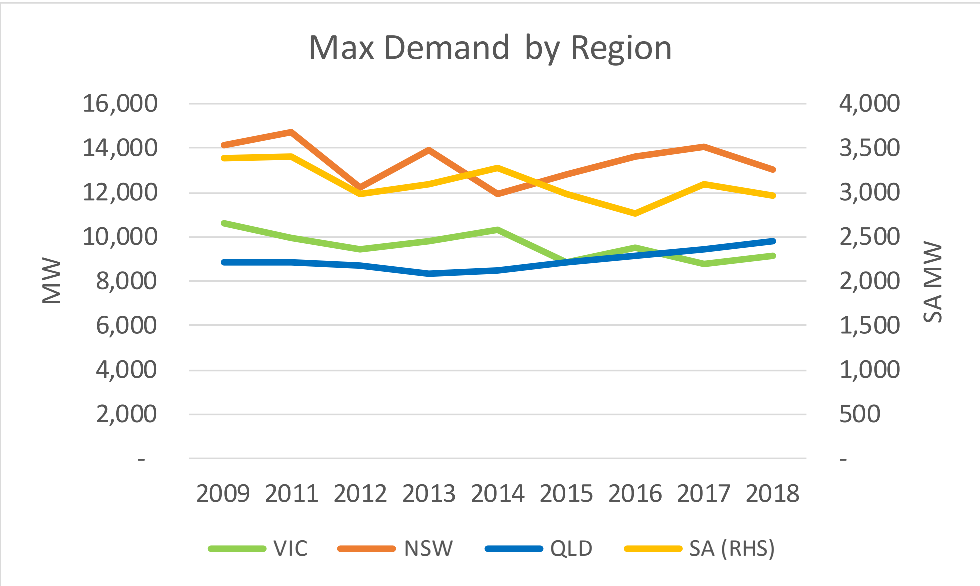
Figure 9 Extract from AEMC Market Performance Report



3.2.1 Maximum demands have been falling

In recent times, with the exception of Queensland, maximum demands have been flat to falling across most NEM regions:

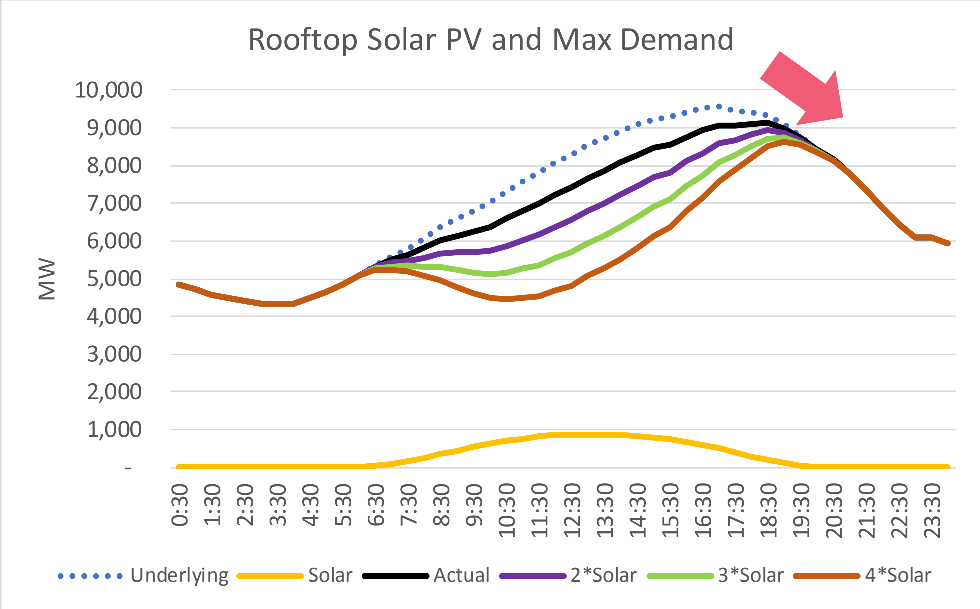
Figure 10 Maximum Demands by Region



The trend in maximum demands has been driven by a combination of two key factors: uptake of rooftop solar PV and entry or exit of large industrial loads.

The growth of rooftop solar PV has led to an increasingly peaky operational demand curve on sunny days. Figure 11 illustrates the impact of increasing penetration of rooftop solar PV on the maximum operational demand. As solar PV increases, the operational demand curve steepens and then hollows out into a "duck-curve" shape. The maximum of the underlying demand is at the 17:00 EST trading interval but is shifted to 18:30 by the impact of solar PV. Future growth in solar PV moves the maximum demand further downwards and later into the evening.

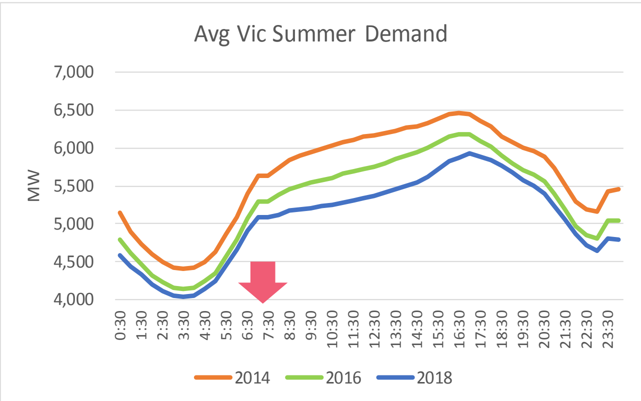
Figure 11 Impact of Rooftop Solar PV on Operational Demand [15]

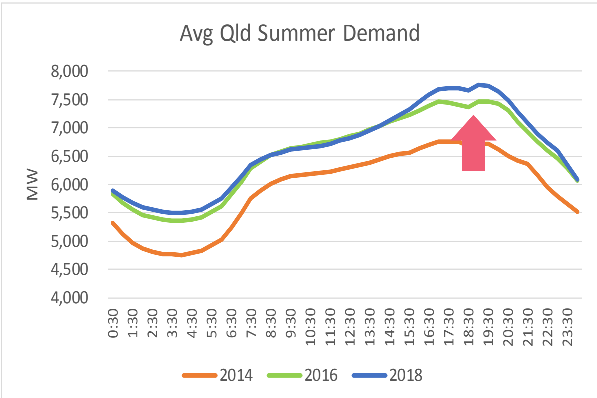


Whilst solar PV is beneficial in reducing the maximum demand it also creates increased risk in managing the system. The impact of cloud cover on a very hot day is to reverse this trend reducing the output from solar PV and causing the maximum demand to shift higher and earlier in the day.

The other key driver of the trend in maximum demands is the impact of large industrial loads. This is most evident in the change in the overnight load shape.

Figure 12 Impact of Large Industrial Loads





Victoria experienced the loss of the Point Henry smelter in 2014 which reduced system load by ~200 MW. By contrast Queensland maximum demands have been increasing and the key reason has been the increase in LNG and other large mining demand which has led to a general increase in demand of 600-700 MW since 2014.

Whilst it is likely that maximum demands will continue to be suppressed by increasing solar PV penetration this will be at a lesser rate than in the past as the time of the peak shifts to later in the day. The impact of changes in industrial demand are hard to predict as it is often driven by commercial decisions of individual loads.

However, set against the declining trends in maximum demands is the continued increase in demand driven by general economic growth and increasing population. Looking forwards, the ESOO projects that maximum demands will increase for Victoria and South Australia whilst they will fall in New South Wales and Queensland before growing again from around 2021.

[15] Chart shows actual Victorian operational demand on 18th Jan 2018. Underlying demand adds back the solar PV to the operational demand. Scenarios are shown for the operational demand with 2, 3 and 4 times the existing solar output.

3.2.2 Until recently the NEM has been over-supplied

For much of the last ten years the NEM has been over-supplied and low wholesale prices were prevalent throughout much of this period. This has contributed to a succession of closures of thermal generation culminating in the 2016 closure of Northern Power Station in South Australia and the March 2017 closure of Hazelwood Power Station.

