Paulo
Brito
Pereira

Future-proof Design of Capacity Mechanisms during the Energy Transition

COMILLAS
UNIVERSIDAD PONTIFICIA

Escuela Internacional de Doctorado
Programa de Doctorado en Energía Eléctrica

Future-proof Design of Capacity Mechanisms during the Energy Transition

Reliability Metrics, Firm Supply Calculation, Cost Allocation and Demand Participation

Author:	Paulo Brito Pereira
Supervisors:	Pablo Rodilla Rodríguez
	Paolo Mastropietro
Tutor:	Carlos Batlle López

MADRID | November 2025

ACKNOWLEDGEMENTS

Although the process of obtaining a PhD can oftentimes be a solitary endeavour, I have been extremely fortunate to have found myself surrounded with a great deal of fantastic people to aid me in this voyage. The first people that bear mentioning in this regard are my thesis supervisors, Dr. Pablo Rodilla and Dr. Paolo Mastropietro, which have constantly aided and guided me, and whom without I would not have embarked on this journey. I would also like to thank Diana Navarrete, Dr. Geovanny Marulanda and Dr. Luis Montero, which have, with their sarcasm and witty humour, helped with the daily struggles, Claudia Dominguez, which has been a constant source of ideas for the thesis, and both Dr. Pedro de Otaola and Antonio Vázquez, for always adding a dose of realism and keeping my feet on the ground. Finally, I would like to thank my parents, which are a constant source of support and my biggest advocates.

SUMMARY

Capacity Remuneration Mechanisms (CRMs) have become a key regulatory instrument in modern power systems, through which regulators and policymakers aim to guide the expansion of the resource mix towards the desired level of resource adequacy. However, the ongoing energy transition is reshaping the resource adequacy problem and the solutions to address it. This PhD dissertation focuses on revising several CRM design elements in this context, specifically the resource adequacy metric used to analyse scarcity conditions, the methodology for calculating firm supply for different resources, and the alternatives for enabling demand resource participation.

The research presented in this document is based on international trends, which were pivotal in determining the topics to address, as well as on model-based analyses, which were used to study the effects of the different reform proposals. The conclusions of this PhD thesis for each of the aforementioned topics can be summarised as follows.

- The most widespread resource adequacy metrics focus on counting the number of loss-of-load events within a given timeframe. However, in the context of the energy transition and increasing demand elasticity, these metrics are becoming obsolete as they cannot capture the severity of each event and therefore cannot distinguish between events with differing amounts of non-served demand. In order to capture these differences and anticipate the rise in electricity demand elasticity, resource adequacy metrics must incorporate the price dimension. Additionally, in light of the increase in extreme weather conditions due to climate change, it would be advisable for resource adequacy metrics to focus on extreme scenarios using statistical measures such as Conditional Value at Risk (CVaR) rather than treating all events equally using the mean.

  Although this proposed resource adequacy metric provides a more accurate representation of scarcity conditions than current practices, regulators may perceive that a single metric cannot capture all facets of potential scarcity conditions. In this case, multiple resource adequacy metrics could be used simultaneously, although this would complicate the process of calculating firm supply. As an alternative, a composite resource adequacy metric could be used, combining several metrics into one and avoiding these complications while still allowing for consideration of several resource adequacy metrics.

- Firm supply recognises how different resources can contribute to overcoming scarcity conditions and improving resource adequacy within a system. Current firm supply calculation methodologies are becoming increasingly inadequate and are being reformed in many power systems. This thesis presents several general recommendations on how firm supply should be calculated in modern CRMs. These can be summarised as follows: i) regulators should use the same resource adequacy metric both to calculate firm supply as for the resource adequacy assessment to help ensure that the desired target is met; ii) firm supply for each resource must be calculated by considering it in conjunction with the rest of the system and with a forward-looking approach to predict how it will perform in

future scarcity conditions, taking into account the system's actual dispatch; iii) firm supply should be calculated according to the marginal contribution of resources to reliability instead of analysing the mean contribution of similar resources, as this does not accurately reflect how new resources would improve system adequacy.

If regulators opt to establish multiple reliability standards, the calculation of firm supply increases in complexity as each standard requires an associated firm supply value to reflect the resources' contribution to it. However, this problem does not arise with a composite reliability standard. This results in a single reliability standard, meaning that only one firm supply value needs to be calculated.

Due to the forward-looking nature of the firm supply calculation process, the effect of the CRM itself on the resource mix must be considered. In this case, establishing a firm supply value for each resource as a qualifying element for a CRM auction will affect the composition of the future resource mix in the system. Consequently, the firm supply of resources could differ significantly from expectations if the resource mix differs substantially from the initial estimate.

- The rising concern for resource adequacy in power systems, coupled with the growing importance of CRMs in electricity market design, has made it increasingly necessary to incorporate demand resources more effectively. Firstly, the increased implementation of CRMs has led to increased costs, which electricity consumers must bear. Therefore, efficient cost allocation methodologies for CRM costs must be designed. CRM costs should be assigned according to each consumer's marginal impact on the system's resource adequacy. This can be estimated by calculating hourly charges according to the scarcity risk of each hour, then multiplying this figure by each consumer's historical consumption.

  On the other hand, to address concerns about resource adequacy most effectively, all potential resources must be adequately integrated into CRMs. This includes demand resources, which are being introduced progressively into different electricity markets. However, integrating demand resources into CRMs is challenging because they demand firm supply and offer it back to the system as demand response. This thesis details the various alternatives for integrating demand response for both implicit and explicit participation. Explicit participation is further subdivided into demand-side participation (i.e. defining the required amount of firm supply demanded) and supply-side participation (i.e. offering demand response). Supply-side participation can lead to a double remuneration problem whereby the demand resource can reduce its associated payments to the CRM while also being paid for participating as demand response. To avoid this problem, demand-side participation should be used.

RESUMEN

Los mecanismos de remuneración de capacidad (CRMs, por sus siglas en inglés) se han convertido en un instrumento regulatorio esencial en los sistemas eléctricos modernos, por el cual reguladores y legisladores intentan guiar la expansión del mix de recursos para lograr un nivel de fiabilidad deseado. Sin embargo, el proceso de transición energética está transformando el problema de la fiabilidad en los sistemas eléctricos y las soluciones para ponerle remedio. En este contexto, esta tesis doctoral se centra en la revisión de varios elementos de diseño de los CRMs, más específicamente la elección de la métrica de fiabilidad para analizar las situaciones de escasez, la metodología para el cálculo del suministro firme de los diferentes recursos, y las diferentes alternativas para permitir la participación de los recursos de demanda.

La investigación presentada en este documento se basa en las tendencias internacionales, que ha sido esenciales para determinar los temas a tratar, y análisis basados en modelos para estudiar los efectos de las diferentes propuestas de reforma presentadas. Las conclusiones de esta tesis doctoral para cada uno de los temas mencionados anteriormente pueden resumirse en los puntos siguientes.

- Las métricas de fiabilidad más extendidas se centran en contar el número de eventos de pérdida de carga en un periodo temporal. Sin embargo, en este contexto de transición energética y creciente elasticidad de la demanda eléctrica, estas métricas de fiabilidad se están convirtiendo en obsoletas, ya que no pueden capturar la severidad de cada evento de pérdida de carga y por ello no son capaces de distinguir entre eventos en los que haya habido una cantidad diferente de demanda que no haya sido servida. Para poder capturar estas diferencias y para anticiparse a la creciente elasticidad de la demanda eléctrica, las métricas de fiabilidad deberían incorporar la dimensión de precio. Adicionalmente, a la luz del aumento de eventos climáticos extremos debido al cambio climático, sería recomendable que las métricas de fiabilidad se centrasen en los escenarios extremos mediante el uso de medidas estadísticas como el Valor en Riesgo Condicional (CVaR, por sus siglas en inglés), en vez de dar igual importancia a todos los escenarios como hace la media.

  Aunque esta métrica de fiabilidad propuesta presenta una mejor representación de las situaciones de escasez que las prácticas actuales, los reguladores podrían percibir que una única métrica no puede capturar todas las posibles facetas de las potenciales situaciones de escasez. En este caso, podrían emplearse varias métricas de fiabilidad simultáneamente, aunque esto complicaría el proceso del cálculo del suministro firme de los recursos. Alternativamente, podría emplearse una métrica de fiabilidad compuesta, en la que se combinan varias métricas de fiabilidad en una y evitando estos problemas pero manteniendo todavía la posibilidad de considerar varias métricas de fiabilidad.

- El suministro firme representa el reconocimiento de como los diferentes recursos son capaces de contribuir a solventar las situaciones de escasez y mejorar la fiabilidad del

sistema. Las metodologías actuales para el cálculo del suministro firme se están volviendo cada vez más inadecuadas y se están reformando en muchos sistemas eléctricos. Esta tesis doctoral presenta una serie de recomendaciones sobre cómo se debería calcular el suministro firme en los mecanismos de capacidad modernos. Estos se pueden resumir en: i) la métrica de fiabilidad empleada para el cálculo del suministro firme debe ser la misma que se utiliza para el análisis de fiabilidad del sistema, para facilitar que se consiga el nivel de fiabilidad deseado; ii) el suministro firme de cada recurso se debe calcular considerándolo en conjunto con el resto del sistema y con una visión prospectiva, para predecir cómo ese recurso desempeñará durante las situaciones de escasez reales del futuro, teniendo en cuenta el despacho del sistema en su conjunto; iii) el suministro firme debe calcularse de acuerdo con la contribución marginal de los recursos a mejorar la fiabilidad, en vez de analizar la contribución media de recursos similares, ya que esto último no considera de forma adecuada como los recursos serían capaces de mejorar la fiabilidad del sistema.

Si los reguladores optan por establecer varios estándares de fiabilidad (objetivos de fiabilidad), el cálculo del suministro firme se vuelve más complejo, ya que cada estándar de fiabilidad requiere de un valor de suministro firme asociado para reflejar la contribución de cada recurso al mismo. Sin embargo, este problema no aparece cuando se utilizan estándares de fiabilidad compuestos. Esto se debe a que existe un único estándar de fiabilidad, por lo que sólo ha de calcularse un único valor de suministro firme.

Dada la visión prospectiva que debe seguirse a la hora de calcular el suministro firme, el efecto del CRM en si mismo, debe ser considerado. En este caso, establecer un valor de suministro firme para cada recurso como elemento de cualificación para una subasta de un CRM afectará en qué recursos terminan siendo parte del mix de recursos del sistema futuro. Por ello, el suministro firme de los diferentes recursos podrían ser substancialmente diferentes de lo esperado, y del suministro firme calculado exante, si el mix de recursos es significativamente diferente de lo que se estimó inicialmente.

- La creciente preocupación por la fiabilidad de los sistemas eléctricos, unido al aumento en la implementación de CRMs, ha llevado a una creciente necesidad de incluir los recursos de demanda de manera más efectiva. En primer lugar, el aumento de la implementación de los CRMs ha llevado a un aumento de sus costes, que deben ser soportados por los consumidores eléctricos. Por ello, es importante diseñar metodologías de asignación de costes de CRMs que sean eficientes. En este caso, los costes deberían ser asignados de acuerdo con el impacto marginal de cada consumidor en la fiabilidad del sistema. Esto puede ser estimado calculando cargas horarias de acuerdo con el riesgo de escasez en cada una de ellas y multiplicándolo con el consumo histórico de cada consumidor.

  Por otro lado, para abordar las preocupaciones sobre la fiabilidad de manera más efectiva, todos los potenciales recursos deben ser integrados adecuadamente en los CRMs. Esto incluye a los recursos de demanda, que se están incluyendo progresivamente en los diferentes mercados eléctricos. Sin embargo, integrar los recursos de demanda en los CRMs es un reto, ya que demandan suministro firme pero pueden ofrecerlo de vuelta al sistema en forma de respuesta a la demanda. Esta tesis detalla las diferentes alternativas para integrar los recursos de demanda, tanto mediante participación implícita y la participación explícita. La participación explícita se subdivide en participación por el lado de la demanda (i.e. definiendo la cantidad de suministro firme demandado) y en el de la oferta (i.e. ofreciendo respuesta de la demanda). La participación por el lado de la oferta puede llevar a un problema de doble remuneración, por el cual el recurso de la demanda puede reducir sus pagos asociados al CRM y también ser remunerado por su participación como respuesta de la demanda. Para evitar este problema, los recursos de demanda deberían participar por el lado de la demanda.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACER	Agency for the Cooperation of Energy Regulators (European Union)
AEMC	Australian Energy Market Commission
AEMO	Australian Energy Market Operator
AEEGSI	*Autorità per l'Energia Elettrica il Gas e il Sistema Idrico* (Italy)
ARERA	*Autorità di Regolazione per Energia Reti e Ambiente* (Italy)
BMWK	*Bundesministerium für Wirtschaft und Klimaschutz* (Germany)
BOG	Belgian Official Gazette
BRC	*Banco de la República de Colombia* (Colombia)
CAISO	California ISO
CCGT	Combined Cycle Gas Turbine
CSMEM	*Comité de Seguimiento del Mercado Mayorista de Energía Eléctrica* (Colombia)
CNMC	*Comisión Nacional de los Mercados y la Competencia* (Spain)
CPUC	California Public Utilities Commission
CRM	Capacity Remuneration Mechanism
CVaR	Conditional Value at Risk
DESNZ	Department for Energy Security & Net Zero
EC	European Commission
EFOR	Equivalent Forced Outage Rate
EIA	Energy Information Administration (United States)
ELCC	Effective Load Carrying Capability
EMRS	Electricity Market Reform Settlement Limited
ENS	Energy Not Served
ENTSO-E	European Network of Transmission System Operator for Electricity
EPE	*Empresa de Pesquisa Energética* (Brazil)

EPRI	Electric Power Research Institute
ERAA	European Resource Adequacy Assessment
ERAWA	Economic Regulation Authority Western Australia
ERCOT	Electric Reliability Council of Texas
ESIG	Energy Systems Integration Group
EU	European Union
EUPHEMIA	EU Pan-European Hybrid Electricity Market Integration Algorithm
EVA	Economic Viability Assessment
FERC	Federal Energy Regulatory Commission (United States)
IAEA	International Atomic Energy Agency
IPCC	Intergovernmental Panel on Climate Change
I-SEM	Integrated Single Electricity Market
ISO	Independent System Operator
LOLE	Loss-Of-Load Expectation
LOLEv	Loss-Of-Load Events
LOLH	Loss-Of-Load Hours
LOLP	Loss-Of-Load Probability
MISO	Midcontinent Independent System Operator
MIT	Massachusetts Institute of Technology
MITEI	Massachusetts Institute of Technology Energy Initiative
MITERD	*Ministerio para la Transición Ecológica y el Reto Demográfico* (Spain)
MTR	Mean Time to Recovery
NERC	North American Electric Reliability Corporation
NPCC	Northwest Power and Conservation Council
NRAA	National Resource Adequacy Assessment
NREL	National Renewable Energy Laboratory
NSE	Non-Served Energy
NWS	National Weather Service
NYISO	New York ISO
ODI	Overseas Development Institute

PJM	Pennsylvania New Jersey Maryland Interconnection
RES-E	Renewable Energy Sources for Electricity
RM	Reserve margin
RTE	*Réseau de Transport d'Électricité* (France)
SEMC	Single Electricity Market Committee
SEPA	Smart Electric Power Alliance
UC	Unit Commitment
UE	Unserved Energy
VaR	Value at Risk
VOLL	Value-of-Lost-Load
WTP	Willingness to Pay

1. INTRODUCTION

1.1. Context

One of the foundational ideas that conceptually prompted the wave of power sector liberalisations in the last decades of the twentieth century was that short-term market prices would drive not only an efficient operation of the system but also efficient investment decisions, to be undertaken by market agents, which would ultimately result in an optimal expansion of the resource mix and guarantee the adequacy[1] of the system (Caramanis, 1982; Caramanis et al., 1982; Schweppe et al., 1988).

However, during their liberalization process, some power systems also introduced regulatory mechanisms that complemented short-term market prices, whose objective was to attract sufficient generation to supply a growing demand. This was especially true in Latin America, where several countries implemented different forms of adequacy mechanisms from the very beginning of their liberalisation process or shortly after, such as Chile, Colombia, Peru, Argentina and Brazil (Estache and Rodriguez-Pardina, 1996; Maurer and Barroso, 2011; Mastropietro et al., 2014).In parallel, the "Promised Land of electricity deregulation" (Oren, 2005), in which short-term market prices are meant to be the only signal that drive investment decisions, was soon questioned by several experts also in the United States and Europe. Although some continued to argue that the market price could drive an efficient system expansion and that all that was needed was to improve the existing short-term market signals (Hogan, 2005), other experts started to highlight some market imperfections that were preventing the market price from conveying these efficient signals (Von der Fehr and Harbord, 1994; Borenstein, 2002).

Two main market imperfections were identified. The first was eventually named the "missing money" problem (Cramton and Stoft, 2006; Joskow, 2006; Joskow and Tirole, 2007; Joskow, 2008; Newberry, 2016). The lack of elasticity of electricity demand, particularly in instances of generation scarcity, and the inability of consumers to define and express in the market their

[1] Usually referred to as long-term security of supply, or more specifically "the existence of enough available generation capability, both installed and/or expected to be installed, to meet efficiently demand in the long term." (Batlle and Pérez-Arriaga, 2008; Batlle and Rodilla, 2010)

utility function, especially in terms of Value-of-Lost-Load[2] (VOLL), caused system operators to administratively establish price caps well below the theoretical whole-system VOLL, with the objective of avoiding the exercise of market power by generators, or perform "out-of-market" interventions during these instances of scarcity, which also limited market prices (Joskow, 2006; Bushnel et al., 2017). Due to these interventions, market agents would be unable to recover their investment costs, eventually leading to underinvestment in the system. This was particularly evident for peaking power plants, which expected to recover their investment costs during a reduced number of instances of generation scarcity.

The second market imperfection was termed the "missing markets" problem (Neuhoff and de Vries, 2004; Newberry, 2016; Peluchon, 2019). In this case, even though both generators and consumers are risk averse, neither can hedge it properly. Although generators require long-term risk-hedging contracts in order to make an investment decision, consumers rarely enter into electricity contracts which are longer than a year. This approach is related to the belief that adequacy could be understood as a public good (Finon and Pignon, 2008, Finon et al., 2008; Keppler, 2017) and that the state will take care to guarantee the electricity supply at affordable prices. As can be observed in Figure 1.1, most of the liquidity of long-term markets is concentrated in the first two years. The inability of generators to hedge this long-term risk eventually leads to an underinvestment from the societal optimum.

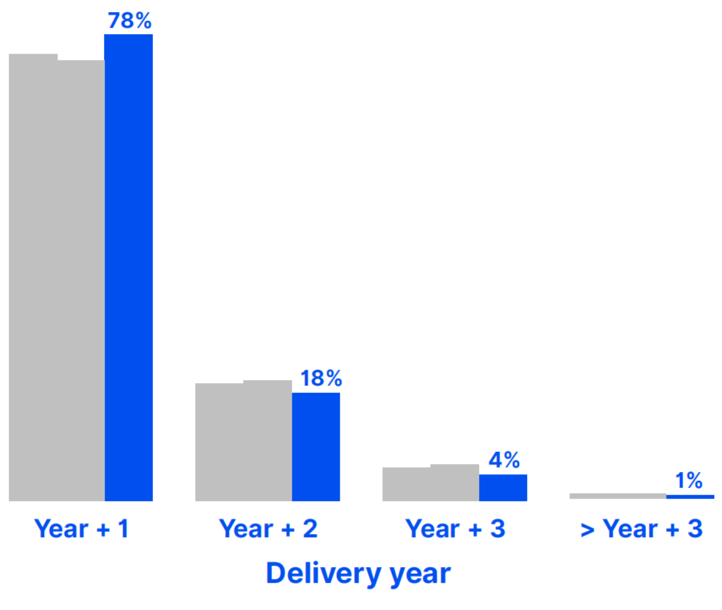

Figure 1.1: Relative shares of traded volume per year in the future for delivery in Germany – 2021-2023 (ACER, 2024a). The grey columns represent the values for both 2021 and 2022, while the columns in blue indicate the values for 2023.

[2] As Gorman (2022) succinctly expressed: "it represents society's willingness-to-pay (WTP) to avoid a power outage". Although Gorman expressed it as society's WTP, each consumer could express their own VOLL.

These two market imperfections, alongside the lumpiness[3] of power plant investment, which can be significant in small power systems (Rodilla and Batlle, 2012), led many regulators and experts to argue that additional price signals were needed to reach a desired level of adequacy. Given that the potential damage from a lack of adequacy (due to underinvestment) could significantly outweigh the risk of overinvestment (De Vries and Heijnen, 2008), this was seen as an important issue to address. Although there was some discussion over the possibility of improving the existing short-term price signals by establishing scarcity pricing and allowing extremely high prices during generation scarcity conditions (Hogan, 2013; Petitet et al., 2017), many experts proposed the introduction of specific regulatory instruments to address these concerns.

Capacity remuneration mechanisms[4] (or CRMs) are long-term regulatory instruments that are implemented to promote investments in power systems to try to achieve a desired resource adequacy level in the system by providing capacity revenues that complement short-term market revenues. CRMs not only drive investments in power systems while solving the aforementioned market imperfections, but they also do so by containing potential abuse of market power (Fabra, 2018). These mechanisms can be classified in a variety of ways, being the most general those that independently analyze the CRM design elements (Batlle et al., 2022).

In this thesis the CRMs used as reference in the different chapters will be a basic design that allows accounting for the ingredients of interest:

- A market-wide mechanism in which the firm supply is the product traded. A market-wide mechanism allows the participation of all resources that are present in the system.
- The regulator aims to ensure a precise quantity of firm supply[5] to reach a desired resource adequacy level through a centralized auction process.
- This auction process will be pay-as-cleared, i.e., all bidding participants will be paid/have to pay according to the price at which supply and demand curves meet and thus remunerating firm supply in €/MWh (or $/MWh).
- No energy contracts are involved in this type of mechanisms.
- There can be penalties in case the firm supply committed in the auction is finally not delivered.

The main objective of the CRM is to drive investment decisions so that the power system in which it is implemented reaches a desired resource adequacy level, which is usually expressed

[3] The fact that power plants present a minimum feasible installed capacity that can be quite significant.

[4] Capacity remuneration mechanisms is an expression used in the European Union to refer to resource adequacy mechanisms. They can be also referred to as capacity mechanisms or capacity markets.

[5] Firm supply refers to electricity supply that is available during scarcity conditions.

by means of a reliability standard[6]. If the regulator performs a resource adequacy assessment[7] that signals that the system will not reach the reliability standard in the future, then a capacity remuneration mechanism could be implemented to drive the expansion of the system towards a reliability-standard compliant mix. Other regulatory instruments, like renewable energy support schemes, also attempt to drive investment decisions in power systems, albeit in a more targeted fashion. If implemented simultaneously, regulators often have to decide how to resolve conflicts between the differing expansion objectives between these mechanisms (Kozlova and Overland, 2022; Kozlova et al., 2023).

Currently, CRMs are very widespread in liberalised power systems. In the United States, they have been introduced by many Independent System Operators (Tarufelli et al., 2022) and even ERCOT, the main example of an energy-only market, is considering the implementation of a regulatory instrument after the 2021 electricity crisis (Busby et al., 2021; Petitet et al., 2024). In the European Union, CRMs were initially seen as a temporary mechanism (EC, 2019); however, this view has changed and they are, as of the time of writing this thesis, considered a more permanent regulatory instrument (EC, 2024). CRMs have become commonplace in Europe, as shown in Figure 1.2, with Spain seeking to reintroduce a CRM (MITERD, 2024) and Germany discussing possible adaptations of its own (BMWK, 2024).

[6] A reliability standard, otherwise known as a reliability target or resource adequacy objective, represents a resource adequacy target that the regulator wants the system to achieve. For example, a reliability standard could be that the system should have less than three hours with loss of load per year.

[7] An analysis performed by regulators to evaluate the resource adequacy level in their system in a timeframe in the future.

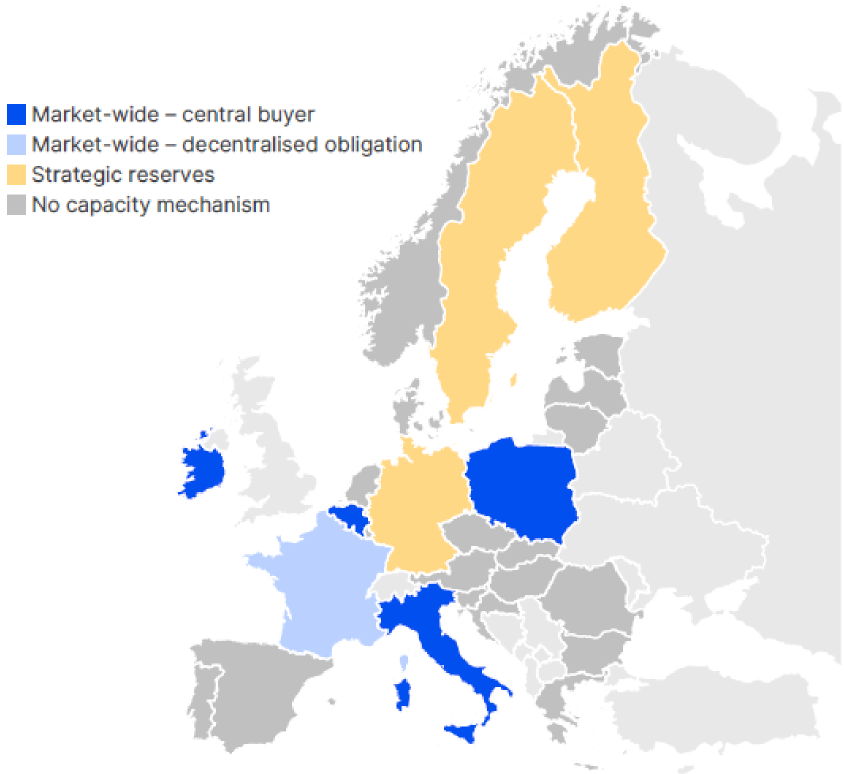

Figure 1.2: Status of Capacity Mechanisms in Europe in 2023 (ACER, 2024b)

However, power systems are now experiencing rapid changes due to the ongoing energy transition, which is causing a shift in the electricity mix of power systems, reducing the relevance of conventional thermal generation, and increasing the importance of renewable energy sources, demand response and electricity storage. Apart from this, climate change is increasing the frequency and intensity of extreme weather events (IPCC, 2023), which could impact power systems significantly (AEMO, 2019; ERCOT, 2021, ACER 2025), both in terms of generation (especially as the penetration of renewable energy sources increases) and in terms of electricity demand. Both elements are causing a change in the scarcity conditions that are usually addressed by CRMs.

These rapid changes are raising questions on whether the current design of resource adequacy assessments and capacity remuneration mechanisms are fit for purpose. This thesis focuses on three main aspects which are crucial for the correct design of CRMs in the current transition that power systems are undergoing:

- Given rapid developments in power systems, scarcity conditions are changing and there is a significant risk that resource adequacy metrics do not capture the inherent risk and damage that these new scarcity conditions could cause.
- The firm supply calculation process, i.e., the evaluation of how different resources are able to improve resource adequacy in the system, needs revision, as it was developed using tools and methodologies which made sense in conventional power systems, but which might not be appropriate for power systems resulting from the energy transition process.

- The increasing prevalence of CRMs in power systems is leading to a rise in the associated costs that consumers must bear. Therefore, proper cost allocation among consumers is becoming increasingly important in sending efficient economic signals. Correctly designing these signals is also important in ensuring that demand response resources are able to participate in CRMs on equal footing with the rest of the potential resources.

1.2. Theoretical framework

These three topics represent the three pillars of this thesis and are further analysed in this theoretical framework. Each element of the theoretical framework describes the need to improve a particular aspect of resource adequacy assessments and CRMs.

As will be described throughout the document, the three parts of the theoretical framework are also highly interrelated, and the decisions made for one part of the framework should and will be considered for the rest. For example, the choice of the resource adequacy metric should affect how firm supply is estimated and also how CRM costs are allocated among consumers. A holistic approach to resource adequacy is needed to ensure the desired level of resource adequacy, for coherence in the different design elements, and to ensure economic efficiency.

1.2.1. The need for resilient resource adequacy metrics

Electricity system operators perform resource adequacy assessments to analyse the security of supply in their power sectors. Resource adequacy assessments are probabilistic analyses that simulate the expected operation of the power system over different scenarios and time horizons and characterise the expected scarcity conditions. These assessments are based on resource adequacy metrics, which are meant to measure the scarcity conditions in the power systems over a given timeframe[8].

In the European Union, the general methodology to elaborate resource adequacy assessments was elaborated by the European Commission and ACER (ACER, 2020) and is executed in the European Resource Adequacy Assessment (ERAA) performed by ENTSO-E and in National Resource Adequacy Assessments (NRAA) done by national system operators. The ERAA consists of a two-step probabilistic analysis. The first is an Economic Viability Assessment (EVA), which simulates the generation expansion (or contraction) in the different countries of the European Union and also considers the effect of CRMs that are already in place. The second step analyses the resource adequacy of each of the countries. This complex analysis has been refined over the years, with the first iteration of the ERAA in 2021 not being

[8] Resource adequacy metrics are also referred to as reliability metrics in the academic literature (Zamani-Gargari et al., 2018; Tapetado and Usaola, 2019; Larsen et al., 2020; Muaddi and Singh, 2022)

approved by ACER (ACER, 2022a). The ERAA of 2022 was likewise rejected by ACER (ACER, 2023a) until the 2023 ERAA was finally approved by ACER (ACER, 2024c).

In the United States, the North American Electric Reliability Corporation (NERC), alongside its regional entities which cover the different power systems operating in North America, performs a probabilistic analysis to evaluate the resource adequacy of the different power systems (NERC, 2016; NERC, 2018). The objective is to estimate the need for additional installed capacity in each system (NERC 2021).

Resource adequacy metrics are divided into two components: the contingency analysed, and the statistical measure used to condense the probabilistic analysis. The contingency refers to the scarcity conditions the metric focuses on, such as loss-of-load hours, the amount of energy not served or served at a very high price, for example. When performing the (probabilistic) resource adequacy analysis, the system operator will determine a probability distribution function for the chosen contingency. The system operator then has to define how to condense this information through a statistical measure, either centring the analysis on a specific part of the probability distribution function (percentile, CVaR) or analysing all the distribution function (mean).

While some power systems use resource adequacy metrics that focus on the difference between peak demand and its installed capacity, such as the reserve margin, many power systems use resource adequacy metrics that focus on loss-of-load instances (ACER, 2022b), either number of hours or number of events. In fact, the discussion between hours and events, in the context of loss-of-load instances, has been historically relevant in the U.S., where the general reliability standard was "1-day-in-10-years" (Stephen et al., 2022). The interpretation of this reliability standard was not homogeneous and has historically led to different resource adequacy levels throughout the U.S. (Pfeifenberger et al., 2013). In order to avoid confusions, experts are now distinguishing more clearly between Loss-Of-Load Events (LOLEv) and Loss-Of-Load Hours (LOLH) (EPRI, 2022). In almost all cases, the statistical measure used is the mean of the probability distribution function, leading to the Loss-Of-Load Expectation[9] (LOLE). Other resource adequacy metrics focus on the amount of load not served, usually referred to as Energy Not Served (ENS), Non-Served Energy (NSE) or Unserved Energy (UE), among other derivations.

Resource adequacy metrics in use in most power systems were designed considering the scarcity conditions that characterised conventional power systems. However, as mentioned in subsection 1.1, these are changing rapidly, given transitioning power systems and the increase in extreme weather conditions. Therefore, resource adequacy metrics must be adapted to

[9] Unless stated otherwise, if the term LOLE is used in this thesis, it will be used to refer to the expected loss-of-load hours over a given timeframe (usually a year). LOLE can sometimes be reformulated as the Loss-Of-Load Probability (LOLP) which reflects the likelihood of not being able to meet demand in a particular timeframe.

capture all the facets of future scarcity conditions, which might be highlighted by different contingencies from those analysed in traditional resource adequacy metrics and might require a statistical measure that focuses on low-probability, high-impact scenarios.

Another alternative, in order to cope with the multiple facets of future scarcity conditions, is the possibility of using several resource adequacy metrics in the resource adequacy assessments performed by regulators, as some experts are recommending (ESIG, 2021; EPRI, 2022). By using several resource adequacy metrics, regulators would be able to distinguish between two different scarcity events, which might be equivalent when observing a particular metric but different according to others, as shown in Figure 1.3. In example 1, where each blue box represents 1 MWh, cases A and B include a single loss-of-load event of a four-hour duration, but the amount of unserved demand is considerably larger in case B. In example 2, both cases have the same amount of unserved demand but differing number of events and of peak unserved demand.

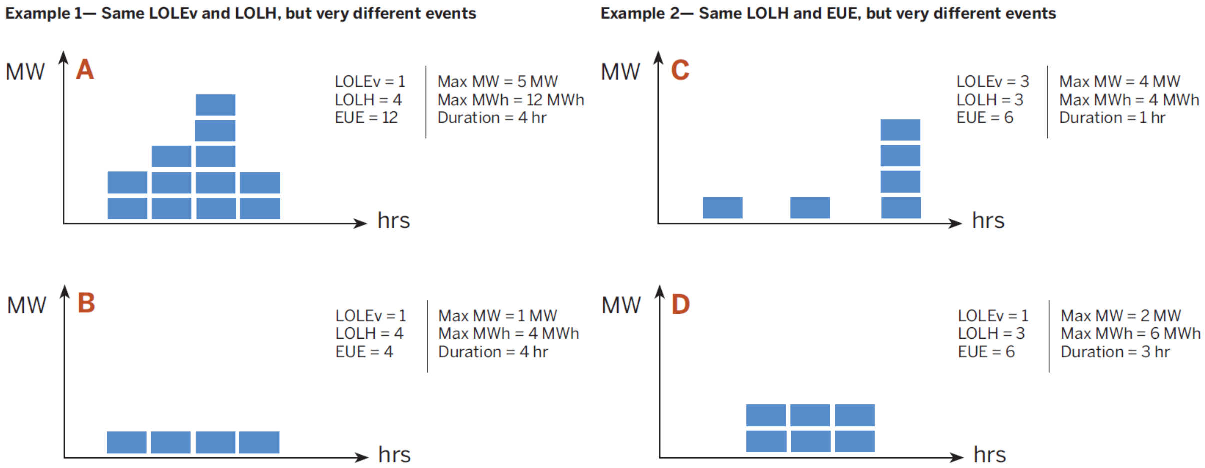

Figure 1.3: Example of the need for a multi-metric approach (ESIG, 2021)

Although this approach enriches the outcomes of the resource adequacy assessment, it could also lead to unwanted consequences, especially regarding the firm supply calculation process, which must be carefully assessed.

1.2.2. The need to properly estimate firm supply

In order to reach the desired resource adequacy level, CRMs attempt to encourage the installation of firm supply in their systems. Firm supply encompasses both firm capacity, which is used in capacity-constrained[10] systems, like those in the European Union (Béres et al., 2024) and in the United States (Hormaza Mejia et al., 2023), and firm energy, which is

[10] In capacity-constrained systems, scarcity problems arise because there is not enough installed capacity available (MW) to satisfy demand at a given moment. Energy-constrained systems, which could certainly satisfy peak demand, but would not be able to supply the demand during the remaining hours of the day/week.

used in energy-constrained systems, like hydro-dominated power sectors, such as Colombia (Isaza Cuervo and Botero Botero, 2016; Saldarriaga-C., and Salazar, 2016) and Brazil (Bastos et al., 2018; Machado et al., 2022). From a theoretical point of view, the firm supply of a resource[11] should reflect how much the resource is able to reduce scarcity conditions in the system, analysing this contribution through the resource adequacy metric used in the resource adequacy assessment. In capacity-constrained systems, firm supply is calculated through de-rating factors[12], which, when multiplied by the installed capacity of the resource, provides the value of firm supply (firm capacity, in this case).

In conventional power systems with low penetrations of intermittent renewable energy sources, demand response and electricity storage, firm supply was calculated using simplistic methodologies. In the case of conventional thermal generation, firm supply has usually been estimated according to their Equivalent Forced Outage Rate (EFOR), which still persists in many cases (National Grid, 2015; Elia, 2019). In the case of intermittent renewable energy sources, firm supply was calculated according to historical generation data (Mastropietro et al., 2019), which implied that regulators expected that these resources would produce the same in future scarcity conditions as they had done in the past.

However, as mentioned in subsection 1.2.1, if scarcity conditions are changing and resource adequacy metrics must be adapted, then the same is true for the firm supply calculation methodologies, as the two are interlinked. To begin with, if scarcity conditions will not be the same as those in the past, then the use of historical information could not only be imprecise but also bring about misleading results. Secondly, the increased presence of renewable energy sources is highlighting interdependencies between these sources of generation and electricity demand (Shaner et al., 2018; Tong et al., 2021), which could change how and when scarcity conditions appear in the system and how different resources may be able to help in solving them.

These increased interdependencies between the availability of different resources and between this availability and peak demand in the system also highlight that the firm supply of a given resource depends on the resource mix in the power system. This means that to calculate firm supply correctly, the methodology should be based on an estimation of the future composition of the electricity mix. In the case of a CRM, this can be problematic, as the mechanism itself significantly impacts the resource mix. If this is not accounted for correctly, the future mix

[11] In the context of this thesis, the term “resource” is meant to encompass generation, storage and demand response resources.

[12] The de-rating factor (Mastropietro et al., 2019) or capacity credit (Ensslin et al., 2008) represents a percentage (from 0 to 100%) of the installed capacity that the resource can trade in a CRM. It reflects that resources will not always be available at full capacity during scarcity conditions. The term de-rating factor will be used throughout this document, instead of capacity credit.

could be different than anticipated and lead to resources contributing less (or more) during scarcity conditions than what was expected when the firm supply was estimated.

Although firm supply should be calculated based on the expected contribution to achieve the reliability standard set by the regulator, in many power systems, the firm supply calculation and the resource adequacy assessment are based on different metrics (Elia, 2019; National Grid, 2019). This approach is not coherent. If the reliability standard is set using one resource adequacy metric but firm supply is calculated using another, the reliability standard may not be achieved, and the incentives provided to resources will not be aligned with reaching the desired reliability standard.

This is aggravated when regulators opt to use multiple resource adequacy metrics and establish various reliability standards to deal with these changing scarcity conditions, as highlighted in subsection 1.2.1. In this case, multiple firm supply values would have to be calculated, one for each resource adequacy metric, and the CRM should procure different reliability products. A possible solution to this problem could be the implementation of composite resource adequacy metrics, which combine these multiple resource adequacy metrics into one.

1.2.3. The need to involve demand to solve the adequacy problem

CRMs complement the energy market by providing resources with remuneration for their contribution to system reliability. In the debate surrounding the introduction of these mechanisms, one of the main criticisms concerns the possibility of CRMs being used to subsidise conventional generation based on fossil fuels (ODI, 2016), whose market share has been reduced by the penetration of new resources.

