National Regulatory Research Institute

The Economics of Resource Adequacy Planning: Why Reserve Margins Are Not Just About Keeping the Lights On

Kevin Carden and Nick Wintermantel
Astrape Consulting

Johannes Pfeifenberger
The Brattle Group

April 2011

11-09

Table of Contents

I. Introduction **2**

II. Resource Adequacy Modeling—Overview **4**

III. Challenges of Relying Solely on the 1-in-10 Standard **6**

- **A.** Absence of a standard definition for 1-in-10 6
- **B.** Absence of standard analytical inputs in resource adequacy studies 8
- **C.** Failure to consider the full customer cost of reliability-related events 8
- **D.** Does not document the full economic value of reserve margins 10
- **E.** Does not reflect customer preferences and changes in the value of reliability 11
- **F.** Does not consider the mix of peaking resources that supply energy during high load periods 11

IV. Evaluating and Defining Resource Adequacy Based on Economic Value **12**

V. A Framework for Economic Resource Adequacy Analyses **13**

- **A.** Defining reliability-related costs 13
- **B.** Determining the cost of emergency purchases 14
- **C.** Determining the lowest-average-cost reserve margin 16
- **D.** Determining risk-adjusted reserve margins 17

VI. Conclusions **20**

Appendix A **A-1**

Appendix B **B-1**

Online Access

This paper can be accessed online at http://www.nrri.org/pubs/electricity/NRRI_resource_adequacy_planning_april11-09.pdf.

The Economics of Resource Adequacy Planning: Why Reserve Margins Are Not Just About Keeping the Lights On[1]

Reliability planning for the bulk power system, referred to as *resource adequacy planning*, has historically been based strictly on Loss of Load Expectation (LOLE), or the number of firm load shed events an electric system expects over a period of one or more years. In fact, the utility industry has for decades used an LOLE of "1 day of firm load shed in 10 years" (here simply referred to as the "1-in-10" reliability standard) as the primary if not sole means for setting target reserve margins and capacity requirements in such resource adequacy analyses.

This paper presents a case study to illustrate the advantages of supplementing the assessment of physical resource adequacy as measured by the 1-in-10 standard with an analysis of the economic costs and benefits associated with a given level of planning-reserve margins. As seen during the California energy crisis, the primary economic consequence of reliability-related events is not necessarily in the frequency or duration of firm load shed events, but rather through market exposure in the form of unanticipated high power costs. Thus, in addition to the value of avoided physical curtailments, the economic value of increased reserve margins includes both the reduction in other reliability-related costs, such as the high cost of emergency purchases, and the insurance value of reducing the likelihood of extremely high-cost outcomes. We recommend that the definition of a reliability event not be limited to firm load shed events but include shortage events that have economic impacts due to unanticipated high costs.

This paper shows that an economic simulation of bulk power reliability events and their costs and benefits can provide a greatly improved understanding of resource adequacy risks, help identify more cost-effective solutions to meet given resource adequacy standards, help clarify the link between economically efficient planning-reserve margins and physical reliability standards such as the 1-in-10 standard, and inform stakeholders about the value customers are receiving from paying for reserve capacity. As we show, sole reliance on physical reliability standards easily results in setting target reserve margins that—depending on system size and characteristics—are either too low or too high to be cost-effective and economically efficient.

---

[1] Kevin Carden is the Director and Nick Wintermantel is a Principal of Astrape Consulting (www.astrape.com). Johannes Pfeifenberger is a Principal and Practice Area Leader of The Brattle Group (www.brattle.com). The authors would like to thank Paul Centolella, Reed Edwards, Philip Hanser, Scott Hempling, Delphine Hou, Kamen Madjarov, John Seelke, and Steven Stoft for helpful comments and discussions. The opinions expressed in this article, as well as any errors or omissions, are solely those of the authors and do not necessarily reflect the views of Astrape Consulting, The Brattle Group, or our clients. A shorter version of this paper was published as an article in the March 2011 issue of *Public Utilities Fortnightly*.

I. Introduction

For decades, the utility industry has been using the 1-in-10 standard as the primary if not sole means for setting target reserve margins and capacity requirements in resource adequacy analyses. While the origination of the 1-in-10 metric is somewhat vague, there are multiple references to it in papers starting with articles by Calabrese from the 1940s.[2] In the literature we surveyed, no justification was given for the reasonableness of the standard other than that it is approximately the level that customers were accustomed to. Because customers rarely complain about the level of reliability they receive under the 1-in-10 standard, few question the 1-in-10 metric as an appropriate standard. In regions with capacity markets, such as PJM, some have questioned whether the 1-in-10 metric results in reserve margins that impose too large a cost on customers.[3] However, to our knowledge little empirical work has been undertaken in recent history to quantify the economic value provided by reserve margins based on the 1-in-10 standard or to confirm that sole reliance on such physical reliability standards produces a reserve margin that reasonably balances the tradeoff between the economic value of reliability and the cost of carrying the amount of planning reserves needed to maintain target reserve margins.

While the 1-in-10 standard may have reasonably satisfied physical resource adequacy needs in the last half century, we believe that structural changes in energy and capacity markets, increased penetration of renewable and demand-side resources, and legislative changes raise the question of whether target reserve margins set solely based on the 1-in-10 standard are either too low or too high to be reasonably cost-effective and efficient today. We also believe that customers and policy makers must have a means to understand the full economic value that additional capacity (i.e., higher reserve margins) provides beyond physical reliability.

We supplement the 1-in-10 standard with an analysis that attempts to balance the economic value that customers receive from reliability with the cost of supplying that level of reliability. Our position is that an economically efficient resource adequacy standard should:

1) Provide a level of bulk power reliability that is meaningful to all customer classes
2) Reasonably balance the economic value, including price-risk mitigation, that customers receive from resource adequacy with the cost of supplying that level of reliability
3) Demonstrate to customers what economic and other benefits reserve margins provide beyond the physical reliability benefit
4) Provide adequate investment incentives for suppliers of capacity-only products
5) Result in a reasonably optimal mix of peaking resources that supply energy during the highest-load periods

---

[2] G. Calabrese, "Determination of Reserve Capacity by the Probability Method," *Transactions of the American Institute of Electrical Engineers*, vol. 69, no. 2, pp. 1681-1689, Jan. 1950.

[3] For example, see James F. Wilson, “Reconsidering Resource Adequacy, Part 1: Has the One-Day-in-10-Years Criterion Outlived Its Usefulness?” *Public Utilities Fortnightly*, April 2010; and “Reconsidering Resource Adequacy, Part 2: Capacity Planning for the Smart Grid,” *Public Utilities Fortnightly*, May 2010.

6) Consider the ability of a system to absorb energy-limited, non-dispatchable, and demand-side resources

The economic reliability study presented in this paper attempts to address and balance these goals. Looking at resource adequacy from both a physical and economic standpoint not only allows planners to derive more than just a target reserve margin, but also provides a comprehensive framework that allows planners to understand tradeoffs between the costs and benefits of adding planning reserves, analyze the cost of renewable resource integration, and measure more accurately the resource adequacy value of renewable and demand-side resources.

II. Resource Adequacy Modeling—Overview

Resource adequacy modeling differs significantly from typical production cost modeling. Production cost modeling is designed to determine "expected average system costs" over one or multiple years with a handful of sensitivities, which makes it well-suited to performing fuel budget studies, RFP evaluations, and resource planning studies. The consideration of resource adequacy and associated uncertainties typically is only a minor component of those studies. However, when resource adequacy is the key concern, reliability modeling is necessary: It makes it possible to simulate the many (potentially thousands of) scenarios needed to ensure that low-probability but possible extreme system conditions, like the weather conditions recently experienced in Texas and the Southwest, are actually captured and assigned the correct probabilities. This analysis is typically done with a bulk power reliability planning tool that can run thousands of hourly scenarios quickly and is designed to handle load uncertainty (*e.g.*, as driven by weather), the stochastic[4] nature of generation-unit and transmission-interface outages, and emergency operating procedures during reliability events. In the case study discussed below, the Strategic Energy and Risk Valuation Model (SERVM) was used for this type of reliability modeling.[5]

Three of the most significant input variables in any resource adequacy study are weather, load growth forecast error, and generator outages. All three of these components need to be considered both inside and outside of the region being analyzed to determine the extent to which neighboring systems can help mitigate forced outages or high costs during scarcity and other reliability events.

Weather impact on load and generation is often the largest driver of resource adequacy uncertainty when looking ahead a single year. For instance, the Southeast had a significant drought in 2007 and also had temperatures substantially above "normal weather," resulting in loads that were greater than 6 percent above the expected peak load and hydro energy that was below normal. Without sufficient planning-reserve margins, reliability would have been a major concern. Weather diversity across regions is also a big driver in resource adequacy. If loads across regions peak at exactly the same time, little support may be available from neighboring systems. Weather will determine the availability of other energy-limited resources such as wind and solar as well as affect the capacities of thermal resources. As seen recently in Texas and the

---

[4] Stochastic modeling is a probabilistic (i.e., non-deterministic) modeling process. For example, this might involve simulations of a large number of scenarios for which a generator's outage state is determined by drawing randomly from a probability distribution of outage states, including both full and partial outages. Each scenario is assigned a probability such that both the average and statistical distributions of possible outcomes can be determined.