Following these retirements, the supply-demand balance has tightened significantly with an associated sharp lift in wholesale prices. The outlook for USE has also changed radically since the 2014 ESOO when AEMO forecast no unserved energy anywhere in the NEM across the whole of the 10 year horizon. For the 2017-18 summer AEMO forecast USE in excess of the reliability standard and procured RERT and dispatched them for the first time in the history of the NEM. AEMO will also be procuring RERT for the 2018-19 summer.

3.2.3 Summer 2017-18 weather was favourable for USE

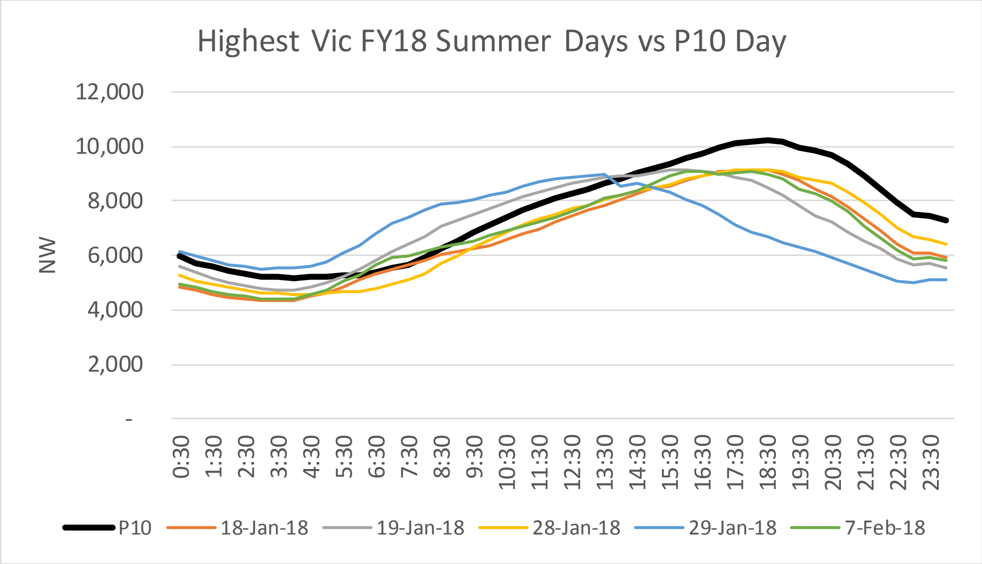
AEMO exercised RERT on two occasions over the summer of 2017-18 on 30 November and on 19 January . Whilst Victoria's summer was warmer than average, the hottest day of 41.7^{o}C occurred on a Saturday early in January so the maximum demand was not high as it would have been on a weekday. The five days of highest demand were as follows:

Table 2 Five Highest Victorian Demand Days Summer 2017-18

Date	Weekday	Max Demand (MW)	Olympic Park Temp ^{o}C
18 Jan 2018	Thu	9,125	40
19 Jan 2018	Fri	9,153	40.3
28 Jan 2018	Sun	9,144	38.1
29 Jan 2018	Mon	8,953	33.5
7 Feb 2018	Wed	9,102	37.4

The load shape over these days is shown on the following chart compared to the P10 load shape which peaks at 10,239 MW.

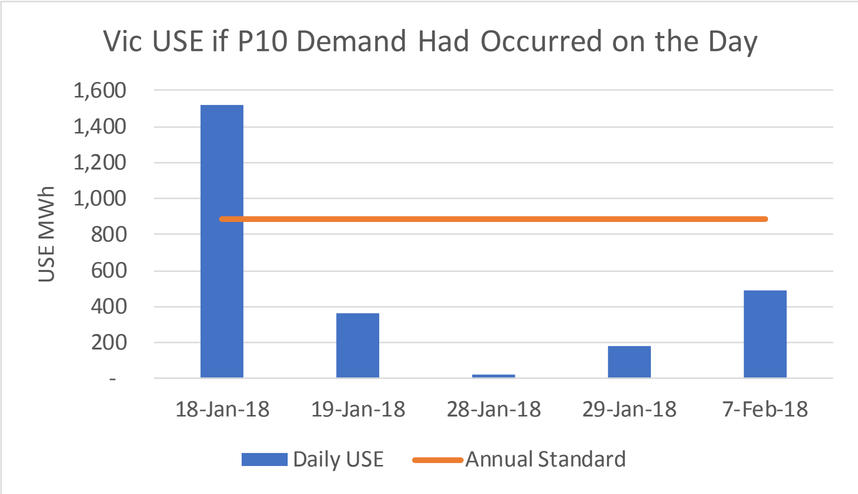
Figure 13 Highest Demand Load Shapes vs P10 Demand



The 28 January saw record demand for a non-workday and 29 January was on track for a high demand before a cool change swept through in the afternoon.

It is possible to restate what the USE would have been if P10 temperatures had prevailed (around 43.5°C) on any one of the 5 highest demand days by comparing the reserve conditions during the day with the additional demand resulting from a P10 demand. This is shown in Figure 14.

Figure 14 Counterfactual USE if POE10 demand occurred



On 18 January the temperature was close to 40°C corresponding to a P50 day. If, however, the day had turned out to be a P10 day the resultant USE would have breached the *annual standard* in just one day. This is driven by the additional P10 demand being more than 1000 MW higher than the actual demand on those days and there being very low reserves available. On 19 January the reserves could have been even lower as Basslink was in danger of being fully de-rated due to the temperature at George Town, Tasmania approaching Basslink's operating limit.

Looking back at the summer of 2017-18 shows how precariously the system is balanced in Victoria. Overall, whilst the summer was warmer than average and there were some hot days it was fortunate that the timing of the hottest days was such that there was no USE.

4. Increasing Tail Risks

4.1 USE is a function of tail events

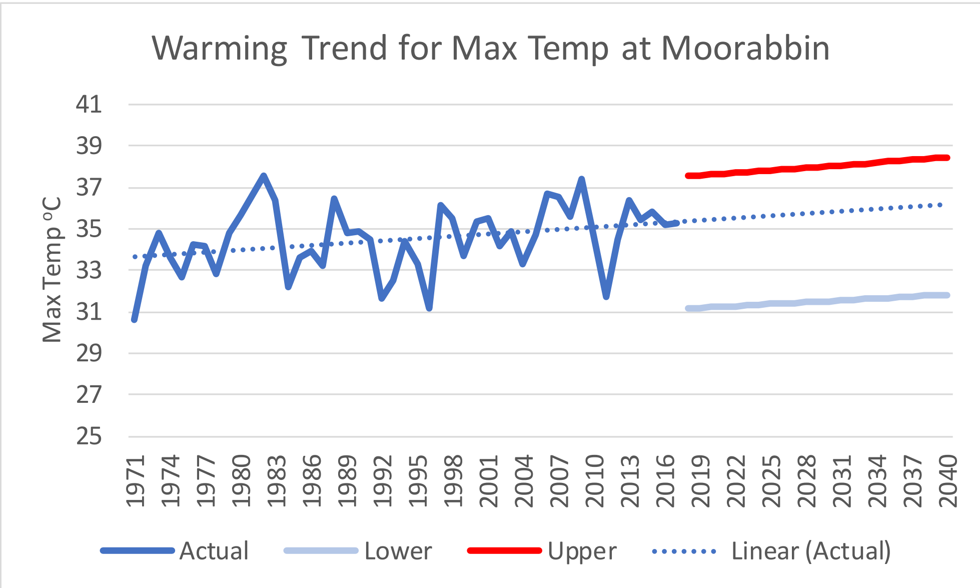
The previous section highlighted that the probability distribution of USE is skewed towards tail events. Whilst USE has a low probability of occurrence there can be high consequences if an event does occur in terms of the magnitude and duration of lost load. Actual outcomes in the NEM are too small a sample to provide a good guide as to the risk of USE occurring in the future. Instead we need to look at the drivers of USE and understand how they are changing over time.

AEMO's forecasts of USE outcomes primarily result from high temperatures leading to high demand coinciding with supply shortfalls. Supply can be impacted through lower generation output and/or lower transmission availability. The key risk is of outages at thermal generation and/or transmission lines and lower than expected wind or solar generation output. At times of system stress, other factors such as the occurrence of bushfires or presence of lightning leading to loss of key transmission lines can also come into play.

