

Clarifying the Interpretation and Use of the LOLE Resource Adequacy Metric

Gord Stephen, Simon H. Tindemans, John Fazio, Chris Dent, Armando Figueroa Acevedo, Bagen Bagen, Alex Crawford, Andreas Klaube, Douglas Logan, Daniel Burke

on behalf of the IEEE Resource Adequacy Working Group (RAWG)

Abstract—The loss-of-load expectation (LOLE) risk metric has been used in probabilistic power system resource adequacy assessment for over 70 years, and today is one of the most recognizable and widely-used measures of system shortfall risk. However, this wide adoption has been accompanied by ambiguities and inconsistencies in its definition and application. This paper provides a unifying reference for defining the metric as it relates to modern analyses, while clarifying a number of common points of confusion in its application. In particular, the paper clarifies that LOLE is not a measure of expected total shortfall duration, that a 2.4 hours per year LOLE target implies a less reliable system than a 1 day in 10 years (0.1 days per year) LOLE target, and that exact conversions between hourly and daily LOLE targets are not generally possible. Illustrative examples are provided to help explain each of these points.

Index Terms—resource adequacy, probabilistic assessment, risk metrics, LOLE, power system planning.

I. INTRODUCTION

The loss of load expectation (LOLE) metric has a long history in power system reliability assessment, and today represents the most common basis for quantifying the resource adequacy of a power system [1]. As adequacy assessment methods have evolved over time, LOLE’s definition and use has been adapted in different and sometimes mathematically inconsistent ways, often leading to confusion and inconsistent applications. While these issues are not generally new, they remain prevalent in industry practice, particularly as they relate to the interpretation of the common “1 day in 10 years” rule of thumb for acceptable shortfall risk as applied in North America.

To help address this problem, the IEEE PES Resource Adequacy Working Group (RAWG, operating under the Reliability, Risk, and Probability Applications Subcommittee of the Analytical Methods for Power Systems Committee) is seeking (through this short note) to clarify the mathematically-correct interpretation of LOLE relative to other related metrics and terms. It should be noted that the RAWG also recognizes the importance of using multiple different metrics to understand system adequacy [2], [3]. For example, while NERC publishes LOLE results (in terms of hours per year) in its biannual reports, it provides expected unserved energy results as well [4]. The intention of this paper is strictly to clarify the best-practice definition and interpretation of LOLE given its wide prevalence today, and not to advocate for or against

its use in any specific future application. Furthermore, this paper only addresses practices for interpreting and communicating probabilistic outcomes: methods and best practices for modeling uncertainty (whether from thermal or variable renewable generation availability, demand, or other factors) are independent of this discussion.

Section II of this paper provides brief historical context for the development of LOLE as a risk metric and the emergence of the “1 day in 10 years” rule of thumb that has been applied in North America. Section III outlines the RAWG’s recommended interpretation of LOLE and related metrics, and Section IV reiterates how commonly-used conversions between those metrics are both mathematically inaccurate and systematically biased. Section V provides concluding remarks.

II. HISTORICAL BACKGROUND

Probabilistic resource adequacy assessment had been discussed as early as the 1930s [5], and in 1947 Calabrese published one of the first papers [6] dealing with the metric that would come to be known as LOLE. That work and many that followed it calculated the expected (mean) count of the number of days in which daily maximum load would exceed available capacity. This mean would be computed with respect to a probabilistic model of generation and demand. Unfortunately, the meaning of that result is described somewhat loosely throughout the paper as “loss of load duration”, “fraction of time during which loss of load may be expected to occur” and “expected total number of days of loss of load” which, while perhaps justifiable in the absence of higher resolution data, would ultimately lead to confusion with the advent of hourly adequacy analysis in recent decades.

Billinton and Chu [1] chronicle the discussion and adoption of this “average count of shortfall days” metric in industry, and the eventual coalescence around 1 day in 10 years (0.1 days per year) as an acceptable level of risk through the 1960s. In that discussion the authors also note that earlier work emphasized that adequacy criteria should be determined based on the operator’s risk tolerance and an appropriate balance between the costs and benefits of avoiding unserved energy, while in later years this target tended to be taken as a given without significant justification.

More recently, improved computational capabilities and increasing penetrations of variable generating resources (which

can shift periods of system shortfall risk away from peak load hours) have motivated adequacy assessments at higher temporal resolution than the historical daily peak analysis, allowing for the quantification of the expected count of hours (rather than days) experiencing shortfall. A common adaptation of the historical “1 day in 10 years” criterion to hourly assessments has been to interpret it as “24 hours in 10 years”, based on the incorrect but understandable premise (given much of the language used to describe the metric) that the original criteria referred to a full day’s duration of shortfall. Section IV will explain why this interpretation is best considered as inaccurate.