However, non-conventional, non-fossil resources, such as electricity storage and demand response, can also participate in CRMs and receive remuneration for their contribution to system reliability. In the case of demand response, CRM can even become their main source of revenues, as in PJM (Monitoring Analytics, 2023), shown in Figure 1.4, and France (ACER, 2022b).

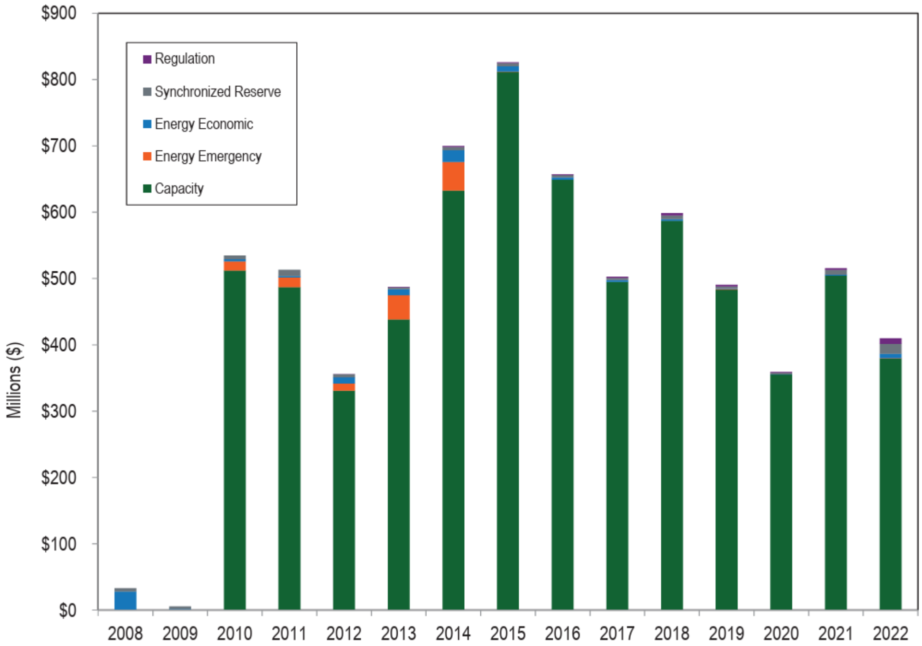

Figure 1.4: Demand response revenues by market in PJM from 2008 to 2022 (Monitoring Analytics, 2023)

Nevertheless, given that an end user selling demand response demands and offers firm supply at the same time, determining how these resources can participate in CRMs can be challenging. The first step in establishing an effective participation of demand resources is designing an efficient cost allocation methodology for CRMs. This methodology should ensure that consumers receive the proper economic signals to discourage consumption during scarcity conditions and pay according to the needs of firm supply they impose on the system. This would mean costs should not be allocated in instances of peak demand but during instances of higher risk of scarcity conditions.

The second step is designing an efficient participation methodology for demand response. In this case, there are two possibilities: i) demand resources participate in the demand side of a CRM, by defining the amount of firm supply to be purchased on their behalf, or ii) demand resources participate in the supply side by offering their firm supply to the system. In the second case, there is a risk of these resources receiving a double remuneration (ACER, 2024b) by acting as demand response, receiving remuneration for their firm supply, and having to pay less in terms of CRM costs due to their reduced consumption during scarcity conditions.

1.3. Research questions

After providing the theoretical background of this thesis and having reviewed the literature on these topics, the research questions that this thesis aims to answer can be determined. There are three main topics, one for each part of the theoretical framework described in subsection 1.2. All of these questions aim to improve the design of resource adequacy assessments and CRMs in the areas described by the theoretical framework and always in the context of rapidly evolving electricity mixes due to the energy transition. These questions can be expressed as follows.

- Resource adequacy metrics
  - How should resource adequacy metrics be reformed in order to analyse system adequacy effectively during the decarbonisation process?
  - If regulators opted to analyse adequacy using several resource adequacy metrics, what would be the implications?
  - What would be the effect if these resource adequacy metrics were combined into a single, composite metric?
- Firm supply calculation
  - How should the methodology to calculate firm supply be reformed in order to accurately reflect the contribution of each resource or technology during expected scarcity conditions?
  - How would firm supply be assessed if several resource adequacy metrics were used to analyse adequacy? And in the case of a composite metric?
  - How does the firm supply estimation impact the outcome of capacity auctions? And what is the impact of capacity auctions on the accuracy of the firm supply estimated *ex ante*?
- Cost allocation and demand participation
  - How should CRM costs be allocated among consumers?
  - What are the different alternatives for demand participation in CRMs?
  - How should CRM cost allocation and demand-response participation in CRM be coordinated in order to avoid double remuneration?

1.4. Document structure

This document is structured following the three pillars of the theoretical framework described in section 1.2: resource adequacy metrics, firm supply calculation and CRM cost allocation/demand participation. These three parts are developed in chapters 1 to 7, with chapter 8 presenting the key findings of the thesis and future work. The structure of the thesis has been schematised graphically in Figure 1.5. Although the three pillars of the thesis are presented separately in Figure 1.5 to classify the contents of the thesis and facilitate interpretation for the reader, these elements are clearly intertwined. The research presented in this thesis has not been developed "in silos" and the three pillars presented in Figure 1.5 have been addressed considering their interactions.

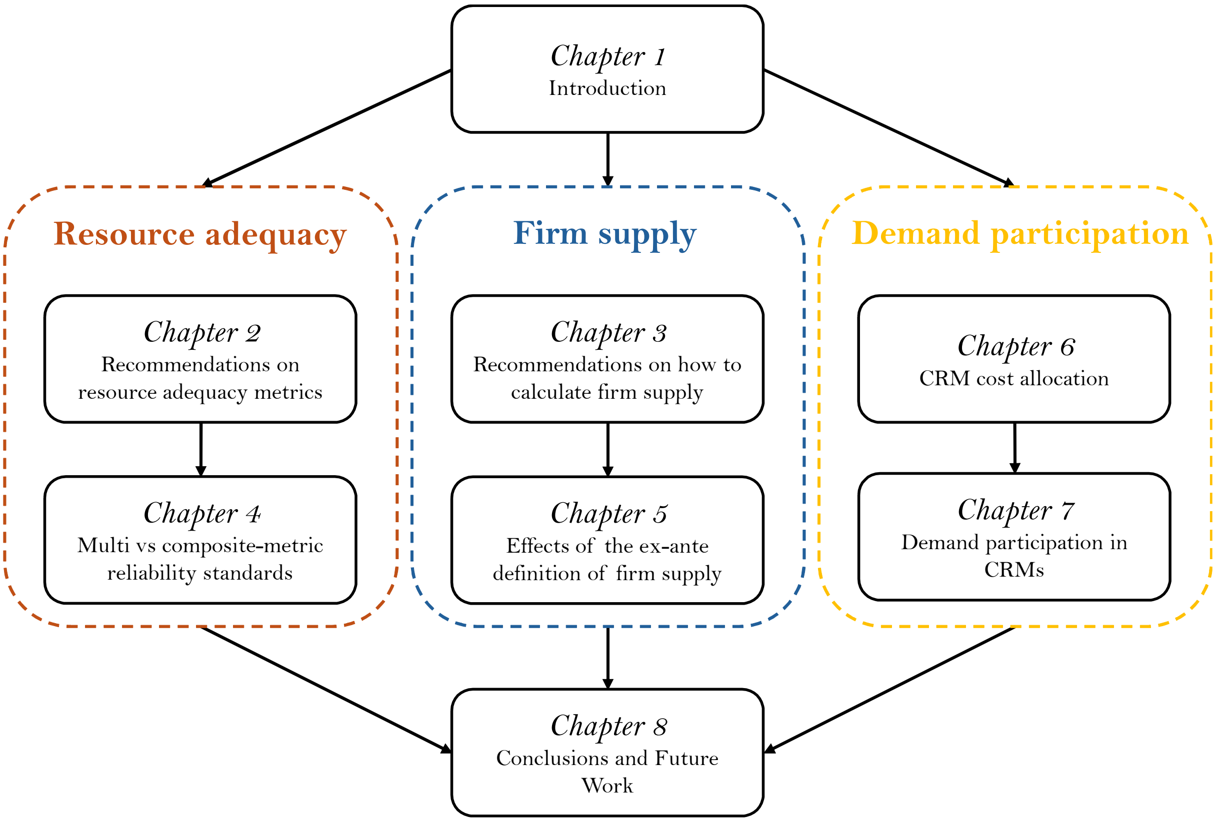

Figure 1.5: Document structure

Chapter 2 details the main regulatory recommendations regarding resilient resource adequacy metrics in the context of the energy transition. This chapter aims to outline the main characteristics that a resource adequacy metric should have to fulfil its purpose during the decarbonisation process. This chapter is based on a published article coauthored by the author of this PhD thesis (Brito-Pereira et al., 2022b). Chapter 3 presents similar recommendations but in terms of firm supply calculation. In this case, several improvements are proposed to current methodologies, and some alternatives to reduce the computational burden of this process are recommended. Likewise, it is based on a published article coauthored by the author of this PhD thesis (Brito-Pereira et al., 2022b).

Chapter 4 analyses the effect of using various resource adequacy metrics to perform resource adequacy assessments and compares its effects against the use of a composite resource adequacy metric. The contents of this chapter are based on a published article coauthored by the author of this PhD thesis (Brito-Pereira et al., 2025). Chapter 5 presents the problem of calculating the firm supply of resources based on an estimation of which resources will be in the system and how it might diverge from the ex-post value calculated with the resource mix resulting from the CRM. The contents of this chapter are also based on a published article coauthored by the author of this PhD thesis (Brito-Pereira et al., 2022a).

With respect to demand participation and CRM cost allocation, chapter 6, based on a working paper coauthored by the author of this PhD thesis[13], presents a methodology on how to allocate CRM costs efficiently among consumers. Chapter 7 details the different alternatives for demand participation in CRMs and how to avoid a double remuneration problem, when demand resources are represented in both the demand and the supply side of the capacity market. The findings presented in this chapter were published in article coauthored by the author of this thesis (Rodilla et al., 2023).

Finally, chapter 8 concludes the thesis by presenting an overview of the main discussions of the thesis and their most important takeaways and outlining possible future lines of work that derive from the research presented in this thesis document.

[13] Brito-Pereira, P., Rodilla, P., Mastropietro, P., 2025. Efficient cost allocation in capacity remuneration mechanisms: applying the causer-pays principle to resource adequacy. Working paper IIT-24-052. Under review at International Journal of Electrical Power and Energy Systems.

2. RECOMMENDATIONS ON RESOURCE ADEQUACY METRICS

2.1. Introduction[14]

Resource adequacy metrics attempt to assess the adequacy of the system by quantifying future security-of-supply problems. In the past, in power systems which were dominated by dispatchable, conventional thermal generation, resource adequacy was often analysed with simplistic methodologies such as the reserve margin, which only represents a relation between the peak demand of the system and its installed capacity.

The use of simplistic methodologies made sense in the context of a prevalence of dispatchable, conventional thermal generation, as security of supply problems usually occurred during peak demand when there was not enough available capacity to supply all of the demand. However, current power systems are diverging significantly from the past, and the penetration of new generation technologies, such as intermittent renewable energy sources, electricity storage and demand response, is also reshaping resource adequacy.

Current changes in the technology mix of power systems, coupled with the impact of climate change, are causing a transformation in the scarcity conditions that power systems are facing. The increased penetration of renewable energy sources for electricity (RES-E) is highlighting interdependencies between different generation resources and between these resources and electricity demand (Shaner et al., 2018; Tong et al., 2021), which could lead to scarcity conditions shifting from instances of peak electricity demand to instances of low RES-E generation. Climate change also influences scarcity conditions, causing a rise in extreme weather situations (IPCC, 2023). The combination of these multiple effects is leading to multi-faceted scarcity conditions in power systems.

Traditional resource adequacy metrics, such as the aforementioned reserve margin or the LOLE, might not be suited for the new paradigm of power systems. These resource adequacy metrics either do not consider that scarcity conditions can appear in instances other than peak

---

[14] This chapter heavily draws from Brito-Pereira, P., Mastropietro, P., Rodilla, P., Barroso, L. A., Batlle, C., 2022, Adjusting the aim of capacity mechanisms: Future-proof reliability metrics and firm supply calculations. Energy Policy, Volume 164, Article 112891.

demand or peak net demand, like the reserve margin, or consider each loss-of-load instance as equivalent, like the LOLE.

The main objective of this chapter is to describe the ideal characteristics of resource adequacy metrics in the current context of energy transition. Future-proofed resource adequacy metrics must be resilient to changes in the characteristics of power systems and be able to capture the shift in scarcity conditions that these changes may provoke. Once a suitable resource adequacy metric that possesses the aforementioned resilience to changes in power systems has been chosen, regulators would then embark on the arduous journey of establishing a reliability standard (or standards) based on this metric. This process should be guided by an economic analysis that balances the costs associated with a lack of adequacy (non-served demand or high electricity supply costs) with the costs associated with achieving the desired level of adequacy (procuring firm supply through a capacity remuneration mechanism) (De Vries and Heijnen, 2008). However, this discussion exceeds the scope of this chapter and this thesis.

Chapter 2 is structured as follows. Section 2.2 will provide an overview of the international experiences regarding resource adequacy metrics. Section 2.2.1 will centre the discussion on the different contingencies that resource adequacy metrics can analyse, presenting and discussing the different alternatives and which are best suited to the context of the energy transition. Section 2.2.2 will discuss the different choices with regard to the statistical measure to be used in the resource adequacy metric and present recommendations. Subsection 2.3 presents the main conclusions of this chapter.

2.2. Review of resource adequacy metrics

The resource adequacy metrics that have been most widely used to assess the resource adequacy of power systems are the 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, 2018b; ACER, 2020b).

- The reserve margin (RM) measures the difference between the installed capacity of a system and its peak demand. Potential insufficient availability has traditionally been accounted for by de-rating the installed capacity of units using rudimentary methodologies (e.g., by multiplying that value by an equivalent forced outage rate). Reliability standards based on reserve margin have 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, standards defined in terms of this resource adequacy 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. The reserve margin is still used as adequacy metric in some European systems (EC, 2016b).
- The loss of load probability (LOLP)/loss of load expectation (LOLE) is a metric used to estimate the likelihood of a system being unable to accommodate the full load demanded at any given time, as opposed to only during peak periods. The two metrics are actually

equivalent, with the only difference being the manner in which reliability is expressed. For the LOLP, reliability is expressed as a probability (in per cent), whereas for LOLE, it is expressed as the cumulative duration (usually in hours per year) of scarcity events during which demand cannot be met. Both are extensively used in the United States and Europe (NERC, 2018; ENTSOe, 2020).

- The expected energy not served (EENS) is a metric used to quantify the expected amount of energy that the system will be unable to supply within a specified time horizon. In contrast to the focus of LOLP and LOLE, which is placed on the mere frequency and/or duration of scarcity events, EENS aims to quantify the depth or magnitude of load loss in terms of energy, thereby providing 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 within the specified time horizon. The EENS is used in Alberta (NERC, 2018) and Australia (AEMO, 2019).
- The 95th percentile of loss of load duration (LOLE95 or LOLD95) is determined by applying that statistic to the probability distribution function for loss of load duration (ENTSOe, 2019). It focuses on extreme scenarios (although it rules out the uppermost 5% of the distribution function) and is always larger than LOLE. Until recently, it was used in Belgium (Elia, 2016), but as of the writing of this document only the LOLE is used for their analyses (Elia, 2024a).
- The analysis of the energy supply in the least favourable hydrological scenario is a typical procedure in energy-constrained systems, such as hydro-dominated power systems (Rodilla et al., 2015). In such systems, 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.

2.2.1. Type of contingency measured

2.2.1.1. Events vs unserved energy

As mentioned in the previous section, LOLP and LOLE pool events where demand was not met in a given period of time without distinguishing the 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 was not met would be assigned the same weight as one where the deficit was 1 GW. This inherent limitation may impede the ability of these metrics to accurately reflect the actual contribution of a resource 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. Conversely, a 100 MW power plant would eliminate the deficit entirely, reducing the LOLP or LOLE, despite the almost equivalent contribution of the two plants to the system. As discrete metrics, LOLP and LOLE are able to measure full compensation for a shortage only, a characteristic that could intensify adequacy assessment volatility.

In contrast, EENS quantifies the severity of each scarcity event, thereby establishing a continuous metric. When applied to the foregoing example, it would reflect that the two plants cover very similar amounts of unserved energy, thereby providing a more accurate measure of the overall contribution to system adequacy. In Europe, several countries have started to analyse resource adequacy using the EENS, such as Cyprus, Estonia and Finland, which marks a shift from the originally widespread use of LOLE for such analyses (ENTSO-E, 2025).

2.2.1.2. Resilient metrics for power sectors with elastic demand

However, the evolution of power systems and the increasing price elasticity of demand will render the concepts of loss of load and unserved energy less significant. This phenomenon can be observed in the simple graphic example in Figure 2.1. In circumstances where demand is found to be inelastic (i.e. on the left-hand side of the graph), it is possible to identify a scarcity event and to measure demand that has not been met[15]. Conversely, at the other extreme, in circumstances where demand is wholly elastic, it is not possible to define unserved energy. This is due to the fact that at least part of the demand that goes unsupplied attributes a lower value to electricity than the market clearing price. It would be erroneous to consider such demand as unmet; rather, it is more accurate to conclude that individuals are choosing not to engage in trade activities that are not beneficial according to their utility function.

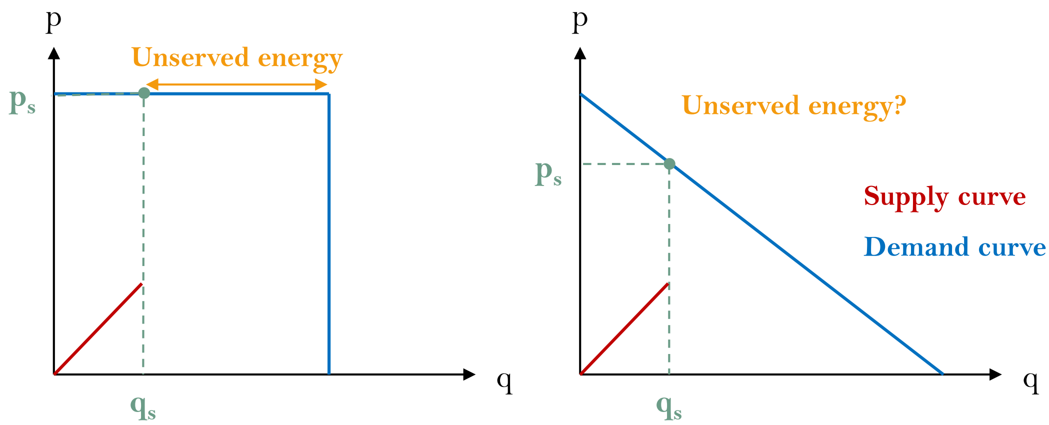

Figure 2.1: Unserved energy with inelastic (left) and elastic (right) demand

[15] 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).

It is important to note, however, that even in scenarios where demand is totally elastic in the short term, there is a possibility that prices remain extremely high over extended periods of time. This could be indicative of a resource mix that deviates substantially from the optimal that would maximise social welfare. Regulators may wish to lower the probability of such events, with a view to move towards a resource mix better suited to maximising welfare by implementing a capacity mechanism.

As demand becomes increasingly price-responsive, it is imperative that reliability criteria should incorporate this dimension, thereby providing regulators with a tool capable of identifying emergent scarcity situations. While demand elasticity remains constrained, particularly among certain consumer groups, capacity mechanisms are long-term regulatory instruments that should be designed to withstand changes in the power sector of the future. It is a widely held view amongst experts that rising demand elasticity will characterise the energy transition (MITEI, 2016; EC, 2019).

Even in power systems that are still characterised by scant demand elasticity, the market price is an excellent barometer of scarcity. It is 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 evidenced in Colombia in 2015/2016 (Mastropietro et al., 2020).

One potential approach to incorporating the price dimension into reliability metrics would be to include high-price events in traditional adequacy assessments. This would necessitate 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 encompass the energy offering cleared at a value higher than a certain reference price. The difference between traditional reliability and price threshold-based metrics is represented graphically in Figure 2.2. The implementation of this price threshold would seek to ascertain the genuine willingness to pay by the part of the demand not currently participating in the market and is consequently undetected.

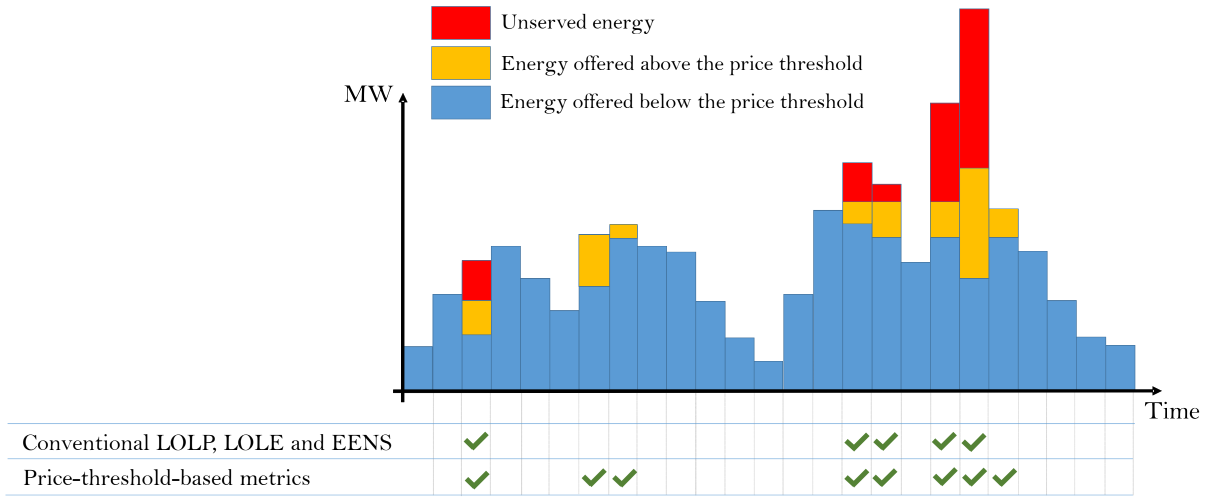

Figure 2.2: Difference between traditional reliability and price threshold-based metrics

The incorporation of the price dimension as a factor in capacity mechanism design is not a novel concept; indeed, it is among the earliest fundamentals of the reliability options[16] approach (Vázquez et al., 2002; Bidwell, 2005; Cramton et al., 2013; Andreis et al., 2020). However, with such mechanisms (reliability options are currently being used in Colombia, Ireland and Italy), 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 neither considered in the adequacy assessment nor in firm supply calculations.

The sole experience in Europe where a power system pondered accommodating the price dimension in adequacy assessment was a proposal put forward by Elia (2019b), the Belgian system operator, for redesigning the country's capacity mechanism. Elia recognised that focusing solely on loss-of-load events was insufficient, as it failed to consider 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, in addition to times when a minor rise in demand would translate into unserved energy. However, although Elia studied the possibility of using a price threshold to identify near-scarcity hours, this approach was ultimately dismissed on the basis that administratively setting such a scarcity price would be arbitrary (Elia, 2019a). However, this assertion may be countered by the observation that defining a demand increment threshold deemed as conducive to unserved energy would also be arbitrary.

Currently, Brazil is the only power system in the world with a reliability standard referencing the price dimension. In 2019, public authorities started taking steps towards an update of its

[16] Reliability options are financial call options enabling buyers of the option to obtain the positive difference between the underlying asset (in this case usually the electricity day-ahead, intraday o reserve market price) and the strike price from the seller. A more detailed explanation of this instrument can be found in the paper by Vázquez et al. (2002).

reliability standards. At the end of the process, four reliability standards were established: two based on capacity requirements (one based on the LOLP and the other limiting the CVaR of the power not served), and the other being energy requirements (one limiting the CVaR of the non-served energy and the other the CVaR of the marginal cost of operation, CMO). The latter reliability standard limits the average monthly electricity market price in the worst scenarios to the established threshold (EPE, 2020a).

Figure 2.3: Infographic describing the four reliability standards used in Brazil (EPE, 2020b)

2.2.2. Statistical measure used to characterise the contingency

2.2.2.1. Deterministic vs. probabilistic approaches

The underlying contingency (number of scarcity events, unserved energy or energy supplied above a price threshold) may be measured in a deterministic manner, 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, it is necessary to employ a statistical measure that will '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). Their computational simplicity is counterbalanced by their steadily declining accuracy in representing modern power system realities, which has resulted in their use in very few jurisdictions. Conversely, the more complex probabilistic approach necessitates a considerable number of assumptions that have the potential to influence the outcome. However, it is important to note that these models deliver more precise results, as they are able to capture the significant interdependence among the variables studied (such as the occurrence of scarcity events and the availability of certain technologies).

2.2.2.2. Mean and median vs. consideration of extreme scenarios

While mean and median values can be considered representative of the entire probability distribution function, they are known to underestimate the weight of the tail of the probability distribution function, i.e. the area where the most severe shortages lie. Since the purpose of capacity mechanisms is to guarantee the security of supply, particularly under the harshest conditions, resource adequacy 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 phenomenon may result in the presence of more pronounced tails in the distribution function of stress events (as illustrated graphically in Figure 2.4 for Australia), which should be internalised in adequacy assessments[17].

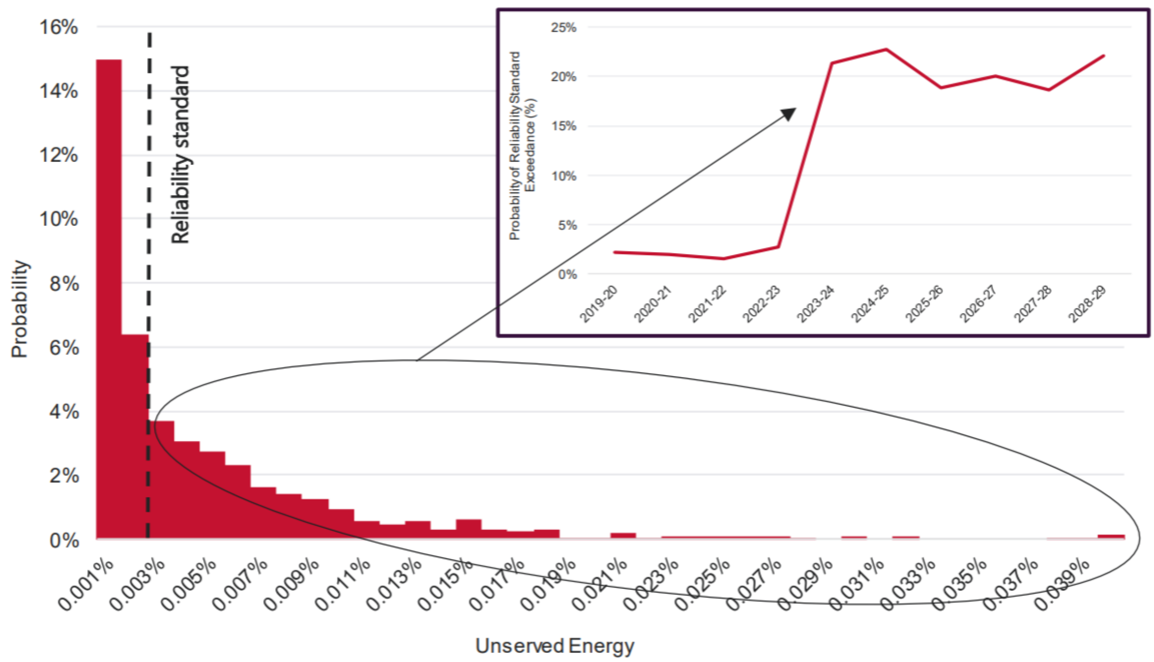

Figure 2.4: Distribution of annual unserved energy in New South Wales, 2023-2024 (AEMO, 2019)

In Europe, as of the writing of this document, countries continue to focus on average scenarios, although the latest ERAAs (ENTSO-E, 2025) provided results related to extreme scenarios (95th percentile). Apart from this, as mentioned at the start of section 2.2, Belgium used to perform analysis related to 95th percentile of the loss of load distribution curve (Elia, 2016).

The Colombian power system provides another example of the significance of statistical parameters. In the Colombian power system, hydropower constitutes 70 % of the total installed capacity. However, every few years, the El Niño phenomenon brings droughts and high temperatures to the region for several months, reducing hydropower reservoir inflows

[17] 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).

and jeopardising security of supply (Mastropietro et al., 2020). The El Niño phenomenon is known to exhibit variability in terms of duration, intensity and periodicity. However, 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. As demonstrated in Figure 2.5, there is a marked increase in the price of electricity in Colombia during the El Niño events[18].

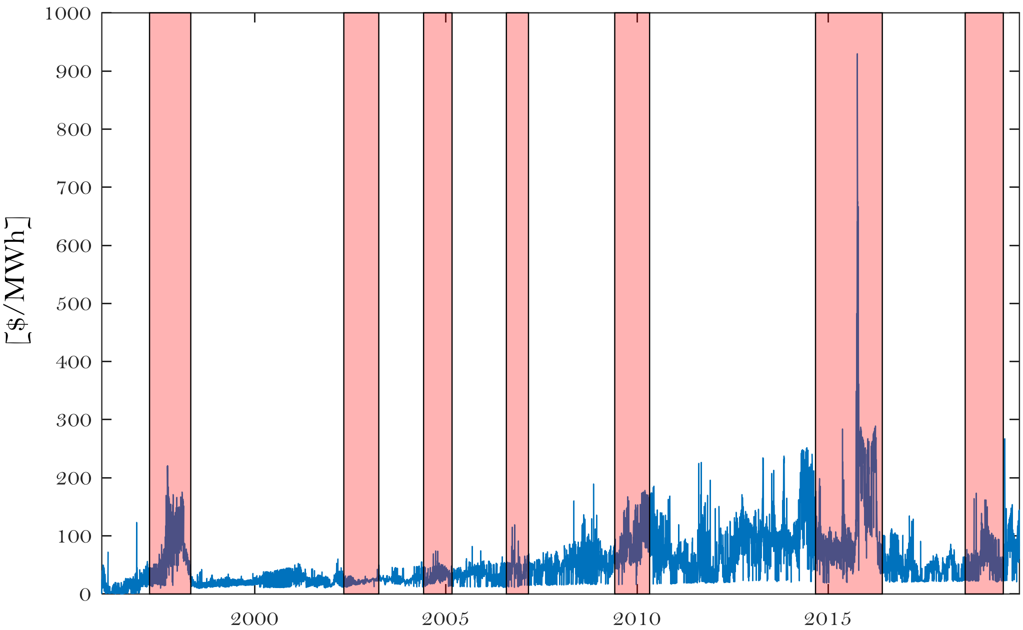

Figure 2.5: Wholesale electricity market prices in Colombia, 1996-2019, inclusive (authors' formulation based on XM data (XM, 2021) and BRC exchange rate information (BRC, 2021))

A subset of the Colombian market prices (2009 to 2019) presented in Figure 2.5 was rearranged to construct the probability distribution function[19] illustrated in Figure 2.6 to exemplify the limitations of some statistical parameters. In addition to the information included on the graph, it should be noted that even very high (95th or 99th, PCTL 95/99) percentiles may not fully capture the full weight of the tail.

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

[19] Assuming equiprobability for all instances.

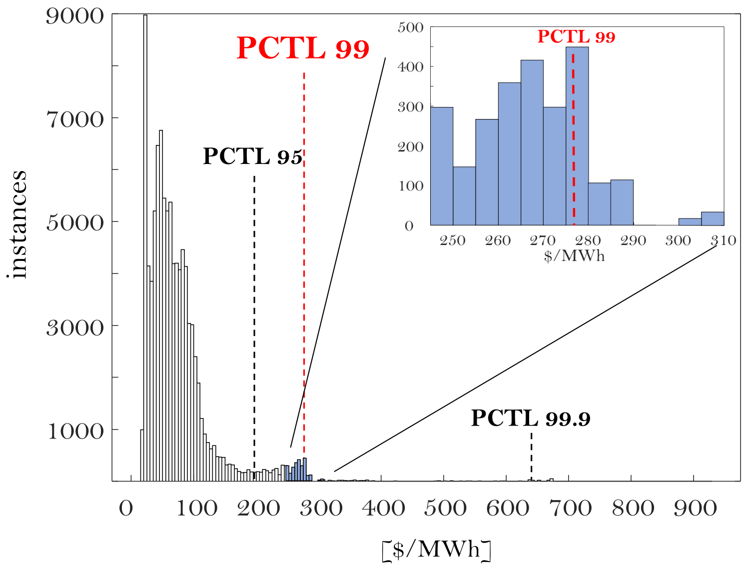

Figure 2.6: Distribution function for wholesale electricity market prices in Colombia (authors' formulation based on XM data (XM, 2021) and BRC exchange rate information (BRC, 2021))

A statistical parameter that is less widely used in adequacy assessments but that might remedy the aforementioned shortcomings is Conditional Value at Risk (CVaR), although the latter is more prevalent to represent risk aversion in generation investment decisions (Fraunholz et al., 2023; Hohl and Lo Prete, 2025). 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 2.7 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 %).

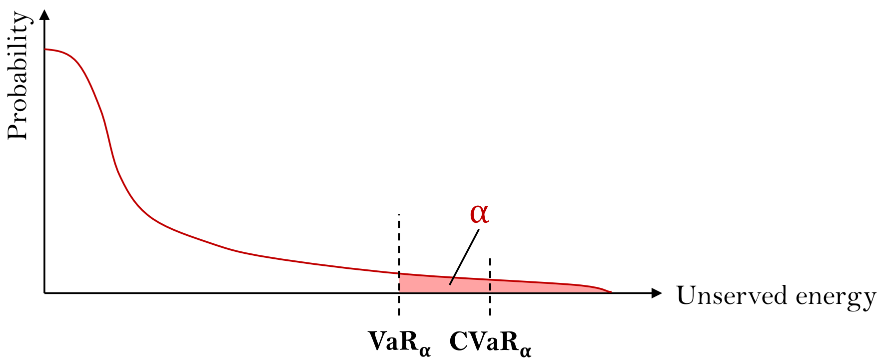

Figure 2.7: Graphical representation of the CVaR applied to unserved energy

As mentioned in subsection 2.2.1.2, Brazil already considers the CVaR in three of its four reliability standards. This emphasises the importance placed by the Brazilian institutions in focusing their adequacy assessment on the harshest conditions that the power system may have to face. Additionally, the discussion of using the CVaR as a statistical measure for resource adequacy metrics is now also present among experts in the field (Dent et al., 2024; Tillmanns et al., 2026).

2.3. Conclusions and policy recommendations

In the context of the energy transition, power systems and the scarcity conditions they face are changing both rapidly and drastically. Traditional resource adequacy metrics widely used in the past, such as the reserve margin and the LOLE, focus on very specific characteristics of resource adequacy that are becoming increasingly irrelevant. This is why power systems need to reconsider the resource adequacy metrics they use when assessing their resource adequacy levels.

This chapter has highlighted the main discussions regarding the design of resource adequacy metrics, both in terms of the contingency analysed and the statistical measure used to condense the probabilistic information of the resource adequacy assessment. The main conclusions and recommendations on both of these elements is presented hereunder.

In terms of the contingency to be measured, given the critical analysis presented and discussed in section 2.2.1, the optimal reliability metric for managing the new present and future electricity resource mix should be the energy cleared and supplied above a price threshold. This definition includes both unserved energy and energy offered and cleared over that price. As with EENS, this reliability metric is continuous, thereby enhancing the consistency and stability of the results of any assessment conducted with it[20]. In contrast to EENS, however, it demonstrates resiliency to higher demand elasticity, as it identifies scarcity conditions based on the market price, which serves as the most accurate barometer of such conditions (see subsection 2.2.1.2).

The literature on reliability options, in conjunction with international experiences of countries where this kind of capacity mechanism is in place, can be used to inform the definition of the price threshold. According to Vázquez et al. (2002), the strike price of reliability option contracts should be set at a sufficiently elevated level to ensure that they do not interfere 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 price rises are driven by other factors (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. This is, in turn, defined as the dispatchable technology that would be included in the optimal generation mix with the lowest unit investment cost (€/MW). It is important to note that while the reference technology does not change, the

[20] To return to the example in subsection 2.2.1.1, 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.

strike price is subject to indexation and varies on a weekly basis. The Italian indexation formula encompasses fuel costs (the most significant 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 resource adequacy metric.

On the other hand, in terms of the statistical measure, the proposed metric should be analysed using a model designed to process the variables at issue sequentially in order to simulate power system operation within the time horizon under study. Although determinist adequacy assessments, although they may present computational advantages, are too simplistic, and a probabilistic assessment is required to capture all possible future scenarios.

In such a probabilistic model, the resource adequacy metric must depend on a statistical parameter that pools the results for all the scenarios envisioned. As contended earlier, the statistical parameter most suitable for contemporary and future power systems is the 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 is focused on extreme scarcity events, those for which capacity mechanisms have been designed to mitigate, and those whose frequency and intensity may be heightened by climate change.

3. Recommendations on how to calculate firm supply

3.1. Introduction[21]

Capacity remuneration mechanisms aim to achieve a desired level of resource adequacy, expressed through the reliability standard. CRMs aim to comply with the reliability standard by attracting firm supply, which can be understood as the amount of capacity (or energy in energy-constrained systems) that is expected to be able to supply during scarcity conditions.

Determining the firm supply of resources is a key element in the operation of CRMs and the most important feature in the design of the reliability product. Firstly, it is essential to evaluate if the system can comply with its reliability standard, and secondly, it determines the amount of reliability product that resources can trade in the CRM, rewarding each resource according to its contribution to solving the resource adequacy concerns in the system.

Similarly to what occurred with resource adequacy metrics, in power systems characterised by a prevalence of dispatchable thermal power plants, simplistic methodologies, based, e.g., on power plant outages or historical generation patterns, were used to estimate firm supply. However, as the presence of RES-E, electricity storage and demand response in power systems has increased, estimating firm supply has become increasingly complex

As highlighted in chapter 1, the increased presence of these resources has caused a shift in scarcity conditions, exacerbated by the increased interdependencies of these resources and demand. In this case, estimating firm supply based on historical generation patterns might be increasingly inaccurate. This is why many experts and regulators have proposed more accurate methodologies to calculate firm supply and an update in the firm supply that can be provided by resources, but these are often centred around RES-E and short-term energy storage (Tapetado and Usaola, 2019; Mertens et al., 2021; Freire-Barceló et al., 2022). However, as has been highlighted by the Texas winter cold snap of 2021, conventional thermal generation could present correlated outages, which was not considered beforehand in

[21] This chapter heavily draws from Brito-Pereira, P., Mastropietro, P., Rodilla, P., Barroso, L. A., Batlle, C., 2022b, Adjusting the aim of capacity mechanisms: Future-proof reliability metrics and firm supply calculations. Energy Policy, Volume 164, Article 112891.

resource adequacy assessments (Busby et al., 2021), and predicting hydro output might become increasingly complex, especially in hydro-dominated systems like Colombia and Brazil if climate change increases drought risks (Elkouk et al., 2022; Santos Borba et al., 2023).