4.2 Drivers of increasing tail risk

Recent data from the Bureau of Meteorology[16] shows that temperatures across the NEM are increasing and AEMO's ESOO forecasts assume this trend will continue. The ESOO's demand forecasts are based on a range of temperature trend scenarios that fall within a lower and upper range. This is illustrated in Figure 15 showing the warming trend for the maximum daily temperatures (averaged over top 5th percentile) at Moorabbin Airport in Melbourne. Other weather stations show similar trends.

Figure 15 Warming Trend for Max Temperatures at Moorabbin[17]

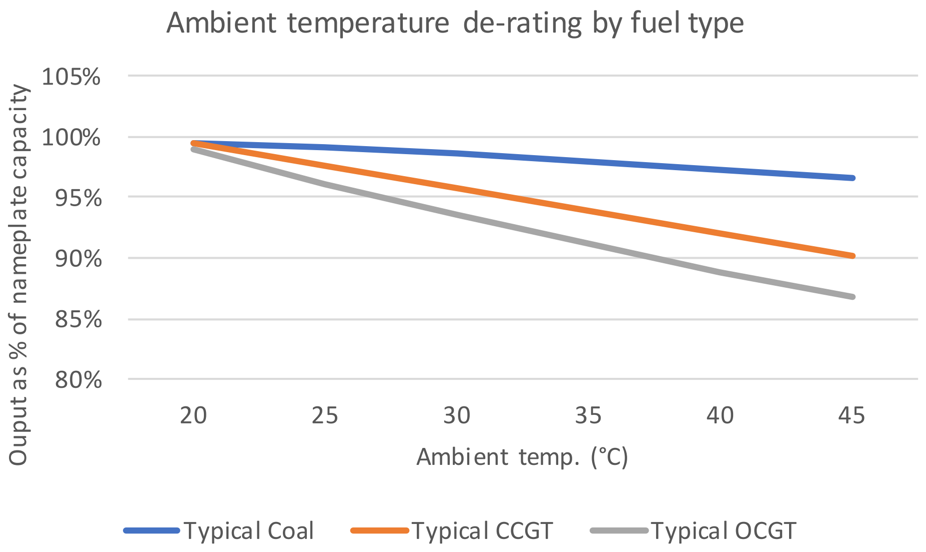


[16] Bureau of Meteorology, *Annual climate statement 2017*, 10 Jan 2018.

[17] From Bureau of Meteorology database

Clearly, rising maximum temperatures can lead to increasing electricity demand but what is less obvious is that they can also reduce supply which magnifies the risk of USE. The impact of temperature on thermal generation is shown in Figure 16.

Figure 16 Thermal Generator De-rating with Temperature

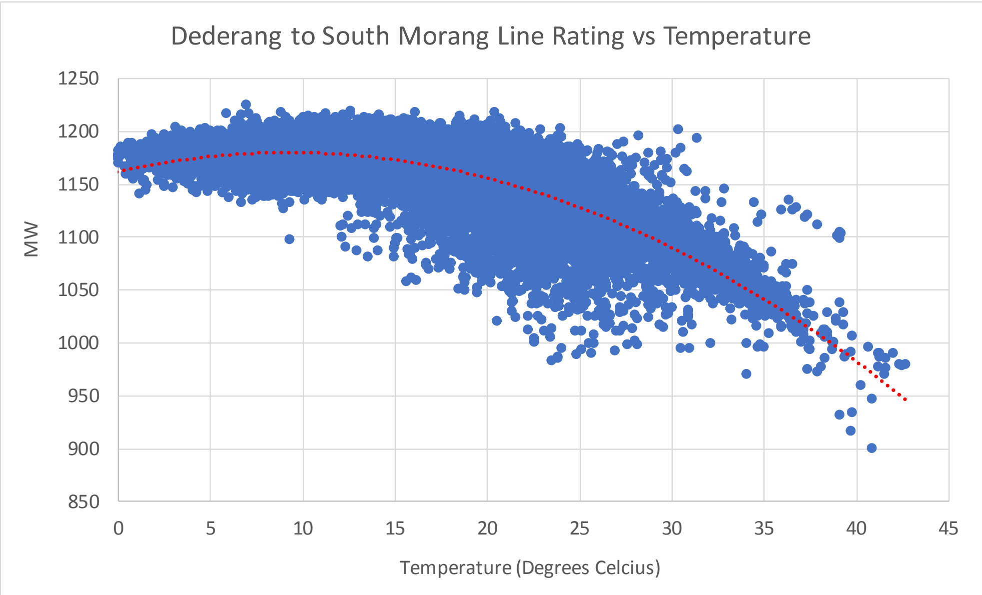


As temperatures increase, gas plant can be de-rated by around 10% on a 40°C day with peaking plant (OCGTs) being the most impacted. In fact, temporary diesel generators, like those installed in South Australia, are even more de-rated by more than 20%.

AEMO's forecasts take into account this de-rating with lower summer capacities used for all generators. For example, the difference in capacity between the Victorian winter and summer ratings can be over 500 MW.

Transmission and distribution lines are also impacted by temperature. Figure 17 shows the actual line rating for the key Dederang to South Morang line in Victoria which is de-rated by ~15% at 40°C.

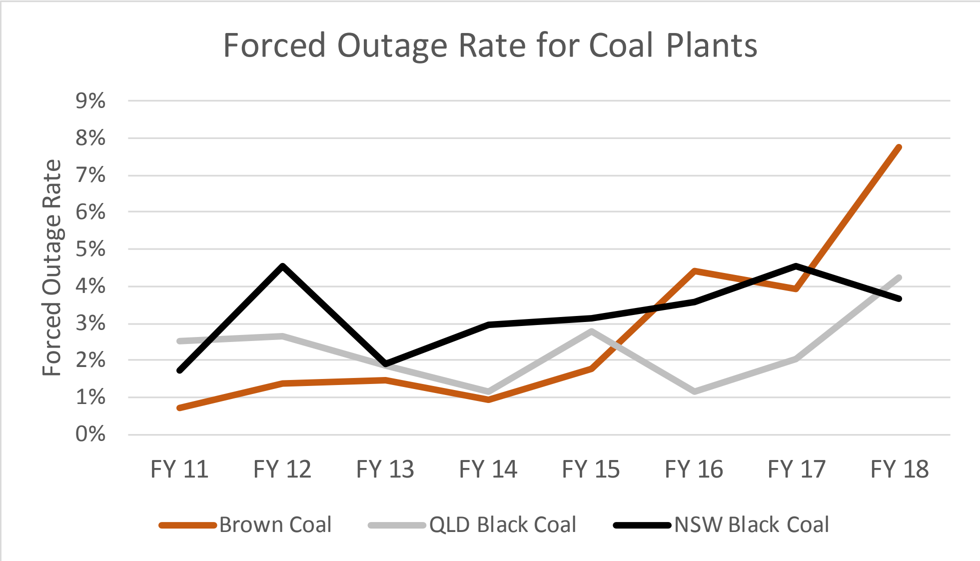
Figure 17 Transmission Line De-rating with Temperature



Basslink is another key transmission line that is impacted by temperature with the line fully de-rated once temperatures at George Town in Tasmania reach 36°C. As an example, on 18 January 2018, high temperatures were predicted for George Town for the following day and this led Basslink to reduce its capacity projection in line with agreed operational procedures. As it transpired the temperatures were lower than forecast on 19 January 2018 and Basslink's capacity reverted to 478 MW.

Whilst there is insufficient historical data to suggest that forced outage rates increase with temperature, there has been a general trend upwards in forced outage rates within the coal fleet over the last few years (Figure 18) and this is considered in AEMO's forecasts. Looking ahead, as thermal plant nears retirement there is the possibility that less capital will be invested in these stations which could lead to higher forced outage rates.

