A New Resource Adequacy Standard for the Pacific Northwest

Background Paper

12/6/2011

A New Resource Adequacy Standard for the Pacific Northwest Background Paper

CONTENTS

ABSTRACT

This paper describes a revision to the Council's 2008 Resource Adequacy Standard, which was adopted by the Council in December of 2011. The standard is used as an early warning system to indicate when resource development does not adequately keep up with demand growth. Like the 2008 standard, the new standard uses the system's loss of load probability (LOLP) as the adequacy metric and keeps the maximum allowable LOLP at 5 percent.

However, instead of calculating separate LOLP values for winter and summer, a single annual value is assessed. This bypasses potential problems with assessing only two of the four seasonal values. In addition, assumptions regarding the use of standby resources have been refined. Those resources are now identified individually and are limited to non-modeled resources and load management operations that are contractually available to regional utilities.

The new standard maintains the 2008 philosophy that a reasonable amount of market supply should be included in adequacy assessments. The amount of these resources to be used in the assessment will be decided by the Forum.

Another major change is that the translation of the probabilistic LOLP metric into static metrics (load/resource balance and capacity planning margins) will no longer be a part of the standard. Of course, this doesn't prevent planners from making these translations for limited use.

In addition to assessing the LOLP, a *State of the System* report will be produced, which will include information about the frequency, duration, size and timing of potential shortfalls. It will also highlight conditions under which major shortfalls could occur and it will provide statistics on the use of non-firm resources. The new standard is designed to be simpler and more informative than the 2008 version and should prove to be much more useful to regional planners, commissioners, policy makers and other involved parties.

SUMMARY

In 2008, the Northwest Power and Conservation Council adopted the Resource Adequacy Forum's proposal, which set a standard for power supply adequacy in the Pacific Northwest. Since then the region has used that standard to assess the ability of the power supply to adequately serve Northwest customers. In the past three years, assessments

have shown the regional supply to be adequate but also that the surplus of resource capability over demand in summer months was growing smaller.

While the 2008 standard has been very useful to planners, it is somewhat cumbersome to calculate and it invariably leads to misinterpretations every year. The 2008 standard is based on a probabilistic simulation of the region's power supply. That probabilistic measure is converted into more commonly used static measures of power supply, namely annual load/resource balance and hourly planning reserve margins. Comparison of these static measures to similar calculations reported in utility publications has always led planners into a quagmire of confusion, mostly because each set of calculations is used for different purposes.

Because of the issue mentioned above and other considerations, the Council chose to have this methodology reviewed by a peer group. Results indicated that while the current standard was sound, it could be improved. This paper describes the 2008 standard and the revisions prompted by the peer review.

There are three elements of an adequacy standard that can be addressed separately. Those three elements are; 1) the methodology, 2) the metric and threshold used to define an adequate supply and 3) assumptions made about resources and loads, in particular about non-firm and standby resources. The peer review assured us that the methodology is appropriate. Assumptions regarding non-firm and standby resources are still being refined and the Forum will have an opportunity at a later date to review and approve those assumptions. *Thus, this paper focuses only on the metric and threshold*.

The 2011 proposed revisions to the current adequacy standard make it simpler and more informative. It keeps the probabilistic approach to assessing adequacy but drops the conversion to static metrics. Its function is focused on being an early warning should regional resource development fall dangerously short. In this sense, it is like a smoke alarm – everyone has to have one but hopes it never goes off. But, while the new standard may be simpler, the annual assessment will provide much more analytical information that should prove to be very useful to utility planners. This additional information will be presented in the *State of the System* report.

The new adequacy metric remains the loss-of-load probability (LOLP) and the associated threshold is similarly kept at 5 percent. In other words, the power supply is deemed to be adequate if its LOLP is 5 percent or less. The LOLP is based on an annual analysis (as opposed to seasonal analyses done in the 2008 standard), in which the operation of the power supply is simulated over each hour of every month. The LOLP is calculated by dividing the number of yearly simulations with significant curtailment events by the total

number of simulations. Significant curtailment events are defined as shortfalls that exceed the capability of standby resources and contractual load management actions.

BACKGROUND

In 2005, the Pacific Northwest Power and Conservation Council (Council) and the Bonneville Power Administration (BPA) created the Pacific Northwest Resource Adequacy Forum (Forum). The Forum includes representatives from the region's electric utilities and utility organizations, public utility commissions and public interest groups, as well as from BPA and the Council. It is made up of a steering committee and a technical committee.

