Reserves in Capacity Planning
A Northwest Approach

PNUCC

System Planning Committee

June 2010

Electronic copies of this report are available on the

PNUCC Website

www.PNUCC.org

101 SW Main Street, Suite 1605

Portland, OR 97204

Memorandum

To: Board of Directors

From: Steve Kern, System Planning Committee Chair

Date: June 4, 2010

Subject: Paper Completed – Reserves in Capacity Planning

Over the last few months a subgroup of the System Planning Committee and PNUCC staff have been on track to document what the Committee has learned regarding the role of operating reserves in regional capacity planning efforts. The question of how individual utilities are incorporating operating reserves in their planning was raised. The attached technical paper, "*Reserves in Capacity Planning, A Northwest Approach*" addresses this question by providing an overview of how operating reserves are considered in assessing adequacy, choosing an adequacy target and how they are incorporated into long-term planning.

The region's utilities have typically planned for new resources based on a need for energy. However, with the significant increase in intermittent resources and the forecast for more large amounts to be integrated into the region's power system, it has become crucial to consider capacity needs in long-term planning as well.

The System Planning Committee concluded in compiling this paper that long-term planning efforts should reflect operating reserves. The nature of intermittent resources and the ability to manage their fluctuating generation must be accounted for. This paper provides a better understanding of how the operating reserves can be reflected and recognizes how subtle differences in approaches can create significant differences in estimates about system reliability and adequacy. It underscores the need to cautiously proceed when making comparisons of different operating reserves estimates and establishing planning margins.

Finally, I would like to acknowledge the subgroup that devoted significant time and took the lead to work through the details of this paper. It is their effort that allowed the full Committee to engage, providing input and review throughout the process. This work will benefit all PNUCC members and enhance Northwest utilities' general understanding of this technical subject matter.

Reserves in Capacity Planning

Table of Contents

Long-term planning to ensure adequate power supply is an essential element of Northwest utilities' work. Providing customers with reliable electric service requires utilities to engage in complex planning processes. Here in the Northwest utilities have traditionally relied on assessments based on energy needs. However, as the region becomes more capacity constrained the need for capacity planning in the Northwest grows. With a focus on capacity planning the issue of what to do about operating reserves comes into play. The PNUCC System Planning Committee developed this paper to addresses one narrow aspect of long-term planning – if and how to include operating reserve in their analysis.

"A discussion of approaches to resource adequacy assessment ….is a challenge in semantics".[1] As this paper was drafted, the Committee members experienced this challenge. In addition to reinforcing the need to account for operating reserve in planning, the writing of this paper highlighted the communication challenge in discussing the topic.

In reviewing the work of Northwest entities, including utilities and other entities across the nation, the Committee found that planners are incorporating operating reserve in various ways. This paper describes the different approaches for reflecting operating reserve in stochastic modeling efforts and it offers cautions about comparing and contrasting reliability metrics. Readers will gain a better understanding of how the operating reserve can be reflected and recognize how subtle differences in approaches can create significant differences in estimates about system reliability and adequacy.

I - Introduction

Electricity is a unique necessity that cannot be stored and must be used the instant it is generated. This means that the regional electric grid must balance demand and supply second-to-second. Given this extreme form of just-in-time delivery, balancing authorities are required to maintain sufficient generation operating reserve to manage unexpected, short-term deviations in demand or supply. Long-term resource planners must consider operating reserve obligations to ensure that system operators will have sufficient resources to meet end-use demands plus reserve obligations required to maintain grid stability.

There are a variety of reasonable methods to establish and apply long-term resource planning targets, and there are different ways of reflecting operating reserve obligations in such analyses. Given the numerous ways analysis can be performed, care must be taken to ensure operating reserve is properly reflected without unintentionally under- or over-stating the obligation. This paper provides a summary of various elements of operating reserve and planning margin, followed by a discussion of methods for long-term resource adequacy assessment. An example of integrating operating reserve and planning margin is also provided.

[1] See p. 41 of NERC's Reliability Assessment Guidebook, Version 1.2, April 2009.

II - Overview of Reliability and Reserve Obligations

NERC's Role in Ensuring Reliability

NERC has a long history of being associated with ensuring the reliability of the bulk power systems on the North American continent. Over time its role has evolved from one of coordinating a voluntary set of reliability metrics to developing and enforcing statutorily-designated reliability standards, an outcome of the Energy Policy Act of 2005. NERC, with oversight by governmental authorities in both Canada and the United States, performs four primary reliability activities: 1) develop and enforce reliability standards; 2) monitor the bulk power systems; 3) assess and report on future adequacy; and 4) offer education and certification services to utility personnel.

With respect to operating reserve, NERC affects Northwest parties primarily through its promulgation and enforcement of the NERC Balancing Standards BAL-001 and BAL-002. These standards specify the performance requirements for control under normal and disturbance conditions.

> **Normal Conditions** – control performance under normal conditions is directly associated with regulating reserve which is often defined to include load following.
>
> **Disturbance Conditions** – control performance under disturbance conditions is most closely associated with contingency reserve.

In addition to the NERC BAL-002 standard, the WECC has its own contingency reserve standard, BAL-STD-002-0.

