Filing Receipt

Received - 2023-03-29 02:27:21 PM
Control Number - 54584
ItemNumber - 19

PROJECT NO. 54584

RELIABILITY STANDARD FOR ERCOT MARKET	**§ §**	**PUBLIC UTILITY COMMISSION OF TEXAS**

TEXAS INDUSTRIAL ENERGY CONSUMERS' COMMENTS

I. INTRODUCTION

Texas Industrial Energy Consumers (TIEC) appreciates the opportunity to provide comments on Commission Staff's questions related to developing a reliability standard for the ERCOT market. The long-term reliability of the ERCOT system is critical for TIEC members, as they are engaged in energy-intensive businesses that depend on a continuous and reliable supply of electricity. Historically, the ERCOT market has had a reliability standard, and that standard was routinely used to calculate a reserve margin target from when the market first opened until a few years ago. Importantly, a reserve margin target or a reliability standard is not the same as a reserve margin mandate, which is a precursor to a capacity market. TIEC supports continuing to calculate a reliability standard for purposes of assessing the health of the market, but opposes a reserve margin mandate.

A reliability standard must be rooted in economic principles to be fair and cost-effective for consumers. As this project moves forward, the Commission should focus on establishing a reliability standard that is rooted in sound economic principles. Absent an economic underpinning, a reliability standard will essentially be plucked out of thin air, which risks creating an arbitrary goalpost, such as the antiquated "one event in ten years" standard. Such a standard would not be based on balancing marginal changes to reliability against the increased costs to customers from applying that standard. In addition, the Commission should not use an "event-based" standard like Loss of Load Expectation (LOLE) or Loss of Load Hours (LOLH) because those standards do not properly reflect the economics of expected outages, and the costs of the reserves needed to avoid those outages. TIEC understands that there may not be an appetite today for the "economically optimal reserve margin" (EORM), as the Commission endorsed in the past, and that there is a desire to take a more conservative approach to reliability. Still, economics and consumer costs must be considered if any reliability standard will be the basis for requiring new charges for customers or their load-serving entities, as contemplated by the Performance Credit Mechanism (PCM).

As discussed below, TIEC believes that the Commission could develop a rational reliability standard by modeling the Expected Unserved Energy (EUE) at various levels of reserves and the imputed Value of Lost Load (VOLL). Examining projected EUE would provide the Commission with a meaningful view of the system's outage risk at various levels of reserves because unlike a LOLE or LOLH standard, EUE considers both the depth and duration of potential outages. When viewed in conjunction with a reasonable VOLL, an EUE metric will allow the Commission to determine the economic cost of outages at various reserve levels as compared to the cost of avoiding those outages. This will enable the Commission to identify an EORM, which is the point where cost of projected outages is equal to the cost of procuring additional reserves, and use that as a starting point to ensure that the standard appropriately balances costs and benefits for customers.

The Commission should be aware that even using the same inputs, different models and different statistical methods to estimating loss of load will result in different estimated reliability outcomes. Therefore, the Commission should develop a standard model to project the amount of EUE that can be expected at various reserve margins. The Commission should also use transparent metrics and assumptions that can be reviewed and scrutinized by all market participants, rather than simply announcing a target reserve margin, a required level of "additional generation needed," or other "black box" metrics that are opaque and highly sensitive to modeling assumptions. Inviting stakeholder participation and input in the modeling process will result in a standard that reflects the best collective understanding of what the ERCOT market needs to ensure that there is sufficient capacity available to meet contingencies without imposing undue costs on customers. In particular, a transparent modeling process using publicly available inputs will allow stakeholders to run their own sensitivity analyses to support a robust debate on this extremely important topic and ultimately support the best policy outcome. Further, a robust modeling process will help the Commission to accurately project the impact of extreme scenarios operational changes like an accelerated rollout of distributed energy resources.

TIEC believes the Commission should focus on setting a single, ERCOT-wide reliability standard. Any deliverability or congestion issues identified by the reliability analysis should be resolved by building additional transmission facilities, not creating locational requirements that will create an economic disadvantage for growing areas of the state.[1] Deregulation in ERCOT was premised on the existence of a single, unified transmission system that would create a level playing field for generators to compete in a single energy market, and that would give customers across the state equal access to low-cost, reliable power. To the extent there are load or generation pockets, PURA grants the Commission ample authority to resolve those issues by building additional transmission facilities. The reliability analysis undertaken here should not be used to create generator entitlements to scarcity pricing or congestion rents in particular areas. Creating separate reliability standards for different areas will necessarily create winners and losers among both customers and generators. For instance, congestion is often linked to robust load growth and geographic challenges that limit new generators' ability to site resources within load pockets. Setting location-specific reliability mandates for areas with constrained transmission will only serve to reward incumbent generators that happen to have facilities inside the load pocket. Additionally, if the Commission examines reliability based on location, that could give certain geographically dominant generation companies an opportunity to exercise local market power. This is not an acceptable outcome when transmission can be built to resolve the issue, and TIEC believes creating additional incentives for this behavior should be avoided.

The Commission should also avoid including a seasonal component in its reliability standard. Recent experience has shown that reliability in ERCOT is driven by whether capacity is available during the intervals with the highest net peak load (meaning demand minus renewable output). If the system resources are sufficient to handle the single highest net peak load interval, the same system resources will be able to manage less impactful intervals, and any operational concerns should be addressed through other mechanisms (such as weatherization requirements or ancillary service procurements).

Finally, TIEC believes it would be reasonable for the Commission to reexamine its reliability standard approximately every three years. Revisiting the standard every three years will allow it to keep pace with changes in ERCOT's load and generation mix without unduly burdening the Commission's limited resources or disrupting market expectations.

---

[1] *See e.g.*, PURA §§ 37.056; 39.203(e).

II. COMMENTS

(1) The Commission has previously considered various reliability metrics, such as Loss of Load Expectation (LOLE), Loss of Load Hours (LOLH), and Expected Unserved Energy (EUE). (a) Which reliability metrics, including those not previously studied, should the Commission consider in establishing a reliability standard for the ERCOT power region? (b) Which reliability metric, or combination of reliability metrics, should the Commission adopt for the reliability standard in ERCOT? (c) What are the advantages of your chosen reliability metrics, and what are the disadvantages of alternative approaches?

i. The Commission should use an EUE metric in conjunction with the VOLL to determine the economically optimal reserve margin and, by extension, the cost that consumers will pay for reliability in excess of that economically optimal target.

The Commission should establish a reliability standard that balances the cost of incentivizing a particular level of reserves against the cost of any projected firm load shed events at that reserve target. Absent such an analysis, the Commission would be selecting a reliability standard in the dark, without any objective economic underpinning. If the reliability standard will be used to support a PCM or any similar capacity requirement, it is essential that customers are receiving appropriate value for the additional charges and not paying exorbitant rate increases for little to no incremental reliability improvement.

To that end, EUE is the only suitable metric to calculate the value of firm load shed because unlike LOLE or LOLH, which are discussed below, EUE captures all elements of outage risk, including both the depth and duration of outages.[2] In addition, it would be arbitrary for the Commission to select an EUE target in isolation without considering the ***cost*** of that unserved energy as compared to the cost required to avoid it. Accordingly, the Commission should pair an EUE metric with an economic analysis of the VOLL to determine what EUE level is economically

---

[2] Using an EUE metric would also allow the Commission to meaningfully compare ERCOT's reliability performance against other systems. *See* Brattle, *Estimating the Economically Optimal Reserve Margin in ERCOT* at 42 (Jan. 31, 2014) (available at: https://www.brattle.com/wp-content/uploads/2017/10/6098_estimating_the_economically_optimal_reserve_margin_in_ercot_revised.pdf) ("We recommend adopting normalized EUE as a preferred reliability metric for setting the reliability standard because it is a more robust and meaningful measure of reliability that can be compared across systems of many sizes, load shapes, and other uncertainty factors. Such a cross system comparison is not meaningful for either LOLE or LOLH because neither metric considers the MW size of the outage endured nor the size of the system itself.").

optimal for consumers, and then make a policy choice on an appropriate reliability standard.[3] This will allow the Commission to make rational, economically grounded decisions when choosing between various policy options to achieve its reliability goals.

TIEC has historically supported setting ERCOT's target reserve margin at an economically optimal level, such that the value of projected outages is equal to the costs that customers would incur to bring additional reserves into the market. Following Winter Storm Uri, TIEC recognizes that it may make sense to take a more conservative approach and incentivize a higher level of reliability. A reserve margin set above the economically optimal level assumes customers should pay more for reserves than is economically justified based on the level of firm load shed that is expected. While this is not an efficient market outcome, it may be a reasonable one, particularly given that VOLL estimates can vary substantially and have an element of subjectivity in their calculation. Still, the Commission should at least start by calculating an economically optimal reserve margin so that the Commission has information on the additional costs versus the reliability benefits of various options as compared to the economically optimal outcome.

One of the benefits of using EUE in conjunction with the VOLL is that it will allow the Commission to conduct transparent sensitivity analyses when establishing reliability targets. For instance, the Commission could perform sensitivities based on different VOLLs. As noted above, VOLL is an estimation of the price that consumers are willing to pay for uninterrupted electricity. However, it is not necessarily correct to apply a single VOLL at all times. For example, a residential customer may be willing to pay more to maintain electric service during extreme weather. Similarly, a customer's VOLL may increase as an outage increases in duration. Generally, the Commission adopts a static VOLL that accounts for these various situations through some type of probabilistic analysis. While the Commission should establish an economically optimal reserve margin using the actual, probabilistic VOLL in conjunction with the EUE, it could also run sensitivities based on higher assumed VOLLs. This would yield a more conservative reliability standard that is still externally verifiable and grounded in economic principles.

---

[3] (Magnitude of Outage in MW) * (Duration of Outage in Hours) * (VOLL in $/MWh) = Economic Cost of Outage

ii. The Commission should not rely on LOLE to judge reliability because it is an arbitrary metric that does not accurately evaluate the reliability impact of outages.

TIEC interprets LOLE to mean a reliability standard that is based on the expected number of loss of load events in a year. For instance, a 0.1 LOLE standard would mean that one loss of load event would be anticipated over the course of ten years. The problem with LOLE is that it only looks at the frequency of outages, but does not consider either their duration or magnitude. For example, a system that produced one Winter Storm Uri sized outage per decade would satisfy a 0.1 LOLE standard, while a system that produced two ten minute outages that impacted ten customers over the same time span would not. As this example illustrates, an LOLE standard is one of the most arbitrary possible reliability metrics. As the Commissioners have noted, regulators around the world have recognized that an LOLE metric is inadequate.[4] Additionally, research by Commission Staff in Project No. 42302 revealed that there was very little, if any, reasoning behind the Commission's traditional reliance on an LOLE standard, even in the regulated utility context.[5] As TIEC has previously commented, the "one in ten years" LOLE standard was likely the product of integrated utilities supporting a reliability standard that would incentivize unnecessary generation investment so they could expand their rate base.

Further, an LOLE standard is not conducive to accurately modeling the reliability needs of the market. LOLE does not properly reflect the economics of expected outages, the costs of the reserves needed to avoid those outages, or the rational reserve level a market should produce based on these factors. Because LOLE has no foundation in economic or market principles, it does not place any rational limit on the costs that the market will incur to avoid outages. Further, LOLE is primarily a supply-side analysis, so it does not take accurately account for demand response

---

[4] *See* Public Utility Commission of Texas, January 26, 2023 Open Meeting at 1:02:41 (available at: https://www.adminmonitor.com/tx/puct/open_meeting/20230126/) ("I know that for SPP's purposes we have a broad based reliability standard of 0.1, 1 in 10. But they are talking about Expected Unserved Energy. Everybody wants to grow into a more granular reliability standard to hold utilities accountable for for the reliability on a regional basis."); Public Utility Commission of Texas, February 16, 2023 Open Meeting at 21:47, (available at: https://www.adminmonitor.com/tx/puct/open_meeting/20230216/) ("As we know through our work in other markets and everything we've been observing throughout the country, other ISOs and RTOs and regions are evaluating other metrics because the one-in-ten reliability standard alone isn't proving to be the best metric."); id. at 25:56 ("As I've said in the past, the loss of load expectation of 0.1 metric is inadequate from my opinion going forward. I think everyone from markets around the world knows that.").

[5] *See Review of the Reliability Standard in the ERCOT Region*, Project No. 42302, Staff Memorandum at Appendix B (Jun. 13, 2014).

resources—particularly price responsive demand. Additionally, as multiple parties discussed in their feedback to the E3 report last fall,[6] the outputs of an LOLE-based model vary widely depending on the assumptions in the model. As a result of the inherent flaws in the LOLE metric, there have been repeated "false alarms" about resource adequacy in ERCOT over the years where claims have been made that we should have "an outage every three years," and yet there have been ***zero*** loss of load events driven by resource inadequacy since ERCOT was deregulated. Rather, all load shed events have been driven by extreme weather and extreme generator forced outages, which is a performance issue not a resource adequacy issue. For all of these reasons, the Commission should reject an LOLE approach as it has in the past, and instead focus on metrics that account for the magnitude, duration and economic impact of outages.

iii. The Commission should not rely on LOLH because it does not consider the magnitude of an outage or its economic consequences.

TIEC understands LOLH to mean a metric that examines the expected combined duration (in hours) of loss of load events per year. LOLH is just a more granular version of LOLE and it suffers from many of the same flaws. For instance, neither LOLH nor LOLE consider the size of an anticipated outage. For example, a one-hour 100 MWh outage event and a one-hour 10,000 MWh outage event would look the same using either LOLE (they are both one outage) or LOLH (they both lasted one hour), even though the impact of the second outage is one hundred times greater. Additionally, like LOLE, LOLH does not provide the Commission with the enough information to compare a projected outage against the size of the ERCOT system or to optimize the costs of potential firm load shed against the costs of incentivizing enough additional capacity to prevent that load shed. In sum, the Commission should not use a LOLH metric because it only focuses on the duration of potential load shed, but does not provide essential information about the magnitude and, more importantly, the economic consequences of an outage.

(2) What is the most effective way that the Commission can include deliverability in the reliability standard?

TIEC opposes having location-specific reliability standards, such that costs for certain areas of the state may be perpetually higher than other areas of the state. This will tend to

---

[6] *See Review of Market Reform Assessment Produced by Energy and Environmental Economics, Inc. (E3)*, Project No. 54335, IMM Comments at 10 (Dec. 15, 2022); Project No. 54335, TIEC Comments at 5 (Dec. 15, 2022).

disadvantage large load centers and economic hubs. If this is what is meant by a "locational" component, this path should be avoided.[7] However, TIEC supports analysis that identifies restrictions on serving load in high-growth areas and seeks to resolve those issues through additional transmission construction.

A reliability standard is meant to assess resource adequacy and is not the appropriate tool for identifying or resolving localized constraints that prevent generators from effectively delivering power to customers in all areas of the state. When ERCOT was deregulated, a policy choice wad made that customers across ERCOT would bear the costs of building out the ERCOT transmission system, which is designed to create a level playing field for competition between generators, and to treat loads equally throughout the state. In turn, that robust transmission system would prevent any one generator from exercising local market power and would result in lower energy prices. That is why ERCOT has a "postage stamp" system that socializes the cost of transmission investment across all load-serving entities and, by extension, all end-use customers in ERCOT. As part of deregulation, the Commission was also given sufficient tools to ensure that transmission buildout will be sufficient to minimize deliverability issues. Not only does the Commission have authority to approve new transmission projects to resolve reliability issues, but PURA § 37.056(d) already requires the Commission to consider deliverability improvements—in particular, "estimated congestion cost savings for consumers"—when issuing transmission CCNs for economic projects. Further, as explained in PURA § 39.203(e), if transmission "constraints are such that they are not being resolved through [PURA] Chapter 37 or the ERCOT transmission planning process," then the Commission has authority to "require a [utility] to construct or enlarge facilities ***to ensure safe and reliable service and to reduce transmission constraints*** within ERCOT in a cost-effective manner."[8] Moreover, as the Commission found in multiple cases related to building additional import capacity into the Houston area, increasing costs within a load pocket is not effective in incentivizing new generation behind the constraint when the ability to construct new generation is restricted by ownership concentration and permitting issues.[9] As a

[7] *See* Question 3(a), *infra*.

[8] PURA § 39.203(e) (emphasis added); *see also Project for Commission Ordered Transmission Facilities*, Project No. 52682, Order (Oct. 14, 2021) (exercising authority to order transmission buildout under PURA § 39.203).

[9] *Application of CenterPoint Energy Houston Electric, LLC to Amend a Certificate of Convenience and Necessity for a Proposed 345-kV Transmission Line within Grimes, Harris, and Waller Counties*, Docket No. 44547,

result, deliverability issues should only be identified in any reliability analysis for purposes of ***transmission planning*** and not to create entitlements to higher profits for generators in certain constrained areas. The Commission should instead ensure that ERCOT's regulated transmission utilities continue to promptly build additional transmission as it becomes justified based on the reliability and economic thresholds described in PURA Chapter 37,[10] or, if that proves inadequate, exercise its authority to directly order new, cost-effective transmission buildout.

(3) Additional considerations in establishing the reliability standard in the ERCOT power region.

a. Should the reliability standard include a locational requirement?

TIEC understands a "locational requirement" to mean some sort of system to separately judge resource adequacy for various geographical subsets of the ERCOT grid. TIEC does not support including a locational component in the reliability standard. As explained above in response to Question 2, one of the fundamental principles of deregulation is that customers would collectively fund a single, unified transmission system to facilitate competition across a single power market. To the extent that transmission constraints create location-specific reliability problems, the solution to those problems should be to build additional transmission, not to reward or incentivize generators who happen to be located on the correct side of a constraint.