The Forum's overarching goal is to *"establish a resource adequacy framework for the Pacific Northwest to provide a clear, consistent, and unambiguous means of answering the question of whether the region has adequate deliverable resources to meet its load reliably and to develop an effective implementation framework."*

Toward that end, the Forum reached a consensus in recommending a Resource Adequacy Standard[1] (Standard) for the Pacific Northwest region. The Forum's proposed standard was officially adopted by the Council in April of 2008.

The Standard serves as a gauge to assess whether the Northwest's electricity supply is sufficient to meet the region's needs now and in the future. It provides a *minimum threshold* that serves as an early warning should resource development fall dangerously short. It does not mandate compliance or imply any enforcement mechanisms. It does not directly apply to individual utilities – because every utility's circumstances differ. What it is intended to do, is to initiate a discussion among appropriate regional entities, should the assessment indicate that resource development is seriously lagging.

The current standard has been in use since 2008. Every year since then, the adequacy assessment has shown the Northwest's power supply to be adequate. However, the most recently published assessment[2] warned that our summer electricity needs were growing rapidly and that increasing amounts of variable generation (such as wind) are adding to the complexity of system operation.

---

[1] The 2008 standard can be found at http://www.nwcouncil.org/library/2008/2008-07.pdf.

[2] The latest published assessment can be found in Chapter 14 of the Council's Sixth Power Plan, at http://www.nwcouncil.org/energy/powerplan/6/final/SixthPowerPlan_Ch14.pdf.

Since its inception, the adequacy assessment has sparked a debate about what it really measures and how it relates to individual utility planning. It was pointed out that an apparent discrepancy existed even in the Council's own power plan. The assessment in the sixth plan indicated that the power supply was adequate at least through 2015, yet the resource acquisition strategy in the same plan encouraged aggressive acquisition of conservation and other resources beyond the adequacy threshold. *The answer, of course, is that an adequacy assessment does not equate to a resource needs assessment*. The latter involves much more than simply "keeping the lights on." While an adequacy assessment indicates the likelihood of curtailment, cost notwithstanding, a resource needs assessment takes economic, political and environmental effects into consideration. For example, the amount of reliance on market supply becomes a policy decision when performing a resource needs assessment. Some utilities may choose to assess their resource needs in the absence of a market supply.

Nonetheless, it became clear to Forum members that the current standard was not as useful to planners as originally believed. Thus, the Forum initiated a peer review[3] for the methodology used to develop the standard. That review was completed in May of 2011 and recommended that the standard be modified to include a measure of the size of potential problems. It also implied that a secondary measure should be assessed; one that measures how often standby resources might be dispatched. When the dispatch likelihood for standby resources grows too high, system operators are likely to become more uneasy, which may be a good indication that new resources are needed. This secondary measure goes beyond assessing whether the supply is adequate and is likely to be more consistent with utility resource need assessments.

METHODOLOGY

The Council uses a Monte Carlo (probabilistic) methodology to assess the adequacy of the region's power supply. The GENESYS[4] computer model was developed in 1999 and performs a detailed hourly chronological simulation of the Northwest's resources, including the hydroelectric system, over many different possible future conditions. Each future's operation is simulated using different assumptions for uncertain variables, namely; 1) river flows (which affect the amount of water for hydroelectric generation), 2)

[3] A summary of the peer review is provided in Appendix B.

[4] *Northwest Power Supply Adequacy/Reliability Study Phase 1 Report,* Council Document 2000-4, March, 2000. http://www.nwcouncil.org/library/2000/2000-4.pdf

temperature (which affects demand for electricity), 3) forced outage conditions for generating resources and 4) wind generation.

GENESYS provides much more information to decision makers than simple deterministic (static) comparisons between resources and demand. Besides the expected values for hydroelectric and thermal generation, it also provides statistical data regarding the operation of each resource. Most importantly, it records hours when the power supply is not able to meet all demand obligations – and conditions under which these shortfalls occur. The frequency, duration, size and timing of curtailment events are recorded and the overall probability of not being able to fully serve load is calculated. This probability, commonly referred to as the loss-of-load probability (LOLP), is the metric used to assess adequacy of the power supply.