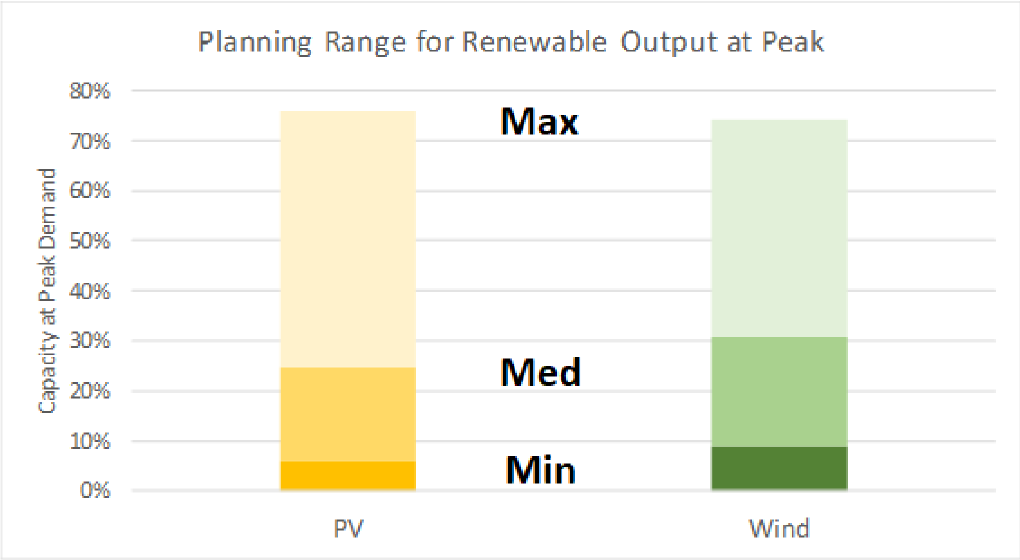
Figure 18 Forced Outage Rates for Coal Generators



The output of variable renewable energy generators at peak demand times is influenced by a wide range of factors. The impact of solar PV in reducing maximum demands and shifting them to later in the day has previously been discussed. Over time this will lead to a lower contribution of PV output to meeting peak demand simply because the maximum is occurring later in the day. AEMO's planning range, based on eight reference years, for rooftop PV output shows outcomes ranging from 6% to 70% of capacity coincident with peak demand.

Wind generation shows a similar wide range of outcomes with approximately 9% to 65% of installed capacity available at peak demand times. Although these results are based on a limited sample of solar and wind generation patterns over the last eight years, the results illustrate that AEMO is now managing a system that is much more uncertain and unpredictable than it was in the past.

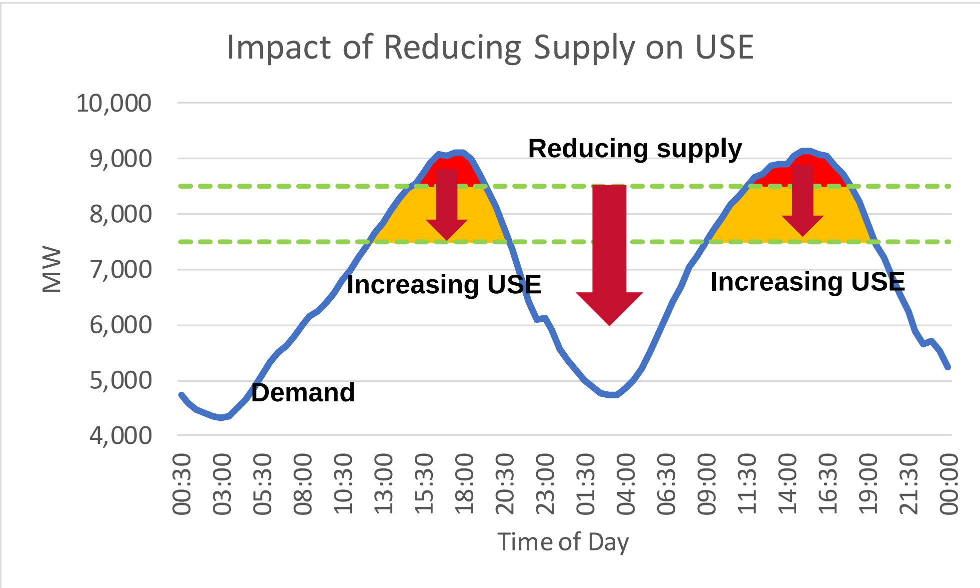
Figure 19 Renewable Capacity at Peak Demand Times for Victoria



4.3 Impact of these trends on USE outcomes

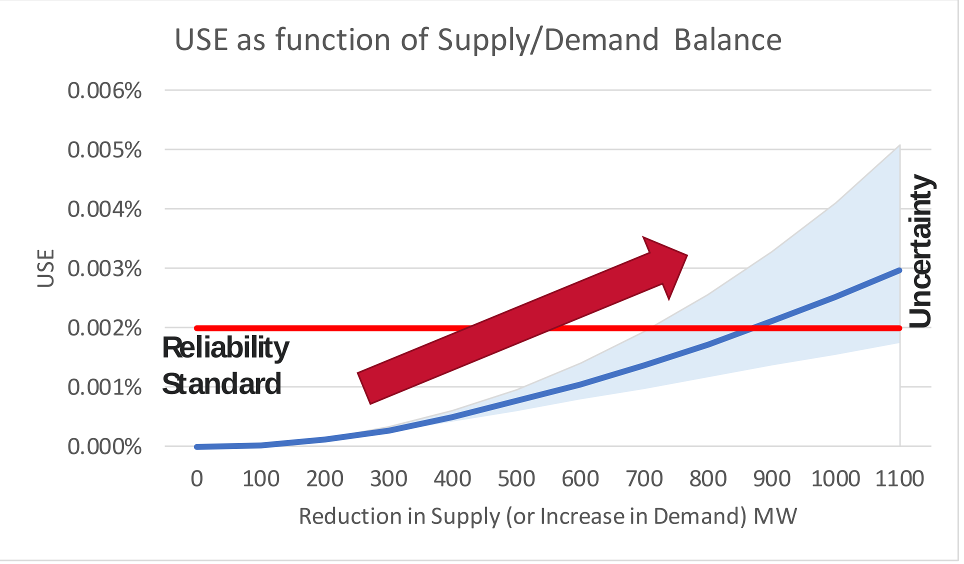
The culmination of all these factors is to increase both the likely amount of USE and the range of potential USE outcomes. As the supply-demand balance tightens there is a *non-linear* increase in the amount of USE as demonstrated in Figure 20. As the supply line is reduced (or alternatively as demand is increased) the amount of USE under the demand line (i.e. the shaded area) increases in a non-linear way.

Figure 20 Increase in USE as Supply-Demand Balance Tightens



Not only does USE increase in a non-linear way but the *uncertainty* of USE increases with a skewed distribution towards higher USE events as illustrated by a widening "funnel" of USE outcomes.

Figure 21 USE and USE Risk Increases as Supply-Demand Tightens



These impacts mean that it is quite plausible to have a very low risk of USE when the supply-demand balance is wide but for USE to quickly emerge as the supply-demand balance tightens.

5. Managing Tail Risks

5.1 Tolerance for load shedding

Managing tail risks invariably means accepting some level of load shedding but how much is acceptable? The usual approach to answering this question is to determine a cost of load shedding and then to compare that with the cost of additional resources to avoid load shedding.

The cost of load shedding is typically measured by the Value of Customer Reliability (VCR), which is the theoretical maximum amount a customer is willing to pay to avoid load shedding. Given there is a finite value for VCR, there exists a threshold so that if the magnitude, or the probability of load shedding is sufficiently small, it will no longer be efficient to build additional supply capacity and incur the upfront capital cost to avoid it.

The current VCR value was estimated by AEMO in 2014[18], and will be updated by the Australian Energy Regulator by the end of 2019. Using a survey-based approach, AEMO found the NEM wide aggregate VCR to be $33,460/MWh in 2014, which varied depending on customer segments, timing of USE and its duration.

Whilst determining the cost of load shedding can be a useful input in determining the tolerance for load-shedding, survey-based results should be used with caution when considering tail events, as people tend to anchor their views on recent experience. As USE is a very rare occurrence, respondents would likely find these events difficult to contemplate and hence may underestimate their impact.

Another way of approaching this is to seek input on the tolerance for load shedding in terms of the maximum acceptable duration and scale of an event. For example, the tolerance for being without air conditioning during a heatwave may be acceptable for a few hours but not beyond 6-8 hours. Similarly, a local event may be tolerated more than an event which blacks out the whole state.