Operating Reserve and Planning Margin

For starters, what is the difference between operating reserve and planning margin? Operating reserve is a formal obligation that balancing authorities must observe under the North American Electric Reliability Corporation (NERC) and Western Electricity Coordinating Council (WECC) standards. This reserve is comprised of generation used to meet varying end-user demands and supply fluctuations caused by generation or transmission outages.

Planning margin differs from operating reserves. It is a buffer above normal weather end-use demand and is generally considered only in long-term resource planning studies. Its purpose is to achieve a desired level of electrical system reliability. Planning margin addresses load and resource balance uncertainties caused by extreme temperature variation and unexpected generation outages. Levels are determined through reliability studies described in Section V. Resources to meet planning margins are intended to be used to meet end-use demand, that is, they are not "reserved". In contrast, operating reserve is not intended to be used to meet end-use demand (other than small fluctuations in demand), but instead to maintain real-time system reliability in the event of outages or load and resource fluctuations.

Operating Reserve Defined

Operating reserve can be best thought of as the capability above firm system demand required by the balancing authority to maintain system balance. Reliable operation of the bulk power system requires that adequate capacity be available at all times to maintain scheduled frequency (60 hertz in the WECC) and to avoid a loss of firm load following due to either a generation or transmission outage. The WECC reliability standard specifies a minimum amount of operating reserve that must be maintained by each balancing authority. Specifically, this capacity is needed to:

- supply requirements for load variations,

- replace generating capacity and energy lost due to forced outages of generating or transmission equipment,
- meet local area protection,
- meet on-demand obligations (e.g. reserve trades with other balancing authorities),
- replace energy lost due to curtailment of non-firm or interruptible imports, and
- compensate for generation deviation of variable energy resources such as wind and solar generation.

Common usage of the term "operating reserve" within the electricity industry has been quite imprecise, and it is often used in reference to only one component of operating reserve – the contingency reserve. However, the proper definition of **operating reserve is the sum of contingency reserve <u>and</u> regulating reserve**.

Operating Reserve–Contingency

The portion of operating reserve held to mitigate the impacts of contingencies is relatively easy to envision. It is defined as reserve generating capacity (or rights to interrupt delivery of generation) necessary to allow an electric system to recover from disturbances such as generation failures and to provide for load following and frequency regulation.

The WECC defines contingency reserve requirements as at least the greater of:

- the loss of generating capacity due to forced outages of generating or transmission equipment that would result from the most severe single contingency (at least half of this reserve requirement must be spinning),

OR

- the sum of five percent of the load responsibility served by hydro-generation and seven percent of the load responsibility served by thermal generation (at least half of this reserve requirement must be spinning);

PLUS

- additional non-spinning reserve equal to amount of interruptible imports, and
- additional reserve for on-demand obligations (these would typically be reserve trades with other balancing authorities).

In this context, load responsibility is defined as balancing authority firm load demand plus any firm balancing authority exports (sales) minus any firm imports (purchases) for which contingency reserve capacity is provided by the supplier balancing authority.

In addition, contingency reserve consists of both spinning and non-spinning reserve.

> **<u>Spinning reserve</u>** is unloaded generation or load synchronized to the system and ready to automatically serve additional demand.
>
> **<u>Non-spinning reserve</u>** is generating capacity capable of being synchronized and ramping to a specified level in ten minutes, or load (generation delivery) capable of being interrupted within ten minutes. In either case the generation addition or load interruption must be sustained for a period of no less than sixty minutes.

Non-spinning reserve capacity can be provided in a number of ways. It can be met within the WECC by use of:

- load which can be interrupted within ten minutes of notification,
- interruptible balancing authority export schedules,
- on-demand rights from other balancing authorities,
- spinning reserve in excess of spinning reserve requirements (regulating reserve plus spinning contingency reserve),and
- off-line generation which can be started, synchronized and loaded within ten minutes.

Balancing authorities must maintain sufficient contingency reserve to meet the performance requirements of the NERC BAL-002 standard 100% of the time. NERC BAL-002 requires area control error (ACE)[2] to return to at least zero where its position is counter to where it was in the period immediately prior to the disturbance. This must be accomplished within fifteen minutes of the initiation of the disturbance.

Contingency Reserves in the Northwest

In the Northwest, the contingency reserve portion of the operating reserve is met by utilities through the Northwest Power Pool Reserve Sharing Group. Balancing authorities within the Northwest Power Pool area operate within a reserve sharing group administered by the Pool. A reserve sharing group is a group of balancing authorities acting together to meet reliability requirements while lowering the overall regional reserve obligation. NERC requires these groups to have written documentation specifying their rules. The rules may add requirements that are more specific or stringent than required by NERC. For example, the Power Pool reserve sharing group requires its members to increase their reserve requirements by five percent of the wind generation.

Within the Power Pool's reserve sharing group, the contingency reserve is provided to render recovery assistance following a disturbance such as the loss of a generating resource, a failure of a generating unit to increase output as instructed, or a transmission-related event impacting ACE in an instantaneous and unexpected change. The contingency reserve recovery assistance period is sixty minutes from the beginning of the contingency. Where the contingency lasts beyond sixty minutes, the balancing authority experiencing it must make arrangements outside of the Pool's sharing group. The individual balancing authority must also restore its own contingency reserve to pre-disturbance levels.

