Capacity Resource Accreditation for New England's Clean Energy Transition

REPORT 2: OPTIONS FOR NEW ENGLAND

PREPARED BY

Samuel A. Newell
Kathleen Spees
John Higham

PREPARED FOR

Massachusetts Attorney General's Office

JUNE 28, 2022

Brattle

NOTICE

This report was prepared for the Massachusetts Attorney General's Office with funding from The Barr Foundation, in accordance with The Brattle Group's engagement terms. It is intended to be read and used as a whole and not in parts. This document is a follow up to our report entitled*, Report 1: Foundations of Resource Accreditation* ("Report 1"), which was issued June 2, 2022 (publicly posted here).The report reflects the analyses and opinions of the authors and does not necessarily reflect those of The Brattle Group's clients or other consultants. There are no third-party beneficiaries with respect to this report, and The Brattle Group does not accept any liability to any third party with respect to the contents of this report or any actions taken or decisions made as a consequence of the information set forth herein.

TABLE OF CONTENTS

Executive Summary

As the New England Power Pool (NEPOOL) and the Independent System Operator of New England (ISO-NE) embark on modernizing resource adequacy accreditation for the clean energy transition, the Massachusetts Attorney General's Office (AGO) of Ratepayer Advocacy has asked us for assistance. Our first report, *Capacity Resource Accreditation for New England's Clean Energy Transition, Report 1: Foundations of Resource Accreditation* ("Report 1"), established a problem statement and outlined the technical questions, conceptual basis, and criteria for evaluating accreditation reform options. It also summarized how several other jurisdictions are addressing similar problems. This second report applies the framework established in the first report to evaluate a range of potential options and provide recommendations for assessing resources' contributions to resource adequacy in New England.

Accurate resource accreditation in New England's Forward Capacity Market (FCM) will have to measure reliability contributions in light of the region's most pressing reliability challenges: meeting summer and winter peak demands; ensuring sufficient energy and capacity supplies throughout winter cold snaps when generators' fuel supply can be limited; and accounting for the region's increasing reliance on intermittent and energy-limited resources.

Among these challenges, ISO-NE has identified winter fuel supply adequacy as an immediate and severe concern.[1] Current resource accreditation methods do not account for the non-availability of resources during cold snaps. But enhancing resource accounting under ISO-NE's current annual (but summer-focused) resource adequacy construct is not the best way to meet winter reliability needs. A more comprehensive solution would: (a) adopt a seasonal capacity market construct with separate reliability requirements for the summer and winter seasons; (b) separately credit resources' summer and winter capacity ratings based on their contribution to reliability in each season; (c) procure capacity commitments to meet each season's need, considering that a different resource mix may apply in each season; and (d) produce separate seasonal prices reflecting the availability or scarcity of cost-effective capacity to meet each season's needs. In this report, we offer analysis and recommendations on resource accreditation for all resources that would be valid under the current annual approach as well as under a seasonal capacity market approach; however, we caution that possible winter reliability challenges will remain

[1] See, for example, Katie S. Dykes, Connecticut Department of Energy and Environmental Protection, "Re: Fuel Security in New England for Winter 2021–2022," Letter to Gordon van Welie, ISO-NE, December 17, 2021; Gordon van Welie, ISO-NE, Letter to Katie S. Dykes, Connecticut Department of Energy and Environmental Protection, December 23, 2021; New England States Committee on Electricity, Letter to ISO New England, January 18, 2022.

unclear and potentially unmet until the winter need is clearly defined, quantified in a technology-agnostic fashion, and formalized within the resource adequacy construct.

The reality of the clean energy transition must also be adopted as a central assumption as New England adopts a resource accreditation framework that can remain relevant for more than a short period. As states and consumers across the region advance toward a fully decarbonized electricity grid over the coming one to two decades, the capacity market must be designed in a way that will be robust to a future that has fossil-fuel resources shifting to a back-up role and eventually retiring. As fossil plants retire, the system will increasingly depend on clean resources with high or complementary resource adequacy value such as batteries, demand response, hybrid resources, and other clean capacity technologies. An accurate capacity accounting system will serve the dual purposes of: (1) ensuring that reliability is maintained throughout all stages of clean energy transition; and (2) providing incentives for innovative players to build and operate resources that will maximize their reliability contributions to the system.

Building on the traditional and innovative approaches utilized or proposed in other jurisdictions' resource adequacy constructs, we evaluate several distinct methods that could be utilized to improve upon ISO-NE's current Installed Capacity (ICAP) MW accreditation system.[2] The options we consider are:

- **Historical Availability Measurements in Pre-Defined Hours:** Measures resources' historical availability (or historical capacity factor, in the case of intermittent resources), for example, based on offers into the real-time energy market during pre-defined windows when shortfalls are most expected to occur. This approach reduces capacity ratings to an unforced capacity (UCAP) MW level that somewhat accounts for resources' outage rates and non-availability during critical times.
- **Historical Tight-Intervals Measurements:** Measures the historical availability and/or performance of resources to deliver energy and ancillary services during the tightest realized hours or scheduling intervals, such as the 200 tightest hours over the past several years, placing a higher weighting on performance during times of realized operating reserve shortages and load-shedding events.
- **Simulated Marginal Effective Load Carrying Capability (ELCC) or Marginal Reliability Impact (MRI):** Uses a reliability modeling platform to estimate the reliability contribution of 1 MW of

[2] In an ICAP accounting convention, traditional thermal generators and energy storage resources have capacity accreditation based on maximum output capabilities during peak demand conditions, without considering outage rates or fuel supply limitations. In contrast, an unforced capacity (UCAP) convention accounts for resource outage rates when determining resource accreditation.

each resource or resource type relative to a MW of "perfectly available" capacity.[3] As a hypothetical example, removing 1 MW of wind would reduce simulated nameplate wind capacity from 5,000 to 4,999 MW and may require 0.2 MW of "perfect" capacity to be added to restore reliability to the standard. In this scenario, 1 MW of nameplate wind has a perfect capacity equivalent of 0.2 MW (or 20% of nameplate), so wind resources would be credited at 20% of nameplate as marginal ELCC value.

- **Simulated Average ELCC:** Similarly, uses a reliability modeling platform but displaces an (arbitrarily defined) entire class or classes of resources to quantify the equivalent in perfect capacity, expressing a total or average value that may differ substantially from the marginal value. For example, 5,000 MW nameplate of wind, if removed from the model, may need to be replaced by 1,500 MW of "perfect" capacity to restore the system to the reliability standard. Each such wind resource would then be credited under average ELCC at 30% (i.e., 1,500 total value divided by 5,000) of nameplate as capacity value, even if the last MW is worth only 0.2 MW of marginal ELCC value.
- **Hybrid Marginal Reliability Value based on Modeling and Empiricism (Recommended Approach):** A final option would be to pursue a marginal value approach that relies on both historical measurements of resources' performance and advanced reliability modeling, with the details tailored to New England's specific needs and the characteristics of each class of resources. For some resource classes (such as wind, solar without integrated batteries, and perhaps thermal resources without firm fuel), simulated performance under marginal ELCC could be used as the initial estimate of reliability contributions, with resource-specific performance factors applied based on historical measurements to express how each resource differs from the average of its technology type. For other resource classes (such as batteries, demand response, solar-battery hybrids, and thermal resources with unlimited access to firm fuel), historical operating data may be of primary importance for estimating future reliability contributions, but with model-based adjustment factors to account for the effects of energy limits, fuel limits, or other availability drivers that are not fully captured in historical data.

In evaluating these options, the key considerations should be to: (1) ensure that reliability will be maintained, avoiding poor reliability outcomes that could arise if resources are systematically credited as providing higher reliability contributions than they can be expected to deliver; (2) provide a reliability-neutral exchange rate among resources, both to enable reliable substitutions and to signal efficient investment—which requires accurately expressing each resource's marginal reliability value; (3) help incent resource owners to enhance, maintain, and operate their facilities to be able to perform when needed most, by reflecting such in their

---

[3] "Perfectly available capacity" refers to a hypothetical resource that is fully dispatchable and not susceptible to outages.

accreditation; and (4) recognize modeling and data limitations that may differ among types of resources. There are however several additional related or derivative considerations. The list of criteria includes reliability, economic efficiency, technology-neutrality, practicality, transparency, and consumer cost, as outlined in Report 1.

Table 1 and Table 2 in the body of this report summarize our assessment of each of these options relative to the defined evaluation criteria. None of the options is perfect, suffering from some combination of conceptual challenges and implementation accuracy challenges.

Modeling-based approaches (simulated marginal and average ELCC) offer the advantage of incorporating information about reliability challenges across a wide range of weather patterns and resource correlations, including conditions that have not been experienced in the past few years. Modeling-based approaches suffer from the disadvantage that accuracy is limited by modeling simplifications and modeler judgement. Simulated ELCC (or MRI) modeling should therefore be calibrated based on rigorous accuracy testing relative to real-world events, empirical analysis of incidence of shortages that differ from modeled outcomes, and resource classes' performance relative to accreditations. Simulation-based estimates also have limited ability to capture the nuances of technically complex and diverse resources whose reliability value is primarily determined by how they are operated (*e.g.*, hybrid and battery resources), and so will be of limited use for assessing these resource types.