[5] SERVM has been used extensively by large utilities in the southeastern U.S. In contrast to several other reliability modeling tools (such as GE-MARS), SERVM allows for the explicit consideration of economic factors such as the cost of emergency purchases, the cost of integrating intermittent or energy-limited resources, the cost of demand-side resource dispatch, and the economic and reliability value of tie-line capacity to neighboring power systems. For more information regarding SERVM, see http://www.astrape.com/index.php?file=products.

Southwest, weather can also impact plant and natural gas availability, causing multiple generators to be unavailable simultaneously. As discussed in more detail in Appendix B to this paper, weather is currently not being addressed sufficiently in many resource adequacy studies.

The second factor, load growth forecast error, is the measure of the extent to which load forecasters will underestimate or overestimate economic growth for the next several years depending on the year being studied. An accurate representation of economic forecast error must be included in any such analysis.

The third major variable is generator outages. It is important to simulate the percent of time that a system will have a significant amount of generation offline due to forced outages, including partial outages. In many production cost models, only an average forced outage rate[6] is used, whereas reliability models seek to identify the impact of peak coincident outages[7]. Reliability models simulate random generator failure using Monte Carlo algorithms[8].

A detailed study that addresses resource adequacy must take into account the full distribution of these three input variables to understand the probability of having reliability events. When considering the impact of these variables on available generation in neighboring systems that could be delivered across an interface, transmission must also be analyzed in detail. Even if a neighboring system has available resources, they may not be able to provide reliability support if adequate transmission is not available.

The load and weather inputs along with system variables are then simulated against a range of reserve margins. The lowest acceptable reserve margin yields reliability results that meet the specific metric that was selected to define the target reserve margin. Often, this metric is an LOLE of 1-day-in-10-years.

---

[6] Average forced outage rates are typically used in production cost models that only allow for full outages to be represented. A distribution of multiple outage states is used in reliability planning models to more accurately represent the distribution of outages across a system.

[7] Peak coincident outages are reflected in reliability models as distributions of the amount of system capacity that could be unavailable due to forced outages during peak periods. When peak coincident outages are above normal, reliability events are more likely to occur.

[8] Monte Carlo algorithms are a class of computational algorithms that rely on repeated random sampling to create a large number of scenarios that are then simulated individually.

III. Challenges of Relying Solely on the 1-in-10 Standard

There are several challenges associated with relying on the 1-in-10 standard for resource adequacy planning purposes. These challenges mean that relying solely on the 1-in-10 standard will not generally result in the identification of economically efficient resource adequacy standards. The following six points discuss these challenges in more detail.

A. Absence of a standard definition for 1-in-10

As recognized in the recent effort by NERC and Reliability First Corporation, the 1-in-10 reliability standard has different interpretations.[9] Most resource adequacy planners define it as *one event* in 10 years and measure this by calculating Loss of Load Expectation in "events per year." For this definition, the 1-in-10 equates to 0.1 LOLE in events per year.[10] However, others define it as *one day* (i.e., 24 hours) of load loss during a 10-year period and measure this by calculating Loss of Load Expectation in "hours per year." For this definition, the 1-in-10 metric equates to an LOLE of 2.4 hours per year, which would generally involve multiple outage events.[11]

Figure 1 shows results from a case study we performed in which we compared the two definitions for a 40,000 MW system with significant interconnections with neighboring systems. The left axis of the figure shows LOLE in *events* per year, while the right side represents LOLE in *hours* per year. The figure shows that relying on the 0.1-event-per-year interpretation results in a higher reserve margin than using the 2.4-hours-per-year interpretation. The target reserve margins that satisfy these two different definitions range from 14.5% under the first definition to only 10% under the second.

While planners recognize that the 2.4-hours-per-year interpretation provides different reliability than the 1-event in-10-years interpretation, current reliability studies do not provide guidance as to which provides an economically efficient level of reliability. Considering that both metrics are system-wide metrics, the actual physical reliability impact per customer is very small. For example, the 2.4 hours-per-year interpretation for a certain system may correspond to approximately 7 minutes per customer per year. Applying the 1-event-in-10-years interpretation to the same system means that each customer can expect approximately 1 minute per year of firm

---

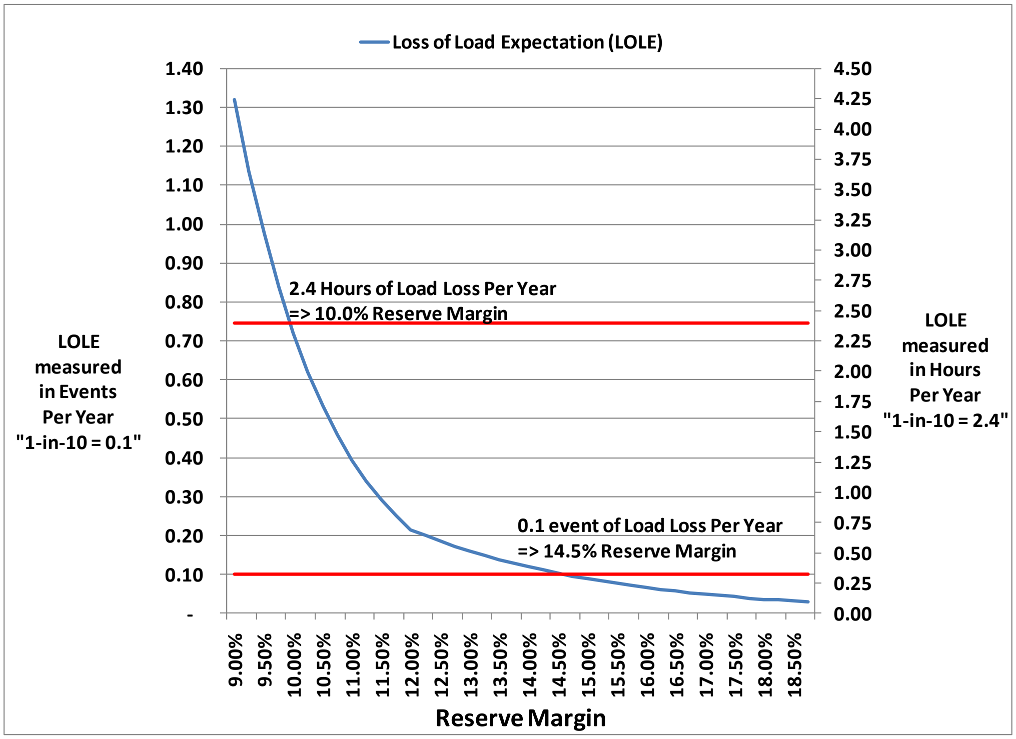
[9] See FERC Notice of Proposed Rulemaking on Planning Resource Adequacy Assessment Reliability Standard, Docket No. RM10-10-000, October 21, 2010 (responding to NERC's filing of the regional reliability standard BAL-502-RFC-02).

[10] As an example for the application of this definition, see Standard BAL-502-RFC-02 of ReliabilityFirst Corporation (RFC) as posted at http://www.nerc.com/files/BAL-502-RFC-02.pdf.

[11] For example, the Southwest Power Pool utilizes the 2.4 loss of load hours definition of the 1-in-10 standard for resource adequacy assessments. See also Milligan and Porter "Determining the Capacity Value of Wind: A Survey of Methods and Implementation," NREL Conference Paper, May 2005, as posted at http://www.nrel.gov/docs/fy05osti/38062.pdf.

load shed.[12] Given that customers experience average distribution outages of approximately 150 minutes per year,[13] the incremental physical reliability benefit of 6 outage minutes gained by moving from 2.4 hours per year to 0.1 event per year is very small. However, while the physical reliability benefits of changing these targets may be small, we demonstrate that the customer-cost and risk-mitigation benefits of the higher physical reliability requirement can be worth billions of dollars.

Figure 1. Two Alternative Interpretations of the 1-in-10 Standard



Moreover, the 1-in-10 standard also does not generally define the magnitude or duration of the firm load shed as measured by the Expected Unserved Energy ("EUE"). EUE for a power system is the energy a system was unable to serve due to capacity shortages likely caused by a combination of events such as generator outages, severe weather, or higher-than-expected load. Based on our experience with modeling different-sized systems, the average magnitude of EUE as a percentage of total load varies from 1% for large systems (greater than 30,000 MW) to

---

[12] These calculations assume 5% of load is shed during firm load shed: 5% x 2.4 hours = 0.12 hours = 7.2 minutes; 1 event in 10 years corresponds to 0.3 hours of lost load per year from Figure 1. 5% x 0.3 = 0.015 hours = 0.9 minutes.

[13] Joseph H. Eto and Kristina Hamachi LaCommare, *Tracking the Reliability of the U.S. Electric Power System: An Assessment of Publicly Available Information Reported to State Public Utility Commissions*, prepared for the Ernest Orlando Lawrence Berkeley National Laboratory, October 2008, p. 15, Table 4.

around 5% for relatively small systems (less than 10,000 MW). Assuming that this relationship holds true across all large and small systems, the 1-in-10 standard does not provide the same level of reliability for customers in different-sized power systems even if both systems interpret the standard in terms of 1 load loss event in 10 years. This is one reason why "normalized EUE" (EUE divided by the "Net Energy for Load") was adopted as a physical reliability metric for the NERC effort under the Generation and Transmission Planning Models Task Force (GTRPMTF).[14]

Other areas where there is uncertainty due to lack of standard definition on the calculation of 1-in-10 include the simulation of only daily peak hours versus simulation of all hours and the definition of the point at which a loss-of-load event is recorded. In some models, a loss-of-load event occurs as soon as the required operating reserves cannot be completely met. In other analyses, loss-of-load events are not recorded until all operating reserves have been depleted and firm load is actually curtailed.