Therefore, AEMO recommends that the AER's work on reviewing VCR also considers seeking consumer, business and government views on the maximum duration or scale of an event they would be willing to accept.

5.2 Optimal resource mix to manage tail risks

The selection of resources required to manage tail risk needs to strike the right balance between the cost/tolerance for load shedding, the performance characteristics of those resources and their cost structure. Different types of resources have different performance characteristics which may be more, or less, suited to the shape of the tail risk. A short, sharp tail could be addressed through a combination of load shedding and demand response (DR) whereas a longer, fatter tail may require resources with long durations like a peaking generator.

The cost structure of the resources is also critical in determining the overall, optimal mix of resources. Different types of resources have different cost structures. Some technologies like peaking generators have high fixed costs (capital and establishment) and relatively low variable, operating costs. Other types of technologies like industrial DR may have relatively low fixed costs but much higher variable costs reflecting the cost of lost production. The features of different types of resources are summarised in the following table.

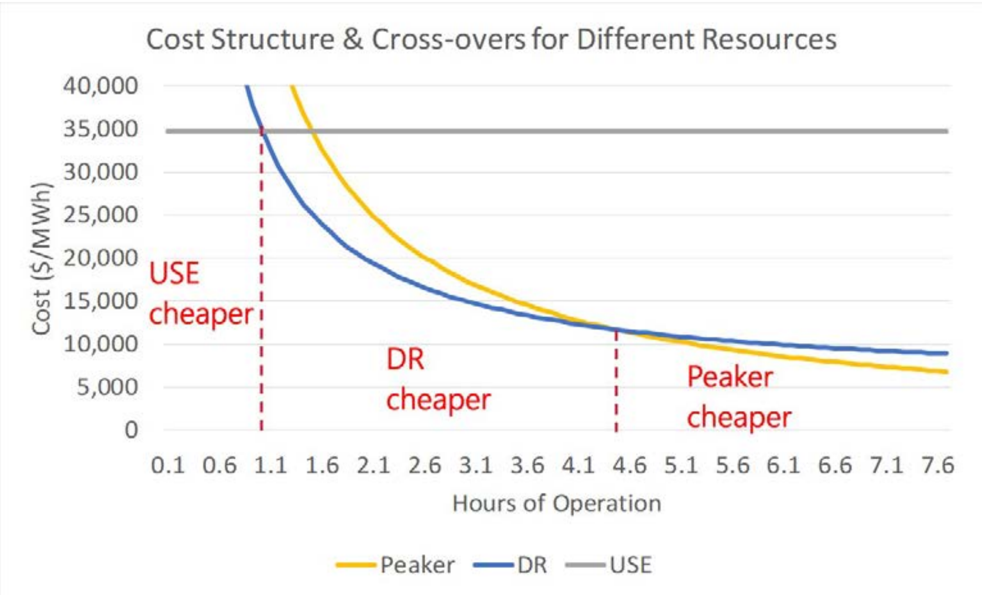
[18] AEMO, *Value of Customer Reliability Review,* September 2014

Table 3 Summary of resource mix to manage USE tail-risks

Possible Resource Mix	Typical Characteristics	Cost structure
Smelter Demand Response	Fast response, large MW Limited duration	potentially high variable cost
Industrial Demand Response	Slower response Limited duration	Low establishment cost. Potentially high variable cost
Batteries, VPPs	Fast response Limited duration.	High capital cost. Low variable cost.
Gas or diesel Peaker	Long duration, slow to build. Capacity de-rated during heat	High capital cost. Low variable cost.

Figure 22 shows the indicative trade-offs among different resources, where the numbers are for illustrative purposes only. As the duration of the potential USE event shortens, the optimal resources switches from technologies with high capital/low variable costs to those with low capital/high variable cost. When the duration is sufficiently short, the cheapest outcome is to leave some very small amount of residual USE.

Figure 22 Standard cross-over points for different resources



Note: Illustrative numbers

5.2.1 Impact of limited duration on the optimal resource mix

The optimal resource mix is dependent on the shape of the tail risk, the tolerance for load shedding and the cost and performance characteristics of the different types of resources. Typically, the types of resources used to mitigate USE are constrained by having a limited duration over which they can be used. For example, batteries are limited by their storage capacity, smelter DR is limited before it affects the potlines and even peaking generation may be limited by the capacity of onsite fuel supplies.

To examine the impact of these factors we have developed a simple model using the indicative cost inputs shown in Table 4. The model takes as input a sequence of days with unserved energy and aims to minimise the overall cost of mitigating this taking into account the cost of USE represented by AEMO's 2014 VCR estimate.

Table 4 Summary of resource mix to manage USE risks

Cost	Peaker	Battery	DR	USE
Variable Cost ($/MWh)	$96.5	$1,000	$31,000	$33,460
Capital Cost ($/MW/yr)	$95,000	$65,000	$3,000	$0

An illustrative year with four USE events was modelled as shown in the time sequential charts below. The events are of different magnitudes, durations and starting times. The model solves the resource mix for each hour to deliver the overall lowest economic cost.

The results of the modelling show that:

- **Unlimited Duration of Resources** - When there is no limitation on the resources, the optimal mix to manage the USE events consists mostly of batteries, supplemented by DR with high variable costs and some residual USE (Figure 23).
- **Limited Duration Batteries** - Figure 24 shows the change in the optimal resource mix if batteries can run for at most 2 hours in each potential USE event. The optimal mix replaces some batteries with peakers, which have higher operating costs, but are unconstrained (in this model).

Figure 23 Optimal resource mix without constraints

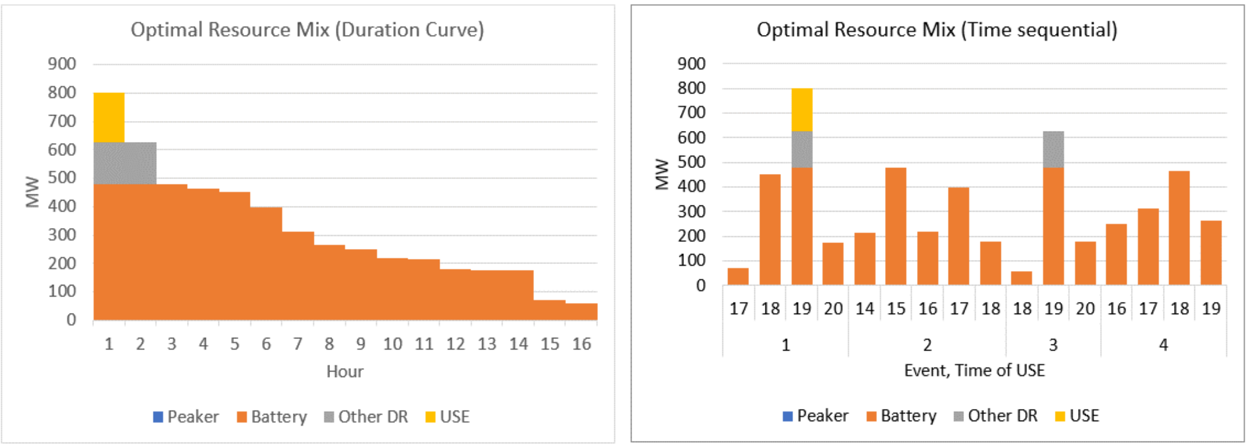
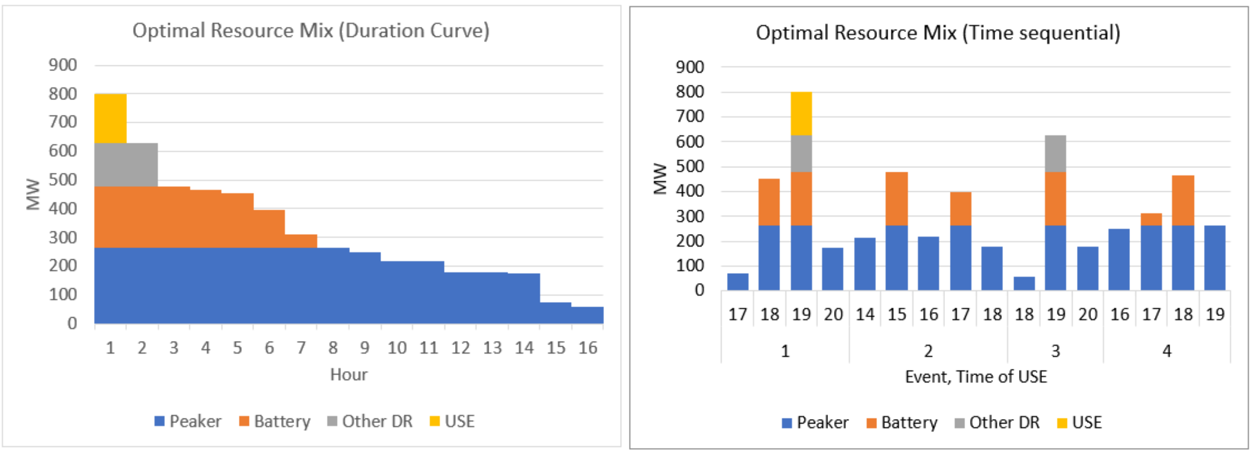