The main objective of this chapter is to present the main design elements of the methodologies used to estimate the firm supply of power system resources and present recommendations for each of them. To do this, the chapter is structured as follows. Section 3.2 details the different design alternatives that exist when calculating firm supply. Sections 3.2.1 to 3.2.6 go in depth through each of them and presents recommendations. Section 3.3 summarises these recommendations and presents the conclusions and policy implications of the chapter.

3.2. Methodologies to calculate firm supply

This section addresses theoretical considerations around methods for calculating firm supply 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 resource adequacy metric used to calculate each resource's contribution
- 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.

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 segment resources, for which no theoretical justification can be alleged. Although defining specific methodologies for new technologies (a course of action adopted by numerous regulators for variable renewable energies) may offer an expeditious and uncomplicated approach to facilitating the integration of these resources into the capacity market without necessitating a comprehensive review of the adequacy assessment, the outcome is suboptimal and may distort competition among competing technologies.

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

The contribution of a resource to system adequacy is contingent upon its output during 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 resources in the system. The duration of shortages can vary significantly depending on the characteristics of the system. In capacity-constrained systems, shortages may persist for a few hours, whereas in energy-constrained systems, they can extend over weeks, months, or even years.

However, in some jurisdictions, the firm supply for specific technologies is calculated on the basis 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[22]. A similar line of reasoning is adopted when a thermal power plant is de-rated based solely on its equivalent forced outage rate, disregarding any potential correlation among such outages and/or between outages and the occurrence of shortfalls. Examples of these types of 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; Busby et al., 2021).

3.2.2. Reliability metric for assessing contributions

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

[22] 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.

Although such reasoning may sound obvious, it is surprisingly often absent from standard practice. In most power systems that rely on capacity mechanisms, such as those found in 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), the firm supply of some, or all, technologies is calculated using a different metric than used to define the reliability standard. For instance, resource adequacy is assessed in the UK using the LOLE, while de-rating factors for renewable resources are calculated on the grounds of their contribution to the reduction of EENS. Belgium also conducts resource adequacy assessments based on LOLE, whereas the operator proposed calculating renewable de-rating factors further to expected output during the near-scarcity hours (Elia, 2019a). However, this approach was modified eventually, and the de-rating of RES-E technologies is calculated according to their contributions during scarcity conditions only, as of the writing of this document. In this case, given that the calculation of firm supply is affected by the choice of the resource adequacy metric (Muaddi and Singh, 2022), the de-rating factors calculated will differ if the resource adequacy metric used to define the reliability standard is also utilised for the calculation of firm supply.

It is evident that this approach is suboptimal, inasmuch as it entails remunerating resources for a service that, while related to, is not exactly the one needed to meet the reliability standard defined. The theoretical justification for this assertion can be found in the mathematical formulation of the optimisation problems provided in Annex I, where welfare maximisation in a centralised context, constrained by a system reliability target, is compared to the decentralised (or liberalized) context, which generates price signals for both the short-term market and firm supply (Pérez-Arriaga, 1994; Schweppe et al., 1988).

3.2.3. Marginal (or incremental) vs. mean historical contributions

The design element that is likely to have the most significant 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 a minor (typically a few MW) increment in resource or technology capacity. By contrast, in the latter, the focus is placed on the mean contribution of the resource or technology capacity as a whole[23].

The difference between mean and marginal contributions is particularly relevant for technologies that might progressively change the scarcity conditions facing the system as their presence in the system increases. One example is solar PV-based resources, which help reduce the likelihood of shortfalls during daylight hours but are unable to generate power late in the evening. This effect is illustrated in Figure 3.1, which provides a comparison between

[23] 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 3.2.5.

the generation 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.

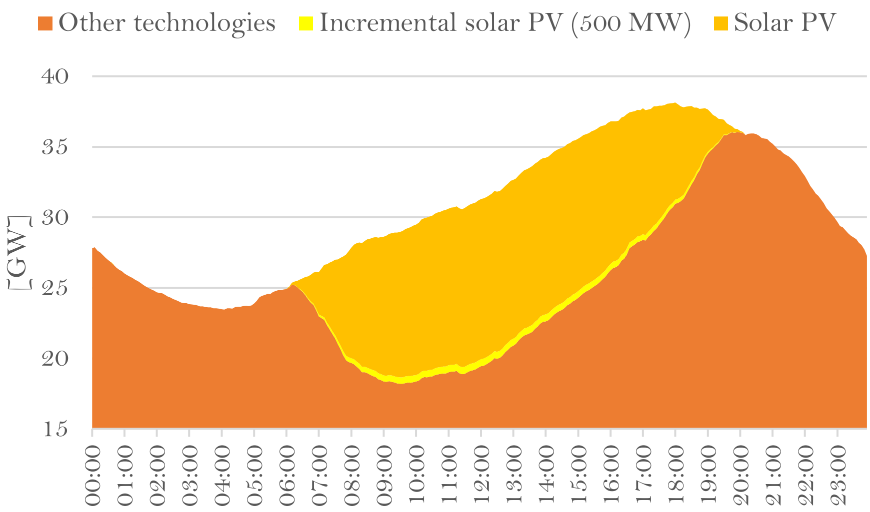

Figure 3.1: Total and net demand for a system with substantial solar PV output (authors' formulation based on California ISO data (CAISO, 2019) for 16 July 2019, including incremental solar power)

As illustrated in Figure 3.1, the hypothetical production of an additional 500 MW of solar capacity (light yellow) is depicted in order to demonstrate the effect of a marginal[24] increment of the installed capacity of solar PV. 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, given that its production at net peak demand is practically nil, its contribution to marginally calculated firm supply would be negligible. Consequently, the capacity mechanism should not remunerate new solar resources installed under such circumstances.

Conversely, if solar technology is taken as a whole, it contributes to improving system adequacy as 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. This is because the total contribution, which rises almost negligibly, would be shared by a larger number of resources. Acknowledging the mean contribution of the technology as a whole to the new resources would credit them for greater firm supply than merited by their actual contribution.

[24] 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.

The downturn in solar PV marginal firm supply with greater system presence is depicted in Figure 3.2, which plots PV solar power ELCC (Effective Load Carrying Capability) against installed capacity in California ISO.

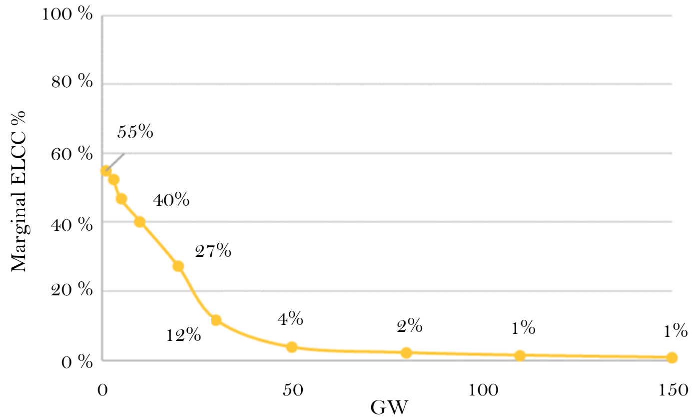

Figure 3.2: Variation in solar PV marginal ELCC vs installed capacity (Energy & Environmental Economics, 2019)

The contribution illustrated in Figure 3.2 is expressed as marginal ELCC, a probabilistic method that estimates the additional demand that would be met by increasing the installed capacity of the resource in question without worsening the adequacy of the system. 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 one that would guarantee the regulator's reliability target (Garver, 1966). Despite the extensive utilisation of ELCC 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 in this chapter (such as basing the adequacy metric used for the ELCC calculation on a continuous and resilient metric).

This discussion in this section has focused up to now on the role of the resource mix in determining each resource or technology's firm supply. However, firm supply obviously depends as well 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 categorised under different 'storage classes', as defined in terms of storage duration. As can be drawn from the data in Table 3.1, the de-rating factor is lowest for the shortest durations, inasmuch as this characteristic limits such resources' contribution during scarcity events.

Table 3.1: 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%

With the marginal approach to determine the firm supply of resources, 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) and is increasingly being used by experts to calculate the firm supply of different resources (Mills and Rodriguez, 2020; Muaddi and Singh, 2022; Ssengonzi et al., 2022; Awara et al., 2023; Wang et al., 2024).

The suitability of using marginal contributions to analyse firm supply is further substantiated by the mathematical formulation set out in Annex I. The resulting optimality demonstrates that the per-unit remuneration of a specific resource (M_{K_i}) for its participation in the capacity mechanism is the derivative of the resource adequacy 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 essence, the latter represents the capacity/adequacy market price:

$$M_{K_i} = \frac{\partial RM}{\partial K_i} \cdot \beta$$

3.2.3.1. A second-best yet efficient approach

As previously stated in chapter 2, the resource adequacy metric can focus on a contingency which is discrete or continuous by nature. If the regulator opts for a continuous metric (such as the EENS or the energy cleared and supplied above a price threshold), a second-best alternative to the marginal contribution may be implemented, as discussed in this section.

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 3.2.4.

If the reliability metric proposed in chapter 2 was used, the calculation would be performed as outlined in Figure 3.3. Critical periods are defined as instances 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 (light green areas on the graph) might contribute to a reduction in some of this 'critical' energy. This 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).

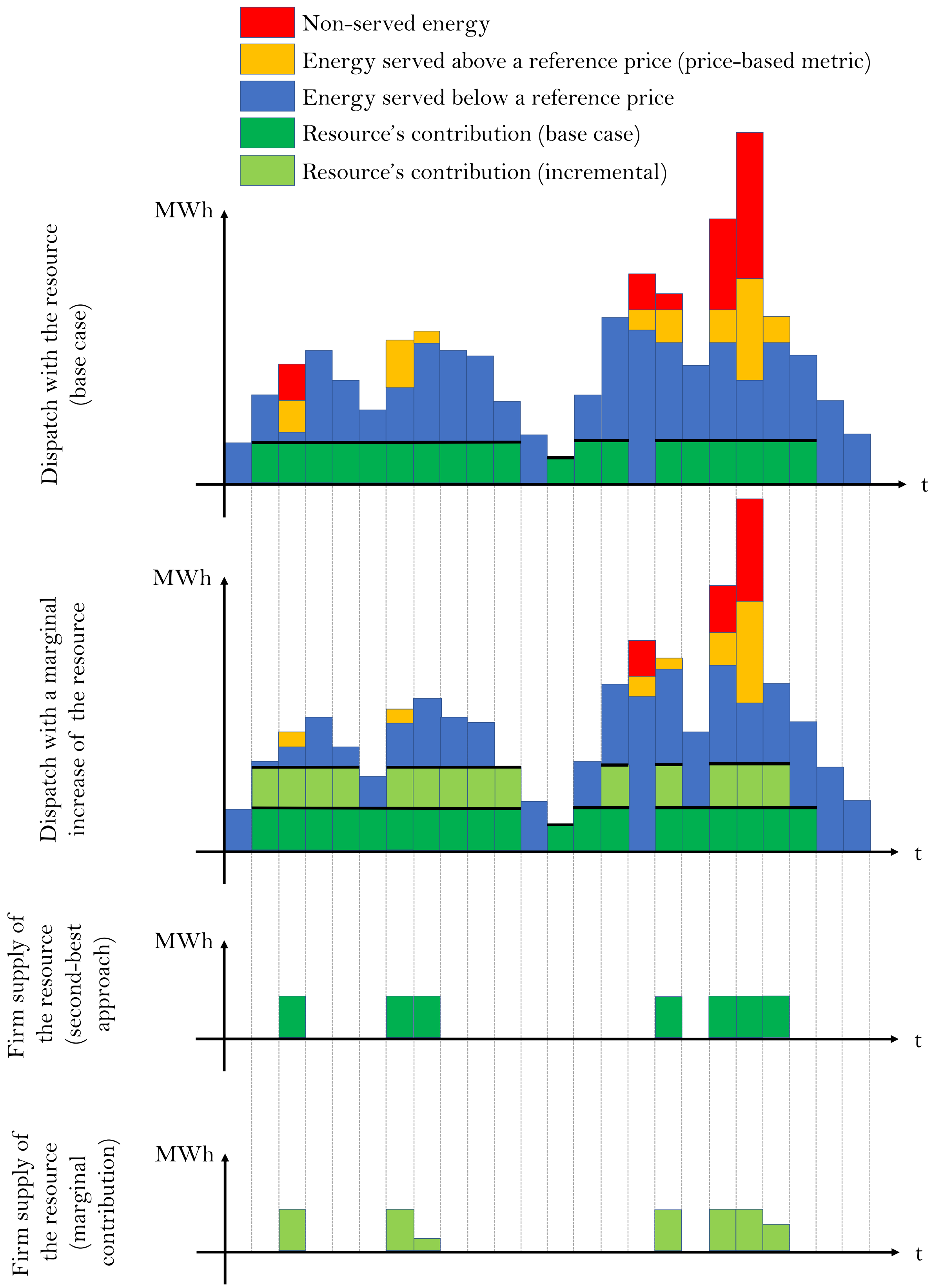

Figure 3.3: Graph showing the best practice (marginal) and second-best methods for calculating firm supply

This methodology delivers an accurate estimate of each resource's contribution to the reliability target. However, it may be subject to computational challenges (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. Thirdly, although a marginal contribution is mentioned, what is usually assessed is an incremental contribution, i.e., a real size resource is added to the system and its impact is evaluated. In this case, although regulators evaluate the impact of a real resource, the economic signals provided might not be as precise.

The aforementioned computational problems may be circumvented 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. This approach would lead to differences with respect to the first-best approach described beforehand when a resource's marginal contribution would eliminate the critical energy during scarcity conditions, as shown in Figure 3.3. 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 chapter 2, the CVaR is used as the statistical parameter to calculate the resource adequacy 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.

Case study to analyse the accuracy of the second-best approach

In order to measure the accuracy of the second-best approach when calculating the firm supply of resources, a case study is briefly presented in this section. In this case study, the precision of the second-best approach was studied for eight resources: three short-term electricity storage resources (1-hour, 4-hour and 8-hour battery storage), three conventional thermal generation resources (CCGTs, coal and nuclear generation), and two RES-E resources (solar photovoltaic and wind power).

A simplified Unit Commitment (UC) model[25] was executed for ten cases: a base case execution, one for each of the resources in the study, in which a small increment (100 MW) of the corresponding resource was added to the system, and a final execution in which a 100 MW of perfect generation (infinite ramping possibilities, no outages and negligible variable operation cost) was added to the system. The base case execution uses an approximation of the Spanish electricity mix of 2022, in which a small amount of the three short-term electricity storage resources has been included (100 MW) as they were not naturally present in the system and

[25] The model formulation as well as the data used can be consulted in Annex II.

the second-best approach requires the resources to be present in the base case system. For each of these cases, in which a small increment of a resource is added, and the base case, a deterministic UC is run 1000 times, each one accounting for different hourly availability scenarios of the different resources, renewable generation patterns and electricity demand levels and a year of hourly operation.

To calculate the firm supply according to the first-best approach, the expected non-served energy (the resource adequacy metric chosen for this analysis) of the base case was computed and measured against the EENS of each of the incremental cases. The de-rating factor of each resource (DF_i) was then calculated according to the following expression:

$$DF_i = \frac{EENS_{Base} - EENS_i}{EENS_{Base} - EENS_{Perf}}$$

Where:

- $EENS_{Base}$ corresponds to the expected energy non-served from the base case.
- $EENS_i$ corresponds to the expected energy non-served when a small increment of resource *i* is added to the system.
- $EENS_{Perf}$ corresponds to the expected energy non-served when a small increment of perfect generation is added to the system.

To calculate the de-rating factor of each resource according to the second-best approach, its expected (average) production during instances of non-served energy ($Prod\_ENS_i$) was compared to its maximum potential output during those instances (its installed capacity, K_i). This can be represented as:

$$DF_i = \frac{\mathbb{E}(Prod\_ENS_i)}{K_i}$$

The de-rating factors resulting from both the first-best and second-best approaches and the differences between them are presented in Table 3.2.

Table 3.2: De-rating factors according to the first-best and second-best approaches

Technology	First-best de-rating factor (p.u.)	Second-best de-rating factor (p.u.)	Relative difference
1-hour battery	0.2605	0.2582	0.89%
4-hour battery	0.7355	0.7288	0.92%
8-hour battery	0.8573	0.8538	0.40%
CCGT	0.9304	0.9314	0.11%
Coal	0.8683	0.8699	0.18%
Nuclear	0.9818	0.9820	0.03%
PV	0.0869	0.0891	2.38%
Wind	0.1321	0.1338	1.27%

As can be observed in the right column of Table 3.2, the differences between the first-best and second-best approaches are small, while the advantages of the second-best approach reduce the computational burden by a factor of N+2, where N accounts for the number of resources being analysed. In this case, ten executions were needed for the first-best approach but only one for the second-best approach to obtain the de-rating factor of eight resources. The differences in precision between both methodologies are explained by the fact that the first-best approach used in this case, and in the majority of power systems, analyses incremental contributions and not true marginal contributions.

3.2.4. Historical vs. predicted future contributions

Many power systems that have implemented capacity remuneration 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). However, power systems are evolving rapidly with the introduction of new generation and storage technologies and a rise in demand elasticity. It is then reasonable to infer that historical data may not be representative of future power system operation. If firm supply is calculated based on historical data, regulators run the risk of remunerating resources that may be unable to contribute to countering expected future scarcity events[26].

If capacity mechanisms aim to meet future reliability targets, firm supply should be based on projections of future power system operation. This involves the utilisation of a model to simulate operation in different scenarios (varying hydro inflows and renewable production, among others) and subsequently analysing each resource's contribution to the established target. This concept, which is hardly new, underlies methodologies such as the aforementioned ELCC (CPUC, 2014).

Evidently, the results of simulations are contingent on the initial assumptions and the scenarios established, although such problems are common to all models and methods. The calculation of firm supply from historical data is predicated on the assumption that scarcity conditions will not change in the future, a premise over which energy transition is casting doubts. As noted in subsection 3.2.1, 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

[26] 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.

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, 2017 and National Grid, 2019b), seem to prefer to calculate firm supply on the grounds of projections.

3.2.5. 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, as the technical characteristics of resources sharing a given technology may differ from one plant to another (such as position in the grid or availability of the primary energy source).

The trouble with this approach is that, while it is theoretically robust, its real-life application encounters a significant drawback. Assessing the marginal contribution of each resource calls for a 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 a tolerance that is too large to evaluate the expected marginal contribution with any precision (see earlier comment in subsection 3.2.3).

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. This approach has been adopted by a number of countries, including the United Kingdom (National Grid, 2019a), Belgium (Elia, 2019a), Ireland (SEMC, 2018) and Italy (Mastropietro et al., 2018). 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.

However, it should be noted that such approximations may be inaccurate for other technologies, particularly those which are non-conventional. 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 system 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[27] (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

[27] 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).

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, however. Recent stress events, such as the extreme weather conditions that affected Texas in February 2021 (ERCOT, 2021), may suggest a high correlation between the outage rates of individual thermal power units, such as combined cycle power plants, which should not be handled separately when calculating their firm capacity (EPRI, 2021). Similarly, during periods of extreme heat, there is a possibility that nuclear power plants may cease operation simultaneously, as in France and Germany in the summer of 2019 (Reuters, 2019).

Conversely, the calculation of a single de-rating factor for an entire technology may be inefficient for hydroelectric power plants. This 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 equally to the reliability target. 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, so that it conveys signals that efficiently incentivise the agents concerned (see Faria et al., 2009, for a discussion of alternatives).

3.2.6. Single vs multiple products

Another element in the design of firm supply calculation methodology revolves around whether the capacity mechanism envisages a single or multiple products. While 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.

In circumstances where two or more reliability products are assumed, the underlying rationale 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. This approach involves the decomposition of the adequacy assessment into sub-problems to perform firm supply calculations based on the reliability standard. 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 range of products in the capacity mechanism, coupled with the subsequent calculation of several firm supply values, could benefit resources that can contribute better to solving scarcity events addressed by one or more of the reliability products. 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[28]). It is also noteworthy that transferring the seasonality risk to market agents may favour large generation companies that own many resources that use a wide array of different technologies, as they would be able to offset the seasonal production of one resource with the output of others in their portfolio.

3.3. Conclusions and policy implications

The calculation of the firm supply of resources is one of the most important design elements of a capacity remuneration mechanism. It is crucial for both the system operator to properly assess how different resources will be able to contribute to solve scarcity conditions in the system, and for investors, as it will determine the possible remuneration for their investments.

This chapter has presented the main design elements of the firm supply calculation process and has drawn recommendations for each so that the firm supply calculation process accurately reflects how different resources will help in solving adequacy concerns in the system.

This leads to the following conclusions:

- The firm supply calculation methodology must be based on the same reliability metric used to establish the reliability standard and ideally be the same for all resources and considering how different resources contribute to solve scarcity conditions, according to this metric, as part of the system.
- 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 Annex I.

[28] 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.

- 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. This has been demonstrated through a case study.
- To facilitate the calculation of firm supply, resources could be grouped by technology to deliver a single de-rating factor for each technology. This 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, which can be true for renewable energy sources, but also for conventional generation (CSMEM, 2016; Freeman et al., 2020).
- The decision of using a single reliability product or multiple should be envisaged in the calculations calls for balancing a number of factors, including the simplicity and transparency of the method, as well as the management of risk and uncertainty. 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.

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 3.2 (projections vs. historical data; marginal vs. mean contribution; per-resource vs. per-technology de-rating; annual vs. seasonal/monthly de-rating).

4. MULTI VS. COMPOSITE-METRIC RELIABILITY STANDARDS

4.1. Introduction[29]

Chapter 2 described the rationale for a deep revision of resource adequacy metrics. Several experts are calling for a fundamental reform of these metrics (EPRI, 2022; ESIG, 2024), both in terms of the type of contingency they assess and the statistical measure to condense the information from different scenarios (De Vries and Sánchez Jiménez, 2022; Stephen et al., 2022). This dynamic is also affecting regulatory activity. Brazil (EPE, 2020b), MISO (MISO, 2024) or Pacific Northwest (NPCC, 2019)) in the United States, the United Kingdom (DESNZ, 2023) or Australia (AEMC, 2022) are just a few examples of power systems that have reformed or are in the process of reforming their resource adequacy assessments and the metrics on which they are based.

One of the most common arguments found in the academic literature to justify a reform of resource adequacy assessments is that a single resource adequacy metric cannot characterise the new and evolving scarcity conditions (EPRI, 2024). A single metric could represent two scarcity events in the same way, which may, in practice, be significantly different and have very different impacts on electricity consumers. For example, Figure 4.1 shows different scarcity conditions, with each blue block representing a 1-MWh shortfall. Cases A and B have the same LOLE, but very different EUEs, while the opposite is true for cases C and D.

[29] This chapter heavily draws from Brito-Pereira, P., Bruninx, K., De Vries, L., Mastropietro, P., Rodilla, P. Future-proofed resource adequacy metrics: a model-based assessment of multi-metric vs. composite-metric reliability standards. Sustainable Energy, Grids and Networks, Volume 44, Article 101957.

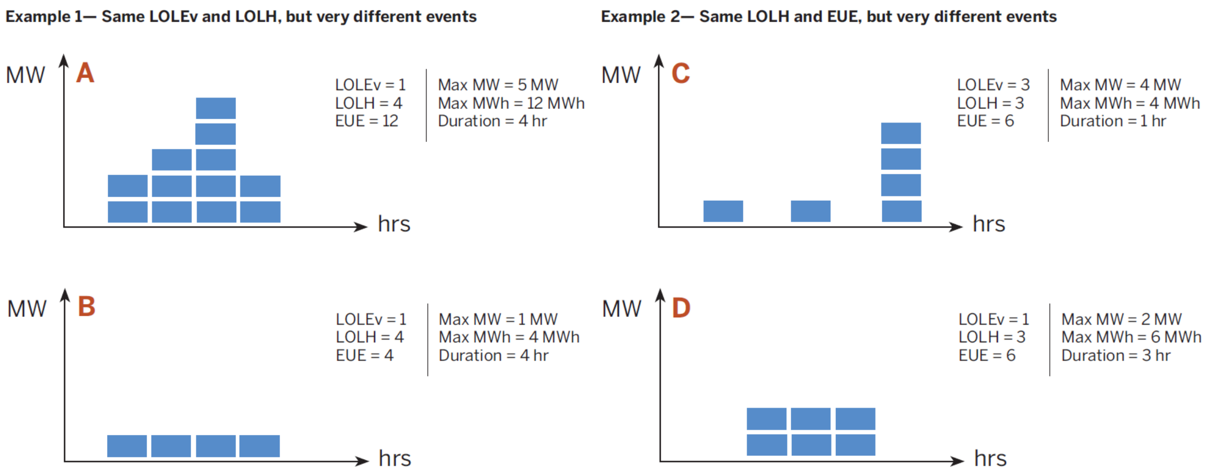

Figure 4.1: Example of the need for a multi-metric approach (ESIG, 2021)

According to several experts, the solution to this problem would be to move to a multi-metric approach (EPRI, 2024). Resource adequacy assessments would use a set of different resource adequacy metrics and evaluate the compliance of the resource mix with multiple reliability standards. A multi-metric approach would make it possible to examine the different facets of security of electricity supply and provide the system operator and the policymaker with more information on the nature of potential shortfall events.

However, when the resource adequacy assessment informs the design of a capacity mechanism, a multi-metric approach may have a significant drawback. In capacity remuneration mechanisms, resources should be remunerated according to their ability to help the system meet the reliability standard by providing their contribution during the expected scarcity conditions. This means that the reliability product to be traded in the capacity mechanism is conceptually related to the reliability standard that the system must comply with. As resources cannot guarantee to be fully available during scarcity events, all resources willing to participate in a capacity mechanism are assigned a de-rating factor that is used to calculate their "firm capacity" (ESIG, 2023). A 100 MW resource with a 30% de-rating factor would only be allowed to trade 30 MW in the capacity mechanism. In theory, the de-rating factor should reflect each resource's contribution to meeting the reliability standard set by the regulator. If the system has to meet multiple reliability standards, then each resource should be assigned multiple de-rating factors (one for each standard), and the capacity mechanism should be based on the procurement of multiple reliability products. However, there is no obvious solution to coordinate the procurement of these different reliability products and inefficiencies may arise[30].

One possible solution may come from the so-called composite metrics. These metrics simply combine different primary resource adequacy metrics through weighting factors to build a

[30] The possibility of introducing a multi-product capacity market in the Chilean power sector was assessed but rejected exactly due to these complexities and potential inefficiencies (Rodilla et al., 2018).

single metric. For instance, a composite metric based on the non-served energy but assessed by two different statistical measures, i.e., the expectation and the CVaR5%, has been proposed by the Australian regulator (AEMC, 2022; Mancarella, 2022) as reflected in the following expression.

$$w \cdot EENS + (1-w) \cdot CVaR_{5\%}(ENS) \tag{1}$$

Composite metrics allow different facets of the security of supply problem to be captured in a single reliability standard. An advantage of this approach is that the associated de-rating factors for participation in the capacity mechanism can be calculated according to the contribution to this single reliability standard. Therefore, each technology or resource would be assigned a single de-rating, simplifying the design of the capacity mechanism. A disadvantage of a composite-metric standard, also analysed in this chapter, is that it may not guarantee that all the single standards that compose the composite-metric standard are met.

The objective of this chapter is to compare multi-metric and composite-metric reliability standards and to quantitatively assess, through some illustrative case studies, the impact of the two approaches on the resulting resource mix and de-rating factors to be used in a capacity mechanism. De-rating factors are analysed as the design element of the reliability product to be traded in the capacity mechanism that is more affected by the selection of the resource adequacy metric. This is achieved through a stochastic expansion planning model[31], where the optimisation is constrained using different resource adequacy metrics, first by a multi-metric approach and then by setting a composite-metric reliability standard. The modelling exercise complements the theoretical analysis presented in this introduction and allows the effects of the choice of metrics for the reliability standard to be clearly demonstrated. The resource adequacy metrics used in the chapter to build the multi-metric and composite-metric standards are the same as those used in the Australian proposal, i.e., expected non-served energy, EENS, and CVaR5% of the non-served energy, CVaR5%(ENS). However, the main findings of the chapter are not affected by the choice of the underlying resource adequacy metrics.

Although experts and regulators have proposed multi-metric and composite-metric approaches, their application has not been independently tested in a model-based comparative analysis. The findings of this chapter could inform the current debate on the reform of resource adequacy frameworks and the discussions raised in the case studies may be relevant to both system planners and regulators.

The remainder of this chapter is structured as follows. Section 4.2 presents the methodology and case studies used to compare multi-metric and composite-metric approaches. Section 4.2.1

[31] In this chapter, the central planner model with resource adequacy constraints is used to simulate an electricity market with perfect competition and an incentive to install firm capacity. For further details on this equivalence, the reader may refer to Annex I or (Pérez-Arriaga and Meseguer, 1997).

presents and discusses the results of the simulations. Section 4.3 summarises the main findings of the chapter and outlines areas of research that could be explored in future work.

4.2. Methodology

The simulation model used to compare different resource adequacy metrics is based on stochastic expansion planning with 500 scenarios, which optimises the resource mix from scratch (greenfield) to supply a given load curve at the minimum expected cost (see Annex III for the detailed mathematical formulation). The load curve is inspired by the real demand of the Spanish power sector and consists of 672 hourly values, corresponding to four weeks, one for each season of the year. The corresponding load duration curve is shown in Figure 4.2.

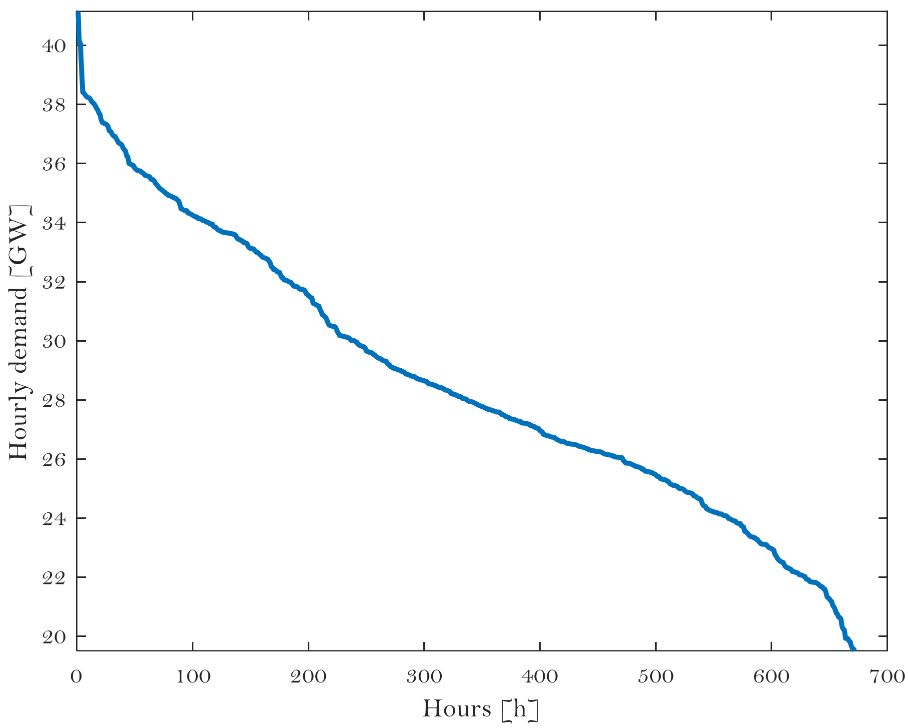

Figure 4.2: Load duration curve of the demand used in the stochastic expansion planning

To serve this demand curve, the model can decide to install and dispatch five different technologies: nuclear, combined cycle gas turbines (CCGTs), wind, solar photovoltaic and diesel[32]. The cost of unserved energy (acting in the model as a price cap) is set at 3 000 €/MWh. The stochastic simulation considers 500 scenarios of availability of thermal resources (nuclear, CCGT and diesel), modelled by availability matrices built by a two-state Markov chain using a Monte Carlo approach (Lisniaski et al., 2012; Mastropietro et al., 2016). These availability matrices have been modified to simulate the effects of a polar vortex similar

[32] Investment and variable costs for these technologies have been taken from Lazard (2023).

to the one that hit the ERCOT power system in 2021 (Busby et al., 2021). In 2.5 % of the scenarios, CCGT availability drops by 80 % for three consecutive days. Similar cold snaps, albeit with less dramatic impacts, have affected other electricity systems in the United States in the past, such as the polar vortex experienced in the Midwest, South Central, and East Coast regions in 2014 (NERC, 2024). The polar vortex is used in the model to represent an extreme weather event that may affect the tail risk in the resource adequacy assessment, decoupling the different resource adequacy metrics assessed in the simulations to better illustrate the findings of the research. However, these findings can be generalised to any type of scarcity condition (see Section 4.3.3.1 for a sensitivity analysis of the impact of the polar vortex). The profiles of demand and availability of renewable resources are the same in all scenarios.

Several simplifications (four weeks to represent one year, deterministic profiles for demand and availability of renewable energy sources, absence of energy-limited resources among the potential technologies, etc.) have been introduced in the modelling exercise to meet computational constraints. However, none of these simplifications are expected to affect the findings of the chapter. The aim of the simulation model is not to represent or predict the real operation of a power system, but rather to compare multi-metric and composite-metric reliability standards through illustrative, yet realistic, case studies that allow their performance to be assessed.

4.2.1. Case studies

Four case studies are generated by the simulation model, as shown in Figure 4.3. In the first case study, the stochastic expansion planning is run without any reliability standard, i.e., without any constraint on resource adequacy. In this base case, the model simply seeks an economic equilibrium between increasing the cost of new entrants and reducing the cost of unserved energy valued at the 3 000 €/MWh price cap. The outcome is evaluated quantitatively in terms of EENS and $CVaR_{5\%}$(ENS), and the resulting resource mix is set as the reference for comparison with the other case studies.

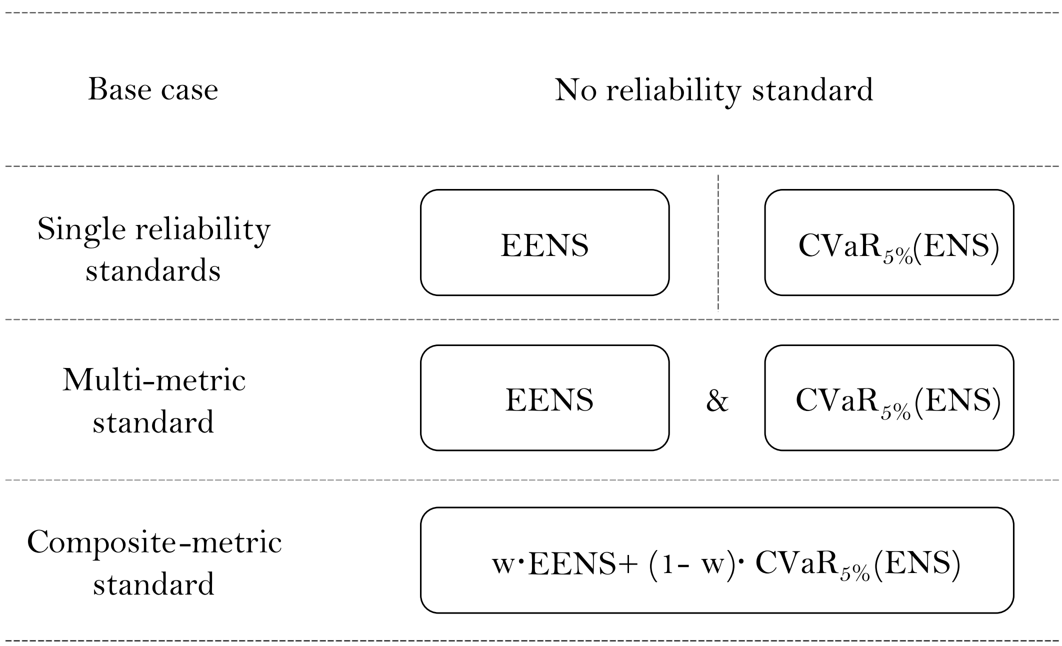

Figure 4.3: Schematic representation of the four case studies

In the second case study, the model is run with single reliability standards. The EENS and $CVaR_{5\%}$(ENS) constraints are imposed separately in the stochastic expansion planning, creating two sub-cases. The outcome of each sub-case is evaluated by examining the variation in installed capacity of different technologies with respect to the base case and their de-rating factors. De-rating factors, in this and other case studies, are calculated as the contribution of each technology to the reliability standard using an ELCC methodology: i) first, the installed capacity of a given technology is marginally incremented and the resulting variation in the resource adequacy metric is calculated; ii) second, the same marginal increment in perfect generation (i.e., a unit with full availability and no outages) is added to the resource mix and the resulting variation in the resource adequacy metric is calculated; iii) the two variations are compared and the de-rating factor is calculated as their ratio; iv) the process is repeated for each technology.

In the third case study, the model is run with a multi-metric reliability standard, i.e., the EENS and $CVaR_{5\%}$(ENS) constraints are imposed together, with reliability standards that result in the simultaneous activation of both constraints. Again, the outcome is evaluated by examining the variation in installed capacity of different technologies with respect to the base case and their de-rating factors. In this case, since two reliability standards are active at the same time, two de-rating factors are calculated for each technology (see Section 4.3.3).

In the fourth case study, the model is run with a composite-metric reliability standard. A weighting factor, w, is used to combine EENS and $CVaR_{5\%}$(ENS) into a single metric and a constraint is applied based on this. The reliability standard and the constraint is formulated as in Equation 2, with $EENS_{RS}$ and $CVaR_{5\%}(ENS)_{RS}$ being the standards used in the second and third case studies. The normalisation of each resource adequacy metric by its standard allows metrics that may have different orders of magnitude, or even different units of measurement, to be combined, as shown in equation 2.

$$w\cdot\left(\frac{EENS}{EENS_{RS}}\right)+(1-w)\cdot\left(\frac{CVaR_{5\%}(ENS)}{CVaR_{5\%}(ENS)_{RS}}\right)\leq 1 \tag{2}$$

The model is run for values of the weighting factor ranging from 0, corresponding to a $CVaR_{5\%}(ENS)$ constraint, to 1, corresponding to an EUE constraint. Again, the outcomes are evaluated by examining the variation in installed capacity of different technologies relative to the base case and their de-rating factors.

4.3. Results and discussion

This section presents the main results of the simulation model, divided according to the four case studies identified in Section 4.2.

4.3.1. Base case – no reliability standard

As discussed in Section 4.2, in the absence of a resource adequacy constraint, the stochastic expansion planning seeks an economic equilibrium between increasing the cost of new entrants and reducing the cost of unserved energy. The resulting resource mix is shown in Figure 4.4, while Figure 4.5 presents the dispatch of a typical day. As can be seen, nuclear is used for baseload, CCGTs are responsible for supporting the variability of the large installed renewable capacity, especially wind, while diesel is used as a peaker, especially in the polar vortex scenarios when CCGTs are not available.