Operating Reserve–Regulating

The regulating portion of operating reserve is often the most difficult to understand for non-operating personnel. A particular difficulty arising in the discussion of regulation is that there is neither a universally accepted time frame associated with it, nor is there a well defined simple rule to state how much must be carried. Moreover, there are also differences of opinion among balancing authority operators as to terminology applied to regulating reserve. That being said, for planning purposes in the Northwest, it is suggested that regulating reserve be considered to be defined as follows:

> **Regulating Reserve** – spinning reserve capacity immediately responsive to automatic generation control (AGC) dedicated to meeting the moment to moment variations in energy demand or generator output. Regulation is the reserve component deployed to accommodate variability on a time scale of four seconds to five minutes.

[2] Area Control Area (ACE) is a real-time indicator of how much generation differs from load within a balancing authority. The calculation of ACE is defined by NERC.

In the Northwest, primarily in Washington and Oregon, hydro generation, particularly on the main stem of the Columbia River, provides the regulating reserve service. However, it can be provided by other generation technologies as well.

The WECC reliability standard requires each balancing authority to maintain regulating reserve sufficient to meet its NERC control performance criteria as specified in NERC BAL-001. The control performance standard sets the limits of a balancing authority's area control error for a specified time period. These include Control Performance Standard 1 (CPS1) and Control Performance Standard 2 (CPS2)[3]. It should be noted that neither WECC nor NERC identify a specific requirement for regulation. Rather, the requirement is stated in terms of meeting the control performance standards, i.e. standards that refer to load-generation balance to keep the ACE and frequency within acceptable limits. Moreover, the regulating reserve requirement is a balancing authority function in the Northwest Power Pool. The Pool reserve sharing group has been set up to deal with the contingency portion of operating reserve only.

Load Following

In addition to the operating reserve components defined above, power planners must ensure adequate resources to cover the full intra-hour timeframe when determining total capacity requirements. This service is known as load following and is defined as follows:

> **Load Following** – available capacity, possibly but not necessarily, immediately responsive to automatic generation control that is dedicated to meeting anticipated schedules, variable resources and load variations over intervals beyond the regulation timeframe up to the end of the prevailing scheduling term (presently sixty minutes). When assessing the need for load following, allowance for forecast uncertainty as well as anticipated load, variable energy resource following and dynamic transfer changes must be made. Similar to regulating reserve, load following capacity must be adequate to provide movement in both the up and down directions.

Load following service requirements have increased greatly in recent times with the addition of large quantities of variable energy resources. However, the load following service has always required significant amounts of capacity to be available primarily during morning load increases and evening load reductions. It should also be realized that there are nomenclature differences associated with the term we have defined as load following. For example, some balancing authorities consider load following to be a component of their operating reserve. Moreover, in balancing authorities governed by independent system operators with organized real-time markets, load following is sometimes known as supplemental or imbalance energy. The salient point for long term power planners is that in addition to the capacity required for operating reserve, an additional capacity requirement for intra-hour load following must be provided for.

Current Issues

A major concern for the Northwest is the regional and extra-regional misconception that it has spinning reserve capabilities greatly exceeding WECC requirements. Typically the Northwest reports between three and five times the spinning reserve levels required by NERC. This reporting contradicts the very real problem of inadequate flexibility to integrate variable energy resources. A snapshot view of available spinning reserve capacity, as it is defined by NERC, does not consider the energy-limited constraints of hydro capacity. The Northwest generally has large amounts of surplus hydro spinning capacity at any

[3] See NERC Reliability Standard Bal-001 for details on CPS 1 & 2 details (http://www.nerc.com/files/BAL-001-0_1a.pdf)

instant, but this capacity is often only sustainable for time periods on the order of minutes. The definition of available spinning reserves should recognize the energy limitations of the resources being used for this service. Given the Northwest's sixty minute scheduling time period, the operating horizon for spinning reserve should be at least sixty minutes.

III - Assessing Adequacy

A number of methods are used in the electricity industry to ensure system reliability. Some track energy while others focus on capacity. Some are deterministic while others are stochastic, using Monte Carlo simulation techniques. Deterministic methods generally rely on earlier stochastic work and therefore are simpler to define. Three stochastic methods for establishing capacity-based reliability metrics are discussed later in this paper.

Data

Each stochastic method for developing a capacity target requires substantially similar information. The study must collect data on loads, forced and maintenance outages, resource operating characteristics, fuel availability, assumed market reliance (assuming one is not modeling the entire marketplace in the analysis), and transmission availability. It also is important to decide how reserve obligations are treated in the model[4]. The models must be robust enough to represent all of the data in a realistic fashion. Because extreme events happen so infrequently and for differing reasons, durations and magnitudes, it is difficult to adequately represent potential operations based on historical events. For example, in the Northwest it is important to be able to represent accurately the capability of the hydroelectric system, not just its historical performance.

<u>Loads</u>

Load distributions generally are developed using a set of recent years' data. Historical variability is represented by analyzing datasets to determine relationships between key factors such as weather and other trends (e.g., schools days, holidays). Once historical variability has been explained, it can be used to develop a larger dataset of theoretical load profiles. Data may be developed stochastically, or chronologically based on actual historical weather and usage patterns.