Figure 24 Optimal resource mix with battery duration constraints



5.3 Incentivising the optimal resource mix

The design of the NEM incentivises investment in resources through the signal provided by the energy price which is capped at the MPC. AEMO intervention via RERT acts as a supplementary mechanism when the market does not deliver sufficient resources to meet the reliability standard.

The reliability framework should enable the delivery of the optimal resource mix through both market and non-market mechanisms. However, there might be obstacles that prevent the market from delivering the optimal mix of resources needed to manage tail risks. In this case, a non-market procurement mechanism is crucial to ensure the optimal resource mix can be made available to deliver the efficient reliability outcome.

5.3.1 Role of demand response and distributed energy resources

DR and distributed energy resources (DER) can play an important role in both preventing and mitigating the risk of load shedding. DR and DER can be used for load-shifting to flatten out the demand curve and prevent USE occurring in the first place. Example of roles that they can play in preventing USE are:

- “Pre-cooling” can be used to cool a residence on a hot day so that less cooling is required later in the day at the time of system peak.
- Control of flexible loads e.g. operation of pool pumps could be shifted to low demand times.
- Distributed storage could be used to flatten out the evening peak on hot days.

Due to its low upfront capital cost, DR also plays an important part in *mitigating* potential USE events. There are two types of DR resources that can be used to improve the reliability of supply: wholesale and emergency DR.[19] Wholesale DR can directly participate in the wholesale electricity market and can be used by market participants to earn spot revenues or manage pool exposure on a commercial basis. Emergency DR, due to their high operating costs (typically above MPC) or other operating restriction, does not participate in the wholesale market, but is centrally controlled by AEMO to avoid involuntary load shedding and procured under RERT contracts[20].

Whilst DR and DER have value in both preventing and mitigating the risk of load shedding, the way these resources are rewarded by market and non-market mechanisms needs to incentivise the right outcomes. For example, pre-cooling to avoid using 1 MWh of electricity at peak demand times is just as valuable as storing 1 MWh for use during the peak, or switching off 1 MWh when the peak is occurring. It is not clear that the signals for each of these outcomes is currently aligned.

5.3.2 Barriers to procuring the optimal resource mix

The optimal mix should consist of a combination of technologies with different cost structures and operating characteristics. However, both the current market and non-market mechanisms may present obstacles to the delivery of the optimal mix. This is due to the following factors:

Gap between VCR and MPC

Currently the maximum price a resource can receive in the energy market (i.e. the market price cap) is less than the cost of load shedding (as measured by VCR). This means that whist there may be further economic benefits in reducing the expected amount of USE, resources with operating costs above the MPC (but below VCR) cannot recover their cost by entering the market. Raising the MPC to the same level as VCR, however, could significantly increase volatility in the energy market and increase the risk exposure of market participants who would seek to pass the additional cost of managing these risks onto consumers.

Uncertain and highly volatile tail events do not provide sufficient market revenue

Even if the MPC was set at the VCR, a potential new entrant still needs assurance that it can recover its capital cost, which can be significant for peaking plant and some types of DR resources. However, as tail events are by nature, rare and difficult to predict, new entrants with significant capital cost might be deterred from entering the market due to uncertainty in their revenue stream.

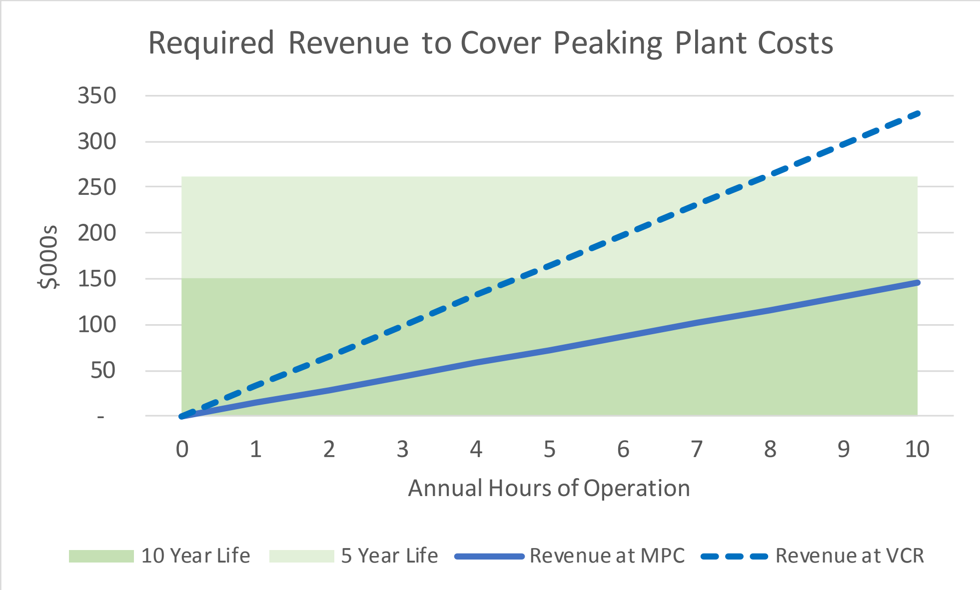
---

[19] The other two types of demand responses, network demand response and ancillary service demand response, are not directly related to reliability of supply and hence not discussed here.

[20] AEMC, *Reliability Framework Review, Interim Report*, p54 and pp 102-103.

Figure 25 illustrates the required hours of operation at different price caps to cover the costs of a peaking plant with either a 5 year or 10 year payback period. A peaking plant such as a diesel generator with a 10 year payback period would need 10 hours of operation at the MPC *every year* to recover its capital costs. Raising the MPC to VCR would reduce this to 5 hours every year. If the payback period were shorter (eg. 5 years) then even more hours of operation at the MPC are required.

Figure 25 Revenue required to recover fixed cost



The alternative to relying on unpredictable spot revenue is to sell cap contracts but these contracts are quite illiquid and the market for caps usually only trades two years out so it would be difficult to rely on a cap income stream to underpin an investment. Hence, the combination of certain fixed costs and uncertain and unpredictable revenues is likely to mean that the market will be reluctant to invest in these types of resources. The same analysis applies not just to peaking plant but also to any DR or DER investments where the fixed costs require several years of certain revenue to justify the investment.

Lack of certainty of RERT income mechanisms

The lack of certainty will also cause similar inefficiency with RERT procurement mechanisms. Currently, AEMO is only able to procure RERT nine months out, and only to ensure that the reliability standard is met (or to maintain power system security). Whilst the reinstatement of long notice RERT is helpful in ensuring there is sufficient lead time to procure the most economic resources that are available, it does not address the issue that providers of resources do not know if they will be required again in subsequent years.

This has two negative consequences:

- Providers of resources with high deployment and demobilisation costs require high availability or usage cost (or a combination of both) in order to recover their cost over a one-year payback period. This will tend to make them uncompetitive versus resources with low fixed costs but high variable costs.
- There may be under-investment in capacity to participate in RERT thus limiting competition, leading to inefficient resource mix and higher costs.

These issues were recognised in 2017 when ARENA and the NSW government committed funding of nearly $36m to promote investment in 200 MW of new DR.