Empirical approaches include the "historical availability measurements in pre-defined hours" and "historical tight-intervals measurements" approaches described above. Such approaches offer the advantages of simplicity and transparency while incenting resources to be available. Yet they suffer from the drawback that there will be limited data from which to accurately measure resources' availability under conditions that have not been observed in the recent past. Moreover, they may not appropriately weight available observations according to risk, which introduces accreditation biases toward certain resource types.[4] We therefore recommend augmenting historical empirical approaches with analytical or simulation-based adjustment factors to adjust for any accreditation biases.

In our view, the most conceptually sound options are the three that aim to measure the incremental or marginal reliability contribution of all resources (marginal ELCC, historical tight intervals, and hybrid modeling-empirical). Of these three options, we recommend the hybrid approach as a means to capture the advantages of both model-based and empirically-based

4 In general, a historical data-based approach will tend to over-credit resources whose non-availability drivers correlate with challenging system conditions (particularly of those conditions have not been observed in the historical data set). Resource types whose unavailability is more correlated with shortage periods will be more overrated, distorting accreditation away from reliability-neutral exchange rates across resource types.

estimates, and to address the respective limitations of each approach. A combination of modeling and empiricism can be used to measure each resource's ability to perform when reliability events actually occur, consistent with the marginal value concept.

The nature of the "best" hybrid may vary by resource type. Historical measurement during critical periods is particularly important for conventional resources, batteries, hybrid resources, and demand response; all resources whose availability depends strongly on their management, which is difficult to model but may be observed through actual performance. But modeling is important too, for example, to relate days of firm fuel storage to reliability during extreme winter weather that has not been observed within the historical measurement timeframe. Wind and solar resources' value, however, depends more on exogenous weather factors than on how the resource is operated. Thus, intermittent resources without integrated storage may be most accurately assessed via modeling, along with resource-specific performance adjustments. For each of these, determining the best hybrid approaches to accreditation will require extensive development and calibration.

RECOMMENDATIONS

Our initial examination of New England's needs and industry-wide practices leads to several inter-related recommendations for consideration during NEPOOL and ISO-NE's resource adequacy and accreditation reforms. These will be necessary to facilitate industry transformation while maintaining reliability at minimum costs to consumers:

- **Improve reliability modeling** to capture the correlated risks of the current and future fleet and to consider a wider array of weather and load scenarios, particularly regarding winter fuel supply and outage risks.
- **Develop seasonal accounting of reliability needs and supply accreditation** to best meet both summer and winter reliability needs; consider transitioning to a seasonal capacity market to separately enforce and pay for summer/winter needs (or other seasons, if needs arise).
- **Implement more accurate accreditation for thermal resources immediately,** rather than delaying application. Subject to confirmation based on granular system analysis, we suspect that thermal resources lacking firm fuel backup are the resources whose capacity ratings are most substantially overstated by current ICAP-based accounting methods in New England and therefore pose the most immediate reliability concern.[5] Delaying application of a marginal value concept to these resources (while expediting the application to other resource types)

[5] Intermittent resources, on the other hand, are already subject to a substantial level of derating and comprise a smaller portion of the total supply mix. Further derating the capacity accreditation of intermittent resources while continuing to overstate the capacity accreditation of thermal resources could have the unintended adverse effect of amplifying present inconsistencies in the capacity exchange rate among resources.

risks exacerbating present reliability concerns by amplifying economic incentives for resources with the most over-stated capacity ratings. The primary goal of a marginal value approach, a reliability-neutral exchange rate across all resources, cannot be achieved unless that concept is applied to all resources, including thermal resources. There is less academic and industry literature on how to properly credit thermal resources as compared to intermittent and battery resources, so developing a robust methodology for thermal supply will require a particular emphasis.

- **Develop hybrid capacity accreditation approaches.** The best approaches will aim to incorporate the most accurate representation of reliability contributions, considering both: (1) simulated future conditions across a wide range of weather possibilities, ensuring representation of supply correlations and limitations, including firm fuel limitations, energy duration limits, and hybrid resource injection limits; and (2) resource-specific measurement of performance/availability and audits of resources' technical capabilities. While different resource types may require different hybrid methods, all resource types would be subject to the general marginal value concept, which should be implemented for all resource types on a similar timeframe to avoid distorting tradeoffs among resources.
- **Use calibration of historical reliability events, performance measurements, and simulated events** to iteratively improve both the empiricism and modeling, understanding that these approaches will need to be continuously improved over time. In particular, carefully review each resource type and the fleet as a whole to identify and explain any systematic discrepancies between simulated and empirically measured risk incidence and marginal reliability value.
- **Guarantee consistency in the resource adequacy framework among capacity accreditation, capacity obligations, and performance assessments** by ensuring (1) that total capacity supply obligations (CSOs) for the fleet as a whole meet the anticipated demand across all potentially concerning system conditions; (2) that resources will be accredited in alignment with their expected deliveries in those concerning conditions; and (3) that no resources (individually or in aggregate) will be required to consistently perform at levels exceeding their CSOs during shortages.[6] Following through on this recommendation in New England will likely require dividing the year into seasons.

---

[6] As an example of a problematic outcome: consider a scenario in which summer and winter seasons are both reliability concerns, winter is identified as having the greatest reliability challenges in simulation modeling, but the region maintains the current annual construct. A marginal ELCC construct if implemented too rigidly in this context would focus all reliability requirements and capacity ratings on the winter season, resulting in total capacity obligations that are too low to meet summer peak demand. Under the current ISO-NE construct, all resources' obligations would be inflated in the summer to meet summer peak needs (*i.e.* Balancing Ratio > 1), effectively placing financial obligations on resources to deliver capacity beyond what they have been paid to

deliver (and for some resources, above what they can technically deliver). This scenario is problematic for reliability, financial, and pragmatic reasons and can be addressed by breaking the capacity market into two seasons, both of which should have reliability requirements above peak load in the relevant season. Our recommendation is to evaluate all resource types and potentially concerning system conditions to identify and address any systematic discrepancies between accredited capacity and the volumes that are counted on to be delivered from any class of resources (or the fleet as a whole) during potentially scarce conditions.

I. Accreditation Options for New England

To date, ISO-NE has utilized an Installed Capacity (ICAP) basis for measuring reliability contributions. This approach consists of a measurement of maximum resource output (for thermal resources)[7] or a calculation of historical capacity factor measurements in pre-defined hours (for intermittent resources).[8] ISO-NE, State entities, and stakeholders in New England and other regions have identified this historical approach as increasingly insufficient for the purpose of accurately measuring resources' reliability and capacity value in light of changing system needs to support winter reliability and the clean energy transition.[9]

Upcoming NEPOOL and ISO-NE efforts on resource accreditation aim to improve the accuracy of resources' capacity ratings in light of the most pressing reliability challenges. The goal should be to apply a technology-agnostic approach to ensure that 1 MW of capacity supply will contribute equal reliability value, regardless of the underlying resource.[10] In our view, improving accuracy will require: (1) transitioning to an unforced capacity (UCAP) MW measurement basis (that accounts for differentiated outage rates among resources); (2) conducting advanced ELCC simulations (to account for the reliability implications under a wide range of system conditions); (3) incorporating a measurement of individual resources' historical performance; and (4) assessing practical accounting realities such as verification of resources' access to firm fuel in winter. Further, we anticipate that all of the available options for improving accreditation will perform best if implemented in the context of a transition toward a seasonal resource adequacy requirement and a seasonal capacity market.

We discuss here several distinct approaches that could be applied to improve the accuracy of both intermittent and thermal resources' capacity ratings, how they could be implemented most effectively in New England, and the relative advantages of each concept. These options are

---

7 ISO-NE, "Qualification: How to Participate in Forward Capacity Auction Lesson 3B: Existing Capacity Qualification and," Forward Capacity Market (FCM 101), October 18–21, 2021, p. 10.

ISO-NE, "Claimed Capability Auditing," Web Broadcast, August 6 & 12, 2013, p. 30.

ISO-NE, Market Rule 1, Section III.13.1.2.2.1.

8 ISO-NE, Market Rule 1, Section III.13.1.2.2.2.

9 Katie S. Dykes, Connecticut Department of Energy and Environmental Protection, "Re: Fuel Security in New England for Winter 2021–2022," Letter to Gordon van Welie, ISO-NE, December 17, 2021.

Gordon van Welie, ISO-NE, Letter to Katie S. Dykes, Connecticut Department of Energy and Environmental Protection, December 23, 2021.

New England States Committee on Electricity, Letter to ISO New England, January 18, 2022.

