Adjusting the Aim of Capacity Mechanisms: Future-Proof Reliability Metrics and Firm Supply Calculation

Paulo Brito-Pereira, Paolo Mastropietro, Pablo Rodilla, Luiz Augusto Barroso, Carlos Batlle

This article is a preprint. Please cite the published version:

10.1016/j.enpol.2022.112891

Working Paper IIT-21-042A, first version: October 2019; this version: November 2021

Submitted for publication to Energy Policy

ADJUSTING THE AIM OF CAPACITY MECHANISMS: FUTURE-PROOF RELIABILITY METRICS AND FIRM SUPPLY CALCULATION

Paulo Brito-Pereira[a], Paolo Mastropietro[a], Pablo Rodilla[a],
Luiz Augusto Barroso[a,b], Carlos Batlle[a,c]

[a] Instituto de Investigación Tecnológica, Universidad Pontificia Comillas, Sta. Cruz de Marcenado 26, Madrid, Spain.

[b] PSR Energy Consulting & Analytics, Praia de Botafogo 370, Rio de Janeiro, Brazil.

[c] MIT Energy Initiative, 77 Mass. Av., Cambridge, US and Florence School of Regulation, Florence, Italy.

Abstract

Capacity mechanisms are now deemed a regulatory mainstay in liberalised power system decarbonisation. These instruments aim to ensure sufficient resource adequacy with a mix able to meet the reliability target set by the re. Remuneration in capacity mechanisms depends on so-called firm supply (calculated from de-rating factors or capacity credits), taken as a proxy for each resource's expected long-term contribution to system adequacy. Most adequacy assessment and de-rating methods used to calculate security of supply were developed for power systems very different from today's and tomorrow's, in which renewables account for a higher share of the mix and demand is more elastic. Regulators the world over are already revising these methods, although that seldom involves an overall rethink of their general approach. Drawing from theoretical considerations and international best practice, this article defines an updated theoretical framework for the resource adequacy problem against the backdrop of the challenges ahead for the power sector. The conclusions include recommendations for resilient reliability metrics and de-rating calculation methods.

Keywords

Reliability; adequacy; capacity mechanisms; firm supply; firm capacity; firm energy; de-rating; security of supply; extreme weather events; flexibility.

1 INTRODUCTION

Capacity remuneration or resource adequacy mechanisms are introduced to reinforce the energy market and attract the power system investments needed to guarantee long-term security of supply (Neuhoff and De Vries, 2004; Joskow, 2008; Cramton et al., 2013; Petitet et al., 2017). The ultimate objective is to maximise social welfare (Ozbafli and Jenkins, 2016; Bublitz et al., 2019). Such instruments have become increasingly prominent on the regulatory agenda in the last two decades (Batlle et al., 2015) and are now used in most liberalised power sectors. The impetus has been fuelled, among others, by the growing presence of intermittent renewables in the resource mix. That has intensified the two market failures normally cited in economic theory to justify the introduction of capacity mechanisms, namely the missing money and the missing markets problems (Newbery, 2016).

Capacity mechanisms should be designed around two essential and interdependent elements: i) The first is long-term adequacy assessment able to identify the security-of-supply problem such mechanisms are intended to solve; it is commonly based on the reliability metric (such as loss of load probability) regulators also use to define a target. ii) The second is a de-rating method able to quantify each resource's expected long-term contribution to system adequacy. That parameter, usually denominated firm supply[1], is a key element for investors, as it represents the amount of reliability product they can trade and be remunerated for under the capacity mechanism. Resources are de-rated in the realisation that they are not necessarily available at full installed capacity in situations of scarcity. A power plant with an installed capacity of 100 MW and a 30 % de-rating factor is qualified to trade a firm supply of only 30 MW in the capacity mechanism[2].

Most adequacy assessments and de-rating methods used to calculate firm supply were developed for power systems with resource mixes very different and much simpler, stabler and more predictable than the ones presently in place or expected to come on stream in the future. In conventional power systems with a prevalence of dispatchable thermal power plants, for instance, resource adequacy could be assessed with very simple metrics such as the reserve margin, since scarcity was expected to be attributable solely to insufficient capacity to cover peak demand. In such systems the availability of different resources was

[1] The generic term 'firm supply' is used in this article to include both firm capacity, a concept applicable to power systems with a predominance of thermal power plants (such as Ireland, ISO New England, PJM), and firm energy, a notion used in hydro-dominated systems (such as Brazil, Colombia, Norway, Canada).

[2] Although several references define de-rating factor otherwise, here it is expressed as per cent: the higher the de-rating factor, the higher the firm supply acknowledged the resource.

only marginally interdependent. The contribution of each resource to system adequacy could therefore be analysed separately based on historical data, since the resource mix was not expected to change drastically at any future time.

In a context of energy transition, however, power systems are undergoing speedy change. More technological alternatives are now being implemented, making predictions about the future mix considerably more complex. The presence of renewable technologies whose performance depends on the energy source that drives them has varied the type and magnitude of scarcity affecting power systems and raised the correlation between resource availability and peak load timing. That precludes estimating each resource's contribution to adequacy separately, for contributions depend on overall system performance.

Electricity demand is also becoming more elastic, in turn modifying the concept of adequacy itself, since scarcity can only be defined where price is taken into consideration as well. The complexity of resource adequacy issues has been further intensified by climate change with its corollary, the appearance of extreme weather events and the concomitant re-composition of scarcity conditions. Some of these new patterns have been addressed in the literature (Bothwell and Hobbs, 2017; Mastropietro et al. 2019; Söder et al., 2020; Lambin, 2020; EPRI, 2021).

An urgent need has been identified to reform adequacy assessments and the de-rating methods now in place to embrace all the new technologies and dynamics ushered in by energy transition. Whilst a number of regulators are presently rising to the challenge, the measures taken are frequently designed to solve specific issues (such as the need for a de-rating method for renewables or storage, which may differ from the procedure used for all other resources) and seldom involve a general methodological overhaul.

This article aims to define an updated theoretical framework for the resource adequacy problem and propose a comprehensive approach to rise to address the new challenges. The proposal thoroughly addresses the elements of the problem ranging from the reliability metric used in the adequacy assessment to the calculation of firm supply. The study drew from recent international experience on the subject at hand. The document is structured as follows:

- Section 2 reviews the reliability metrics underlying both adequacy assessments and de-rating. Particular importance is attached to the need to properly factor the price dimension into the reliability metric and make allowance for extreme events.

- Section 3 classifies the methodologies for calculating firm supply based on six design elements and discusses several alternatives, including resources'/technologies' marginal vs. mean contributions and historical vs. forward-looking approaches. The authors propose recommendations for each element of the procedure described, supporting their arguments with theoretical considerations and recent international experience.
- The recommendations are pooled and summarised in section 4, where conclusions are drawn and policy implications defined.

2 RELIABILITY METRICS

The present review of the long-term reliability metrics most widely used in the power sector, especially for capacity and adequacy mechanisms, is followed by an analysis of the general theoretical background on reliability metrics. Two major design dimensions (the type of contingency and the statistical measure used in its study) are identified and the advantages and disadvantages to the procedural alternatives in each are discussed.

2.1 Most widely used long-term reliability metrics

The reliability metrics most widely used to assess power system long-term resource adequacy are reserve margin, loss of load probability, loss of load expectation, expected unserved energy and the 95th percentile loss of load duration (IAEA, 1984; Billinton and Allan, 1994; EC, 2016a; NERC, 2018; ACER, 2020).

- Reserve margin (RM) measures the difference between a system's installed capacity and its peak demand. Potential insufficient availability has traditionally been accounted for by de-rating units' installed capacity further to simple methods (e.g., by multiplying that value by an equivalent forced outage rate). The reserve margin target has commonly been expressed in terms of capacity or as a percentage of the load, dividing the difference between de-rated installed capacity and peak demand by peak demand. Alternatively, targets defined in terms of this reliability metric have also been based on the N-1 criterion, i.e., the capacity reserve margin should be larger than the installed capacity of the largest power plant in the system. Reserve margin is presently the metric used in Spain, Poland and Sweden, as well as in many other systems (EC, 2016b).
- Loss of load probability (LOLP)/loss of load expectation (LOLE) estimate the likelihood of system inability to cover the full load demanded at any given time (rather than peak time only). The two metrics are actually equivalent, differing only in the way reliability

is expressed: as probability (in per cent) in LOLP and as the cumulative duration (in hours/year) of scarcity events during which demand cannot be met in LOLE. Both are widely used in the United States and Europe (NERC, 2018; ENTSOe, 2020).

- Expected energy not served (EENS) quantifies the expected amount of energy that the system will not be able to supply in a certain time horizon. Unlike LOLP and LOLE, which focus on the mere frequency and/or duration of scarcity events, EENS attempts to measure the depth or magnitude of load loss in terms of energy as an indication of severity. Its expression can also be normalised, for example by dividing that cumulative value by the total amount of energy demanded by the system in the given time horizon. EENS is used in Alberta (NERC, 2018) and Australia (AEMO, 2019).
- The 95th percentile of loss of load duration (LOLE95 or LOLD95) is found by applying that statistic to the probability distribution function for loss of load duration (ENTSOe, 2019). It envisages extreme scenarios (although it rules out the uppermost 5 % of the distribution function) and is always larger than LOLE. It is used in Belgium (Elia, 2016).
- Energy supply in the least favourable hydrological scenario is typically used in energy-constrained, such as hydro-dominated, power systems (Rodilla et al., 2015), where the annual electricity demand must be covered in all potential hydrological scenarios, defined on the grounds of historical inflow records. It is used in Colombia.

Two main design elements inform all these metrics: i) the underlying contingency measured (either the number of scarcity events or the energy that could not be served during these events); and ii) the methodology used to estimate this contingency (either deterministic or probabilistic) and the statistical measure used to characterise it. These elements are analysed hereunder.

2.1.1 Type of contingency measured

Events vs. unserved energy

As noted earlier, LOLP and LOLE pool events where demand went unmet in a given period of time without distinguishing degree of severity, i.e., the amount of unserved energy in each event. In other words, with LOLP and LOLE an event where 1 MW of demand went unmet would be accorded the same weight as one where the deficit was 1 GW. That shortcoming may prevent these metrics from capturing a resource's actual contribution to system adequacy.