Potential mis-alignment of risk preferences

In a decentralised energy market, participants undertake commercial decisions to manage their own risk, including investing in, and operating generation assets and trading hedging contracts. In general, participants will seek to make the appropriate decisions to achieve the optimal risk and cost trade-offs for themselves. However, it is unclear whether the optimal resource mix for managing the retailers’ individual portfolio risk automatically translates to the optimal mix for managing the risk of USE for the whole society. When the risk of tail events increases, the potential misalignment between the retailers' and society's risk preference could potentially lead to a sub-optimal outcome if the market is the only source for delivering resources.

This is recognised by SW Advisory & Endgame Economics[21]:

> **… [I]n some circumstances, the market may be unable to deliver an acceptable outcome. At times of scarcity, a decision that is economic for a single generator or load may lead to an outcome that is not optimal for the entire power system.**

5.3.3 Enhancing the RERT procurement mechanism

The RERT procurement mechanism should be strengthened so that it can supplement the energy market to ensure delivery of the efficient resource mix due to the issues already described including:

- Emergency DR resources not being able to participate in the energy market;
- The energy market alone not being able to deliver the efficient resource mix.

To date, the adoption of Wholesale DR has been slow in the NEM, possibly due to the combination of obstacles discussed in Section 5.3.2. To assist its uptake, the AEMC has recommended a few reforms in its Reliability Framework Review including[22]:

- a short-term forward market (STFM) to provide revenue certainty
- recognising DR aggregators and providers on an equal footing with generators
- allowing multiple trading relationships at the same connection point to promote competition.

AEMO understands that a rule change request has been submitted to the AEMC on 31 August 2018 in relation to the second and third recommendations above[23]. While the AEMC's recommendations could facilitate the development of DR resources, their effect on the market is yet to be seen, and it is unknown how long it will take for the market to mature.

As tail risk increases, the cost of the market not delivering the efficient amount of resources becomes larger, particularly in those extreme, but plausible, USE events. It has become more important to ensure that the RERT, as a supplementary procurement mechanism, is able to act as an effective safety net and provide insurance to society when the energy market alone does not deliver the efficient resource mix. AEMO has submitted a rule change proposal to enhance the RERT mechanism.

AEMO's **Enhanced RERT proposal** is aimed at strengthening the supplementary procurement mechanism to deliver more optimal outcomes. We have proposed to achieve this through:

- Extending the procurement lead time to one year and having the option to sign a multi-year contract if it leads to lower costs. The multi-year contract effectively extends the payback period and will provide more certainty to potential suppliers.
- Delinking the RERT procurement trigger from the current reliability standard and determining RERT volume through an economic cost assessment framework. The assessment will take into account both the cost and risk of resource procurement and residual USE.
- Standardisation of RERT products to streamline the procurement process

As the RERT is part of the reliability framework, the discussion about its procurement methodology needs to be held within the context of the reliability standard. Therefore, we will first discuss the appropriateness of the current reliability standard in Section 6.1 before moving onto the RERT procurement methodology in Section 6.2.

---

[21] SW Advisory & Endgame Economics, *Review of Intervention Pricing, Final Report*, 4 Oct 2017, p1

[22] AEMC, *Reliability Framework Review Final Report,* 26 July 2018, p v.

[23] See https://www.aemc.gov.au/rule-changes/wholesale-demand-response-mechanism

6. Appropriateness of the Reliability Standard

6.1 Appropriateness of the current reliability standard

Increasing tail end risks raises a question about the appropriateness of the current reliability standard which compares an *average* USE measure to a single 0.002% threshold. AEMO considers that the current reliability standard does not adequately address the following two important issues:

1. **It assumes a *single cost* of VCR** – if VCR varies by customer segment, timing and magnitude of load shedding then the average USE metric will not be proportional to the cost of load shedding.
2. **It ignores the *value of insurance and risk mitigation*** – the current reliability standard approach does not recognise the value of reducing the risk of USE (i.e. reducing the range, or extremity of potential USE outcomes), hence the insurance value and benefit of a safety net that limits costs in extreme USE events is not considered in the trade-off.

Failure to incorporate these two aspects means the current reliability standard will not incentivise the optimal resource mix. In addition, linking RERT procurement in a binary form to the current reliability standard can lead to an on-again, off-again procurement trigger which could increase overall RERT costs.

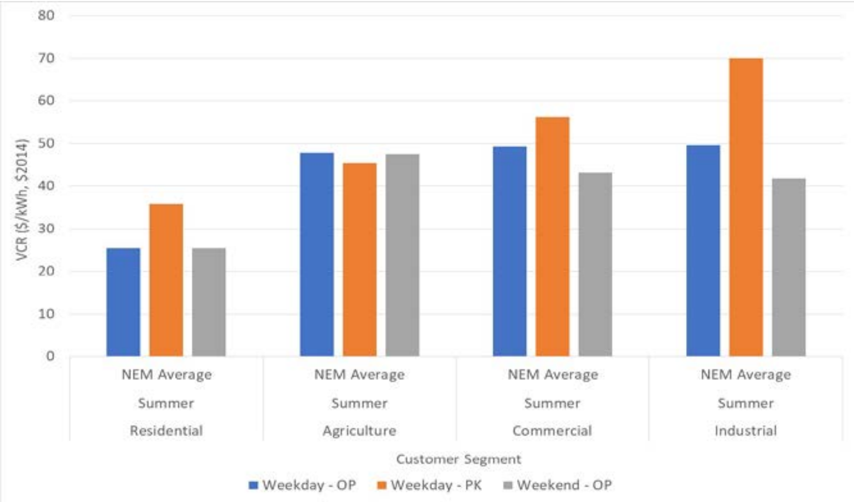
6.1.1 Impact of cost structure of VCR

As described in section 1 the reliability standard balances the cost of USE against the cost of new resources to avoid load shedding. The current average USE metric implicitly assumes that every MWh of load shedding carries the same cost, irrespective of the customer segment, time and duration of occurrence. However, this does not seem to be the case. AEMO's 2014 VCR study[24] shows that VCR in summer is higher during weekdays and peak times[25] than on weekends, as shown in Figure 26. This result is to be expected for businesses where load shedding during normal hours means lost sales and lost production. Even for households there is a higher disruptive cost of load shedding if it occurs during the morning and evening peaks on working days.

[24] AEMO, *Value of Customer Reliability Review,* September 2015

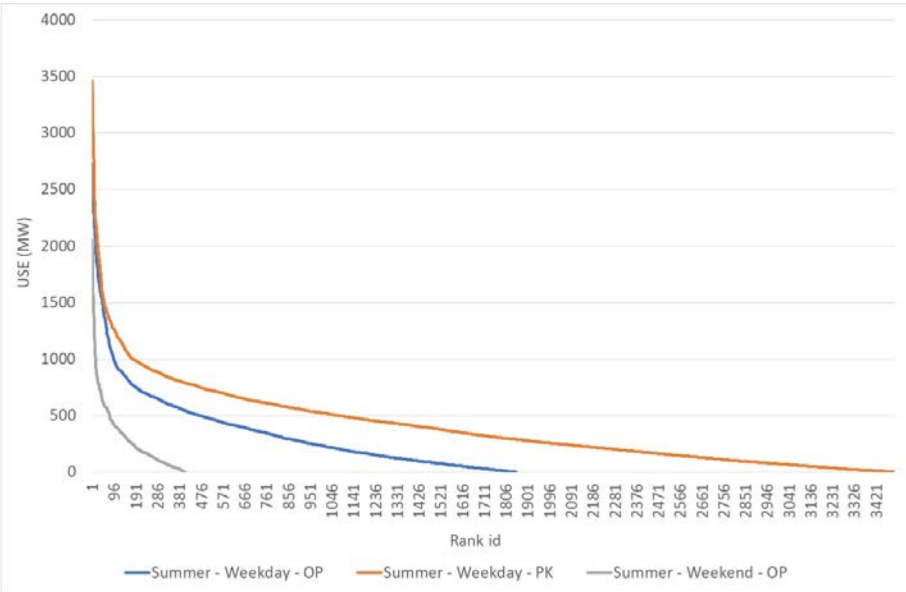
[25] Note in the study peak is defined between 7-10 am and 3-6pm.