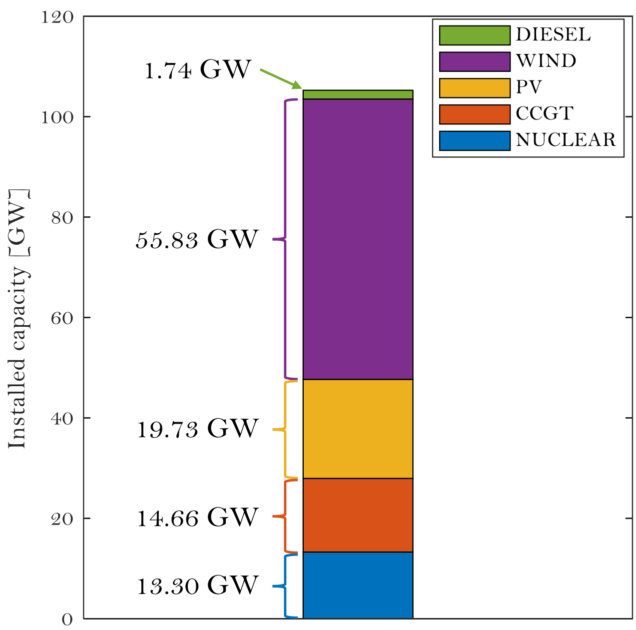

Figure 4.4: Resource mix in the base case

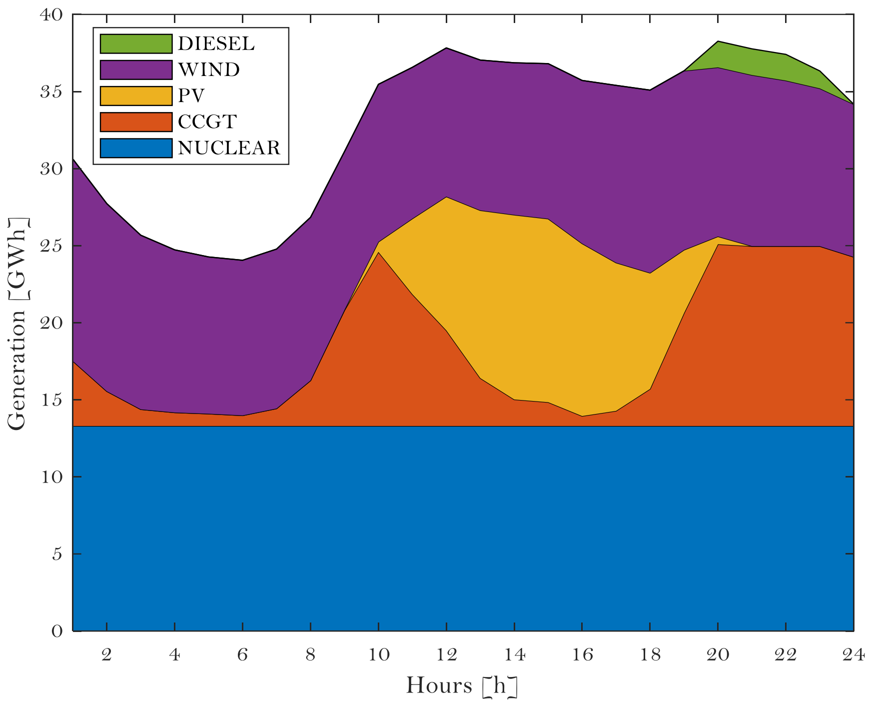

Figure 4.5: Dispatch of different technologies in a typical day

The resource adequacy of the base case is assessed using the EENS and the $CVaR_{5\%}$(ENS). In the upper chart of Figure 4.6, these parameters are plotted together with the unserved energy for each scenario, ordered by increasing values. It can be observed how the unserved energy grows slowly until the polar vortex scenarios are reached, where it becomes much higher. In the lower chart of Figure 4.6, the same information is presented as a probability density distribution function, which makes it easier to visualise the so-called tail risk, represented by events with low probability but high magnitude (very high unserved energy).

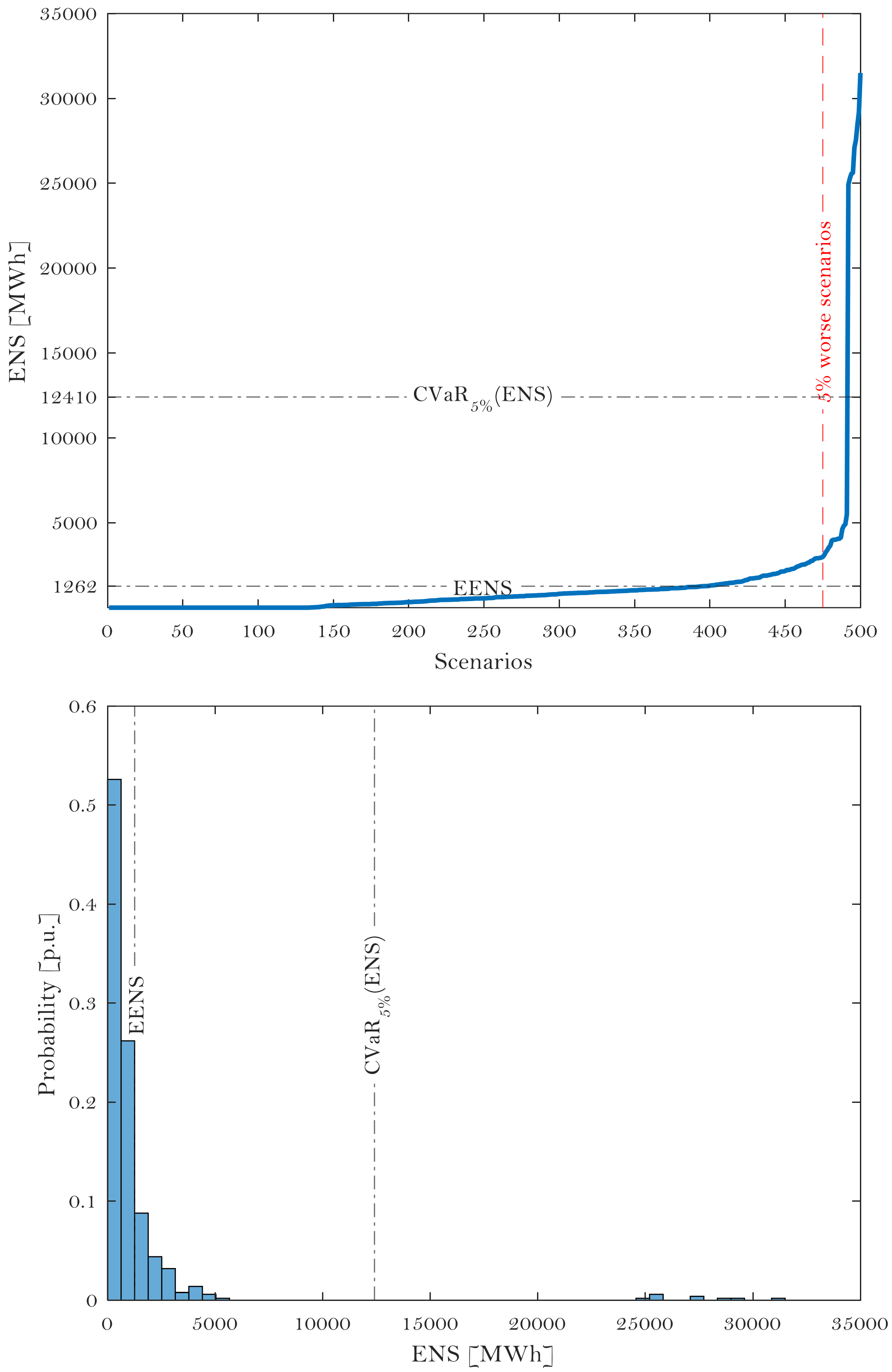

Figure 4.6: Resource adequacy assessment for the base case

The CVaR$_{5\%}$(ENS) is ten times higher than the EENS because it focuses on the 5% scenarios with the highest unserved energy.

4.3.2. Single metric reliability standard

In this case study, the stochastic expansion planning is run twice with different resource adequacy constraints, one applied to the EENS and one to the $CVaR_{5\%}$(ENS). The outcomes are shown in Figure 4.7, with blue representing the results of applying the EENS constraint and red representing the results of applying the $CVaR_{5\%}$(ENS) constraint. Here, as in other charts, attention is focused on the two technologies that are more affected by a change in the resource adequacy metric used to set the reliability standard, i.e., CCGTs, which suffer severe unavailability in polar vortex scenarios, and diesel, which is the peaking technology called upon by the model to make up for the missing CCGT capacity.

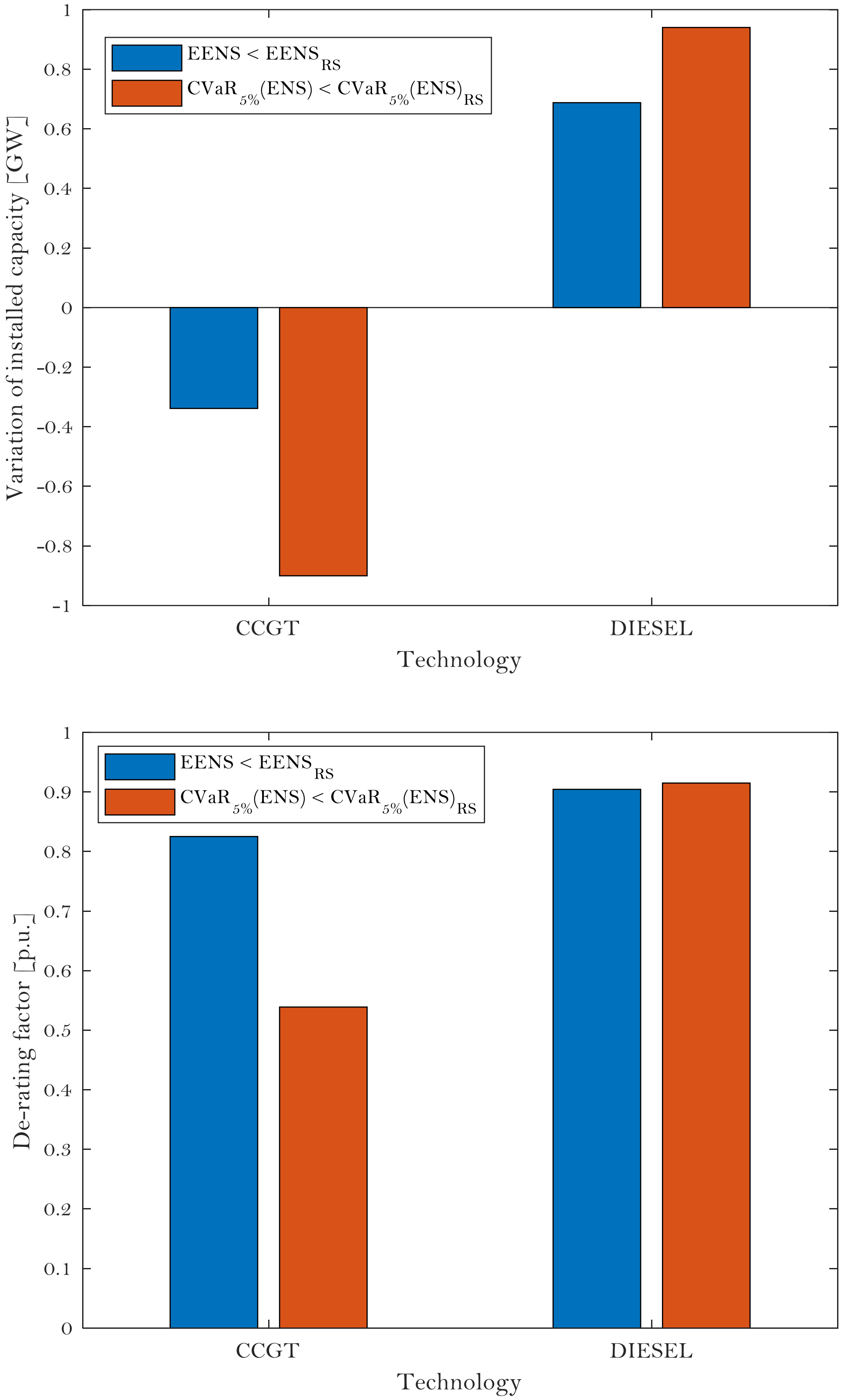

Figure 4.7: Outcomes of the imposition of single-metric standards

When a resource adequacy constraint is imposed, i.e., when the model seeks a more reliable resource mix, the installed capacity of CCGTs decreases, due to their poor performance in the polar vortex scenarios. However, the drop is much larger for the $CVaR_{5\%}(ENS)$ constraint than for the EENS constraint. In fact, when a $CVaR_{5\%}(ENS)$ constraint is imposed, the model is asked to reduce the unserved energy in the worst scenarios, which encompass the polar vortex. Therefore, more CCGTs are replaced by diesel units, as shown in the upper chart in Figure 4.7. For the same reason, the de-rating factor assigned to CCGTs drops from over 80% when the reliability standard is based on EENS to 55 % when the constraint is based on $CVaR_{5\%}(ENS)$. The diesel de-rating is high in both sub-cases but is slightly higher when the reliability standard is based on $CVaR_{5\%}(ENS)$.

4.3.3. Multi-metric reliability standard

In this case study, the stochastic expansion planning is required to meet two reliability standards simultaneously. The first step is to find values for the two resource adequacy constraints that allow this simultaneous activation. To find these values, the multi-metric domain was characterised as shown in Figure 4.8. Each point of the multi-metric domain represents an EUE-$CVaR_{5\%}(ENS)$ pair that could be imposed in a multi-metric reliability standard. The starting point is the base case, which defines the initial values for the EENS and the $CVaR_{5\%}(ENS)$. From this point, the blue line is drawn by constraining the EENS and assessing the $CVaR_{5\%}(ENS)$ of the resulting mix, repeating this operation for small decrements in EUE. The red line is drawn in a similar way by constraining the $CVaR_{5\%}(ENS)$ and assessing the EENS of the resulting mix.

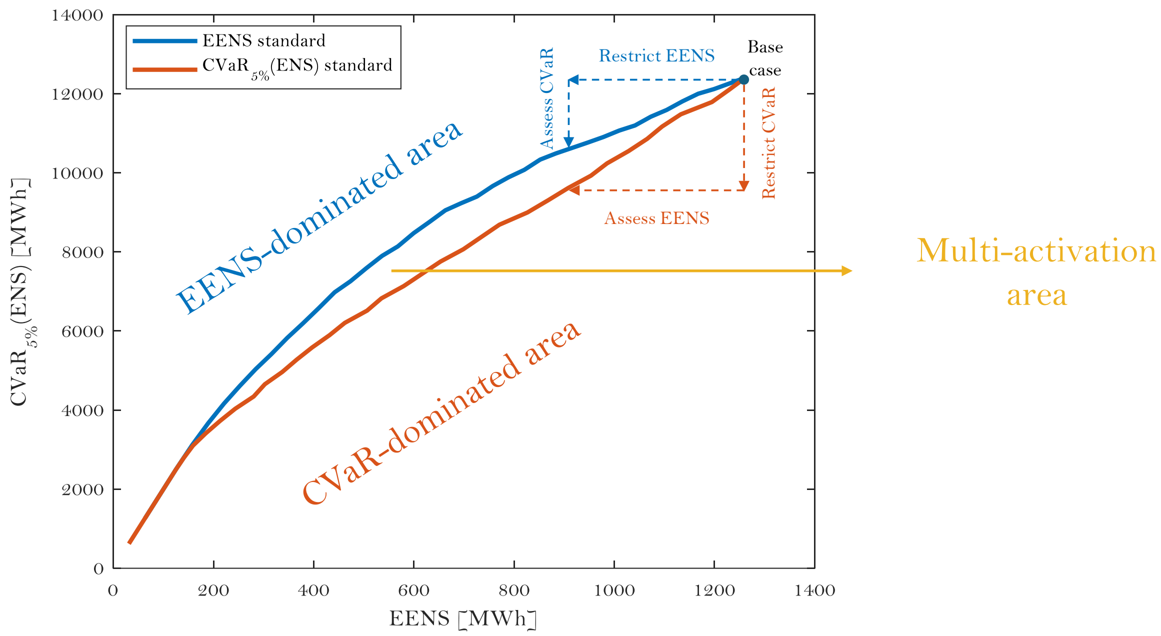

Figure 4.8: Characterisation of the multi-metric domain and identification of the multi-activation area

The two lines allow three areas to be identified in this double-metric domain[33]. In the EENS-dominated area, the imposition of two reliability standards would make the $CVaR_{5\%}$(ENS) standard redundant as only the EENS constraint would be activated. In the CVaR-dominated area, only the $CVaR_{5\%}$(ENS) constraint would be activated and the EENS standard would be redundant. However, in the area between the blue and red lines, referred to here as the multi-activation area, both resource adequacy constraints would be activated simultaneously. To demonstrate this multiple activation, the stochastic expansion planning was run with two reliability standards corresponding to the point[34] shown in Figure 4.9, which is in the multi-activation area.

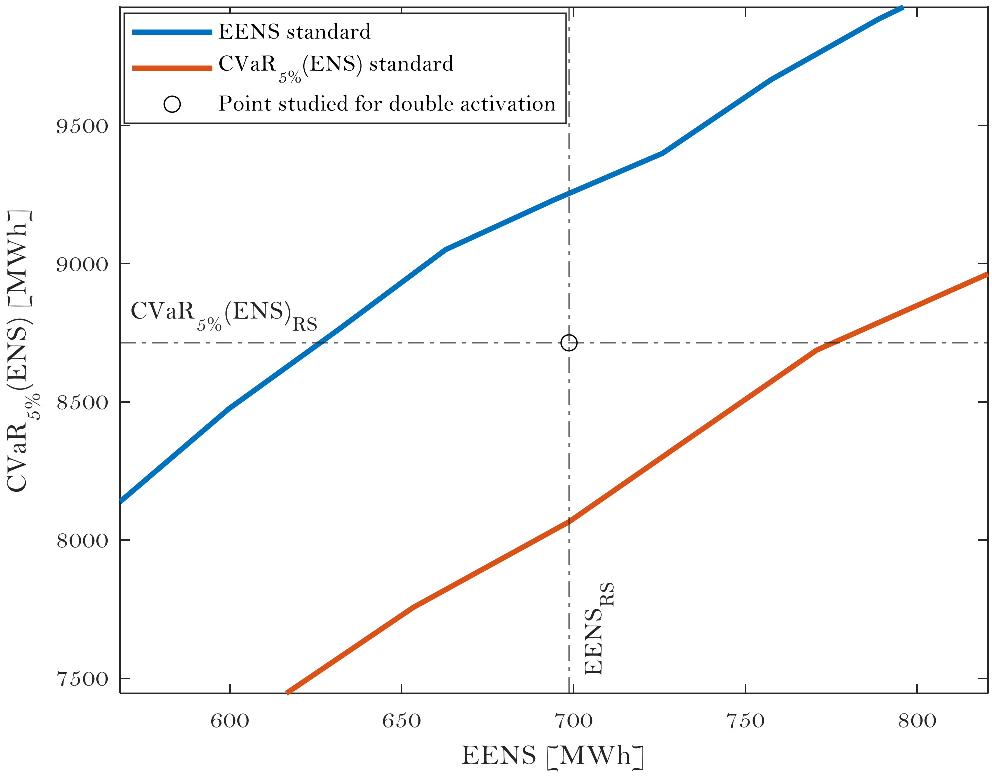

Figure 4.9: Multi-activation area and reliability standards under study

The outcomes of applying this multi-metric standard are shown in yellow in Figure 4.10, where they are compared with the results of applying single-metric standards. From a

[33] In this chapter, the multi-metric domain is composed by two resource adequacy metrics and this results in a bi-dimensional multi-activation area. Multi-metric or composite-metric reliability standards can also consider three or more resource adequacy metrics, resulting in multi-activation volumes or hypervolumes.

[34] This point was also used to define the reliability standards in the previous case study, where the resource adequacy constraints were applied separately. The same reliability standards are also the starting point for the following case study, where the constraints are combined. This is done to ensure comparability of results.

computational point of view, the activation of both constraints has been demonstrated by their dual variables, both of which assume a non-zero value.

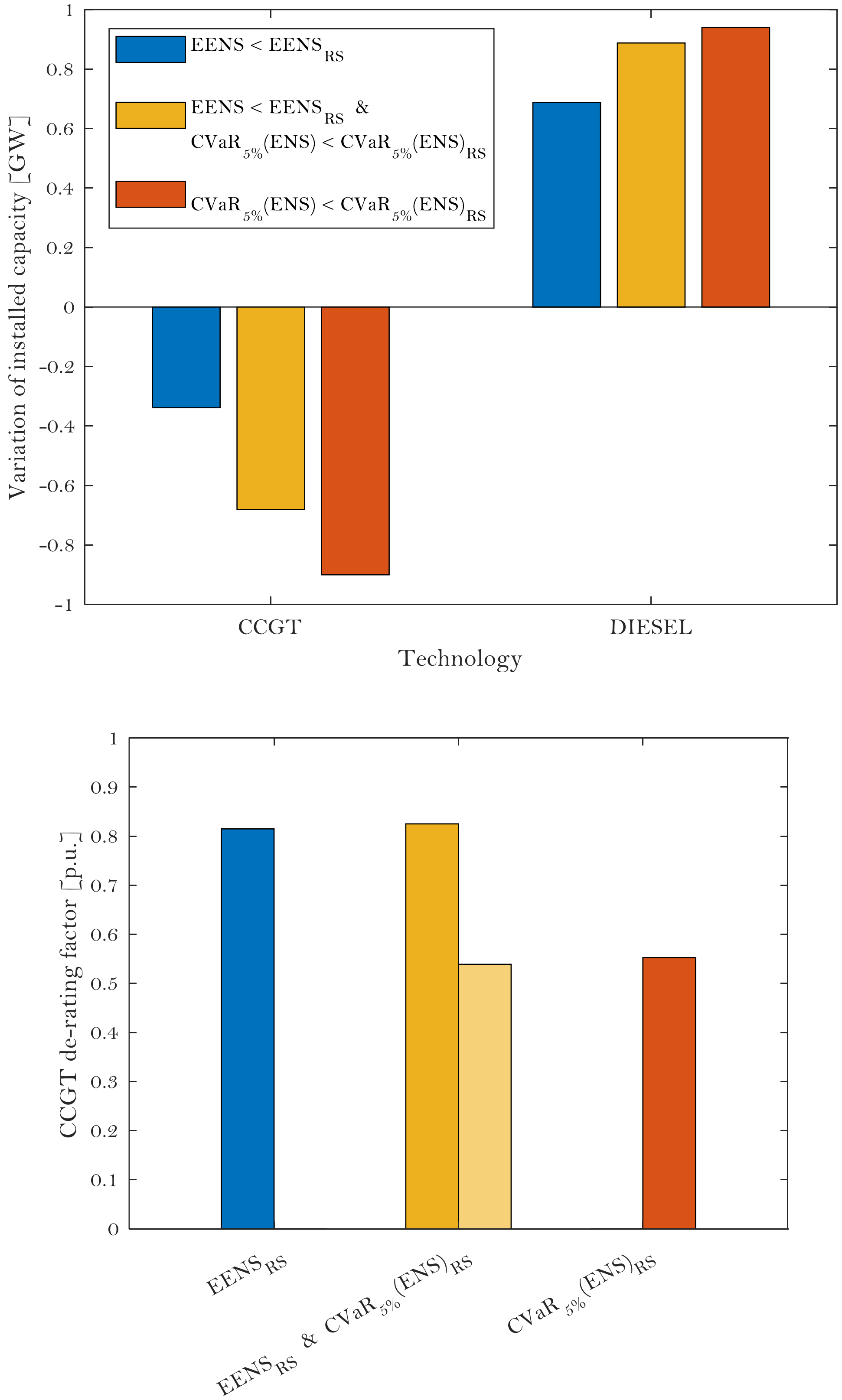

Figure 4.10: Outcomes of the imposition of a multi-metric standard

The resource mix resulting from the application of the multi-metric standard shows an intermediate variation in installed capacity compared to the single-metric standards. However, as shown in the lower chart in Figure 4.10, the application of a multi-metric standard results in the calculation of two different de-rating factors for a single technology, in this case CCGT, with a higher de-rating factor corresponding to its contribution to the EUE standard and a lower de-rating factor corresponding to its contribution to the $CVaR_{5\%}$(ENS) standard. In the capacity mechanism, these two de-rating factors should be reflected in the procurement of two reliability products. Procurement of multiple products in the capacity market forces the regulator to define different requirements and may lead to different prices for different reliability products, increasing the complexity of the design of the regulatory instrument (ERAWA, 2024)

4.3.3.1. Sensitivity analysis without polar vortex

It should be remarked that resource adequacy metrics tend to be highly correlated, meaning that a scarcity event is likely to affect different metrics in a consistent way. In the case studies presented here, the polar vortex creates two types of scarcity conditions in this particular power system: shortages simply related to sporadic outages of thermal power plants, and shortages related to an extreme weather event that mostly affects a single technology (CCGT). This approach made it possible to reduce the correlation between the two resource adequacy metrics under study, EENS and $CVaR_{5\%}$(ENS). This has an impact on the multi-activation area, which becomes larger when metrics in the multi-metric domain are less correlated. To investigate the impact of the correlation between metrics on the multi-activation area, as well as to study the generalisation of the concepts introduced in this chapter, the model was run without polar vortex scenarios. The result is shown in Figure 4.11, which has the same axis ranges as Figure 4.8 for comparison.

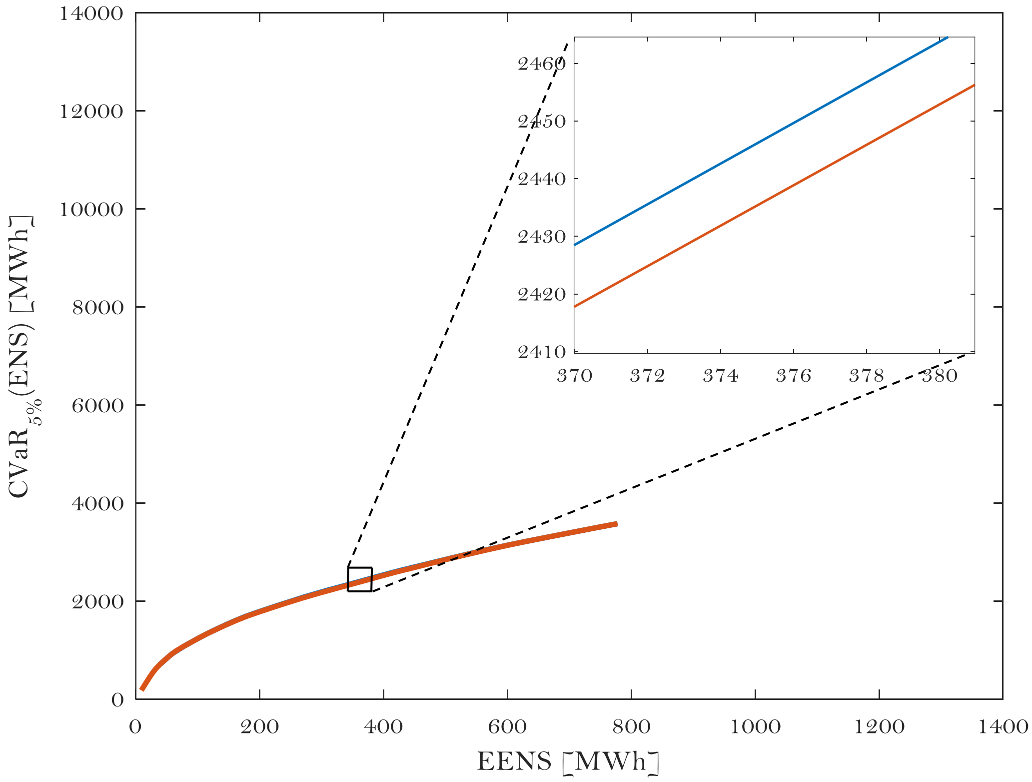

Figure 4.11: Characterisation of the multi-metric domain for a case study without polar vortex

The base case without polar vortex obviously shows lower values for both EENS and $CVaR_{5\%}$(ENS). However, it can also be observed that the multi-activation area, although present (as shown in the zoomed-in graph), is much thinner than in the polar vortex case study. This means that, without the polar vortex, the two resource adequacy metrics are much more correlated. The imposition of a multi-metric standard, in this case, is much more likely to result in the activation of only one of the two initial standards.

Of course, the size of the multi-activation area depends on the type of scarcity conditions expected in the system and the type and number of metrics selected to assess resource adequacy in the system. However, any discussion on the use of multi-metric standards should consider the possibility of several redundancies and the eventual activation of a single reliability standard. Additionally, the persistence of the multi-activation area, despite the lack of purposeful differentiation of scenarios in this sensitivity analysis, lends weight to the generalisability of the case study presented in this chapter.

4.3.4. Composite-metric reliability standard

The last case study examines the application of a composite-metric standard. The resource adequacy constraint is imposed by the normalised formula presented in equation 2. The model is run for different values of the weighting factor and the results in terms of the adequacy of the resulting resource mix are shown in the multi-metric domain in Figure 4.12.

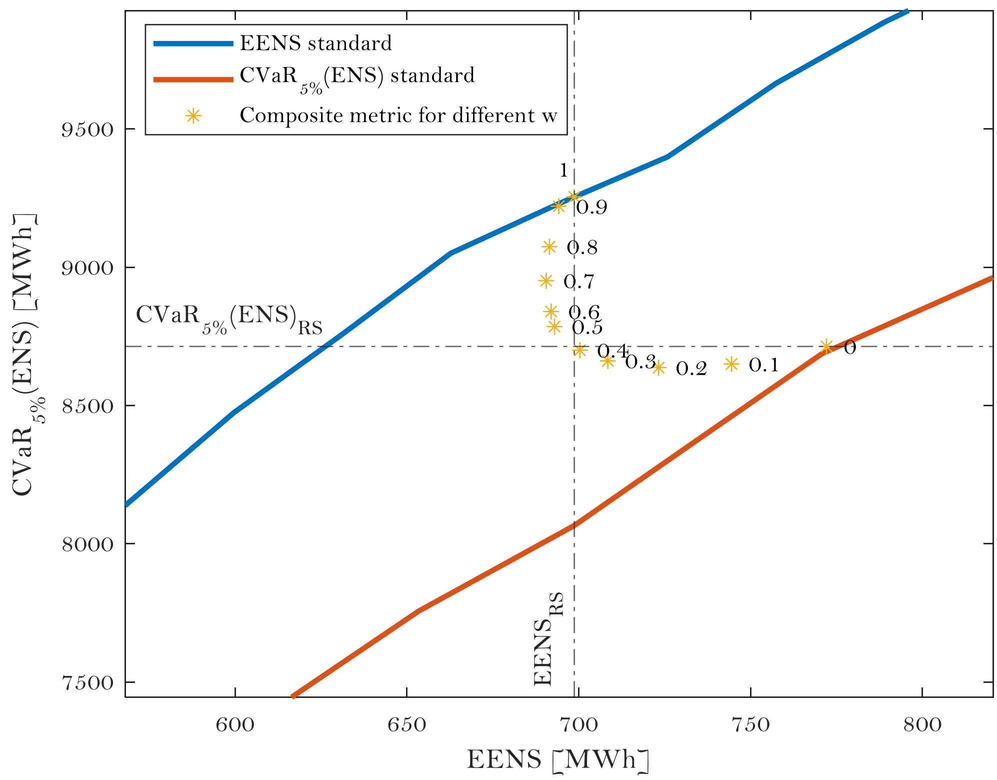

Figure 4.12: Outcomes of the imposition of a composite-metric standard represented in the multi-activation area for different values of the weighting factor

When the weighting factor is equal to 0, the composite-metric standard coincides with the $CVaR_{5\%}$(ENS) constraint and the resulting point in the multi-metric domain lies on the red curve. When the weighting factor is equal to 1, the composite-metric standard coincides with the EENS constraint and the resulting point lies on the blue curve. For values of the weighting factor between 0 and 1, the optimisation model is allowed to seek an equilibrium between the two initial reliability standards in the pursuit of cost minimisation and compliance with the composite-metric standard. For low values of the weighting factor (between 0 and 0.4 in this case study), the $CVaR_{5\%}$(ENS) constraint is still dominant and the resulting resource mix exceeds the initial $CVaR_{5\%}$(ENS) standard (being below that reference) but is above the initial EENS standard. For high values of the weighting factor (between 0.5 and 1 in this case study), the EENS constraint becomes dominant and the resulting resource mix exceeds the initial EENS standard but is above the initial $CVaR_{5\%}$(ENS) standard.

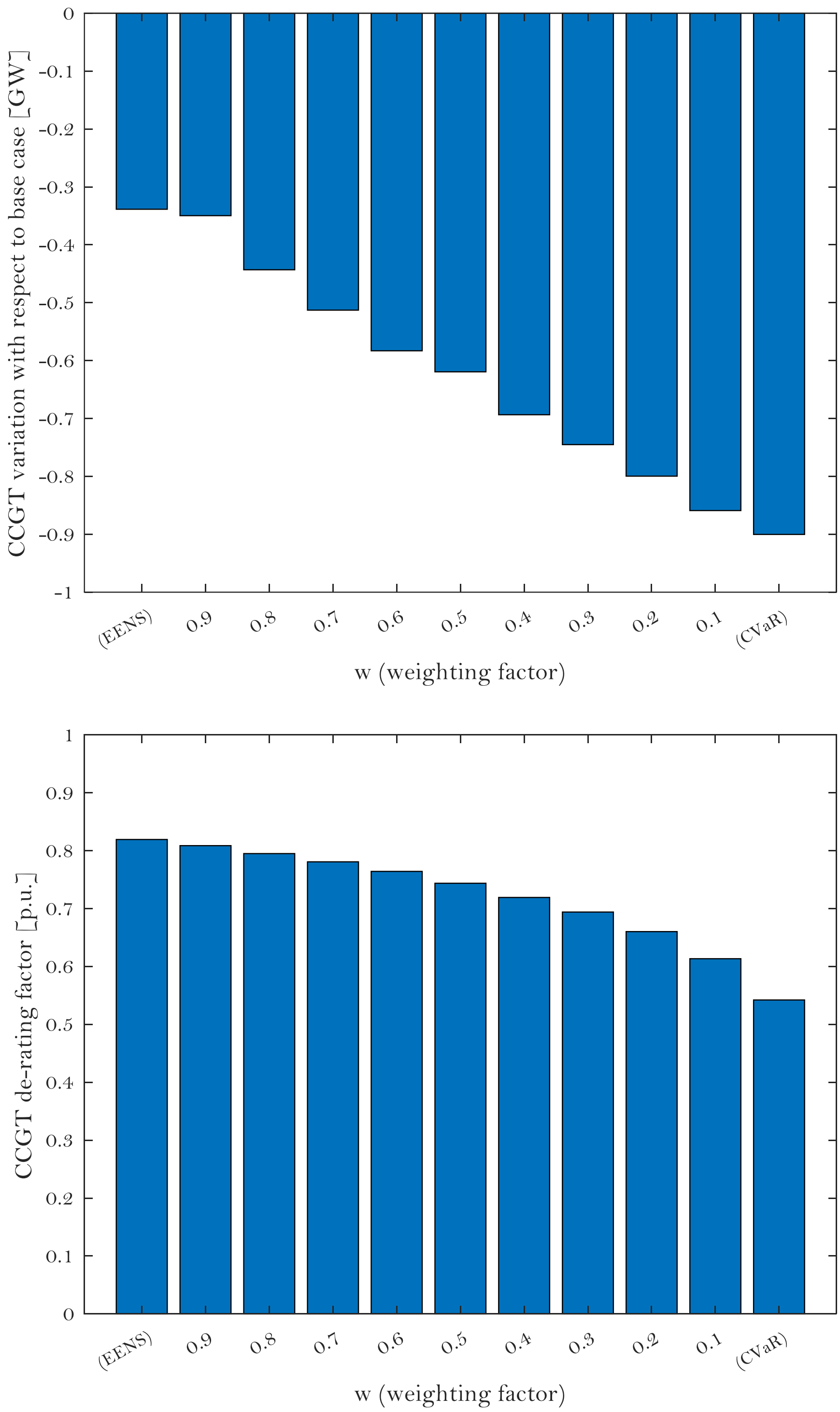

Figure 4.13: Outcomes of the imposition of a composite-metric standard on the CCGT technology for different values of the weighting factor

The outcomes in terms of installed capacity and de-rating factors with the application of a composite-metric standard are shown in Figure 4.13, focusing on combined cycles for the sake of clarity. Both the installed capacity of CCGTs and their d-rating factor (calculated as their contribution to the composite-metric reliability standard) gradually decrease as the weighting factor goes from 1, corresponding to the EENS constraint, to 0, corresponding to the

$CVaR_{5\%}$(ENS) constraint. This behaviour is consistent with the effects shown in the previous case studies.

The main difference between the application of a composite-metric and a multi-metric reliability standard is that the composite-metric standard allows a single de-rating factor to be calculated for each technology. If a capacity mechanism is implemented, this feature allows the procurement of a single reliability product, avoiding the complexities of a multi-product capacity market.

4.4. Conclusions and policy implications

The traditional resource adequacy framework is being challenged by the decarbonisation of the power sector. New correlations in the resource mix and extreme weather events are rapidly changing the expected scarcity conditions in the electricity system, and conventional resource adequacy metrics are revealing their limitations in this new environment. Several experts are calling for a reform of resource adequacy metrics and reliability standards, and many policymakers have already started to update their regulations. Some experts argue that a single resource adequacy metric cannot characterise the new and evolving scarcity conditions and call for the introduction of multi-metric reliability standards. Other experts have proposed the introduction of composite-metric reliability standards that combine different resource adequacy metrics through weighting factors.

This chapter compares these two approaches using case studies obtained from a simulation model. The latter is based on stochastic expansion planning, which is run i) without any reliability standard, ii) with single reliability standards, EENS and $CVaR_{5\%}$(ENS), separately, iii) with a multi-metric standard that imposes both constraints simultaneously, and iv) with a composite-metric standard that combines the two resource adequacy metrics in a single standard. The main findings are summarised below.

The outcomes of the case studies show the impact of the choice of resource adequacy metric(s) on the resource mix, as well as on the associated de-rating factors to be used in a capacity mechanism to guide the system towards that mix. In the case studies, the selection of a metric that focuses on the scenarios with the highest unserved energy, such as $CVaR_{5\%}$(ENS), leads to a reduction in the installed capacity and de-rating factor of those technologies that perform poorly in these scenarios (CCGTs in this model).

The case studies also show that resource adequacy metrics can be highly correlated. Strong correlations result in a very small multi-activation area. This can significantly reduce the scope of multi-metric reliability standards, as they may often result in the activation of a single standard, with the others being redundant. However, where a multi-activation area exists and is large enough for a set of reliability standards to be active at the same time, multiple de-rating factors can be calculated for each technology, reflecting the contribution of that technology to the different reliability standards imposed simultaneously. In the case studies, this was demonstrated by the calculation of two different de-rating factors corresponding to CCGTs. In the capacity mechanism, this approach should be reflected in the procurement of multiple reliability products (one for each metric included in the multi-metric standard), but this may increase the complexity of the design of the regulatory instrument (multiple capacity requirements, multiple clearing prices, etc.).