Reliability models should consider load forecasting error, either explicitly or implicitly. No load forecast is perfect and can be off in any given hour by a significant amount. Some models assume that forecast error is inherent to the variability underlying the forecast. However, most load models base their forecasts on actual metered sales levels and thereby ignore forecast error. It is reasonable to evaluate a system's forecast error and include it in the load distributions.

A peak-hour integrated demand of 10,000 MW is in actuality the average of many values above and below 10,000 MW. Reliability modeling should use the maximum load within an hour to account for such intra-hour variability as these deviations are not covered by operating reserves.

<u>Forced and Maintenance Outages</u>

Forced and maintenance outages affect system reliability. Forced outages generally are represented in a stochastic model based on actual plant history. Depending on the level of detail desired for the study,

---

[4] Reserve as used in this context is broadly defined to include operating reserve, regulation, load following, and any other product deemed necessary to ensure short-term (within-hour) system reliability.

one can represent forced outages either as a simple de-rating of all plants, a de-rating of plants by technology type, or a more detailed representation using a random outage by plant with a mean time to repair. One can also model outages by time of year to the extent that historical operating performance warrants it.

Maintenance is another key element in a reliability study. It can be represented in a manner similar to forced outage. The entire fleet can be de-rated for maintenance using historical operating procedures or maintenance can be represented in more detail by technology type or at the plant level using specific dates of the scheduled outage.

Depending on when it occurs, reliability events can be driven by maintenance alone. It is therefore useful in a reliability study to develop scenarios that quantify the impacts of the maintenance on system reliability. It is possible that planning maintenance during a shoulder period would make sense from an economic perspective, but when viewed from a reliability perspective taking a large number of generating plants offline at the same time could create a reliability concern. Maintenance might need to be strategically scheduled across more periods during the year to reduce the chances of a system emergency.

Resource Operating Characteristics

The makeup of a system's generation resources can greatly affect system reliability. Each resource category has differing capabilities and limitations that benefit or detriment system reliability. Any reliability study should adequately represent all such capabilities and limitations. Some generic operating characteristics associated with the various resources are detailed below.

Hydroelectric Generation

In the Northwest a large and historically very flexible hydroelectric system has provided the region with a vast surplus of capacity generation. However, given reliance on the water flowing in the river systems, the energy-limited nature of hydroelectric generation resources can create unique deficiencies. Hydro resources fall into three basic categories: run-of-river, storage, and pumped storage.

Run-of-river systems have little to no ability to shape generation over extended periods of time. River inflows are released as they arrive at the project. Some run-of-river projects have short-term (e.g., intra-day) storage, but in most cases these systems' capacities are driven by expected real-time inflows.

Storage systems can hold water in a river back for more valuable periods of the day or year, or push more water downstream, augmenting natural flows into the projects. Capturing the flexibility inherent to these projects is essential to Northwest reliability studies. Any such study should include logic to reasonably reflect the flexibility and the storage limitation of these projects.

Pumped storage systems enable water in a lower reservoir to be pumped up to a higher reservoir during periods of lower loads. When loads increase, water in the upper reservoir can be released to meet them.

Hydroelectric technologies have the capability of providing very fast responses to changing system conditions. Many hydroelectric units can move from zero to full output in a few minutes or seconds. However, many projects have strict operating requirements placed upon them that limit this capability and must be represented in the reliability modeling.

Coal-Fired Generation

Coal-fired generators generally are much less flexible than hydroelectric generation resources, taking hours to days to move from a cold start to full output. In addition, there are delays to move generation from one level to another, and a coal plant cannot be operated across its full range.

Gas-Fired Generation

Gas-fired generators come in varying forms including simple-cycle peaking plants, combined-cycle base load plants, reciprocating engines, and boilers. Each plant has different characteristics and should be accurately represented in a reliability model. For example, simple-cycle peaking and reciprocating plants generally can reach their full output in a few minutes, whereas boilers and combined-cycle plants operate in some ways similar to coal plants and can only be ramped up and down more slowly.

Gas-fired generators capacity is dependent on ambient atmospheric conditions. That is, in colder temperatures output of the plant increases and in hotter temperatures output decreases. A reliability model must reflect the same temperatures that correspond with the demand level used within simulation.

Renewable Energy

Renewable energy resources have varying characteristics depending on their fuel availability and operating characteristics. Wind plants operate when wind conditions are ripe, oftentimes driven by thermal or storm activities. Solar photovoltaic generation is somewhat more predictable though most technologies are unable to provide significant generation except during periods of intense sunlight; no generation is available at night. Solar technology may present integration challenges very different from those being experienced presently in our wind fleet. Variations in solar energy output are much more abrupt when clouds pass overhead than when the wind generation picks up or dies off within an hour. Solar thermal plants may have limited storage that enables generation into the late afternoon and early evening periods and provides steadier output than either wind or solar photovoltaic technologies. Geothermal generation is capable of providing output across most hours of the day, but generally these facilities are not designed to provide significant levels of ramping capability. Biomass has similar operating characteristics to geothermal energy and provides stable and dependable energy as long as the fuel supply is sufficient.