Assuming a 1-hour, 100-MW loss of load, for instance, inasmuch as a 99-MW power plant would not cover the shortfall, its contribution would not be captured by loss-of-load metrics. In contrast, a 100-MW power plant would offset the deficit entirely, lowering LOLP or LOLE, despite the nearly identical contribution of the two plants to the system. As discrete metrics, then, LOLP and LOLE are able to measure full compensation for a shortage only, a characteristic that could intensify adequacy assessment volatility.

EENS on the contrary measures the severity of each scarcity event and as such constitutes a continuous metric. Applied to the foregoing example, it would show the two plants to cover very similar amounts of the unserved energy, providing a more accurate measure of the overall contribution to system adequacy.

Resilient metrics for power sectors with elastic demand

Nonetheless, power system evolution and increasingly energy price-sensitive demand will render loss of load and unserved energy less meaningful. That is illustrated in the simple graphic example in Figure 1. When demand is totally inelastic (left), a scarcity event can be identified and unmet demand measured intuitively[3]. Conversely, at the other extreme where demand is wholly elastic, unserved energy cannot be defined so readily, for at least part of the demand that goes unsupplied attributes a lower value to electricity than the market clearing price. Such demand should not be deemed unmet, but rather simply not participating in a trade not found to be beneficial by its utility function.

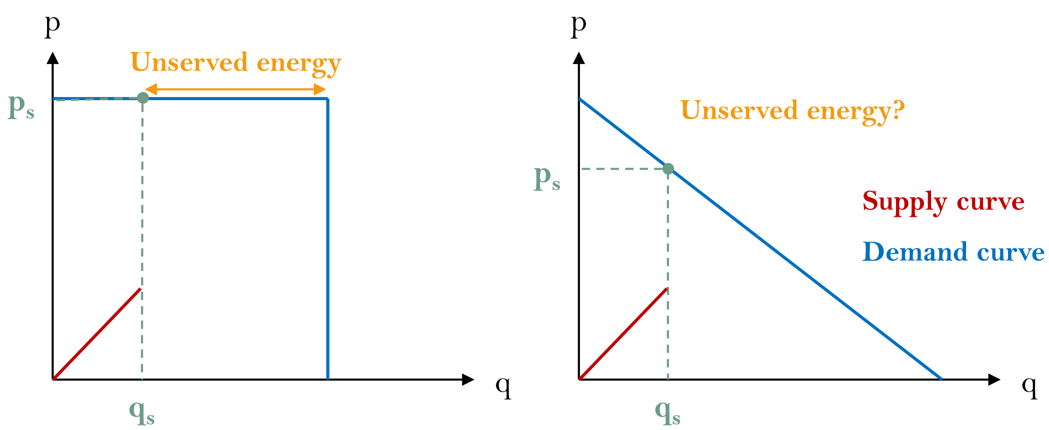

Figure 1: Unserved energy with inelastic (left) and elastic (right) demand

[3] The value of lost load (VOLL) may not be constant, however, and is likely to depend on the agents involved and the duration and frequency of scarcity events (EPRI, 2021).

Be it said, however, that even where demand is totally elastic in the short term, scenarios with extremely high prices might prevail over long periods of time, perhaps denoting a resource mix that deviates significantly from the ideal that would maximise social welfare. Regulators may wish to lower the probability of such events to move toward a resource mix better suited to maximising welfare by implementing a capacity mechanism.

As demand becomes more price-responsive, reliability criteria should take that dimension into consideration, providing regulators with a tool able to identify new scarcity situations. Whilst demand elasticity is still very constrained, especially among certain communities of consumers, capacity mechanisms are long-term regulatory instruments that should be designed to be resilient to changes in the power sector of the future. Rising demand elasticity is a feature many experts believe will characterise the energy transition (MITEI, 2016; EC, 2019).

Even in power systems still characterised by scant demand elasticity, however, the market price is an excellent barometer of scarcity able to accommodate all the dimensions of the security-of-supply problem and provide a neutral approach to their identification. The market price would reveal insufficient generation capacity-related scarcity events attributable to highly-correlated thermal fleet outages, such as in the cold snap that affected PJM in 2014 (PJM, 2018) or Texas in 2021 (ERCOT, 2021). The market price would also identify scarcity situations stemming from a lack of system flexibility, a major concern in many systems where renewables are acquiring a growing presence. The quintessential example is California's rolling blackouts in August 2020 (Joskow, 2020; CAISO, 2021). Market pricing would likewise reveal the long-term power shortages that affect hydro-dominated systems in dry years when hydropower reserves decline, as in Colombia in 2015/2016 (Mastropietro et al., 2020).

One possible way to factor the price dimension into reliability metrics would be by including high-price events in traditional adequacy assessments. That would mean broadening loss-of-load metrics to take account of hours when the market price exceeds a certain threshold. Unserved energy metrics, in turn, would have to envision the energy offering cleared at higher than a certain reference price. The difference between traditional reliability and price threshold-based metrics is represented graphically in Figure 2. Such a price threshold would aspire to determine the actual willingness to pay by the part of the demand not participating in the market and consequently undetected.

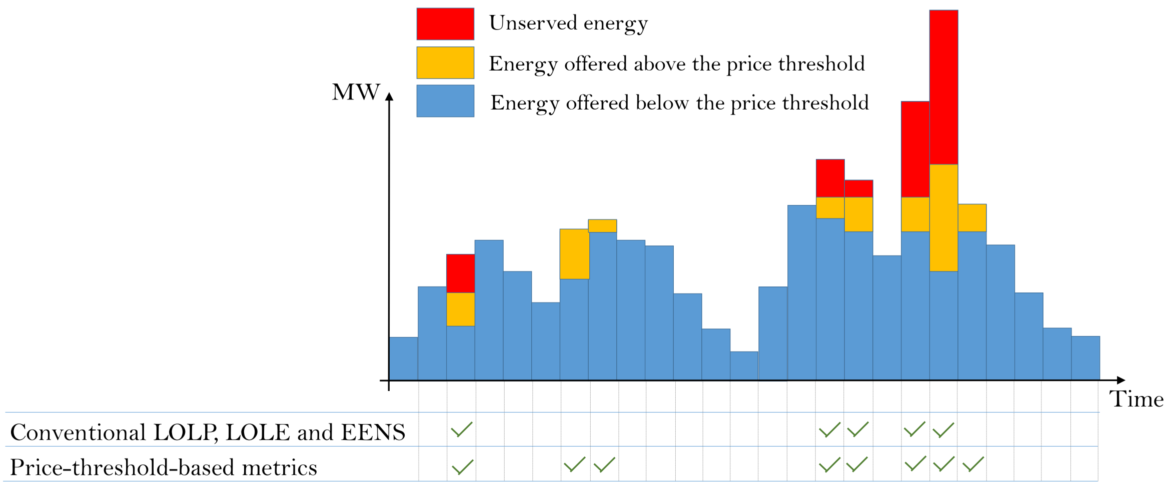

Figure 2: Difference between traditional reliability and price threshold-based metrics

Factoring in the price dimension is not a new idea in capacity mechanism design, for it is among the earliest fundamentals of the reliability options approach (Vázquez et al., 2002; Bidwell, 2005; Cramton et al., 2013; Andreis et al., 2020). With such mechanisms (reliability options are currently being used in Colombia, Ireland and Italy), however, the market price is used only to define the reliability product and the rules activating verification of actual performance by a resource party to a reliability contract. It is taken into consideration neither in the adequacy assessment nor in firm supply calculations.

The sole experience anywhere in the world where a power system pondered accommodating the price dimension in adequacy assessment was a proposal put forward by Elia (2019a), the Belgian system operator, for redesigning the country's capacity mechanism. Elia realised that focusing solely on loss-of-load events was insufficient, for it would not rule out near-scarcity events that would also be critical to system adequacy. The institution consequently contended that firm supply calculations should take account of resource contributions during near-scarcity hours, defined as periods with unserved energy plus times when a minor rise in demand would translate into unserved energy. Whereas Elia studied the possibility of using a price threshold to identify near-scarcity hours, it disregarded that pathway on the grounds that administratively setting such a scarcity price would be arbitrary (Elia, 2019b). That argument may be countered, however, with the observation that defining a demand increment threshold deemed as conducive to unserved energy would be equally arbitrary.

2.1.2 Statistical measure used to characterise the contingency

Deterministic vs. probabilistic approaches

The underlying contingency (number of scarcity events, unserved energy or energy supplied above a price threshold) may be measured deterministically, i.e., based on a single scenario, or probabilistically, in which a number of scenarios, each with a given probability of occurrence, are envisioned. The underlying contingency for each scenario is then computed and a probability distribution function is built. In the probabilistic approach a statistical measure is then needed to 'condense' the function into a single value.

Both deterministic metrics and probabilistic metrics based on overly simplistic assumptions fail to capture the probabilistic nature of the adequacy problem (EC, 2016a). As their computational simplicity is offset by their steadily declining accuracy in representing modern power system realities, they are used in very few jurisdictions. By the same token, the more complex probabilistic approach calls for a fair number of assumptions that may sway the outcome. It delivers more precise results, however, for it can capture the significant interdependence among the variables studied (such as the occurrence of scarcity events and the availability of certain technologies).

Mean and median vs. consideration of extreme scenarios

Whereas mean and median values may be representative of the entire probability distribution function, they underestimate the weight of the tail of the probability distribution function, where the most severe shortages lie. Since the purpose of capacity mechanisms is to guarantee security of supply, particularly under the harshest conditions, reliability metrics that cover extreme scenarios may provide very valuable information. That has been acknowledged by Australian institutions, which have drawn attention to the rising likelihood of extreme events due to climate change (AEMO, 2019). This procedure may lead to heavier tails in the distribution function of stress events (as shown graphically in Figure 3 for Australia) that should be internalised in adequacy assessments[4].

---

[4] The importance of encompassing extreme weather events in adequacy assessments is further highlighted by the dramatic supply crisis faced by ERCOT in February 2021 (ERCOT, 2021).