B. Absence of standard analytical inputs in resource adequacy studies

As discussed above, the full range of possible values for the three main drivers of reliability (weather, load growth forecast error, and unit performance) must be taken into account. Many planners rely on different processes to generate study input assumptions with respect to these variables and their uncertainty. Based on our experience with resource adequacy studies from across the country, target reserve margins can vary by 8 percentage points (*e.g.*, from 4 percentage points too low to 4 percentage points too high) based simply on different approaches to the selection of input variables. As shown in Appendix B, choosing between severe or mild historical weather load shapes alone can significantly impact LOLE results. The assumptions regarding load uncertainty due to economic growth will also shift LOLE results considerably. This uncertainty in reserve margins is in addition to the range of different results associated with the alternative interpretations of the 1-in-10 standard discussed above.

C. Failure to consider the full customer cost of reliability-related events

Like any solely *physical* reliability standard, the 1-in-10 standard assumes that a "reliability event" occurs only if firm load is shed. However, reliability-related costs realistically also include costs associated with events such as calling on interruptible loads, dispatching high-cost emergency resources such as older inefficient oil turbines, and making unanticipated costly emergency purchases from neighbors.

---

[14] See the NERC Generation and Transmission Reliability Planning Models Task Force (GTRPMTF) "Final Report on Methodologies and Metrics – September and December 2010 with Approvals and Revisions," posted at http://www.nerc.com/docs/pc/gtrpmtf/GTRPMTF_Meth_&_Metrics_Report_final_w._PC_approvals,_revisions_12.08.10.pdf.

During the 2000-2001 California energy crisis, for example, only approximately 8,000 MWh of firm load was shed during a total of 6 days.[15] Even if these load drops are priced at $10,000/MWh, the economic cost of the curtailments is only $80 million, which is a small fraction of the estimated $50 billion in total costs attributed to the crisis.[16] Similarly, in the case study presented below, a number of individual simulation outcomes had zero unserved energy but the possibility of total reliability-related costs that exceeded one billion dollars a year.

Determining resource adequacy based only on the frequency of unserved energy completely ignores these economic risks. Resource adequacy studies should consider the full range of reliability-related events, including economic impacts of high-cost events without curtailments, such as emergency purchases, in addition to the frequency of unserved energy.

A typical criticism of the 1-in-10 standard is that it provides for more reliability than customers are willing to pay for, the argument being that even if the value of lost load is $20,000/MWh, the "last CT" would need to displace 5 hours of lost load per year to be economically justifiable (assuming the carrying cost of a CT is $100/kW-yr). However, our analysis shows that the majority of customer-side reliability costs may not be incurred in the form of lost load, which means the last CT actually provides substantially more value than just offsetting the cost of the firm load shed event, including the option value to dispatch the unit at cost whenever other dispatched or purchased resources would be more expensive. When the full range of reliability-related impacts and costs is quantified, the 1-in-10 standard can actually result in target reserve margins that are too low in some regions of the country from an economic-efficiency and overall cost-effectiveness perspective.

Figure 2 is an illustration showing how providing the same 1-in-10 reliability level for three systems that differ significantly in terms of size and resource mix yields substantially different cost exposure. It shows that a smaller system with weak interconnections to neighboring systems has much less cost exposure at exactly the same level of physical reliability than a system with significant interconnections. The reason is that for the small system with limited neighbor assistance to achieve 1-in-10 reliability, it must carry much higher reserves (a 23% reserve margin). However, the economic risk associated with this high reserve margin is very low, which means it is not likely that the high reserves would be justified on an economic basis. For a larger system with a substantial amount of energy limited resources[17] and significant tie-line assistance, the 1-in-10 standard yields a cost exposure that is much higher than for the

---

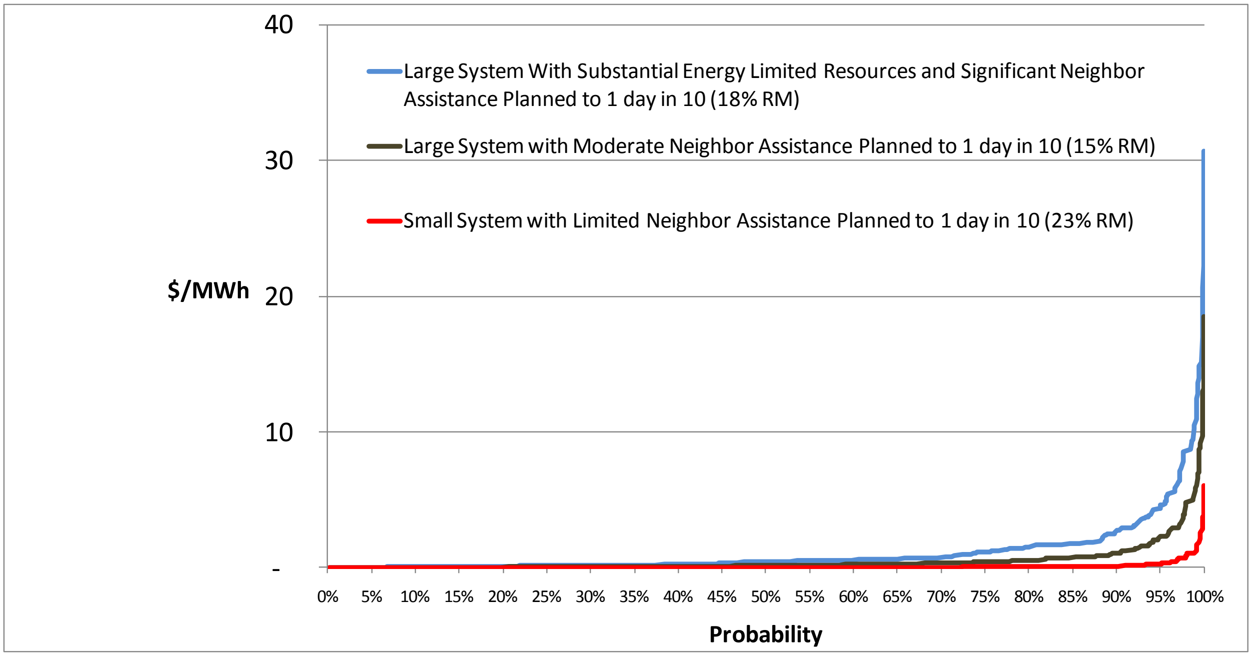
[15] Sweeney, James (2002), *The California Electricity Crisis* (Hoover Institution Press), ISBN 978-0817929121, p. 171.

[16] Weare, Christopher (2003), *The California Electricity Crisis: Causes and Policy Options* (San Francisco: Public Policy Institute of California), ISBN 1-58213-064-7. http://www.ppic.org/content/pubs/report/R_103CWR.pdf., pp. 3-4.

[17] That is, resources that cannot generate at the plant's full capacity whenever needed (such as hydro plants, wind plants, combustion turbines with environmental operating constraints and demand response).

other systems, as it is expected that a system of this type would have more hours in which it had to rely on expensive market purchasers.

Figure 2. Illustration of Economic Risk for Different Systems
(Average annual reliability-related costs)



D. Does not document the full economic value of reserve margins

The 1-in-10 standard defines a target reserve margin and an associated level of physical bulk power reliability that ultimately means relatively little to customers or in terms of the overall cost and cost uncertainty of power supply. Regardless of the interpretation of the 1-in-10 metric, each consumer will expect to experience far less than one firm load shed event in 10 years. The practical benefits to customers of having a resource-adequacy-related firm load shed event once every 20 years versus experiencing one only every 100 years are negligible. However, as discussed earlier, even without any actual curtailments, the cost exposure differences that exist at varying levels of reserve margins can have a significant impact on customers' monthly bills. By using an economic approach to quantify the economic value of different reserve margin targets, planners and regulators will gain a better understanding of the trade-off between higher capital costs, improved physical reliability, and reduced high-cost exposures in power markets.

E. Does not reflect customer preferences and changes in the value of reliability

While most customers would likely agree that the 1- in-10 standard has been "adequate" in the past, they might also be willing to accept a lower level of reliability if the savings from doing so were adequate compared to risks. Or they may be willing to pay for a higher level of reliability if they highly value reliability or if doing so reduces energy cost uncertainty. Given that the value of lost load is constantly changing due to new technology and evolving customer preferences, it is important that we periodically reevaluate customers' perception of the value of reliability.

F. Does not consider the mix of peaking resources that supply energy during high load periods

Most reliability studies based on the 1-in-10 standard treat all capacity as the same from a resource adequacy perspective. However, all capacity is not equal. If combustion turbines are built for all peaking capacity needs, some of the turbines will only be required to be dispatched for 20 hours per year on average. This is not cost-effective when other opportunities exist to meet peak load conditions. For instance, demand-response resources can typically be procured for substantially lower costs if they are dispatched infrequently. Energy storage technologies with only a few hours of storage may also be available to meet some peaking needs at comparatively lower capital costs.

In addition, the economic value of energy-limited, intermittent, or non-dispatchable peaking resources will depend on their share of total system capability and the system's mix of other resources. A system with 8% demand response and energy-limited resource capacity will have a very different cost of reliability than a system with a 2% share of these resources, even though they both meet the 1-in-10 standard. For example, the 8% system may have to incur much higher power-purchase costs during hours adjacent to the system peak to keep their energy-limited resources available for peak load conditions. Similarly, intermittent resources will tend to have more capacity value in a system with energy-limited or hydro storage resources. A methodology that focuses only on firm load shed events would not account for these subtleties. A physical reliability standard such as 1-in-10 would not even recognize expensive reliability-related purchases during near-peak conditions as a reliability event. This reiterates the point that reliability events begin well before firm load is shed.