Conservation and Demand-Side Management

Conservation and demand-side management can play a vital role in system reliability. Conservation programs can be modeled based on expected savings. However, the potential for measures like space conditioning must be considered carefully. Some evidence suggests that these programs, while offering savings during most periods, do not cycle off at extreme temperatures and therefore are running at full capacity across the system peak. Well-designed programs joining efficient appliances and improved shell insulation measures typically allow the installed space conditioning capacity (and potential peak load) to be reduced. Some programs, such as lighting programs, are probably unaffected by temperature.

Utilities sometimes have contracts with large commercial and industrial customers whereby these customers obtain a rate reduction in exchange for being interrupted during times of system deficiency. The frequency and duration of these contracts can be input into a robust reliability model to reflect their system benefits.

In recent years utilities have offered more demand-side programs outside of the traditional commercial and industrial classes. Early evidence indicates that during prolonged system events many customers choose to override the control equipment. Utilities must therefore be diligent in monitoring these systems so system reliability benefits are properly reflected.

Market Purchases

Many utilities rely on the wholesale short-term marketplace to serve a portion of their peak demand. Doing so can be an efficient means to meet customer requirements. When market reliance is planned by a utility performing a reliability study, it is a critical requirement to have a full understanding of the load-resource balance in the larger region. Absent this discipline, a region-wide crisis could limit market purchases as many utilities scramble to meet planned deficiencies at the same time.

Long-term purchases must be accurately reflected in a reliability study. For one utility to count an off-system purchase toward meeting its firm loads, the selling utility must reflect the firm obligation in its reliability planning.

Fuel Availability

As with market purchases, resources should be considered available to serve firm loads only to the extent that there is a very strong likelihood that fuel will be available during that period. A firm transportation contract for natural gas might well provide a firm generation resource during the winter months. Similarly a summer-peaking utility might be able to rely on non-firm gas in the summer when natural gas is not in high demand for heating. However, where gas transportation is tight there is the possibility that some gas-fired plants could not procure fuel during a severe cold period. Any resource relying on spot deliveries for fuel should reflect the probability of such supplies not being available during peaking periods.

Treatment of Reserve in Reliability Modeling

Reliability modeling has the potential to assist utility systems in determining the operating reserve margin. Operating reserve, as discussed in Section II, consists of two components—contingency and regulating reserves. In the Northwest contingency reserve is addressed through a reserve sharing program under the auspices of the Northwest Power Pool Reserve Sharing Group. It provides a benefit to the utility by covering the first hour of an outage, and thus in modeling one could omit the first hour of forced outages in reliability assessments. Contingency reserve is also an obligation to the balancing authority (i.e. under the Northwest Power Pool agreement, 5% of hydro plus 5% of wind plus 7% of thermal capacity). A reasonable way to model this obligation would be to reduce the pool of generation assets by the reserve obligations of that agreement.

The regulating and load following portions of reserves require the utility to be able to flex generation up or down within the hour to compensate for fluctuations in load and/or generation. For systems where upward regulation/following is the primary concern, regulating and load following services can be modeled by modifying the load forecast, including those specific obligations in the algorithm, potentially by emulating control area performance criteria in the reliability model, or by reducing the peak generating capability of the pool of resources. For systems with constraints on minimum generation and constraints on the ability to decrease generation, regulation/following reserve can be modeled by increasing the minimum generation levels to provide a buffer for reducing generation in real-time operations.

Reliability Metrics

The main aspects of reliability include the frequency, duration and magnitude of interruptions to electricity service. There are many approaches and metrics for evaluating system reliability throughout the utility industry. The nation-wide trend appears to be the use of probabilistic metrics, which typically use Monte Carlo simulation techniques to simulate future system performance under various uncertain variables (such as load variations and thermal resource availability). Three common probabilistic metrics are described below. The basic ideas and conceptual formulas are provided for each, although many variations are being used by planners throughout the industry.

- **Loss of Load Probability (LOLP)** in units of *percent*, measures the probability that at least one loss of load (or shortfall) event will occur in the time period being evaluated (typically a year). It is calculated as the number of simulations in which a shortfall event occurs divided by the number of simulations. LOLP provides no information regarding duration, quantity, or magnitude of resource shortfalls.
- **Loss of Load Expectation (LOLE)** in units of *hours per year*, measures the amount of time that shortfall events are expected. It is calculated as the total number of hours in which a shortfall occurs in all simulations divided by the total number of years simulated.
- **Expected Unserved Energy (EUE)** in units of *megawatt-hours per year*, measures the expected amount of energy (in megawatt-hours) not served in a given time period. It is calculated by adding up all of the unserved energy for all simulations and dividing by the total number of years simulated. EUE provides some indication of the magnitude of shortfalls but only in aggregate. It does not reflect the frequency, duration or magnitude of individual shortfall events.

Overall, since one single metric does not provide all meaningful information, consideration should be given to all aspects of reliability that customers value. The following provides further details on calculating these reliability metrics[5]. Planners should understand the details of the various metric calculations when discussing and comparing analyses.

Loss of Load Probability

Loss of Load Probability(LOLP) focuses on the probability of the load loss events and is calculated by dividing the number of simulations with at least one shortfall event by the total number of simulations. If each simulation represents a potential future for customers, then LOLP is the probability that customers will face an interruption at some point in the future. The analysis is typically performed with an hourly level of granularity for an entire year. That is, a Monte Carlo study with 3,000 simulations means that the year's operation is simulated 3,000 times for each of the 8,760 hours in the year. In some cases a utility or region may only want to focus on its most critical part of the year, such as winter. In this case, hourly simulations only need to be done over the winter months, and the resulting LOLP is thus a winter reliability measure only.