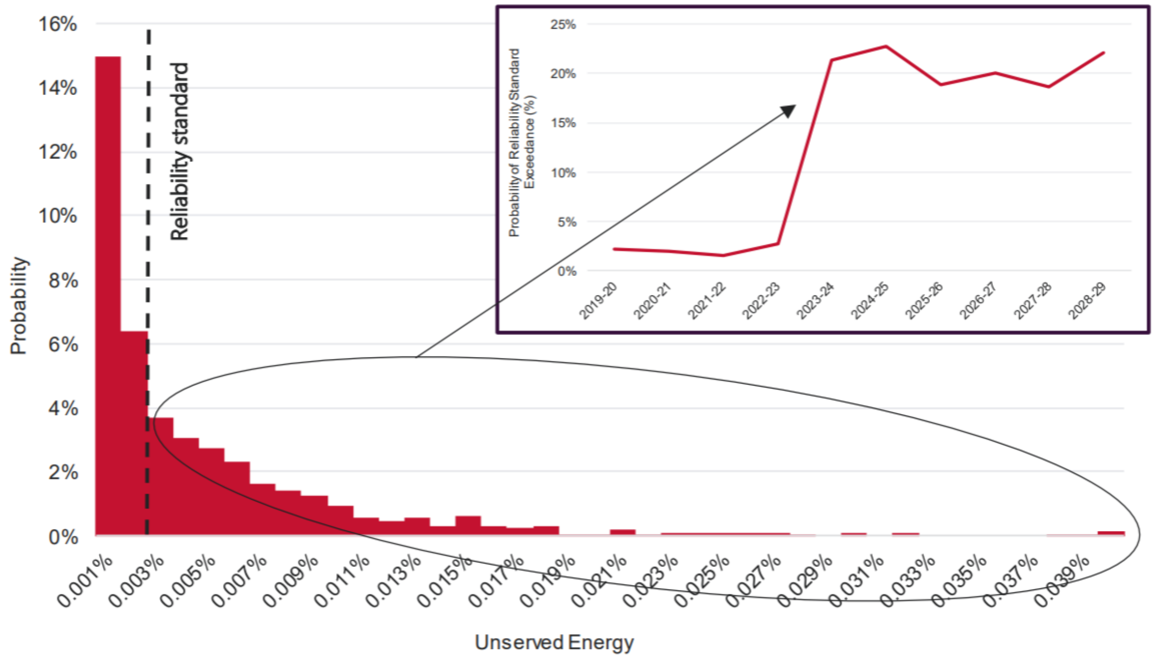

Figure 3: Distribution of annual unserved energy in New South Wales, 2023-2024 (AEMO, 2019)

The Colombian power system affords another example of the importance of statistical parameters. In that system, hydropower accounts for 70 % of total installed capacity. Every few years, however, the El Niño phenomenon brings droughts and high temperatures to the region for several months, reducing hydropower reservoir inflows and jeopardising security of supply (Mastropietro et al., 2020). The El Niño effect may vary in duration, intensity and periodicity, but it always has a long-term impact on the Colombian electricity market price, which may remain high for months and is used here as an indicator of scarcity conditions. The graph in Figure 4 of the price of electricity in Colombia over the last 25 years illustrates its rise during El Niño events[5].

[5] El Niño periods are formally defined in the Oceanic Niño Index (ONI) (NWS, 2021).

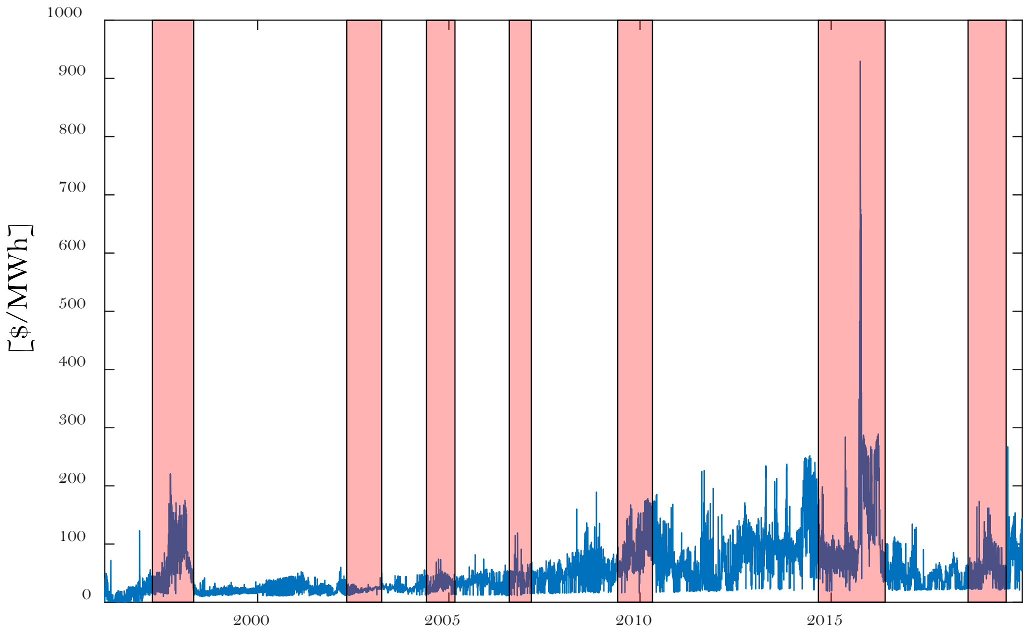

Figure 4: Wholesale electricity market prices in Colombia, 1996-2019, inclusive (authors' formulation based on XM data (XM, 2021) and BRC exchange rate information (BRC, 2021))

A subset of the Colombian market prices (2009 to 2019) given in Figure 4 was rearranged to build the probability distribution function[6] graphed in Figure 5 to exemplify the limitations of some statistical parameters. Further to the information included on the graph, even very high (95th or 99th, PCTL 95/99) percentiles may fail to capture the full weight of the tail.

[6] Assuming equiprobability for all instances.

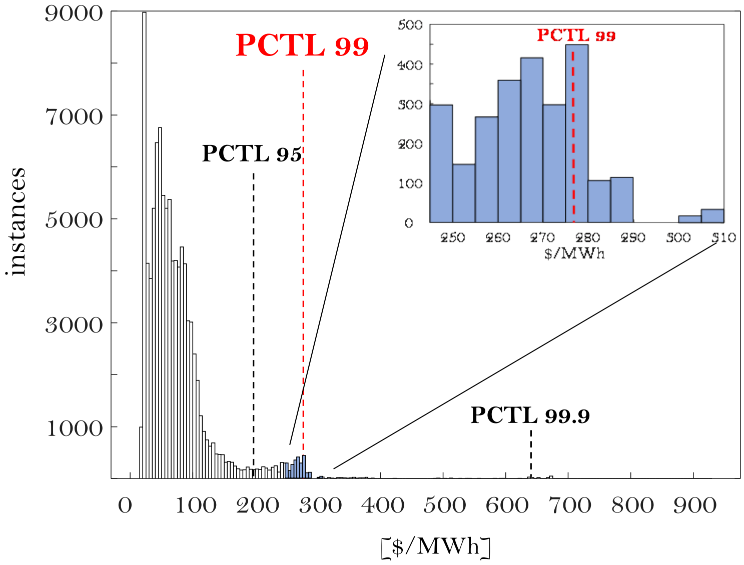

Figure 5: Distribution function for wholesale electricity market prices in Colombia (authors' formulation based on XM data (XM, 2021) and BRC exchange rate information (BRC, 2021))

A statistical parameter that is less widely used in adequacy assessments but that might remedy the aforementioned shortcomings is Conditional Value at Risk (CVaR). CvaR focuses on extreme scenarios, isolating the tail of the probability distribution function, defined as a percentage (α) of the worst cases and calculating its mean value (NERC, 2018). As Figure 6 shows, the CVaR of a system's unserved energy would be the weighted mean of that energy in the least favourable scenarios (the upper 5 % of the probability distribution function, for instance, with α=5 %).

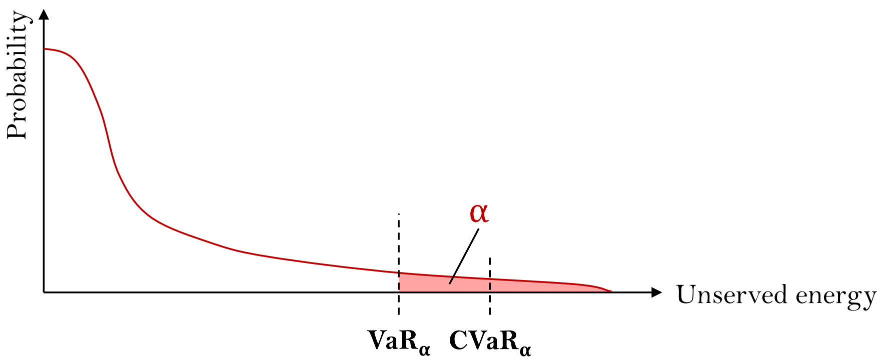

Figure 6: Graphical representation of the CVaR applied to unserved energy

2.2 Best practice

2.2.1 Type of contingency measured: introducing the price dimension in the metric

Further to the critical analysis discussed above, the authors believe the optimal reliability metric for managing the new present and future electricity resource mix is the energy cleared and supplied above a price threshold (a definition that includes both unserved energy and energy offered and cleared over that price). As this reliability metric, like EENS, is continuous, it enhances the consistency and stability of the results of any assessment conducted with it[7]. Unlike EENS, however, it is resilient to higher demand elasticity, since it identifies scarcity conditions on the grounds of market price, the most accurate barometer of such conditions (see subsection 2.1.1).

The literature on reliability options and the international experiences of countries where this kind of capacity mechanism is in place can be used as a reference to define the price threshold. According to Vázquez et al. (2002), the strike price of reliability option contracts should be high enough to elude interference with market operation under normal circumstances. That is applicable to any price threshold used in the reliability metric, which should identify only those situations where security of supply is at risk but not where prices rise for other reasons (such as a spike in fuel prices).

The threshold for the strike price of reliability options may be defined from the variable cost of peaking units. In Italy for instance, the strike price is set as the variable cost (€/MWh) of the reference peak technology, in turn defined as the dispatchable technology that would be included in the optimal generation mix delivering the lowest unit investment cost (€/MW). Whereas the reference technology does not change, the strike price is subject to indexation and varies weekly. The Italian indexation formula includes fuel costs (the most prominent item), energy imbalance costs, CO_2 costs and green-certificate costs, among others (Terna, 2018a). In the current context of ever greater fuel and CO_2 price long-term volatility, threshold indexation would be an essential feature of the reliability metric.

---

[7] To return to the example in item 2.1.1, the 99-MW and the 100-MW power plants would be assigned very similar de-rating factors and no discontinuity would be observed between them.