IV. Evaluating and Defining Resource Adequacy Based on Economic Value

We recommend that physical reliability standards be supplemented and target reserve margins be validated with an analysis of economic value and cost effectiveness. The point at which the costs of providing additional capacity start to exceed all reliability-related economic benefits of the additional capacity, taking into account the full range and uncertainty of possible outcomes, can be used as a reference point to set economically efficient and cost-effective target reserve margins.

This economically efficient and cost-effective target reserve margin may differ from the target reserve margins derived with the 1-in-10 standard—it can either be below or above the reserve-margin targets based solely on physical reliability. Setting target planning reserves to include economic considerations achieves the above-stated goals of economic reliability planning. Consumers will enjoy a level of reliability they are willing to pay for while also taking cost uncertainty into account. They will be protected not only from excessive firm load shedding, but also from the high energy costs frequently associated with reliability-driven extreme market conditions.

Economic resource adequacy planning also informs how renewable and other energy limited resources can be integrated into electric systems economically without adversely affecting reliability. In regulated markets, this approach will also allow planners to develop a cost-effective mix of peaking resources for high load periods, such as the optimal mix of CT and demand-response capacity.

V. A Framework for Economic Resource Adequacy Analyses

To illustrate the application of economic considerations to bulk power reliability analysis, Astrape Consulting performed a case study using an actual (but for the purpose of this paper generalized) power system that includes approximately 40,000 MW of capacity. The system consists of a typical resource mix with about 35% base load resources, 30% intermediate resources, and 35% peaking resources including natural gas turbines, oil turbines, demand response resources, and hydro. The weather-normalized peak load is approximately 36,000 MW. The system, which also includes approximately 10,000 MW of interties with multiple neighboring systems, was modeled using the SERVM stochastic reliability simulation tool designed specifically for economic reliability analysis.

Each generating unit was modeled based on its capacity, heat rate, fuel price, dispatch constraints (such as minimum up and minimum down times), startup times and costs, and historical outage events. SERVM allows multiple states for each generator and uses Monte Carlo simulations to determine the outage state of every resource for each hour within each of the annual simulations. 40 annual weather shapes were constructed by modeling 40 individual annual load shapes based on 40 historical weather years and 40 historical hydro years. A load growth forecast error distribution was created based on the accuracy of historical load forecasting. Appendix B documents in more detail how the weather and load modeling was performed.

SERVM commits and dispatches the system economically to meet load plus operating reserves during all 8,760 hours of a year and then calculates reliability costs and other reliability metrics such as LOLE and LOLH. SERVM is a multi-area model that models directly interconnected neighboring regions to simulate out-of-region purchases over tielines when necessary for reliability.

The resource adequacy analysis involves thousands of full annual simulations to yield an accurate picture of a system's physical reliability and reliability-related costs. For example, for each target reserve margin level simulated, all combinations of weather years (40 years) and load forecast errors (7 points) are simulated, resulting in 280 (40 x 7) scenarios. Then each scenario runs for 400 iterations to achieve convergence on unit outage draws, resulting in 112,000 full-year simulations (each for 8,760 hours) for each reserve margin level analyzed. The results from these simulations are then used to determine the average and distribution of reliability-related costs for different reserve margin levels. Simulating a sufficiently large range of reserve margins thus allows for both the identification of (1) the reserve margins that yield the lowest average costs and (2) an assessment of the cost uncertainty, including the risk (probability) that actual outcomes significantly exceed these average costs.

A. Defining reliability-related costs

Setting target reserve margins based on economic reliability simulations requires balancing the costs of adding new capacity against the benefit of adding that capacity. For our case study, we assumed that new capacity would be a combustion turbine (CT). In other regions, that may not be the appropriate marginal new resource. It is also possible to evaluate a supply curve of new capacity that stretches from lower-cost demand-response resources to higher-cost additions of new physical generation.

As we change the level of installed capacity resources, we need to capture the total benefit of the additional capacity as well as the costs of that capacity. This means the analysis must keep track of all production and purchase costs above the marginal cost of the new capacity resource as well as the fixed costs of the added new capacity. Our analysis breaks these customer reliability costs into the following four categories:

a. Production-related Reliability Costs – defined as any costs of the system's physical generation above the dispatch cost of the new capacity resource. This includes the dispatch of higher-cost generators such as oil-fired turbines and old natural gas turbine units. The addition of a new capacity resource would offset some but not all of these costs.
b. Emergency Purchase Costs – defined as the costs of any purchases at prices higher than the cost of the marginal capacity resource. In our simulations, these emergency purchase costs, including purchases associated with demand-side resources, can range from $1/MWh above the dispatch cost of a CT to the cost of unserved energy (*e.g.*, well in excess of $1,000/MWh) under extreme conditions.[18]
c. Unserved Energy Costs – The value of lost load to customers. This value typically is derived from customer surveys.
d. Capacity Resource Carrying Costs – The annual carrying cost of adding additional capacity in $/kW-yr.

B. Determining the cost of emergency purchases

The production-related reliability costs and unserved energy costs are easily tracked in most reliability models. However, a portion of reliability-related costs are costs associated with unit dispatch and power purchases during reliability and emergency events. As noted earlier, the majority of the costs seen from the California energy crisis were due to expensive market purchases. A scarcity pricing model is thus necessary to model purchase costs during capacity shortages.

SERVM explicitly forecasts market prices during emergency and reliability-related shortage conditions. Depending on the capacity available in surrounding markets and transmission constraints, SERVM uses scarcity pricing algorithms to forecast these prices. Prices can easily exceed $1,000/MWh during scarcity conditions. For this case study, 10 years of actual historical prices from bilateral reliability and emergency purchases in the region were analyzed to estimate scarcity pricing curves that vary with reserve margin and the amount of capacity needed. SERVM was then calibrated to an actual historical year to ensure that the model is accurately projecting the cost of such purchases. In this calibration run, SERVM simulations were based on actual historical load and generator outages for the simulated year so

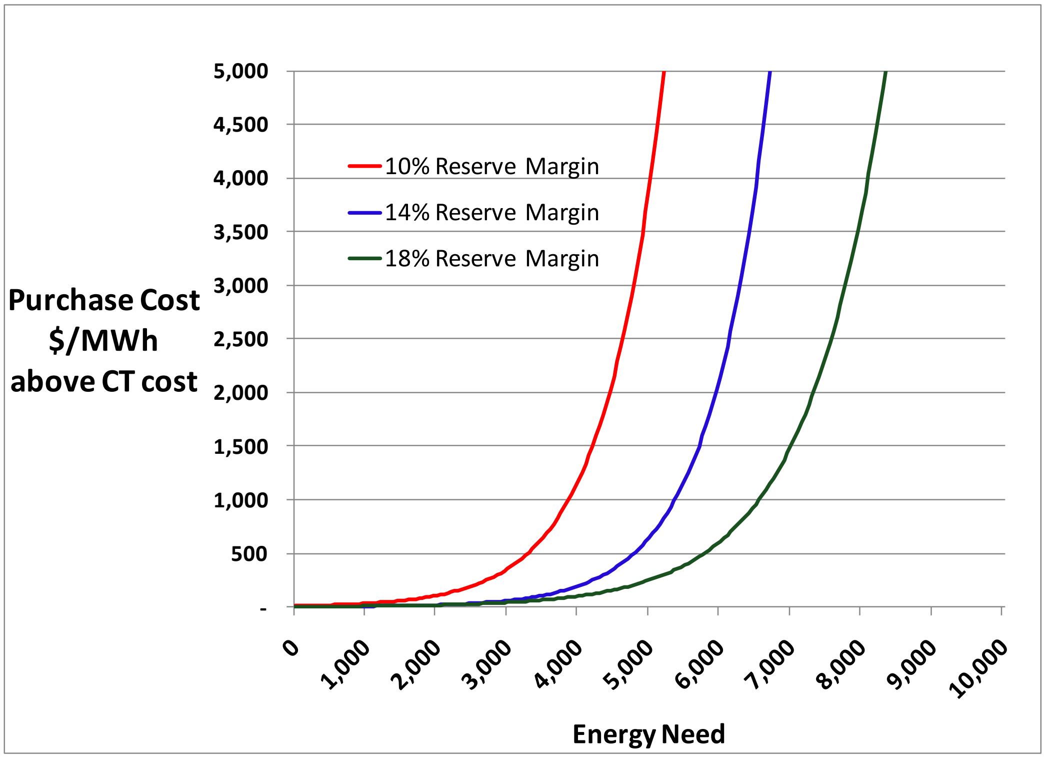
[18] Purchase prices during reliability or emergency events may also include premiums associated with high opportunity costs of energy-limited resources, emergency assistance available from high-dispatch-cost demand-side resources in neighboring systems, or markups related to the exercise of market power by suppliers during scarcity events.

that total annual reliability costs could be compared to actual historical costs to ensure reasonable simulation results.

It would be very difficult to calibrate a model to only unserved energy events because most regions do not have recent experience in which they incurred lost load that resulted from resource inadequacies. In contrast, almost all regions have had market conditions during which capacity was very scarce and high dispatch and emergency purchase costs, including dispatch of high-cost demand-side resources, were incurred that could have been avoided or lowered if additional physical-capacity resources had been available in the region.