For the Northwest, the Council has defined an adequate power supply to have an LOLP no greater than 5 percent. This means that out of all the simulations run, no more than 5 percent of them had significant curtailments. In other words, the likelihood of the region experiencing a significant shortfall during any time of the year must be 5 percent or less in order for the supply to be deemed adequate.

THE 2008 STANDARD

The 2008 standard uses the 5 percent LOLP threshold as a basis to calculate minimum thresholds for three deterministic metrics, one for winter energy, one for winter capacity and one for summer capacity. These deterministic metrics act as a surrogate for the LOLP metric and can be calculated without the use of a Monte Carlo simulation program. It was believed that by defining thresholds for these deterministic metrics, the adequacy of the power supply could more readily be assessed.

The deterministic adequacy metric for energy is the annual balance between resources and demand. The resources include reasonable assumptions about the availability of non-firm resources and market supplies. The demand forecast includes assumptions regarding the anticipated amount of efficiency gains in the year to be analyzed. The threshold for the energy metric (annual load/resource balance) is zero, meaning that resources (including market supply) must at least equal the amount of firm demand, on an annual basis, in order to provide an adequate supply.

The deterministic adequacy metric for capacity is the planning margin or the surplus generating capability over a sustained peak period. The sustained peak period is defined to be the 6 highest demand hours of the day over 3 consecutive days. This 18-hour non-contiguous period is intended to capture the highest demand hours over a winter cold

snap or summer heat wave event. The minimum thresholds for the winter and summer sustained-peak planning reserve margins are 23 and 24 percent respectively and equate to an LOLP of five percent. Resource assumptions used to calculate the system's planning margins are similar to those used for the energy metric, in that they include some non-firm and market resources.

Unfortunately, the use of deterministic adequacy metrics became problematic. For one thing, each time the system changes, deterministic thresholds have to be recalibrated to the 5 percent LOLP standard. Secondly, it was difficult to compare the annual load/resource balance and the planning margins to similar metrics published in utility reports[5] – because the purposes of each are different. Thus, the Forum spent much time after every assessment explaining why these measures should be different from those in other reports.

Because of issues with the deterministic metrics, the Forum chose to focus on the probabilistic LOLP metric for its assessment. However, this was also problematic because there was not just one LOLP to calculate but three. LOLP assessments were made for winter energy and capacity needs and for summer capacity needs. The 2008 standard implies that as long as all of these LOLP values are 5 percent or less, the power supply is adequate. Unfortunately, this approach is faulty and could yield false negative results because each LOLP value is assessed independently of the others. Situations could easily occur when all three LOLP values are less than 5 percent but the overall likelihood of experiencing a problem in either winter or summer is greater than 5 percent. This happens if winter and summer shortfalls occur in different simulations. The Forum now concedes that using multiple LOLP metrics is faulty and suggests using a single annual LOLP value, which identifies both energy and capacity problems.

Another concern with the current standard is that it provides no indication of the size of potential problems. Just knowing the likelihood of curtailment is important but also knowing the size of the problem can be extremely beneficial to utility planners. In fact, the Forum recognizes that there exists a wealth of analytical data that can be very useful and suggests that it be provided along with each annual adequacy assessment.

THE PROPOSED REVISIONS

[5] These reports include the Pacific Northwest Utility Conference Committee's Northwest Regional Forecast of Loads and Resources and the Bonneville Power Administration's White Book.

The Forum's 2011 proposed revisions to the 2008 adequacy standard make it simpler and more informative. It keeps the LOLP as the metric of choice and also keeps the 5 percent threshold for adequacy. However, it differs in many other ways. The first difference is that the LOLP is now calculated on an annual basis (i.e. every hour of the year is simulated, not just winter and summer periods). Secondly, capacity and energy shortfalls are counted together (rather than separately as in the current standard). The assessment will only be made five years into the future and will count only existing resources (or those expected to be operational within the 5-year assessment period) and expected levels of efficiency gains.

Another difference is that capacity and energy shortfalls are defined differently. In the current standard, an energy "event" was counted only if the total curtailed energy in the winter or summer period was greater than 28,800 megawatt-hours. A capacity "event" was counted if any hour's curtailment in winter or summer exceeded 3,000 megawatts. These curtailment thresholds were used as a surrogate for additional actions (not modeled in GENESYS) that the region could take to avoid a shortfall during times of stress. These additional actions include dispatching standby resources or executing contractual load reduction agreements.