Additionally, creating a locational reliability requirement would necessarily balkanize the ERCOT grid and create winners and losers among ERCOT's customers because it is more difficult to site generation in some areas of the state. For instance, the Houston area has traditionally been a load pocket due to inherent limitations, like difficulty obtaining air permits, that make it hard for

---

Order on Rehearing at FoFs 71-72 (Apr. 28, 2016) ("Historical experience has shown that construction of new generation in the Houston area has not kept pace with generation plant retirements. Relying on "just in time" generation is not an appropriate alternative to planning an adequate and reliable transmission system."); *id.* at FoFs 101-102 (explaining that "[b]ased on the ERCOT independent review, ERCOT concluded that additional transmission capacity is needed to reliably serve the Houston area" because while it is possible that additional generation capacity could materialize "ERCOT did not recommend deferring [transmission buildout] because of the risk of retirement of existing generating units and the fact that ERCOT cannot compel generation or demand response to locate in a certain area and participate in the ERCOT market."); *id.* at FoFs 161-162 ("The HIP will facilitate robust wholesale competition by increasing the transfer capacity between the northern and coastal regions, thereby reducing constraints and allowing potentially more efficient generation in each zone to compete in the other. A new import path into the Houston area will open the market for new, more efficient generation sources to be constructed outside of that area and sell power into Houston, which will introduce additional competition for the legacy generation resources in the area.").

[10] *See* PURA § 37.056.

new resources to site close to the city's ever-expanding load.[11] As explained above, the solution to this problem should be to build additional transmission. If the Commission were to create a location-specific reliability requirement, that would necessarily create differential impacts for customers in different areas of the state. If localized reliability standards meant that customers in resource constrained areas would pay more to incentivize generation to site locally, then those customers would not share equally in the benefits of funding a unified transmission system. Yet these additional payments would not result in additional capacity in the load pocket, as the incumbent generators are unlikely to build if doing so will reduce prices, and may not be able to do so at all if the ability to get an air permit is restricted.

Further, incorporating locational factors into a reliability standard could give rise to localized market power issues. This is not a theoretical concern, NRG and Calpine own nearly 80% of the dispatchable generation capacity in the Houston area.[12] A localized reliability standard would provide those companies with an opportunity to engage in anti-competitive practices, like withholding production, to create artificial scarcity that could drive up prices in the local areas they dominate. As such, localized reliability standards would uniquely advantage these incumbent thar are located within load pockets, and do nothing to actually address the issue. As discussed above, the solution to localized reliability issues is to build transmission, and not to reward local generators with substantial market power within a load pocket.

b. Should the reliability standard include a seasonal component?

It is not clear that including a seasonal component in the reliability standard is necessary or beneficial. As TIEC has discussed in prior comments, recent experience has shown that reliability risk is highest during intervals with the highest net peak load (total demand minus intermittent generation output).[13] If the system has sufficient reserves to meet demand during the highest anticipated net peak load hour for the year, then that should be sufficient to ensure reliability throughout the year, regardless of what season it is when the tightest days occur. If there

---

[11] *See* Docket No. 44547, Order on Rehearing at FoF 62 ("The ability to locate new generation in the Houston area is constrained by both geography and environmental regulations relating to ozone.").

[12] *See* ERCOT Capacity, Demand and reserves Report (Nov. 2022) (available at: https://www.ercot.com/files/docs/2022/11/29/CapacityDemandandReservesReport_Nov2022.pdf) (showing the resources for counties of Brazoria, Chambers, Fort Bend, Galveston, and Harris).

[13] Project No. 54335, TIEC Comments at 3-4 (Dec. 12, 2022).

are insufficient reserves committed and available in real-time, that is not a resource adequacy issue but an operational/commitment issue that should be directly addressed in other ways with more targeted operational tools, such as weatherization requirements and ancillary services.

The Commission's reliability standard should focus on ensuring that there is sufficient capacity available to satisfy demand across the intervals with highest net peak load, regardless of how those intervals are distributed throughout the year. While those intervals may be concentrated in particular seasons, it is ***net peak load***, and not the seasons themselves, that should drive any reliability-based incentives to generators. If a net peak load event occurs in the winter, then that should be factored into the reliability analysis, but "winter" should not be given any special weight just because this is a possibility. Put more directly, the Commission should not artificially assign risk to the various seasons because the seasons themselves are a poor proxy for the actual reliability concerns facing the ERCOT system. Focusing the reliability standard around the seasons rather than net peak load could create irrational incentives. For instance, generators may defer maintenance during the shoulder months in order to chase seasonal incentives, which could impact their availability at more critical times of the year. Loads may also unnecessarily curtail in relatively low-risk circumstances, reducing the state's economic output for no reason. A year-long assessment is more appropriate and will capture net peak load events that occur in any season.

c. How can extreme events be captured in a reliability standard?

Extreme events should be probabilistically factored into the analyses that are performed to develop the reliability standard. In doing so, it is critical that extreme outlier events are either excluded or assigned a realistic probability so they do not skew the analysis. This is consistent with prior treatment of the summer of 2011, the hottest summer on record, which was discounted in subsequent resource adequacy analyses in the 2012-2015 timeframe.

The purpose of a reserve margin is to serve as a "cushion" to provide reliability across the expected distribution of outcomes. If the Commission places undue weight on extreme events and then adds a reserve margin on top of what is necessary to serve "expected" needs in that scenario, then that would effectively double count the risk to the market and lead to overly conservative (and overly expensive) reserve targets.

d. **How can the value of distributed energy and load resources be captured in a reliability standard?**

To the extent these resources are registered, distributed energy resources and load resources should show up as a reduction in demand in the load forecasts that are used to model the ERCOT system. That said, if the Commission has reason to worry that its model is not fully capturing the impact of demand response, it could incorporate adjustments into the model to tweak the anticipated level of price responsive demand.

(4) How frequently should the Commission update the calculation of the requirement necessary to meet the reliability standard? (a) What criteria should help determine the frequency of the update?

It is important for the Commission to periodically update its reliability standard to ensure that it is based on an accurate representation of the ERCOT market. Historically, ERCOT has reassessed the reserve margin required to meet the current reliability standard every two to four years.[14] TIEC believes that reassessing the reliability standard roughly once every three years would strike an appropriate balance between keeping the standard fresh, preserving the Commission's limited resources, and providing the market with a certain level of stability. The exact timing of updates should depend on whether there have been significant changes in the inputs that go into the Commission's model, such as shifts in ERCOT's generation mix or significant variations in load forecasts, or if the Commission has reason to believe that its current standard is proving to be either too conservative or not conservative enough.

(5) If you have any industry or academic papers on the topic and best practices that you believe the Commission should review while establishing the reliability standard for the ERCOT power region, please provide them.

[14] *See e.g.*, *Activities Related to the Projected ERCOT Reserve Margin and Methodology for Calculating the ERCOT Reserve Margin*, Project No. 30715, Memorandum (May 19, 2005); Global Energy Decisions, *ERCOT Target Reserve Margin Analysis* (January 18, 2007) (available at: https://www.ercot.com/files/docs/2007/02/14/ercot_reserve_margin_analysis_report.pdf); *Commission Proceeding to Determine Resource Adequacy in Texas*, Project No. 40000, ERCOT Report on Reserve Margin Analysis (Sep. 19, 2012).

TIEC has identified a report by the Redefining Resource Adequacy Task Force, which is attached as Attachment A for the Commission's review.[15] Notably the report includes the following key findings:

- "LOLE is an inadequate metric in a world of more varied shortfall events because it provides limited information on shortfall events' size and duration. This makes it difficult to know the true impact of potential shortfalls and nearly impossible to determine the type of resources necessary to reduce the number of shortfalls."[16]
- "Increased use of EUE is a good first step [to extracting more information from a reliability standard], as it quantifies the expected aggregate size (amount of energy) and duration of shortfall events, as opposed to only quantifying the probability or frequency of one occurring."[17]
- "[Regulators should report] a broader set of resource adequacy metrics than simply an average LOLE, including hourly EUE and additional information on the distribution of outages. Metrics should also be used to develop detailed statistics on the shortfall events themselves in order to better characterize the size, frequency, duration, and timing of events so that mitigation measures can be properly sized."[18]
- "The proliferation of energy storage, demand response, electric vehicles, and dynamic rate design bring with them new options for load flexibility and should be evaluated in a similar context as generation resources, including uncertainty and availability."[19]
- "As a result [of a more dynamic modern power system], new resource adequacy analysis should be designed to increase cost transparency so that regulators, policymakers, and consumers understand the relative costs of different levels of and approaches to reliability and can make informed investment decisions."[20]
- "Grid planners and regulators should have a clear understanding of the costs associated with achieving different reliability targets in different ways, to ensure that the value provided to the customer is worth the cost of a given investment—that the resource adequacy for which customers are being asked to pay is actually the type of reliability needed on the grid."[21]

---

[15] ESIG Energy Systems Integration Group, *Redefining Resource Adequacy for Modern Power Systems* (2021).

[16] *Id.* at 10.

[17] *Id.* at 12.

[18] *Id.* at 29.

[19] *Id.* at 21.

[20] *Id.* at 26.

[21] *Id.* at 28.

III. CONCLUSION

TIEC appreciates the opportunity to provide these comments and looks forward to further discussion on developing a reliability standard in ERCOT.

Respectfully submitted,

O'MELVENY & MYERS LLP

/s/ Katie Coleman
Katherine L. Coleman
State Bar No. 24059596
Michael A. McMillin
State Bar No. 24088034
John Russ Hubbard
State Bar No. 24120909
303 Colorado St., Suite 2750
Austin, TX 78701
(737) 261-8600
kcoleman@omm.com
mmcmillin@omm.com
jhubbard@omm.com
ommeservice@omm.com

ATTORNEYS FOR TEXAS INDUSTRIAL ENERGY CONSUMERS

PROJECT NO. 54584

RELIABILITY STANDARD FOR	**§**	**PUBLIC UTILITY COMMISSION**
ERCOT MARKET	**§**	**OF TEXAS**

TEXAS INDUSTRIAL ENERGY CONSUMERS'
EXECUTIVE SUMMARY

- The Commission should establish a reliability standard that is rooted in sound economic principles. Absent an economic underpinning, a reliability standard risks creating an arbitrary goalpost, such as the antiquated "one event in ten years" standard.

- Expected Unserved Energy (EUE) is the only suitable metric to calculate the value of firm load shed because it captures all elements of outage risk, including the depth and duration of outages. When EUE is paired with an economic analysis of the value of lost load (VOLL), the Commission can identify an economically optimal reserve margin (EORM), and use it as a starting point to ensure the standard balances the costs and benefits for consumers. While the Commission may choose to adopt a reserve margin that is more conservative than the EORM, the EORM should be the starting point for any analysis so that the Commission has information on the additional costs versus the reliability benefits of various options as compared to the economically optimal outcome.

- Loss of Load Expectation (LOLE) only considers frequency of outages, not their duration or magnitude. Further, LOLE has no foundation in economic or market principles and fails to accurately account for demand response. Loss of Load Hours (LOLH) is a more granular version of LOLE, and it similarly fails to provide essential information about the magnitude and, more importantly, the economic consequences of an outage.

- Deliverability or congestion issues should be resolved by building additional transmission facilities, not adding locational requirements to the reliability standard because doing so will create winners and losers among both customers and generators. Further, the Commission should avoid including a seasonal component in its reliability standard because if the system resources are sufficient to handle the single highest net peak load interval (demand minus renewable output), the same system resources will be able to manage less impactful intervals. The Commission can address operational concerns through other mechanisms, such as weatherization requirements or ancillary service procurements.

- In developing a reliability standard, the Commission should use transparent metrics and assumptions that can be reviewed and analyzed by all market participants, rather than simply announcing "black box" metrics that are opaque and highly sensitive to modeling assumptions. Inviting stakeholder participation and input in the modeling process will result in a standard that reflects the best collective understanding of what the ERCOT market needs to ensure that there is sufficient capacity available to meet contingencies without imposing undue costs on customers.

Attachment A

Redefining Resource Adequacy for Modern Power Systems

A Report of the Redefining Resource Adequacy Task Force

2021

ESIG

ENERGY SYSTEMS INTEGRATION GROUP

ENERGY SYSTEMS
INTEGRATION GROUP

About ESIG

The Energy Systems Integration Group is a nonprofit organization that marshals the expertise of the electricity industry's technical community to support grid transformation and energy systems integration and operation, particularly with respect to clean energy. More information is available at https://www.esig.energy.

ESIG's publications

This report, as well as other reports and briefs by ESIG, is available online at https://www.esig.energy/reports-briefs.

Get in touch

To learn more about the recommendations in this report, please send an email to resourceadequacy@esig.energy.

To learn more about the Energy Systems Integration Group, send an email to info@esig.energy.

Redefining Resource Adequacy for Modern Power Systems

Report of the Redefining Resource Adequacy Task Force

Authors

Prepared by **Derek Stenclik**, Telos Energy

The core members of the Redefining Resource Adequacy Task Force are:

Aaron Bloom, Energy Systems Integration Group
Wesley Cole, National Renewable Energy Laboratory
Armando Figueroa Acevedo, Black & Veatch
Gord Stephen, National Renewable Energy Laboratory and University of Washington
Aidan Tuohy, Electric Power Research Institute

Acknowledgments

The task force collaborated closely with a project team made up of industry experts. The task force would like to acknowledge the valuable input and support of the following individuals regarding the concepts discussed in this report.

Chris Dent, University of Edinburgh
Rob Gramlich, Grid Strategies
Elaine Hart, Moment Energy Insights
Karin Matchett, consulting writer/editor
Michael Milligan, independent consultant
Mark O'Malley, Energy Systems Integration Group
Nick Schlag, Energy and Environmental Economics

Suggested citation

Energy Systems Integration Group. 2021. *Redefining Resource Adequacy for Modern Power Systems*. A Report of the Redefining Resource Adequacy Task Force. Reston, VA. https://www.esig.energy/reports-briefs.

Design: David Gerratt/NonprofitDesign.com

Contents

Evolving Reliability Needs for a Decarbonized Grid

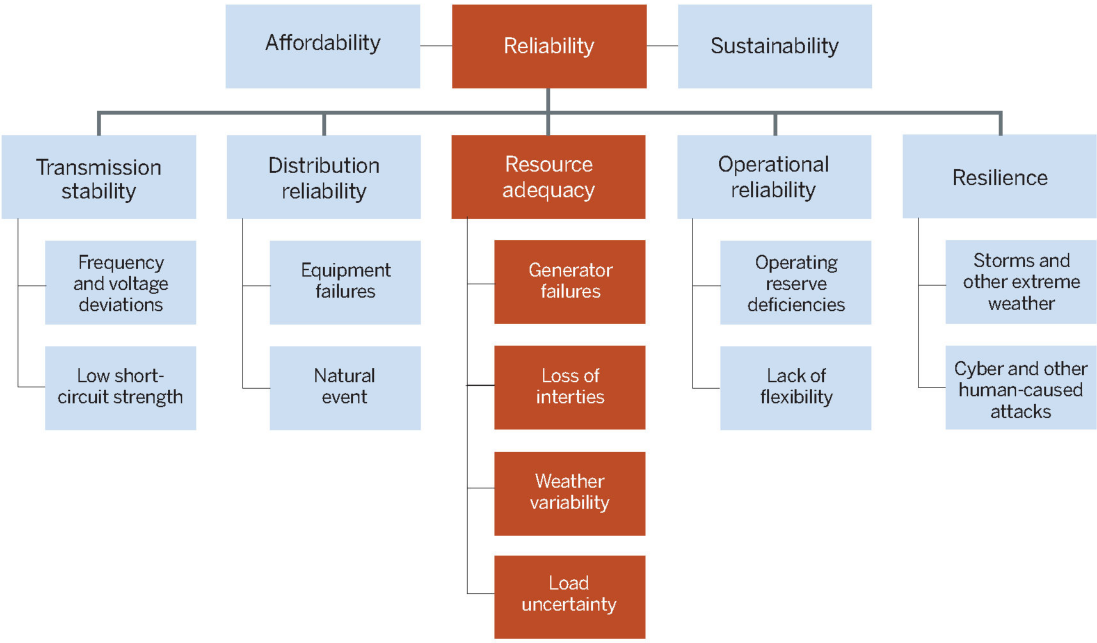
As grids around the world continue to decarbonize and integrate renewable energy, it is critical that power system planners, policymakers, and regulators continue to balance three pillars of power system planning: affordability, sustainability, and reliability (Figure 1). While some stakeholders may have different priorities across the three pillars, each one is critical to ensuring a smooth clean energy transition.

Ask any grid operator their top priority and the answer is simple: reliability. Our society has come to expect, and require, uninterrupted power—even on the hottest days and coldest nights and through the longest storms. These expectations remain as the grid transitions to high variable renewable energy; reliability is paramount. With increased variability and uncertainty, how can we ensure there are enough resources to serve electricity customers whenever and wherever they need power?

Elements of Resource Adequacy Under Rising Levels of Renewables

One dimension of grid reliability, that taking the longest view, is resource adequacy: having enough resources in the bulk power system available to the system operator to meet future load, while accounting for uncertainty in

FIGURE 1
The Elements of Grid Reliability



Source: Energy Systems Integration Group.