Figure 3 shows the scarcity pricing curves developed for this case study. As reserve margins decrease and the system needs for purchases increase, purchase prices increase as well.

Figure 3. Scarcity Pricing Curves for Emergency Purchases
(based on actual historic purchase cost data for the power system studied)

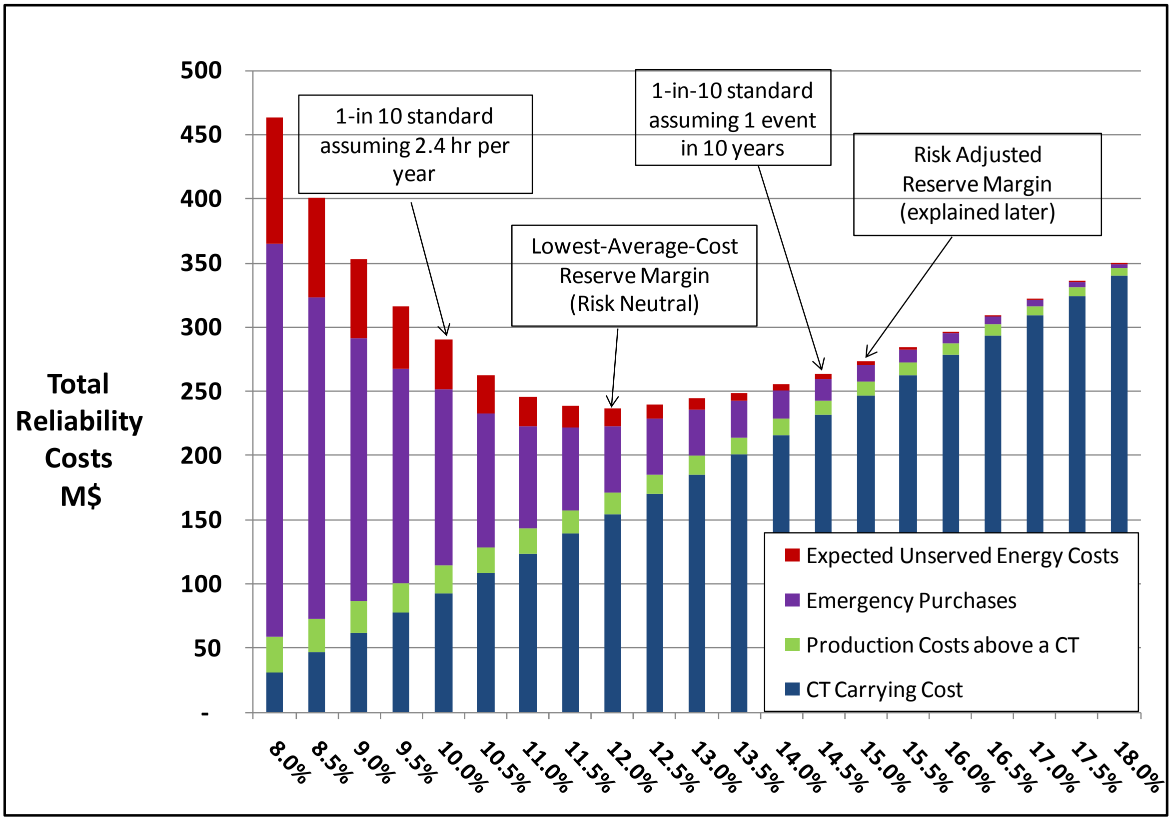


This curve does not indicate the amount of purchases that will actually be available; it shows only the price one expects to pay for energy during peak load conditions and for this study is not allowed to exceed the value of EUE. In some systems, purchase prices may not reach the levels shown here due to a variety of factors, such as regulatory restrictions, contractual cost-based assistance from neighboring systems, or large amounts of demand-side resources that can be called upon at lower dispatch prices. Systems also differ significantly as to the extent to which they rely on reliability or emergency purchases before firm load is shed.

C. Determining the lowest-average-cost reserve margin

Figure 4 summarizes one set of results from this case study. The figure shows the probability-weighted average cost of various reliability-related cost elements as a function of planning-reserve margin. The lowest-average-cost reserve margin can be determined, for example, based on the point at which total reliability-related costs plus the cost of carrying additional reserves is the lowest, ignoring the uncertainty of costs around the weighted average costs shown in the chart. In our case study, this lowest-average-cost reserve margin is 12%. But this result will vary significantly across regions based on their size, load shape, resource mix, and many other factors.

Figure 4. Risk-Neutral Optimal Reserve Margin



Our analysis also shows that, for the system studied here, the primary driver of reliability costs is expensive market purchases or "emergency purchases" as defined in our simulations. In contrast, and contrary to common assumptions, the value of lost load is not the most important factor in determining optimal reserve margins. Even if the value of lost load is changed by $5,000/MWh, the lowest-average-cost or risk-neutral optimal reserve margin shifts by only approximately 0.5 percentage points.

Importantly, as Figure 4 illustrates, because the cost of reliability events (in particular emergency purchases) increases quickly as reserve margins decline, omitting some of these costs in reserve margin evaluations can lead to greatly understated estimates of risk-neutral optimal reserve margins. If one considered only the installed cost of peaking capacity and the value of lost load, the reserve margin that yields the lowest average costs would appear to be only 9%, while it is 12% when all reliability-related costs are considered (and before even attributing any insurance value to risk mitigation, as discussed further below).

D. Determining risk-adjusted reserve margins

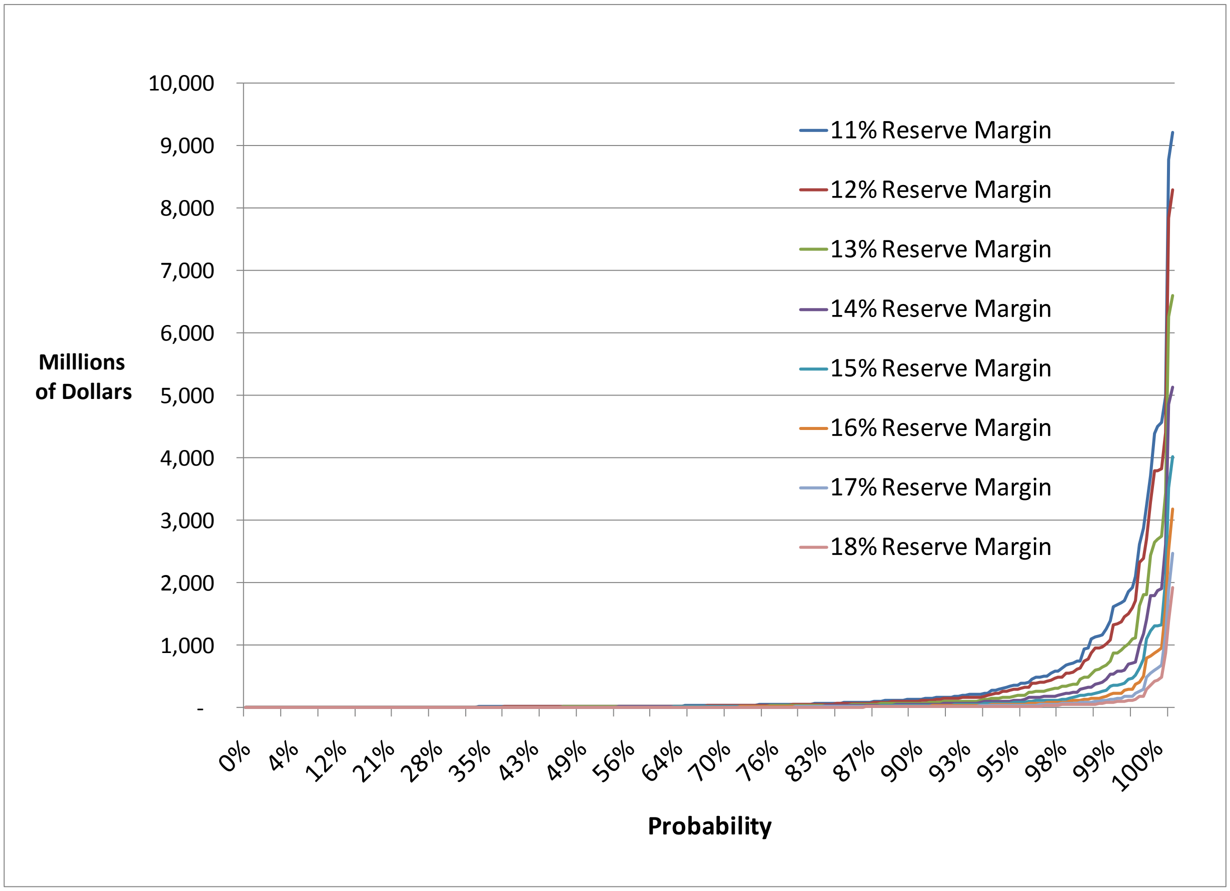
In the presence of risk aversion, the value of higher reserve margins also includes the insurance value of avoiding infrequent high-cost outcomes. While Figure 4 is informative, it over-simplifies the problem by only comparing fixed capacity costs with the long-term averages of very uncertain market exposures. To perform a more informed comparison, the uncertainty of market exposure needs to be considered as well.

The probability distributions of the total annual reliability-related costs (excluding the more certain CT carrying costs) are shown in Figure 5. The figure shows that substantial annual cost uncertainty exists at any given level of reserve margin. Most of this cost uncertainty is associated with the risk of very infrequent high-cost outcomes.