The new standard keeps the idea of using emergency actions to set the thresholds for counting curtailment events. However, the surrogate values are replaced by the actual aggregate capabilities of these emergency actions. These actions contain resources or load management actions that utilities have contractual rights to but are generally not used on an annual basis. These resources can have both capacity and energy components. This group of emergency actions will be referred to as "standby resources."

The new standard is based on an annual LOLP calculated by adding up all simulations in which curtailments exceed the capabilities of standby resources and dividing by the total number of simulations. If the LOLP is 5 percent or less, the power supply is deemed to be adequate.

The Forum kept the new standard simple because it wanted keep the function of the adequacy assessment as an early warning (e.g. like a smoke alarm). The alarm may never go off. However, should resource development lag, for whatever reason, the alarm should sound and appropriate regional entities should gather to discuss whether the alarm was justified and, if so, what actions should be taken. Had this system been in place in the 1990s, the alarm would have sounded in about 1995.

However, the Forum also recognizes that a lot of very useful information is provided by an adequacy assessment study. For that reason, the Forum has chosen to provide

additional information along with the adequacy assessment (e.g. the LOLP value). This additional information is described in more detail below but basically provides statistical data on the frequency, duration, size and timing of potential problems. It also provides information regarding the use of standby resources and non-firm resources. Even if the power supply is deemed to be adequate, this additional information, provided in a new *State of the System* report, should prove to be very valuable to utility planners.

In summary, the revisions adopted by the Council are listed below:

- Keep the current probabilistic methodology, which includes the use of market resources, to assess power supply adequacy.
- Keep the LOLP as the adequacy metric and keep five percent as its maximum threshold.
- Remove the seasonal LOLP assessments and instead calculate only one LOLP value for the entire year.
- Change the assumptions regarding the thresholds used to screen shortfalls. The new thresholds should only count the aggregate capability of standby resources.
- Eliminate the translation of the LOLP measure into deterministic values such as load/resource balance and winter and summer sustained-peak planning margins.
- Include a *State of the System* report will all future assessments.

STATE OF THE SYSTEM REPORT

The Forum plans to assess the adequacy of the Northwest's power supply each year, determining whether it can adequately provide for our needs through a five year period. In addition to that assessment, a *State of the System* report will be produced, which will include information about the frequency, duration, size and timing of potential shortfalls. It will also highlight conditions under which major shortfalls could occur. It will provide statistics on the use of market resources. The new standard is designed to be simpler and more informative than the 2008 standard and should prove to be much more useful to regional planners, commissioners, policy makers and other involved parties. The following section describes some of the additional adequacy metrics that will be reported in the *State of the System* report.

CONDITIONAL VALUE AT RISK

This metric measures the magnitude of curtailment and is recommended in the peer review of our methodology[6]. It reflects the average magnitude of the worst potential curtailments. It is calculated by summing up the total curtailed energy (in megawatt-hours) for each simulated year. This data is then sorted and the total curtailment for the worst five percent of simulations is averaged. As stated in the peer review report, this measure is a coherent measure that is easily incorporated into resource planning optimization models (however not currently used by the Council).

An alternative or supplement is to calculate a CVaR measure for the single worst curtailment *events*. A curtailment event is defined as a series of contiguous hours in which demand is not met. To calculate this variation of CVaR, the largest curtailment event is extracted from each simulated year instead of the total annual curtailment energy.

A second variation is to calculate a CVaR measure for the worst single-hour curtailment. The process is similar to that for calculating the worst-event CVaR. However, neither of these variations is recommended by the peer review and so is not included.

Loss of Load Expectation

This metric measures the frequency of curtailments and is commonly used in other parts of the United States and the world. Unfortunately, it not always well defined. In principle it calculates the expected number of hours of curtailment per year (more aptly named the loss of load hours or LOLH). However, many regions use different sets of random variables (future uncertainties) to calculate LOLH. A region that only uses forced outages to calculate LOLH cannot properly compare its value to another region's LOLH that includes a different set of random variables in its simulation. In the Northwest, for example, water supply, temperature, forced outages and wind generation are modeled as future uncertainties.