The increased role of wind, solar, storage, and load flexibility requires the industry to rethink reliability planning and resource adequacy methods.

both generation and load. Some uncertainties are becoming more important, such as correlated generator outages and changes in the weather. By evaluating these uncertainties statistically, grid planners project their resource needs to reach an acceptably low level of risk of capacity shortages. Risk metrics can then be used to determine how much investment our power grids require, how much new generation should be built, what type of generation should be built, and which generation can retire.

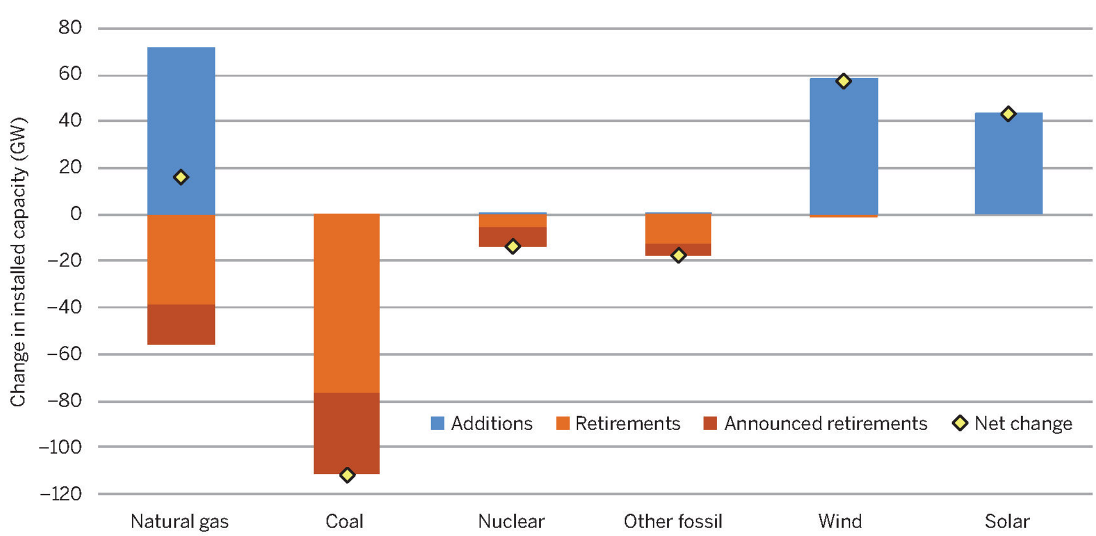
The power system has always been heavily influenced by the weather. Extreme temperatures determine the timing of peak demand, winter cold snaps can limit natural gas supply, gas turbine reliability and output are affected by ambient conditions, and hydro output varies seasonally and annually. However, as the grid increasingly relies on variable renewable energy such as wind and solar, the attention to reliability and weather conditions is increasingly important.

The industry has more than two decades of experience integrating variable resources while maintaining—and even improving—grid reliability. However, a notable trend is occurring. While early wind and solar capacity constituted incremental expansions of the grid's installed capacity, the industry is now seeing a large swath of fossil generator retirements, including coal, nuclear, and legacy gas assets (Figure 2). As a result, portfolios of wind, solar, storage, and load flexibility are increasingly used as replacements to conventional fossil capacity.

These new resources are being utilized not only for energy, but also for the grid services required to maintain grid reliability. The increased role of wind, solar, storage, and load flexibility requires the industry to rethink reliability planning and resource adequacy methods and to reconsider analytical approaches. Computational approaches developed in the 20th century are limiting our collective ability to evaluate reliability and risks for modern power systems. The confluence of changes requires new data, methods, and metrics to better characterize evolving reliability risks.

FIGURE 2

Generation Additions and Retirements from 2014 through 2020, Plus Planned Retirements



Source: Energy Systems Integration Group; data from APPA (2021).

Wake-Up Calls from California and Texas

Two recent events underscore the importance of modernizing our thinking on resource adequacy in an era of changing generation mixes and changing weather. The first occurred in California in August 2020, when a heat storm resulted in two separate days of involuntary rolling outages across California's power system. The second occurred in Texas in February 2021, when extreme winter weather resulted in very high electricity demand while also causing natural gas fuel supply shortages, low wind output, and widespread equipment failures across all generation types. Both of these reliability failures showed how susceptible the grid can be to inadequate supply as well as the economic, political, and social fallout that can occur when grid reliability is jeopardized.

Steve Berberich, the chief executive officer of the California Independent System Operator (CAISO) during the August 2020 events, summarized the changing needs of resource adequacy analysis, stating:

> There doesn't have to be a tradeoff between reliability and decarbonization. What caused the [August blackouts] was a lack of putting all the pieces together. You have to rethink these old ways of doing things, and I think that's what didn't happen. . . . The resource adequacy program in California is now not matched up with the realities of working through a renewable-based system, and, in a nutshell, needs to be redesigned (Hering and Stanfield, 2020).

Reassessing the Resource Adequacy Methods

The objective of this report is to move that redesign forward by providing an overview of key drivers changing the way resource adequacy needs to be evaluated, identifying shortcomings of conventional approaches, and outlining first principles that practitioners should consider as they adapt their approaches.

This report focuses specifically on the resource adequacy analysis and methods that measure system reliability and risk, and it intentionally stops short of translating that analysis into procurement decisions. Ultimately, system planners, regulators, and policymakers need to ensure

Resource adequacy methods have not changed considerably in the past few decades, despite rapid changes of the resource mix. The central message for practitioners, regulators, and policymakers is, what got us here won't get us there.

there are enough resources to serve load. Resource adequacy analysis provides the tools to determine whether there are enough resources and, if not, what type of resource is needed to meet reliability needs. While this report is comprehensive in its treatment of resource adequacy methods, it intentionally does not address capacity accreditation, which determines how to assess reliability contributions of specific resources, or capacity procurement and market mechanisms, which require further analysis.

Resource adequacy methods have not changed considerably in the past few decades, despite rapid changes of the resource mix. The central message for practitioners, regulators, and policymakers is, what got us here won't get us there.

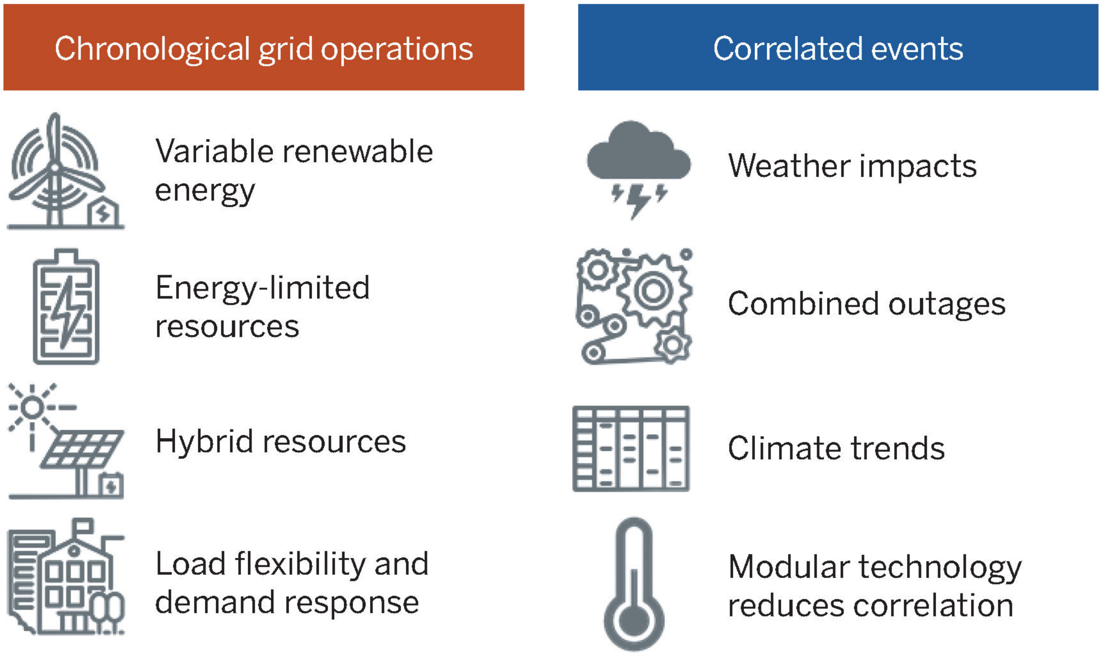
Traditional Resource Adequacy Analysis Problems and Their Causes

At its core, the challenge with resource adequacy analysis is that, while the methods and metrics used by the industry today originated in the last several decades of the 20th century, they have only been improved incrementally while the resource mix transitioned appreciably. For example, early tools evaluated only single peak load periods and did not assess risk of shortfalls across an entire year. These tools often assumed static loads and did not consider energy limitations of most resources. Traditional resource adequacy analysis also made a simplifying assumption that reliability events were uncorrelated and that mechanical failures of generating equipment occurred at random, thus assuming that the probability of multiple failures occurring simultaneously was low.

However, for a grid with high levels of renewables, energy-limited resources, and load flexibility, reliability is strongly affected by chronological operations and weather-influenced correlated events (Figure 3). These are two driving factors requiring the industry to modernize frameworks for resource adequacy analysis.

Chronological grid operations: Traditional resource adequacy analysis often evaluates only individual peak load hours and does not consider the full year of

FIGURE 3
Two Driving Factors That Require New Approaches to Resource Adequacy



Source: Energy Systems Integration Group.

FIGURE 4

Example of Capacity Outage Probability Table

Units Out	Number of Units in Group																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	18
0	980000	960400	941192	922368	903921	885842	868125	850763	833748	817073	800731	784717	769022	753642	738569	723798	695135
1	20000	39200	57624	75295	92237	108471	124018	138900	153137	166749	179756	192175	204027	215326	226093	236342	255356
2		400	1176	2305	3764	5534	7593	9921	12501	15316	18345	21571	24980	28566	32299	36175	44297
3			8	31	77	151	258	405	595	833	1123	1467	1869	2332	2856	3445	4821
4					1	2	5	11	18	30	46	68	95	131	175	228	369
5									1	1	1	2	4	5	8	11	21
6																1	1

The table shows the probability of multiple units being on outage simultaneously. Looking down a single column, the probability of multiple units on outage simultaneously drops precipitously. As the total fleet size increases (moving from left to right along a row), the probability of a large percentage of the overall fleet on outage (e.g., six units out of 18) is a one-in-a-million event.

Source: Calabrese (1947).

operation. This has two problems: it presupposes that the highest risk period occurs during peak load, and it fails to account for the sequential operating characteristics of resources. For example, the usefulness of battery storage as a resource depends on the weather (and resulting generation) in the preceding days and expectations for needs in subsequent hours. Likewise, the use of demand response as a resource depends on how long the system has already been asking customers to provide demand response.

Correlated events: While historical resource adequacy analysis focused on probabilities of discrete independent mechanical or electrical failures (modeled with randomly occurring forced outages), weather-influenced correlated events should now be recognized as a driving factor of reliability.

Why Reliability Events Occur Is Changing

Historically, determining whether there were enough resources available to meet load was a straightforward analysis, with the foundation rooted in probabilistic assessment. With the power system made up of many large, centralized fossil fuel generators for which fuel availability was rarely a concern, the availability of a generator was largely based on discrete maintenance and mechanical failures (forced outages). Each generator could be characterized with a maintenance outage rate (%) and a forced outage rate (%), which were used to determine the likelihood the unit would be unavailable to serve load. Because these were mechanical failures and largely uncorrelated (with one another, the weather, or other factors), probabilistic assessments could quantify

While historical resource adequacy analysis focused on probabilities of discrete independent mechanical or electrical failures, weather-influenced correlated events should now be recognized as a driving factor of reliability.

the likelihood that many generators would be on outage at the same time, thus increasing the risk of a shortfall and failure to meet load.

An example of this analysis, referred to as the convolution method, can be found in a capacity outage probability table from a seminal work on resource adequacy from the mid-20th century (Figure 4). This table shows the probability, in millionths, that the indicated number of units would be out simultaneously for fleets having a given number of units when the outage rate is 2 percent. It shows that as the number of units goes up, the likelihood of a large portion of the fleet being on outage decreases quickly—so for an interconnected power system, the probability of capacity shortfall events diminishes noticeably as system size increases.

A probabilistic approach may have been appropriate for the historical power system, where reliability risk stemmed largely from mechanical failures of large generating units that could mean many hundreds, or even thousands, of megawatts (MW) lost due to a single failure. Coal and nuclear generation were the primary fuel sources and had weeks' worth of fuel storage on site, so fuel availability was not a concern and output was not variable.

However, while randomly occurring forced outages are still important to consider, it is increasingly important to consider correlated generator failures and outages, due to either the underlying weather or other root causes.

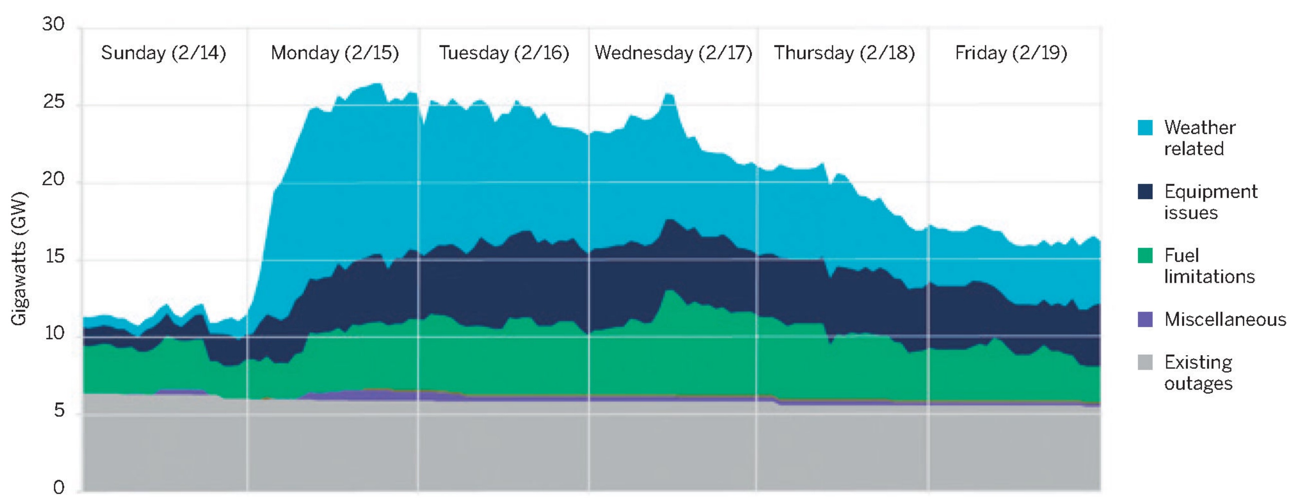
First, a large shift from coal to gas capacity has increased risks associated with fuel supply. The electric power sector is now tightly coupled with the natural gas delivery system, which delivers fuel on demand, with little or no storage located at the power plant. As a result, correlated outages due to fuel supply failures are now a key reliability risk, especially during the winter months when multiple power plants may experience interrupted fuel supplies simultaneously. These same time periods see significant increases in load and mechanical failures. This confluence of factors is leading some system operators, like the New York Independent System Operator (NYISO), to require dual-fuel capability for natural gas generators and others, like the Electric Reliability Council of Texas (ERCOT), to discuss potential winterization requirements.

Second, the gas turbine technology in wide use today is more dependent on ambient temperature than are steam turbine technologies. This is especially true at high summer temperatures. Both extreme high and extreme low temperatures derate the maximum output of the machines, correlating their availability to the underlying temperature. Mechanical failures are also more likely during extreme cold events for most technologies and fuel types in common use today. The correlation in these types of outages was clearly evident in the February 2021 event in Texas, as shown in Figure 5 (ERCOT, 2021c).

Third, for a grid with higher levels of wind, solar, storage, and load flexibility, the actual events that are correlated have very different characteristics. Unlike a fossil fuel–powered generator, which can lose hundreds or even thousands of MWs of capacity to a single failure, the loss of capacity from the disconnection or failure of small, modular resources is much smaller and more geographically dispersed. Wind, solar, and storage plants are made up of many independent inverter-controlled resources. While any individual wind turbine may fail, the probability of an entire plant failing is much lower. This modularity shifts the analysis from discrete generator forced outages to evaluations of the likelihood of correlated events and common mode failures.