10 Accrediting resources according to the reliability value they provide supports resource adequacy regardless of fleet composition, awards economic incentives in proportion to reliability contributions, and guides investments toward the resources that can deliver the greatest reliability benefit at the lowest cost.

arranged in order of complexity. *Historical Availability Measurements in Pre-Defined Hours* is essentially a status quo option that expands ISO-NE's current renewable accreditation approach to cover all resource types. *Historical Tight-Intervals Measurements* represents a best-practices empirical approach. *Marginal ELCC* and *Average ELCC* are far more complex simulation-based approaches. Finally, we discuss opportunities for hybrid methods that combine elements of both empirical and modeled approaches. None of these approaches is perfect. Each of the options presents a number of nuanced technical and implementation challenges and a different balance of tradeoffs relative to the evaluation criteria we have identified in Report 1.

A. Historical Availability Measurements in Pre-Defined Hours

CONCEPT: HOW IT WOULD WORK IN NEW ENGLAND

The most common approach historically used in other capacity markets, and how ISO-NE currently accredits intermittent resources, is to accredit resources according to their average availability in a set of pre-defined hours (*e.g.,* summer hours ending 14–19). Since the hours used in the accreditation are pre-defined, the determination of when scarcity risk exists is exogenous to the accreditation process, serving as an input to the accreditation calculation. Those pre-defined hours could correspond to the hours when peak load typically occurs, or they could correspond to net peak load.

Since ISO-NE already employs a historical availability measurements approach with pre-defined hours for intermittent resources, pursuing this option would require ISO-NE to expand the existing approach to dispatchable resources. To improve the accuracy of an availability-based approach, we would recommend applying historical and forward-looking simulation analysis to identify the most likely times (hours of day, of week, and weeks within each season) that reliability events are likely to occur. The analysis of the most valid availability windows should also consider the tradeoff between: (a) having a larger number of hours considered in these windows (which will encourage availability across a wider range of potentially challenging conditions and offer more data from which to develop a statistically valid availability measurement); and (b) having a narrower set of hours considered (which can more specifically target the measurement to the most challenging conditions). We would recommend updating the analysis of proper availability windows on a regular schedule so as to proactively align with emerging reliability threats.

A historical availability measurements approach with pre-defined hours can be used to capture the differences in outage rates experienced between resources as well as the possibility that some resources may perform better/worse during the most likely critical times. Measured availability for each resource would be tabulated from 3–5 years of historical operating experience, while new resources would be subject to a class-average or administratively-reviewed estimate that would be replaced with resource-specific operating data over time.

CONSIDERATIONS FOR SPECIFIC TECHNOLOGIES

The same availability windows would be used for measuring availability for all resources. However, the approach to measuring availability will necessarily differ based on the practicalities of data availability for each resource type. At a minimum, we recommend the following considerations be taken into account for certain resource types:

- **Intermittent Resources:** As zero- or negative-marginal cost resources, intermittent resources generate any time they are available to do so. Availability is synonymous with performance for these resources, so they may be accredited according to their average generation output in the relevant hours. This is consistent with how ISO-NE already accredits intermittent resources. A concern with this approach is that a small number of recent historical years may fail to capture the wider range of weather conditions that can be statistically predicted across 15–30 years, some of which could coincide with prominent reliability risks. To address this concern we would recommend comparing outcomes that would be estimated from a subset of years, relative to a broader analysis across many years (*e.g.*, by comparison to ELCC analysis). This comparison could illustrate either an acceptable level of consistency or else indicate that some adjustment is needed to reflect risks across many weather years.
- **Dispatchable Generation and Demand Response:** Dispatchable resources will only generate when they are available to do so *and* when it is economic to do so. For these resources, an availability-based approach does not require that a resource actually produces energy or ancillary services in the relevant hour, just that it demonstrates availability through its energy market offers. Dispatchable resources would be accredited based on their average availability in the relevant hours. Availability could be measured based on offers in the real-time energy/ancillary market, day-ahead energy/ancillary market, reported outage data, or on effective forced outage rate demand (EFORd).[11] EFORd measures availability only if the resource is actually called for dispatch (given that apparently available resources may face forced outages when attempting startup). Among these options we recommend real-time and day-ahead offers as the starting point for examination, with an adjustment based on the EFORd metric. Resources with high availability at the day-ahead timeframe but poor availability in real-time dispatch would have a reduction applied to their measured day-ahead availability to account for the proportion of reliability events that are not yet predicted at the time of the day-ahead market (such as from day-ahead forecasting error).

[11] EFORd is a metric expressing the probability that a particular resource will be offline due to forced outage at any given time.

- **Infrequently-Dispatched Resources:** We further recommend examining the extent to which infrequently-dispatched resources' availability can be overstated if their availability is measured using energy market offers. High-variable-cost resources such as peaking generation and demand response can be incentivized under an availability mechanism to overstate their MW of availability in the energy and ancillary markets at a high offer price, since the overstated offer would be infrequently subject to the "reality check" of being dispatched. For resources with relatively low dispatch frequency such that there are insufficient data to measure EFORd, it may be appropriate to utilize surprise dispatch tests to supplement economic dispatch data.
- **Batteries and Other Use-Limited Resources:** As with other dispatchable resources, batteries' availability could be measured based on energy/ancillary offer data and EFORd. However, short-duration batteries may need an additional downward adjustment to the extent that shortfall events are expected to have a longer duration than the resource's energy storage capability. Batteries of different durations (*e.g.*, 15-min, 1-hour, 4-hour, 8-hour) may be subject to different duration adjustments based on some combination of historical event durations, simulated future event durations, and ELCC simulations.
- **Resources Subject to Winter Fuel Availability Restrictions:** ISO-NE has stated repeatedly that fuel availability in winter is the most pressing reliability risk in New England. However, it cannot be measured by using historical availability or performance data alone (unless a critical winter shortfall event has actually occurred in the most recent several years). Natural gas resources that are subject to interruptible supply contracts would have their availability-based UCAP ratings reduced based on an analysis of realized and simulated winter fuel supply shortfall events; oil resources would be adjusted based on the committed fuel inventory (including specified refilling commitments); dual fuel resources would similarly be subject to fuel inventory commitments and adjustments. As these fuel availability concerns primarily arise during extreme cold weather, downward UCAP adjustments would likely apply only in the winter season, not in the summer.
- **Hybrid Resources (*e.g.,* Solar + Battery):** Hybrid resource accreditation could follow a constrained sum-of-parts approach, where for each hour in the defined availability window, the Regional Transmission Organization (RTO) would calculate the sum of: (1) the output of the intermittent component (net of output used to charge the dispatchable component); and (2) the availability of the dispatchable component, not to exceed the total inverter capacity of the combined resource. Those constrained sums would then stand in for the hybrid resource's availability in a given hour.

ADVANTAGES AND DISADVANTAGES OF HISTORICAL AVAILABILITY MEASUREMENTS IN PRE-DEFINED HOURS APPROACH

Advantages: An availability-based historical measurements approach can offer the benefits of being relatively simple, transparent, and technology-neutral. If relying only on historical operational data, this historical approach would not require complex simulated dispatch modeling. The result would be an accreditation based on real-world conditions, real events, and individual resources' real operations. The close relationship between resources' UCAP ratings and their physical operational data should encourage better performance and availability during the defined windows. There is also an advantage to using realized historical data (as opposed to simulated data, as under an ELCC simulation approach like those described in Sections I.C and I.D) in that a historical measurement will not be subject to simulated modeler error.

Disadvantages: A pre-defined hours approach will have limited effectiveness if shortage events cannot be accurately predicted to fall within the specified windows. Broadening the windows can capture more reliability events and incentivize resources to be available across a broader set of conditions, but the incentive to remain available and earn higher UCAP ratings will be diluted in the most critical hours.[12] The approach will also be limited to a relatively small number of recent historical years and therefore may fail to capture weather patterns, fuel supply risks, and other availability drivers outside of those recently experienced in the region. This limitation could become pronounced in light of a changing climate with increasingly extreme weather patterns.

Though this approach is the simplest and most similar to ISO-NE's current approach, the use of pre-defined windows is an unnecessarily crude way to measure resources' reliability contributions. The historical tight-intervals measurements approach, described in Section I.B, significantly improves upon the pre-defined hours approach with essentially no downside. As such, even an optimally designed pre-defined hours approach is likely not an ideal candidate for meeting ISO-NE's resource accreditation challenges.

B. Historical Tight-Intervals Measurements

CONCEPT: HOW IT WOULD WORK IN NEW ENGLAND

An improved variation of a historical data-based accreditation approach could be devised through two key changes: (1) using performance-based (rather than availability-based) measurements whenever possible; and (2) using "tight-intervals" rather than pre-defined hours.

[12] Performance in the most critical hours is the best indicator of reliability because those are the hours when shortages are most likely to arise. However under this convention, these most critical hours are given equal weight to all other hours in the window, including hours with little to no shortage risk. The importance of the most critical hours is therefore diluted and under-weighted.

These two changes could be implemented independently, but in this section, we assume that both changes are implemented in tandem. This approach may be thought of as a best-practices approach for utilizing historical resource data.