2.2.2 The statistical parameter: shifting the focus to extreme events

The metric proposed should be studied with a model designed to process the variables at issue sequentially to simulate power system operation in the time horizon under study. In such a probabilistic model the reliability metric must depend on a statistical parameter that pools the results for all the scenarios envisioned. As contended earlier, the statistical parameter best suited to today's and tomorrow's power systems is CVaR, which identifies the mean value of the tail in the probability distribution function for energy supplied above the price threshold. In that approach attention can be focused on extreme scarcity events, the ones capacity mechanisms are designed to mitigate and the ones whose frequency and intensity may be heightened by climate change.

3 METHODOLOGIES TO CALCULATE FIRM SUPPLY

By way of follow-up on the alternative ways to determine reliability metrics discussed above, this section addresses theoretical considerations around methods for calculating firm supply (also termed de-rating methods) as an element in capacity mechanism operation. Firm supply, the product traded in capacity mechanisms, is designed to acknowledge and incentivise resource contributions to system adequacy, defined as their ability to produce (or fail to consume) in periods when the system is strained.

The design alternatives addressed in this analysis are as follows.

- Firm supply may be calculated taking each resource or technology separately or in conjunction with the rest of the mix on the grounds of the expected dispatch of the entire system, which internalises potential inter-technology synergies.
- The reliability metric used to calculate each resource's contribution must be chosen from among those available.
- Firm supply may be based either on mean or marginal contributions.
- Firm supply calculations may be based on historical or projected data.
- Firm supply may be calculated for each resource individually or pooling the data for all the resources sharing a given technology.
- As either one or several products may be considered in the capacity mechanism, the impact of the latter option, which would call for calculating several firm supply values, must be determined.

As in the preceding section, a list of best practices is provided for each of the items discussed, further to which an ideal methodology for calculating firm supply is proposed in the conclusions.

Before proceeding to conduct the analysis, however, one initial recommendation is in order: the methodology for calculating firm supply should be the same for all the resources potentially participating in the capacity mechanism. Any other approach would segregate resources, for which no theoretical justification can be alleged. Although defining specific methodologies for new technologies (as many regulators did and some still do for variable renewable energies) may be a quick and easy way to open the capacity market to these resources without revising adequacy assessment as a whole, the outcome is suboptimal and may distort competition among rival technologies.

3.1 Measuring resource contribution separately or as part of the system as a whole

The contribution of a resource to system adequacy depends on its output in scarcity events. As such events result from the balance between demand and power availability, electricity system adequacy depends on the combined performance of all its component resources. Shortages may differ in duration, lasting but a few hours in capacity-constrained systems but much longer (weeks, months or years) in energy-constrained systems.

In some jurisdictions, however, firm supply for certain technologies is calculated on the grounds of the performance of each resource or technology separately (Mastropietro et al., 2019), irrespective of the conditions prevailing in the system. For instance, a wind power plant's firm supply may be defined as a certain percentile of its capacity injections[8]. Similar reasoning is deployed when a thermal power plant is de-rated based only on its equivalent forced outage rate, ignoring any possible correlation among such outages and/or between outages and the appearance of shortfalls. Examples of that correlation include the cold snap that affected the eastern United States in 2014 (Mastropietro et al., 2017) and the more recent extreme weather event in Texas in 2021 (EPRI, 2021).

[8] That is not the same, however, as using a certain resource's capacity injection, or a percentile of it, during scarcity conditions (defined, for instance, as events with unserved energy or very high prices). In such cases, the correlation between resource output and system dispatch is associated with the definition of the scarcity conditions.

Best practice

Theoretically speaking, any methodology that calculates a resource or technology's firm supply based solely on its performance should be avoided. Such readily implemented approaches may be used when a certain technology is first introduced but as they are extremely imprecise they induce significant error and inefficiency. Resilient methodologies should define each resource's firm supply deemed as a component of system dispatch as a whole.

3.2 Reliability metric for assessing contributions

Theoretically, capacity mechanisms are introduced when adequacy assessment shows the regulator's reliability target to be at risk. Resources are presumably remunerated for contributing to reaching that target. Each resource's firm supply should consequently be defined as its contribution to the reliability target set by the regulator and calculated based on the metric used in the assessment.

Although such reasoning may sound obvious, it is surprisingly often absent from standard practice. In most power systems relying on capacity mechanisms, including for instance the United Kingdom (National Grid, 2019a), Belgium (Elia, 2019a and Elia, 2019b) and many of the systems in place in the United States (NYISO, 2019; PJM, 2019), firm supply of some or all technologies is calculated using a different metric than used to define the reliability target. By way of example, adequacy is assessed in the UK using LOLE, while de-rating factors for renewable resources are calculated on the grounds of their contribution to lowering EENS. Belgium also conducts adequacy assessments based on LOLE, whereas the operator proposed calculating renewable de-rating factors further to expected output during the near-scarcity hours, as discussed in item 2.1.1.

That approach is clearly suboptimal, inasmuch as it entails remunerating resources for a service that while related to is not exactly the one needed to meet the target defined. Theoretical justification for this assertion can be found in the mathematical formulation of the optimisation problems given in Annex I, where welfare maximisation in a centralised context, constrained by a system reliability target, is compared to the marketplace, which generates price signals for both energy and firm supply (Pérez-Arriaga, 1994; Schweppe et al., 1988).

Best practice

Adequacy should be addressed holistically. In other words, the same metric should be used both to define the reliability target in the adequacy assessment and to calculate each resource or technology's firm supply.

3.3 Marginal (or incremental) vs. mean historical contributions

The design element bearing what is likely to be the heaviest impact on de-rating factor calculations is the decision to use marginal or mean contribution values. In the former the analysis focuses on the impact on system adequacy of minor (typically a few MW) increments in resource or technology capacity, whereas in the latter it is the mean contribution of resource or technology capacity as a whole[9].

The difference between mean and marginal contributions is particularly relevant to technologies whose presence is liable to progressively change the scarcity conditions facing the system. One example is solar PV-based resources, which help reduce the likelihood of shortfalls during the daylight hours but are unable to generate power late in the evening. The curve graphed in Figure 7 compares the contribution between solar generation and that of all other technologies. Whereas absolute demand peaks early in the evening, net peak demand (total demand minus solar production) is recorded in the late evening. Solar power consequently shifts and lowers net peak demand, thereby impacting the likelihood of scarcity events during the day.

---

[9] The decision on whether to calculate a homogeneous firm supply for the entire technology (e.g., through a common de-rating factor) or an individual firm supply for each resource is another design element, which is discussed in section 3.5.

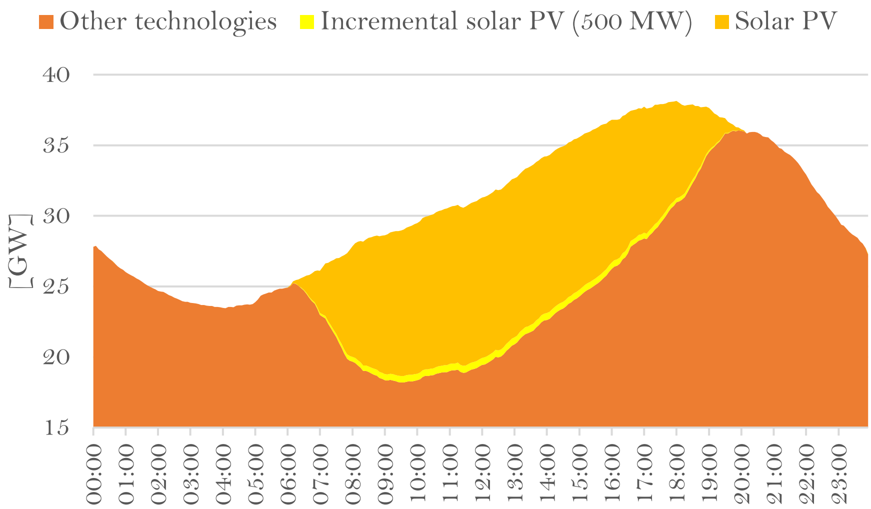

Figure 7: Total and net demand for a system with substantial solar PV output (authors' formulation based on California ISO data (CAISO, 2019) for 16 July 2019, including incremental solar power)

Figure 7 includes the hypothetical production of an additional 500 MW of solar capacity (light yellow) to depict the effect of a marginal increment in solar PV installed capacity[10]. Simplifying the matter by assuming the system under study to be purely thermal with no energy storage, the contribution of the additional capacity to system adequacy is minimal. In other words, as its production at net peak demand is practically nil, its contribution to marginally calculated firm supply would be close to naught. As a result, new solar resources installed under such circumstances should not be remunerated by the capacity mechanism.

Conversely, where solar technology is taken as a whole, it contributes to raising system adequacy for it lowers net peak demand. Calculations based on the mean contribution would accord solar resources positive firm supply, although its value would decline with each new solar plant installed. (The total contribution, which rises almost negligibly, would be shared by a larger number of resources). Acknowledging the mean contribution for the technology as a whole to the new resources would credit them for greater firm supply than merited by their actual contribution.

The downturn in solar PV marginal firm supply with greater system presence is depicted in Figure 8, which plots PV solar power ELCC against installed capacity in California ISO.

[10] Computationally speaking, the marginal increment should be lower than 500 MW, a value selected here only to make the new area visible in the chart.

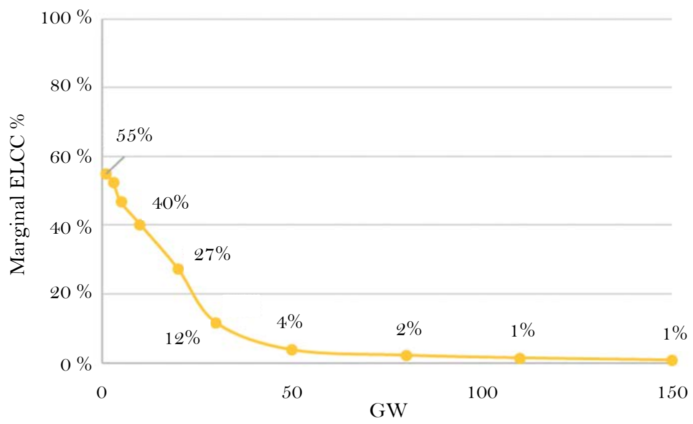

Figure 8: Variation in solar PV marginal ELCC vs installed capacity (Energy & Environmental Economics, 2019)

The contribution graphed in the figure is expressed as marginal ELCC (Effective Load Carrying Capability), a probabilistic method that estimates the additional demand that would be met by raising the installed capacity of the resource at issue with no detriment to reliability. An alternative approach would be to lower resource installed capacity and analyse the amount of 'perfect' generation (with a de-rating factor of 100 %) that would need to be installed to ensure the reliability target established. The baseline generation mix used in ELCC calculations is as required to guarantee the regulator's reliability target (Garver, 1966). Although ELCC is widely used in capacity mechanisms, it must be understood as a generic computational method that can be implemented in a number of ways. Its practical implementation may benefit from the recommendations made hereunder (such as basing the parameter on a continuous and resilient metric).