FIGURE 5

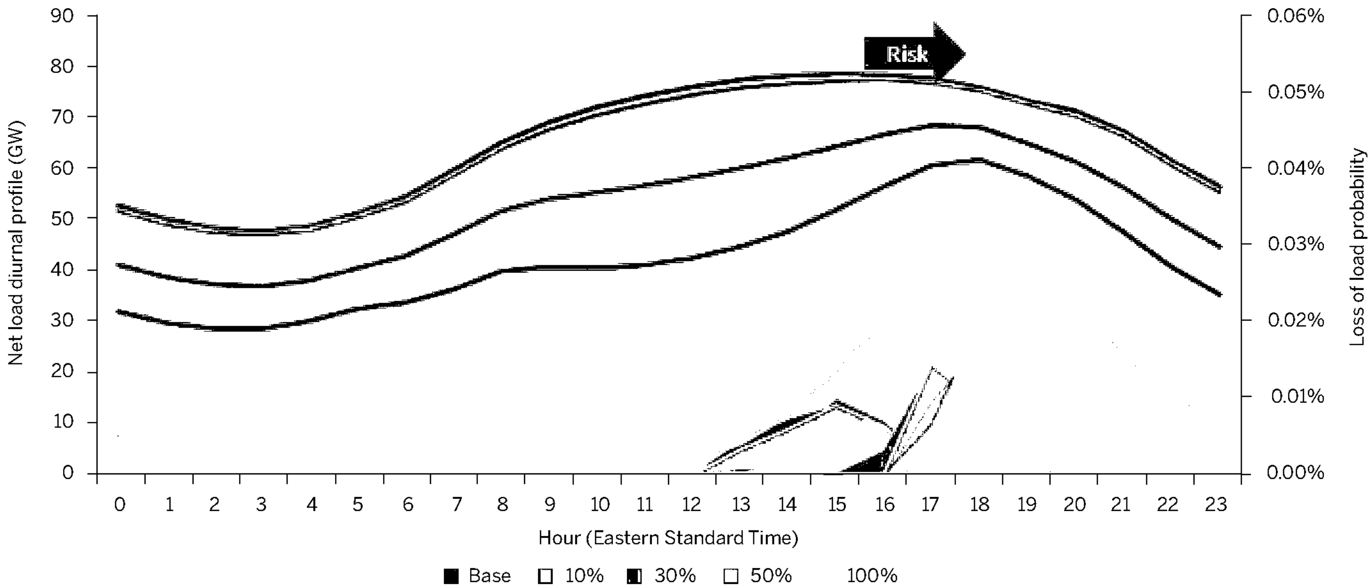
Correlated Outages for Natural Gas Generators by Cause During the ERCOT February 2021 Event



Note: Extreme cold temperatures began on Monday morning.

Source: Electric Reliability Council of Texas (2020c).

FIGURE 6
Shifting Periods of Risk in MISO with Increasing Levels of Solar Photovoltaics



Source: Midcontinent Independent System Operator (2021).

When Reliability Events Occur Is Changing

The changing resource mix is also affecting when reliability events are more likely to occur. In traditional resource adequacy analysis, periods of higher probability of a shortfall were almost always associated with peak loads. Because generator outages were assumed to be random and variable resources constituted a small part of the resource mix, generator availability was assumed to be relatively uniform across the year. As a result, peak risk occurred during periods of higher demand. Across most of North America, this usually occurred during hot summer afternoons or cold winter mornings or evenings.

However, time periods with a risk of shortfall are shifting. The periods of risk we're used to keeping our eye on may no longer be the most challenging. In the case of solar, the diurnal pattern causes a drop in solar production at the end of the day correlated among all solar plants in the area, and extended cloud cover can reduce output as storms pass through a region. For wind generators, wind speeds can be correlated as different atmospheric conditions or storm fronts pass through a region. As a result, in a system with high levels of wind and solar resources, there are both predictable lulls in production as well as other, weather-influenced times during which production across the fleet is well below average.

These dynamics were evident in the involuntary rolling outages in California in August 2020, which occurred late in the evening after the sun had set and solar resources dropped off, several hours after peak load occurred in the middle of the day. The shifting periods of shortfall risk are illustrated in Figure 6 (MISO, 2021). As levels of solar generation increase, the periods of risk shift from 3 p.m. to 6 p.m. due to changes in resource availability.

Winter cold snaps are also increasingly challenging, as seen in the Texas event in February 2021, which occurred in a historically summer-peaking system that has a high winter reserve margin. While load is higher than normal during periods of extreme cold, for most summer-peaking systems these winter loads still tend to be lower than the annual peak. However, the challenge also manifests itself on the supply side with increased probabilities of equipment failures, wind turbine icing, and natural gas supply that is stressed by heating demand. Thus, in these winter periods, shortfalls do not have a single root cause, but are rather a correlation of multiple challenges.

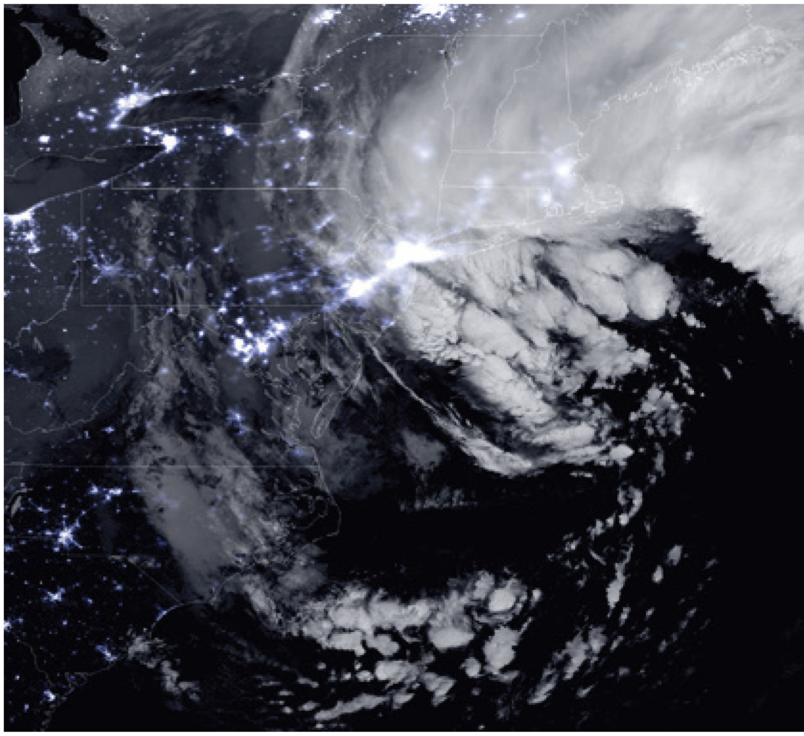
Periods of risk can also be common during off-peak periods if extended cloud cover across a region reduces solar availability and weather patterns reduce wind speeds. These declines in solar and wind production can align with extreme cold and heat, and therefore higher loads. Finally, even periods once characterized as low risk, like the spring and fall seasons, may have increased risk. Given that large fossil generators are typically taken offline for maintenance during these periods, if an outlier weather event occurs, the probability of a shortfall can increase significantly even though loads are considerably lower than they are in peak periods.

Taken together, the shifting periods of risk mean that planners can no longer bypass analysis and evaluate only peak load periods. A broader evaluation across all hours of the year is necessary to accurately capture shifting periods of risk of shortfall. Given the energy limitations of storage and demand response and the operational characteristics of other resources like start-up times and ramp rate limitations, the all-hour approach must be combined with a chronological assessment of grid operations across an entire year. From a modeling perspective, the disciplines of production cost modeling and resource adequacy modeling are increasingly blurred.

How Reliability Events Occur Is Changing: It's All About the Weather

The power system has always been heavily influenced by the weather—extreme temperatures determine the timing of peak demand, winter cold snaps can limit natural gas supply, gas turbine reliability and output are affected by ambient conditions, and hydro output varies seasonally and annually. However, as already discussed, as the grid increasingly relies on variable renewable energy, like wind and solar, the attention to reliability and weather conditions is increasingly important.

Traditional resource adequacy analysis typically evaluated weather as a driver of system load. Weather changes could move peak demand periods and created uncertainty in planners' load forecasts. There was some recognition that weather could lead to correlated outages of the fossil fleet, but rarely was this trend evaluated explicitly. Instead, the outage rate assigned to generators was based only on forced outages for unexpected mechanical failures and planned maintenance.



The increased dependence on weather that accompanies the shift to more wind and solar on the system causes multiple issues. The first and most obvious is that weather variability affects the availability of these generation resources. Hour-to-hour changes in weather and electricity generation mean that a system's probability of a capacity shortfall is constantly changing. Given that serving load in a high-renewables power system also involves the use of energy-limited resources such as storage and demand response, a chronological perspective on system modeling and simulation is required, rather than the static analysis used in traditional analysis.

In addition, while weather is constantly changing, so is climate—the weather conditions prevailing in an area in general or over a long period. If a changing climate leads to changes in weather, temperature, and extreme events, it changes the overall resource adequacy risk profile. Traditional resource adequacy analysis relied solely on historical weather data; however, the use of historical data to characterize load and renewable resources may not be appropriate for gauging future risks affected by climate change.

The Need for a Modified Approach

To overcome the limitations in traditional resource adequacy analysis, a fresh look is required. While decades of resource adequacy analysis can be used as a reference point for reliability planning moving forward, future methods will need to evolve, and a set of first principles can be a useful guide. While each region and system require a unique process, guiding principles can help ensure a consistent approach in terms of the objectives, structure, and process of resource adequacy planning. Consistency in approaches to resource adequacy can better allow for sharing of insights and best practices, interregional resource coordination, and a smoother regulatory process for resource procurement.

The first principles listed below are based on a few simple questions: if the approaches to resource adequacy analysis started from scratch, without a backdrop of 100 years of power system planning and conventional approaches, how would resource adequacy be evaluated for modern power systems? How should risk and reliability be evaluated in a power system with large shares of wind, solar, storage, and load flexibility? How can methods be developed in a technology-neutral manner, to ensure the methods evolve with a changing resource mix and new technologies? Responses to these questions point to six principles of resource adequacy for modern power systems.

The objective of these principles is to clearly articulate evolving resource adequacy concepts to system planners, regulators, and policymakers in order to encourage a consistent approach to complex challenges. The principles are not meant to be overly prescriptive; instead, they are designed to provide a guiding framework that can

be used by system planners around the world, regardless of the unique system attributes. These principles are designed to help system planners do three things: first, to better understand and quantify the reliability shortfalls that a modern power system is more likely to experience; second, to identify ways that such shortfalls can be mitigated and responded to; and third, to understand what the resource adequacy analysis means for resource procurement.

Principles 1 and 2 address the new needs in our understanding of capacity shortfalls. Principles 3, 4, and 5 focus on new understandings of capacity and resource types in a modern power system. Principle 6 calls for the inclusion of economic considerations in reliability analyses.

PRINCIPLE 1: Quantifying size, frequency, duration, and timing of capacity shortfalls is critical to finding the right resource solutions.

As the power system's resource mix changes, resource adequacy metrics need to be transformed as well. The conventional resource adequacy metric, loss of load expectation (LOLE), quantifies the expected number of days when capacity is insufficient to meet load. A common reliability criterion is one day of outage in 10 years, often simplified to 0.1 days per year LOLE.

But LOLE is an opaque metric when used in isolation. It only provides a measure of the average number of shortfalls over a study period and does not characterize the magnitude or duration of specific outage events. It also does a poor job of differentiating shortfalls, which, depending on their length and duration, can have unequal impact on consumers and can require different mitigation options. For example, since LOLE only quantifies frequency, a shortfall of 1 percent of load for 10 hours is measured the same way as a shortfall of 10 percent of load for 10 hours. In addition, there is very little consistency in this metric's application, as different planners in different regions interpret the criterion differently, and each region has different institutional and regulatory requirements that determine what probability of unserved energy is acceptable.

Similar metrics also provide information on the probability of a shortfall event but limited information regarding what the shortfalls look like. Loss of load hours (LOLH) counts the average expected number of hours of shortfall, loss of load events (LOLEv) is similar to LOLE but allows for multiple "events" to occur in a single day or a single event to span multiple days, and expected unserved energy (EUE) calculates the average amount of energy unserved.

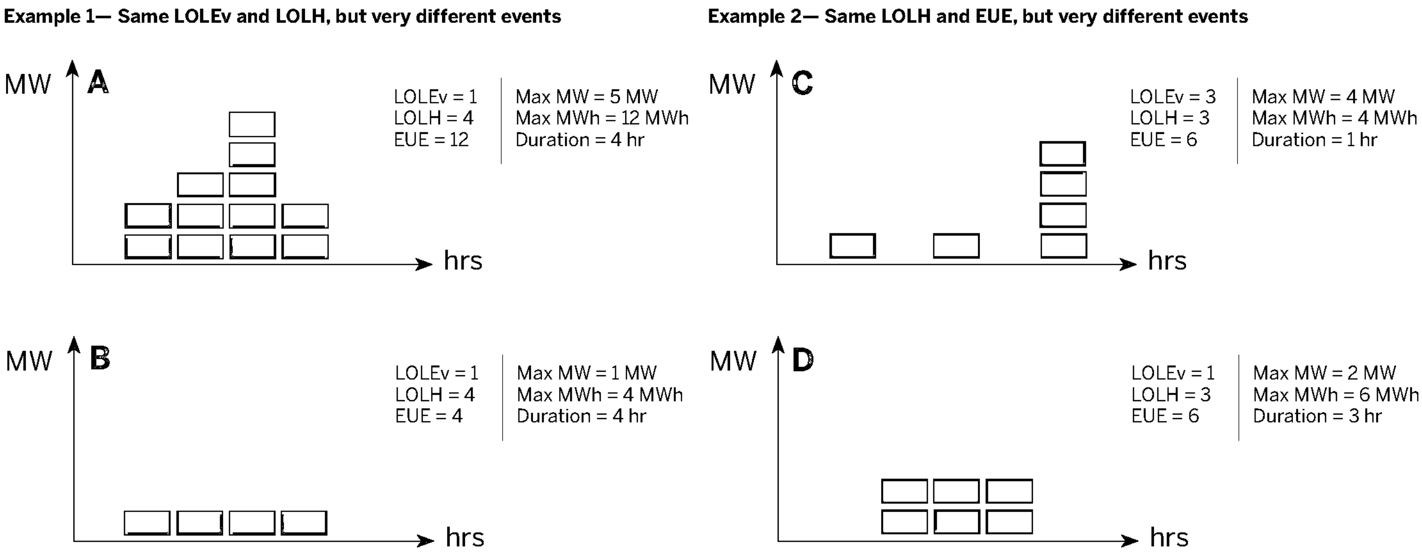
Looking Beyond LOLE

The reliance on the LOLE metric was adequate in traditional resource adequacy analysis because shortfalls tended to share similar characteristics, largely occurring during peak load events and caused by randomly occurring forced outages of the conventional fossil fleet. In addition, the resource solutions implemented when the LOLE measure was exceeded were one size fits all. The combustion turbine was the de facto resource used to meet reliability needs, as it was the lowest capital cost way to get more "steel in the ground," and operating costs (based on fuel efficiency) were not a concern because the units were rarely utilized. However, the resource options available to system planners today are numerous. Energy storage, demand response, and load flexibility provide competitive alternatives to the combustion turbine approach for many types of shortfall events.

Energy storage, demand response, and load flexibility provide competitive alternatives to the combustion turbine approach for many types of shortfall events.

In addition, the reliability events are now more varied; therefore, understanding the size, frequency, duration, and timing of potential shortfalls is essential to finding the right resource solutions. LOLE is an inadequate metric in a world of more varied shortfall events because it provides limited information on shortfall events' size and duration. This makes it difficult to know the true impact of potential shortfalls and nearly impossible to determine the types of resources necessary to reduce the number of shortfalls.

FIGURE 7
Building Blocks of Resource Adequacy Metrics



Each block represents a one-hour duration of capacity shortfall, and the height of the stacks of blocks depicts the MW of unserved energy for each hour. A: a single, continuous four-hour shortfall with 12 MWh of unserved energy; B: a single, continuous four-hour shortfall with 4 MWh of unserved energy; C: three discrete one-hour shortfall events with 6 MWh of unserved energy; D: a single, continuous three-hour shortfall with 6 MWh of unserved energy.

Source: Energy Systems Integration Group.

Differentiating Capacity Shortfalls

Systems with the same LOLE and LOLH can have very different risk profiles, types of shortfalls, and mitigation options. Figure 7 illustrates four different capacity shortfall events. On the x-axis of each chart is time and on the y-axis is the MW of a shortfall event. Each block represents a one-hour duration of capacity shortfall, and the height of the stacks of blocks measures the amount of unserved energy. These building blocks show how different shortfall events can be and thus how easily traditional metrics can fail to capture them.

The two charts on the left (Figure 7A and B) show how simple expected value metrics can fail to distinguish between very disparate events. These charts show a single continuous capacity shortfall event of equal duration (four hours). Both of these events would count toward the aggregate loss of load events (LOLEv) metric as one event, since they occur within the same day, and both would count toward LOLH with four hours. From an LOLEv and LOLH perspective, then, the events are indistinguishable. However, the top event is three times larger in terms of unserved energy and five times larger in terms of the maximum unserved energy at a single point in time. The events have very different impacts on customers and may require different mitigation strategies on the part of system operators.

While EUE is better at differentiating individual events, this metric too can have challenges. The charts on the right (Figure 7C and D) show consistent unserved energy and loss of load hours, but the top plot shows three distinct events (LOLEv of 3), whereas the bottom plot shows a single event. In this case, the corresponding EUE and LOLH metrics are identical, but the LOLEv metrics are three times larger in the top example. Separate events could be mitigated by energy storage that can re-charge between events, but may be further challenged by demand response programs that may be limited by the number of allowable calls.

Without the use of multiple metrics, as well as additional information on the size (both in MW and megawatthours (MWh)), frequency, and duration of individual events, determining appropriate mitigation actions is difficult. For example, the event on the top left would require at least three times more energy storage and demand response than the event on the bottom left. For the events on the right, a battery resource of 4 MWh could avoid all of the unserved energy in the top right event (provided it could recharge between events), but would be insufficient to avoid the bottom right event (where we would need 6 MWh of storage). This information would be impossible to ascertain by LOLE, LOLH, and EUE metrics alone. Resource adequacy metrics that can quantify size, frequency, duration, and timing of shortfall events are critical to finding the right resource solutions.