Composite-metric reliability standards allow different facets of the security of supply problem to be captured in a single reliability standard by means of weighting factors. This chapter proposes a normalised formulation for these composite-metric standards, which allows resource adequacy metrics of different orders of magnitude, or even different units of measurement, to be combined. In the case study, this formulation was effective in combining the EENS and $CVaR_{5\%}$(ENS) constraints, with the solution, both in terms of installed capacity and de-rating factors, moving between these two extremes for different values of the weighting factor. The advantage of a composite-metric approach is that it allows a single de-rating factor to be calculated for each technology, which could simplify the design of the capacity mechanism.

5. EFFECTS OF THE EX-ANTE DEFINITION OF FIRM SUPPLY

5.1. Introduction[35]

In their goal to reach a desired level of resource adequacy, CRMs become the entry point of new investments when implemented in power systems. Each of the design choices of a CRM will impact investment decisions by market agents and, therefore, affect the future electricity mix's evolution. For example, penalizations for underperformance by market agents during scarcity conditions will lead to a prioritization of resources which are more reliable and, possibly, a change in maintenance and dispatch decisions by resources (Mastropietro et al., 2016). Alternatively, capacity price cap levels will determine if CRMs can ensure an adequate expansion of the system, or if there will be an underinvestment scenario if this price cap is too low (Creti and Fabra, 2007). Out of the different design elements of CRMs that might affect investment decisions, the recognition of the firm supply of resources, i.e., the amount of the reliability product they can trade in the CRM, is one of the most relevant, as already mentioned in chapter 3.

Among other characteristics, firm supply should be calculated with a forward-looking approach, estimating how the different resources will contribute to solving the scarcity conditions within the context of the future electricity mix, as argued in chapter 3. This approach, which is gradually being implemented in an increasing number of CRMs (I-SEM, 2018; Elia, 2019a; National Grid, 2019b), presents the hurdle of having to estimate the composition of the future electricity mix, hereby anticipating possible technology advancements, changes in market conditions and regulatory adaptations.

Among the many aspects that CRMs need to anticipate in order to correctly determine the firm supply of resources is the impact of the CRM itself as a driver of future investment. More specifically, the determination of the firm supply that resources might trade in the CRM will likely affect the composition of the future electricity mix. This can create a sort of catch-22 situation in which the regulator, when forecasting the operation of different resources in order

[35] This chapter heavily draws from Brito-Pereira, P., Rodilla, P., Mastropietro, P., Batlle, C., 2022a, Self-fulfilling or self-destroying prophecy? The relevance of de-rating factors in modern capacity mechanisms. Applied Energy, Volume 314, Article 118939.

to calculate firm supply, is inevitably influencing the expansion of the system and, therefore, also the future reliability performance of each resource. On the one hand, we may have a self-fulfilling prophecy[36], i.e., a situation in which the firm supply that each generator can offer determined by the regulator drives the system right in the direction outlined by the forecasts on which their calculation is based and allowing it to comply with the reliability standard originally defined. On the other hand, there is the risk of a self-destroying prophecy, where the firm supply calculation process drives the system right in the opposite direction than the forecasts behind their computation, therefore shifting the resource mix towards inefficient solutions.

This chapter focuses on addressing how the ex-ante determination of firm supply, which corresponds to both the regulator's estimation of resources contributions to scarcity conditions and how much reliability product can be traded by resources in a CRM, affects the composition of the electricity mix and the actual ex-post firm supply provided by resources, as the real contribution to solving scarcity conditions.

In particular, this exercise is carried out for solar PV by studying how the de-rating factor that is defined ex-ante, i.e., before the capacity market is cleared, conditions the results of the capacity auction and, consequently, what is the real contribution of solar PV to reliability and whether there is a risk of producing a mismatch between the expectations and the outcomes. This technology has been selected because the effect is more evident for solar PV, but the same findings can be applied to other technologies, e.g., to wind power.

A two-step model is used to simulate the energy and the capacity market. The de-rating factors of the potential new entrants in the mix are considered as exogenous variables defined by the regulator, and the outcome of the model is studied for a range of different values that these parameters can assume. This model is described in detail in section 5.2, while section 5.3 presents the results of the simulation model and discusses them. Section 5.4 concludes and provides the main policy implications of this study.

5.2. Materials and methods

The analysis presented in this chapter is based on a two-step model that replicates the participants' behaviour in a capacity auction, in which the bids are based on the results of a simulated future short-term market.

The simulation model is based on the one presented in Mastropietro et al. (2016), which has been adapted to illustrate the influence of de-rating factors on the results of a capacity auction. The model mimics the market agents' auction bids building process: they estimate what their

[36] In this chapter, the expression self-fulfilling prophecy is not used with a negative meaning, but rather it refers to a situation in which a prediction causes itself to become true. Of course, this sociological/psychological notion is not used in this chapter in a strict sense.

future income in the short-term market will be, and then, on that basis, they evaluate the income they need to get from the capacity remuneration to make the investment decision sufficiently profitable. The ultimate objective is to illustrate how the ex-ante allocation of the firm capacities heavily conditions the outcomes and how the latter might not necessarily match the expectations.

A direct-search approach is applied by means of a two-step model that seeks to attain the least-cost capacity market result, in which agents are able to perfectly anticipate the future mix (and, therefore, the result of the auction):

- In the first step, all potentially feasible future generation mixes are identified, and the future performance of the short-term market is simulated for each of them, evaluating income to be collected by the different resources. This step consists of a centralised deterministic Unit Commitment (UC) that aims to simulate a fully competitive short-term market through a minimization of electricity supply costs.
- Then, the second step consists of clearing the capacity auction. To do so, the bids for the capacity market are first calculated based on the result of the short-term market and the de-rating factors. Based on those bids, the capacity market is cleared following a pay-as-cleared mechanism. Finally, the mix resulting from the auction is compared to the initial mix used to simulate the short-term market for validation. Only those generation mixes for which the mix resulting from the auction matches the forecasted one are considered as valid.
- Finally, once all valid solutions are identified, the model selects the one that minimises the price in the auction.

The second step of the previous model is executed for several different values of the de-rating factor for the solar PV technology. The simulations allow to analyse the impact that the definition of de-rating factors has on the capacity auction results and, consequently, on the evolution of the generation mix.

The model methodology is graphically represented in Figure 5.1, while the remainder of this section describes in detail the modelling of the two different steps.

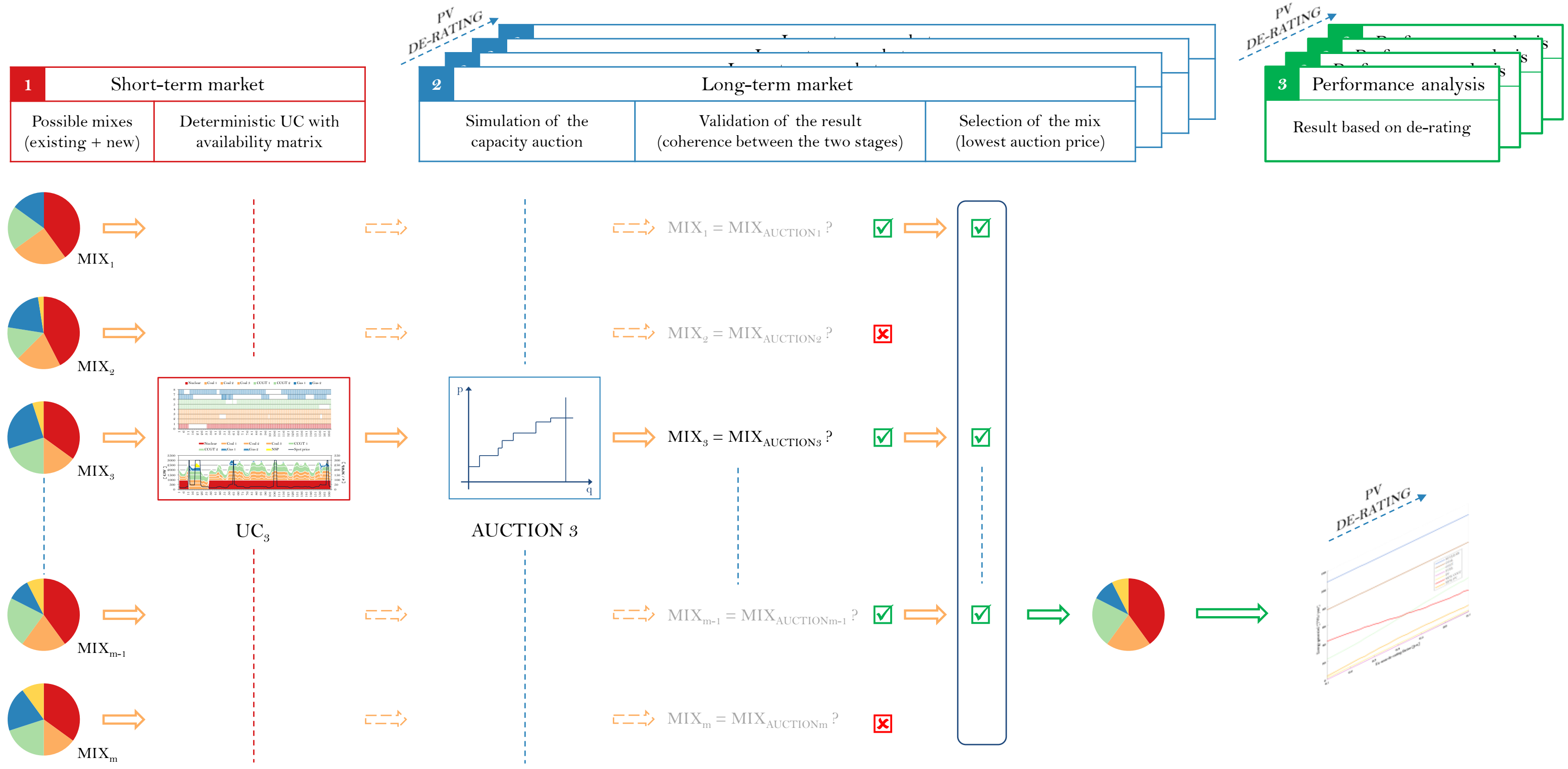

Figure 5.1: Schematic representation of the model that simulates the capacity auction for different PV de-rating factors

5.2.1. First step: the short-term market

The first step is the creation of the different scenarios of the future generation mix that feed agents' calculations to determine their bid in the capacity auction. Generation mixes are based on a predefined set of existing generation units (representing a capacity-constrained system dominated by thermal power plants), which are kept constant in all scenarios, and all plausible combinations of new entrants[37]. For the sake of simplicity, the case example only considers two potential technologies for new investments, namely Combined Cycle Gas Turbines (CCGT) and solar photovoltaic (PV) power plants.

For each scenario of the generation mix, a thousand deterministic unit commitments, each with its own distinct power plant outage pattern, are run for a time horizon of one year, as represented on the left side of Figure 5.1. This UC reproduces a centralised hourly day-ahead market with perfect competition, inelastic demand and the objective function minimises electricity supply costs. In order to maintain computation time in an acceptable range, units of the same technology have been clustered, as performed in Mastropietro et al. (2016), which was also proposed by Gollmer et al. (2000), and is solved as a Relaxed Mixed Integer Programming (RMIP) problem. The revenues of market agents are based on the hourly marginal spot price, plus non-linear side payments (Vazquez et al., 2017; Herrero et al., 2015) with daily settlements for those units that do not recover their start-up, no-load or shut-down costs.

Thermal power plants in the model are subject to outages. As in Mastropietro et al. (2016), these events are represented through a vector (one per plant) that determines the hourly availability for each thermal power plant and period in the model. These availability vectors are built considering the equivalent forced outage rate (EFOR) of each thermal plant (both existing and new entrants) and are computed by applying a two-state Markov chain with a Monte Carlo approach[38].

PV power plants are considered not to suffer any outage, and their hourly power output is purely deterministic[39], modelled through a load profile that replicates the typical

[37] Generation mixes are built considering all possible combinations of new entrants, with the number of new PV units and new CCGT units varying from zero to the maximum number of units of these technologies as expressed in the input data in Annex IV.

[38] For details on the modelling of the availability of thermal plants, please refer to Mastropietro et al. (2016)

[39] This is of course a simplification, and more precise results may be obtained through a probabilistic modelling of the availability of solar PV, which could result in a reduction of solar PV availability during scarcity conditions. This technology is characterised by a significant variability along the year; however, during the summer period, coincident with peak demand, solar output tends to be high and

meteorological year. Such hourly load profile has been obtained from the System Advisor Model (SAM) of the National Renewable Energy Laboratory (NREL) (Sengupta et al., 2018). The annualised investment costs for new CCGT and new PV generation units have been obtained from EIA (2019) and NREL (2018), respectively, considering a discount factor of 7 % and a payback period of 20 years. The cost of non-served energy, in this case, is the same as the model price cap, which is set at 3000 €/MWh, which is the price cap established in the EUPHEMIA algorithm used to clear the European regional day-ahead market. The detailed formulation of the UC model is presented hereunder.

5.2.2. Second step: the capacity auction

Once the expected incomes from the short-term market are computed for each generating unit, their bids, expressed as €/MW-year, are computed based on the methodology presented in this subsection, the auction is cleared at the marginal price, and all units are remunerated at the price offered by the last accepted bid. De-rated capacities from different technologies compete on equal terms to cover the demand for firm capacity in the auction.

The reliability product for this case study does not consider any performance incentive or penalty[40]. Therefore, bids presented by different agents only depend on the value of investment costs that are not recovered through the short-term market revenues (perfect information on the future behaviour of the market is considered). In the case of existing generators, investment costs are sunk, so their bids are set at zero price. On the other hand, the bids from new entrants can be represented by the following expression (Mastropietro et al., 2016):

$$bid_i = \text{Max}\left[0; \frac{icos_i + ocos_i - mrev_i}{capn_i \cdot drf_i}\right] \tag{1}$$

Being these the terms:
bid_i is the bid of generating unit i.

$icos_i$ is the annualised investment cost of generating unit i.

$ocos_i$ is the total operation cost of generating unit i throughout the year.

$mrev_i$ is the total short-term market revenue of generating unit i throughout the year.

$capn_i$ is the nameplate capacity of generating unit i.

stable according to the typical meteorological year data used. This should reduce the divergence between a deterministic and a probabilistic modelling of its availability.

[40] Performance incentives are a pivotal element of the design of capacity mechanisms. However, applying a penalty would not have an impact on the effect which this chapter is focusing on; therefore, it is not modelled here.

drf_i is the de-rating factor of generating unit i

The numerator represents the sum of all costs, investments plus operational costs, minus the revenues from the short-term market. The denominator represents the firm capacity of the power plant, which depends on the de-rating factor assigned by the regulator to each generation unit or each technology.

In this model, the de-rating factor used in the second step is an exogenous variable For the sake of simplicity, and in order to focus only on the effect that is being studied, the de-rating factors for thermal generators are set according to the EFOR of their technology and do not vary. In contrast, the de-rating factor for PV power plants varies between 10 % to 70 %[41]. The goal is to analyse the impact that the regulatory decision regarding the de-rating factor of solar energy may have on the outcome of the capacity auction and the resulting generation mix.

For each value of the PV de-rating factor, the auction is cleared for all the mixes initially considered, using the results of the unit commitment as an input for the bid calculation. Then, the mix resulting from the auction is compared to the initial mix and only the mixes where the expectations match the auction results are considered valid solutions from the second step (right side of Figure 5.1). This validation phase eliminates all infeasible and incoherent mixes, i.e., scenarios in which some of the new resources considered in the unit commitment are eventually not cleared in the auction and, therefore, would not be installed, leading to the bids presented not being accurate. Finally, the model selects the one with the lowest auction price among all feasible mixes. Since the model is run for several PV de-rating factors ranging from 10 % to 70 %, a different optimal mix will be identified for each PV de-rating factor and this allows the study of the impact of this pivotal element of the capacity market, as discussed in section 5.3.

5.2.3. Data for simulations

The existing installed capacity of the power system considered in the simulations is 46 GW, with a preponderance of CCGTs, nuclear and coal power plants, and a small installed capacity of renewable energy sources (namely PV power plants). The system demand is represented as a continuous profile of 8760 hours with an annual peak demand of 44.35 GW in summer, while demand is lower in winter. Three instances of daily demand profiles (low, medium and high) are shown in Figure 5.2, together with a graphical representation of the generation mix.

[41] In real capacity mechanisms, the de-rating assigned to solar PV may vary between 5% (e.g., Ireland) and 50% (e.g., MISO), as analysed in Mastropietro et al. (2019).

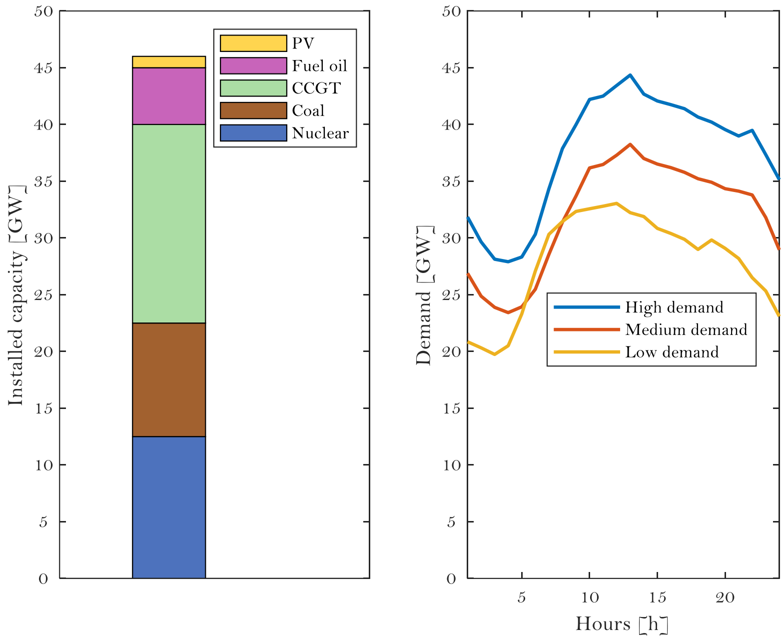

Figure 5.2: Pre-existing installed capacity (left) and daily demand representation (right)

In this case study, the demand of firm capacity in the capacity auction is another exogenous variable and it is set at 49 GW. This value may reflect the will of the regulator to improve the reliability of the system with respect to the current level, and it has been selected to leave some space for new entrants[42], who will determine the price in the capacity auction. Since the demand in the auction is an exogenous variable, the modelling exercise included a sensitivity analysis to study how this parameter affects the outcomes of the model. However, no significant impact was found, nor could any peculiar interdependence between the demand in the auction and the ex-ante PV de-rating factor be identified[43].Additional information on the model formulation and the data used can be found in Annex IV.

5.3. Results and discussion: the role of the de-rating factor

This section presents the main results that have been extracted from the simulation model. The main focus is on how the resource mix evolves for different values of the PV de-rating

[42] The demand in the auction (or, using different terminologies, the capacity requirement or the target volume) is another central element in the design of capacity mechanisms; it should be set considering the reliability target that the regulator wants to achieve. A theoretical discussion on the topic exceeds the scope of this chapter, but it can be found in Zachary et al. (2022) or in Byers et al. (2018), and in chapter 7.

[43] The sensitivity analysis was not included in the final version of the chapter, since no relevant additional finding could be extracted from it. Its goal was only to check that the findings of the chapter are not dependent on the value selected for the demand in the auction.

factor, analysing several elements, such as the new installed capacity (CCGTs and solar PV) cleared in the auction, the annual production of different technologies, the actual PV de-rating factor that is registered ex-post, the non-served energy, the price of the short-term market, and the clearing price of the capacity auction.

5.3.1. Installed capacity of new entrants

The mix under study needs new firm capacity to meet future demand reliably. Two technologies are competing in the auction, new PV power plants and new CCGTs. In this case example, when the de-rating factor that is recognised to PV units increases, the PV installed capacity in the resource mix increments to the detriment of new CCGTs. The variation of the installed capacity of new PV power plants and new CCGTs as a function of the de-rating factor is represented on the left of Figure 5.3. The graph on the right of Figure 5.3 represents the combined firm capacity of PV, new CCGT and new PV, obtained as the product of the de-rating factor and the installed capacity of each of these technologies.

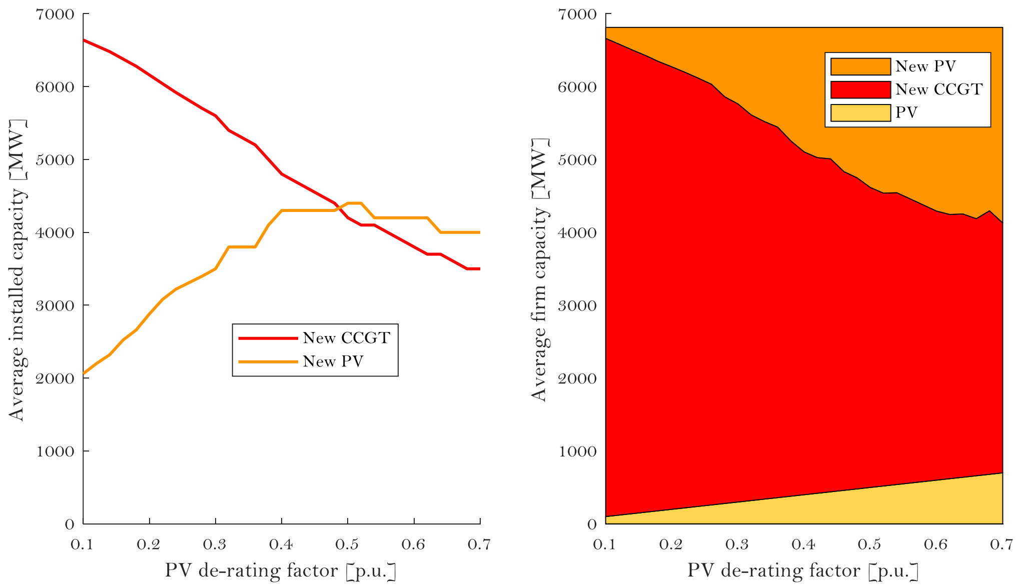

Figure 5.3: Variation of resulting installed nameplate capacity of new entrants depending on the PV de-rating factor (left) and variation of the firm capacity (right); demand for firm capacity is 49 GW, of which 6.8 GW have to come from new CCGTs and solar PV.

Despite this general positive correlation between the PV de-rating factor and its installed capacity in the optimal mix, when the de-rating factor of PV increases up to 0.5, the resulting installed capacity begins to decline. The decrease is caused by the reduction in the installed nameplate capacity of PV needed to cover the auction demand as the de-rating factor grows larger. On the other hand, there is a counterintuitive outcome for those small intervals in which the new PV installed capacity remains constant while the CCGT installed capacity decreases. In those intervals, CCGTs, whose de-rating is fixed, cover a particular share of the demand for firm capacity, while the rest must be covered through PV power plants. If the latter are recognised a higher de-rating factor, a lower CCGT installed capacity will be needed to provide the same amount of firm capacity. All these effects can be better understood by comparing the left and right graphs of Figure 5.3.

5.3.2. Annual production

The annual production of each technology is obviously influenced by the outcomes of the capacity auction. As analysed in the previous subsection, higher values of the PV de-rating factor provide a competitive advantage in the auction to new PV units, which increases their installed capacity and their yearly generation, while new CCGTs suffer the opposite effect, i.e., lower capacity cleared in the auction and, consequently, lower annual production.

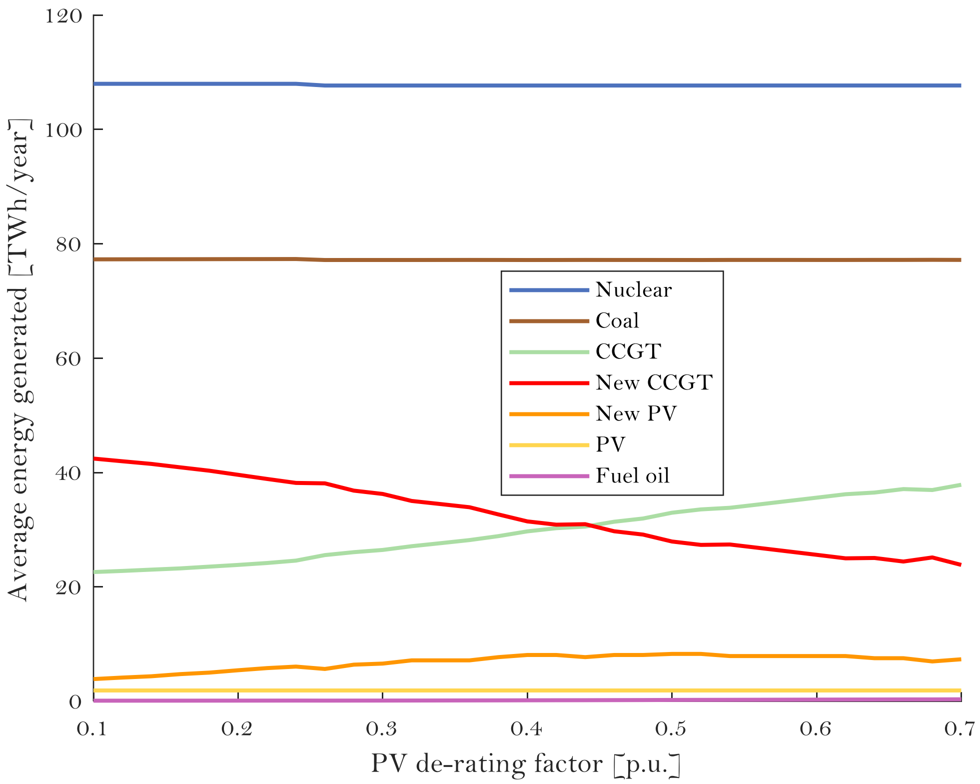

Figure 5.4: Variation of technology production depending on the PV de-rating factor

However, new PV plants are not able to produce throughout the whole day; therefore, the annual production that is provided by new CCGTs for low PV de-rating factors has to be substituted by existing technologies with higher variable costs (in this case example, existing CCGTs and fuel oil power plants). Since, in the model, existing CCGTs have lower unitary operating costs than fuel oil power plants, except for start-up and shut-down costs, the former experience most of the increase in the annual production (green line) to cover the reduced production by new CCGTs (red line).

5.3.3. Actual PV contribution to reliability and its divergence with its de-rating

As explained in chapter 3, the firm capacity of a resource, as calculated through the corresponding de-rating factor, must reflect its expected marginal contribution during scarcity conditions. Therefore, if the objective is to compare the expected contribution with the actual one, the performance of solar PV during scarcity conditions in the modelled power system must be studied.

There are several metrics that can be used to represent reliability and identify scarcity conditions (IAEA, 1984; Billinton and Allan, 1994; NERC, 2018b). As discussed in chapter 2, in the future, the growth in the elasticity of demand will not allow for the identification of scarcity conditions using only technical parameters and reliability metrics will have to internalise the price dimension in order to be resilient. In this case study, and in line with chapter 2, scarcity conditions are identified through the market price, which is used as a critical period indicator, following a basic feature of the reliability options design (Oren, 2005; Batlle et al., 2007; Elia, 2019a; Batlle et al., 2021). A price threshold equal to 300 €/MWh is set and scarcity conditions are defined as those instances when the short-term market price exceeds such threshold[44].

This reliability metric is used to compute the actual contribution of solar PV through a marginal methodology, following the principles described in chapter 3. For each value of the ex-ante PV de-rating, the corresponding generation mix is considered and its dispatch is set as the reference dispatch. Then a PV unit is removed from the generation mix[45] and a new dispatch is run. Scarcity conditions are defined as those instances when the price exceeds 300 €/MWh in the new dispatch. The marginal contribution of the PV unit is computed by measuring its production in the reference dispatch during scarcity conditions identified in the new dispatch, and dividing it by its installed capacity.

Before showing the results of this assessment, it is convenient to clarify how the operation of the system varies depending on the ex-ante PV de-rating factor. As already observed in this case example, low values of the PV de-rating factor lead to low capacity additions of PV. For low penetrations of solar PV, the peak net demand (defined as total system demand minus solar PV production) still occurs in the central hours of the day, causing higher short-term

[44] It must be remarked that analogous conclusions could be extracted from the case study presented in this chapter if the contribution to the system reliability were assessed through a different metric, for instance, by identifying scarcity conditions as those hours with non-served energy and assessing the contribution of each generating unit in those hours (Figure 5.8 shows how the two metrics present the same behaviour for growing PV de-rating factors).

[45] This PV unit represents here the marginal increment of the methodology.

market prices in these hours (Figure 5.5). In this case, PV power plants provide a valuable contribution to reliability, since they produce when the system is tight.

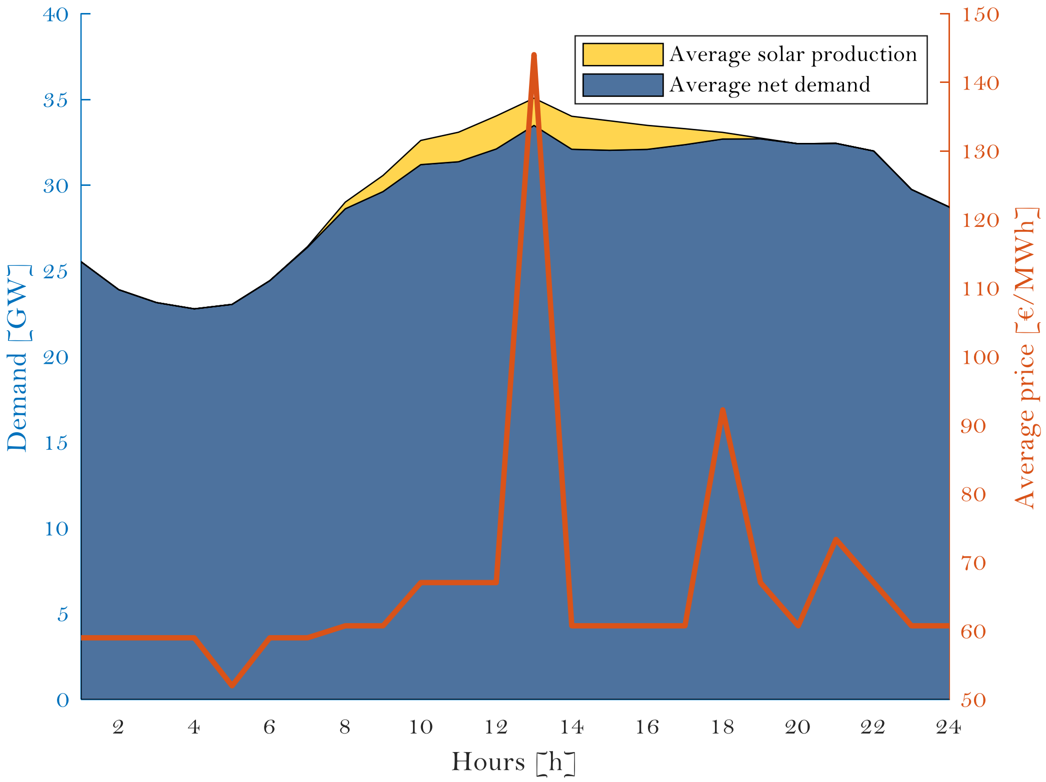

Figure 5.5: Total demand, net demand, and short-term market price for a PV de-rating factor of 0.1

On the other hand, higher PV de-rating factors increase the new PV capacity cleared in the auction. A higher PV installed capacity provokes a shift in the peak net demand towards the evening, when solar generation declines and thermal power plants are called to ramp up, thus resulting in higher short-term prices and higher risk of scarcity in a time period when solar PV cannot produce (Figure 5.6). In these conditions, the contribution to reliability from PV power plants is reduced.

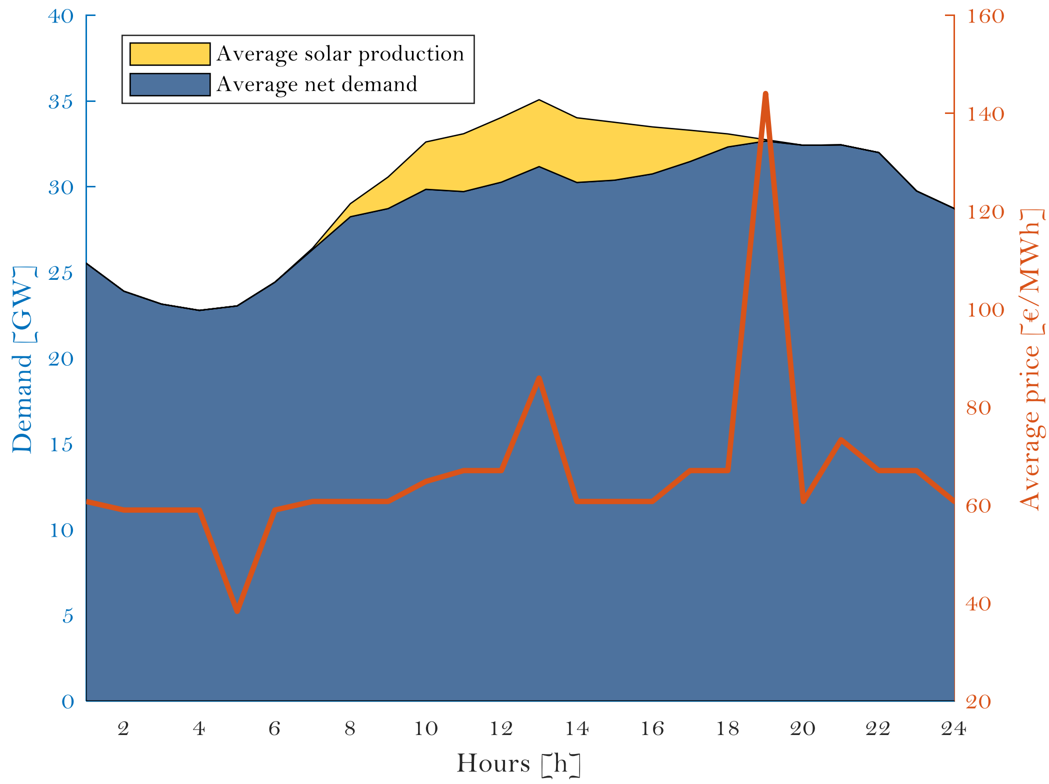

Figure 5.6: Total demand, net demand, and short-term market price for a PV de-rating factor of 0.7

These changes in the operation of the system, in terms of peak net demand and short-term price, affect the actual contribution to reliability from PV power plants. For higher ex-ante PV de-rating factors, these plants will not be able to produce when the system is tight and the short-term price is abnormally high. Therefore, their actual marginal contribution to reliability will be lower, as shown in Figure 5.7.

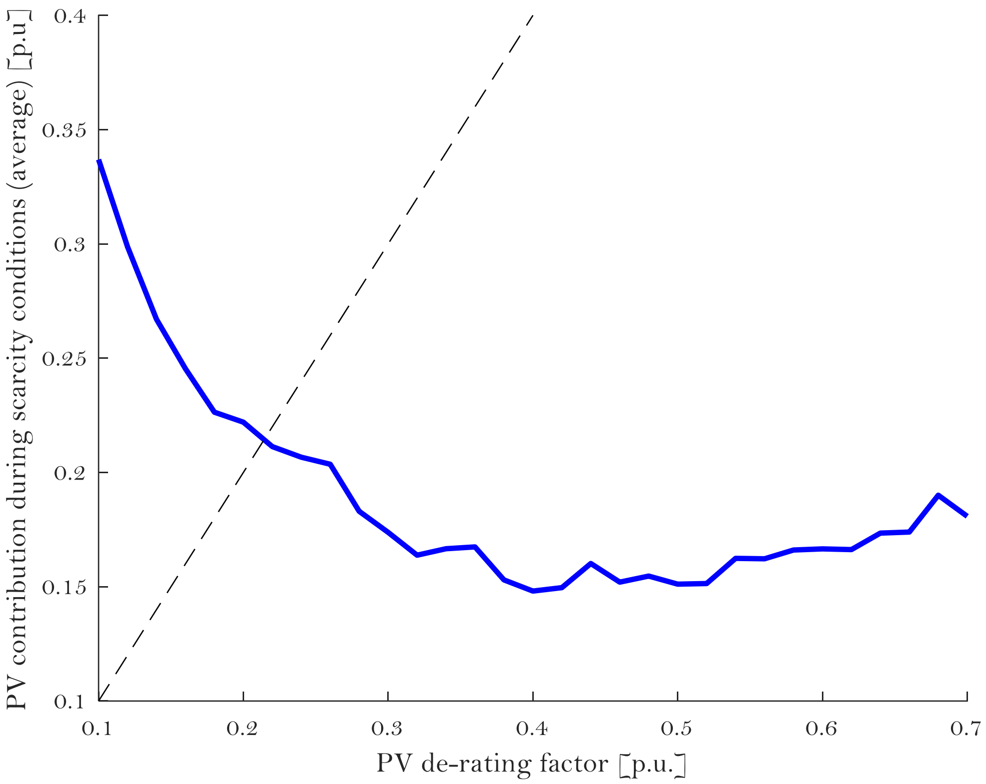

Figure 5.7: Variation of the actual marginal contribution to reliability from solar PV depending on its de-rating factor[46]

Figure 5.7 presents a key outcome of this discussion. On the top-left of this line, the actual contribution from PV is higher than the one recognised ex-ante through the de-rating factor assigned by the regulator. Thus, the technology ends up providing a larger contribution than the one it is being remunerated for, based on the original allocation of firm capacity. In this case, an excessive amount of firm capacity from new CCGTs will be procured in the auction, increasing the cost of the capacity mechanism.

On the bottom-right of the chart, the actual contribution from PV ends up being lower than the one recognised in the de-rating phase, meaning that the PV technology does not provide the reliability contribution it was remunerated for in the auction. This result also implies that the demand for reliability cannot be expected to be covered, leading to undesired levels of non-served energy (subsection 5.3.4) and short-term market prices (subsection 5.3.5).

The discrepancy between the assumptions behind the model used to compute de-rating factors (or the demand for firm capacity in the auction) and the actual performance of the resource mix resulting from the auction itself is one of the main challenges in the design of capacity mechanisms. The ideal solution would be to clear capacity markets through an iterative

[46] The dotted line in Figure 5.7 represents points in which the actual contribution to reliability from PV matches the PV de-rating factor defined by the regulator.

process that allows the validation of the outcome of the auction only if such discrepancy is below a certain tolerance, as also proposed in Zachary et al. (2022). However, no capacity mechanism implemented to date is based on an iterative process of this type[47]. Therefore, the model presented in this chapter is based on a conventional clearing of the capacity auction but highlights the importance of carefully defining the assumption on which de-rating factors are computed.

5.3.4. Reliability and scarcity conditions

The PV de-rating factor also influences the overall reliability of the system. As mentioned in the previous subsection, if the de-rating factor of PV power plants is larger than the real reliability contribution that these units provide, the demand for reliability is not satisfied, since many new CCGTs were pushed out of the auction by new PV power plants. This effect can be observed through different reliability metrics. The chart on the left in Figure 5.8 shows how the non-served energy registered in the system increases for growing values of the PV de-rating factor. A similar increase can be observed in the number of hours with high short-term market prices, which is the reliability metric used to assess the actual contribution to reliability by PV units, as presented in the right chart in Figure 5.8.