This discussion has focused up to now on the role of the resource mix in determining each resource or technology's firm supply. Firm supply obviously depends as well, however, on the technical characteristics of the resource studied (such as ramping capability or energy constraints). That notion is particularly relevant to storage technologies. For their capacity markets, both Ireland (I-SEM, 2018) and the United Kingdom (National Grid, 2017) attribute variable de-rating values to resources grouped under different 'storage classes' defined in terms of storage duration. Further to the data in Table i the de-rating factor is lowest for the shortest durations, inasmuch as that characteristic limits such resources' contribution during scarcity events.

Table i. Capacity market de-rating factors proposed for duration-limited storage classes in the 2018/19 T-1 and the 2021/22 T-4 auctions in the UK (National Grid, 2017)

De-rating per storage duration	2018/19	2021/22
0.5 hours	21.3%	17.9%
1.0 hour	40.4%	36.4%
1.5 hours	55.9%	52.3%
2.0 hours	68.1%	64.8%
2.5 hours	77.3%	75.5%
3.0 hours	82.6%	82.0%
3.5 hours	85.7%	85.7%
4.0+ hours	96.1%	96.1%

Best practice

With the marginal approach, the contribution to system adequacy is estimated more accurately and the economic signal emitted by the capacity mechanism is more efficient (Bothwell and Hobbs, 2017). The advantages of using marginal contribution have been acknowledged and implemented by some regulators, such as in Ireland (SEMC, 2018) and the United Kingdom (National Grid, 2017).

Marginal contribution optimality is also inferred by the mathematical formulation set out in Annex I. The resulting optimality shows that the per-unit remuneration of a give resource (M_{K_i}) for its participation in the capacity mechanism is the derivative of the reliability metric (RM) used in the adequacy assessment relative to the installed capacity of that resource (K_i), multiplied by the dual variable of the constraint associated with the reliability target in the centralised optimisation problem (β). In a way, the latter parameter represents the capacity/adequacy market price:

$$M_{K_i} = \frac{\partial \mathrm{RF}}{\partial \mathrm{K_i}} \cdot \beta$$

A second-best yet efficient approach

As noted in item 2.2.1 above, two types of reliability metrics can be defined depending on the contingency measured. Where the regulator opts for a continuous metric (such as EENS or energy cleared and supplied above a price threshold) a second-best alternative to the marginal contribution may be applied, as discussed in this item.

Marginal contributions to resource adequacy are computed by modelling the resource mix and analysing the variation in the reliability metric resulting from a marginal increase in a resource or technology's installed capacity. Such modelling should simulate the future resource mix based on the best forecasts available, as described in greater detail in section 3.4.

Taking the reliability metric proposed in subsection 2.2, calculations would be performed as outlined in Figure 9. Critical periods are defined as times when energy, including unserved energy (red areas in the chart), is cleared and supplied above a price threshold (yellow areas in the graphs). A marginal rise in a resource or technology's installed capacity (green areas on the graph) might contribute to a reduction in some of this 'critical' energy. That reduction would be the firm supply the resource under study is able to add to the system during the time horizon defined in the model, calculated from its marginal contribution (best practice method).

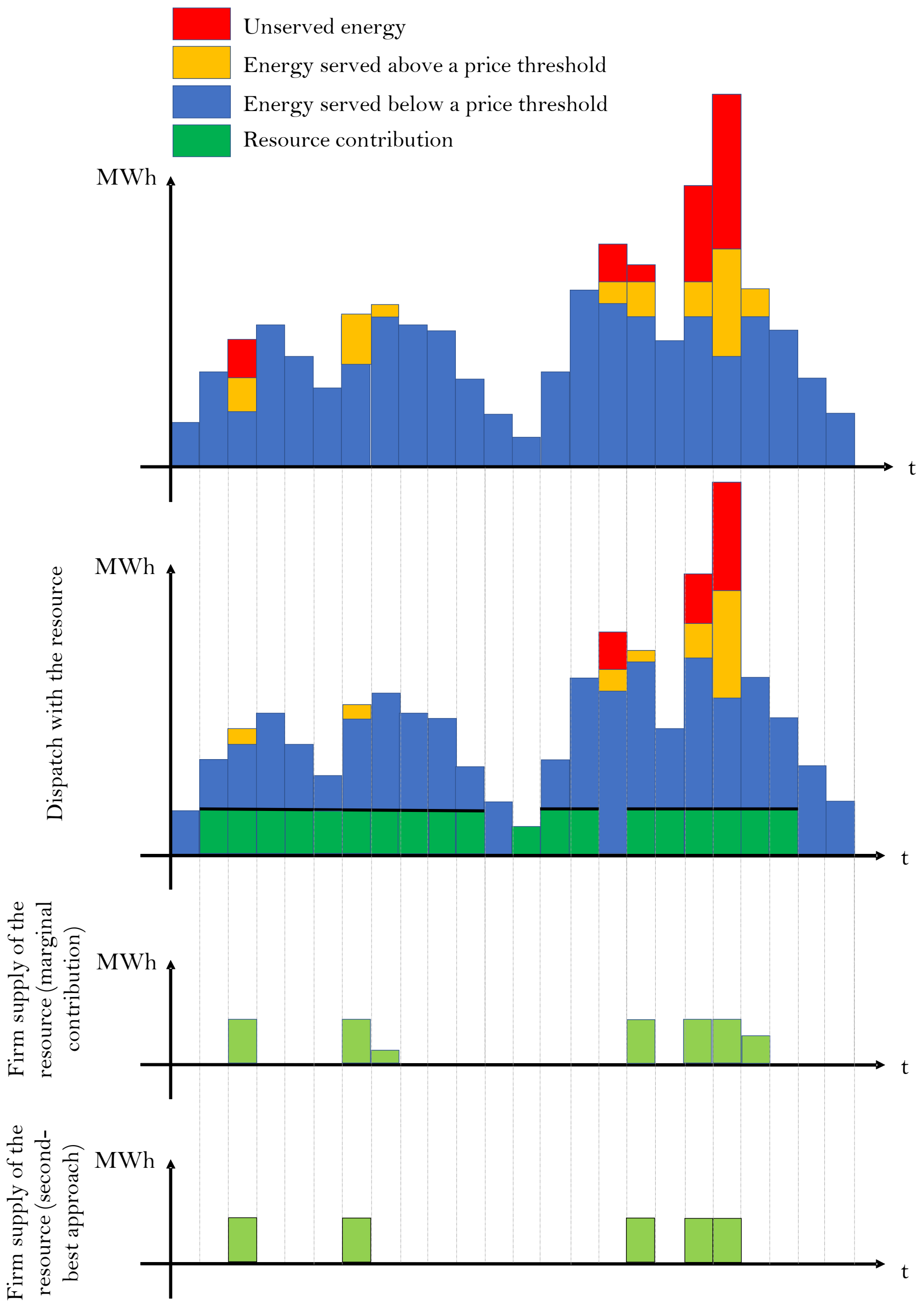

Figure 9: Graph showing the best practice (marginal) and second-best methods for calculating firm supply

This methodology delivers an accurate estimate of each resource's contribution to the reliability target. It may be subject to computational problems, however (Faria et al., 2009). To begin with, the probabilistic nature of the model and the need to repeat the process for each resource or technology may involve a very large number of simulations. Secondly, the algorithms used to solve these optimisation problems may be unstable, with a minor rise in the installed capacity of a certain technology possibly resulting in significantly different dispatch arrangements.

Such computational problems may be eluded by running a single probabilistic simulation with the resources expected to comprise the future generation mix and evaluating each one's production during critical periods as defined above. Such an approach disregards critical periods when a resource's marginal contribution would eliminate the critical energy, as shown in Figure 9. The error would be minor, however, provided the marginal increase involved is small.

Given that such a model would be probabilistic, production during critical periods would be assessed under more than one scenario. If, as recommended in subsection 2.2, CVaR is used as the statistical parameter to calculate the reliability metric, the focus will be on scenarios with a higher total critical energy. Under those circumstances a resource's firm supply would be its mean output during critical periods in such scenarios.

3.4 Historical vs predicted future contributions

Many power systems that apply adequacy mechanisms calculate resources' firm supply as their historical output in past scarcity situations (RTE, 2014; ISO New England, 2016; Terna, 2018b; NYISO, 2019; PJM, 2019). Power systems are evolving rapidly, however, with the introduction of new generation and storage technologies and rises in demand elasticity. The inference is that historical data might not be representative of future power system operation. If firm supply is calculated from historical data, regulators run the risk of remunerating resources that may be unable to contribute to countering expected future scarcity events[11].

If capacity mechanisms aim to meet future reliability targets, firm supply should be based on projections of future power system operation. That entails using a model to simulate operation in different scenarios (varying hydro inflows and renewable production, among others) and analysing each resource's contribution to the established target. That idea, which is hardly new, lies at the base of methods like the aforementioned ELCC (CPUC, 2014).

The results of simulations depend of course on the initial assumptions and the scenarios established, although such problems are common to all models and methods. Calculating

[11] Some power systems may be characterised by a resource mix and scarcity conditions not be subject to major energy transition-mediated alteration. If scarcity conditions follow the same pattern, data on historical output during stress events may continue to afford constitute a good approximation for calculating firm supply.

firm supply from historical data is tantamount to assuming that scarcity conditions will not change in the future, a premise over which energy transition is casting doubts. As noted in subsection 3.1, given that a resource's firm supply depends on system operation as a whole, future operation forecasts should be based on the best information available rather than assuming that operation will remain constant in the future.