As Figure 5 shows, for 90% of possible annual outcomes, the reliability-related reliability energy cost exposure is quite low for reserve margins in the 11% to 18% range. It is only the last 10% of possible annual outcomes (i.e., conditions likely experienced less than once in ten years) in which a combination of factors occur that cause substantial reliability-related costs. For example, while the expected average of annual reliability-related costs at a 12% reserve margin is only $240 million, Figure 5 shows that there is a very small chance that total annual reliability-related costs could be as high as $8.3 billion. Assuming total retail rates are 10 cents/kWh, this maximum cost exposure would raise consumers' annual costs by 50% for the system analyzed. These numbers are not out of line with estimates that the California Energy Crisis would have doubled retail rates if all costs had been passed through to customers.[5]

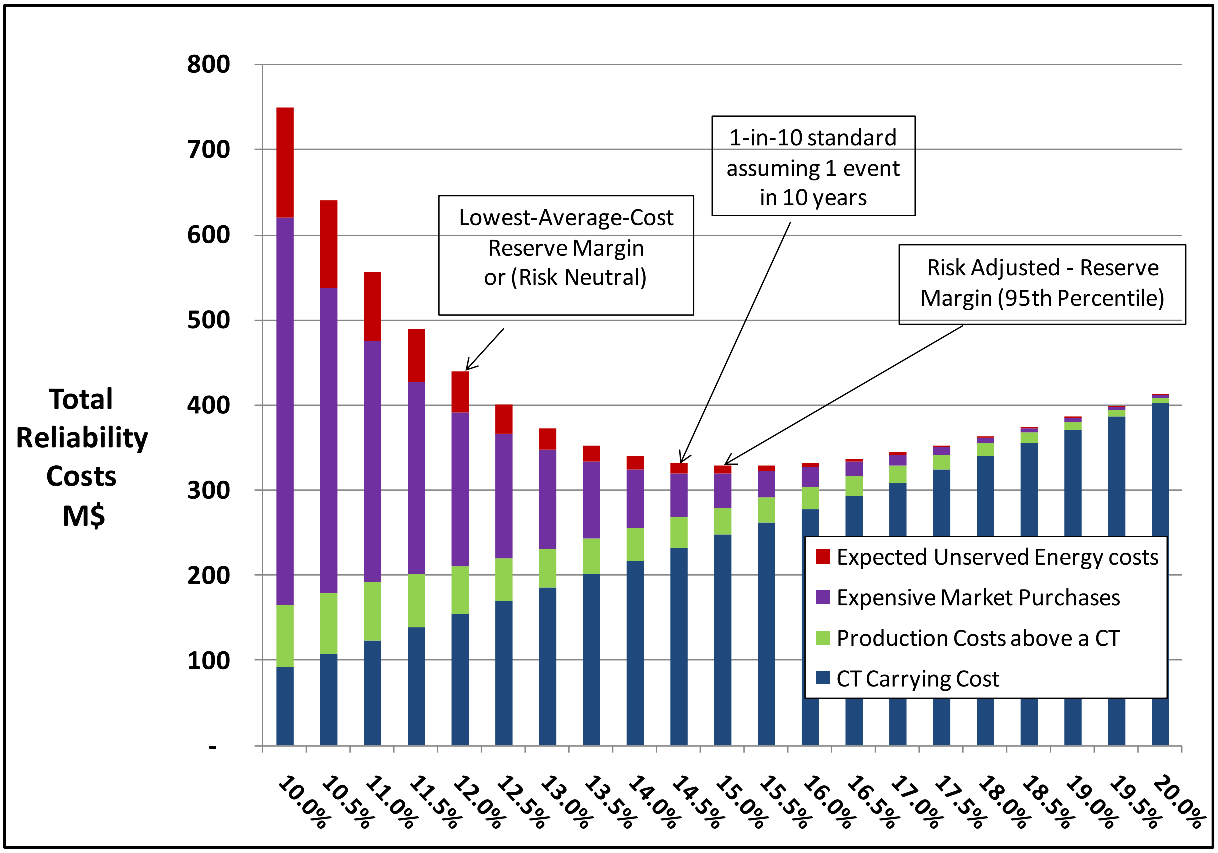
Considering that customers, utilities, regulators, and policy makers all tend to be risk-averse to high-cost outcomes, the "optimal" target reserve margin should consequently not be based solely on the lowest-average cost reserve margin, shown as 12% in Figure 4. While a 12% reserve margin would offer the cheapest option for customers in terms of long-run average costs, the highest-cost outcomes that load-serving entities and customers would be exposed to might be unacceptable.

Figure 5. Distribution of Reliability Cost Exposure



In the insurance industry, premiums are frequently set using a 95% confidence level that the insurance company will be covered in the long term. A similar calculation for determining the appropriate risk adjustment can be used for setting the target reserve margin. Assuming that substituting the 95th percentile cost for the weighted average cost is a proper risk adjustment, the "optimal" target reserve margin increases from 12% to 15% as shown in Figure 6.

Figure 6. Lowest-Cost Reserve Margin at 95th Percentile Confidence Level



As Figure 5 previously showed, installing additional CT capacity to increase the target reserve margin from 12% (the lowest-average cost shown in Figure 4) to 15% (the lowest cost risk-adjusted reserve margin shown in Figure 6) decreases reliability cost exposure. At a 12% reserve margin, there is a 1% probability that costs could exceed \$1.3 billion (with a maximum of \$8.3 billion), while at a 15% reserve margin there is a 1% probability that costs could exceed \$359 million (with a maximum of only \$4.0 billion). The change in reserve margin increases the incremental carrying costs of CTs by approximately \$110 million per year (from approximately \$150 million per year to about \$260 million per year, as shown in Figure 6),[19] which increases average retail rates by less than 1%. However, the maximum possible reliability-cost-related annual retail rate impact is reduced from 50% to only approximately 25%. While any such risk adjustment is subjective, using an economic framework nevertheless allows system planners and state regulators to consider the trade-offs between the costs and benefits of reliability.

[19] The incremental CT carrying costs shown in the two figures are relative to a 7% target reserve margin at which incremental CT carrying costs are zero for the purpose of these illustrations.

VI. Conclusions

Our analysis shows that an economic simulation of bulk power reliability allows for a more informed determination of cost-effective and economically efficient target reserve margins. It also allows planners and policy makers to gain a much better understanding of the resource adequacy question, associated risks, and costs-versus-risks tradeoffs. It further helps to answer the question of what customers are "getting" for the cost of the additional reserves. As the analysis shows, this value of additional reserves includes both (1) the reduction in expected average reliability-related costs such as the high cost of emergency purchases or the cost of curtailments; and (2) the insurance value associated with the reduction of infrequent but extremly high-cost outcomes.

Setting planning-reserve margins solely based on a physical reliability metric, such as the often inconsistently applied 1-in-10 standard, does not allow stakeholders to understand these costs and risks and their tradeoffs. Our case study and past work for utility clients show that planning-reserve margins set based on an economic analysis of reliability costs and values tend to be in the 12% to 18% range for a medium-to-large system. The range of these results is not fundamentally different from the range of planning-reserve margins that are derived solely with physical reliability measures, such as the 1-in-10 standard. However, based on our experience, it is also often the case that reserve margins determined solely with physical reliablity analysis can differ by several percentage points from target reserve margins that are informed by an economic assessment of reliability. These differences, which can be negative or positive, are not systematic and depend greatly on factors such as system size, interconnections, and regional resource mix.

In summary, we believe that an economic simulation of reliability costs and benefits for different levels of reserve margins can (1) provide a significantly improved understanding of resource adequacy risks, (2) help determine more cost-effective solutions that consider the tradeoff between the expected level and uncertainty of reliability-related costs, (3) help us understand the link between economically efficient target reserve margins and physical reliability standards such as the 1-in-10 standard, and (4) inform stakeholders about the value customers are receiving in exchange for paying for reserve capacity. Sole reliance on physical reliability standards, in particular the often vaguely defined 1-in-10 standard, easily results in setting target reserve margins that—depending on system size and characteristics—are either too low or too high to be cost-effective and economically efficient.

Appendix A

Questions to Ask About Resource Adequacy Analyses

Every substantial bulk power reliability event, whether or not caused by a natural disaster, has been subjected to intense scrutiny by the industry and its regulators. For events caused by inadequate generation or import capabilities, the result of these reviews is typically a recommendation to address resource adequacy standards, the operational processes, or equipment that contributed to the event. However, it is impractical and cost-prohibitive to attempt to address every process and every piece of equipment that could impact bulk power reliability.

An analytical approach to study the economics of resource adequacy can quantify the cost and benefit of, as well as prioritize, proposed mitigation procedures. In the current environment of a 1-day-in-10-years physical resource adequacy standard and an N-1 reliability standard for transmission, the industry has the difficult task of determining whether reliability events are simply due to inadequate resource adequacy or whether they are a necessary and efficient trade-off to avoid the higher cost of stricter standards. Reliability studies that also evaluate the economics of resource adequacy standards would give regulators, utilities, and consumers visibility of the economic trade-offs between the costs of additional reserves and the risk exposure of carrying fewer reserves.

At the request of NRRI, we have compiled this set of questions that utilities, RTOs, and commissions can ask to arrive at a better understanding of the costs and benefits of reserve capacity.