Achieving Deeper Insights into Resource Adequacy Metrics

One of the biggest limitations of LOLE, LOLH, and EUE metrics is that they provide only an *average* measure of system risk across many hundreds or thousands of samples. They do not provide information on the full distribution of shortfalls. New methods in resource adequacy analysis should expand to provide additional insights into not only the average (expected value) resource adequacy events, but also the characteristics of the individual events themselves. System planners require this type and quantity of information to ensure that they can right-size mitigations to meet the system's specific reliability needs.

New methods in resource adequacy analysis should expand to provide additional insights into not only the average (expected value) resource adequacy events, but also the characteristics of the individual events themselves.

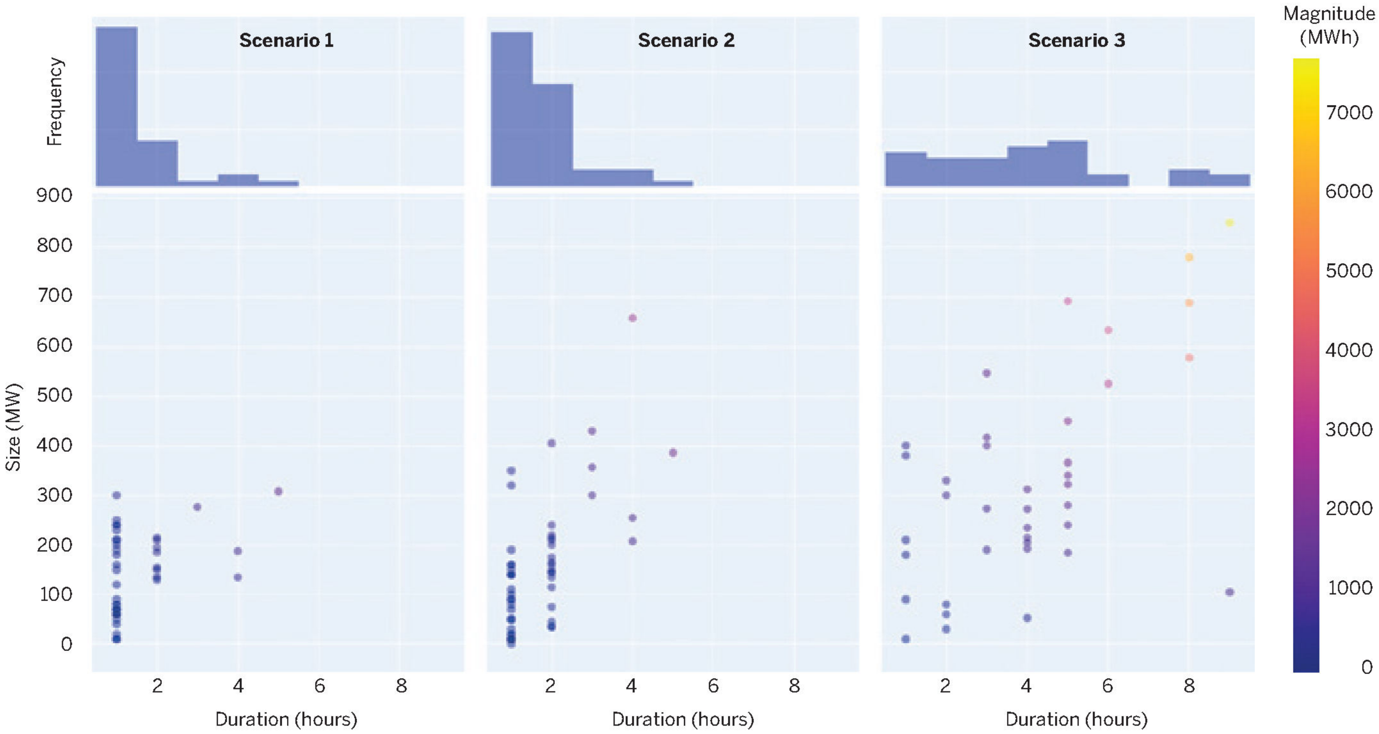
It is too early to tell whether entirely new metrics need to be developed, but what is certain is that planners need to extract more information and details from existing ones. Increased use of EUE is a good first step, as it quantifies the expected aggregate size (amount of energy) and duration of shortfall events as opposed to only quantifying the probability or frequency of one occurring. However, EUE still provides only a single average metric that cannot distinguish between the individual events.

In addition, resource adequacy analysis should pay attention not just to the expected values, but to potential tail events. While high-impact, low-probability events are very rare—and system planners intentionally do not plan to mitigate all potential risk—these events' impact on a high-renewables grid is important to assess given their potentially devastating impact on customers.

Unfortunately, traditional resource adequacy metrics' simplification of hundreds or thousands of power system simulations into a single average offers little insight into the distribution of potential resource adequacy shortfalls that the system could experience. A system that has rare but very large events could appear to have the same level of reliability as a system with more frequent, smaller events, causing current metrics to fail to account for the much greater impact on consumers—and society in general—of the large events. Future resource adequacy analysis should move beyond expected values and provide information on the distribution of individual events.

The chart in Figure 8 (p. 13) quantifies the number of shortfall events (each represented as a dot) for a single system simulated across three different resource mixes.

FIGURE 8
Scatter Plot of Size, Frequency, and Duration of Shortfall Events with Energy-limited Reliance on Energy Limited Resources



Source: Energy Systems Integration Group.

Each resource mix has very different underlying resources but the same LOLE of 1-day-in-10-years (i.e., the same number of dots). However, despite having the same LOLE, the systems have very different risk profiles. An improperly planned high-renewables grid may experience much larger shortfall events than those we are used to planning for due to sustained periods of low renewable production. This could cause longer and larger disruptions—even if the probability of these events occurring is lower than historical norms. Improved use of resource adequacy metrics can avoid this challenge.

Improved utilization of existing metrics and visualizations must move beyond average values. They must provide information on the distribution of events as well as provide emphasis on individual, rather than aggregate, event characteristics. Relying on multiple metrics and visualizations of the size, frequency, duration, and timing of shortfall events will allow planners to select mitigations and resources that are appropriately sized to fit system needs and avoid over-procurement of resources.

PRINCIPLE 2: Chronological operations must be modeled across many weather years.

Historically, traditional resource adequacy analysis evaluated only periods of peak demand for reliability risk. This was in part due to the more limited computational capabilities of the time as well as to a resource mix that did not fluctuate much seasonally or hourly, making the fluctuations of load the main variable. In addition, there was limited energy storage on the system with which to smooth out demand. The small amount that was installed was pumped hydro, often with 12 or more hours of energy storage, and energy limitations were less of a concern. As a result, systems included few short-duration and energy-limited resources that would not be able to provide extended support during reliability events. Therefore, if generation on the system was adequate during the period of highest load, it would be adequate during the rest of the year as well.

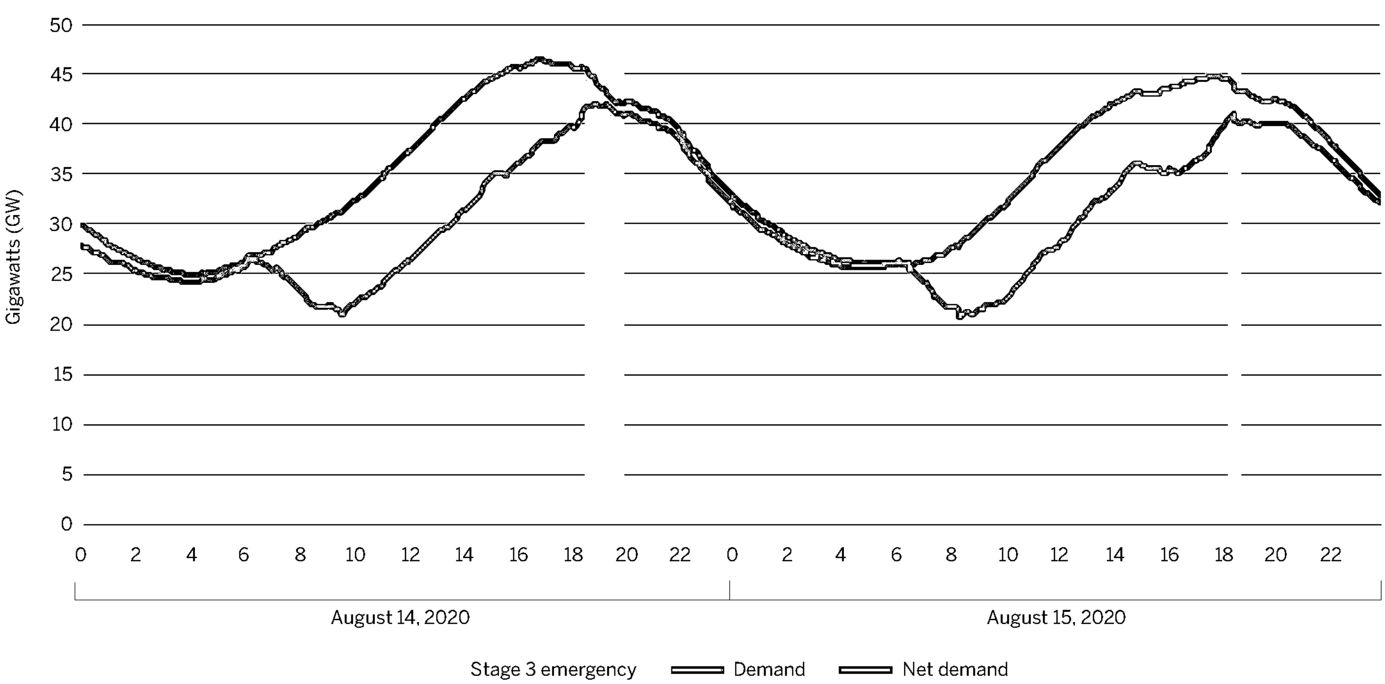
Today, the increased reliance on variable renewable energy and energy-limited resources is changing the resource adequacy construct. Periods of risk are no longer confined to peak load conditions, but are shifting to other time periods due to abnormal weather events, the daily setting of the sun, and the fossil fleet undergoing increased maintenance during fall and spring. The increased levels of variable renewable energy mean that resource analysis requires specific attention to hourly, seasonal, and inter-annual resource variability. The sequence of the variability is key, given that energy-limited resources such as batteries or demand response require either a preceding period or subsequent period of high production to be useful for grid reliability. This will require increased reliance on weather and power forecasting and integrated storage scheduling that considers forecast uncertainty to ensure that storage can be available when needed.

As a result, the conventional approach of designing a system solely to meet peak load conditions—via a static planning reserve margin—is no longer appropriate. A simple planning reserve margin that is used to procure a certain amount of capacity above and beyond peak load does not ensure that the system will be reliable during other times of the year given changes in the resource mix.

Importance of Chronological Evaluation of All Hours

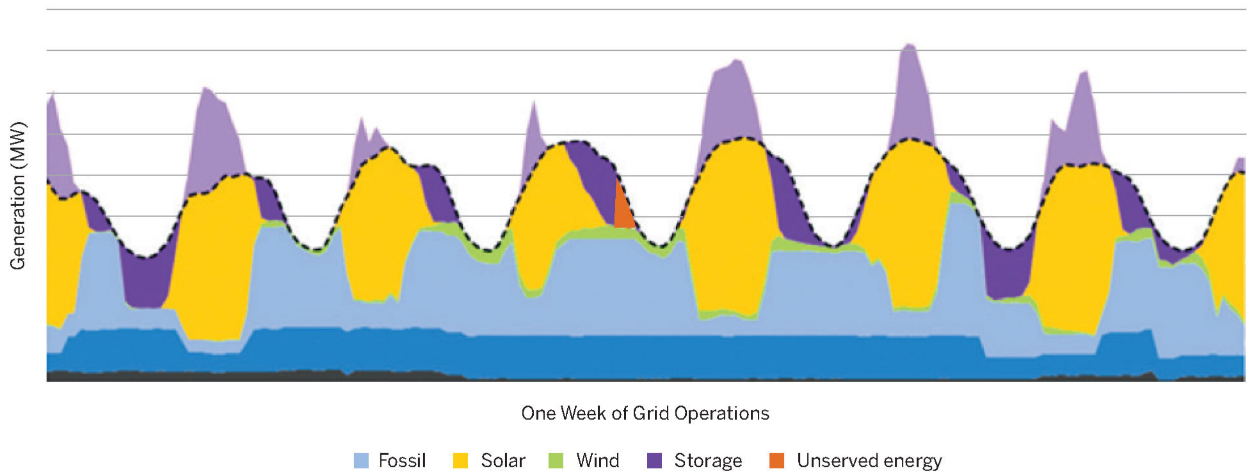
The California rolling blackouts in 2020 are a good example. California's resource adequacy construct and planning reserve margin are based on the peak gross load, which occurs in the middle of the day during summer months. However, periods of peak risk in California now occur in the evening hours as solar resources decline and loads remain relatively high. This is clearly illustrated in Figure 9, which shows the gross and net load for CAISO for the August days when rolling blackouts occurred (CAISO, 2021). The conventional assumption that peak risk is aligned with peak load is no longer true, requiring a chronological evaluation of all hours of the year so that the times of risk of shortfall can be accurately identified.

FIGURE 9
Gross and Net Load During the 2020 California Reliability Event



Source: Energy Systems Integration Group; data from California Independent System Operator (2021).

FIGURE 10
Example of Chronological Resource Adequacy Simulations with a Shortfall Event



Source: Hawai'i Natural Energy Institute (2020).

Chronological operations and scheduling ensure that energy storage and demand response will be around long enough, and can fully recharge, to support the system through reliability challenges.

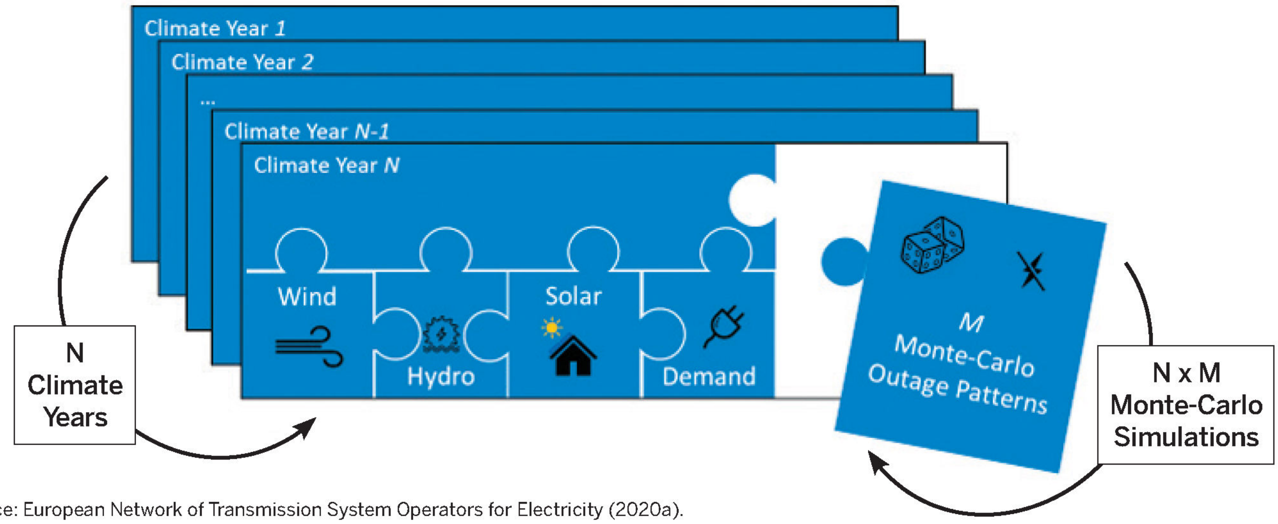
This chronological assessment is required to ensure that the energy storage and demand response will be available for enough hours to get the system through periods of scarce supply. Energy-limited resources may reduce reliability risk in some periods (when the storage is discharging or when load is reduced), but only if they increase risk in other periods (when the storage is charging or when load is shifted to earlier or later times). Hour-to-hour operations and scheduling ensure that energy storage and demand response will be around long enough, and can fully recharge, to support the system through reliability challenges. Chronological assessment is essential to highlight resource adequacy needs and necessary procurement of long-duration storage resources.

Modeling sequential grid operations is critical to capture the whole picture: the variability of wind and solar resources along with the energy limitations of storage and load flexibility. Chronological stochastic analysis is thus increasingly important, simulating a full hour-to-hour dispatch of the system's resources for an entire year of operation across many different weather patterns, load profiles, and random outage draws. An example is shown in Figure 10, which illustrates a week of chronological commitment and dispatch of a power system, and a shortfall that occurs when there is insufficient storage available to extend through the late evening hours. Despite load being significantly lower in the late evening hours, the probability of a shortfall is higher (HNEI, 2020).

Need for Many Years of Weather Data

In addition to modeling chronological grid operations, resource adequacy analysis for modern power systems requires the incorporation of many years of weather data. Many years of synchronized hourly weather and load data are necessary to understand correlations and inter-annual variability between wind and solar generation, outages, and load. The same weather conditions can affect wind and solar output, whose probabilities are driven by irregular and complex weather patterns, and load and thermal unit derates—requiring that the

FIGURE 11
ENSTO-E Example of Monte Carlo Simulation Principles



Source: European Network of Transmission System Operators for Electricity (2020a).

weather data be consistent across these inputs. The California event in August 2020 stemmed, at least in part, from a widespread heat wave that seemed highly improbable based on historical patterns but may be more likely now and into the future due to climate change. More changes to resource adequacy analysis and modeling are needed to address both potential conditions and resource availability during these conditions.