[47] However, some regulators are recognizing the problem explicitly. In Ireland, de-rating factors are computed for different generation mixes and then averaging the de-rating factor obtained in each of them for a given technology (EIRGRID, 2016).

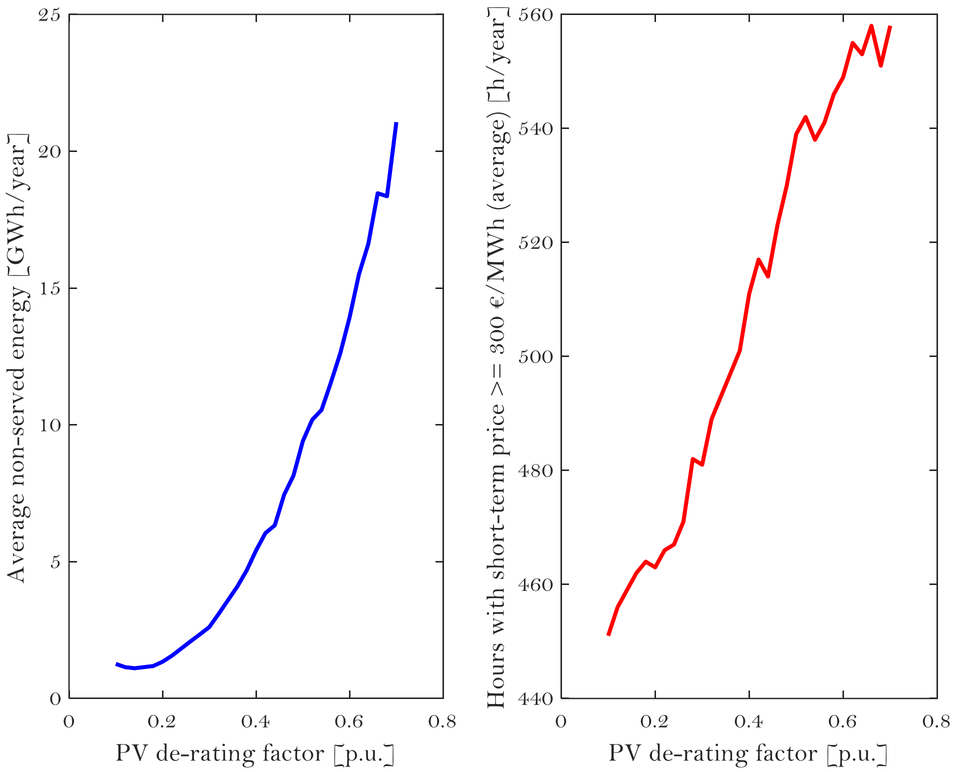

Figure 5.8: Variation of the non-served energy and the number of hours with high short-term prices

5.3.5. Short-term prices in the long-term and capacity market prices

5.3.5.1. Short-term market price

The short-term market price[48] evolves according to the generation mix and the annual production of different technologies. For higher PV de-rating factor values, the yearly production of new CCGTs decreases in favour of more costly existing thermal plants, which leads to an increase in the average short-term market price. This effect is aggravated by the occurrence of scarcity events with non-served energy, during which the short-term price reaches the administrative price cap (see section 5.2.1).

[48] The short-term market price is obtained as the dual variable associated to the generation-demand balance constraint of the optimization problem, see section 5.2.1 for details.

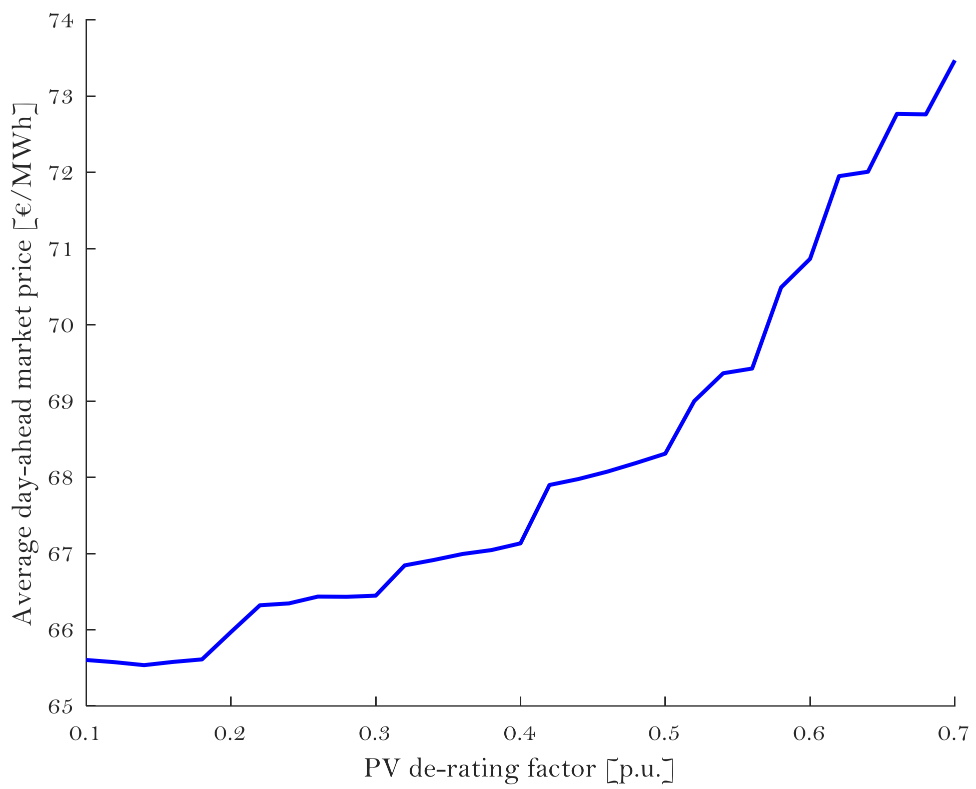

Figure 5.9: Variation of the average short-term market price depending on the PV de-rating factor

5.3.5.2. Clearing price of the capacity auction

As explained in subsection 5.2.2, the bids from generation units in the capacity auction depend on the de-rating factor granted to them and their market revenues in the short-term market. As the de-rating factor of solar PV increases, the bids submitted by these power plants decrease, following the equation introduced in section 5.2.2. Additionally, the growth in the short-term market price (Figure 5.9) increases the market revenues for all technologies, thus provoking an overall reduction also in the bids presented by new CCGT units. The variation in the yearly non-recovered investment costs (the numerator of the bid calculation formula presented in section 5.2.2) for solar PV and the bids presented by these new power plants can be observed for two values of the PV de-rating factor in Table 5.1.

Table 5.1: Comparison between the yearly non-recovered investment costs and the bids by new PV power plants for a de-rating factor of 0.2 and 0.6

	PV de-rating factor [p.u.]	
	0.2	**0.6**
Non-recovered investment costs [k€/MW]	5.446	4.552
Resulting Bids [k€/MW]	27.232	7.587

The decrease in new PV and new CCGTs bids leads to a reduction in the capacity auction clearing price as the PV de-rating factor rises, which can be observed in Figure 5.10.

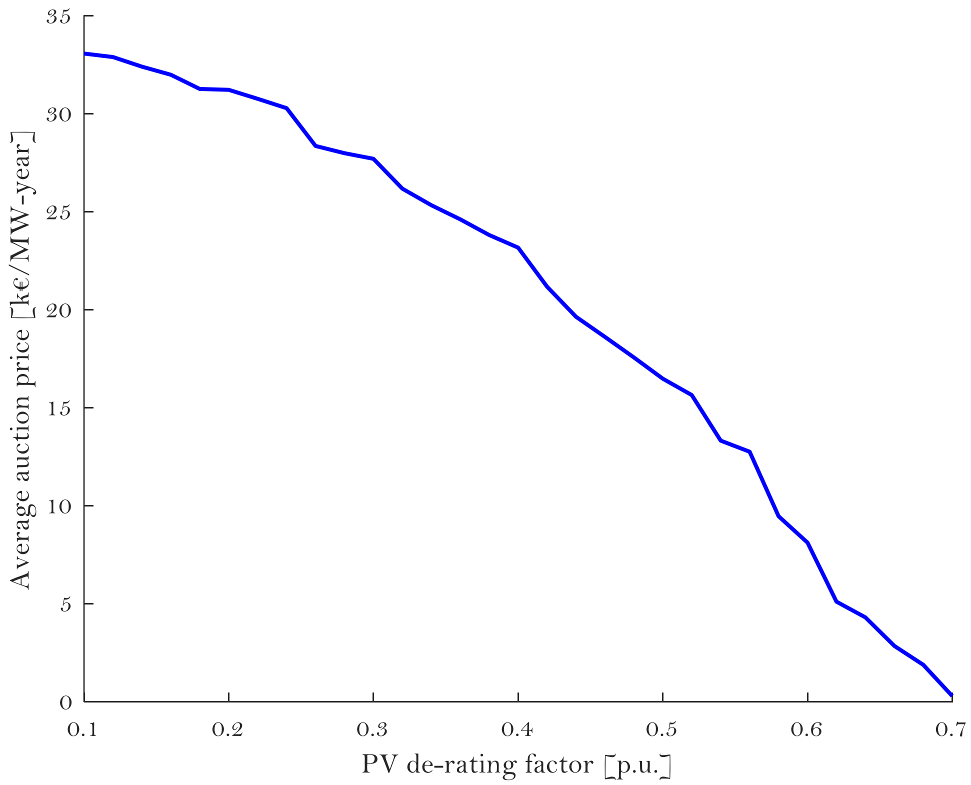

Figure 5.10: Variation of the capacity auction price depending on the PV de-rating factor

A low clearing price in the capacity auction (due to a generous/optimistic allocation of firm capacity credits) may look like a positive effect since it decreases the overall cost of the capacity mechanism. However, it must be remarked that this might come at the expense of an increase of the non-served energy (Figure 5.8) and the short-term market price (Figure 5.9). A capacity mechanism cannot be isolated from the rest of the market design. Even a small change in an individual parameter, such as the de-rating factor of PV units, may significantly impact the outcome of the capacity auction and the performance of the system as a whole.

5.4. Conclusions and policy implications

Capacity mechanisms have become a pillar of the electricity market design while power sectors undergo the energy transition. When they are in place, these mechanisms, together with auctions for renewables, become the main entry point for new resources and their outcome drives the evolution of the entire mix (Battle et al., 2021). As presented in this chapter, a pivotal feature in the design of these instruments is the calculation of the firm capacity that each resource or technology can trade in the capacity market, commonly obtained through the application of de-rating factors.

In this chapter, the impact of the ex-ante definition of de-rating factors on the resource mix emerging from the capacity market has been quantitatively studied and presented. This analysis is based on a two-step model that simulates the capacity market. The case study presented in this chapter focuses on a power system dominated by thermal generation, with a summer peak demand, and with only two potential new entrants, i.e., CCGT and solar PV. The de-rating factor assigned to solar PV generation strongly influences the outcome of the capacity market. The larger the PV de-rating factor, the greater the competitive advantage of this technology with respect to CCGTs and the greater the capacity of PV power plants cleared in the auction. The outcome of the auction affects the resource mix and, consequently, the operation and the performance of the system. A larger PV de-rating factor increases solar penetration and this may provoke a shift in the scarcity conditions that the system has to face. However, this also affects the actual contribution of solar PV to reliability. Conversely, an insufficient allocation of firm capacity for solar PV leads to unnecessary extra costs in the capacity auction. These effects, together with the main findings of the chapter, are represented graphically in Figure 5.11.

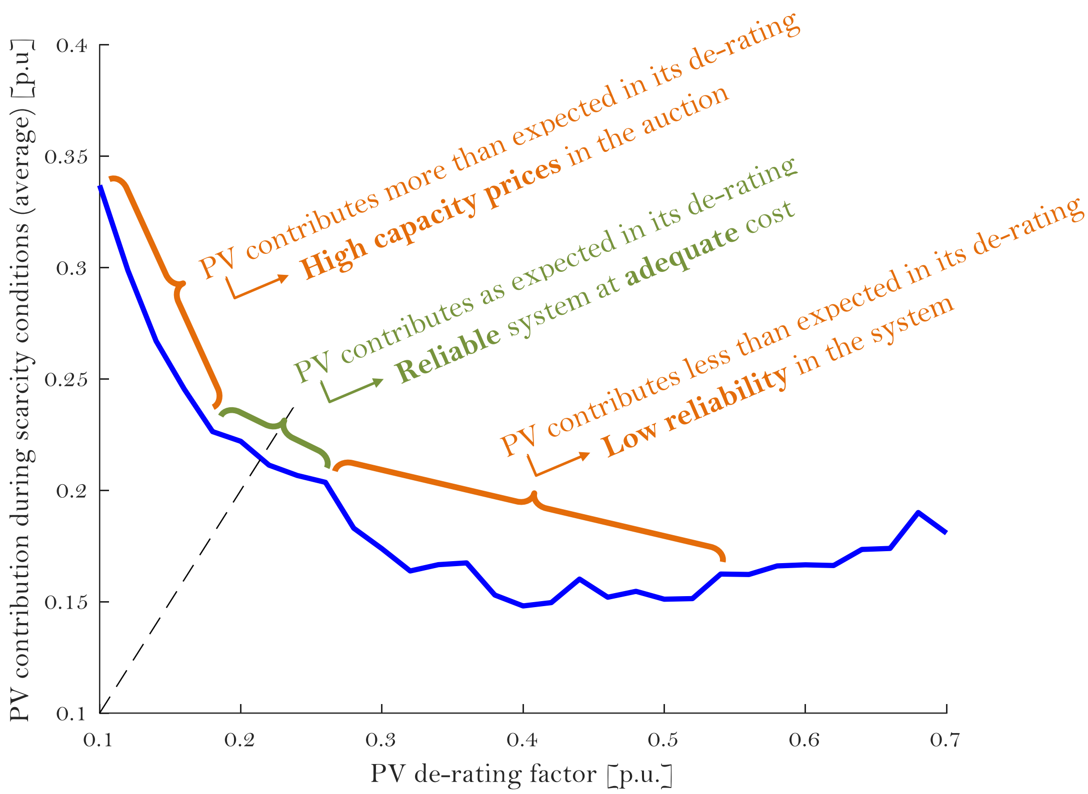

Figure 5.11: Graphical representation of the main findings of this chapter

Although the simulation model presented in this chapter shows these effects for solar PV, its findings also apply to other technologies, especially intermittent resources, whose availability is more likely to present some correlation with demand. In a more general sense, the same effect could occur with any technology whose marginal contribution to scarcity conditions declines with increasing capacity of said technology, like the examples presented in section 3.2.3. With current and upcoming resource mixes, the regulator has a very tight margin to guarantee efficiency when defining a de-rating factor. This decision influences the resource mix and the real contribution that each resource can provide to the adequacy of the system. If the de-rating factor is too low, it will undercompensate the technology for its adequacy contribution and may potentially increase the overall cost of the capacity mechanism. If the de-rating factor is too high, it will distort the outcome of the auction and result in a resource mix that is not able to achieve the reliability standard, causing non-served energy and higher short-term market prices.

Capacity mechanisms represent a regulatory intervention that aims at impacting the resource mix to come in order to guarantee the adequacy of the system. However, this discussion reflects the significant complexities that these mechanisms entail, and thus it calls for a careful design process, in which the actual impact of the different design elements, in particular the estimation of firm supply, needs to be cautiously assessed.

6. CRM COST ALLOCATION

6.1. Introduction[49]

In the last decade, CRMs have become a mainstay regulatory instrument in modern power systems, especially in the European Union. As these mechanisms become increasingly widespread, their global associated costs grow, with the consequent burden for electricity consumers. In the European Union, for example, the surge of these mechanisms has led to an increase in their weight in the total cost of electricity supply (ACER, 2024b; ACER, 2025), as shown in Figure 6.1.

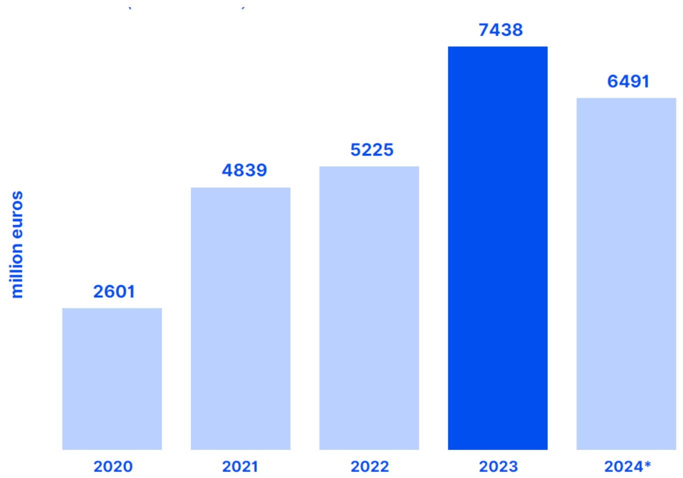

Figure 6.1: Incurred and projected costs of CRMs in the EU (ACER, 2024b)

Therefore, it is crucial to allocate these growing CRM costs efficiently to consumers and send them economic signals that discourage consumption patterns that force otherwise unnecessary investment in new firm capacity. However, in many electricity systems, cost allocation methodologies for CRMs are often simplistic and do not focus on the true cost drivers of these mechanisms. In Ireland, for example, CRM costs are recovered through volumetric charges applied to consumption between 07:00 and 23:00 each day (SEMC, 2024). The economic signal given by these charges is that consumption between these hours is

[49] This chapter heavily draws from Brito-Pereira, P., Rodilla, P., Mastropietro, P. Efficient cost allocation in capacity remuneration mechanisms: applying the causer-pays principle to resource adequacy. Working paper IIT-24-052. Under review at International Journal of Electrical Power and Energy Systems.

equally detrimental to the reliability of the system and that consumption outside this period has no effect on it. In Belgium, CRM costs are recovered through a series of volumetric consumption charges that vary according to the category of consumer. The result is that some consumers face CRM charges almost 100 times higher than other consumers, although their impact on the reliability of the system may be similar (BOG, 2021). Italy seeks to recover 70 % of its CRM costs through charges applied to demand during the 500 hours of the year when the reliability risk is expected to be highest. The remaining 30 % is recovered through lower charges applied to consumption during the remaining hours of the year (Terna, 2021; ARERA, 2024). In Spain, where the implementation of a CRM is underway (MITERD, 2024), the cost of the CRM is expected to be recovered through volumetric charges that vary according to the tariff segment (determined by the connection voltage and the contracted power) and the tariff period (peak, flat, and valley) defined in the Spanish regulation for network tariffs. The CRM costs are allocated to each tariff period according to the risk of suffering a shortfall of firm capacity, using an index calculated as the ratio of the contracted firm capacity and the whole-system demand.

Despite the increasing importance of CRM costs and the inefficiencies that can result from the application of simplistic CRM cost allocation methods, only a few articles have been published in the academic literature on the subject. For example, Wang et al. (2023) argue that the majority of CRM costs should be allocated to consumption during peak demand periods. At the same time, there has been considerable research on how to calculate the contribution of demand response to security of supply (Nolan et al., 2014; Zhou et al., 2015, 2016; Nolan et al., 2017; Zeng et al., 2020; Freire-Barceló et al., 2022) or how to properly integrate demand response resources into CRMs (Lynch et al., 2019; Lambin, 2020). However, calculating the firm supply of DR resources and allocating CRM costs to consumers are two sides of the same problem, and the methodologies behind them should be coordinated.

In order to fill this void in the literature research, this chapter proposes an easy-to-implement methodology that allocates CRM costs according to consumers' marginal negative impact on system adequacy. The proposed cost allocation formulae allow the calculation of a set of hourly factors that reflect the marginal impact of a load variation in each hour on the reliability standard. This methodology is then quantitatively evaluated using a convolution-based probabilistic model in two case studies, a stylised example to illustrate the results of the calculation and a full-scale case study inspired by the Spanish power system.

The chapter is structured as follows. Section 6.2 describes in detail the proposed theoretical cost allocation methodology and shows how it can be applied to calculate hourly charges for consumers. Section 6.3 details the materials and methods used to evaluate the application of the methodology. Section 6.4 presents the case studies used to quantitatively assess the methodology and discusses the results. Section 6.5 draws the main conclusions and policy recommendations from the results of the case studies and the theoretical discussion presented in this chapter.

6.2. Efficient CRM cost allocation

Capacity remuneration mechanisms seek to address adequacy concerns in power sectors by providing incentives for resources to help the system meet a certain reliability standard. These incentives encourage market participants to implement both new investments and operational strategies that are consistent with meeting that reliability standard.

As presented in Annex I, the remuneration that each resource receives from the CRM should ideally be based on its marginal contribution to meeting the reliability standard. More specifically, the remuneration should be proportional to the following expression[50]:

$$\frac{\partial \mathrm{RM}}{\partial \mathrm{K_i}} \mathrm{K_i} \tag{1}$$

Where:

- The sub-index *i* refers to the generation asset under consideration.
- K_i is the installed capacity of resource *i* (expressed in MW)
- *RM* is the resource adequacy metric used by the regulator (or system operator) to set the reliability standard.

The marginal contribution of each resource depends on the characteristics of the system (demand, existing generation, storage technologies, etc.), the characteristics of the resource (availability, production profile, etc.), and the chosen reliability standard. If, for the same system, a different metric (e.g., LOLP instead of EENS) is used to define the reliability standard, the marginal contribution will most likely change and therefore the remuneration of different resources would also change.

Just as resources providing adequacy services should be remunerated in proportion to their marginal contribution to the reliability standard, consumers should ideally be allocated CRM costs in proportion to their negative marginal contribution (causer-pays principle). Annex V presents an alternative approach to the one described in chapter 3 which is more suitable in the context of demand response and cost allocation methodologies for CRMs. This formulation replaces the installed capacity variable (the key variable on the generation side) with hourly consumption (the key variable on the demand side). With this new formulation, we derive the expression providing the cost to be allocated to any consumer i:

$$\sum_{\mathrm{h}} \frac{\partial \mathrm{RM}}{\partial \mathrm{D^h}} \mathrm{D_i^h} \tag{2}$$

[50] This expression is also proportional to the firm supply that is recognised to each resource in the mechanism.

Where:

- The super-index h refers to hourly values.
- D_i^h is the hourly, h, demand of consumer i (expressed in MWh)

We refer to the term $\frac{\partial RM}{\partial D^h}$ as the "hourly marginal reliability factor" for hour h, which reflects the impact of a marginal increase in demand in that hour on the system reliability metric. For example, the marginal reliability factor of an hour in which the system is under stress may be twice that of a valley hour. This means that consumption in the first hour should be allocated twice the share of CRM costs as consumption in the second hour.

If each consumer's demand profile is multiplied by the hourly marginal reliability factors and then all the hourly values obtained are summed, the result is what we refer to in this chapter as the "cost allocation weight" for consumer i (equation 2). These weights indicate how much a consumer should pay relative to other consumers and can be expressed as percentages ("cost allocation shares") to be directly multiplied by the total cost of the CRM.

6.3. Materials and methods

The theoretical approach presented in section 6.2 is applied to a convolution-based probabilistic production cost (PPC) model, which allows us to demonstrate the advantages of the proposed methodology in a rather illustrative and straightforward way.

PPC models are well-known classic tools that have traditionally been used to assess power system reliability (Sutanto et al., 1989). These models attempt to calculate the loss of load and the non-served energy in the system through probabilistic functions of the electricity demand and the energy output from generators. To account for all possible combinations of power plant outages, these models use a convolutional process (Lu et al., 2018; Ma et al., 2023). PPC models have several strengths, including low modelling complexity and computational time, while they can accurately represent power plant outages and the stochastic nature of load. However, these models lack a time-sequential representation of system dispatch and do not incorporate some technical aspects of generators, such as ramping constraints and minimum down-times. However, for the purpose of these analyses, convolution-based PPC models can be used to illustrate and demonstrate the methodology proposed in this chapter and its application to CRM cost allocation. It should be noted, however, that the theoretical approach described above can be applied to more sophisticated simulation models.

In PPC models, the probabilistic distribution of electricity demand is usually obtained by analysing historical demand patterns over a given time period. This can be done in two steps: first, by transforming the historical hourly electricity demand into a load duration curve, where the hourly demand is ordered from higher to lower values (Figure 6.2), and second, with the subsequent transformation of the load duration curve into a complementary distribution function (Figure 6.3). The complementary distribution function will reflect the probability of the demand exceeding a certain level for any generic time interval (in this case, hours).

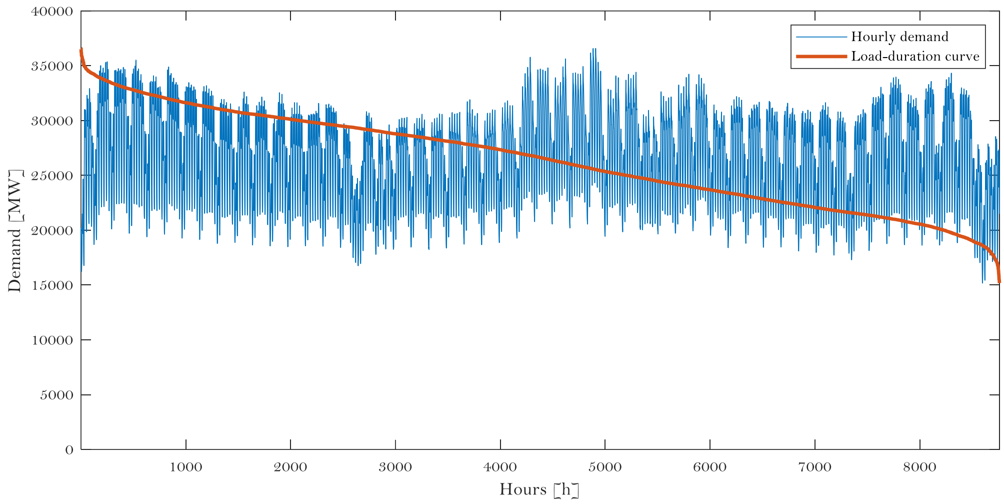

Figure 6.2: Spanish 2019 hourly electricity demand and its transformation into a load duration curve

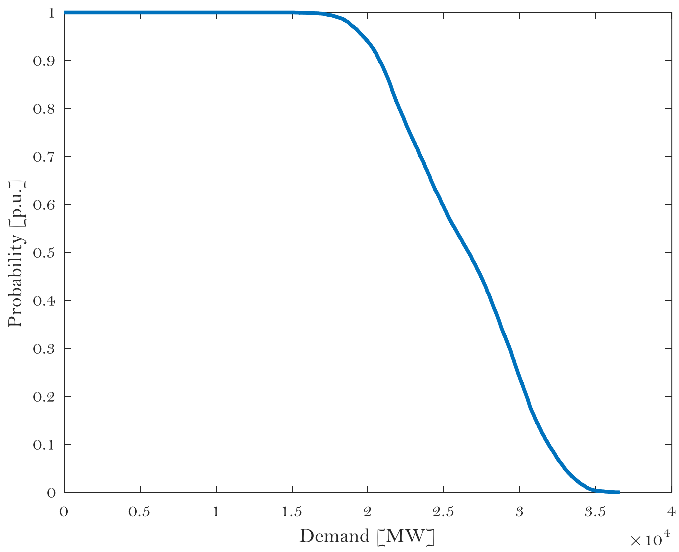

Figure 6.3: Complementary distribution function of electricity demand in the system

In most PPC models, generators are either available at full capacity or completely unavailable. In this case, each generator (i) has a probability p_i to be fully available and a probability q_i ($q_i=1-p_i$) of being unavailable. In order to consider all possible combinations of availability and unavailability of generators, a convolution process is applied to the probability distributions of the generators. The convolution process consists of calculating the equivalent demand after probabilistic dispatching of n generators. If no generator is considered (i.e., $n=0$), then the equivalent demand would be equal to the complementary distribution of electricity demand. As generation assets are progressively included in the assessment according to an economic merit order, the equivalent demand reveals the probability distribution function of the non-served energy in the system, which would progressively decrease as more generators are added to the convolution, as shown in Figure 6.4. Once all N generators are included in the convolution process, the equivalent demand represents the distribution function of the non-served energy. Therefore, the Expected Energy Non-Served, or EENS, can be obtained as the area defined by the distribution function of the non-served energy, while the Loss of Load Probability, or LOLP, can be obtained as the intercept of the same function on the y-axis.

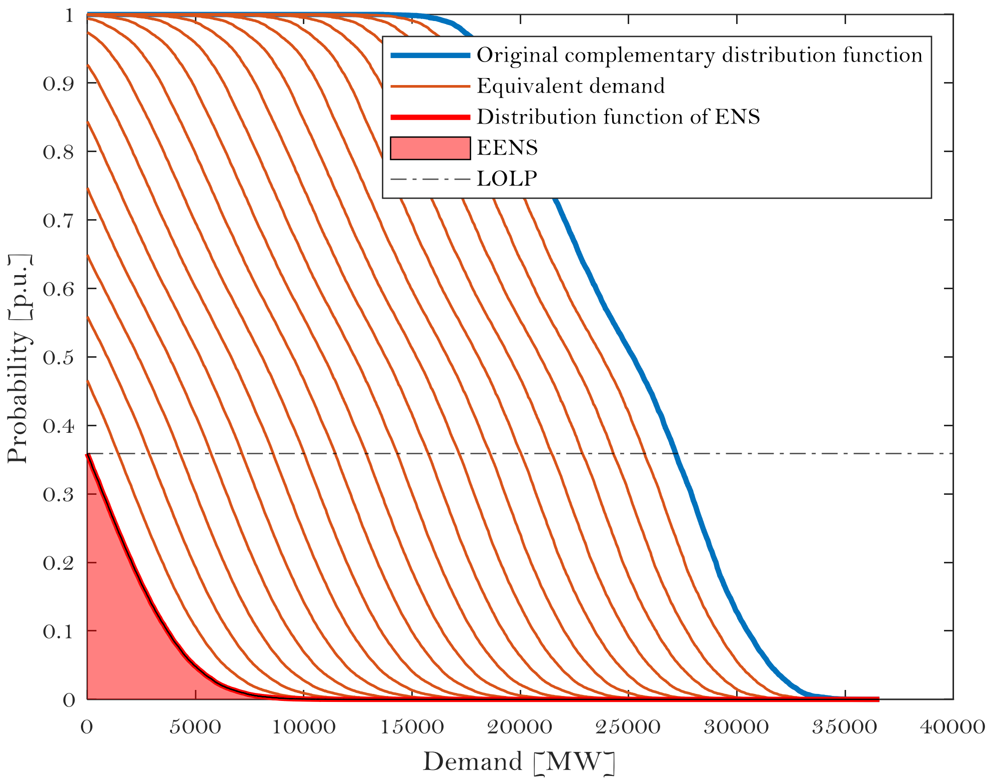

Figure 6.4: Application of the convolution process to the complementary distribution function of electricity demand

PPC models can be used to calculate the hourly marginal reliability factors as defined in section 6.2, which reflect the impact of a marginal increase in demand in a given hour on the reliability metric. The original complementary distribution function of electricity demand is modified to account for a marginal increment in hour h. The convolution process is repeated and a new value of the reliability metric is calculated. The change in the reliability metric after adding a marginal increment of demand in hour h, is the hourly marginal reliability factor for hour h. The process must be repeated for all the hours in the time horizon, in this case for 8760 hours, resulting in 8760 hourly marginal reliability factors.

6.4. Case studies: Results and discussion

The methodology described in section 6.3 is applied here to two case studies. The first is a stylised case study used to better illustrate the methodology and its outcomes. The second is a full-scale case study, inspired by the Spanish power system, which is used to discuss how CRM costs should be allocated to consumers. Both case studies are based on an EENS reliability metric.

6.4.1. Stylised case study

In the stylised case study, the methodology is applied to a 24-hour load, which is divided between six consumers with different demand profiles, as shown in Figure 6.5. The generation mix serving this demand consists of ten 1-MW CCGTs, each with an Equivalent Forced Outage Rate (EFOR) of 0.1.

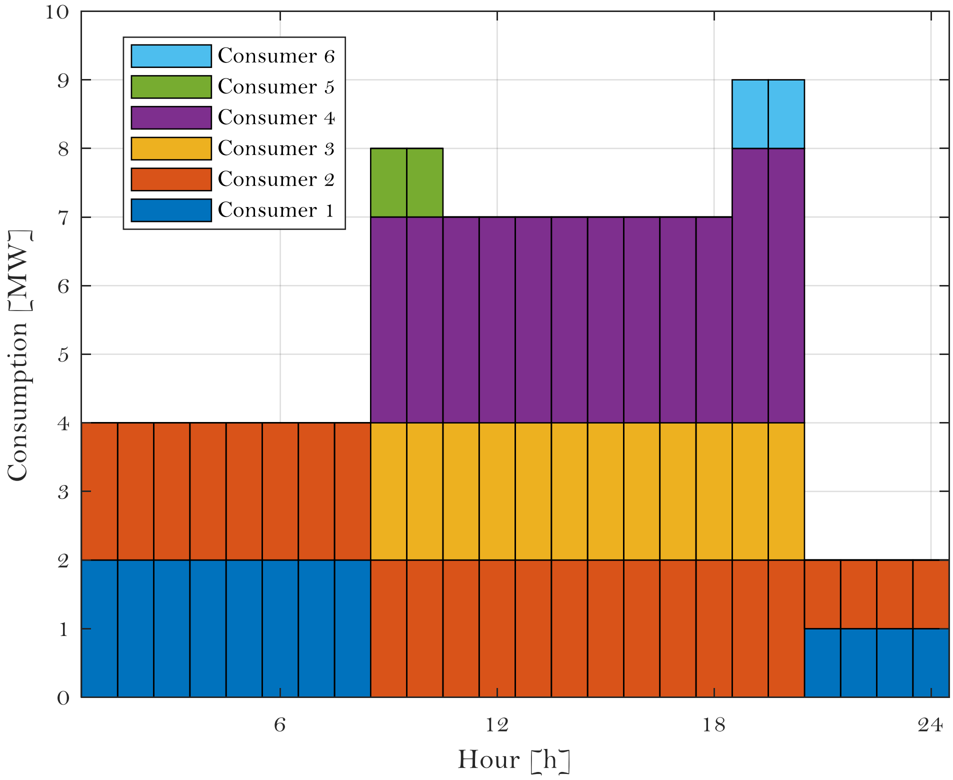

Figure 6.5: Consumer demand profiles in the stylised case study

The first step in the methodology is to calculate the hourly marginal reliability factors using the procedure presented in section 6.2. The results are shown in Figure 6.6. As we use EENS as the reliability metric, both the variation in the reliability metric and the marginal increase in the demand are expressed in MWh and the factors, which are the ratios of these two terms, are then expressed on a per-unit basis.

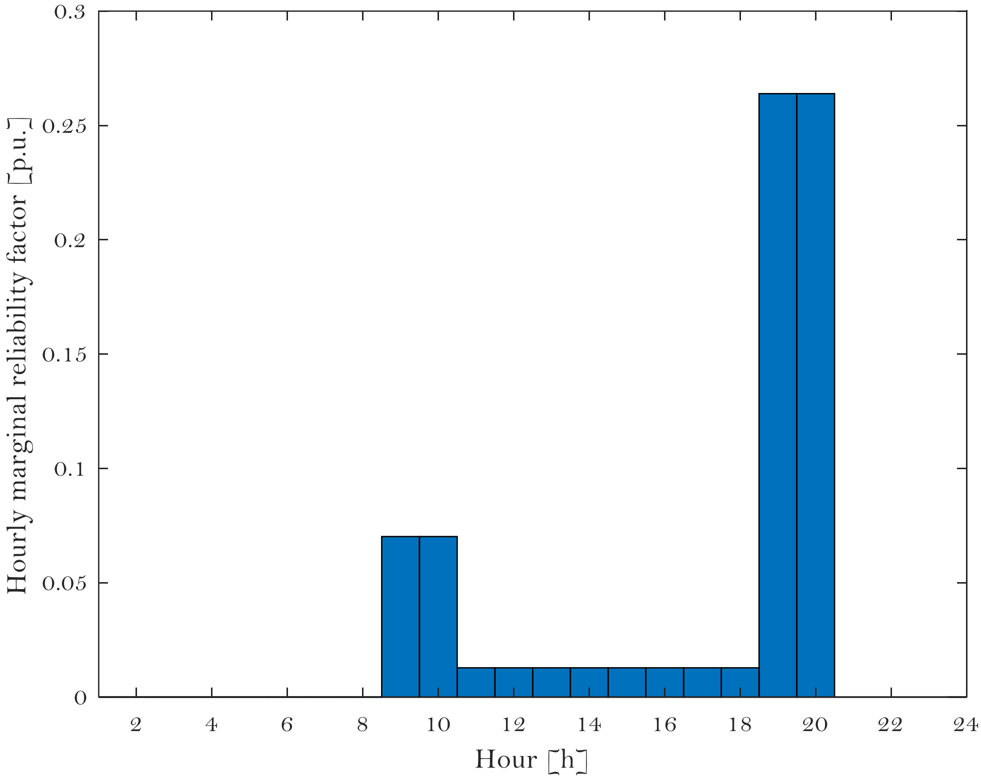

Figure 6.6: Hourly marginal reliability factors for the 24-hour time horizon of the stylised case study

The highest marginal reliability factors are recorded during periods of higher demand, such as between 9:00 and 10:00 and 19:00 and 20:00. It should be noted that the factors are not linearly proportional to electricity demand. For example, although the evening peak is only 1 MW higher than the morning peak (12.5 % higher), the marginal reliability factor for the evening peak is almost four times higher than that corresponding to the morning peak. Additionally, the marginal reliability factors for the first eight hours and the last three hours of the day are very low (1.18×10^{-5} and 1.18×10^{-8} p.u., respectively), but they are not zero. These values reflect that, although most of the CRM costs should be allocated to the hours with the highest risk of scarcity (in this case, the peak-demand hours), no hour has a zero marginal EENS or a zero marginal reliability factor, and some of the CRM costs should be allocated to the off-peak hours.

To determine the share of the CRM costs to be borne by each of the six consumers, the hourly marginal reliability factors should be multiplied by the consumption profile of each consumer. For the sake of clarity, both variables are shown together in Figure 6.7.