The definition of a shortfall or event is critical to the analysis and can vary among utilities and regions of the country. Planners may use different thresholds on the size or the duration of the shortfall before it will be considered an event. Whether consecutive hours are treated as individual events or as a single event

[5] The concepts illustrated here are based on "Use of Monte Carlo Simulation In Teaching Generating Capacity Adequacy Assessment," by R Billinton and L. Gan, published in Transactions on Power Systems, Vol. 6, No. 4, November 1991.

for calculating the metric is also an important detail. For example, a continuous eight hour shortfall could be considered as eight events, or just a single event, or not an event at all if the minimum threshold is twenty four hours. Understanding the definition of event is important to comparing various LOLP studies. The following equation is a generalization of how LOLP is calculated:

$$\text{LOLP} = \frac{\sum_{i=1}^{N} \left(\text{Events}_i\right)}{N}$$

Where:

LOLP = Loss of Load Probability (%)

Events = a simulation when load and operating reserve obligations exceed resources based on the definition of event. These simulations could have one or more times within the year that resources are inadequate to meet load and reserve obligations.
N = the number of Monte Carlo simulations for the period, which is typically one year using hourly level of granularity

Loss of Load Expectation

Loss of Load Expectation (LOLE) is based on the same kind of framework as Loss of Load Probability. The difference is that LOLE focuses on the total number of hours that load plus operating reserve obligations exceed resources. In LOLE, the number of hours that events occur is counted, not just a summation of events. It does not perfectly incorporate the concept of duration, but reflects more information than just probability like the LOLP approach. As with the LOLP metric, LOLE can be limited to counting only significant shortfall events, that is, those that exceed minimum capacity, energy and/or duration thresholds. This metric is expressed in hours per year, and may also be referred to as LOLH or HLOLE. It is often expressed in terms of days per year as well.

LOLE is generally calculated as follows:

$$\text{LOLE} = \frac{\sum_{i=1}^{N} \left(\text{Hours Short}_i\right)}{N}$$

Where:

LOLE = Loss of Load Expectation (hours/year)

Hours Short = the number of hours that load plus operating reserves exceed resources (based on the definition of event)
N = the number of years simulated in the Monte Carlo analysis

Expected Unserved Energy[6]

Loss of Load Probability and Loss of Load Expectation are metrics more closely associated with the frequency aspect of reliability. Neither provides any indication of the size or magnitude of potential shortfalls nor of the duration of shortfalls. Expected Unserved Energy (EUE) provides some measure of the magnitude of shortfall events. It is calculated as the sum of all unserved energy (in megawatt-hours) for all hours of the simulation divided by the total number of years simulated. The metric is expressed in megawatt-hours per year. It is also often expressed in megawatt-hours per hour as well.

EUE is calculated as follows:

$$EUE = \frac{\sum_{i=1}^{N} \left(\text{MWh Short}_i\right)}{N}$$

Where:

EUE = Expected Unserved Energy (megawatt-hours/year)

MWh Short = the sum of the amount of energy (megawatt-hours) not served in each hour that load plus operating reserves exceed resources
N = the number of years simulated in the Monte Carlo analysis

Other Metrics

Numerous other metrics can be calculated to address different aspects of reliability. Capacity-based probabilistic metrics can be used to examine magnitude of short-falls on a more single outage basis. Duration and magnitude can also be addressed through different ways of averaging the expected size, duration, and frequency of short-fall events.

[6] Also see “Justification for a NERC Resource Adequacy Assessment Model, A NERC Staff White Paper,” Bob Cummings, Mark Lauby, John Seelke, February 28, 2007, Revised July 31, 2007 pg 3 http://ewh.ieee.org/cmte/pes/rrpa/RRPA_files/2.28.07%20rev%2007.31.07%20Justification%20for%20a%20NERC%20Resource%20Adequacy%20Model.pdf.

IV - Choosing an Adequacy Target

Capacity targets, expressed as a planning margin, can be derived in numerous ways as described above. In many parts of the United States, regional reliability councils have adopted resource adequacy metrics and related planning margins. These metrics help guide decisions about reasonable levels of resource adequacy. The following high-level summary, based on the Westwide Resource Assessment Team Resource Adequacy Briefing Paper (March 23, 2004), was included in PacifiCorp's 2004 Integrated Resource Plan[7]:

(See note below)	WECC	MAPP	SPP	ERCOT	MAIN	ECAR	FRCC	NPCC	SERC	MAAC
Planning Margin	Not Specified	15%, 10% if hydro system Planning Margin	12%, 9% if 75% hydro Planning Margin	12.5% Planning Margin	15-20% Planning Margin	0.1 day/yr LOLE	15% Planning Margin	Not specified	Varies by member system	Based on LOLE criterion
Regional Resource Adequacy Criteria	Not Specified	1-in-10 yr LOLE	1-in-10 yr LOLP	Not specified	1-in-10 yr LOLP	Use of supplemental capacity for 1-10 d/yr (DSCR)	1-in-10 yr LOLP	LOLE by disconnecting firm load due to resource deficiency no more than 0.1 d/yr	No uniform criterion for entire region	LOLE of 1 day/10 years or 0.1 day/yr
Methodology	Not Specified	Margin derived using LOLE 1-in-10	LOLE analysis	Reserve Margin based on LOLE studies	Based on LOLP and LOLE studies	1 to 10 days DSCR consistent with LOLE 1 in 10 yr	Periodic analysis of LOLP for reserve margin	Based on LOLP and LOLE studies	Not specified	PJM approval of the required margin

More recently, in August of 2008, NERC provided a more comprehensive summary of regional resource adequacy analysis[8].