A. General questions about reliability standards and target planning-reserve margins

1. How are planning-reserve margins determined?
2. Based on the 1-day-in-10-year or some other reliability standard?
3. How is the 1-day-in-10-year standard defined and what approach and assumptions are used to calculate it?
4. When have the economic and risk implications of planning-reserve margins based on the applicable resource adequacy standard last been evaluated?
5. What economic and reliability value do customers receive from paying for reserve capacity at the current target reserve margin levels? What benefits did customers receive by having a specific reserve margin in the past?
6. By how much would the risks of high-cost outcomes (e.g., due to emergency purchases) and load shed events increase (or decrease) if the target reserve margin was reduced (or increased) by 1 percentage point?
7. Does the market structure and applicable regulatory framework offer proper incentives to provide effective levels of reliability and cost stability?

8. What is the impact of price caps on setting an economic optimal reserve margin?
9. Can generation and transmission planning be coordinated by considering the economic interaction between the two disciplines? Or should they always be performed independently and with differing standards?
10. To what extent is your system reliant on support from neighboring systems?
11. How will the implementation of renewable portfolio standards affect reliability and cost exposure?

B. Specific questions about reliability studies

1. Does the system in your balancing area or jurisdiction contain a high percentage of energy limited resources (i.e., solar, wind, hydro, pump storage, demand response)? How do these impact reserve margin determinations?
2. What levels of penetration of demand side resources would require increasing dispatchability of these resources to provide the maximum economic and reliability benefit?
3. How will price-responsive loads impact the economics of reliability?
4. How are the impact of extreme weather conditions and their impacts on loads and generation availability modeled? How is the diversity of weather across neighboring systems modeled in terms of its impact on reliability and price risk mitigation?
5. How will potential load growth from technology advancements such as electric vehicles affect reliability planning?
6. What impact does transmission availability have on reliability costs to consumers and how is transmission availability modeled in reliability studies?
7. What impact will significant base load coal retirements have on resource adequacy planning?

Appendix B:

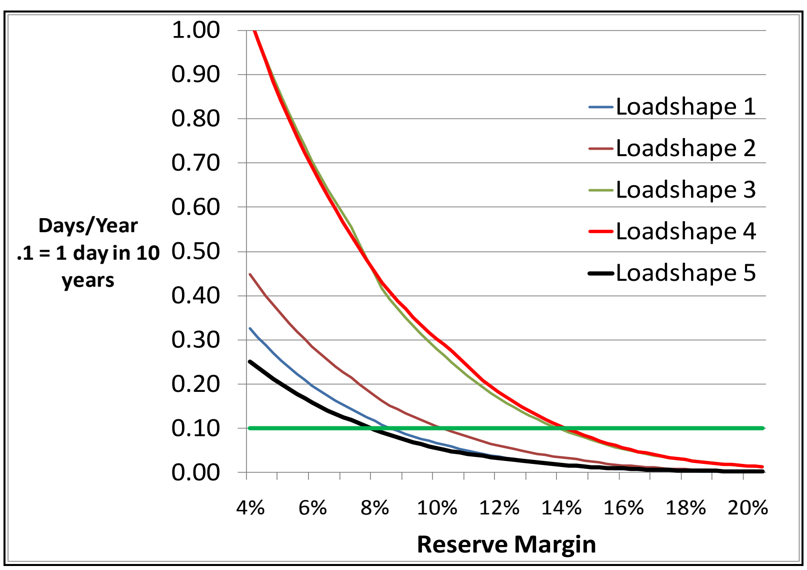
Additional Discussion of Resource Adequacy Study Inputs

A. Weather

Weather is the most difficult variable for planners to get their hands around but accurate representation of how weather affects load is critical to achieving meaningful results in reliability studies. The examples below show why it is important to model weather more robustly than what is done in many studies today.

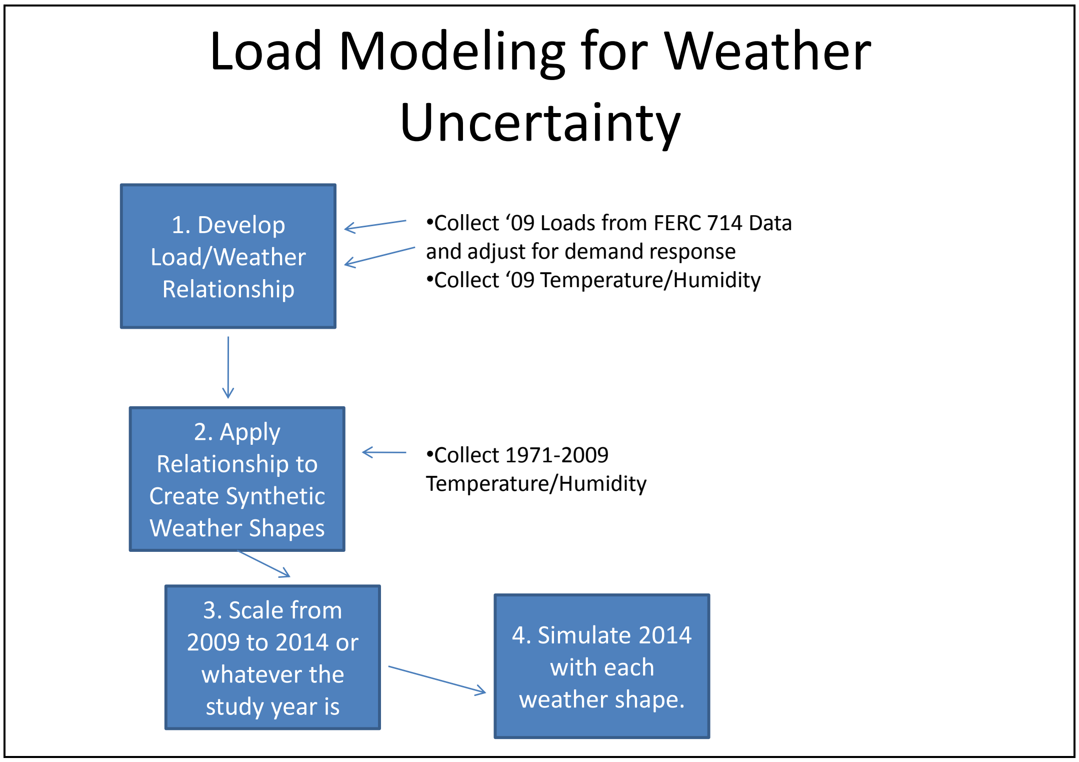
1. Weather impact on load shape

Depending on weather, the load shape in a given year or season can vary significantly. Some years will have two to three weeks of severe weather, while other years will have only one day of severe weather. LOLE calculations change drastically depending on the selected historical load shape. To address the weather impact on load, most planners will select a load shape from a single historical year and then scale that load shape up or down with multipliers to simulate both weather and economic forecast error. The problem with this approach is that the selected load shape significantly affects the LOLE results. To document this impact, we took load shapes from five individual historical years and scaled them to exactly the same peak load and annual energy. The same multipliers were then used for each of these five load shapes and each of them simulated in a reliability model. Based on the 1-in-10 standard, the target reserve margin determined with these five load shapes varied by 6 percentage points—from a low of 8% to a high of 14%—as shown in the figure below.



Most resource adequacy studies are not properly addressing this issue and are introducing a significant amount of error by relying only on a load shape taken from a single historical year without considering actual weather-related uncertainties.

The chart below summarizes the alternative approach we have taken in the case study presented in the main body of this paper. It represents a methodology that ensures that weather is being modeled accurately in load shapes. The best representation of what weather may look like in a future year is to evaluate actual historic information. To simulate a future year, we derive "synthetic" load shapes by analyzing weather for the last 40 years. This process develops a load-weather relationship based on the most recent load and weather information and then applies this relationship to 40 years of historical weather. This creates a realistic weather-dependent distribution of load shapes that can be scaled to the future year being analyzed so that the average of the system peaks will be equal to the actual "weather normalized" load forecast. All of the 40 derived individual load shapes are given equal probability of occurrence in the simulation. This also allows the planner to analyze what would happen if we had "2006 weather" again and will capture the true weather-driven load uncertainty from one year to the next.

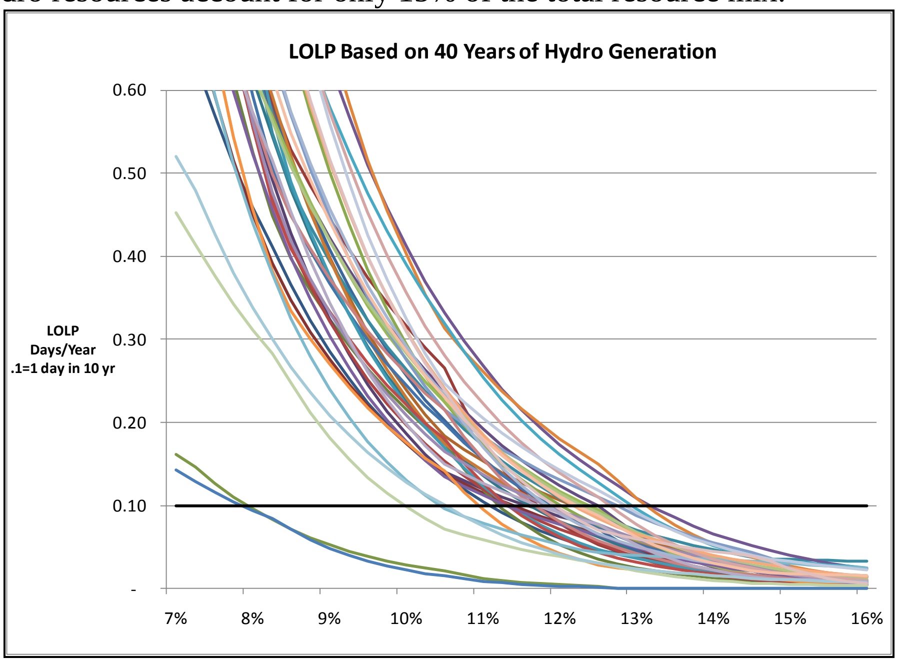


b. Weather impact on resources

Temperatures, rainfall, wind speeds, hours of direct sun, and other weather parameters also significantly impact the capacity available from generating resources. A portion or all of these parameters should be taken into account depending on the system's resource mix. Most models ignore the impact of temperature on thermal generating units and provide only a static representation for hydro, wind, and solar generation.