In contrast, a number of European countries [7] use LOLE metrics that are directly expressed in hours per year. In a recent decision [8], the EU Agency for the Cooperation of Energy Regulators prescribes a reliability standard in terms of LOLE (in hours per year) that is based on the value of lost load and cost of new entry of resources that contribute to system adequacy. Such an argument is loosely based on the interpretation of LOLE as the expected fraction of time during which loss of load occurs.

III. LOLE AND RELATED DEFINITIONS

In the following, we attempt to provide a definition of LOLE that is compatible with current usage — to the extent that this is possible — and avoids the common pitfall that results from the adaptation of units in resource adequacy computations. For the purposes of this paper, the following definitions will be used:

- A “horizon” is the period of time over which a resource adequacy risk is reported.
- An “event-period” is a general period of time during which, at some point, system resources are insufficient to meet all demand¹. By definition, the duration of a simulated event-period must equal or exceed the length of simulation timestep used in the study.
- An “event-hour” is an event-period lasting one hour².
- An “event-day” is an event-period lasting one day (during which at least one event-hour occurs).
- An “event-year” is an event-period lasting one year (during which at least one event-day occurs).
- An “adequacy event” (event) is a set of event-periods that are contiguous at the highest available level of temporal resolution.

Using these definitions, we define LOLE as a *counting measure*: the expected count of event-periods per horizon (e.g. 1 event-day per 10 years, 3 event-hours per year, 1 event-year per 20 years). This definition emphasizes that one cannot simply perform a unit change on the event periods

¹A discussion of what defines “insufficient resources” is reserved for another paper. It obviously can mean an actual loss of service but can also mean the implementation of emergency actions, which could be the use of expensive backup resources or contractual buy-back provisions for certain customers. It could also be interpreted as periods when market electricity prices soar to extreme levels.

²Resource adequacy assessment is not typically performed at greater than hourly resolution, but if it were to be, one event-hour would contain at least one event-minute, etc

(e.g. by equating 1 event-day with 24 event-hours), thus avoiding the pitfalls discussed in Section IV. Note that it is the responsibility of study authors to report the event-period and horizon used for the LOLE definition.

In addition, one can define metrics for special cases of LOLE that make the choice of event-periods more explicit:

- Loss of load hours (LOLH), the expected count of event-hours per horizon (e.g. 3 event-hours per year).
- Loss of load days (LOLD), the expected count of event-days per horizon (e.g. 1 event-day per 10 years)
- Loss of load years (LOLY), the expected count of event-years per horizon (e.g. 1 event-year per 20 years)

As mentioned, in North American studies, “LOLE” is often implied to be LOLD, with results in terms of event-hours labelled as LOLH instead [9]. In European studies, “LOLE” typically implies the LOLH criterion [8]. The LOLY criterion is used by the Northwest Power and Conservation Council [10]³.

The RAWG does not consider one of these definitions to be more “correct” than any other. LOLE can be reported in terms of any kind of event-period (and horizon), as long as those units are clearly specified.

Finally, we distinguish LOLE from similar metrics:

- LOLE is *not* a measure of total shortfall duration, as shortfalls may be shorter than the event-periods in which they occur.
- LOLE does *not* count the number of adequacy events (as defined above). For this, one should calculate the loss of load events (LOLEV, sometimes called loss of load frequency, LOLF), the expected count of adequacy events per horizon (e.g. 1 event per 10 years, which is not the same as 1 event-day per 10 years – see Figure 1).
- LOLE does *not* directly express a probability an event takes place. The link between LOLE and event probability is further elaborated below.

A. LOLE vs LOLP

The loss of load probability (LOLP) for a period is the probability that at least one shortfall occurs during this period (i.e. the probability that this period is an event-period). While LOLE and LOLP are conceptually distinct, if the same definition of periods is used for both metrics, it is possible to calculate one in terms of the other. Because LOLP can be interpreted as an identically-valued LOLE in units of event-periods per period, the sum of LOLP over all periods in a horizon yields the expected total count of event-periods within that horizon – the horizon LOLE.⁴ For example, if an LOLP

³The NWCC uses annual loss of load probability as its adequacy metric, which in its application is equivalent to loss of load years.

⁴Conversely, an LOLE can be divided by the total number of periods within the relevant horizon to obtain an average LOLP (also known as “snapshot LOLP”) for the periods. For example, if an LOLE in event-days (LOLD) has been calculated for some horizon, dividing the LOLE value by the number of days in the horizon yields the average daily LOLP across all days. Note that this is simply an average value, and does *not* in general correspond to the actual LOLP of any individual day, nor to the LOLP of the full horizon.

is calculated for each individual hour in a year, the sum of those values yields the LOLE (LOLH) for the year.