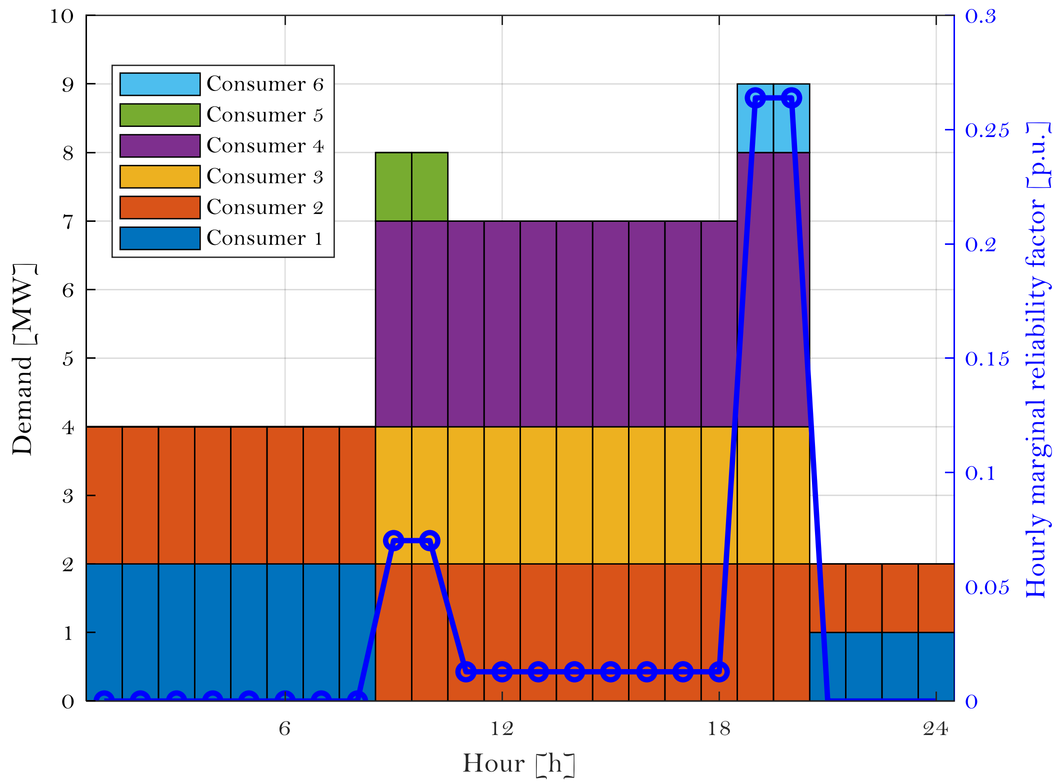

Figure 6.7: Hourly marginal reliability factors and consumption per consumer

The result of these multiplications is the cost allocation weight for each consumer profile, as presented in Table 6.1. These cost allocation weights can also be expressed as percentages, representing the proportion of CRM costs that should ideally be allocated to that consumer. In addition, if the cost allocation weights are divided by the total demand of each consumer, they can also be expressed on a unit basis, allowing for a more straightforward comparison between them.

Table 6.1: Cost allocation weights and shares for the stylised case study

Consumer	Cost allocation weights [p.u.]	Cost allocation share	Unitary cost allocation weights [p.u.]
1	1.46 x10^{-4}	0.002 %	1.75 x10^{-4}
2	1.5412	23.387 %	0.8407
3	1.5411	23.385 %	1.5411
4	2.8394	43.086 %	1.7933
5	0.1404	2.130 %	1.6846
6	0.5278	8.009 %	6.3336

The results show that consumer 6 has the highest unitary cost allocation weight because it only uses electricity during the hours with the highest risk of shortage, when the marginal reliability factor is the highest. On the contrary, consumer 1, who only consumes during off-peak hours, has the lowest unitary cost allocation weight. In terms of cost allocation share, the largest share of CRM costs should be borne by consumer 4, who has a high demand and withdraws electricity during both the morning and evening peaks. Consumer 1, who only has off-peak demand, would have to pay a small but still positive percentage of the total CRM costs.

6.4.2. Full-scale case study

The full-scale case study is based on the hourly demand of the Spanish power system in 2019, as shown above in Figure 6.2, together with the corresponding load duration curve. For the sake of simplicity, the generation mix modelled in the full-scale case study includes only conventional thermal generation units, and the installed capacity of these generators has been calibrated to achieve a normalised EENS in the system equal to 0.002 %[51]. The total installed capacity in the system to achieve this reliability standard (assessed with the PPC model) is 38 750 MW, with the technology mix[52] described in Table 6.2.

Table 6.2: Electricity mix used for the full-scale case study

Technology	Number of units	Installed capacity per unit [MW]	EFOR [p.u.]
Nuclear	15	500	0.015
Fuel Oil	10	500	0.100
CCGT	50	500	0.050
OCGT	5	250	0.200

Unlike the stylised case study, which has a 24-hour time horizon, the full-scale case study evaluates an entire year of hourly data. Using the same methodology described for the stylised case study, it is possible to calculate the marginal reliability factors for each hour of the year. This information is presented as a heatmap in Figure 6.8. As intermittent renewable resources are not included in the model, the hours with the highest marginal reliability factors (representing the highest risk of scarcity) appear in the middle hours of summer days, when demand is higher. On the other hand, the weekends (corresponding to the last two vertical sections of the heatmap) and the early hours of the weekdays have the lowest marginal reliability factors and, therefore, the lowest adequacy risk for the system.

---

[51] This is the reliability standard used in AEMO (AEMO, 2019), which is expressed as normalised EENS and can be easily transposed from one system to another. The normalised EENS is obtained by dividing the EENS by the total demand in the system in the same time horizon.

[52] Many generation mixes could meet the same target, but this would not significantly change the results of the simulation or the key messages of this paper.

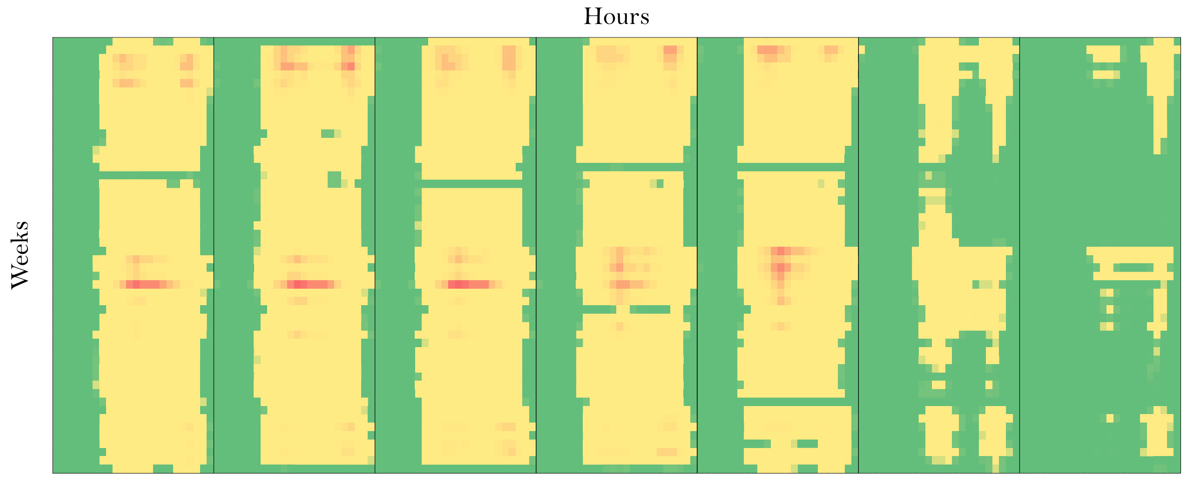

Figure 6.8: Heatmap of the hourly marginal reliability factors for the full-scale case study

Following the same reasoning as in the stylised case study, the hourly marginal reliability factors can be used to determine the cost allocation share of any load profile. For this case study, three different demand profiles are evaluated, reflecting very different consumption patterns, as shown in Figure 6.9. Consumer 1 only withdraws electricity in the early hours of the day, when the reliability risk is lower, consumer 2 only consumes during the peak hours of the day, and consumer 3 has a more balanced load, which may resemble that of a residential consumer. Each of these consumer profiles is repeated over the 365 days of the year and, for ease of comparison, has an identical annual demand of 8760 MWh.

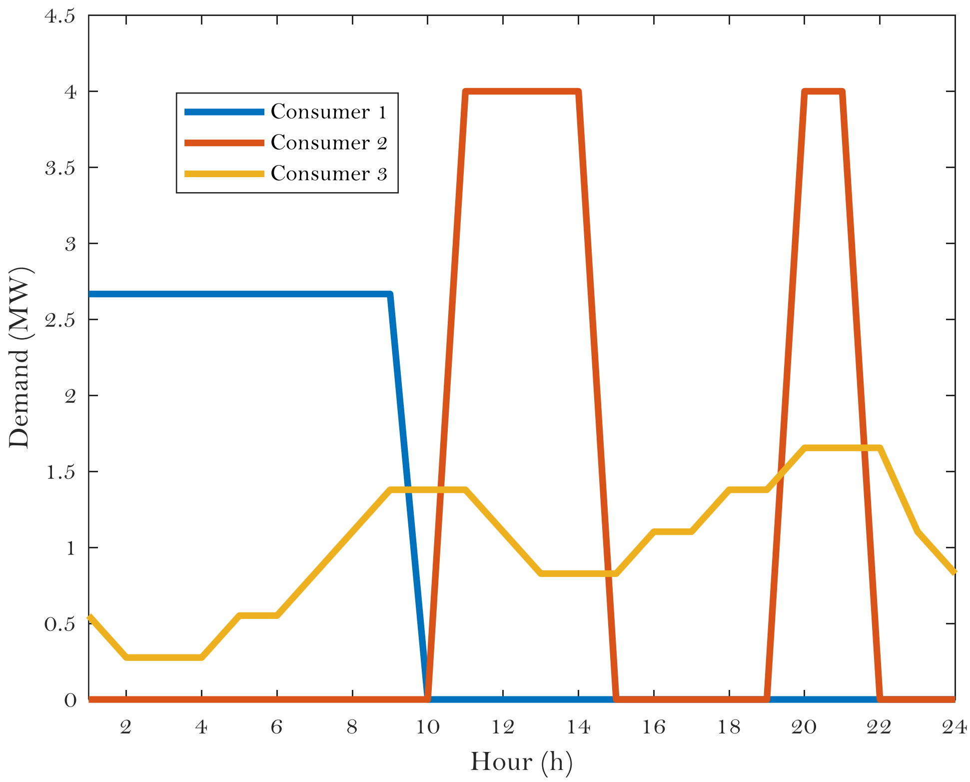

Figure 6.9: Consumer profiles to determine cost allocation in the full-scale case study

Again, each consumers cost allocation weight and share can be obtained by multiplying its hourly consumption by the hourly marginal reliability factors. The results are presented in Table 6.3. The three consumption profiles considered in the full-scale case study represent only a small fraction of the demand in the system, and therefore their cost allocation shares are very small in percentage terms. However, these shares are used here to show the comparative difference between the profiles.

Table 6.3: Cost allocation shares for the full-scale case study with a 0.002 % normalised EENS

Consumer	Cost allocation share
1	$0.011 \cdot 10^{-3}$ %
2	$10.051 \cdot 10^{-3}$ %
3	$3.066 \cdot 10^{-3}$ %

These results show how consumer 1, who only withdraws electricity at night, would have to pay a very small but still positive share of the CRM cost. Consumer 2, who only consumes during peak hours, has a cost allocation share that is more than three times higher than that calculated for consumer 3, even though they have the same total demand during the year.

It should be emphasised that the cost allocation shares also depend on the reliability standard. The latter has an impact on the expected scarcity conditions in the system and the time intervals in which they are more likely to occur. As a sensitivity analysis, the full-scale case study was run for a reliability standard of normalised EENS equal to 0.05 % (i.e., with a resource mix with fewer generation units and a higher risk of scarcity than in the previous run). The new cost allocation shares of the three consumers are shown in Table 6.4.

Table 6.4: Cost allocation shares for the full-scale case study with a 0.05% normalised EENS

Consumer	Cost allocation share
1	$0.072 \cdot 10^{-3}$ %
2	$9.114 \cdot 10^{-3}$ %
3	$3.473 \cdot 10^{-3}$ %

In this case, consumer 1 has a higher cost allocation weight, more than six times higher than with a 0.002 % EENS, and would have to pay a larger share of the CRM costs. In fact, with a normalised EENS of 0.05 %, the system registers many more scarcity conditions and they are more likely to occur during the off-peak hours, when consumer 1 withdraws electricity. For the same reason, consumer 2 has a lower cost allocation share (10% lower with a 0.05 % normalised EENS than with a 0.002 % normalised EENS) and would have to pay a lower share of the CRM costs.

6.5. Conclusions and policy implications

As CRMs become increasingly prevalent in modern power systems, their costs, which electricity consumers must bear, continue to rise. Therefore, it is becoming more and more important to establish efficient cost allocation methodologies that respond to the cost drivers of these mechanisms. Efficient allocation of CRM costs should send appropriate economic signals to consumers, encouraging them to shift their load away from hours when the system is at a higher risk of scarcity. In the long term, this would also reduce the demand for firm supply to be covered by the capacity mechanism and the overall cost of electricity supply.

This chapter proposes an efficient cost allocation methodology for CRMs based on the causer-pays principle. This is achieved by calculating hourly marginal reliability factors that reflect the impact on the reliability metric of a marginal increase in load in that hour. These factors can then be used to calculate cost allocation weights for each consumption profile, reflecting the proportion of the total CRM costs that should be allocated to that consumer. The results of the quantitative analysis presented in this chapter show that CRM costs should not be allocated exclusively to peak demand (as is the practice in many electricity systems) but should be allocated to each time interval based on the impact that a marginal increase in demand in that interval has on the system reliability metric. Peak demand hours (or peak residual demand hours in those systems with a high penetration of renewables) may well have a greater impact on the allocation of CRM costs, but off-peak hours also have a non-zero probability of scarcity conditions.

The effects of this methodology were analysed through two distinct case studies, the first being a stylised example used to illustrate how the methodology operated, and a second real-size case study using the 2019 hourly demand from mainland Spain. Although, due to limitations imposed by the convolution methodology, the generation mix was adapted to include only conventional generation, the methodology presented here would be valid with a generation mix which includes renewable energy and storage resources.

In fact, although it would require more computational effort, the same methodology presented in this chapter could be applied to more complex modelling techniques, such as those commonly used for resource adequacy assessments (e.g., Unit Commitment, or UC, models), to integrate all technologies with their constraints and obtain more accurate results. Future work should investigate how hourly marginal reliability factors and cost allocation weights can be calculated with UC models and/or based on reliability metrics other than EENS.

The assessment presented in this chapter is also closely related to the participation of demand resources in capacity mechanisms. The model allows the calculation of the cost allocation weight for each consumption profile, which can also be considered as the demand for firm supply that the profile imposes on the system. It should be evident that such a value represents a baseline for any kind of consumer participation in the CRM. If the same methodology is applied to the negative demand profile that a consumer is willing to offer to the system as demand response, it would also be possible to calculate the firm supply that the demand-response resource can offer to the system.

7. DEMAND PARTICIPATION IN CRMS

7.1. Introduction[53]

As already mentioned in the introduction, CRMs have been often criticised for subsidising fossil-fuel-driven conventional generation (ODI, 2016). In power systems with a large thermal fleet run on fossil fuels, these conventional resources have actually captured the majority of CRM revenues (Komorowska et al., 2023).

However, CRMs also present an opportunity for new technologies and resources, such as electricity storage and demand response (as the recent participation of these resources in European CRMs demonstrates; ACER, 2023b). The integration of these resources is essential, given the valuable contribution that they can provide to reliability of many power systems. International experiences have shown that, when allowed to participate, demand response resources can receive a significant share of their revenues from CRMs, as is the case of PJM, which is shown in Figure 7.1.

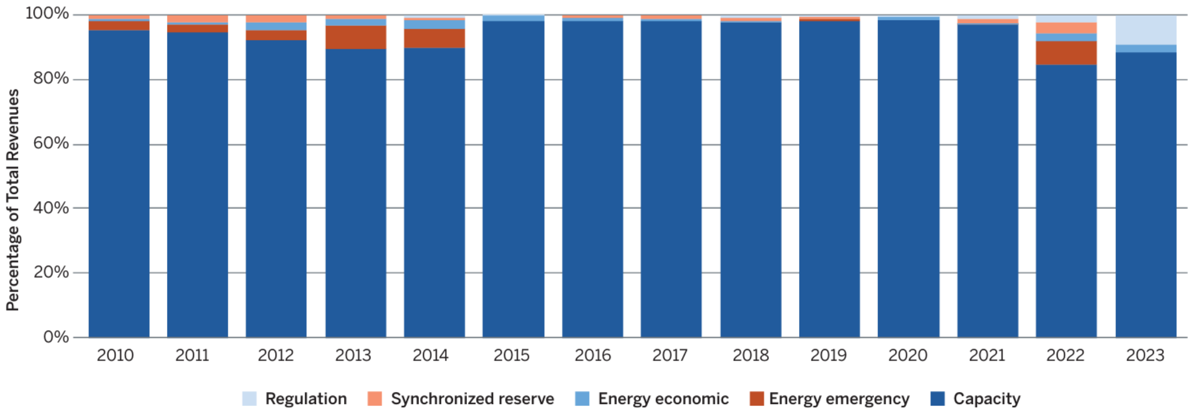

Figure 7.1: Demand Response revenue by source in PJM between 2010 and 2023 (ESIG, 2025)

However, integrating demand resources into CRMs effectively presents additional challenges than other resources, given its unique characteristics. Unlike conventional generation resources, demand resources demand firm supply but can also offer it to the system if

[53] This chapter heavily draws from Rodilla, P., Mastropietro, P., Brito-Pereira, P., 2023, The Challenge of Integrating Demand Response in Capacity Remuneration Mechanisms: Providing a Comprehensive Theoretical Framework. IEEE Power and Energy Magazine, Volume 21, Number 4, Pages 67-71.

participating as demand response. This particularity can create complications if demand resource participation is not accounted for correctly. As mentioned in chapter 6, one of the main design elements of CRMs to enable the participation of demand resources is the allocation strategy of CRM costs. A proper cost allocation strategy provides the correct economic signals to consumers, ensuring that they perceive the incentives to reduce consumption during instances where the risk of scarcity conditions is higher.

Alternatively, demand resources could participate actively in CRMs, either determining the amount of firm supply they require or offering firm supply to the system as demand response. However, integrating demand resources in this manner is challenging, given their dual nature of consumers and potential providers of firm supply. The goal of this chapter is to define a comprehensive theoretical framework for the participation of demand resources in capacity mechanisms, identifying all potential participation modes, but also highlighting the inefficiencies that could arise from certain designs, to provide guidelines for regulators who are currently addressing this issue.

The chapter is structured as follows. Section 7.2 commences by establishing the first necessary element for demand resource participation in CRMs, the definition of firm supply demand. Section 7.3 details a classification of all potential participation modes, listing the benefits and the potential inefficiencies of each of them, but also highlighting which are being used in real CRMs. Finally, section 7.4 draws regulatory recommendations.

7.2. Demand for firm supply

CRMs seek to acquire sufficient firm supply in the system in order to achieve a desired level of resource adequacy. In most CRMs, the demand for firm supply is estimated in a very aggregate way. For instance, the demand curve in capacity auctions is defined through an estimation of the whole-system demand and its evolution in the future, potentially applying a least-worst regret approach (e.g., in the United Kingdom, National Grid ESO, 2023). This approach significantly constrains the kind of participation that can be expected by demand resources, as analysed in the next section.

The most efficient way for demand to participate would be the involvement of end-users in this initial phase, letting consumers define their own demand for firm supply without administrative interventions; this demand for firm supply would become the upper limit of their consumption during scarcity conditions in the system. Alternatively, the regulator could estimate an initial requirement for each consumer or consumer group and then allow them to increase or decrease such value. These approaches would also simplify and improve the efficiency of the cost allocation, since such self-declared demand for firm supply is the best cost driver on which to apply CRM charges. This process would be symmetric to the de-rating process for generation resources. Each consumer (or consumer group/category) would pay the costs of the capacity mechanism according to the consumer's expected "negative" contribution to the reliability of the system (since procuring 1 MW of firm capacity entails contributing "negatively" for that amount).

7.3. Different modes of participation in CRMs

After having defined the necessary background on the methodologies for estimating the demand for firm supply and the strategies for CRM cost allocation, it is possible to classify the participation for demand resources in capacity markets into different participation modes. Using the standard terminology on demand response, two broad categories are identified hereunder.

- Explicit participation: consumers explicitly take part in some phase of the capacity market and assume binding commitments; as analysed in section 7.3.1.1, they can do that in the demand side of the capacity market, by defining their demand for firm supply, or in the supply side, selling demand-response services that are equated to the reliability services offered by generators.
- Implicit participation: consumers do not explicitly participate in the capacity market, they do not assume any binding commitment to reduce their load, but they react to CRM charges during its operation, modifying their demand to reduce their contribution to the coverage of CRM costs (and, if charges are designed properly, their contribution to scarcity conditions). This participation mode is addressed in section 7.3.2.

7.3.1. Explicit participation

7.3.1.1. Demand side (opt-in or opt-out)

Although very infrequently used in practice, the most obvious way to involve consumers in capacity mechanisms would be to conceive an active role for them in the calculation of the demand for firm supply. Ideally, consumers could be asked to define beforehand their demand for firm supply (the selection of this demand for firm supply could be informed by some brief report from the system operator with estimations on the number of stress events expected in the system and on the range of the charges which this capacity demand would be subject to). In power systems where smart meters have already been deployed, this approach could encompass the entire demand, including residential or regulated demand. In a few countries (e.g., Spain), consumers are already asked to specify different contracted capacities, e.g., for peak or valley hours, which are subject to different charges (CNMC, 2020). Widening this approach to include resource adequacy would only require asking consumers to specify an additional contracted capacity that would be used to limit consumption during scarcity conditions (or to impose sanctions on the withdrawals exceeding it).

This theoretical approach would move the responsibility of defining the demand for firm supply fully onto consumers' shoulders. Although technically feasible, this may be challenging from a regulatory and political point of view. However, there are other approaches that mimic this first alternative and partially achieve its benefits. For instance, the regulator or the system operator could estimate an aggregated demand for firm supply but compute a disaggregated estimation for certain consumer categories (e.g., large commercial or industrial end-users). The latter would then be given the chance of opting out, i.e., of reducing or directly setting to zero the demand for firm supply assigned to them. The opt-out would generate a commitment that allows the system operator to limit withdrawals during scarcity conditions, but it would also exempt the consumer from paying CRM charges for the opted-out capacity. This approach is represented graphically in Figure 7.2, for a centralised capacity auction (the same reasoning could be applied, however, to decentralised capacity markets).

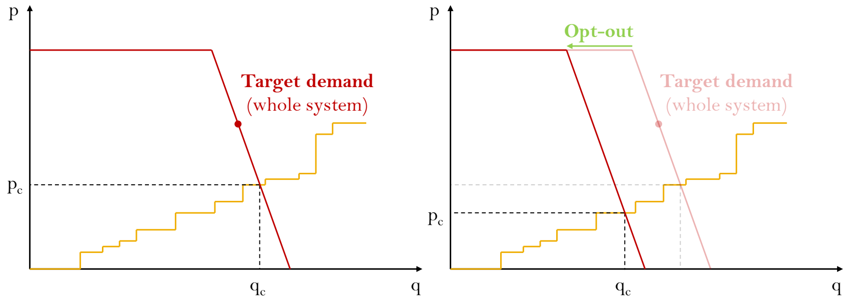

Figure 7.2: Explicit participation of demand resources in the demand side of the CRM through an opt-out (the demand curve mimics curves used in real CRMs, which represent demand elasticity)

A similar approach would consist of estimating the demand for firm supply only for certain consumer categories (e.g., residential or regulated demand) and requiring the rest of consumers to define their own demand through an explicit opt-in in the capacity market, which would generate the same commitments that have already been mentioned above. This opted-in capacity could be used to simply shift the demand curve (chart in the middle in Figure 7.3), or these consumers could be asked to present price-quantity demand bids, specifying also the value that they assign to the firm supply (chart at the right in Figure 7.3).

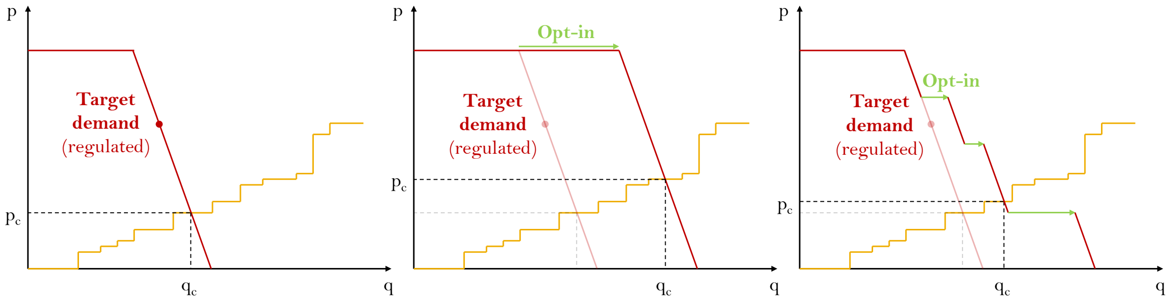

Figure 7.3: Explicit participation of demand resources in the demand side of the CRM through an opt-in

It must be remarked that these approaches would also simplify the allocation of CRM costs. The demand for firm supply is the real driver of these costs; therefore, if some or all consumers have a certain demand for firm supply earmarked to them, either estimated by the regulator/system operator or self-defined by the end-user, CRM charges could be easily applied to this capacity during each delivery period.

7.3.1.2. Supply side (demand response)

The disaggregation of the demand for firm supply is a complex task that, as mentioned in section 7.2, is hardly found in real CRMs. Especially in centralised capacity markets, the demand is commonly defined for the entire system and no opt-in or opt-out is allowed. In this context, consumers can still participate in the capacity market by offering demand-response services. These services are offered through price-quantity supply bids that go into the supply curve of the market, as shown in Figure 7.4. However, it must be highlighted that the consumers involved in such demand response are represented twice in the auction, both in the demand curve (since they are part of the whole-system demand for firm supply) and in the supply curve. This feature is prone to arbitrages and other inefficiencies, as analysed next.

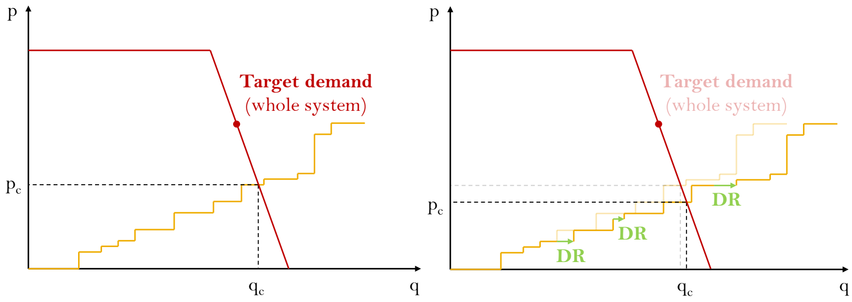

Figure 7.4: Explicit participation of demand resources in the supply side of the CRM through DR services

For consumers to offer demand-response services in the CRM, the regulator must design a reliability product that these resources will be allowed to trade in the capacity market. In principle, the reliability product should be the same for all the resources competing in this market segment and should reflect the ability of each agent to contribute to the reliability target in force in the power system. However, many regulators, both in Europe and the United States, have defined specific reliability products that are tailored somehow to the characteristics of demand resources and are meant to foster their participation. This is the case, for instance, of the reliability-option CRMs introduced in Belgium (EC, 2021) and Italy (EC, 2018) and of the decentralised capacity obligations traded in France (EC, 2016c). For instance, in Belgium, demand resources are allowed to bid their own strike price, a key element for the settlement of the reliability option (EC, 2021; Elia 2024b).

Baselining

Another element required to allow this kind of demand participation and the assessment of its performance is a methodology to identify a demand baseline. The latter may be used to define the firm supply of these demand resources, in conjunction with a de-rating factor. De-rating of demand resources usually depends on the self-declared or tested duration of the service that the demand aggregator can provide. The lower the duration, the lower the expected contribution to scarcity conditions, usually, and, consequently, the de-rating assigned to the resource. The demand baseline is also essential to verify and quantify the compliance of these resources to their capacity commitments, by comparing the actual withdrawal with the one that would have been registered if the service had not been activated. As for any other DR programme, several baselining approaches are possible. Some recent studies, such as the ones developed by Elia, the Belgian TSO, found that the most widely adopted methods for capacity mechanisms are historical and control-group baselining (Elia, 2021). The former uses historical data to estimate the expected demand in the activation period, by applying exclusion rules and rankings. For instance, high-X-of-Y methodologies focus on the last Y days of the same kind as the activation day (e.g., working days) and, within this group, select the X days with the higher load. For each settlement period, the baseline is defined by averaging the load during these X days. Historical baselines may also rely on some sort of same-day adjustment, i.e., a methodology that modifies the baseline according to the load registered during the day of activation (with expedients to avoid gaming from the demand resource to overestimate its contribution). A typical example of a historical baseline methodology with same-day adjustment used in CAISO is shown in Figure 7.5.

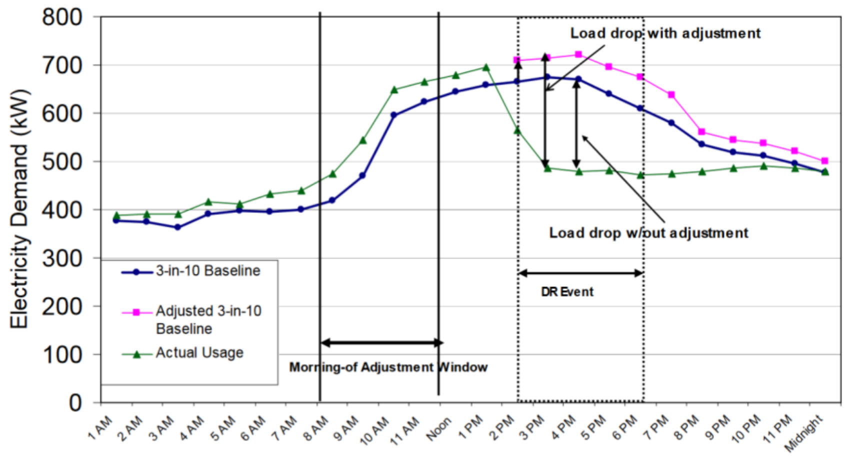

Figure 7.5: Historical baselining with morning-of adjustment used in CAISO (CPUC, 2009)

Control group baselining does not rely on historical data but estimates the load that the demand resource would have withdrawn if it was not activated based on the withdrawals of a control group of consumers. These consumers may be selected among those who are providing demand response (randomised controlled trial over a small number of active consumers), or they may be consumers with similar characteristics as those in the demand-response programme, but who do not provide DR services (Nexant, 2017). This type of demand baselining is used especially in the United States.

Other baseline methodologies, which are rarely used in the framework of capacity mechanisms, are MBMA (Meter Before - Meter After, which defines the baseline based simply on the load registered before the DR service is activated), declarative baselining (the baseline is estimated directly by the DR service provider, who communicates it to the operator), and regression-based baselining (the baseline is computed through a complex formula with several parameters, as temperature and daylight, whose coefficients are defined based on historical data) (Elia, 2021).

The double-remuneration (or double-benefit) problem

The supply-side explicit participation of demand resources in CRMs has a significant disadvantage that stems from the presence of a certain group of consumers both in the demand and the supply curve of the capacity market. A demand resource taking part in the capacity mechanism is remunerated for reducing its load, especially during scarcity conditions. However, by doing so, it can also reduce its contribution to the recovery of CRM costs, if the charges are designed to allocate these costs to the withdrawals during scarcity conditions. This way, the demand resource is remunerated twice and, more importantly, its net position in the CRM could be larger than zero, i.e., it could have a net revenue from its participation.

However, what the demand resource is actually doing is just to reduce or to set to zero its load during scarcity conditions, with the final goal of avoiding the payment of CRM charges (and without benefitting from the coverage of the mechanism). Therefore, its net position in the CRM should be, at maximum, null. This was clearly stated by the Italian regulator ARERA during the design of its capacity mechanism. According to its criteria, a demand resource involved in a capacity market provides a service that can only be used by itself (through the consumers it is formed by, which are also represented in the demand curve), but which cannot be provided to a third party, differently to the service provided by a generator, whose position in the CRM is of course expected to be larger than zero (AEEGSI, 2017).

The risk of a double remuneration for demand resources depends on the design of CRM charges. Paradoxically, the double-remuneration problem has been avoided so far, thanks to the inefficient cost-allocation strategies adopted in most capacity mechanisms. Volumetric charges covering a very large number of hours reduce the benefit that can be achieved by demand resources reacting during scarcity conditions. However, cost-allocation strategies based on capacity charges during scarcity conditions could increase the risk of double remuneration for demand resources. The most efficient way to deal with this problem would be, once again, to introduce fixed CRM charges (based, for instance, on historical consumption during scarcity conditions). With this approach, the demand providing DR services would pay a fixed amount of CRM costs and would offset this quantity by the revenues it receives from the capacity market. If all the elements of the CRM are properly harmonised, this should result in a net position close to zero, although deviations are possible.

7.3.2. Implicit participation

Once the capacity market is cleared and commitments assigned, there is still some space for implicit demand participation. The potential for this kind of participation mode clearly depends on the design of the charges introduced to recover the costs of the mechanism and on the signals they convey. Demand resources can basically shift their load to minimise payments derived from these charges, moving their consumption out of potential scarcity conditions.

As mentioned in chapter 6, the real driver of CRM costs is the contribution of each consumer to the demand for firm supply and a good proxy parameter to estimate this contribution is the historical load during scarcity conditions, with a moving average. Using historical data with a moving window certainly dilutes but does not eliminate the signal for consumers to reduce their load. If data from the last five years are used, an end-user who manages to reduce her load significantly during all scarcity conditions registered in the system would considerably reduce her CRM payments after five years. The signal could be further strengthened if charges are applied on consumption during scarcity conditions in the delivery period, although this strategy may affect cost recovery.

In real CRMs, however, implicit participation of demand resources has always been almost inexistent, since, in the majority of cases, CRM costs are recovered through volumetric charges applied over a large number of hours, which impede an efficient reaction by the load. The United Kingdom offers a paradigmatic example. Since the introduction of the capacity market, CRM costs were recovered through a charge on electricity suppliers that was applied to the net demand (gross demand minus embedded generation) they served from 16:00 to 19:00 in the working days from November to February. This prompted suppliers to sign agreements with embedded generation (mainly diesel gensets) to produce in those hours, thus reducing the net demand. This kind of demand response was not efficient from an adequacy point of view, since the load reduction was taking place in hours where no scarcity was registered, and it was having a harmful environmental impact. For this reason, the cost-allocation strategy was modified in 2018 (and later on in 2024) and CRM charges are, as of the time of writing this document, applied on the gross demand minus the demand of energy intensive industries of each supplier (EMRS, 2024a and 2024b), in order to avoid this inefficient implicit participation of demand resources. However, the inefficiency stemmed from an inefficient cost-allocation strategy, which was only partially amended by the 2018 reform.

7.4. Conclusions and regulatory recommendations

As flexibility needs rise in power systems, demand resources become increasingly important to achieve resource adequacy. Most capacity mechanisms in place today allow the participation of demand resources, although with different rules and different outcomes. While demand response covers a larger share of the demand for firm supply in some power systems in the United States, Europe is lagging behind and DR only accounts for less than 5% of capacity markets in the region (ACER, 2024b).

This chapter presents a comprehensive theoretical framework for the participation of demand resources in capacity mechanisms. The first element of this framework is defining the firm supply demand to achieve the adequacy needed in the system. Consumers can then participate explicitly in the CRM by determining their firm supply needs, in the demand side of the CRM, reducing the payments they would have to disburse.

If consumers are not allowed to take part in the definition of the demand for firm supply, then they can only be allowed to participate in the supply side of the capacity market, where they could sell demand-response services. This approach is actually the most widely adopted in capacity mechanisms. However, it presents several complexities, stemming from the fact that the same demand is represented twice in the capacity market, both in the supply and the demand side. This situation could result in the so-called double-remuneration problem, when a demand resource is remunerated for reducing its load during scarcity conditions, but by doing so it also reduces its contribution to the coverage of CRM costs.

The participation on the supply side also requires methodologies for the definition of a baseline. As mentioned in the chapter, the most widespread methodologies are historical (e.g., high-X-of-Y) and control-group baselining. In theory, demand resources should be required to provide exactly the same reliability product as other resources, since they compete in the same market. However, several regulators have defined specific products that are tailored to DR services and are meant to reduce the risk perceived by these agents and to incentivise their participation.

All these complexities for the participation of demand resources in the supply side of the capacity market, if not properly addressed, may result in significant inefficiencies in the operation of the CRM. Regulators should strike the right balance between supporting demand response and ensuring the performance of the capacity mechanism.

8. CONCLUSIONS AND FUTURE WORK

8.1. Introduction

More than forty years after the first wave of power system liberalisations, capacity remuneration mechanisms have proved to be a key regulatory instrument to guide the expansion of the resource mix towards the desired level of reliability. As highlighted in chapter 1, CRMs address the main market imperfections that prevent an efficient expansion of the resource mix by introducing an economic signal that complements the energy market price.

However, since their inception, CRMs have had to be adapted and their rules amended to be able to fulfil their objectives efficiently. Given their central role in achieving a desired level of resource adequacy and the impact of their design in the resource mix, it is important to constantly guarantee that CRMs are properly designed and fit for purpose.

The energy transition that power systems worldwide are undergoing to mitigate climate change is causing a drastic shift in both the composition of the resource mix and electricity consumption patterns. In turn, this has led to a shift in the scarcity conditions experienced by power systems, which poses a challenge for resource adequacy assessments that aim to characterise these conditions and for CRM that aim to prevent them.

This thesis aims to contribute to the ongoing debate about how CRMs should be reformed. It addresses several important aspects that have recently attracted the attention of experts and policymakers, particularly with regard to the selection of resource adequacy metrics, the methodologies to determine firm supply, the allocation of CRM costs among consumers and the methods for an efficient participation of demand resources in CRMs.

8.2. Recommendations on resource adequacy metrics

As highlighted in chapter 2, power systems undergoing an energy transition are experiencing rapid changes that are also affecting the scarcity conditions they face. Resource adequacy metrics that have been used historically, such as the reserve margin and the LOLE, are becoming increasingly inefficient to represent current and future scarcity conditions.

Therefore, system operators must reconsider the resource adequacy metrics being used to assess resource adequacy, both in terms of the contingency and the statistical measure used to characterise scarcity conditions.

Chapter 2 detailed some key principles that resource adequacy metrics should follow, according to academic literature, international practices and regulatory principles. In order to be resilient to future developments, system operators should avoid the use of discrete resource adequacy metrics, such as LOLE, which focuses on individual events. Therefore, system operators should focus on continuous metrics, such as the resource adequacy metric proposed in chapter 2, the CVaR of non-served energy plus the energy cleared and supplied beyond a price threshold. The proposed resource adequacy metric has two main benefits: i) as it internalises the price dimension, it is resilient to an increase in the elasticity of electricity demand, which might render non-served energy metrics obsolete, and ii) the CVaR allows the system regulator to focus on extreme events and to represent the regulator's risk aversion. Although both elements of the recommended resource adequacy metric are used in other contexts and are not new to power sector regulation and economics, the inclusion of these elements in a resource adequacy metric is a main contribution of this thesis and an improvement from the more widespread metrics, given recent and expected developments of power systems.