The shift from availability-based to performance-based measurements would aim to measure the actual energy or ancillary service production of the resource, which is the most accurate measurement of reliability contributions (while availability measurements tend to be overstated). When it is possible to use them, performance-based measurements are superior to availability-based metrics because certain resources may submit offers into the real-time energy market even when they are not actually available. Unlike energy market offers, coming online and supplying energy or ancillary services provides proof of contributions during reliability events. Shifting to performance-based measurements rather than availability-based measurements would apply only during true scarcity intervals (*i.e.,* 5-minute intervals when load shed events, operating reserve shortages, or other defined emergency events materialize). This is because scarcity event intervals are the only times when all available resources will be dispatched, regardless of economics. A limitation of the performance-based approach is the lack of a sufficient number of performance intervals within which to develop a statistically robust measurement, which will require supplementing performance measurements with availability measurements.[13]

A tight-intervals approach involves selecting specific historical periods when realized real-time system conditions suggest elevated scarcity risk. In this approach, the RTO would identify 5-minute intervals across recent historical years when the real-time market supply cushion (*i.e.*, the difference between net load and available dispatchable capacity) was lowest. The criteria for selecting those hours could take a number of forms, for example: the 100 tightest hours (1,200 tightest 5-minute intervals) from each of the last 3 years, or the 200 tightest hours (2,400 tightest intervals) from the last 5 years. As we explained in Section IV.C of Report 1, the more logical approach is to select the tightest intervals across multiple years rather than selecting a set number of intervals from each year. There is a tradeoff that informs the decision of how many hours to include. All else equal, the system is inherently exposed to greater shortage risk in the

[13] In the future, we anticipate that New England, like all markets with high penetrations of intermittent and demand resources, will need to add more types of ancillary services and increase the total volume of operating and ramping reserves. In a scenario with much larger ancillary services procured according to a robust operating reserve demand curve, it would be possible to transition to greater and potentially exclusive reliance on performance-based measurements. For more discussion of the topic, see Spees, Newell, "Efficiently Managing Net Load Variability in High-Renewable Systems: Designing Ramping Products to Attract and Leverage Flexible Resources," filed before the Federal Energy Regulatory Commission, February 4, 2022, in Docket No. AD21-10-000.

1st tightest interval than the 200th tightest interval.[14] The RTO could decide to include only a very small number of intervals (say the 300 tightest intervals) in order to focus on only the highest risk intervals, but that results in a very small sample size that may become a less statistically relevant measure of some resources' performance. Broadening the number of intervals creates a larger statistical sample but dilutes the relationship to tight system conditions. A potentially more precise approach could involve selecting a larger number of intervals but placing the greatest weight on the intervals demonstrating the tightest supply cushion. The idea behind this approach is that hours are not equally risky and therefore should not be weighted equally. The number of tight supply intervals and weighting method to use in this approach would be decided after a comprehensive examination of several variations and the consistency in identified tight intervals outcomes as compared to simulations and historical reliability event data.

CONSIDERATIONS FOR SPECIFIC TECHNOLOGIES

As with the historical availability measurements in pre-defined hours approach, some resources would require technology-specific considerations to offer the most precise measurement of resource availability. For fuel-limited and intermittent resources, adjustments may need to be applied to account for reliability concerns that have not materialized in recent years. Otherwise, the technology-specific considerations for this approach are largely consistent with those described in Section I.A.

ADVANTAGES AND DISADVANTAGES OF HISTORICAL TIGHT-INTERVALS MEASUREMENTS APPROACH

Advantages: The tight-intervals approach maintains several of the benefits of the historical availability measurements in pre-defined hours approach in that it is simple, transparent, technology neutral, relies on historical data, and minimizes the impact of administrative modeling errors. The tight-intervals methodology has the additional advantage of not requiring the system operator to accurately predict when shortage conditions *might* occur, but instead narrows in on whenever tight conditions *actually* occur. Reliability events will automatically coincide with the hours used to determine future capacity accreditation. Additionally, by utilizing performance-based metrics wherever possible (*i.e.,* during actual shortage events), this approach minimizes its reliance on energy market offer data, which may not always be representative of resources' true availability.

[14] Broadly, the risk of shortage is inversely related to the amount of available supply cushion. If the supply cushion is very small, a relatively modest change in system conditions could trigger a shortage. If there is an ample supply cushion available, it would require very significant outages and/or an unexpected load increase in order to trigger a shortage. Therefore, the tightest hours pose the highest risks, with relatively lower risk in less tight hours.

Another inherent benefit of a historical tight-intervals measurements approach is the direct emphasis on individual resource performance. Resource accreditations depend on resource-specific performance rather than class averages. This approach will therefore produce strong incentives to maximize performance during reliability events and to pursue performance-improving upgrades or operational changes. In addition to normal energy market signals and performance incentives, resources are incentivized to perform in order to maximize their going-forward capacity accreditation. The approach incentivizes strong performance not just in the 200 or so hours when availability is retrospectively measured, but also in the broader set of hours that sophisticated market participants could have reasonably predicted might become tight.

Disadvantages: Like the availability-based approach, a tight-intervals approach relies on a limited number of recent years to measure resources' ability to address shortage risks that will occur in the future. This measurement may not be accurate if certain reliability challenges such as a winter fuel shortage or a wind drought have not occurred in the relevant historical period. Accounting for these reliability risks would require additional adjustments to some resources' reliability ratings based on either simulation results or resource-specific technology capabilities.

The tight-intervals approach introduces a challenge (as do all accreditation approaches focused on marginal reliability contribution) in that there is a possibility that accredited fleet-wide Capacity Supply Obligations (CSOs) could be lower than system peak load.[15] For example in New England if we observe that tight supply hours are primarily realized in winter season, then resources would be primarily accredited relative to winter capacity value and total fleet-wide CSO awards would be sufficient to meet winter peak demand (but not high enough to meet summer peak demand). To ensure reliability in this scenario, another step will need to be taken to develop a separate reliability requirement in the summer season when gross peak demand will occur. After analyzing and addressing any inadequacies in estimation of the reliability standard and resource accreditations, the corrected framework would ensure that total CSO commitments in aggregate across the fleet will be sufficient to serve load in all potentially concerning system conditions.[16]

---

[15] A more detailed description of this phenomenon, along with a set of illustrative diagrams, is provided in Section III.B of Report 1. To restate in brief here: in a tight-intervals approach, resource accreditation and the Installed Capacity Requirement (ICR) are set according to resource performance/availability demand during periods of greatest shortage risk. In a traditional system (supplied primarily by dispatchable firm-fuel generation), shortage risk is typically highest at the summer gross peak load. However, in New England it is possible that increased winter risks and high reliance on non-firm fuel generation may reveal that the greatest shortage risk is in winter rather than summer. In that case, resources will be accredited according to performance/availability in winter conditions, and the ICR will be set according to system needs in winter.

[16] In addition, under the Pay-for-Performance (PFP) framework, there are obligations and incentives to produce energy and ancillary services that can exceed resources' CSO levels in a scenario where demand exceeds aggregate fleet-wide CSOs; in that case each resource's obligation would be inflated so that the total obligation

C. Simulated Marginal ELCC

CONCEPT: HOW IT WOULD WORK IN NEW ENGLAND[17]

ELCC is a capacity accreditation approach designed to estimate the perfect capacity equivalent of a given resource or resource class. ELCC is estimated using complex simulated dispatch modeling designed to quantify, for a given resource fleet, the total scarcity risk (Loss of Load Expectation or LOLE) that would occur in a particular weather year. For a more detailed explanation of the mechanics of ELCC modeling, see Section III.B of Report 1.

Fundamentally, ELCC models are designed to estimate the level of shortage risk that exists in each hour of a given weather year and then to quantify how much the addition of a given resource or resource class can offset that shortage risk. In theory, this is more nuanced than an availability-based approach, producing a comprehensive assessment of the performance attributes of each resource type. Additionally, as a modeled approach, simulated ELCC enables the RTO to analyze *many* weather years, rather than just a few recent years.

However, as with any model, an ELCC study is only as good as its inputs. The inputs required for an ELCC approach include:

- **Projection of target year fleet**. An *ex-ante* estimate of which resources will clear the capacity auction for the target year. This is potentially quite difficult to estimate because the *results* of the ELCC study will directly influence the accreditation (and therefore the competitiveness) of certain resource classes.
- **Historical weather data and accompanying load and intermittent resource output profiles.** Demand profiles and intermittent resource output profiles should be aligned with particular historical weather years, in order to ensure that the strong correlation (and anti-correlation) between supply and demand can be captured. The modeling should include as many weather years as possible, while recognizing that a changing climate may limit the predictive power of historical weather data, especially for older weather years. If the sample size is large (30+

matches the system need (Balancing Ratio > 1). But the PFP system is insufficient to induce performance that systematically and regularly exceeds awarded CSO levels because: (1) implied obligation would be larger than the MW quantity for which resources have been paid and in some cases greater than what they are technically capable to deliver; (2) the size of the incentives at play are too small given the infrequency of performance events; and (3) the PFP stop-loss guarantees would limit the scale of both penalties and over-performance payments.