An example of this process is shown in Figure 11, which depicts the European Network of Transmission System Operators for Electricity's (ENTSO-E) regional grid planning methodology. In this approach, random unit outages are sampled across many years of synchronized weather data and across many years of annual variations in wind, hydro, solar, and load (ENTSO-E, 2020a). Each weather year is simulated against the same number of stochastic generator outage profiles to create a matrix of weather years and outage draws. The total number of samples is the product of the two, and average resource adequacy statistics are calculated across them. This process allows system planners to identify whether certain weather conditions lead to increased probabilities of shortfall events.

The methodology also helps ensure resource adequacy across an entire range of potential operations, as opposed to just the peak load periods or average weather conditions. Using stochastic production cost methods—combining both chronology and varying weather across a full 8,760-hour analysis—is necessary to help identify times and situations of peak risk. Given that low-probability events drive resource adequacy challenges, a long historical record of weather data is necessary to identify the probability of potential extremes. With higher renewable energy and storage capacity on the grid, these periods are likely to be made up of more combinations, across more variables, than planners were accustomed to in the past.

Data Limitations in Weather Modeling

Analysts and policymakers should be cognizant, however, of data limitations. This methodology is data-intensive and requires a convergence of power systems and meteorological expertise. System planners often have access to long historical records of solar and hydro resources, but may be limited on wind data. In addition, historical data may be available for system load, but underlying changes to consumer behavior, load growth, and distributed energy resources may limit the usefulness of legacy load data from several years in the past. Where long historical records of correlated wind, solar, hydro, and load are not available, planners will need to either use a limited data sample or develop methods that can bootstrap a larger dataset based on correlation of a smaller, but complete, dataset to a longer dataset such as temperature.

In addition, past observations may no longer be good predictors of future conditions with a changing climate. Part of reliability planning is ensuring that the system can maintain reliability during potential—and credible—weather events. The California heat wave saw some of the highest average temperatures in the past 35 years, spread across most of the western United States. Similarly, during the 2021 winter events, Texas saw temperatures well below the near-term historical record, sustained for many days longer than a similarly cold event in 2011. Just because our recent weather data do not include a weather event doesn't mean system planners do not have to prepare for one in the future.

To assume that historical trends continue into the future can also be problematic due to climate change, for two reasons. A changing climate will likely cause weather conditions to diverge from their historical norms and may shift load and renewable generation away from expectations. And climate change may increase the frequency of extreme weather events, which can increase the probability of resource adequacy shortfalls. The European members of ENTSO-E, for example, have identified climate change as a key contributor to resource adequacy risk and are planning to incorporate a climate change trend as a baseline assumption in their resource adequacy process:

> The impact of climate change on adequacy assessments can be significant, considering that an important element of the adequacy models is the underlying climate-dependent data used as input. ENTSO-E is working with climate and data experts to prepare a database that will reliably reflect the impact of climate change on climate variables and, thus, on adequacy simulation results. . . . Our efforts will continue during the upcoming three years, targeting to reliably incorporate in our models the impact of climate change by the end of 2023 (ENTSO-E, 2020b).

Given the uncertainty in the weather, limited data across a long historical record, and potential climatic changes, system planners should identify and evaluate potential drivers of resource adequacy risk, even if they have not occurred or stressed the system in the past. While it will be impossible to assign probabilities to these events, and thus use them to quantify conventional resource adequacy metrics, these drivers can be used to understand potential periods of risk for further investigation and contingency planning.

More planning should be focused on identifying potential situations where the traditional data-driven statistical modeling has limitations, and on testing system reliability. Future resource adequacy analysis should evaluate potential situations that may not have occurred in the past but could reasonably occur in the future. Identification of these high-impact, low-probability events can then be evaluated in isolation to determine whether and how they should be mitigated.

PRINCIPLE 3: There is no such thing as perfect capacity.

As Principle 1 suggests, some capacity shortfalls may consist of frequent but short-duration events, while others may be infrequent but long-duration events. Mitigation strategies will need to be specified accordingly, because different resources bring different capabilities. Battery energy storage may be well suited to solve frequent, short-duration shortages, while demand response may be better suited for infrequent, but challenging, events. Additional resources like long-duration storage, hydro, and thermal generation may be required for long-duration capacity shortages spanning days or weeks.

Different resources bring different capabilities. Battery energy storage may be well suited to solve frequent, short-duration shortages, while demand response may be better suited for less frequent events.

Unfortunately, traditional resource adequacy analysis is designed around a one-size-fits-all approach to resource adequacy additions. Conventional system planning has often treated a natural gas combustion turbine as peaking "firm capacity" and, therefore, a near-perfect capacity resource that could be added to improve reliability. If a system was determined to be short of capacity, combustion turbines were often used as the default resource to bring the system to the reliability criteria. This is because these represented a low-installed-cost resource and could effectively put more "steel in the ground" for reliability. Under this construct, resources like wind, solar, and storage are given partial "firm capacity" credit.

However, gas plants are not always available on demand, as they experience planned as well as weather-related outages. The false dichotomy between the perfect resource and resources with only partial "firm capacity" is due to be replaced by analysis applying the effective load-carrying capability (ELCC) metric to all resource types. ELCC measures the amount of load that can be added to a system given the addition of a resource, while maintaining the same level of reliability as the system prior to the resource addition.

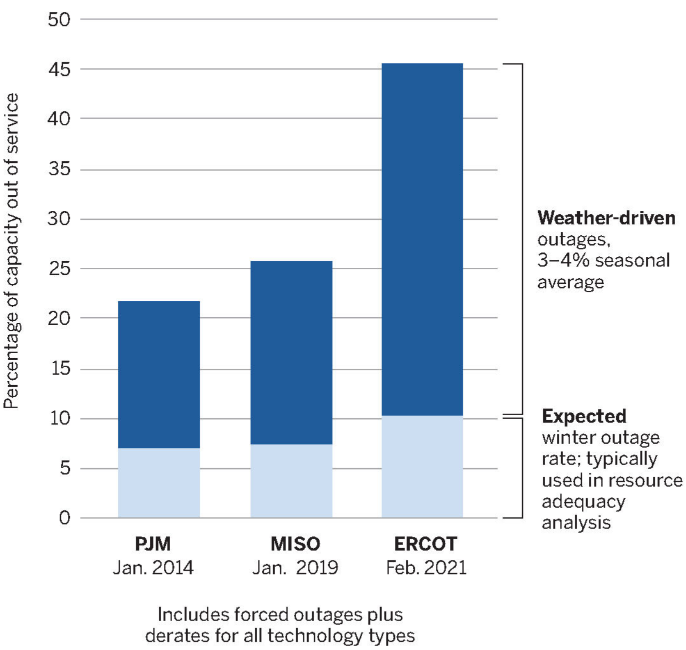
Weather-Dependence of Thermal Generators

The bias toward centering resource adequacy around "firm capacity" and treating a gas turbine as a perfect capacity resource (having an ELCC of 100 percent) causes several problems. First, it assumes that combustion turbines and similar fossil technology are available on demand, and rarely assigns an ELCC to these technologies in a similar manner as wind, solar, storage, and demand response technologies. In some cases, the fossil technology is discounted, but only based on the equivalent forced outage rate on demand (EFORd). For example, a gas turbine with a 5 percent forced outage rate would receive 95 percent capacity credit toward the planning reserve margin.

However, as discussed above, there are times when correlated outages occur on the gas fleet, which increases reliability risk substantially. All generation sources are weather-dependent to some degree. The light blue segments of the bar chart in Figure 12 (p. 19) provide the average forced outage rate of resources throughout the year, whereas the dark blue bar segments show the increase in forced outage rates during extreme cold conditions. Thermal generators, including nuclear, require a water supply which can be threatened by extended drought conditions, and extreme temperatures can force reduced operations. Gas turbines have ambient derates due to high temperatures, forced outage rates that are considerably higher during extreme cold conditions, and a fuel supply that can be jeopardized by competition with gas heating demand. Coal piles can freeze solid. Availability considerations due to weather, supply, and intra-resource correlations should be applied to all resource types. If ELCC is used for capacity accreditation, the methodology should be applied to all resource types, not just variable renewable energy and energy-limited resources.

Second, the conventional "perfect capacity" approach assumes that a resource is needed during all hours and must be dispatchable at any moment to be effective for reliability. In reality, a balanced portfolio of resources, including wind, solar, storage, load flexibility, and fossil technologies can be used to ensure that enough resources are available when needed, despite the inherent uncertainty of individual resources.

FIGURE 12
Total Unplanned Outages During Recent Cold Weather Events

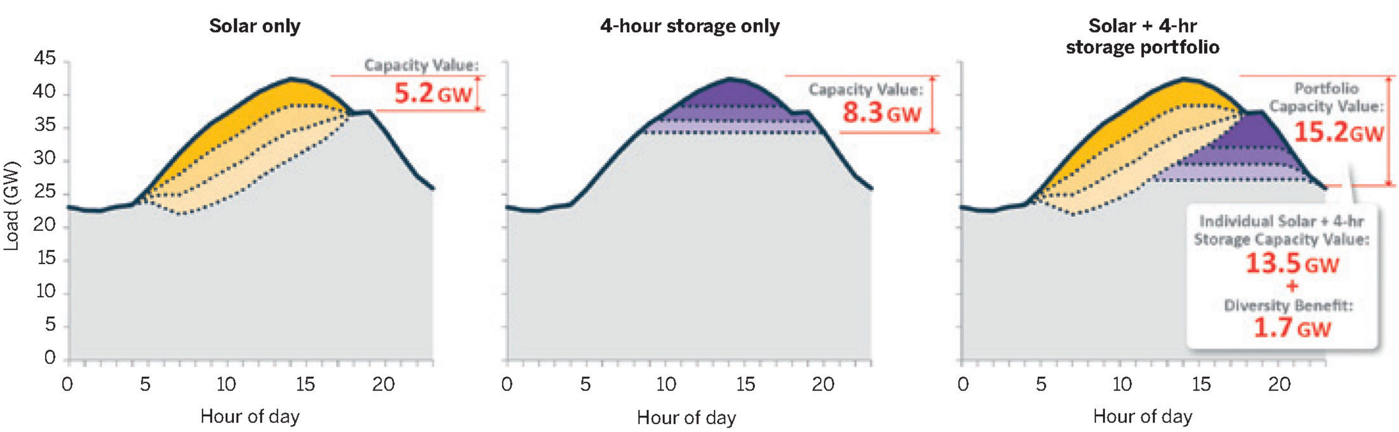


Source: Energy Systems Integration Group.

Recognition of Resources' Limitations and Strengths

ELCC is a useful metric to evaluate reliability contributions and correlated output within and between resource types. Combinations of resources and the interactions between them are important to understand, though currently difficult to quantify. Numerous studies suggest that the ELCC of a resource type is highly dependent on the underlying resource mix and the load profile—both of which change continuously. Figure 13 shows how the ELCC of solar and storage are both higher when evaluated in combination than when evaluated separately. These interactive effects may be either antagonistic, where each increment of solar provides successively lower capacity value, or synergistic, where a portfolio of solar and storage likely provides more value than the sum of its parts (Schlag et al., 2020). In addition, the way in which a system is operated can have an impact on ELCC, especially for energy-limited resources. It is possible, for example, to operate a storage system to maximize resource adequacy, which could differ at times from operation that maximizes revenues. The ability to accurately forecast system conditions can also change ELCC accreditation.

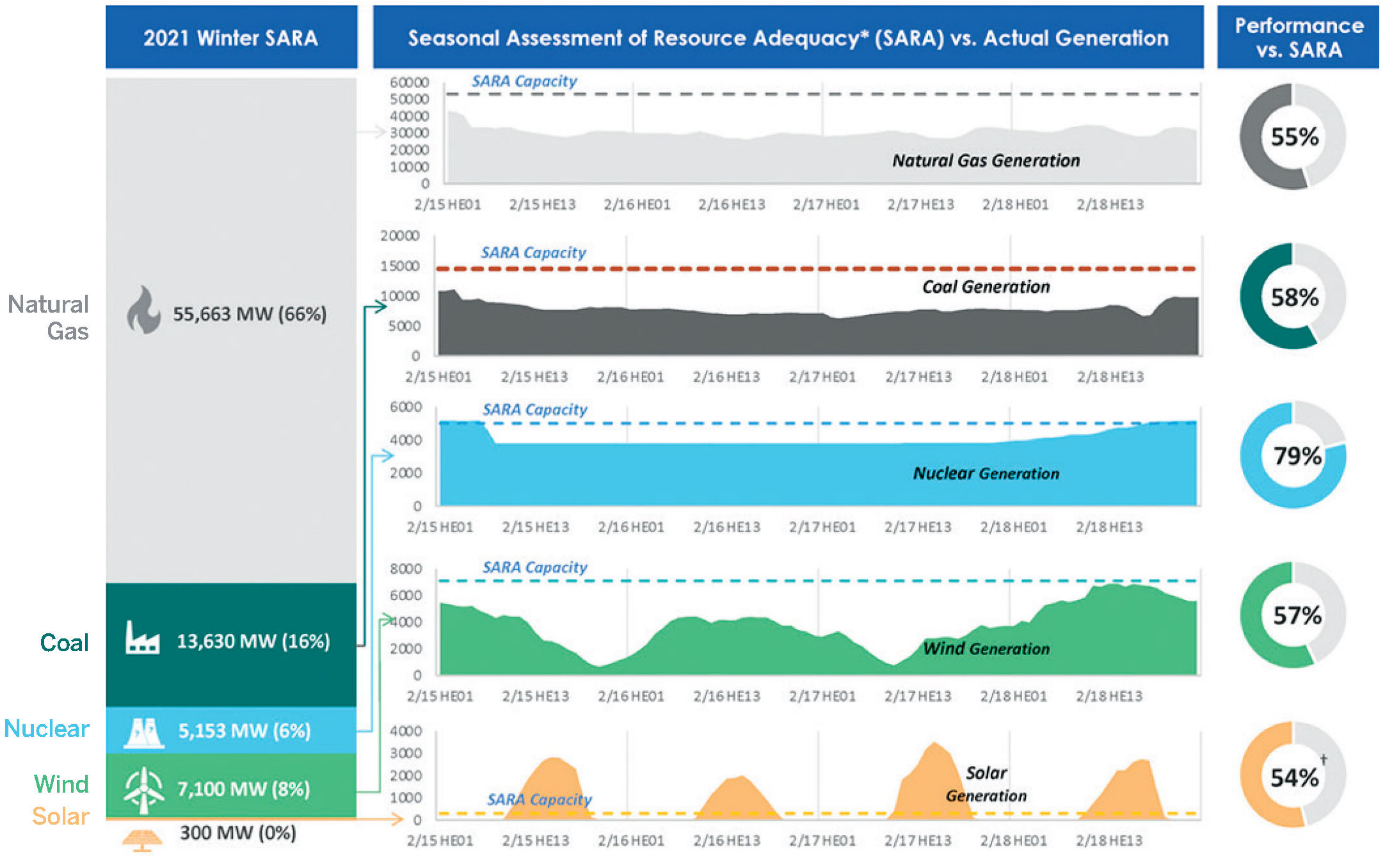
FIGURE 13
The ELCC of Solar Alone, Storage Alone, and the Two Resources in Combination



Source: Energy and Environmental Economics (E3) / Schlag et al. (2020).

FIGURE 14

Resource Adequacy Capacity Credit vs. Actual Generation During the Texas 2021 Event



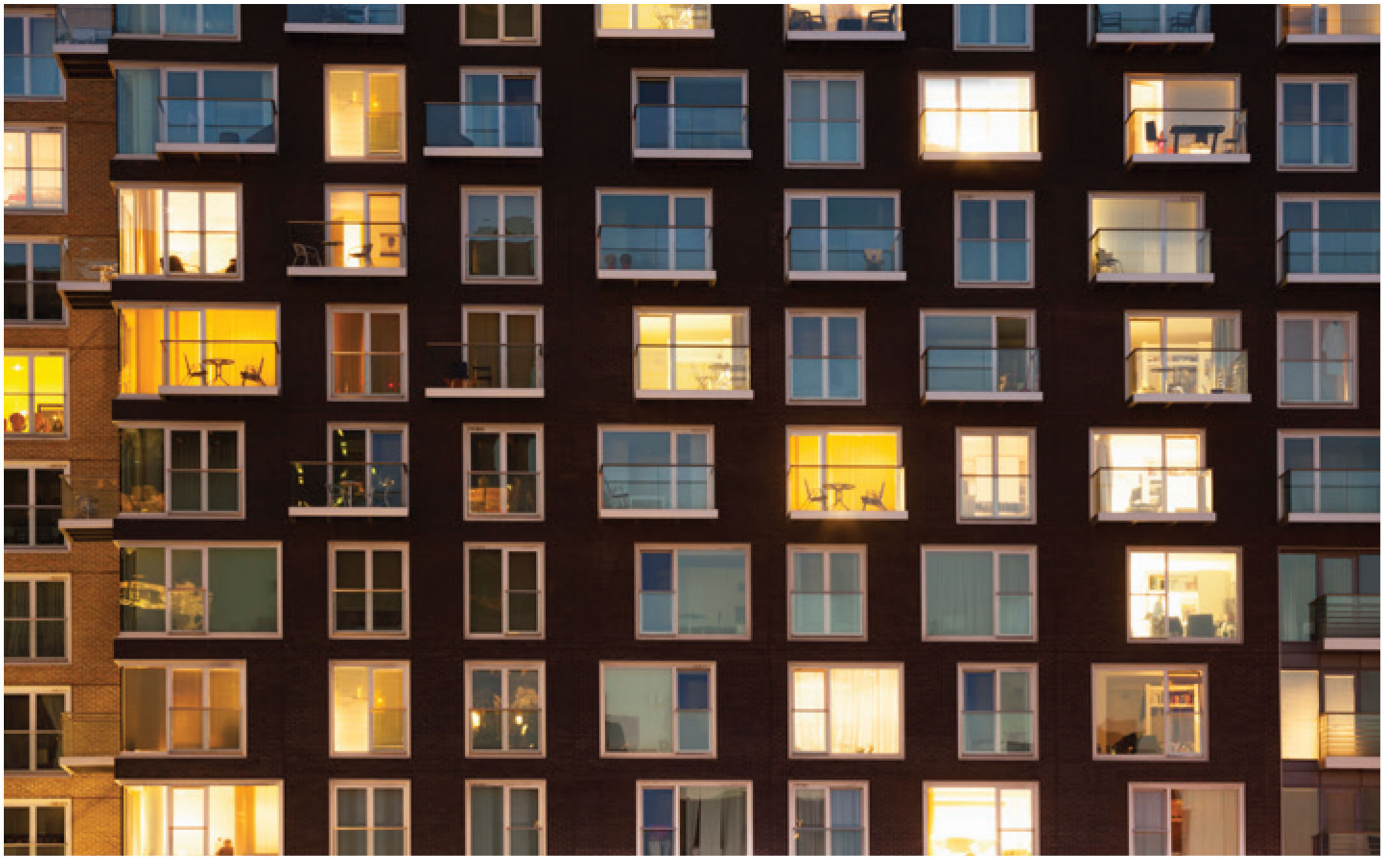
ERCOT electricity generation versus seasonal expected availability (February 15-18, 2021).

