Study Committee C5

Electricity Markets & Regulation

Paper 10240

BEYOND EXPECTED VALUES: EVOLVING METRICS FOR RESOURCE ADEQUACY ASSESSMENT

Derek STENCLIK[1], Aaron BLOOM[2], Gord STEPHEN[3,6], Wesley COLE[3], Armando FIGUEROA ACEVEDO[4], Aidan TUOHY[5], Genevieve DE MIJOLLA[5]

[1]Telos Energy, [2]Energy Systems Integration Group, [3]National Renewable Energy Laboratory, [4]Midcontinent Independent System Operator, [5]Electric Power Research Institute, [6]University of Washington

Motivation

As the power system's resource mix changes, resource adequacy metrics need to transform as well. The conventional resource adequacy metric, loss of load expectation (LOLE), quantifies the expected amount of time when capacity is insufficient to meet load in a given year. LOLE is an opaque metric when used in isolation. It only provides a measure of the average amount of shortfalls over a study period and does not characterize the magnitude or duration of specific outage events.

This is true for loss of load hours (LOLH) and loss of load events (LOLEv) as well. In addition, these metrics are expected values and simply quantify an average risk across all of the replications evaluated and provide no information on the distribution of shortfall events or insight into extreme, or tail events.

Four Steps Forward for Better Shortfall Risk Metrics

1. Place greater emphasis on normalized unserved energy metrics
2. Report a suite of metrics
3. Study full outcome distributions and quantify tail risks
4. Examine the nature of individual shortfall events

The communications challenge: maintaining simplicity without sacrificing rigor

All of these steps forward for resource adequacy metrics will require expanding, or at least adapting, the quantitative results that must be communicated to non-technical regulators and other stakeholders. This presents a distinct challenge and likely constrains the degree of technical complexity that is possible when attempting to describe the increasingly complicated nature of power system adequacy.

Comparing risk metrics to other fields

Power system engineering is clearly not the only domain concerned with risk and reliability. While the power system engineering field has a well-established literature on this topic, many other disciplines have been undertaking their own parallel efforts to address these issues and it would be wise to learn from these domains, particularly as we enter into a period of power system transformation where traditional approaches and heuristics may no longer be effective.

System Uptime

Many engineering domains are able to express system reliability in terms of a single percentage value, representing a measure such as 'uptime' or 'availability'. This measure typically represents a fraction of time which is not sufficient to fully capture the nature of adequacy risk in power systems. However, this general approach provides an intuitive and widely recognized measure of reliability which may be useful to the resource adequacy community. Percentages are self-contextualizing, can be readily compared across different systems, and have established "class of nines" logarithmic language available for accessible communication of the exponentially increasing costs of reliability (e.g. "five nines" for 99.999% and "six nines" for 99.9999%)

Coherent Financial Risk Metrics

The issue of describing risk tolerance in terms of tail events is so important in quantitative finance that a body of literature has developed around so-called 'coherent' risk metrics. Some requirements for coherence that are relevant to power system adequacy metrics include:

Monotonicity: If, in every scenario considered, one resource portfolio performs better (i.e. more reliably) than another, the risk metric assigned to the first portfolio should be lower than the second. Traditional resource adequacy metrics generally satisfy this property.

Sub-additivity: Diversifying a portfolio should never increase risk. In financial applications, this means that if several investments each satisfy some risk criteria individually, a portfolio combining those investments must as well. Unserved energy outcomes (i.e., EUE) satisfy this criterion, but event and event-period counts (i.e. LOLEv and LOLE) do not.

This research was partially supported by the Energy Systems Integration Group, through the Redefining Resource Adequacy Task Force.

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Strategic Analysis. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

Study Committee C5
Electricity Markets & Regulation
Paper 10240

BEYOND EXPECTED VALUES: EVOLVING METRICS FOR RESOURCE ADEQUACY ASSESSMENT

Derek STENCLIK[1], Aaron BLOOM[2], Gord STEPHEN[3,6], Wesley COLE[3], Armando FIGUEROA ACEVEDO[4], Aidan TUOHY[5], Genevieve DE MIJOLLA[5]

[1]Telos Energy, [2]Energy Systems Integration Group, [3]National Renewable Energy Laboratory, [4]Midcontinent Independent System Operator, [5]Electric Power Research Institute, [6]University of Washington

Four Steps Forward for Better Shortfall Risk Metrics

#1: Place greater emphasis on normalized unserved energy metrics

If system adequacy information must be reduced to a single comprehensive indicator, unserved energy metrics such as EUE may be the better choice, as such metrics combine the frequency, duration, and magnitude of shortfall events into a single value.

While EUE expressed in natural energy units (e.g., MWh) can complicate comparisons across different systems, or within a single system at different times, the measure can easily be normalized by the total energy demand of the system (in the region or time period of interest). This normalized expected unserved energy (NEUE) value can be expressed as a percentage or parts-per-million (ppm). For example, Australia's National Electricity Market uses a 0.002% (20 ppm) NEUE standard

In many jurisdictions, transitioning to an energy-based primary adequacy criterion will take time. The industry has decades of experience and familiarity with LOLE. Fortunately, EUE and LOLE are currently well correlated in most regions. It is not until variable renewable energy and energy-limited resources become an ascendant source of a system's capacity that the relationship between EUE and LOLE may begin to misalign and cause issues. A first step towards greater emphasis on unserved energy for a system currently focused on event-period performance may be to simply calculate and report both types of metrics.