[17] Marginal Reliability Impact (MRI) may be considered a sub-set of marginal ELCC and can be expected to yield similar results. It relies on the same conceptual framework as marginal ELCC but employs a methodology that requires fewer model runs. Since we consider these two options to be functionally synonymous, we only use the term “marginal ELCC” in this report.

years), it may be possible to weight all years equally. For smaller sample sizes, it may be necessary to assign unique probabilities to each year based on broader weather trends.

- **Adjusting demand profiles.** Historical demand profiles shapes may need to be scaled to account for load growth or adjusted to account for growth of Distributed Energy Resources (DERs) and electrification. For example, electrification of heating demand will increase both the absolute size of winter demand as well as the correlation of high demand with relatively poor winter resource performance.
- **Correlated thermal outages in summer and winter periods.** Thermal resources tend to face higher outage rates in both extreme hot and extreme cold weather when demand is high; this correlation can be either statically measured (as in statistically correlated forced outage rates) or captured explicitly via modeling techniques (as with quantifiable fuel supply limits).
- **Intermittent Resource Output Profiles.** Depending on resource type and data availability, the modeling may involve some combination of historical resource production profiles and simulated profiles derived from wind speed and solar insolation data (again, tied to a self-consistent weather year).[18]

Once a robust modeling analysis has been conducted and extensively vetted for accuracy against historical data, it can be used to estimate simulated Marginal ELCC values. Marginal ELCC accredits each resource as if it were the last resource in the fleet to come online (unlike Average ELCC, as discussed in Section I.D). Theoretically this could be calculated on a unit-specific basis, but computational limitations likely require that resources be grouped by class. This could involve grouping by technology or it could involve more granular classes based on attributes such as geography, fuel storage, or weatherization technology. Certain resources (*e.g.,* hydro resources with non-pumped storage) have highly unique characteristics and may need to be modeled on a resource-specific basis.

For example, the RTO might assign all solar resources to a class. They would then take a small increment, say 50 MW, of a hypothetical generic solar resource and use LOLE simulation modeling to calculate the perfect capacity equivalent of that 50 MW of solar, when removed from the otherwise full target fleet. If removing that 50 MW of solar required 10 MW of perfect capacity be added to keep LOLE unchanged, then all solar resources would be accredited at 20%

[18] For example, if modeling a weather year from 15 years ago, there may not be significant historical wind or solar output data available. Or if there is significant data available, it may be that the fleet from 15 years ago is not representative of the current fleet. In either case, this situation will likely require the RTO to use historical weather data to predict how the current fleet would have performed in that year (*i.e.,* putative output data). Where significant historical output data is available for a given resource type, the RTO may opt to simply scale that historical data to the size of the current fleet (and possibly modify the data to reflect things like technological improvements). Alternatively, the RTO could still employ putative output calculations instead, and it will be up to the RTO to determine the best approach.

(10 MW perfect capacity divided by 50 MW nameplate solar) of nameplate. Optionally the RTO may employ a subsequent resource-specific performance adjustment so that over-performing solar resources may receive accreditations higher than 20% (offset by under-performing solar resources receiving accreditations lower than 20%). For a more detailed explanation of the mechanics of marginal ELCC, see Section III.B of Report 1.

CONSIDERATIONS FOR SPECIFIC TECHNOLOGIES

In contrast to availability-based approaches, simulated ELCC approaches require substantially different treatment for each resource type. In New England, we recommend that an ELCC-based approach should be applied to thermal resources as the most immediate priority (given our expectation that fuel-limited thermal resources are presently the most substantially over-accredited and hence pose the most immediate reliability risk).

Considerations for applying marginal ELCC to various technologies include:

- **Thermal Resources:** As no other North American RTO has yet implemented ELCC for these resources, there is less experience and literature to draw from for developing a valid methodology.[19] Promising research and methods have been studied by Astrapé Consulting that could be adapted in the New England context.[20] Applying marginal ELCC concepts in New England will require careful focus on the underlying drivers of resource non-availability and the level of correlation with other reliability risks (namely, the correlation of high demand and poor resource performance with extreme hot and cold weather). Implementing this analysis in New England will require substantial analysis to refine the most valid methodology. A first step is to assess the level of correlation in fleet-wide, technology-specific, and resource-specific outage rates with extreme hot and cold temperatures, which can be modeled by applying a correlation factor between resource outages and temperature (noting that this will automatically implement correlation with high demand periods, given that demand profiles will also have been derived from the same weather data). An additional requirement in New England will be to explicitly model resource fuel limitations for resources that have not demonstrated access to firm fuel, which could be analyzed by applying some combination of a cap on the total volume of non-firm gas supply fuel that can be delivered to

[19] Ireland does employ ELCC-based accreditation for thermal resources. *See* Eirgrid, SEMO, and SONI, "I-SEM Capacity Market: Methodology for the Calculation of the Capacity Requirement and De-rating Factors," June 2018, pp. 18-21.

[20] As an example of the impact of applying ELCC methods to thermal resources, Astrapé found that applying ELCC concepts to thermal resources could reduce winter capacity ratings from 95% (based on a historical availability approach) to 76% (based on an average ELCC approach); summer ratings could decrease from 95% to 85%. These results should not necessarily be expected to align with similar analyses that could be conducted in the New England region due to differences between marginal and average ELCC, resource mix, weather correlation measurements, and other regional differences. See Dison, Dombrowsky, Carden, *Accrediting Resource Adequacy Value to Thermal Generation*, May 30, 2022.

certain power plants, and/or explicitly modeling the inventory of back-up fuel oil available. Applying marginal ELCC to thermal resources may require the RTO to group thermal resources into granular resource classes according to the nature of their fuel supply arrangements, the number of days of firm fuel, and weatherization attributes or even modeling some thermal resources on a resource-specific basis. The accuracy of the resulting reliability modeling would need to be extensively back-tested relative to realized historical performance and system-wide reliability/scarcity metrics during tight winter periods.

- **Intermittent Resources:** In many ways, renewable resources are a natural fit for ELCC because renewable generation can be almost entirely predicted by prevailing weather conditions. This enables an RTO to compute renewable generation data even for weather years *without* substantial renewable generation data. For recent weather years with substantial renewable generation data, the RTO can scale observed historical output to the size of the current renewable fleet. Additional adjustments may be required to account for technological or geographical differences between the historical and current fleet. Alternatively, the RTO may pursue a bottom-up approach by generating output shapes for each individual resource, leveraging historical output data where applicable, and relying on putative calculations elsewhere. For weather years predating significant renewable penetration, the RTO must rely on putative output (either fleet-wide or unit-specific) as a function of historical weather data.
- **Batteries and Other Use-Limited Resources:** The output of dispatchable, use-limited resources at any given time is highly dependent on the prevailing system conditions and associated economic incentives that will drive batteries' charging and discharging behaviors. Since ELCC simulations involve modifying those system conditions across scenarios, historical storage output data is of limited use in ELCC modeling, even in recent historical years with significant storage operational data. As such, calculating ELCC for storage resources requires the use of simulated operations data in all years, as estimated endogenously within the reliability model (and in a fashion that is self-consistent with expected system dispatch and associated economic incentives). The resulting battery resource simulated operations (and hence estimated marginal ELCC) will likely be highly sensitive to model assumptions such as resource management (state of charge), level of foresight, relative position in supply stack, treatment of energy-limited resources during emergency events, and ability for batteries to provide ancillary services. Accordingly, it would be especially crucial to thoughtfully design these model inputs and validate the model outputs through historical back-casts. The reliability value of a storage resource depends heavily on the duration of that storage resource. Longer duration resources can address wider peaks and are inherently more valuable (on a per-MW basis) than shorter duration resources. If resources are not modeled on a resource-specific basis, it will be imperative that storage resources be grouped by duration.

- **Hybrid Resources (*e.g.,* Solar + Battery):** Hybrid resources (assumed here to be comprised of a renewable component and a storage component) often have unique characteristics (*e.g.,* inverter and interconnection limits, grid charging capabilities) that lead them to perform differently than if they were two separate resources. Calculating ELCC for hybrid resources would likely involve calculating separate ELCC values for the renewable and storage component and then calculating a sum-of-parts ELCC that respects any resource constraints. The hybrid resource itself would not be grouped in a resource class, but each of its components would.

ADVANTAGES AND DISADVANTAGES OF SIMULATED MARGINAL ELCC APPROACH

Advantages: Marginal ELCC is designed to accredit resources based on the incremental or marginal reliability benefits that each individual resource contributes to the overall fleet. Similar to the historical tight-intervals approach, marginal ELCC aims to accredit resources according to their marginal contribution, determine a reliability-neutral exchange rate among resources that may be substituted for each other between the study and the auction, and provide economically efficient investment and retirement signals. New resources are compensated according to the estimated reliability benefits they contribute, focusing investment incentives toward resources that contribute more to system reliability. Existing resources are compensated according to the reliability benefits that would be lost if they retired. An existing resource will remain online only if the value provided to the system exceeds the cost required to stay online.