* ERCOT's Seasonal Assessment of Resource Adequacy (SARA) attributes an expected available capacity to each generation type, considering seasonal factors.
† During peak demand (18-23h); over all hours: 259%

Source: Eamonn Lannoye, Electric Power Research Institute; data from Electric Reliability Council of Texas (2021a, 2021b).

Future resource adequacy analysis should explicitly recognize that all resources have limitations based on weather-dependence, potential for outages, flexibility constraints, and common points of failure. This was abundantly evident in the extreme cold events in Texas in February 2021. In this case, all of the resources on the system were similarly strained, and no resource was able to contribute as planned for reliability, demonstrating that there is no such thing as perfect capacity. As shown by the Electric Power Research Institute's analysis in Figure 14, natural gas, coal, nuclear, and wind resources all provided energy well below their expected levels when compared against the ERCOT *System Assessment of Resource Adequacy* report.

Resource adequacy analysis for modern power systems should also recognize that each resource brings different capabilities that may work best in specific situations. For example, frequent but short-duration events could be best addressed by battery storage or load flexibility (via time-of-use rates), whereas infrequent, short events could be best addressed by load-shed demand response programs. Frequent and long events may require long-duration resources like fossil generation, long-duration storage, or hydro resources, and infrequent, long events may be best handled through coordination with neighboring grids or emergency procedures.



PRINCIPLE 4: Load participation fundamentally changes the resource adequacy construct.

In traditional resource adequacy analysis, load was treated as a static—that is, uncontrollable—input into the modeling and simulations. Load was increasing year over year, and the purpose of resource adequacy analysis was to determine whether there was enough generation capacity to serve a fixed load. The question for planners was simply, do we have enough generation to meet our load requirement? While there was uncertainty in the load forecast, a higher load forecast just meant additional supply-side resources were needed.

But the historical notion that a specific amount of generation capacity is required to meet a static load is no longer relevant. Load flexibility is increasing quickly. The decreasing costs of distributed sensors and controls, increased proliferation of distributed energy resource aggregation, and increased visibility into behind-the-meter load consumption have all made loads more flexible, price responsive, and intelligent.

Using Load Flexibility for Resource Adequacy

The proliferation of energy storage, demand response, electric vehicles, and dynamic rate design bring with them new options for load flexibility and should be evaluated in a similar context as generation resources, including uncertainty and availability. However, future load flexibility is based on customers' economic decisions. Real-time markets, with a high degree of participation from price-responsive demand, may place more attention on economic considerations rather than reliability needs, as customers can determine and differentiate which loads matter most. It is therefore important to ensure that reliability benefits of flexible loads are not lost, and that various forms of load flexibility—and their associated reliability benefits—are included in resource planning assessments.

Two examples in Hawaii include procurement of capacity grid services from virtual power plants and price-responsive loads. A recent utility procurement is leveraging a virtual power plant of 6,000 residential photovoltaic+battery systems (25 MW, 80 MWh) to provide both "load build" and "load reduce" grid services; this allows the utility to call on behind-the-meter batteries to charge or discharge when grid conditions require. The utility also recently announced another 60 MW procurement target for virtual power plants and price-responsive loads to provide resource adequacy benefits for an upcoming coal retirement (Pickerel, 2021).

Resource adequacy and power system planning should consider load flexibility as a supply-side resource capable of reducing system risk of shortfalls.

Another potential mechanism for increased load participation for resource adequacy benefits is through energy-only markets with price scarcity driving customer behavior. For a full energy-only market to work, value of lost load–based scarcity pricing is needed, along with a market structure that ensures that market participants have both the incentive and ability to procure power in advance or can fully handle any risk of paying scarcity-based prices if they wind up with a short position.

Scarcity pricing, even without an underlying resource adequacy construct, creates two incentives for resource adequacy. First, it provides a clear incentive to reduce loads or switch to behind-the-meter generation during scarcity events. Second, it provides a clear incentive for load-serving entities to enter into bilateral contracts for capacity as a hedge against price volatility.

Additional mechanisms exist to increase load flexibility. This can be done dynamically, via real-time pricing and direct distributed energy resource aggregation and control, or more passively via time-of-use rates, critical peak pricing, energy efficiency programs, and education. Regardless of the method, resource adequacy and power system planning should consider load flexibility as a supply-side resource capable of reducing system risk of shortfalls.

System Planners' Data Needs for Load Resources

As load becomes more flexible, the options to balance both the supply and demand sides of the resource adequacy equation become much more dynamic. However, the industry does not have the same institutional knowledge base and experience with demand-side resources as it does with supply-side generation. If system planners and operators are going to rely on load flexibility to the same extent as supply-side resources, more information is needed on potential unavailability of load flexibility, uncertainty in participation, and scheduling constraints that may affect load resources' utilization.

In other words, the modeling of demand-side resources will require the same level of inputs used to model generation. This includes an equivalent to planned and forced outage rates, seasonal derates, energy and duration limitations, constraints on the number of starts per year/month/day, variable operating costs, and other characteristics typically used to simulate a generation resource. Fortunately, as demand-side load flexibility continues to proliferate, experience and data are also growing. These resources constitute a flexible, modular, and dynamic resource for solving resource adequacy challenges without installing more generation that would be used sparingly, if ever, for reliability needs.

PRINCIPLE 5: Neighboring grids and transmission should be modeled as capacity resources.

Resource adequacy modeling can be complex and is often computationally challenging; a large power system must typically be simulated across hundreds or thousands of Monte Carlo samples. This challenge is further amplified by the increasing need to model full chronology across an entire year of operations (Principle 2). To make this problem tractable, simplifications are required. Often that means only limited representation of neighboring power systems and the transmission network in general.

However, resource sharing can be a significant, low-cost alternative to procuring new resources. Imports from neighboring regions are likely to become more valuable for resource adequacy due to the increased diversity of chronological wind, solar, and load patterns over a much larger area. A typical wind plant output tends to have little correlation with other wind plants a few hundred miles away. Solar output varies with cloud cover and time zones. Load diversity is greater across large areas. While extreme weather can happen anywhere, it does not happen everywhere at once.

Neighboring systems are often simplified in resource adequacy analysis because the system planners' and regulators' perspective has traditionally been that the power system should be self-reliant and able to serve load without requiring imports from neighbors during critical time periods. This mindset is not unreasonable; ultimately the utility or system operator is responsible for reliably meeting its customers needs, regardless of what happens in neighboring regions.

However, the preference for self-reliance leads to a potentially large and expensive overbuild of capacity. If every region is carrying its own margin, which is only used sparingly for reliability, the cost of surplus resources is amplified across the interconnected power system. The value of sharing between adjacent regions was a major driver for the formation of independent system operators and regional transmission organizations, which allowed many smaller, vertically integrated utilities to pool their resources and reduce coincident peak load (achievable when the peak regional load is lower than the sum of individual localities' peaks due to load diversity).

Ultimately, it is up to regulators, policymakers, and system planners to determine the level of reliance on neighbors that is acceptable, given the local conditions and resource mix. There is no right or wrong answer.

An Economic Opportunity Too Large to Ignore

There is a very large economic opportunity in increasing regional coordination, sharing of resources, and relying on imports to meet reliability needs. Major benefits include:

- **Staggered peaks.** Load diversity increases with large geographies, varied weather patterns, multiple time zones, and demographic differences. The larger the system, the less likely peak loads are to occur simultaneously.

Key to unlocking this economic opportunity is transmission, to enable flows between regions and create interregional resource diversity.

- **More consistent renewable generation.** As the planning footprint increases in size, the wind and solar variability diminishes. While the skies can be cloudy and winds can be calm anywhere, it will not likely be cloudy and calm everywhere.
- **Less chance of simultaneous outages.** A larger portfolio of resources means lower probability of simultaneous outages across a large portion of the resource mix, as each region has access to a larger number of generating units and higher installed capacity.
- **Less chance of outages caused by fuel shortages.** There is a lower probability of outages due to fuel shortages because a larger region likely has multiple fuel delivery paths, such as natural gas pipelines.

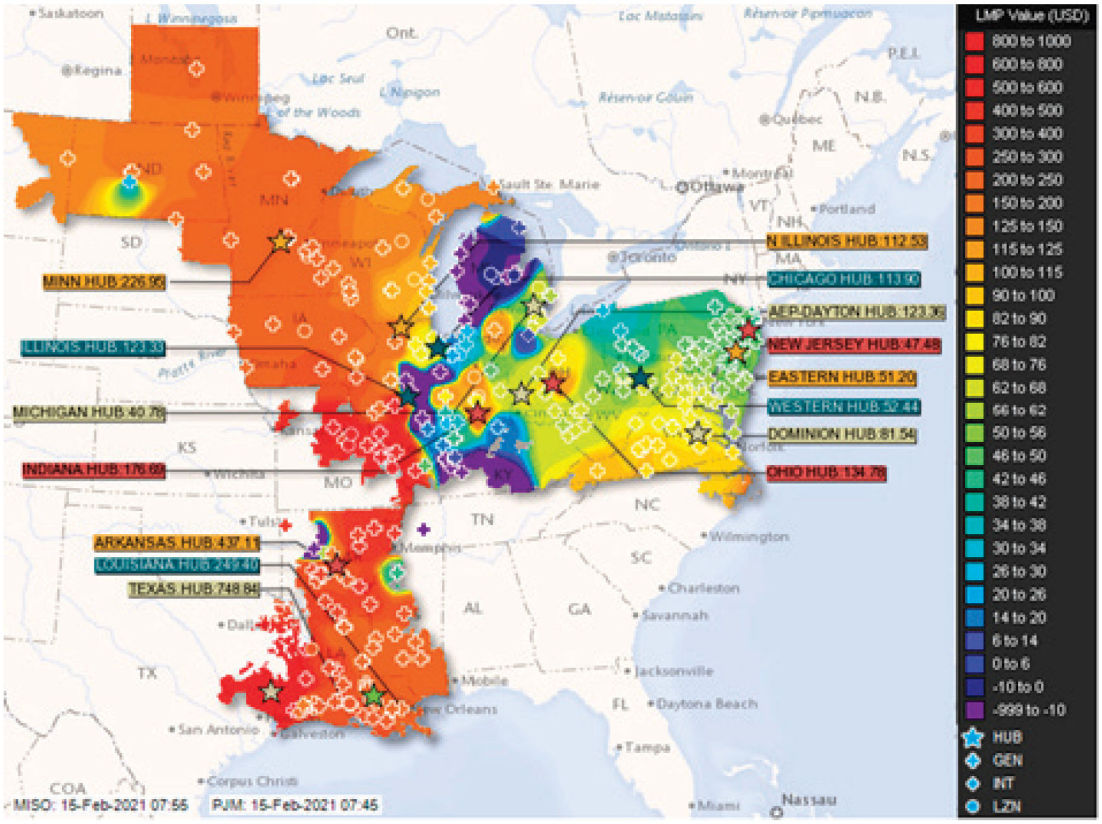
Key to unlocking this economic opportunity is transmission, to enable flows between regions and create interregional resource diversity. Transmission assets should therefore be clearly identified as having resource adequacy benefits.

This principle is clearly illustrated by market data from the February 2021 extreme cold weather event when ERCOT and MISO experienced capacity shortfalls due to cold temperatures (Figure 15, p. 24). As the middle of the country struggled to meet load, much of the east coast was experiencing normal temperatures and had surplus capacity, indicated by significantly lower energy prices and a clear gradient across the MISO-PJM seam (JCM, 2021). Additional transmission capacity between regions could have mitigated some of the resource adequacy failure.

The same benefits can be had locally, available to zones within a single balancing authority. For example, NYISO and PJM both have nested capacity zones. In the New York example, the lower Hudson Valley (Zones G-J),

FIGURE 15

Real-Time Energy Prices During a MISO Resource Adequacy Shortfall Event



Source: Joint and Common Market (2021).

New York City (Zone J), and Long Island (Zone K) have local capacity requirements because there is limited transmission capability into the zones, which also have the highest loads. As a result, from a resource adequacy perspective these zones have higher probability of a shortfall and thus require additional local capacity. If increased transmission capability was constructed, these zones could share resources with neighboring zones, thus decreasing the resource adequacy risk and lowering the amount of local generating capacity needed for reliability.

Modeling and Policy Needs for Transmission Coordination

While the pooling of resources improves reliability, it raises questions about how to appropriately share resources during times of resource adequacy risk. This introduces a policy and regulatory challenge about how to balance reliance on neighbors and self-sufficiency. Politicians and system operators are beholden to their own constituents and customers for reliability—not those in other regions. However, constituents and customers also rely on them for affordable reliability. The desire for self-reliance must be informed by the affordability offered by the option of using neighbors' resources rather than investing in redundant resources.

Integrating neighboring areas into resource adequacy analyses requires some advances in the modeling and policy/regulatory arenas. First, conventional resource adequacy analysis tends to do a poor job of modeling neighboring systems due to the preference of self-reliance and to computational limitations. So, some modeling simplifications may be necessary to make the problem size tractable, and this will need to be done with care and deliberation. But ultimately, more precise modeling of neighboring systems can lend confidence that, statistically speaking, some resources in neighboring systems will likely be available during times of highest reliability risk.

To address these issues, resource adequacy analysis for modern power systems should include two things. First, transmission assets should be evaluated as a capacity resource if they allow additional flow to enter into a capacity- and transmission-constrained region. This provides an alternative resource, beyond just local generation and load flexibility, to meet resource adequacy requirements. Second, resource adequacy analysis should provide a detailed representation of the neighboring systems so that the same probabilistic assessments can be made in neighboring regions in order to provide more fidelity in the availability of imports. This will be computationally challenging but necessary, given the degree to which the transition to higher levels of variable renewable energy is enabled by taking advantage of geographic diversity.

This type of transmission coordination is not just a matter of increased interties between regions, but just as importantly requires market mechanisms that allow for transparent sharing of resources across balancing areas, regional transmission organizations, and other jurisdictions. There are a wide variety of market mechanisms available, ranging from voluntary capacity agreements and bilateral contracts to formal capacity markets. Regardless of the form, establishing clear market rules for capacity sharing is a critical regulatory and policy need for the coming years.

PRINCIPLE 6: Reliability criteria should be transparent and economic.

Conventional resource adequacy analysis largely ignores economic principles and excludes the financial impacts on consumers of a system's reliability choices. In this context, reliability is binary: a system's resource mix is considered either reliable or not when compared to a reliability criterion. The cost of achieving this reliability has not typically been taken into account.