Like LOLE, LOLP may be defined with respect to any time period, and the choice and interpretation of that period should be explicitly stated when reporting LOLP metrics.

IV. CONVERTING BETWEEN TIME SCALES

A common source of confusion is how an LOLE adequacy criterion in terms of event-days per year is translated into a criterion for event-hours per year. For example, a common conversion equates the “1 day in 10 years” shortfall threshold to “24 hours in 10 years” or “2.4 hours in 1 year”. While it is not inherently wrong to use 24 event-hours per 10 years for an adequacy criterion, doing so implies a less reliable system than targeting 1 event-day per 10 years.

To see why, consider a system satisfying the original 1 event-day criterion. Over an “average” ten-year span, the system would experience shortfall during at most one day, in at least one event spanning at least one event-hour and at most 24 event-hours. Now, consider a system satisfying the supposedly “equivalent” 24 event-hour criteria: over the same “average” ten years, the system could experience up to 24 different shortfalls over 24 different days, a massive violation of the original 1 event-day requirement. In the best-case scenario (in terms of meeting the original criteria), the 24 hours of shortfall all happen in a single day, matching the worst-case reliability associated with the original scenario. The two criteria are only equivalent under the worst-case scenario still satisfying the 1 event-day threshold *and* the best-case scenario matching the 24 event-hour threshold – in any other situation, the 24 event-hour system is less reliable. More broadly, exact conversions between criteria involving different types of event-periods and event counts are not generally possible without making certain assumptions or empirical estimates, because the quantity and distribution of events across the different timescales are ambiguous.

Figure 1 provides some examples of these potential ambiguities. In case A, a single adequacy event occurs that lasts less than an hour in “true” (continuous) time (green) and is also fully contained within a single discretized hourly timestep (red) and discretized daily timestep (blue). The single event could therefore be described as lasting less than an hour in duration, while also being reported as both one event-hour and one event-day of shortfall. In case B, two shortfall events occur within a few hours of each other. Their total “true” duration may only be three hours, but they span four discrete hourly timesteps and so correspond to four event-hours. Since the events occur within the same discrete daily timestep (calendar day), the two events would also be reported as a single event-day of shortfall. Case C depicts a single event that also spans four hourly timesteps, and so would correspond to four event-hours. However, since the event spans two calendar days, it would also be reported as two event-days of shortfall. Case D describes a single 24-hour event that occurs within a single calendar day, and so would be reported as 24 event-hours and one event-day. Case E describes a series of 24 events, each

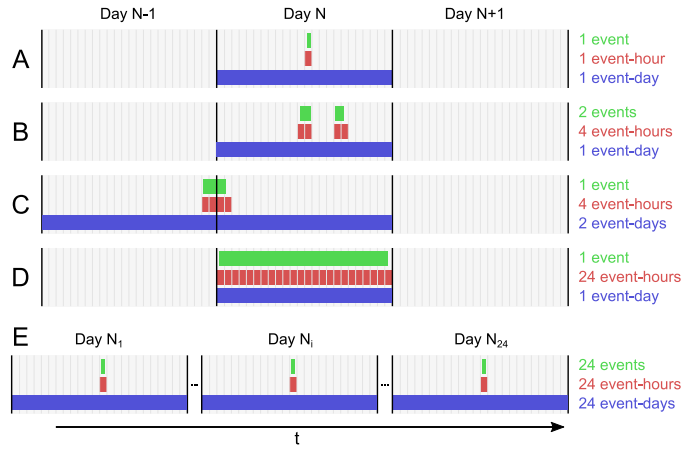


Fig. 1. Examples of adequacy event sets with different event, event-hour, and event-day counts.

occurring within a single hour on separate days, such that together they correspond to both 24 event-hours and 24 event-days.

Additional examples of why an event-day criteria cannot be translated directly into an event-hour criteria are provided in [11] and [12].⁵ Empirical examples of varying event-day / event-hour relationships across different systems are available in [13].

V. CONCLUSION

In spite of (or perhaps because of) its wide application across industry, the LOLE resource adequacy risk metric has been interpreted and used in many different and at times inconsistent ways over its 75-year history, with some ambiguity persisting to this day. With this letter, the IEEE PES RAWG has proposed a clear, consistent, and transparent recommendation to industry regarding this sometimes-confusing metric.