Disadvantages: The primary disadvantages of the simulated marginal ELCC approach derive from its reliance on simulation modeling that will inevitably incorporate errors and omissions. Administrative modeling approaches cannot hope to foresee with perfect accuracy precise levels of supply, demand, non-availability drivers, and their underlying correlations; these accuracy concerns can be mitigated (but not eliminated) by extensive model validation, analysis, back-testing, and continuous refinement.

Estimated marginal ELCC values will also depend on the composition of the assumed future supply mix (which cannot be known until after the capacity auction has been conducted). Marginal ELCC values can, in some cases, be highly sensitive to these assumptions with the estimated marginal ELCC dropping quickly as the modeler increases the assumed penetration level of a particular technology. (By comparison, average ELCC estimations may be less conceptually valid, but tend to be less sensitive to modeling assumptions.)

Another potential challenge with a marginal ELCC approach is that, as in the tight-intervals approach described above, this approach can lead to an installed capacity requirement below gross peak load. As with the tight-intervals approach, it is therefore necessary to adjust the way that the CSO is defined or the modeling approach (or both) to ensure that aggregate CSOs will exceed demand during all potentially concerning hours.

A final limitation of simulated ELCC approaches (both marginal and average) relates to the limited ability to capture individual resources' unique technical characteristics and operational behaviors. If assessed on a class basis, each resource within a given class will receive the same accreditation (as a % of nameplate capacity). This approach ignores any differences between resources within a class. By not rewarding individual over-performance, this approach provides no additional incentive (beyond energy market price signals and Pay-for-Performance incentives) for resources to pursue technological upgrades or operational changes to maximize performance during shortage events. More accurate accreditation can be achieved through the addition of a resource-specific performance adjustment. This adjustment involves using a demonstrated performance metric to inform the allocation of each class total ELCC to the individual resources in that class. That demonstrated performance metric should ideally be a measure of capacity factor or availability in certain high-risk hours (*e.g.,* tight supply-cushion hours or peak net load hours). Where actual performance/availability data is not available for a particular resource, the RTO may rely on putative output data or simply accredit that resource at its respective class average. A well-designed adjustment creates the incentive for resources to maximize their performance during high-risk hours, and it will sharpen the capacity market's investment and retirement signals. This results in greater economic efficiency and reduced customer costs.

D. Simulated Average ELCC

CONCEPT: HOW IT WOULD WORK IN NEW ENGLAND

Average ELCC involves similar modeling to marginal ELCC and should be subject to the same detailed model validation and back-casting described above, but accredits resources for their share of the total value provided by a defined group of resources (such as all renewable resources) relative to a scenario in which that entire group were absent. The group value is measured as the MW of perfectly-available resources needed to restore the system to the 1-in-10 LOLE criterion. The value of sub-classes and individual resources can be apportioned in various ways (such as in PJM's modified "delta method" described in Report 1, Section IV.B), which recognize interactions across the resource fleet. Such interactions can cause the sum of all class total ELCC values to differ from the ELCC value of the overall resource fleet. This difference, often called a "diversity benefit," may be positive or negative and must be allocated to the various resource classes in order to have the sum of individual ELCC values equal the total portfolio ELCC.

CONSIDERATIONS FOR SPECIFIC TECHNOLOGIES

Considerations for specific technologies in an average ELCC approach are the same as in a marginal ELCC approach (see Section I.C). The key observation for average ELCC is that, for resource classes with high correlations among resources, the marginal value declines more rapidly with penetration and can cause the marginal value to fall far below the average value. As we have discussed above and in Report 1, the lower marginal value is the more conceptually valid approach to estimating a resource-neutral exchange rate among capacity resources, as long as a marginal reliability concept is applied to all resources. For average ELCC (and even more so for marginal ELCC), the approach should be applied consistently across all technologies including thermal resources. If ELCC is applied only to a subset of technologies such as batteries and intermittent renewables, this could risk exacerbating present inaccuracies in the exchange rate that may be over-accrediting fuel-limited thermal resources by the greatest amount.

ADVANTAGES AND DISADVANTAGES OF THE SIMULATED AVERAGE ELCC APPROACH

Advantages: Average ELCC can support resource adequacy as long as the methodology for calculating the resource adequacy requirement is consistent with the way resources within the study fleet are accredited. As with marginal ELCC, the average ELCC approach can capture a large number of weather years, system conditions, and supply-demand correlations that extend beyond what has been observed in recent history. Compared to the marginal ELCC approach, the average ELCC can be less sensitive to assumptions on resource penetration levels (though it can be more sensitive to the specifics of apportionment). Finally, if ELCC is applied to only a subset of technologies, the higher average ELCC value may be less likely than a marginal ELCC approach to exacerbate misalignment in the exchange rate with resources not subject to the same consistent ELCC concept.

Disadvantages: Like marginal ELCC, average ELCC will suffer from inevitable modeling errors, will be less tightly connected to resources individual performance, and has limited ability to capture unique technology/operational characteristics.

Unlike the tight-intervals and marginal ELCC approaches, average ELCC does not provide a reliability-neutral exchange rate among resources. This can allow reliability shortfalls if, after the study, resources with lower marginal value substitute for those with higher marginal value (but with the same ratings under the average ELCC construct).[21] For the same reason, average ELCC provides less efficient investment and retirement signals compared to marginal ELCC, which can

[21] For example, gas-fired resources without firm fuel could have zero marginal value in winter cold snaps when no or limited non-firm gas is available. Crediting at the average ELCC would not recognize that and would accredit them at a value greater than zero, such that a large number of them could displace highly-reliable nuclear generation.

impose greater societal and consumer costs over time. For example, resources with declining marginal ELCCs would be accredited above their individual reliability contributions, leading developers to overbuild these types of resources. At the same time, resources with constant marginal ELCCs (*e.g.,* firm fuel dispatchable generation) or complementary ELCC (*e.g.*, battery storage complementing solar PV) would be at a relative disadvantage, leading developers to under-build these types of resources. Another issue with average ELCC is the somewhat arbitrary determination of how to select a group of resources whose total value will be shared among the sub-classes and individual resources.[22]

Finally, like marginal ELCC, the average ELCC approach will need to determine how to adjust class ELCC values to account for resource-specific performance differences. A resource-specific adjustment will be required in order to incentivize resources to pursue technological upgrades and operational changes to maximize performance during shortage events. This challenge is not insurmountable, as PJM has found and as we discuss below.

E. Recommended: Empirical-Modeling Hybrid

The options presented thus far are not a comprehensive list. There are many ways in which ISO-NE could combine elements of different approaches to create a more effective hybrid approach. Ideally, a hybrid approach would employ a marginal accreditation concept, capture the benefits of both historical and simulation-based approaches (while compensating for the limitations), and apply the concept to all resource types.

The starting point would be to develop the most accurate possible accreditations based on historical tight-interval measurements, a separate set of accreditations based on marginal ELCC, and then use a robust and iterative process to refine and validate each methodology for each technology type. For some technologies the simulation-based estimate may prove the most reliable (with resource-specific individual performance adjustments); while for other technologies the historical tight-intervals approach may be the most important input (with an additional adjustment based on marginal ELCC modeling).

Simulated Marginal ELCC may be the most appropriate starting point for technology classes with relatively uniform and weather-correlated performance, and whose operational decisions have limited impact on their performance. The estimated marginal ELCC value would be calibrated to

---

[22] For example, one can conduct an average ELCC assessment on: (1) a subset of wind within a particular geography; (2) all wind in the system; (3) wind plus solar; (4) wind, solar, and hydro; (5) on all renewables plus batteries; (6) on all renewables, batteries, and thermal; or (7) some other defined group of resources. Each of these approaches can produce different results. PJM has chosen “all ELCC resources,” which has some intuitive appeal but would not be possible to apply to the entire fleet

observed shortage events and combined with a resource-specific performance adjustment. This is a modeling-based approach that would maximally incorporate observed empirical data to improve accuracy and optimize resource performance incentives. This approach is well-suited for resource types whose performance depends primarily on external factors such as weather and is thus conducive to modeling. The resource-specific performance adjustment is relatively straightforward, and other jurisdictions (*e.g.,* PJM, Midcontinent Independent System Operator) have implemented similar adjustments. To align with a marginal accreditation framework, this performance adjustment should be based on individual performance during the tightest intervals across a set of recent years. As with a historical tight-intervals measurements approach, the RTO must decide how many intervals to observe, how many years to include in the sample, and if/how to weight those intervals.

Historical tight-intervals measurements may be the most important for resources whose availability is not correlated to other resources or to underlying weather conditions, that have unique technical capabilities, and whose operational decisions greatly affect their availability during shortage conditions. Battery resources may be one such example. Even for these resources, a modeling-based marginal ELCC adjustment would be required if there are substantial drivers for non-availability that cannot be expected to occur on a frequent enough basis to be observed in the historical dataset.