A number of regulators around the world have acknowledged the benefits of projection-based firm supply calculations (Moreno et al., 2010). California's regulator has recently changed its approach to firm supply calculations for intermittent renewable resources from the use of historical data to the deployment of a simulation model (CPUC, 2017). Other power systems that have introduced capacity mechanisms in recent years, including Ireland (SEMC, 2018), Belgium (Elia, 2019b) and the United Kingdom (National Grid, 2019b and National Grid, 2017), seem to prefer to calculate firm supply on the grounds of projections.

Best practice

Based on the foregoing theoretical arguments and the international experiences cited, and in keeping with the proposal for the reliability metric, the authors believe that each resource or technology's firm supply should be determined with a simulation model that processes the variables of interest sequentially. Input, in turn, should be drawn from the best forecast available on future system operation.

3.5 Pooled calculations of firm supply for all plants using the same technology

Another issue to be taken into consideration is whether firm supply calculations should be performed for each resource separately or pooled for all resources using the same technology. Theoretically, the former would be the optimal approach, for the technical characteristics (such as position in the grid or availability of the primary energy source) of resources sharing a given technology may differ from one plant to another.

The trouble with this approach is that while theoretically robust, its real-life application encounters a major drawback. Assessing the marginal contribution of each resource calls for de-rating methodology based on an optimisation model able to simulate each one's future performance. The optimisation software embedded in modelling tools typically computes the optimal solution with too wide a tolerance to evaluate the expected marginal contribution with any precision (see earlier comment in subsection 3.3).

To address such computational issues and the volatility of the respective outcomes, regulators tend to define firm supply by pooling all the plants that share a given technology and assessing their combined performance. That approach has been adopted by the United Kingdom (National Grid, 2019a), Belgium (Elia, 2019b) Ireland (SEMC, 2018) and Italy (Mastropietro et al., 2018), among others. Calculating firm supply by technology may yield an acceptable approximation if all resources using each could be assumed to contribute equivalently or similarly to the reliability target. A case in point would be nuclear power plants, which are likely to contribute similarly to the reliability target unless they are subject to very different forced outage rates or the constraints on their fuel supply vary significantly.

Such approximations may be inaccurate for other, particularly non-conventional, technologies, however. Wind farms, for instance, not only come in all manner of sizes and configurations, but their output depends on the availability of the wind resource, which may vary geographically within the power system. Attributing a single firm supply value to the entire wind fleet on the grid would not capture those differences, with the risk of giving project developers inefficient incentives. The United Kingdom de-rates onshore and offshore wind facilities differently for precisely that reason[12] (Ofgem, 2021). Ireland, in contrast, where offshore wind is not expected to be connected any time in the near future, establishes the same de-rating factor for its entire onshore wind capacity (SEMC, 2018), since given the size of the island and its prevailing winds output is closely correlated across its entire wind fleet.

Such correlations are not limited to non-conventional technologies. Recent stress events such as the extreme weather conditions that hit Texas in February 2021 (ERCOT, 2021) may suggest high correlations between the outage rates of individual thermal power units such as combined cycle power plants, which should not be handled separately to calculate their firm capacity (EPRI, 2021). Similarly, in scorching weather even nuclear power plants may go offline simultaneously, as in France and Germany in summer 2019 (Reuters, 2019).

Conversely, calculating a single de-rating factor for an entire technology may be inefficient for hydroelectric power plants. That may be attributed to two factors. On the one hand, as no two hydroelectric resources are totally equivalent (in light of the interrelations among installed capacity, reservoir size, hydro inflows and similar), they are unlikely to contribute

[12] In for the 2019 T-3 auction, for instance, the de-rating factors for onshore and offshore wind were 8.2 % and 12.3 % respectively (Ofgem, 2021).

equally to the reliability target. And on the other, the output of certain plants may be interdependent, such as where several hydropower facilities are sited on the same river basin. In such cases, attributing the most efficient firm supply value to each resource would be a very complex, if not impossible, endeavour. A more robust solution would be to calculate a single firm supply value for the entire hydropower capacity on a given river basin and subsequently design a method to divide that value among them.

Best practice

The authors deem that resources sharing a technology should be grouped when they contribute similarly to the reliability target. The model for calculating firm supply would deliver a single de-rating factor for the technology, which could then be used to calculate each resource's firm supply. That simplification should not be adopted, however, when resources with the same technology contribute very differently to system adequacy. This would be true for renewable resources sited in different areas of the system and subject to very different primary energy availability conditions. Geography might also affect thermal power plants, for example, due to differences in constraints on fuel availability between one area of a country and another (CSMEM, 2016; Freeman et al., 2020).

Similarly, the proposal for hydropower plants is to calculate a single firm supply value for all those located along the same river basin. That value could then be divided among the plants in keeping with specific rules designed to generate signals that efficiently incentivise the agents concerned (see Faria et al., 2009, for a discussion of alternatives).

3.6 Single vs multiple products

Yet another element in the design of firm supply calculation methodology revolves around whether the capacity mechanism envisages a single or multiple products. Whilst standard practice is to address a single product, several may also be defined by:

- factoring time criteria into the reliability product by calculating firm supply seasonally (winter/summer) or even monthly;
- establishing different products for tackling short-term and long-term scarcity events, which would also entail establishing firm capacity and firm energy values or using flexible firm capacity in the calculations.

Where two or more reliability products are assumed, the reason is most often to accommodate time-related issues. A number of power systems define seasonal (winter and summer; ISO New England, 2016; NYISO, 2019) or even monthly (CPUC, 2017) firm supply values. In this approach adequacy assessment is broken down into sub-problems to perform firm supply calculations based on the reliability target. If monthly reliability products are defined, irrespective of whether they are purchased in separate auctions or jointly in the same auction, they rule out the existence of a single product with a single price.

The inclusion of a number of products in the capacity mechanism and the concomitant calculation of several firm supply values could benefit resources that meet the respective requirements most fully. That is clearly illustrated by solar power units. Such resources have higher output in the summer. If they are de-rated monthly, their firm supply would be lower in the winter and higher in the summer months. Breaking firm supply up into shorter periods would translate into lower risk for resources with seasonal output, whereas using a yearly value would require them to provide the same firm supply for the full 12 months (in certain jurisdictions similar effects may be observed for wind and hydropower plants as well as in terms of demand response[13]). Be it said also that transferring the seasonality risk to market agents may favour large generation companies with resources using different technologies, for they would be able to offset the seasonal production of one resource with the output of others in their portfolio.

Best practice

The decision around how many products should be envisaged in the calculations calls for balancing a number of factors, including method simplicity and transparency and risk and uncertainty management. As optimal balance also depends on system characteristics and the resource mix, no one-size-fits-all recommendation can be advanced for this design element.

4 CONCLUSION AND POLICY IMPLICATIONS

A mainstay in market design, capacity mechanisms will constitute a vehicle for energy transition in the power sector. They purport to enhance security of supply during the paradigmatic change in the resource mix to be implemented in the decades to come by enabling all resources that can actually contribute to system adequacy to participate. Such

[13] According to SEPA (2019), PJM approved a summer-only DR proposal to accommodate demand response in connection with the cycling of air conditioning possibly ineligible for annual capacity payments.

participation must be based on modern de-rating methods that efficiently quantify the amount of firm supply that each resource can trade in the capacity market.

This article discusses a theoretical framework for the adequacy problem and analyses the direction regulator tool development should take in the new circumstances. The most prominent result is a comprehensive proposal that addresses both adequacy assessment and the method for calculating firm supply.

- As the reliability metric proposed (CVaR of unserved energy plus energy cleared and supplied beyond a price threshold) is based on market price, it is resilient to the rise in electricity demand elasticity expected in the near future. CVaR accommodates extreme weather events, whose frequency and intensity are likely to grow in the wake of climate change (dramatically illustrated by the 2021 Texas crisis).
- Firm supply calculation methodology must be based on the same reliability metric as used to establish the adequacy target and ideally be the same for all resources.
- Firm supply should be based on the marginal contribution of each resource or technology to ensure the efficacy of the signals sent by de-rating methodology, which should not attract technologies not expected to improve system adequacy. The mathematical substantiation of this recommendation is provided in the Annex.
- Firm supply should be determined with a probabilistic model to simulate power system operation for a future resource mix in anticipation of the significant changes envisioned in the decades to come, which may also alter the nature of the scarcity conditions the system will need to handle.
- Marginal contributions can be estimated from the energy generated by each resource or technology during the critical periods identified by the simulation model.

In addition to this comprehensive proposal, the foregoing discussion may also prove useful for regulators presently introducing or revising a capacity mechanism or revising the associated de-rating methodology when analysing the pros and cons of the dichotomic alternatives addressed in section 3 (projections vs. historical data; marginal vs. mean contribution; per-resource vs. per-technology de-rating; annual vs. seasonal/monthly de-rating).

ACKNOWLEDGEMENTS

This article has been elaborated in the framework of the research project *Flexener*, financed by the Centre for the Development of Industrial Technology (CDTI) of the Spanish Ministry of Science and Innovation, in the context of the call *Misiones CDTI* (grant MIG-20201002). The authors would like to express their gratitude to CDTI. Be it said, the opinions presented in this article are soley the authors view point and do not necessarily reflect the funding entity or the institutions where the authors carry out their research activity.

REFERENCES

ACER, Agency for the Cooperation of Energy Regulators, 2020. Decision no. 23/2020 of 2 October 2020 on the Methodology for Calculating the Value of Lost Load, the Cost of New Entry, and the Reliability Standard.

AEMO, Australian Energy Market Operator, 2019. 2019 Electricity Statement of Opportunities. Report.

AESO, Alberta Electric System Operator, 2017. Resource Adequacy Criteria Overview and Alberta Historical Performance, Adequacy & Demand Curve Workgroup. Technical presentation.

Amelin, M., 2009. Comparison of Capacity Credit Calculation Methods for Conventional Power Plants and Wind Power. IEEE Transactions on Power Systems, vol. 24, iss. 2, pp. 685-691.

Andreis, L., Flora, M., Fontini, F., Vargiolu, T., 2020. Pricing Reliability Options under Different Electricity Price Regimes. Energy Economics, vol. 87, art. 104705.