In particular, we hope to remind practitioners that a 1-day-in-10-year adequacy requirement for LOLE (LOLD) is not equivalent to a 2.4 event-hours per year requirement for LOLE (LOLH), and that a system satisfying the latter will be in general less reliable than a system satisfying the former. We recommend expressing LOLE results in terms of expected counts of “event-periods” (event-hours, event-days, etc) per horizon in order to avoid the common misconception that LOLE and related metrics provide a measure of expected total shortfall duration – they do not.

⁵In general, this conflation becomes less problematic when simulating at higher temporal resolutions. For example, in Great Britain, the security standard is specified as a LOLE of 3 event-hours per year, but studies are performed using half-hours. We recommend against this, as it results in ambiguity – does this mean 3 event-hours per year, or 6 event-half-hours per year? We thus recommend that the metric is reported in terms of count of periods based on the time resolution of the calculation, or at minimum that the time resolution of the calculation is reported clearly. However, taking a pragmatic perspective, the difference is expected to be relatively small when typical shortfall events are longer than one hour – this is in contrast to regarding 1d/10y as equivalent to 2.4h/y, which is both wrong in principle and always makes a big difference in practice.

In closing, we note that the 1-day-in-10-years criterion is arbitrary, and that appropriate adequacy criteria may vary across different systems. While there is nothing inherently wrong with targeting a 2.4 event-hours per year criterion, that threshold should not be used solely as an attempt to recreate other historical standards.

ACKNOWLEDGMENT

The authors would like to thank the IEEE RAWG membership for their direction and review in preparing this work, with particular thanks to Murty Bhavaraju, Michael Milligan, Gene Preston, and Derek Stenclik. The authors also appreciate the additional feedback provided by attendees of the 2021 NERC Probabilistic Analysis Forum and four anonymous reviewers.

REFERENCES

- [1] R. Billinton and K. Chu, "Early Evolution of LOLP: Evaluating Generating Capacity Requirements [History]," in *IEEE Power and Energy Magazine*, vol. 13, no. 4, pp. 88-98, July-Aug. 2015, doi: 10.1109/MPE.2015.2417475.
- [2] S. Zachary, A. Wilson, and C. Dent, "The integration of variable generation and storage into electricity capacity markets", arXiv preprint, 2019, arXiv:1907.05973.
- [3] J. Fazio, D. Hua, "Three probabilistic metrics for adequacy assessment of the Pacific Northwest power system," in *Electric Power Systems Research*, vol. 174, pp. 105858, 2019, doi: 10.1016/j.epsr.2019.04.036.
- [4] "2020 NERC Long Term Reliability Assessment," North American Electric Reliability Corporation, 2019. https://www.nerc.com/pa/RAPA/ra/Reliability%20Assessments%20DL/NERC_LTRA_2020.pdf
- [5] R. Billinton, "Bibliography on the Application of Probability Methods In Power System Reliability Evaluation," in *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-91, no. 2, pp. 649-660, March 1972, doi: 10.1109/TPAS.1972.293251.
- [6] G. Calabrese, "Generating Reserve Capacity Determined by the Probability Method," in *Transactions of the American Institute of Electrical Engineers*, vol. 66, no. 1, pp. 1439-1450, Jan. 1947, doi: 10.1109/T-AIEE.1947.5059596.
- [7] "Annex C: Reliability Standard Methodology," Department of Energy and Climate Change, United Kingdom, 13D/190, July 2013.
- [8] "ACER Decision 23-2020 on the Methodology for Calculating the Value of Lost Load, the Cost of New Entry, and the Reliability Standard," EU Agency for the Cooperation of Energy Regulators, October 2020.
- [9] "Probabilistic Adequacy and Measures", Technical Reference Report, North American Electric Reliability Corporation, April 2018. <https://www.nerc.com/comm/PC/Probabilistic%20Assessment%20Working%20Group%20PAWG%20%20Relat/Probabilistic%20Adequacy%20and%20Measures%20Report.pdf>
- [10] "Pacific Northwest Power Supply Adequacy Assessment for 2024", Technical Report, Northwest Power and Conservation Council, October 2019. <https://www.nwccouncil.org/sites/default/files/2024%20RA%20Assessment%20Final-2019-10-31.pdf>
- [11] R. Billinton and D. Huang, "Basic Concepts in Generating Capacity Adequacy Evaluation," 2006 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), 2006, pp. 1-6, doi: 10.1109/PMAPS.2006.360431.
- [12] "Tracking the origin of the 0.1 day/year LOLE criterion," Slide presentation, Chi-Hung Kelvin Chu, General Electric International, Inc., October 30, 2014.
- [13] E. Ibanez and M. Milligan, "Comparing Resource Adequacy Metrics and their Influence on Capacity Value," 2014 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), 2014, pp. 1-6, doi: 10.1109/PMAPS.2014.6960610.