These hybrid approaches leverage the advantages of both modeling and empiricism. Hybrid approaches may be especially useful if an RTO determines that there is no single approach they can apply to all resource types. Even if unique features of certain technologies must be accounted for (e.g., specific fuel inventory levels or battery durations), the aim would be to achieve a consistent reliability-neutral exchange rate across all resources, even while the pragmatic implementation may differ by technology.

II. Evaluation of Accreditation Options

The objective of this report is to inform the selection and design of resource accreditation options for New England, while recognizing an even more pressing need to split ISO-NE's annual resource adequacy construct into two (or more) separate seasonal components. A seasonal approach is essential to enable ISO-NE to address the fundamentally different winter and summer resource adequacy challenges. Nevertheless, accurate resource accreditation is a foundational element of the capacity market, regardless of whether or not ISO-NE transitions to a seasonal construct. In Section VI of Report 1, we outlined the following criteria for evaluating possible resource accreditation options:

- **Reliability:** Will the reliability modeling accurately represent shortage risks when establishing resource adequacy requirements and associated resource accreditations, so as to ensure that the system maintains reliability? Will ISO-NE, New England states, and consumers be confident that the total of accredited resources is sufficient to meet system resource adequacy needs, without backstop interventions?
- **Economic Efficiency:** Will the design incentivize an efficient investment level, resource mix, and level of demand resources? Will the design incentivize operational decisions that minimize societal cost while supporting system reliability?
- **Technology-neutrality:** Will the design appropriately reward all resource types for their reliability contributions? Will classes of resources be compensated for their full contribution to the reliability objective? Will resources be fairly accredited for performance/capabilities exceeding their class average? Will implementation timing ensure that no resources face inconsistent early application of ELCC, even while others are over-compensated due to implementation delays?
- **Performance:** Will individual resources be incentivized to maximize their performance during reliability events? Will new resources/investments be incentivized to pursue upgrades or operational changes that result in improved going-forward performance?
- **Practicality:** What are the modeling and data requirements of each design? What barriers to implementation may exist?
- **Transparency:** Is the design reasonably easy to understand? To what extent might market participants be able to predict their resource accreditation? Is the approach overly sensitive to administrative judgement, modeler choices, and administrative error?
- **Consumer Cost:** Will consumer costs be contained, considering both risk of out-of-market actions and accuracy of incentives and accounting toward least-cost reliability?

The remainder of this section applies these criteria to the various accreditation options.

RELIABILITY

The primary objective of resource accreditation and the broader capacity market is to ensure the fleet is capable of meeting the reliability standard. ISO-NE's current accreditation approach is not equipped to provide adequate reliability assurances—neither under today's winter reliability and fuel supply risks, nor under a decarbonized future with high levels of intermittent resources. Current methods over-value non-firm gas resources that are subject to fuel interruptions and thermal resources whose capacity ratings are not adjusted to account for outage rates. These methods also imprecisely measure intermittent resources based on pre-defined reliability hours.

ISO-NE could substantially improve upon its current approach by transitioning to a tight-intervals approach, where scarcity risk hours are identified using actual historical data (tight realized supply cushion). Ideally a tight-intervals approach would also include a mechanism to assign higher weight to the most important (*i.e.,* tightest) intervals. Though it is a significant improvement, the tight-intervals approach is not without limitations. By using only recent historical data, this approach is blind to weather patterns not seen in recent years. A simulated marginal ELCC approach would also offer substantially improved resource accreditations by incorporating a more robust assessment of a wide range of potential system conditions across many weather years.

A hybrid historical-simulation hybrid likely offers the most robust reliability support by leveraging a combination of historical and modeled performance data across a wide array of weather years. However, the most reliable approach would be a hybrid that utilizes ELCC-based modeling while prioritizing historical data wherever possible (thus recognizing the limitations of simulation models).

ECONOMIC EFFICIENCY

A resource accreditation approach may be considered economically efficient if it sends the right investment signals on the margin. In other words, new resources should be compensated for the amount of incremental reliability value they provide by coming online, while existing resources should be compensated for the amount of incremental value they provide by not retiring. Simulated marginal ELCC, historical tight-intervals, and simulation-historical hybrid approaches all offer the greatest efficiency benefits based on a resource-neutral exchange rate.

Average ELCC and a pre-defined hours historical approaches would produce a less accurate exchange rate and therefore produce either lower reliability or higher costs for the same reliability (or both).

TECHNOLOGY NEUTRALITY

Under either a tight-intervals or a pre-defined hours historical approach, all resources receive the same treatment, regardless of technology; however the result of the pre-defined hours approach will be less technology-agnostic because it will tend to overstate the reliability value of some resources more than others. In a marginal ELCC approach, all resources are compensated for their individual reliability contributions, regardless of technology. The average ELCC approach also aims to achieve technology neutrality, but will overvalue the contributions of resource types with declining marginal ELCCs (while undervaluing the contributions of complementary resources).

The most important consideration with respect to technology neutrality is likely ensuring that the approach is applied consistently to all resources including thermal resources.

PERFORMANCE INCENTIVES

A tight-intervals approach will incentivize individual resource performance during future tight intervals and shortage events. A pre-defined hours approach will incentivize resource performance during that pre-defined window, but as explained above, that window may not always align with actual periods of heightened scarcity risk. Marginal and average ELCC approaches will not inherently provide these performance incentives. However, if either of these ELCC-based approaches is modified to include a resource-specific performance adjustment, it will incentivize resources to maximize performance during shortage events and pursue beneficial upgrades and operational changes.

PRACTICALITY AND TRANSPARENCY

Due to their simplicity, historical availability/performance approaches are arguably the most transparent and implementable options. The design is easily comprehensible, and market participants will be able to independently calculate their own capacity accreditation. Overall, there are minimal barriers to implementation since ISO-NE already employs a type of historical availability approach for intermittent resources. As a result, this approach can likely be applied uniformly and simultaneously across all resource types in a technology-neutral manner.

By comparison, ELCC approaches are far more complex than historical availability/performance approaches. Calculating ELCC values requires sophisticated modeling software and is often highly sensitive to modeling and administrative judgements. There are significant data requirements to calculate load profiles and putative generation. The approach may be difficult to understand for market participants, who may be unable to accurately predict their resource accreditation. In many ways, it is a black box. Furthermore, it will be critical to apply ELCC to thermal resources, which will require the development of new modeling techniques that can accurately capture the most critical reliability drivers in New England. If applicability of ELCC to thermal resources were delayed, then renewables, storage, and hybrids would be at a relative disadvantage. They would be subject to a stringent ELCC methodology while thermal resources continue to receive a favorable ICAP-based accreditation (*i.e.*, not penalized for correlated outages). As such, it is imperative that implementation occur simultaneously for all resource types. If ISO-NE determines that it cannot implement marginal ELCC effectively and simultaneously for thermal resources, it would be necessary to pursue an alternative but conceptually consistent approach such as a modified tight-intervals method that accounts for the likelihood of extreme winter events.

CUSTOMER COSTS

With any approach, customer costs will be minimized if the approach: (1) ensures the reliability target is met (thereby minimizing risk of load shed and out-of-market interventions), (2) sends economically efficient investment signals, and (3) incentivizes resources to maximize performance during shortage events. A tight-intervals approach will incentivize individual resource performance during future tight intervals and shortage events. We believe that a tight-intervals approach is relatively well-positioned to ensure reliability and send efficient investment signals. A pre-defined hours approach may in theory meet these three objectives, but only to the extent that the pre-defined hours align with realized shortage risk. Even if ISO-NE were to proactively define those hours each year, it would still fall short of a tight-intervals approach. We therefore conclude that a tight-intervals approach is more likely to minimize customer costs.

An average ELCC approach may be designed to ensure the reliability target is met, but it will not send economically efficient investment signals. Should an average approach be modified to include a resource-specific performance adjustment, it will also incentivize resources to maximize performance during shortage events and pursue beneficial upgrades and operational changes. A marginal ELCC approach sends accurate investment signals on the margin, thereby incentivizing economically efficient investment that will produce the lowest-cost reliable fleet. If a marginal ELCC approach also includes a resource-specific adjustment, it will produce efficient performance incentives as well. Coupled with the fact that marginal ELCC produces a reasonable assurance of meeting the reliability target, these efficient investment and performance incentives mean marginal ELCC is likely to succeed in minimizing consumer costs.

OVERALL ASSESSMENT

Table 1 is a scoring matrix that provides our assessment of how each accreditation option performs on each of the above criteria. Table 2 provides a high-level summary of the advantages and disadvantages of each approach.

Overall, the best approach will be one that leverages a marginal accreditation concept (*i.e.,* either marginal ELCC or a tight-intervals approach) while incorporating both empiricism *and* modeling. A marginal ELCC approach can be improved by the addition of a resource-specific performance adjustment. A tight-intervals approach can be improved by utilizing ELCC modeling to calibrate the results for a broader context with more weather years.