The common LOLE (more precisely LOLH) threshold used to define an adequate supply is one-day-in-ten years, meaning that in a ten-year study no more than one day can show a failure to meet demand. However, this threshold is also commonly misinterpreted. Since most utilities now do single-year analysis, the one-day-in-ten-year threshold has been redefined as 2.4 hours per year by many regions. This redefinition is not equivalent to the one-day-in-ten-year threshold – because to make them equal, the "one day" would have to show curtailments for all 24 hours.

---

[6] See the Resource Adequacy Forum web site at http://www.nwcouncil.org/energy/resource/Default.asp .

Nonetheless, the recommendation is to report an LOLH value for the Northwest's power supply. The suggested calculation is to take the total number of hours of curtailment and divide by the total number of years simulated. The question remains, however, whether the 2.4-hours-per-year threshold should apply to the Northwest because additional random variables are used in our simulation. The Forum recommends that this issue be discussed in more detail at national forums.

Expected Unserved Energy

This metric measures the size of curtailment. As with the LOLE and LOLH, the EUE is also commonly used across the country. To calculate EUE for the Northwest, the total amount of unserved load (in megawatt-hours) over all simulations is divided by the number of simulations. This yields the average amount of annual energy shortfall.

A variation of this metric is to normalize it by dividing the EUE by the expected load, which yields the average percentage of load that is expected not to be served. As with the LOLE and LOLH metrics, comparing EUE measures across regions could be problematic because of the same issues, namely the number and types of random variables used.

Expected Use of Standby Resources

This is a newly defined metric that measures how often standby resources are likely to be used. Standby resources are resources or load management actions that utilities have contractual rights to and are used only during times of stress to avoid shortfalls. The expected use of standby resources metric or EUSR is the number of simulations in which standby resources are used (at least in one hour) divided by the total number of simulations. It is an LOLP type of metric but will always be larger than the LOLP (assuming that standby resources exist).

Expected Use of Market Resources

The Forum also recommends providing statistics on the use of market resources. Market resources include independent power producers within the region, out-of-region market supplies and the use of "borrowed hydro" (water below drafting rights elevations that is used for short periods and then replaced).

At this time it is not clear how much of this statistical information will prove to be helpful. However, as a starting point, the Forum will report the expected annual and monthly use of these resources. If necessary, probability duration curves, which show the range of use for these resources, can also be reported.

Curtailment Statistics

The Forum will also provide statistical information regarding curtailment events (as defined above). The recommended information may include;

- Frequency of events (events/year)
- Average magnitude of events (megawatt-hours)
- Average duration of events (hours)

The information above will be based on "raw" curtailment events, in other words, curtailment is *not* reduced by the capability of standby resources. However, the same type of information can be provided for events that have been adjusted by subtracting standby resource capabilities.

MONTHLY STATISTICS

The Forum will also provide a number of monthly statistical values, especially for LOLP, EUSR and the expected use of market and non-firm resources. Monthly values are calculated in the same manner as annual values. For example, the monthly LOLP value for January is calculated by adding up the total number simulated Januarys that had curtailment (in excess of the standby resource capability) and dividing by the total number of Januarys simulated.

Non-modeled small resources and standby resources have limited capabilities that can be "used up" prior to the end of the operating year. The raw curtailment record produced by GENESYS will be adjusted to account for the use of non-modeled resources by dispatching them as needed beginning in October (start of the operating year). The same assumption is made for standby resources. Once the capabilities of these resources are used up, they are no longer available to help out during periods of shortfall. Thus, we may see higher impacts to summer months, especially if these resources were dispatched during the winter.

An issue regarding monthly statistics, however, is that all chronological links across months will be lost. In other words, monthly statistics will be based on each month's simulation results independent of other month's results. Thus, although a monthly report (or graph) will help identify periods of the year that are more susceptible to curtailments, it will have no mathematical connection to the annual values. In a sense, it will show a non-realistic "picture" of the adequacy of the power supply (at least in aggregate) because the worst curtailments for each month do not all occur in the same year. Nonetheless, it is a report (using proper caveats) that the Forum believes will be very informative to planners.

HOURLY STATISTICS

Hourly statistics can also be very useful to system planners. Probability curves for single-hour curtailments by month will show how big a single hour shortfall can get. This data provides more detail than the CVaR value for peak events.