Batlle, C., Mastropietro, P., Rodilla, P. and Pérez-Arriaga, I.J., 2015. The System Adequacy Problem: Lessons Learned from the American Continent. In Hancher, L., Houteclocque, A. and Sadowska, M. (Eds.), Capacity Mechanisms in the EU Energy Market: Law, Policy and Economics. Oxford University Press. ISBN: 9780198749257.

Bidwell, M., 2005. Reliability Options: A Market-oriented Approach to Long-term Adequacy. Electricity Journal, vol. 18, iss. 5, pp. 11-25.

Billinton, R., Allan, R. N., 1994. Reliability Evaluation of Power Systems. Springer, ISBN: 9780306452598.

Bothwell, C. and Hobbs, B., 2017. Crediting Wind and Solar Renewables in Electricity Capacity Markets: The Effects of Alternative Definitions upon Market Efficiency, The Energy Journal, vol. 38, pp. 173-188.

BRC, Banco de la República de Colombia, 2021. Tasa Representativa del Mercado (TRM – Peso por dólar). Data published online in https://www.banrep.gov.co/es/estadisticas/trm.

Bublitz, A., Keles, D., Zimmermann, F., Fraunholz, C., Fichtner, W., 2019. A Survey on Electricity Market Design: Insights from Theory and Real-World Implementations of Capacity Remuneration Mechanisms. Energy Economics, vol. 80, pp. 1059-1078.

CAISO, California ISO, 2019. Today's Outlook for July 16th 2019. Data published online in http://www.caiso.com/TodaysOutlook/Pages/default.aspx.

CAISO, California ISO, 2020a. Today's Outlook for August 15th 2020. Data published online in http://www.caiso.com/TodaysOutlook/Pages/default.aspx.

CAISO, California ISO, 2020b. Real-Time Daily Market Watch Report for August 15th 2020. Data published online in http://www.caiso.com/Documents/Real-TimeDailyMarketWatchAug15-2020.html.

CAISO, California ISO, 2021. Root Cause Analysis: Mid-August Extreme Heat Wave, Final. Technical report

CPUC, California Public Utilities Commission, 2017. Decision Adopting Local and Flexible Capacity Obligations for 2018 and Refining the Resource Adequacy Program. Decision 17-06-027.

CPUC, California Public Utilities Commission, 2014. Effective Load Carrying Capacity and Qualifying Capacity Calculation Methodology for Wind and Solar Resources. Working document.

Cramton, P., Ockenfels, A., Stoft, S., 2013. Capacity Market Fundamentals. Economics of Energy & Environmental Policy, vol. 2, iss. 2, pp. 27-46.

CSMEM, Comité de Seguimiento del Mercado Mayorista de Energía Eléctrica, 2016. Lecciones del Niño para gestionar el riesgo inherente al cargo por confiabilidad. Informe no. 110.

D'Annunzio, C., Santoso, S., 2008. Noniterative Method to Approximate the Effective Load Carrying Capability of a Wind Plant. IEEE Transactions on Energy Conversion, vol. 23, iss. 2, pp. 544-550.

DOE, Department of Energy of the United States, 2016. Demand Response and Energy Storage Integration Study. Report.

EIA, Energy Information Administration, 2017. California Wholesale Electricity Prices Are Higher at the Beginning and at the End of the Day. https://www.eia.gov/todayinenergy/detail.php?id=32172. Visited on 03/04/2021.

Elia, 2016. Adequacy Study for Belgium: The Need for Strategic Reserve for Winter 2017-18. Report.

Elia, 2019a. Overview of Belgian CRM Design: Introduction Note. Report

Elia, 2019b. CRM Design Note: Derating Factors. Report.

Ensslin, C., Milligan, M., Holttinen, H., O'Malley, M., Keane, A., 2008. Current Methods to Calculate Capacity Credit of Wind Power: IEA Collaboration. IEEE Power and Energy Society 2008 General Meeting: Conversion and Delivery of Electrical Energy in the 21st Century, PES.

EC, European Commission, 2019. Regulation (EU) 2019/943 of the European Parliament and of the Council of 5 June 2019 on the Internal Market for Electricity.

EC, European Commission, 2016a. Identification of Appropiate Generation and System Adequacy Standards for the Internal Electricity Market. Report.

EC, European Commission, 2016b. Commission Staff Working Document Accompanying the Final Report of the Sector Inquiry on Capacity Mechanisms. Document SWD(2016) 385 final.

Energy + Environmental Economics, 2019. Long-Run Resource Adequacy under Deep Decarbonisation Pathways for California. Report.

ENTSOe, European Network of Transmission System Operators for Electricity, 2019. Mid-term Adequacy Forecast Executive Summary. Report.

ENTSOe, European Network of Transmission System Operators for Electricity, 2020. Proposal for a European Resource Adequacy Assessment Methodology. Stakeholder webinar presentation.

EPRI, Electric Power Research Institute, 2021. Exploring the Impacts of Extreme Events, Natural Gas Fuel and Other Contingencies on Resource Adequacy. Technical report.

ERCOT, 2021. Review of February 2021 Extreme Cold Weather Event – ERCOT Presentation. Technical presentation.

Faria, E., Barroso, L. A., Kelman, R., Granville, S., Pereira, M. V., 2009. Allocation of Firm-Energy Rights Among Hydro Plants: An Aumann–Shapley Approach, IEEE Transactions on Power Systems, vol. 24, iss. 2, pp. 541-551.

Freeman, G. M., Apt, J., Moura, J., 2020. What Causes Natural Gas Fuel Shortages at U.S. Power Plants? Energy Policy, vol. 147, art. 111805.

Garver, L., 1966. Effective Load Carrying Capability of Generation Units. IEEE Transactions on Power Apparatus and Systems, vol. 85, iss.8, pp. 910-919.

Hasche, B., Keane, A., O'Malley, M., 2011. Capacity Value of Wind Power, Calculation, and Data Requirements: The Irish Power System Case. IEEE Transactions on Power Systems, vol. 26, iss. 1, pp. 420-430.

IAEA, International Atomic Energy Agency, 1984. “Chapter 7. Generating System Reliability”, Expansion planning for electrical generating systems, A guidebook. International Atomic Energy Agency Technical report no. 241.

IRENA, International Renewable Energy Agency, 2019. Redesigning Capacity Markets: Innovation Landscape Brief. Report.

I-SEM, Integrated Single Electricity Market, 2018. I-SEM Capacity Mechanism: Consultation on Additions and Modifications to the Capacity Requirement & De-rating Factor Calculation Methodology. Technical report.

ISO New England, 2016. Section III Market Rule 1: Standard Market Design – Section 13.

Joskow, P., 2008. Capacity Payments in Imperfect Electricity Markets: Need and Design. Utilities Policy, vol. 16, iss. 3, pp. 159-170.

Joskow, P., 2020. California's Blackouts, Near Blackouts, and Fires. Technical presentation.

Lambin, X., 2020. Integration of Demand Response in Electricity Market Capacity Mechanisms. Utilities Policy, vol. 64, art. 101033.

Lueken, R., Apt, J., Sowell, F., 2015. Robust Resource Adequacy Planning in the Face Of Coal Retirements. Energy Policy, vol. 88, pp. 371-388.

Mastropietro, P., Rodilla, P., Escobar Rangel, L., Batlle, C., 2020. Reforming the Colombian Electricity Market for an Efficient Integration of Renewables: A Proposal. Energy Policy, vol. 139, art. 111346.

Mastropietro, P., Rodilla, P., Batlle, C., 2019. De-rating of Wind and Solar Resources in Capacity Mechanisms: A Review of International Experiences. Renewable and Sustainable Energy Reviews, vol. 112, pp. 253-262.

Mastropietro, P., Fontini, F., Rodilla, P., Batlle, C., 2018. The Italian Capacity Remuneration Mechanism: Critical Review and Open Questions. Energy Policy, vol. 123, pp. 659-669.

Mastropietro, P., Rodilla, P., Batlle, C., 2017. Performance Incentives in Capacity Mechanisms: Conceptual Considerations and Empirical Evidence. Economics of Energy & Environmental Policy, vol. 6, iss. 1, pp. 149-163.

MITEI, Massachusetts Institute of Technology Energy Initiative, 2016. Utility of the Future: An MIT Energy Initiative Response to an Industry in Transition. Report developed in collaboration with IIT Comillas.

Moreno, R., L.A. Barroso, B. Bezerra, S. Mocarquer, H. Rudnick. Auction Approaches of Long-term Contracts to Ensure Generation Investment in Electricity Markets: Lessons from the Brazilian and Chilean Experiences. Energy Policy. vol. 38, iss. 10, pp. 5758-5769.

National Grid, 2017. Duration-Limited Storage De-Rating Factor Assessment – Final Report.

National Grid, 2019a. Electricity Capacity Report.

National Grid, 2019b. De-rating Factor Methodology for Renewables Participation in the Capacity Market. Report.

NERC, North American Electric Reliability Corporation, 2012. Pilot Probabilistic Assessment. Report.

NERC, North American Electric Reliability Corporation, 2018. Probabilistic Adequacy and Measures, Technical Reference Report Final.

Neuhoff, K., De Vries, L., 2004. Insufficient Incentives for Investment in Electricity Generations. Utilities Policy, vol. 12, iss. 4, pp. 253-267.

Newbery, D., 2016. Missing Money and Missing Markets: Reliability, Capacity Auctions and Interconnectors. Energy Policy, vol. 94, pp. 401-410.

NWS, National Weather Service, 2021. Cold & Warm Episodes by Season. https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php. Visited on 03/04/2021.

NYISO, New York ISO, 2019. Installed Capacity Manual.

Ofgem, Office of Gas and Electricity Markets, 2021. Annual Report on the Operation of the Capacity Market in 2019/20.

Ozbafli, A., Jenkins, G., 2016. Estimating the Willingness to Pay for Reliable Electricity Supply: A Choice Experiment Study. Energy Economics, vol. 56, pp. 443-452.

Pérez-Arriaga, I.J., 1994. Principios económicos marginalistas en los sistemas de energía eléctrica (in Spanish). Technical Report IIT-93-044.

Pérez-Arriaga, I.J. and Meseguer, C., 1997. Wholesale Marginal Prices in Competitive Generation Markets. IEEE Transactions on Power Systems, vol. 12, iss. 2, pp. 710-717.

Petitet, M., Finon, D., Janssen, T., 2017. Capacity Adequacy in Power Markets Facing Energy Transition: A Comparison of Scarcity Pricing and Capacity Mechanism. Energy Policy, vol. 103, pp. 30-46.