TABLE 1: PERFORMANCE RELATIVE TO ASSESSMENT CRITERIA

Approach	Reliability	Economic Efficiency	Technology-Neutrality	Performance Incentives	Practicality	Transparency	Consumer Cost
Historical Availability in Pre-Defined Hours	◔	○	●	◑	◕	●	○
Historical Tight-Intervals Measurements	◑	◕	●	●	◕	●	◕
Simulated Average ELCC	◑	◑	◑	◔	◕	◑	◑
Simulated Marginal ELCC	◕	◕	◕	◔	◕	◑	◕
Hybrid: Marginal Value Based on Modeling and Empiricism	●	●	◕	●	◕	◑	●

Scoring Key: ○ ◔ ◑ ◕ ● Worst ⟷ Best

TABLE 2: RELATIVE ADVANTAGES OF ALTERNATIVE ACCREDITATION OPTIONS

Approach	Advantages	Disadvantages
Historical Availability in Pre-Defined Hours	• Simple, uniform, transparent • Recognizes individual resources' historical availability	• Availability (inferred from energy market offer data) may not accurately measure performance when called to perform • Observed reliability hours may not align with actual reliability events • May over-value poor performing resources
Historical Tight-Intervals Measurements	• Simple, uniform, transparent • Recognizes individual resources' historical performance • Focus incentives in the most relevant hours, when scarcity risk is highest	• Will not capture reliability risks that have not materialized in recent years • Simple averages equally weight all observed hours, irrespective of actual risk in each hour.
Simulated Marginal ELCC	• Capture many weather years and scenarios • Individual performance adjusts resource class marginal ELCC • Most accurate investment and retirement incentives for supporting resource adequacy	• Subject to modeler judgement and error • Lack of transparency in modeling details • Limited ability to capture unusual configurations or operations (*e.g.,* batteries, hybrids)
Simulated Average ELCC	• Capture many weather years and scenarios • Individual performance adjusts resource class marginal ELCC	• Does not provide reliability-neutral "exchange rate" among resources; sends incorrect investment signals • Subject to modeler judgement and error • Lack of transparency in modeling details • Limited ability to capture unusual configurations or operations (*e.g.,* batteries, hybrids)
Hybrid: Marginal Value Based on Modeling and Empiricism	• All the advantages of Historical Tight-Intervals measurements • All the advantages of simulated marginal ELCC but more accurate, so provides better reliability, economic efficiency and costs	• Much work to develop the most effective use of modeling and empiricism, perhaps varying by resource type

III. Recommendations

Our initial examination of New England's needs and industry-wide practices leads to several inter-related recommendations for consideration during NEPOOL and ISO-NE's resource adequacy and accreditation reforms. These will be necessary to facilitate industry transformation while maintaining reliability at minimum costs to consumers:

- **Improve reliability modeling** to capture the correlated risks of the current and future fleet and to consider a wider array of weather and load scenarios, particularly regarding winter fuel supply and outage risks.
- **Develop seasonal accounting of reliability needs and supply accreditation** to best meet both summer and winter reliability needs; consider transitioning to a seasonal capacity market to separately enforce and pay for summer/winter needs (or other seasons, if needs arise).
- **Implement more accurate accreditation for thermal resources immediately,** rather than delaying application. Subject to confirmation based on granular system analysis, we suspect that thermal resources lacking firm fuel backup are the resources whose capacity ratings are most substantially overstated by current ICAP-based accounting methods in New England and therefore pose the most immediate reliability concern.[23] Delaying application of a marginal value concept to these resources (while expediting the application to other resource types) risks exacerbating present reliability concerns by amplifying economic incentives for resources with the most over-stated capacity ratings. The primary goal of a marginal value approach (*i.e.,* a reliability-neutral exchange rate across all resources) cannot be achieved unless that concept is applied to all resources, including thermal resources. There is less academic and industry literature on how to properly credit thermal resources as compared to intermittent and battery resources, so developing a robust methodology for thermal supply will require a particular emphasis.
- **Develop hybrid capacity accreditation approaches.** The best approaches will aim to incorporate the most accurate representation of reliability contributions, considering both: (1) simulated future conditions across a wide range of weather possibilities, ensuring representation of supply correlations and limitations, including firm fuel limitations, energy

[23] Intermittent resources, on the other hand, are already subject to a substantial level of derating and comprise a smaller portion of the total supply mix. Further derating the capacity accreditation of intermittent resources, while continuing to overstate the capacity accreditation of thermal resources, could have the unintended adverse effect to amplify present inconsistencies in the capacity exchange rate among resources.

duration limits, and hybrid resource injection limits; and (2) resource-specific measurement of performance/availability and audits of resources' technical capabilities. While different resource types may require different hybrid methods, all resource types would be subject to the general marginal value concept, which should be implemented for all resource types on a similar timeframe to avoid distorting tradeoffs among resources.

- **Use calibration of historical reliability events, performance measurements, and simulated events** to iteratively improve both the empiricism and modeling, understanding that these approaches will need to be continuously improved over time. In particular, carefully review each resource type and the fleet as a whole to identify and explain any systematic discrepancies between simulated and empirically measured marginal reliability value.
- **Guarantee consistency in the resource adequacy framework among capacity accreditation, capacity obligations, and performance assessments** by ensuring that CSOs for the fleet as a whole meets the anticipated demand across all potentially concerning hours, that resources will be accredited in alignment with their expected deliveries in those concerning hours, and that no resources (individually or in aggregate) will be required to consistently perform at levels exceeding their CSOs during shortages.[24]

[24] As an example of a problematic outcome: consider a scenario in which summer and winter seasons are both reliability concerns, winter is identified as having the greatest reliability challenges in simulation modeling, but the region maintains the current annual construct. A marginal ELCC construct if implemented too rigidly in this context would focus all reliability requirements and capacity ratings on the winter season, resulting in total capacity obligations that are too low to meet summer peak demand. Under the current ISO-NE construct, all resources' obligations would be inflated in the summer to meet summer peak needs (*i.e.* Balancing Ratio > 1), effectively placing financial obligations on resources to deliver capacity beyond what they have been paid to deliver (and for some resources, above what they can technically deliver). This scenario is problematic for reliability, financial, and pragmatic reasons and can be addressed by breaking the capacity market into two seasons, both of which should have reliability requirements above peak load in the relevant season. Another similar scenario could arise if gross and net peak load are both concerning times from a reliability perspective, but only net peak is identified as a concern due to modeling imprecisions. Our recommendation is to evaluate all resource types and potentially concerning hours to identify and address any systematic discrepancies between accredited capacity and the volumes that are counted on to be delivered from any class of resources (or the fleet as a whole) during potentially scarce conditions.

List of Acronyms

AGO	Attorney General's Office
BR	Balancing Ratio
CSO	Capacity Supply Obligation
DER	Distributed Energy Resource
DR	Demand Response
EFORd	Equivalent Forced Outage Rate Demand
ELCC	Effective Load Carrying Capability
FCM	Forward Capacity Market
ICAP	Installed Capacity
ICR	Installed Capacity Requirement
ISO	Independent System Operator
ISO-NE	ISO New England or Independent System Operator of New England
LOLE	Loss of Load Expectation
MA AGO	Massachusetts Attorney General's Office
MRI	Marginal Reliability Improvement
MW	Megawatt
NEPOOL	New England Power Pool
PFP	Pay-for-Performance
PJM	PJM Interconnection
RTO	Regional Transmission Organization
UCAP	Unforced Capacity

Authors

Samuel A. Newell leads Brattle's Electricity Group of over 50 consultants addressing the most challenging economic questions facing an industry transforming to clean energy. His 23 years of consulting expertise centers on electricity wholesale markets, market design, generation asset valuation, integrated resource planning, and transmission planning. He frequently provides testimony and expert reports to Regional Transmission Organizations (RTOs), the Federal Energy Regulatory Commission (FERC), state regulatory commissions, and the American Arbitration Association. Dr. Newell earned a Ph.D. in Technology Management & Policy from the Massachusetts Institute of Technology, an M.S. in Materials Science & Engineering from Stanford University, and a B.A. in Chemistry & Physics from Harvard College.

Sam.Newell@brattle.com

Kathleen Spees is a Principal at The Brattle Group with expertise in wholesale electricity and environmental policy design and analysis. Dr. Spees has worked in more than a dozen international jurisdictions supporting the design and enhancement of environmental policies and wholesale power markets. Dr. Spees earned her PhD in Engineering and Public Policy within the Carnegie Mellon Electricity Industry Center and her MS in Electrical and Computer Engineering from Carnegie Mellon University. She earned her BS in Physics and Mechanical Engineering from Iowa State University.

Kathleen.Spees@brattle.com

John Higham is a Senior Research Analyst with The Brattle Group, specializing in wholesale power market design. He holds a bachelor's degree in Mathematical Economics from Colorado College.

John.Higham@brattle.com