Our simulations link the thermal capacity to weather years similarly to the approach for load shapes discussed above. By doing this, we can ensure that the 40 load shapes that are being simulated can be correlated correctly with resource availability. Utilities planners in the Pacific Northwest, for example, have long recognized this issue with respect to their hydro fleet. They model weather years for both load and resources similarly to how we have outlined. The variability in their hydro fleet is significant and must be taken into account. As the recent experiences in Texas and the Southwest have shown, extreme weather can lead to substantial simultaneous outages of resources.

To illustrate this point, we analyzed a system that contained approximately 15% hydro resources. The analysis shows how the 1 day in 10 year reserve margin would change depending on the selected hydro year used in the study. We forced all other variables (*i.e.*, except hydro) to remain constant and chose a static single-year representation of hydro generation as most planners do. The results, presented in the following figure, show that the target reserve margins determined in reliability studies can change by more than 5 percentage points—from approximately 8% to over 13%—even if hydro resources account for only 15% of the total resource mix.



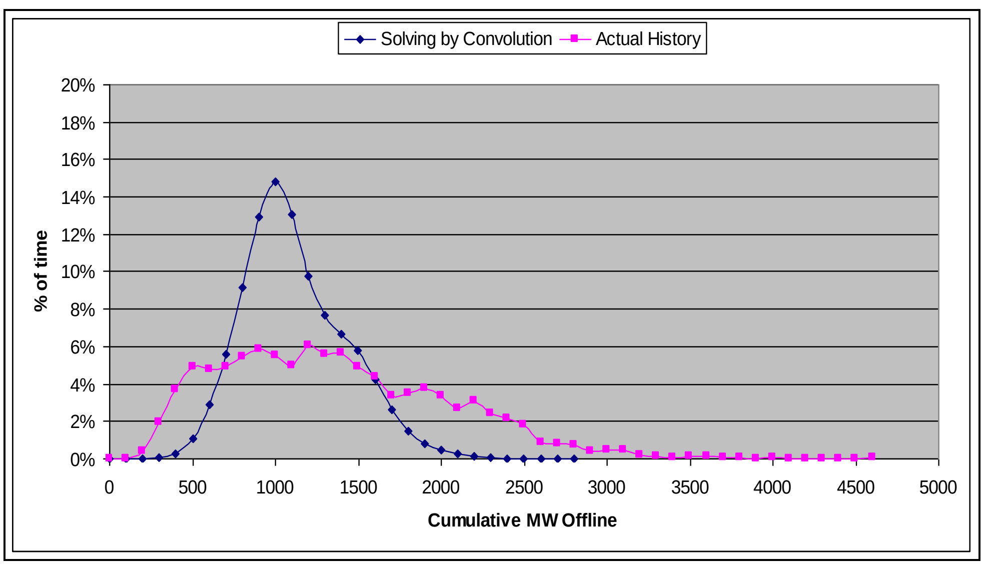
As the chart shows, relying on static representations of such energy limited resources that are significantly dependent on weather introduces a significant amount of error in reliability studies.

B. Economic forecast error modeling

In addition to correctly representing weather-related uncertainties, reliability studies need to consider the economic forecast error on anticipated future loads. As stated previously, many planners tie weather and economic uncertainty into one multiplier and adjust load shapes by this multiplier. As shown, this places too much emphasis on the chosen load shape. A better way to handle economic forecast error is to separate it from weather-related components. If weather uncertainty is represented through 40 weather years, the load forecast error multipliers and probabilities can be applied to each of the 40 weather years. Typically, 6-7 load forecast error point estimates and associated probabilities are sufficient, which means that the analysis has to consider 40x6 to 40x7 (i.e., 240-280) simulations to capture the effect of weather and economic forecast error on loads. This is not a problem given the computing power available today and the ability of fast dispatch-reliability planning models designed for this type of analysis.

C. Unit outage modeling

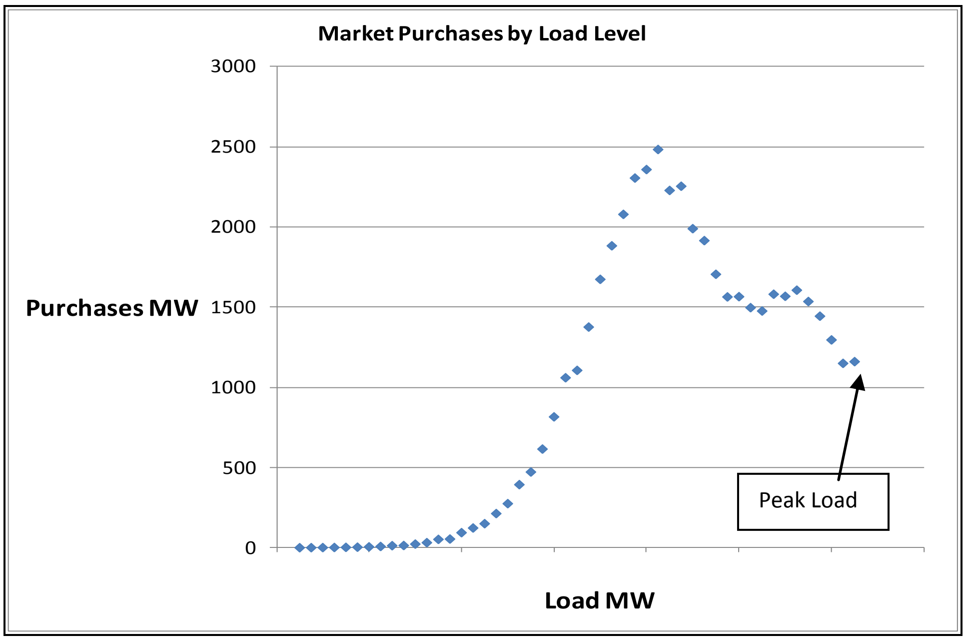
Most reliability analyses place sufficient emphasis on simulating generation outages. For reliability planning, the tails of the distribution are most important. Figure 4 shows a chart of two different distributions representing the amount of generating capacity offline for a given system. One of the distributions is based on the "convolution method" which is used by most production cost models. But this distribution is sufficient only if a planner wants to understand the expected average outage pattern and is not particularly concerned with the full range of the underlying distribution. In comparison, the second curve represents actual historical outages. As this comparison shows, actual outage patterns show a much higher probability that more than 2,500 MWs will be offline compared with the probabilities provided by the convolution method. A reliability model should be able to simulate a distribution that represents actual outage probabilities because it is in these low-probability outage situations that reliability events occur. A Monte Carlo method based on frequency and duration of outages should be used to mimic the more accurate distribution.



D. Reliability assistance from neighboring systems

Some planners assume that their system is an island without the possibility of assistance from neighboring system. Others only model a static representation of what is available in the market from neighbors. Modeling reliability assistance from neighboring system correctly is another important factor in any resource adequacy study.

In our experience, treating a system as an island results in target reserve margins that are 4 to 8 percentage points higher than if interties with neighboring systems were considered. Understanding how much capacity is going to be available from neighboring systems consequently is an important factor in the analysis. A multi-area transportation model that dispatches the system being studied as well as its neighbors is needed to truly understand what will be available in the region. We all realize that during peak hours, when the entire region is experiencing similar weather and similar economic growth uncertainty, that there will be much less capacity available than during off peak hours. Nevertheless, most reliability analyses do not model the implications of this factor. We recommend modeling different weather years for each neighboring system to accurately capture the region's weather diversity. For the case study discussed in the main part of this paper, the following figure shows how much capacity could be purchased at different load levels. As expected, it shows that available purchases from neighboring systems drop off significantly as load reaches its ultimate peak. This phenomenon needs to be captured because, as we have seen, many purchases are made during high-load hours close to system peak conditions to conserve the capacity of energy-limited resources for the peak periods.



E. Scenario modeling

The typical reliability study runs 10 cases representing 10 load forecast multipliers and then runs 500 unit outage draws for a total of 5,000 hourly simulations at a defined reserve margin level. In the case study analyses discussed above we simulated 280 combined weather and load forecast cases with 400 unit outage draws for a total of 112,000 hourly, year-long simulations. This is easily handled with today's computing resources and ensures that weather as well as the other key variables is being taken into account accurately. For this case study, SERVM performed these 112,000 annual simulations (with 8,760 hours each) in approximately 2 hours using a machine with eight processors. This could be cut in half with the use of two computers.

Reliance on single-year load shapes, static representation of energy limited resources, and only high-level assumptions about reliability assistance available from neighboring systems will introduce substantial error in the analysis. Again, with the computing power available today, there is no longer any reason why planners modeling large interconnected regions would need to rely on such simplifications.