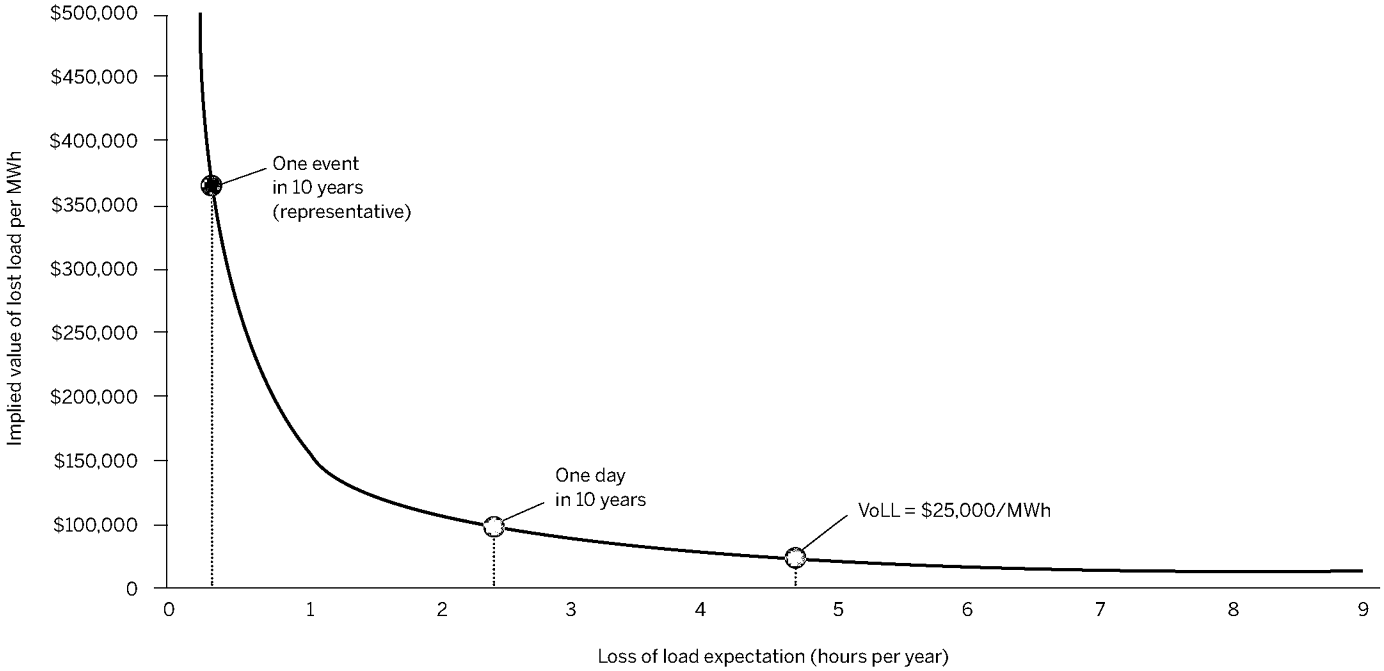
Today, there are many more pieces to the reliability puzzle. Mitigations now include fossil resources, solar, wind, various storage technologies, hybrid resources, various configurations of demand response and load flexibility, and transmission. Different combinations of resources offer distinct reliability profiles, with reliability trade-offs and different costs.

Lack of Transparency Around an Arbitrary Criterion

As discussed above, a common reliability criterion used by many system planners is 1-day-in-10-years LOLE. However, this criterion was developed in the middle of the 20th century, with limited rationale as to its selection and limited evaluation of the costs and benefits of achieving this definition of reliability. The arbitrary nature of the 1-day-in-10-year LOLE criterion is concerning, despite its use as the de facto reliability standard across a wide range of different systems having heterogeneous resource mixes, consumer needs, regulatory structures, and markets.

Simply put, the 1-day-in-10-year LOLE criterion is an arbitrary line in the sand. System planners and regulators set the criteria and determine a portfolio to be reliable or not, regardless of the costs incurred to ratepayers. Decisionmakers are left without knowledge of the costs necessary to achieve the target reliability, and they rarely consider the costs and benefits of measures taken to increase reliability.

FIGURE 16
Comparing the Cost with the Value of Adding Resources for Reliability



Source: Regulatory Assistance Project / Hogan and Littell (2020).

The implications of this lack of awareness are great. Resource adequacy analysis sets the foundation for resource procurement and investment decisions by vertically integrated utilities, and it sets the quantity needs for competitive capacity markets. Although the financial impact of meeting the reliability criteria is large, the current lack of transparency around the costs of different approaches to reliability makes it impossible to perform a rigorous cost-benefit analysis.

Nonlinear Relationship Between Reliability and Costs

The modern power system is much more dynamic than systems of the past. For example, as consumers become increasingly aware of their energy consumption, costs, and alternative objectives like environmental impact, load flexibility becomes an important resource. As a result, new resource adequacy analysis should be designed to increase cost transparency so that regulators, policymakers, and consumers understand the relative costs of different levels of and approaches to reliability and can make informed investment decisions.

Figure 16 shows an example of the relationship between cost and reliability. On the x-axis is LOLH per year for a given system. To make the system more reliable (moving from right to left), additional gas turbine capacity is added, reducing LOLH but leading to an increase in costs. On the y-axis is the implied (implicit) value of lost load, or the incremental change in cost relative to the change in reliability (Hogan and Littell, 2020).

The chart shows a highly non-linear relationship between reliability and cost, and illustrates that a 1-day-in-10-year reliability criterion could be much more expensive to consumers than higher levels of reliability achieved by other means. While value of lost load (VoLL) is a highly debated metric—and varies considerably based on customer type—transparency in the costs of the reliability criterion is critical. Although it may be impossible to identify an economically efficient reliability level because it is hard to speculate how much reliability is worth to a diverse group of customers, there needs to be a clear understanding among policymakers, regulators, and system planners of what incremental reliability costs consumers. Such transparency could reveal that, in some cases, incremental reliability is relatively affordable and worth the investment, while in other cases it is extremely expensive but purchased anyway because it's hidden beneath an arbitrary reliability requirement. Transparency in the cost versus reliability relationship will allow wiser decisions around reliability improvements going forward.

Understanding Resource Adequacy's Share of Overall Reliability

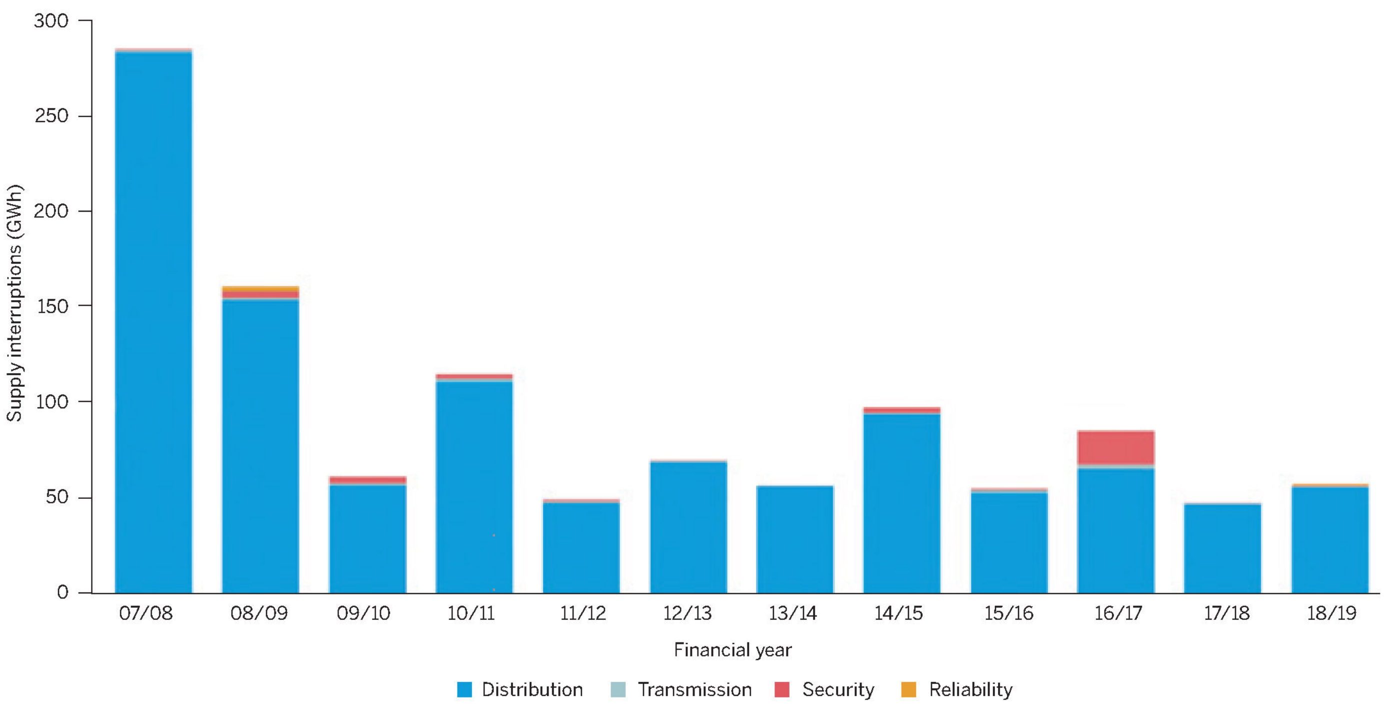
Factors other than resource adequacy also play a role in power system reliability, of course. For example, failures can be due to distribution outages, transmission outages, network instability, and cyber attacks. Setting reliability requirements for resource adequacy must be balanced with allocating resources toward other forms of reliability. Ultimately, the consumer does not differentiate between reasons for lost power. System planners need to make sure that the benefits assumed from a resource adequacy requirement and capacity procurements are actually those needed to ensure grid reliability, as opposed to investments in transmission and distribution infrastructure, grid hardening, and cyber defense.

Without translating the reliability requirements to economic costs as a common comparison, it is impossible to know whether reliability dollars are being allocated efficiently.

Without translating the reliability requirements to economic costs as a common comparison, it is impossible to know whether reliability dollars are being allocated efficiently. Without this economic consideration, system planners risk over-procuring capacity without significantly increasing system reliability. Given limited resources, system planners need to allocate investment appropriately across other facets of reliability.

Figure 17, showing the sources of supply interruptions in the Australian National Energy Market from 2007 through 2019, indicates that only a tiny fraction (0.3

FIGURE 17
Lost Load Energy in Australia by Reliability Type from 2007 to 2016



Source: Australian Energy Market Commission Reliability Panel (2020).

Grid planners and regulators should have a clear understanding of the costs associated with achieving different reliability targets in different ways.

percent of all lost load) was due to capacity shortfalls (AEMCRP, 2020). The vast majority of lost load and customer outages were due to failures and outages on the distribution system, and a single security-related event (South Australia blackout) in 2016. This situation is not uncommon, indicating that the industry may be focusing too much on reliability based on resource adequacy and too little on distribution-system reliability. Transparently showing the economic costs of incremental resource adequacy improvements is critical to understanding the different sources of reliability for each system.

The costs of achieving 100 percent resource adequacy on a high-renewables grid would be infinite, and senseless for most consumers when the same money could be spent on other reliability mitigations. A single resource adequacy criterion centered solely on the number of MW, absent economic considerations, is therefore unjustified. Grid planners and regulators should have a clear understanding of the costs associated with achieving different reliability targets in different ways, to ensure that the value provided to the customer is worth the cost of a given investment—that the resource adequacy for which customers are being asked to pay is actually the type of reliability needed on the grid.

Looking Forward

If the rolling blackout events in California and Texas in 2020 and 2021 teach us anything, it is that the industry cannot continue to approach resource adequacy as we have in the past. These were not failures of the evolving resource mix, but rather failures of planning. Existing methods that have served the industry well historically are not adequate as the resource mix changes to one of variable renewable energy, energy storage, and flexible loads, and as power systems experience increased correlation of generator outages due to weather. Today, the industry, and ultimately consumers, are paying the price of limited planning and analytical shortcuts that do not capture chronological operations and weather-influenced correlated events.

Consistency in resource adequacy analysis and reporting will provide the necessary data and better insight on what shortfall events look like across many systems.

Many of the metrics currently used, such as the traditional planning reserve margin, are not adequate for modern power system planning. The industry needs new options. For now, we must rely on more in-depth analysis of real systems, or general rules that can be applied. As electricity system stakeholders, we need to roll up our sleeves and do the hard analytical work. What we learn will help us develop heuristics and new rules to make resource adequacy analysis easier and less costly to conduct and simpler to understand.

While considerable work is needed to fully define what robust resource adequacy looks like, some basic first steps can lead to improved resource adequacy analysis now. These steps include:

- Considering how the first principles of resource adequacy should be applied to the specific system being examined
- Making the resource adequacy analysis public and easily accessible, so that the community of stakeholders can benefit from seeing a diverse set of case studies from regions around the world with different resource mixes, load profiles, and characteristics of system risk
- Collecting as much chronological and correlated hourly historical weather and load data as possible, and then considering whether the available historical data are sufficiently representative of possible future events, including consequences of climate variability and change
- Reporting a broader set of resource adequacy metrics than simply an average LOLE, including hourly EUE and additional information on the distribution of outages. Metrics should also be used to develop detailed statistics on the shortfall events themselves in order to better characterize the size, frequency, duration, and timing of events so that mitigation measures can be properly sized.

Consistency in resource adequacy analysis and reporting will provide the necessary data and better insight on what shortfall events look like across many systems. Such consistency will help the industry better understand how resource adequacy risk shifts with changes in the underlying resource mix of increased variable renewable energy, energy storage, and load flexibility of modern power systems.

References

AEMCRP (Australian Energy Market Commission Reliability Panel). 2020. *Annual Market Performance Review*, 2020. Sydney, Australia. https://www.aemc.gov.au/data/annual-market-performance-review-2020/sources-of-supply-interruptions-in-the-nem-2008-09-to-2018-19.

APPA (American Public Power Association). 2021. *America's Electricity Generation Capacity, 2021 Update.* Washington, DC. https://www.publicpower.org/system/files/documents/Americas-Electricity-Generation-Capacity_2021-update.pdf.

Calabrese, G. 1947. "Generating Reserve Capacity Determined by the Probability Method." *AIEE Transactions* 66:1439-1450.

CAISO (California Independent System Operator). 2021. *Final Root Cause Analysis: Mid-August 2020 Extreme Heat Wave.* Folsom, CA. http://www.caiso.com/Documents/Final-Root-Cause-Analysis-Mid-August-2020-Extreme-Heat-Wave.pdf.

ENTSO-E (European Network of Transmission System Operators for Electricity). 2020a. *Mid-term Adequacy Forecast 2020: Appendix 2; Methodology.* Brussels, Belgium. https://eepublicdownloads.entsoe.eu/clean-documents/sdc-documents/MAF/2020/MAF_2020_Appendix_2_Methodology.pdf.

ENTSO-E. 2020b. *Resource Adequacy Implementation Road Map.* Brussels, Belgium. https://eepublicdownloads.entsoe.eu/clean-documents/cep/implementation/BZ/201210_ERAA_implementation_Road_Map.pdf.

ERCOT (Electric Reliability Council of Texas). 2021a. Daily Balancing Authority Operations Report. Austin, TX. http://mis.ercot.com/misapp/GetReports.do?reportTypeId=13458&reportTitle=Daily%20Balancing%20Authority%20Operations%20Report&showHTMLView=&mimicKey.

ERCOT. 2021b. *Final Seasonal Assessment of Resource Adequacy for the ERCOT Region (SARA) Winter 2020/2021.* Austin, TX. http://www.ercot.com/content/wcm/lists/197378/SARA-FinalWinter2020-2021.pdf.

ERCOT. 2021c. *Update to April 6, 2021, Preliminary Report on Causes of Generator Outages and Derates During the February 2021 Extreme Cold Weather Event.* Austin, TX. http://www.ercot.com/content/wcm/lists/226521/ERCOT_Winter_Storm_Generator_Outages_By_Cause_Updated_Report_4.27.21.pdf.

Hering, G., and J. Stanfield. 2020. "'You have to rethink these old ways': Parting Advice from CAISO's Retiring CEO." S&P Global Market Intelligence, September 25. https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/you-have-to-rethink-these-old-ways-parting-advice-from-caiso-s-retiring-ceo-60481529.

HNEI (Hawai'i Natural Energy Institute). 2020. *Report to the 2021 Legislature, Annual Report from the Hawai'i Natural Energy Institute.* Honolulu, HI. https://www.hnei.hawaii.edu/wp-content/uploads/2021-HNEI-Annual-Report.pdf.

Hogan, M., and D. Littell. 2020. *Get What You Need: Reclaiming Consumer-Centric Resource Adequacy.* Montpelier, VT: Regulatory Assistance Project. https://www.raponline.org/knowledge-center/get-what-you-need-reclaiming-consumer-centric-resource-adequacy.

JCM (Joint and Common Market). 2021. JCM Contour Map. Carmel, IN. https://www.miso-pjm.com/markets/contour-map. Accessed February 15, 2021.

MISO (Midcontinent Independent System Operator). 2021. *MISO's Renewable Integration Impact Assessment, Summary Report.* Carmel, IN. https://cdn.misoenergy.org/RIIA%20Summary%20Report520051.pdf.

Pickerel, K. 2021. "Swell Energy to Establish 6,000-battery Virtual Power Plant for Hawaiian Electric." *Solar Power World,* January 19. https://www.solarpowerworldonline.com/2021/01/swell-energy-to-establish-6000-battery-virtual-power-plant-for-hawaiian-electric.

Schlag, N., Z. Ming, A. Olson, L. Alagappan, B. Carron, K. Steinberger, and H. Jiang. 2020. *Capacity and Reliability Planning in the Era of Decarbonization: Practical Application of Effective Load Carrying Capability in Resource Adequacy.* San Francisco, CA: Energy and Environmental Economics. https://www.ethree.com/wp-content/uploads/2020/08/E3-Practical-Application-of-ELCC.pdf.

Attachment A

Redefining Resource Adequacy for Modern Power Systems

A Report of the Redefining Resource Adequacy Task Force of the Energy Systems Integration Group

---

The report is available at https://www.esig.energy/reports-briefs.

To learn more about the recommendations in this report, please send an email to resourceadequacy@esig.energy.

The Energy Systems Integration Group is a nonprofit organization that marshals the expertise of the electricity industry's technical community to support grid transformation and energy systems integration and operation, particularly with respect to clean energy. More information is available at https://www.esig.energy/reports-briefs or info@esig.energy.

ESIG
ENERGY SYSTEMS INTEGRATION GROUP