Figure 26 Cost of VCR at different time periods



As might be expected the duration curves in Figure 26 shows that the frequency of load shedding is skewed towards working days and peak periods with very little USE expected on weekends. This means that there is a positive correlation between the cost of USE and the time that it occurs which means that the cost structure of USE is not symmetrical i.e. not all USE is the same. Therefore, a reliability standard based on an average USE will underestimate of the true cost and lead to an inefficient level of reliability.

Figure 27 POE10 USE duration curve at different times (VIC 2018/19 based on ESOO 2018 forecast)



6.1.2 Impact of risk and uncertainty

Although AEMO applies a great deal of rigour in quantifying expected USE, it will always be an estimate of future outcomes that is subject to uncertainty. The *shape* of the USE distribution and the *sensitivity* to different supply/demand inputs are also important pieces of information that should be taken into account in considering the risk of load shedding. However, these are not factored into the average USE metric and so supplementary metrics such as Conditional Tail Expectation or USE at Risk (see Figure 8) should be incorporated into the reliability framework.

The current reliability standard based on an average USE measure does not signal the value of risk management in mitigating against extreme USE outcomes. Using such a metric to balance the reliability trade-off assumes that society is **risk neutral**. That is, it weighs upside and downside outcomes equally and places no explicit value in avoiding extreme USE outcomes.

This is inconsistent with the well-established risk-aversion behaviour in economic theory as well as the prevalence of insurance products, *including hedging contracts in the energy market itself*. Just as retailers are willing to pay a risk premium over the *expected spot prices* for hedging contracts, and more so when the volatility of in the market increases, it would be natural to expect that the value and demand for insurance will increase in society overall when the USE tail risk becomes larger. The current reliability standard based on an average USE measure, however, does not reflect the risk aspect of the USE outcomes and hence does not explicitly signal the value of its mitigation.

Viewed in this light AEMO questions the role of the reliability standard in the current framework. Given the reliability framework is about managing risk (of load shedding) it is useful to draw lessons from other risk management approaches. The most usual way of managing tail-risk where there is a low probability of a high consequence event is to take out **insurance** and RERT can be considered as a form of insurance. The decision to procure insurance involves consideration of a number of factors which can be related to the procurement of RERT:

- The sum to be insured = quantity of RERT.
- The excess or deductible = usage cost of RERT.
- The premium of insurance = availability cost of RERT.

Hence, AEMO believes that RERT should be regarded as providing the insurance role in managing tail risks. The reliability framework, including the reliability standard, setting and response, should explicitly reflect the benefit associated with reducing the range of USE outcomes including limiting the extent of extreme events. The reliability framework should guide the decisions about the amount of insurance to be procured and incentivise the optimal mix of resources to achieve efficient reliability outcomes that take into account both the true cost and risk mitigation aspects.

6.1.3 Impact on efficient procurement

Based on the discussion above, AEMO considers that the current reliability standard will not deliver the efficient reliability outcome as it underestimates the true cost of USE and does not recognise the benefit of reducing its risks. It follows that linking RERT procurement to this standard cannot result in an efficient outcome. In addition, linking the procurement decision to the current standard results in a binary procurement decision, which can create uncertainty about whether RERT will need to be procured in the future. This effectively shortens the required payback period for resources with higher fixed costs and can increase the overall procurement cost.

6.2 RERT procurement and the reliability standard

In its March 2018 rule change request and July submission to the AEMC’s consultation paper, AEMO advocated for a number of changes to the RERT framework including delinking the RERT procurement trigger from the current reliability standard. AEMO considers that due to the increasing volatility and tail risks, RERT plays an important safety-net role by providing insurance against uncertainties and reducing costs in extreme USE events.

In light of its important insurance value AEMO considers that the RERT should be a **standing reserve** in the overall reliability framework and there should be an assessment framework for RERT procurement that takes into account both the cost and risk aspects of USE mitigation.

The reliability framework should set the *level* of the required standing reserve over a defined horizon (akin to determining the sum to be insured) by taking account of:

- the nature of the tail risk - using a range of supplementary metrics;
- the risk appetite for different levels of load shedding expressed both in cost and limits terms;
- the cost structure and optimal mx of resources that can prevent or mitigate load shedding.

This approach would lead to a more stable investment environment that provides greater certainty to developers of resources as well as reducing the risk of load shedding. In fact, having a standing RERT procurement mechanism will improve overall efficiency. Not only does it provide transparency and certainty to participants in the RERT market, leading to lower procurement costs, it also provides certainty to the participants in the main energy market, so that they can undertake the appropriate responses in their operation with confidence.

7. Recommendations

In this paper AEMO examined the appropriateness of the existing NEM reliability framework in the context of the observed trends in the drivers of unserved energy. AEMO's findings can be summarised as follows:

The tail risk of load shedding in the NEM is increasing

The NEM has experienced significant tightening in its supply- demand balance in recent years following the retirement of significant thermal generation. The trend of increasing maximum temperatures not only leads to higher demands, but also lowers supply due to de-rating of generation and transmission. At the same time, the large amount of renewable uptake increases the volatility and unpredictability in the system, which needs to be supplied by the remaining thermal fleet with ageing assets and increased chance of forced outage.

Whilst historically the NEM has experienced very little USE, the past lack of USE is not a good guide to future USE outcomes, as the historical outcome is a very small sample and the system has recently experienced significant changes in its supply demand balance.

The reliability framework should incentivise the optimal resource mix to manage tail USE risks

Managing tail USE risks requires a mix of resources with different cost structures and operating characteristics. Along with traditional generation technologies, demand response and distributed energy resources can play an important role.

Both the in-market and the RERT mechanism should be utilised efficiently to procure the optimal resource mix. However, there are a few challenges in ensuring the optimal mix can be delivered, including:

- Having a market price cap that is lower than VCR prevents resources with high operating cost from operating in the energy market.
- Short and uncertain operating hours means an uncertain revenue stream that could prevent resources with higher fixed cost from entering the market. This challenge can be present in both the energy market and RERT procurement mechanism.
- The market might not deliver the socially optimal level of risk management if the market participants' risk preference regarding USE mitigation does not align with that of the society.

The reliability framework should incentivise the deployment of the efficient resource mix by removing these barriers. Part of the solution is to strengthen the RERT mechanism so that it becomes a standing reserve and acts as a safety net which provides the vital insurance function in the overall reliability framework.

The current reliability standard is not suited to managing risk and uncertainty and cannot deliver the efficient resource mix

The current reliability standard, based on an average USE , does not describe the nature of the tail risk. It also fails to capture two important aspects of USE costs:

- The average USE metric is only related to the cost of load shedding if all USE events are equally regarded. To the extent that there is a limit on the duration of an event that is acceptable or if there is a higher VCR for larger USE events, this metric underestimates the cost of load shedding.
- The average USE metric ignores risk aversion which is counter to most evidence of human behaviour. To the extent that society is risk averse then it will prefer to pay a premium to avoid downside outcomes.

As the current reliability standard does not reflect the true cost of USE and the value of insurance against uncertain but plausible extreme events, it cannot incentivise the optimal resource mix to deliver the efficient reliability outcome. AEMO recommends that the current standard be replaced or supplemented with measures that appropriately reflect the trade-offs in USE mitigation.

Linking the RERT procurement trigger to the current standard leads to inefficient procurement outcomes. AEMO considers that the RERT should be established as a standing reserve to provide insurance to the market with the level of the reserve established over a defined horizon that takes account of the costs and value of mitigating USE risks.

Glossary

This document uses many terms that have meanings defined in the National Electricity Rules (NER). The NER meanings are adopted unless otherwise specified.

Term	Definition
AEMC	Australian Energy Market Commission
AEMO	Australian Energy Market Operator
CPT	Cumulative Price Threshold
DER	Distributed Energy Resources
DR	Demand Response
ESOO	Electricity Statement of Opportunities
LOLE	Loss of Load Expectation
LOLP	Loss of Load Probability
MPC	Market Price Cap
MTPASA	Medium Term Projected Assessment of System Adequacy
NEM	National Electricity Market
NEO	National Electricity Objective
NER	National Electricity Rules
RERT	Reliability and Emergency Reserve Trader
USE	Unserved Energy
VCR	Value of Customer Reliability