One must be cautious when interpreting these standards, because the details are important. Different utilities may have slightly different interpretations or use different analytical assumptions in applying these terms within or across regions. Such slight differences in interpretation can result in vastly different capacity targets and thus different costs. NERC in explaining LOLP in the Reliability Assessment Guide[9], acknowledges this challenge by using words such as "typically" and "in some areas" in their definition. They say:

> "a. **LOLP:** Loss of Load Probability (LOLP) is the building block of probabilistic analyses. LOLP is typically defined as the probability of firm load demand not being met in any given time period.[59] In some areas, the determination is whether firm load demand plus operating reserves, or a portion thereof, can be met in a given time period.[60] When the probabilities of events are summed over time, the result is an expectation." *(footnotes not provided)*

---

[7] See p. 198 of the Technical Appendix to PacifiCorp's 2004 Integrated Resource Plan at http://www.pacificorp.com/content/dam/pacificorp/doc/Environment/Environmental_Concerns/Integrated_Resource_Planning_4.pdf

[8] See http://www.nerc.com/docs/pc/ris/RRO_Adequacy_Assessment_Practices-Survey_Responses_08_14_08.xls

[9] See p. 122 of NERC's Reliability Assessment Guidebook, Version 1.2, April 7, 2009, at http://www.nerc.com/files/Reliability_Assessment_Guidebook-08-24-09-clean.pdf

The Northwest Resource Adequacy Forum was created in 2005 by the Northwest Power and Conservation Council and the Bonneville Power Administration. Its purpose is to develop a consensus-based resource adequacy framework for the region to provide early warning to regional planners should resource development fall short of the region's preference for electric system reliability. In 2008, the Council adopted the Forum's proposed resource adequacy standard, based on a five percent Loss of Load Probability reliability target[10].

It should be noted that the Forum's standard and corresponding minimum thresholds for resource development represent the bare minimum level of reliability that the region will tolerate. They do not represent, nor are they intended to represent, resource planning targets. Those planning targets must take into account each utility's individual situation, including its access to market resources and tolerance for economic risk. Each utility chooses its own individual level of reliability by considering tradeoffs between system reliability and cost.

Some utilities have performed an explicit benefit-cost analysis to identify economically-efficient reliability levels. This entails estimating the incremental cost of improving reliability and comparing it with its incremental value. The goal of such analysis is to find a reliability level where the incremental benefit of reliability equals its incremental cost. An example of this methodology is found in PacifiCorp's 2004 Integrated Resource Plan[11]. Puget Sound Energy's 2005 resource plan included a similar analysis to establish its natural gas planning standard[12].

Estimating the incremental cost of improving reliability is a reasonably straight-forward resource planning analysis. It uses a resource planning analytical framework to minimize the costs of raising the planning standard. Estimating the incremental "value" of increased reliability is much more conceptually challenging. Customers value reliability, but defining that value is not simple. One method is to estimate customer value by using market research for different customer segments. This approach was used in setting Puget Sound Energy's gas planning standard, referenced above. The drawback to this approach is that energy reliability has the attributes of public goods. An example of such public good is traffic lights. Society as a whole may benefit by more than just the sum of individual utility customers. Economic research methods provide another approach to value public good benefits. This approach was used in PacifiCorp's 2004 resource plan.

A carefully designed demand-response or demand buy-back program may reduce the level of generation capacity needed to meet the planning margin. If demand response programs were designed to target specific changes in reliability, and pricing was designed to reflect the incremental cost savings, individual customers would be able to fine-tune their own reliability levels. This would likely require sophisticated metering, billing technologies and contract structures that each utility may or may not possess or desire to develop.

---

[10] The full Resource Adequacy Standard document is available at http://www.nwcouncil.org/library/2008/2008-07.pdf

[11] See pp. 191-223 of the Technical Appendix to PacifiCorp's 2004 Integrated Resource Plan at http://www.pacificorp.com/content/dam/pacificorp/doc/Environment/Environmental_Concerns/Integrated_Resource_Planning_4.pdf

[12] See Appendix I of Puget Sound Energy's 2005 Least Cost Plan at http://www.wutc.wa.gov/rms2.nsf/177d98baa5918c7388256a550064a61e/dcd6a571d635709d88256ff6005be2ad!OpenDocument

Utilities should explicitly or implicitly choose a capacity planning target through their resource planning process. Being aware of the concept of reliability and planning margin will serve to inform their future resource decisions.

V - Integrating Operating Reserve and Planning Margin

This section illustrates one approach to apply the reserve concepts outlined in this paper. There are three steps in the analytical process where the operating reserve obligations are reflected.