PJM, 2018. PJM Cold Snap Performance. Technical report.

PJM, 2019. Manual 21 – Rules and Procedures for Determination of Generating Capability.

Reuters, 2019. Hot weather cuts French, German nuclear power output. July 25, 2019. Available at https://www.reuters.com/article/us-france-electricity-heatwave-idUSKCN1UK0HR

Rodilla, P., García-González, J., Baíllo, A., Cerisola, S., Batlle, C., 2015. Hydro Resource Management, Risk Aversion and Equilibrium in an Incomplete Electricity Market Setting. Energy Economics, vol. 51, pp. 365-382.

RTE, Réseau de transport d'électricité, 2014. French Capacity Market - Report Accompanying the Draft Rules. Technical report.

Schweppe, F.C., Caramanis, M.C., Tabors, R.D., Bohn, R.E., 1988. Spot Pricing of Electricity. Springer, ISBN: 978-1-4613-1683-1.

SEMC, Single Electricity Market Committee, 2018. Capacity Remuneration Mechanism (CRM) 2019/20 T-1 Capacity Auction Parameters and Enduring De-Rating Metodology. Decision Paper SEM-18-030.

SEPA, Smart Electric Power Alliance, 2019. 2019 Utility Demand Response Market Snapshot. Presentation report.

Söder, L., Tómasson, E., Estanqueiro, A., Flynn, D., Hodge, B. M., Kiviluoma, J., Korpås, M., Neau, E., Couto, A., Pudjianto, D., Strbac, G., Burke, D., Gómez, T., Das, K., Cutululis, N. A., Van Hertem, D., Höschle, H., Matevosyan, J., von Roon, S., Carlini, E. M., Caprabianca, M., de Vries, L., 2020. Review of Wind Generation within Adequacy Calculations and Capacity Markets for Different Power Systems. Renewable and Sustainable Energy Reviews, vol. 119, art. 109540.

Terna, 2018a. Disciplina del sistema di remuneazione della disponibilità di capacità di energia elettrica - Fase di prima attuazione. Consultation document, released in March 2018.

Terna, 2018b. Definizione dei parametri per il calcolo della CDP. Annex 3: 'Disciplina del sistema di remunerazione della disponibilità di capacità di energia elettrica'.

Vázquez, C., Rivier, M., Pérez-Arriaga, I.J., 2002. A Market Approach to Long-Term Security of Supply. IEEE Transactions on Power Systems, vol. 17, iss. 2, pp. 349-357.

XM, 2021. Colombia Historic Electricity Wholesale Market Price, available online at http://portalbissrs.xm.com.co/trpr/Paginas/Historicos/Historicos.aspx.

Zachary, S. and Dent, C. J., 2012. Probability Theory of Capacity Value of Additional Generation. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, vol. 226, iss. 1, pp. 33-43.

Zachary, S., Wilson, A., Dent, C., 2019. The Integration of Variable Generation and Storage into Electricity Capacity Markets. Working paper.

ANNEX I

This section demonstrates the link between the reliability metric (which is used to set the reliability target) and the methodology to measure the firm supply. This section also explains how this contribution should be remunerated.

In order to do this, it is first formulated and solved the benchmark optimization problem which is a stylised version of the ideal central planner problem with an adequacy constraint. This adequacy constraint is expressed by means of a reliability metric (RM). Then it is formulated and solved the problem of the individual agents, where the ingredients of interest are their participation in the energy and in the capacity market. By comparing the optimality conditions of both problems we draw conclusions about what should be remunerated in a capacity mechanism and get to the formulation of the firm supply.

The formulation of both problems will be based on the stylised models described in Pérez Arriaga and Meseguer (1997).

Centralised problem

This subsection uses a stylised version of the regulator's model presented by Pérez-Arriaga and Meseguer (1997). In this problem, the objective is to maximise the net social benefit, NSB, related to the supply and consumption of electricity. This NSB is represented by the following expression:

$$\underset{Q\,K}{\mathrm{Max}}\ NSB = U\left(Q_1+Q_2+Q_3+\ldots Q_n\right) - C\left(Q_1,Q_2,Q_3,\ldots Q_n\right) - I(K_1,K_2,K_3,\ldots K_n) \qquad (1)$$

Where:

- Q_i, represents the production of each generating plant, i= 1,2,3,4,…n,
- $Q_1+Q_2+Q_3+\ldots Q_n$ represents the total production and therefore, the supplied demand.
- $U(Q_1+Q_2+Q_3+\ldots Q_n)$ is the demand utility function, which depends on the power consumed by the demand. This demand utility function is assumed to be strictly increasing and concave.
- $C(Q_1, Q_2, Q_3,\ldots Q_n)$ is the generation cost function, which aggregates all the generation units in the system, and also depends on the power consumed, or power produced, Q_i, of each generation unit, i= 1,2,3,4,…n. In this case, the function is assumed to be strictly increasing and convex.

- $I(K_1, K_2, K_3, \ldots K_n)$ is the investment cost function which depends on the amount of investment, K_i, of each generation unit (i). In this case, the function is assumed to be strictly increasing and convex.

This stylised representation only considers two constraints:

$$Q_i \leq K_i \quad \alpha_i \tag{2}$$

$$RM(K_1, K_2, K_3, \ldots K_n) \geq RT \quad \beta \tag{3}$$

The first constraint represents the upper limit of the power produced by each generation unit i, which corresponds to the installed capacity of that unit, K_i.

On the other hand, the second constraint forces the reliability metric RM, which is assumed to be dependent only on the mix $RM = RM(K_1, K_2, K_3, \ldots K_n)$, to fulfil a certain reliability target RT, which is set as a parameter. The RM is assumed to be strictly decreasing and convex.

We have therefore discarded other operation constraints such as ramps, minimum power outputs etc., for simplicity.

In order to obtain the first-order necessary conditions we formulate the Lagrangian function, *L*, and compute its first partial derivatives with respect to the decision variables.

$$L(Q_1, Q_2, Q_3, \ldots Q_n, Q_i, K_1, K_2, K_3, \ldots K_n, K_i, \alpha_i, \beta) = U(Q_1, Q_2, Q_3, \ldots Q_n) - C(Q_1, Q_2, Q_3, \ldots Q_n)$$

$$- I(K_1, K_2, K_3, \ldots K_n) + \textstyle\sum_1^n (Q_i - K_i) \cdot \alpha_i + (RF(K_1, K_2, K_3, \ldots K_n) - RO) \cdot \beta \tag{4}$$

If we compute the first partial derivative of this expression with respect to the decision variable K_i, which is the installed capacity of unit i, we obtain the following expression:

$$\frac{\partial L}{\partial K_i} = -\frac{\partial I(K_1, K_2, K_3, \ldots K_n)}{\partial K_i} - \alpha_i + \frac{\partial RF(K_1, K_2, K_3, \ldots K_n)}{\partial K_i} \beta = 0 \tag{5}$$

The first two terms in equation 5 represent the classical equilibrium between the short term savings (reduction in the value of the objective function by α_i) and the increase in long term costs (increase in the value of the objective function by the increase in investment costs) in the optimality point. The additional term of equation 5 will only be present if the constraint described by equation 3, regarding the adequacy of the system, is binding, which will alter the equilibrium described beforehand.

Decentralised problem

This subsection uses a stylised version of the generators viewpoint of the competitive market model presented by Pérez-Arriaga and Meseguer (1997). In contrast with the centralised problem, the objective of each generation unit, i, is to maximise its own profit, P_i, which is represented by the following expression, where it is assumed that there is both a spot market and a capacity market:

$$\max_{Q_i\ K_i} P_i = SMP_{Q,i} \cdot Q_i + CMP_{K,i} \cdot K_i - C_i(Q_i) - I_i(K_i) \quad (6)$$

Where:

- $SMP_{Q,i}$ is the spot market price perceived by generation unit i, which when multiplied by Q_i results in the spot market revenues.
- $CMP_{K,i}$ is the capacity market price perceived by generation unit i, which when multiplied by K_i results in the capacity market revenues.
- $C_i(Q_i)$ and $I_i(K_i)$ are the generation cost function and the investment cost function of generation unit i, respectively, with the same characteristics as the centralised problem.

The only constraint present in this problem is the following:

$$Q_i \leq K_i \qquad \alpha_i \quad (7)$$

This constraint is equivalent to the first constraint in the centralised problem. The second constraint found in the centralised problem is only present through the regulators perspective and is therefore only translated through the capacity market price in the objective function in this decentralised problem.

In order to obtain the first-order necessary conditions we formulate now the Lagrangian function of this second problem, $\mathcal{L}$, and compute its first partial derivatives with respect to the decision variables.

$$L(Q_i, K_i, \alpha_i) = SMP_{Q,i} \cdot Q_i + CMP_{K,i} \cdot K_i - C_i(Q_i) - I_i(K_i) + (Q_i - K_i) \cdot \alpha_i \quad (8)$$

When computing the first partial derivative of this expression with respect to the decision variable K_i we obtain the following optimality condition:

$$\frac{\partial L}{\partial K_i} = CMP_{K,i} - \frac{\partial I(K_1, K_2, K_3, \ldots K_n)}{\partial K_i} - \alpha_i = 0 \quad (9)$$

Equation 9 is very similar to equation 5, without the global constraint described by equation 3, which is only present in the centralised problem, but with the additional term $CMP_{K,i}$.

Unification of both problems

Comparing equation 9 to equation 5 we obtain the following expression:

$$CMP_{K,i} = \frac{\partial RM(K_1,K_2,K_3,\ldots K_n)}{\partial K_i} \cdot \beta \tag{10}$$

Which leads to the following remuneration in the capacity market:

$$CMP_{K,i} \cdot K_i = \frac{\partial RM(K_1,K_2,K_3,\ldots K_n)}{\partial K_i} \cdot K_i \cdot \beta \tag{11}$$

This allows us to draw several conclusions:

1. The firm supply depends on the marginal contribution to the reliability metric $RM(K_1, K_2, K_3, \ldots K_n)$.
2. $CMP_{K,i}$, expressed in equation 10, represents the price of 1MW of unit i. However, β is the price for the firm supply, which is a value that could be obtained through competitive means, such as a capacity auction.
3. The installed capacity of generation unit i, K_i, multiplied by the variation of RM with respect to it, in the optimality point, represents the firm supply of generation technology i.