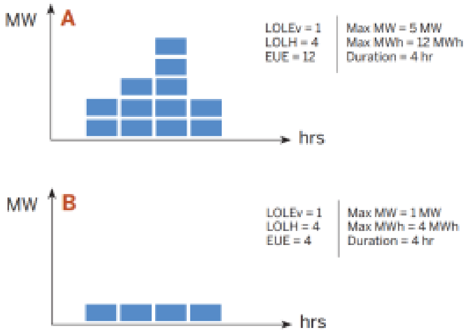
#2: Report a suite of metrics

While unserved energy metrics may be better positioned to capture many dimensions of system performance, they still ultimately only report a single number. Overreliance on a single adequacy indicator has the potential to obfuscate elements of system risk that may be of interest to stakeholders (see Examples 1 and 2 below).

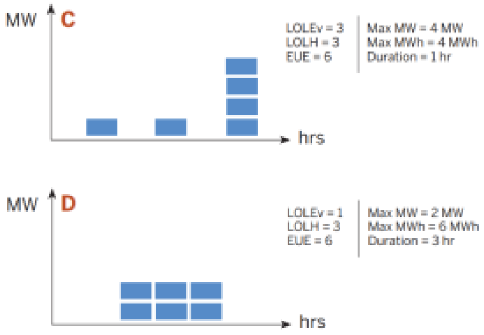
Consider the rolling blackouts experienced in California during August 2020 and those experienced in Texas in February 2021. Both were resource adequacy events but can be characterized very differently depending on the metric used. On a days-of-shortfall basis (LOLE), the California and Texas events are somewhat similar, but when using LOLH or EUE, there is a clear difference, illustrating how overreliance on a single metric can skew the characterization of an event and have serious implications for decision-making around mitigations.

Event Property	Risk Metric	Event		Difference
		California, Aug 2020	Texas, Feb 2021	
Number of days	LOLE	2 days	4 days	+200%
Number of events	LOLEv	2 events	1 event	-50%
Number of hours	LOLH	6 hours	71 hours	+1,200%
Unserved energy	EUE	2.7 GWh	990 GWh	+36,700%
Maximum shortfall		1,072 MW	20,000+ MW	+1,766%

Example 1— Same LOLEv and LOLH, but very different events



Example 2— Same LOLH and EUE, but very different events



Study Committee C5

Electricity Markets & Regulation

Paper 10240

BEYOND EXPECTED VALUES: EVOLVING METRICS FOR RESOURCE ADEQUACY ASSESSMENT

Derek STENCLIK[1], Aaron BLOOM[2], Gord STEPHEN[3,6], Wesley COLE[3], Armando FIGUEROA ACEVEDO[4], Aidan TUOHY[5], Genevieve DE MIJOLLA[5]

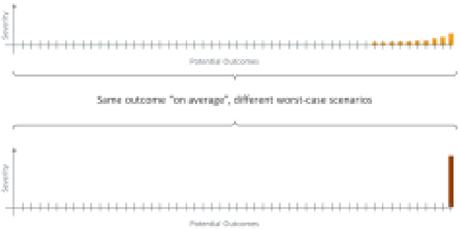
[1]Telos Energy, [2]Energy Systems Integration Group, [3]National Renewable Energy Laboratory, [4]Midcontinent Independent System Operator, [5]Electric Power Research Institute, [6]University of Washington

Four Steps Forward for Better Shortfall Risk Metrics

#3: Study full outcome distributions and quantify tail risks

One limitation of all conventional resource adequacy risk metrics is that they are expected values. While resource adequacy analysis may consider a wide range of probabilistic outcomes, these results are usually averaged and reported as a single value. This can mask the potential for unacceptable system performance under outlier tail events.

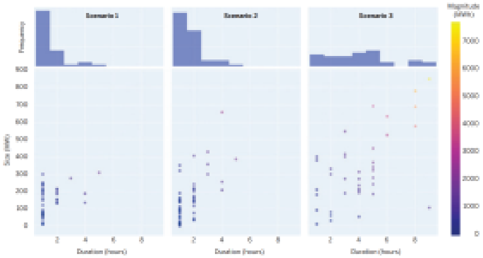
For example, Belgium uses a two-part reliability criterion based on event-hours, which tests both the average count (LOLE of at most 3 event-hours/year) as well as the 95th percentile count ('LOLE95' of at most 20 event-hours/year) and is intended to ensure that the system has enough resources to perform acceptably even in a statistically abnormal (1-in-20) year. Both the average and 95th percentile criteria need to be satisfied for compliance with the regulators' requirements.



Resource adequacy analysis is analogous to an insurance policy in that we voluntarily incur an expense that only delivers a service in rare, extreme circumstances. Planners, regulators, and ratepayers are accustomed to paying a premium for surplus capacity resources that are available to operate only on rare occasions. However, insurance is not intended to cover 'average' operations, but rather worst-case scenarios. It seems reasonable to suggest that resource adequacy analysis should consider performance under these kinds of scenarios as well.

#4: Examine the nature of individual shortfall events

While aggregate risk metrics are useful for understanding the overall adequacy of a power system, they provide very little information about the nature of individual events and the kinds of mitigation strategies that may be appropriate if system risk is deemed too high. Characterizing the size, frequency, duration, and timing of discrete shortfall events can provide this information, better informing planning activities or capacity market design.



Scatter plots and histograms of size, frequency, and duration of shortfall events under three LOLE-equivalent scenarios

Equally important to understanding the size, frequency, and duration of resource adequacy shortfall events is additional understanding into the timing of shortfalls. The chart below shows that the periods of risk can change as the system resource mix changes. It also highlights that mitigation strategies can be narrowly targeted to focus on specific hours and seasons, which is especially important when comparing load flexibility and variable renewable energy resources against other alternatives.