1. **Meet Reliability Target** – Use a loss of load probability analysis to meet a target level of reliability.
2. **Calculate Planning Margin** – Convert the results of the LOLP to a metric that can be used in a resource portfolio analysis.
3. **Apply Planning Margin** – Use planning margin metric in analyzing different resource portfolios.

Meet Reliability Target

In Step 1, a probabilistic approach is used to estimate the required amount of capacity needed to ensure a certain level of reliability–measured as a loss of load probability. Section III provides details on the data and calculation of LOLP. While many different metrics can be used in this step, in this discussion it is presumed that LOLP is used.

One approach to reflect reserves is to define "load" to include its operating reserve obligations. In other words, an "event" would reflect the need to have sufficient resources to meet sales, line losses and operating reserve requirements.

In addition, the <u>use</u> of contingency reserve should be reflected in the analysis. All parties to the reserve sharing program in the Northwest Power Pool have generation available to them for up to sixty minutes in the event of a generator forced outage. This one hour loss of capacity due to a forced outage does not count toward an event for the first hour—but it does count in events lasting two or more hours. In simulated events where loads exceed generation and there are no forced outages, the first hour counts as a loss of load.

Calculate Planning Margin

In Step 2, the planning margin is calculated from the results of the loss of load probability analysis into a metric useful for long-term resource portfolio analysis. The planning margin is expressed as the ratio of the amount of generating capacity in excess of normal-weather peak load to the load.

The following table illustrates these calculations. The LOLP results reflect changing probabilities resulting from adding capacity to the existing resource portfolio and recalculating the LOLP with the additional capacity. In this example an existing system of 5,260 MW has a loss of load probability of 56%, a level much higher than the target (in this example 5%). Additional capacity is added systematically until the loss of load probability is reduced to 5%—an additional 1,125 MW of capacity are needed. This completes Step 1.

LOLP Results				Planning Margin Calculations				
Existing Resource Capacity (MW)	Additional Capacity (MW)	Total Capacity (MW)	Resulting LOLP	Required Operating Reserve (MW)	Capacity Net of Op Reserve (MW)	Normal Peak Load (MW)	Planning Margin Net of Op Reserve	Planning Margin
A	b	c=a+b	D	e	f=c-e	g	h=(f/g)-1	i=(c/g)-1
5,260	-	5,260	56%	250	5,010	5,236	-4.3%	0.5%
5,260	150	5,410	38%	261	5,149	5,236	-1.7%	3.3%
5,260	300	5,560	24%	271	5,289	5,236	1.0%	6.2%
5,260	450	5,710	17%	282	5,428	5,236	3.7%	9.1%
5,260	600	5,860	14%	292	5,568	5,236	6.3%	11.9%
5,260	750	6,010	11%	303	5,707	5,236	9.0%	14.8%
5,260	900	6,160	9%	313	5,847	5,236	11.7%	17.6%
5,260	1,050	6,310	6%	324	5,986	5,236	14.3%	20.5%
5,260	1,125	6,385	5%	329	6,056	5,236	15.7%	21.9%
5,260	1,200	6,460	4%	334	6,126	5,236	17.0%	23.4%

In Step 2 the results of the analysis are converted to a simpler portfolio analysis metric. Using the example in the table, two forms of planning margin are calculated (columns h and i), one net of and another including operating reserves.

The **planning margin net of operating reserves** is calculated as:

Planning Margin Net of Operating Reserves =

(Total Capacity – Required Operating Reserves – Normal Peak Load) / Normal Peak Load

This analysis indicates a planning margin of 15.7% plus operating reserve are needed to meet the 5% loss of load probability target in the example study above.

Another approach is to reflect the **planning margin** including operating reserve. This is calculated as

Planning Margin = (Total Capacity - Normal Peak Load) / Normal Peak Load

In this same example, the planning margin is calculated as 21.9% and includes the required operating reserve.

Applying Planning Margin for Long-Term Resource Planning

Using either the Planning Margin value or the Planning Margin Net of Operating Reserves produces similar results in the resource portfolio analysis, when operating reserve is accounted for properly. Using Planning Margin (without adjusting for operating reserves) in effects locks the ratio of required operating reserves to peak load. When a utility resource mix changes over time, which is highly likely, then the embedded operating reserve amount will not properly account for the operating reserve. For example, if a utility's thermal plants, where seven percent are held for contingency reserve, are replaced with market purchases, the contingency reserve would be zero because the 3^{rd} party utility provides the reserve. Without properly accounting for the change in operating reserve, the capacity requirements would be overstated.

The following table provides a summary of the differences between planning margin with and without adjustment for operating reserve using the values in the example study.

	Planning Margin Net of Operating Reserve	**Planning Margin**
Results from Example	15.7%	21.9%
Operating Reserve	Deducted from Capacity Target to derive Planning Margin	Embedded in Calculation
Application in Planning Analysis	Reflect Operating Reserve obligations in planning analysis.	Do not reflect operating reserve obligation separately. Be aware of changes in resource stack that may affect operating reserve significantly.
Peak Load	5,236 MW	5,236 MW
Capacity for Margin	6,056 MW	6,385 MW
Operating Reserve	329 MW	Embedded
Total Capacity	6,385 MW	6,385 MW