There might also be benefits to analysing resource adequacy using multiple metrics instead of a single one, given the multiple facets of current and future scarcity conditions. In this case, the use of several resource adequacy metrics can provide system operators with a clearer understanding of the scarcity conditions that the system may experience. However, as shown in the different case studies in chapter 4, if these resource adequacy metrics are used to set several reliability standards, there is a risk that some of them end up being redundant. The possibility of redundancy between these reliability standards increases depending on how they are correlated, leading to a small multi-activation area (or volume/hypervolume if more than 2/3 reliability standards are used).

Alternatively, to avoid the redundancy problem while analysing the different facets of scarcity conditions, a composite reliability standard can be used, which combines several reliability standards into a single one using a weighting factor. A normalization process can be used to combine several reliability standards that are radically different and even use different units of measurement. Chapter 4 showed a normalisation process which allowed for the combination of EENS and CVaR(ENS) metrics into a single composite metric. The disadvantage of composite reliability standards is that it is not possible to guarantee that the embedded reliability standards used to build the composite one are fulfilled.

8.3. Recommendations on firm supply

The firm supply calculation process is one of the most important design elements of a CRM, since it determines the remuneration that each technology can obtain from the mechanism, with the consequent impact on the resource mix. In order to achieve the desired level of resource adequacy, power systems must acquire sufficient firm supply to cover demand. To determine the amount of firm supply that each resource can offer, system operators must estimate the contribution from these resources to solve or avoid scarcity conditions.

Historically, the firm supply was calculated through simplistic methodologies, which reflected the characteristics of power systems at the time. However, the rapid deployment of non-conventional resources, such as RES-E, electricity storage and demand response resources, is rendering those methodologies increasingly obsolete. This thesis has provided, in chapter 3, a set of recommendations on how the firm supply calculation methodology should be adapted to the new context. Firstly, the firm supply calculation methodology should be based on the same resource adequacy metric as the one used to define the reliability standard in the system. Secondly, firm supply should be calculated according to the marginal contribution of each resource (or technology) to achieve the reliability standard. Additionally, chapter 3 also presented an efficient alternative to this procedure which is a novel contribution of this thesis and consists of analysing the generation of resources during the expected scarcity conditions, which yields similar results with a reduced computational time. In any case, these estimations should be performed using a probabilistic model that simulates the future operation and composition of the power system.

The firm supply calculation process can be further complicated if several reliability standards are established simultaneously. As highlighted in chapter 4, implementing two or more reliability standards leads to the necessity of calculating the same number of firm supply values, one for each reliability standard, and procuring different reliability products. In the second case study of chapter 4, this leads to two de-rating factors being calculated for each resource, one for the contribution to the EENS and the other one for the contribution to the CVaR(ENS) reliability standard. In the case of CCGTs, which were affected by the polar vortex event simulated in the case studies of chapter 4, this led to a significant difference in the calculated de-rating factor for the EENS and the CVaR(ENS) reliability standards.

To avoid these hurdles, a composite reliability standard can be used instead of multiple single-metric reliability standards. This alternative avoids the calculation of multiple de-rating factors and the consequent procurement of different reliability products. In this case, the firm supply would represent the expected contribution of each resource to achieve the composite reliability standard, as shown in the third case study of chapter 4.

Once the firm supply is calculated, considering the reliability standard defined for the system, its actual implementation in the auction process of a CRM can present additional complications. The de-rating factors established before the auction process of a CRM might lead the system to a composition which is not the one envisioned when calculating the de-rating factors, as shown in chapter 5. In this case, the firm supply actually provided by certain resources could end up being significantly different from the one estimated beforehand and which they are being remunerated for.

The effect of varying the ex-ante de-rating of solar PV was studied in chapter 5 using a two-step model which simulates the auction process of a CRM and its effect on the resource mix. In this case, the higher the de-rating factor of solar PV, the higher its competitiveness and presence in the resource mix resulting from the CRM. In turn, this affects the operation of the system, its scarcity conditions and the actual contribution of PV resource to solve them. For low PV de-rating factor values, the system experienced high adequacy levels at a high CRM cost, while for high PV de-rating values, the system experienced low adequacy at a low CRM cost. This analysis highlights the profound impact that the ex-ante definition of firm supply can have on the system.

8.4. Recommendations on cost allocation and demand participation

The rise in the importance of CRMs is provoking increasing costs to guarantee the desired level of resource adequacy, and this requires new methodologies to allocate these costs among consumers. An efficient cost allocation methodology should send appropriate economic signals to consumers, dissuading consumption during time periods with higher risk of scarcity conditions, leading to a reduction in the future demand for firm supply.

Chapter 6 presented an easy-to-apply methodology for CRM cost allocation. This methodology is based on hourly charges that reflect the likelihood of having scarcity conditions. These charges should be applied to historical consumption patterns to calculate a cost allocation weight for each consumer, which can then be used to determine how the costs of the CRM should be distributed among end-users. The analyses presented in chapter 6 highlight that costs should not only be allocated during peak (or net-peak) demand hours but rather based on the impact that a marginal increase in demand would have in each period.

This cost allocation procedure is the first necessary step for an efficient participation of demand resources in CRMs, as it implicitly represents the firm supply demanded by each consumer. The establishment of the cost allocation methodology allows consumers to participate implicitly in the CRM. By changing consumption patterns, consumers can progressively reduce the CRM charges they have to pay, as described in chapter 7.

However, power systems should aim to allow the explicit participation of demand resources to fully integrate them and capture all their capabilities, such as their flexibility capacity. Chapter 7 highlights the different forms of explicit participation that can be established for demand resources. The first is the participation in the demand side of a CRM by determining their firm supply needs, either through an opt-in or an opt-out process.

The second alternative for explicit participation is to allow demand resources to participate in the supply side of a CRM as demand response. However, as shown in chapter 7, this approach can present inefficiencies and the risk of a double remuneration problem. This occurs because the same demand resource is represented twice, both demanding firm supply in the demand side of the auction and offering demand response services in the supply side of the auction. In this case, if not accounted for correctly, the demand response resource could receive remuneration for the demand response services provided and simultaneously reduce its CRM payments. Additionally, the participation of demand resources in the supply side of the auction requires the definition of a baseline, which can be quite complex. To ensure the correct participation of demand response resources in the supply side of a CRM, regulators must be aware of these complexities and avoid potential inefficiencies.

8.5. Future work

This thesis has covered a wide array of research topics that aim to improve the design of capacity remuneration mechanisms, particularly regarding resource adequacy metrics, the firm supply calculation process, and the cost allocation and participation alternatives for demand resources. However, this research has also revealed, on the one hand, several research lines that could be explored in the future but could not be covered in this thesis and, on the other hand, several research topics that have been addressed in this thesis, but which could be further elaborated.

Although chapter 2 presents a proposal for a price-based resource adequacy metric, which is also used in chapter 5 to analyse scarcity conditions in its case study, it is not used in the rest of the case studies of this thesis. Although the author believes that using a price-based resource adequacy metric guarantees that it remains resilient as electricity demand becomes increasingly elastic, regulatory trends in CRMs have not led in that particular direction, with the exception of the Brazilian resource adequacy scheme. This is why, apart from the exception of chapters 2 and 5, most of the discussions regarding resource adequacy metrics have been centred around non-served energy to better engage with current discussions on the matter. However, further work should be conducted on the best approach to include price-based resource adequacy metrics in CRMs, given that the proposed metric in chapter 2, the energy offered and served above a price threshold, is just one possible example of a price-based metric. Additionally, including price-based metrics in a multi-metric or composite-metric approach could lead to interesting insights that have not been able to be explored in this thesis.

Chapter 4 presented a case study to compare multi-metric and composite-metric approaches to reliability standards. As mentioned in that chapter, to the best of the author's knowledge, no similar analysis has been carried out to compare these approaches. However, although the analysis provided interesting insights to help regulators decide the most efficient approach, it was limited to two different reliability standards. In order to obtain a deeper understanding of the dichotomy between a multi-metric and a composite-metric approach, more reliability standards should be studied, and different alternatives to combine them into the composite reliability standard analysed. Combinations of different resource adequacy metrics could lead to a better understanding of the different factors that shape the multi-activation area/volume/hypervolume, while different normalisation approaches for the definition of a composite metric could provide better qualities for both the reliability standard and the firm supply calculation process.

Electricity storage is becoming increasingly prevalent in modern power systems, especially as a flexibility provider. However, including storage in power system analyses can increase complexity, both computationally and in terms of understanding the results of the analyses. This is why, with the exception of the case study introduced in chapter 3, electricity storage resources have not been included in the remainder of the case studies presented in this thesis. Given the relevance of electricity storage in modern power systems, future research should aim to enhance the analyses presented in this thesis with the inclusion of these resources. In particular, the effect of the ex-ante definition of the firm supply of energy storage could be of interest, as it may exhibit similar patterns to that observed for solar PV in chapter 5 or reduce these effects for other resources, if electricity storage is present in the system.

Energy systems integration is reshaping modern power systems. As different elements and energy vectors become intertwined with the electricity sector, analysing resource adequacy will become increasingly complex. In this case, it may not be possible to analyse the adequacy of a power system without considering the adequacy of other energy systems simultaneously or to disregard the operation of these systems when analysing the adequacy of the power system. A future research line could be dedicated to the analysis of the discussions introduced in this thesis, regarding resource adequacy metrics, firm supply definition and demand participation, and how these discussions are being reshaped by energy systems integration.

As renewable energy sources, battery storage technologies and demand response resources become more prevalent, the need for flexibility in power systems will increase. In this case, concerns about resource adequacy could arise due to a lack of flexible resources instead of insufficient generation capacity to meet peak demand, as was usually the case in the past in capacity-constrained systems. While the principles and conclusions presented in this thesis are agnostic with respect to the causes of resource adequacy concerns in power systems, the proposed resource adequacy metric will reflect scarcity conditions regardless of the root cause. However, this thesis has not addressed the particularities that may arise as power systems become more dependent on flexible resources. Therefore, future research could examine in more depth how capacity remuneration mechanisms are affected by an increase in the need for flexibility in power systems, and how capacity mechanisms may overlap with flexibility support schemes.

Resource adequacy analyses primarily concentrate on electricity generation and demand, often oversimplifying or disregarding the electricity grids (and other energy vectors). In some cases, limited interconnection capacity in the electricity grid can hinder the ability of resources to contribute to solving adequacy concerns. In these instances, resources of the same technology will provide different levels of firm supply, due not only to the availability of the primary energy source (such as wind), but also to grid capacity and their location. Although this effect was mentioned in section 3.2.5, none of the case studies provided in this thesis have considered the presence of an electricity grid. Future work on capacity remuneration mechanisms should incorporate electricity grids and other energy grids, such as natural gas grids, to improve both resource adequacy assessment and the process of calculating firm supply.

Chapter 2 proposed that future-proofed resource adequacy metrics should incorporate the price dimension, in order to be resilient to an increase in demand elasticity. However, it did not consider possible nuances that could arise when using price-based metrics, such as whether the price threshold should be static or dynamic with indexation of some kind. If the price threshold is updated dynamically by indexing it to an energy commodity, such as natural gas, scarcity conditions could arise that would not be reflected in the resource adequacy metric if the indexed commodity became scarce and its price increased significantly. Conversely, the underlying commodity's price could rise significantly due to factors other than real scarcity. In some cases, as in the 2022 European energy crisis, the distinction between these two issues can be unclear. Future research should address more precisely how the price dimension can be effectively incorporated into resource adequacy metrics.

Finally, chapter 6 addressed the question of how CRM costs should be allocated among consumers, proposing the design of hourly charges according to their negative marginal contributions to adequacy. However, the practical application of these charges in a CRM was not discussed in detail. The implementation of these charges must strike a balance between the effectiveness of the economic signals provided to consumers and the cost recovery of the mechanism. For example, if the charges are designed ex-ante, they provide effective signals to consumers at the expense of cost recovery, if actual consumption is used instead of historical consumption. Conversely, ex-post charges guarantee cost recovery, but do not provide effective economic signals to consumers. Further research is required to determine the most effective implementation of the hourly charges outlined in chapter 6.

List of Publications

Journal articles

Brito-Pereira, P., Mastropietro, P., Rodilla, P., Barroso, L. A., Batlle, C., 2022a. Adjusting the aim of capacity mechanisms: Future-proof reliability metrics and firm supply calculations. Energy Policy, Volume 164, Article 112891.

Brito-Pereira, P., Rodilla, P., Mastropietro, P., Batlle, C., 2022b. Self-fulfilling or self-destroying prophecy? The relevance of de-rating factors in modern capacity mechanisms. Applied Energy, Volume 314, Article 118939.

Rodilla, P., Mastropietro, P., Brito-Pereira, P., 2023. The Challenge of Integrating Demand Response in Capacity Remuneration Mechanisms: Providing a Comprehensive Theoretical Framework. IEEE Power and Energy Magazine, Volume 21, Number 4, Pages 67-71.

Brito-Pereira, P., Bruninx, K., De Vries, L., Mastropietro, P., Rodilla, P., 2025. Future-proofed resource adequacy metrics: a model-based assessment of multi-metric vs. composite-metric reliability standards. Sustainable Energy, Grids and Networks, Volume 44, Article 101957.

Conference presentations

Brito-Pereira, P., Rodilla, P., Mastropietro, P., Batlle, C., 2023. Firm supply of demand resources and CRM cost allocation, 18th IAEE European Energy Conference, Milan (Italy). 24-27 July 2023.

Brito-Pereira, P., Rodilla, P., Mastropietro, P., Barroso, L.A., Batlle, C., 2022. Efficient de-rating in modern capacity mechanisms and the interdependence with reliability metrics, 17th IAEE European Energy Conference, Athens (Greece). 21-24 September 2022.

Mastropietro, P., Rodilla, P., Brito-Pereira, P., Batlle, C., 2024. Future-proofed resource adequacy metrics, IEEE Power & Energy Society General Meeting - IEEE PES GM 2024, Seattle (United States of America). 21-25 July 2024.

Working papers

Brito-Pereira, P., Rodilla, P., Mastropietro, P., 2025. Efficient cost allocation in capacity remuneration mechanisms: applying the causer-pays principle to resource adequacy. Working paper IIT-24-052. Under review at International Journal of Electrical Power and Energy Systems.

Rodilla, P., Brito-Pereira, P., Mastropietro, P., Batlle, C., 2021. Consideraciones previas al diseño de un mecanismo de capacidad en el sistema eléctrico español, Working paper IIT-20-202A.

ANNEX I

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

In order to do this, the benchmark optimization problem, which is a stylised version of the ideal central planner problem with an adequacy constraint is first formulated and solved. This adequacy constraint is expressed by means of a resource adequacy metric (RM). The problem of the individual agents is then formulated and solved, where the areas of interest are their participation in the energy and in the capacity market. By comparing the optimality conditions of both problems we can 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:

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

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.

- $C(Q_1, Q_2, Q_3,\dots Q_n)$ 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.
- $I(K_1, K_2, K_3,\dots K_n)$ 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:

$$Q_i \leq K_i \qquad (\alpha_i) \qquad (2)$$

$$RM(K_1,K_2,K_3,\dots K_n) \leq RS \qquad (\beta) \qquad (3)$$

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 resource adequacy metric RM, which is assumed to be dependent only on the mix $RM=RM(K_1, K_2, K_3,\dots K_n)$, to fulfil a certain reliability standard RS, 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 the Lagrangian function, *L*, is formulated and its first partial derivatives with respect to the decision variables is computed.

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

$$- I(K_1,K_2,K_3,\dots K_n) + \textstyle\sum_1^n\left(Q_i-K_i\right)\cdot\alpha_i + (RM(K_1,K_2,K_3,\dots K_n) - RS)\cdot\beta \qquad (4)$$

If the first partial derivative of this expression with respect to the decision variable K_i, which is the installed capacity of unit *i*, is computed the following expression is obtained:

$$\frac{\partial L}{\partial K_i}=-\frac{\partial I(K_1,K_2,K_3,\dots K_n)}{\partial K_i}-\alpha_i+\frac{\partial RM(K_1,K_2,K_3,\dots K_n)}{\partial K_i}\beta=0 \qquad (5)$$

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:

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

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:

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

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 the Lagrangian function of this second problem, *L*, is formulated and its first partial derivatives with respect to the decision variables is computed.

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

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

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

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 the following expression is obtained:

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

Which leads to the following remuneration in the capacity market:

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

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.

ANNEX II

This section details the UC model used to perform the case study presented in Section 3.2.3.1. The formulation presented in this section represents the UC model ran for each case (base case and marginal executions) and for each of the availability, renewables and demand scenarios (1000 in total for each of the cases). Each of these individual executions are deterministic, with generators having perfect foresight throughout the entire execution window (8760 hours). This UC model is formulated and solved as a Relaxed Mixed Integer Programming (RMIP) problem.

This UC model reproduces a centralised hourly day-ahead market with inelastic demand and seeking to minimise operational and non-served energy costs. Power plants in the model are subject to unplanned outages, which are represented through a vector that determines the hourly availability. These are constructed by implementing a two-state Markov chain. Additionally, power plants of the same technology are considered in an aggregate manner, to keep execution times manageable. The specifics on how unavailabilites are considered, and how power plants are aggregated can be found in greater detail in chapter 5, which uses the same approach, and in the article by Mastropietro et al. (2016).

<u>Indexes and sets</u>

$g \in G$	Generating technologies
$B \subseteq G$	Battery storage subset
$H \subseteq G$	Hydro subset
$t \in T$	Hourly periods
$sc \in SC$	Scenarios

<u>Parameters</u>

C_g^V	Variable cost of a unit of technology g [€/MWh]
C_g^{NL}	No-load cost of a unit of technology g [€]
C_g^{SU}	Start-up cost of a unit of technology g [€]

C_g^{SD}	Shut-down cost of a unit of technology g [€]
C^{NSE}	Non-served energy price (in this case 3000 €/MWh) [€/MWh]
$\overline{P}_g$	Maximum power output of a unit of technology g [MW]
$\underline{P}_g$	Minimum power output of a unit of technology g [MW]
$AV_{sc,g,t}$	Availability of technology g available in period t in scenario sc [p.u.]
D_t	Demand in period t [MWh]
N_g	Number of units installed of technology g [p.u.]
$StoCap_g$	Storage capacity for technology g (only for batteries)
$\overline{EFOR}_g$	Maximum equivalent forced outage rate of a unit of technology g [p.u.]
$\underline{EFOR}_g$	Minimum equivalent forced outage rate of a unit of technology g [p.u.]
MTR_g	Mean time for recovery for units of technology g [h]
$StoIni_g$	Initial storage for hydro [MWh]
Eff_g	Round cycle efficiency for battery storage, set at 0.9 [p.u.]

Variables

nse_t	Non-served energy in period t [MWh]
$p_{g,t}$	Power output above minimum output of all technology g units in period t [MW]
$u_{g,t}$	Number of units of technology g committed in period t [p.u.]
$v_{g,t}$	Start-up decision of technology g in period t [p.u.]
$w_{g,t}$	Shut-down decision of technology g in period t [p.u.]
$binj_{g,t}$	Battery charge-up in period t [MW]
$bsto_{g,t}$	Battery storage level in period t [MW]
$hsto_{g,t}$	Hydro storage level in period t [MW]

$nse_{sc,t}, p_{sc,g,t}, binj_{g,t}, bsto_{g,t}, hsto_{g,t} \in \mathbb{R}_{\geq 0}$

$u_{g,t}, v_{g,t}, w_{g,t} \in \mathbb{Z}_{\geq 0}$

Data

For this case study, the initial electricity mix tries to be similar to that of the Spanish electricity mix of 2022. Several adaptations have been made to this end. Solar thermal generation has not been considered, and instead it has been assimilated as solar PV. Similarly, cogeneration has been assimilated as CCGTs. Potential hydro inflows have not been considered and instead the initial storage values correspond to the usual electricity generation by hydro in Spain in a year. Costs are not intended to be representative but are just used as a tool to establish merit order in the electricity dispatch.

The base electricity demand corresponds to the hourly electricity demand of mainland Spain for 2019. From this base electricity demand, 100 demand scenarios are created by scaling the base demand by values ranging from 0.9 to 1.1, following a random draw. These 100 scenarios are then allocated among the 1000 total scenarios randomly. The solar and wind availabilities were obtained as the Spanish capacity factors (hourly generation divided by the installed capacity of each technology) of each technology between 2007 and 2011. These 5 renewable scenarios were allocated randomly among the 1000 final scenarios.

All in all, these simplifications do not hinder the results of the accurateness of the second-best methodology introduced in section 3.2.3.1, and results should be somewhat similar regardless of the electricity mix and cost structure introduced.

The data considered for the base case execution are presented hereunder, with the exception of an additional 100-MW perfect power plant with an equivalent forced outage rate of 0, no minimum power requirements and zero operating costs.

	Nuclear	Coal	CCGT	PV	Wind	1-Hour Battery	4-Hour Battery	8-Hour Battery	Hydro
No. of units	7	5	30	55	108	1	1	1	17
$\overline{P}_g$ [MW]	1000	1000	1000	250	250	100	100	100	1000
$\underline{P}_g$ [MW]	1000	600	400	0	0	0	0	0	0
C_g^V [€/MWh]	5.27	36	59.5	0	0	0	0	0	0
C_g^{NL} [€]	0	1050	3150	0	0	0	0	0	0
C_g^{SD} [€]	50	2250	3500	750	0	0	0	0	0
C_g^{SU} [€]	50	22500	35000	7500	0	0	0	0	0
$\overline{EFOR}_g$ [p.u.]	0.02	0.15	0.06	0.00	0.00	0.04	0.04	0.04	0.04
$\underline{EFOR}_g$ [p.u.]	0.01	0.05	0.04	0.00	0.00	0.06	0.06	0.06	0.06
MTR_g [h]	20	20	20	20	20	20	20	20	20
$StoIni_g$ [MWh]	0	0	0	0	0	0	0	0	26852017

For each of the marginal executions, a single 100 MW unit of the corresponding technology is added to the system (with the exception of hydro), with the same exact technical characteristics and costs as the ones presented in the previous table.

Formulation

Equation 8 is the only not applied during the execution of the UC model but rather is applied ex-ante to limit the availability of resources in terms of outages and RES-E output.

$$\min \sum_{t\in T}\left[\sum_{g\in G}\left[C_g^{NL} u_{sc,g,t} + C_g^{V}\left(\underline{P}_g u_{sc,g,t} + p_{sc,g,t}\right) + C_g^{SU} v_{g,t} + C_g^{SD} w_{g,t}\right] + C^{NSE} nse_{sc,t}\right] \qquad (1)$$

s.t.

$$\sum_{g\in G}\left[\underline{P}_g u_{g,t} + p_{g,t} - binj_{g,t}\right] = D_t - nse_t \qquad \forall t\in T\ (\lambda_t) \qquad (2)$$

$$p_{g,t} \le \left(\overline{P}_g - \underline{P}_g\right) u_{g,t} \qquad \forall g\in G,\ \forall t\in T\ (\bar{\rho}_{g,t}) \qquad (3)$$

$$u_{g,t} - u_{g,t-1} = v_{g,t} - w_{g,t} \qquad \forall g\in G,\ \forall t\in T\ (\tau_{g,t}) \qquad (4)$$

$$\underline{P}_g u_{g,t} + p_{g,t} - (binj_{g,t} \mathrm{Eff}_g) = bsto_{g,t-1} - bsto_{g,t} \qquad \forall g\in B,\ \forall t\in T\ (\varphi_{g,t}) \qquad (5)$$

$$\underline{P}_g u_{g,t} + p_{g,t} = hsto_{g,t-1} - hsto_{g,t} \qquad \forall g\in H,\ \forall t\in T\ (\omega_{g,t}) \qquad (6)$$

$$bsto_{g,t} \le \mathrm{StoCap}_g N_g \qquad \forall g\in B,\ \forall t\in T\ (\gamma_{g,t}) \qquad (7)$$

$$v_{g,t}, w_{g,t} \le N_g \qquad \forall g\in B,\ \forall t\in T\ (\beta_{g,t}) \qquad (8)$$

$$u_{g,t} \le AV_{sc,g,t} \qquad \forall sc\in SC,\ \forall g\in G,\ \forall t\in T \qquad (9)$$

Annex III

This annex presents the mathematical formulation of the stochastic expansion planning model used for the case studies presented in chapter 4. The model was solved as a RMIP (Relaxed Mixed Integer Program) problem.

Indexes and sets

$g \in G$	Generating technologies
$t \in T$	Hourly periods
$sc \in SC$	Scenarios

Parameters

C_g^V	Variable cost of a unit of technology g [€/MWh]
C_g^{NL}	No-load cost of a unit of technology g [€/MWh]
C^{NSE}	Non-served energy price (in this case 3000 €/MWh) [€/MWh]
$\overline{P}_g$	Maximum power output of a unit of technology g [MW]
$\underline{P}_g$	Minimum power output of a unit of technology g [MW]
$AV_{sc,g,t}$[54]	Availability of technology g available in period t in scenario sc [0-1]
AIC_g	Annualised investment cost of units of technology g [€/MW]
D_t	Demand in period t [MWh]
$EENS_{RS}$	EENS reliability standard [MWh]
$CVaR_{RS}$	CVaR(ENS) reliability standard [MWh]

[54] In order to avoid instability when deciding whether to invest more or less into a particular technology, the availability will be considered homogenous for each technology. This way, investing into an additional unit of a particular technology will yield the same availability as the rest of the units of said technology.

α	Percentile threshold for the CvaR [p.u.]
w	Weighting factor for the composite standard [p.u.]
$EFOR_g$	Equivalent forced outage rate [p.u.]

Variables

n_g	Number of units installed of technology g
$nse_{sc,t}$	Non-served energy in period t and scenario sc [MWh]
$p_{sc,g,t}$	Power output above minimum output of all technology g units in period t and scenario sc [MW]
$u_{sc,g,t}$	Number of units of technology g committed in period t and scenario sc
Ω_{sc}	Auxiliary variable used to characterise the CVaR (amount of non-served energy in scenario sc above the VaR) [MWh]
VaR	Value at risk (if and only if the CVaR constraint is active, else it will not represent the VaR)[55] [MWh]

$nse_{sc,t}, p_{sc,g,t}, VaR, \Omega_{sc} \in \mathbb{R}_{\geq 0}$

$n_g, u_{sc,g,t} \in \mathbb{Z}_{\geq 0}$

Input data

	Nuclear	CCGT	PV	Wind	Diesel
$\overline{P}_g$ [MW]	1000	1000	1000	1000	1000
$\underline{P}_g$ [MW]	1000	400	0	0	200
C_g^V [€/MWh]	6.52	155.6	0	0	710
C_g^{NL} [€/p.u.]	0	7200	0	0	4000
$EFOR_g$ [p.u.]	0.05	0.1	0.00	0.00	0.12
AIC_g [M€]	626.25	78.57	90.10	128.61	50.77

[55] For a detailed explanation on both the auxiliary variable and the Value at Risk in the context of this model the reader can refer to Uryasev (2000).

Formulation

$$\min \sum_{sc}\left[\frac{1}{|SC|}\sum_{t\in T}\left[\sum_{g\in G}\left[C_g^{NL}u_{sc,g,t}+C_g^{V}\left(\underline{P}_g u_{sc,g,t}+p_{sc,g,t}\right)\right]+C^{NSE}nse_{sc,t}\right]\right]+\sum_{g\in G} n_g AIC_g \quad (3)$$

s.t.

$$\sum_{g\in G}\left[\underline{P}_g u_{sc,g,t}+p_{sc,g,t}\right]=D_t-nse_{sc,t} \qquad \forall sc\in SC,\ \forall t\in T\ (\lambda_{sc,t}) \quad (4)$$

$$p_{sc,g,t}\leq\left(\overline{P}_g-\underline{P}_g\right)u_{sc,g,t} \qquad \forall sc\in SC,\ \forall g\in G,\ \forall t\in T\ (\bar{\rho}_{sc,g,t}) \quad (3)$$

$$u_{sc,g,t}\leq AV_{sc,g,t}\ n_g \qquad \forall sc\in SC,\ \forall g\in G,\ \forall t\in T\ (\bar{\mu}_{sc,g,t}) \quad (4)$$

$$\sum_{sc\in SC}\left[\frac{1}{|SC|}\sum_{t\in T}nse_{sc,t}\right]\leq\ EENS_{RS} \qquad (\bar{\varepsilon}) \quad (5)$$

$$VaR+\frac{\sum_{sc\in SC}\left[\frac{\Omega_{sc}}{|SC|}\right]}{\alpha}\leq\ CVaR_{RS} \qquad (\bar{\delta}) \quad (6)$$

$$\Omega_{sc}\geq\sum_{t\in T}\left[nse_{sc,t}\right]-VaR \qquad \forall sc\in SC\ (\tau_{sc}) \quad (7)$$

$$w\frac{\sum_{sc\in SC}\left[\frac{1}{|SC|}\sum_{t\in T}nse_{sc,t}\right]}{EENS_{RS}}+(1-w)\frac{VaR+\frac{\sum_{sc\in SC}\left[\frac{\Omega_{sc}}{|SC|}\right]}{\alpha}}{CVaR_{RS}}\leq 1 \qquad (\bar{\theta}) \quad (8)$$

Models used

Without reliability standards: equations (1)-(4).

With an EENS standard: equations (1)-(5).

With a CVaR(ENS) reliability standard: equations (1)-(4), (6) and (7).

With both an EENS and a CVaR(ENS) reliability standard: equations (1)-(7).

With a composite EENS-CVaR(ENS) reliability standard: equations (1)-(4), (7) and (8).

ANNEX IV

This Annex contains detailed information about the modelling used in chapter 5. The model was solved as a RMIP problem. The data used for these case studies is included at the end of this section.

Indeces and sets

$g \in G$	Generating technologies
$t \in T$	Hourly periods
$sc \in SC$	Scenarios

Parameters

C_g^{V}	Variable cost of a unit of technology g [€/MWh]
C_g^{NL}	No-load cost of a unit of technology g [€]
C_g^{SU}	Start-up cost of a unit of technology g [€]
C_g^{SD}	Shut-down cost of a unit of technology g [€]
C^{NSE}	Non-served energy price (in this case 3000 €/MWh) [€/MWh]
$\overline{P}_g$	Maximum power output of a unit of technology g [MW]
$\underline{P}_g$	Minimum power output of a unit of technology g [MW]
$AV_{sc,g,t}$	Availability of technology g available in period t in scenario sc [p.u.]
D_t	Demand in period t [MWh]
N_g	Number of units installed of technology g [p.u.]
AIC_g	Annualised investment cost of units of technology g [k€/MW]
$\overline{EFOR}_g$	Maximum equivalent forced outage rate of a unit of technology g [p.u.]
$\underline{EFOR}_g$	Minimum equivalent forced outage rate of a unit of technology g [p.u.]
MTR_g	Mean time for recovery for units of technology g [h]

Variables

nse_{t}	Non-served energy in period t [MWh]
$p_{g,t}$	Power output above minimum output of all technology g units in period t [MW]
$u_{g,t}$	Number of units of technology g committed in period t [p.u.]
$v_{g,t}$	Start-up decision of technology g in period t [p.u.]
$w_{g,t}$	Shut-down decision of technology g in period t [p.u.]

$nse_{t}, p_{g,t} \in \mathbb{R}_{\geq 0}$

$u_{g,t}, v_{g,t}, w_{g,t} \in \mathbb{Z}_{\geq 0}$

Input data

The data used in this modelling exercise are not meant to be realistic and may not reflect the reality of some of the technologies included in the mix.

	Nuclear	Coal	CCGT	Fuel oil	PV	New CCGT	New PV
No. of units	25	20	35	10	5	75	65
$\overline{\mathrm{P}}_{g}$ [MW]	500	500	500	500	200	100	100
$\underline{\mathrm{P}}_{g}$ [MW]	500	300	200	200	0	20	0
C_{g}^{V} [€/MWh]	6.5	37.25	60.75	189.5	0	59	0
C_{g}^{NL} [€/MWh]	0	525	3150	6750	0	1575	0
C_{g}^{SD} [€]	50	2250	3500	750	0	700	0
C_{g}^{SU} [€]	50	22500	35000	7500	0	7000	0
$\overline{\mathrm{EFOR}}_{g}$ [p.u.]	0.02	0.15	0.06	0.20	0.00	0.02	0.00
$\underline{\mathrm{EFOR}}_{g}$ [p.u.]	0.01	0.05	0.04	0.10	0.00	0.02	0.00
AIC_{g} [k€/MW]	-	-	-	-	-	75	135
MTR_{g} [h]	10	10	10	10	10	10	10

Formulation of the UC model

$$\min \sum_{t \in T} \left[\sum_{g \in G} \left[\mathrm{C}_{g}^{NL} u_{sc,g,t} + \mathrm{C}_{g}^{V} \left(\underline{\mathrm{P}}_{g} u_{sc,g,t} + p_{sc,g,t} \right) + \mathrm{C}_{g}^{SU} v_{g,t} + \mathrm{C}_{g}^{SD} w_{g,t} \right] + \mathrm{C}^{NSE} nse_{sc,t} \right] \quad (6) \quad (5)$$

s.t.

$$\sum_{g \in G} \left[\underline{\mathrm{P}}_{g} u_{g,t} + p_{g,t} - binj_{g,t} \right] = \mathrm{D}_{t} - nse_{t} \qquad \forall t \in T \qquad (7)$$

$$p_{g,t} \leq \left(\overline{\mathrm{P}}_{g} - \underline{\mathrm{P}}_{g} \right) u_{g,t} \qquad \forall g \in G, \forall t \in T \qquad (3)$$

$$u_{g,t}-u_{g,t-1}=v_{g,t}-w_{g,t} \qquad \forall g\in G,\ \forall t\in T \qquad (4)$$

$$v_{g,t},w_{g,t}\leq N_g \qquad \forall g\in G,\ \forall t\in T \qquad (5)$$

$$u_{g,t}\leq AV_{sc,g,t} \qquad \forall sc\in SC,\ \forall g\in G,\ \forall t\in T \qquad (6)$$

ANNEX V

This section demonstrates the link between the reliability metric, the computation of the firm supply and how this firm supply should be remunerated. Additionally, this section demonstrates how to determine how much firm supply each consumer is demanding and how much it should pay for it. This section is based on Annex I, in which a stylised version of the ideal central planner problem is solved and compared to the individual agent problem. Comparing both optimality conditions allows us to draw several important conclusions.

The main difference between what was presented in Annex I and the developments presented in this section is that the reliability target set in the centralised problem is presented as a function of generation and demand, not only as a function of the installed capacity (since the focus is now expanded to the demand side and, therefore, the hourly consumption becomes a variable). Furthermore, the time dimension is considered explicitly in this formulation. Finally, the problem from the perspective of an individual consumer is also presented.

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:

$$\underset{D\ Q\ K\ NSE}{\mathrm{Max}}\ NSB = U(D) - C(Q) - I(K) - NSEC(NSE) \tag{1}$$

Where:

- *D* represents the full set of elements $\{D_{j,h}\}$, which in turn correspond to the hourly, h=1,2,3,4,…p, demand of each consumer, j= 1,2,3,4,…m.
- *U(D)* is the demand utility function, which depends on the hourly power consumed by the demand. This demand utility function is assumed to be strictly increasing and concave.
- *Q*, represents the full set of elements $\{Q_{j,h}\}$, which in turn correspond to the hourly production of each generating unit, i= 1,2,3,4,…n.

- *C(Q)* is the generation cost function, which depends on the hourly generation of each of the generation units. In this case, the function is assumed to be strictly increasing and convex.
- *K* represents the full set of elements $\{K_i\}$, which in turn correspond to the installed capacity of each generating plant, i= 1,2,3,4,…n.
- *I(K)* is the investment cost function which depends on the installed capacity of each generation unit. In this case, the function is assumed to be strictly increasing and convex.
- *NSE* represents the full set of elements $\{NSE_h\}$, which in turn correspond to the non-served demand of each hour, h=1,2,3,4,…p.
- *NSEC(NSE)* represents the hourly cost associated with the hourly non-served demand due to limited generation capacity. In this case, the function is assumed to be strictly increasing and convex.

This stylised representation only considers three constraints:

$$\sum_{j=1}^{m} D_{j,h} \leq \sum_{i=1}^{n} Q_{i,h} + NSE_h \qquad \forall h \in H \ \ (\pi_h) \qquad (2)$$

$$Q_{i,h} \leq K_i \qquad \forall h \in H, \forall i \in N \ \ (\alpha_{i,h}) \qquad (3)$$

$$RM(K,D) \geq RT \qquad (\beta) \qquad (4)$$

The first constraint represents the balance between the hourly demand and generation, with the inclusion of non-served energy allowing for some part of the demand to not be satisfied if this was economically efficient.

The second constraint represents the upper limit of the power produced by each generation unit *i* in hour *h*, which corresponds to the installed capacity of that unit, K_i.

On the other hand, the third constraint forces the reliability metric, *RM*, to fulfil a certain reliability target, *RT*, which is set as a parameter. The *RM* is assumed to be strictly decreasing and convex. In Annex I of chapter 3, this reliability metric was assumed to be dependent only on the electricity mix, $RM(K_1, K_2, K_3,\dots K_n)$. In this case, this formulation is expanded and is dependant also on the system demand, which is coherent because the reliability of the system will depend on its installed capacity in relation to the electricity demand in the system.

Other operation constraints such as ramps, minimum power outputs, etc., have been discarded for simplicity.

In order to obtain the first-order necessary conditions the Lagrangian function, *L*, is formulated and its first partial derivatives with respect to the decision variables is expressed.

$$L(Q,D,K,\pi_h,\alpha_{i,h},\beta) = U(D) - C(Q) - I(K) - NSEC(NSE) + \sum_{h=1}^{H}\left[\left(\sum_{j=1}^{m} D_{j,h} - \sum_{i=1}^{n} Q_{i,h} - NSE_h\right)\pi_h + \sum_{1}^{n}\left(\left(Q_{i,h}-K_i\right)\cdot\alpha_{i,h}\right)\right] + (RM(K,D)-RT)\cdot\beta \quad (5)$$

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

$$\frac{\partial L}{\partial K_i} = -\sum_{h=1}^{H}\alpha_{i,h} - \frac{\partial I(K)}{\partial K_i} + \frac{\partial RM(K,D)}{\partial K_i}\beta = 0 \quad (6)$$

If the same is done with respect to the demand of consumer *j*, $D_{j,h}$, the following expression can be obtained:

$$\frac{\partial L}{\partial D_{j,h}} = \pi_h + \frac{\partial U(D)}{\partial D_{j,h}} - \frac{\partial RM(K,D)}{\partial D_{j,h}}\beta = 0 \quad (7)$$

The difference in signs of the partial derivative of the reliability metric is because generation and demand have opposing effects on the reliability metric. An increase in generation capacity would improve the reliability metric, but an increase in demand would worsen the reliability metric.

Decentralised problem

Generator 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:

$$\max_{Q_{i,h}\, K_i} P_i = \sum_{h=1}^{H}\left[SMP_{Q,i,h}\cdot Q_{i,h}\right] - C_i\left(Q_i\right) + CMP_{K,i}\cdot K_i - I_i(K_i) \quad (8)$$

Where:

- $SMP_{Q,i,h}$ is the hourly spot market price perceived by generation unit *i* in hour *h* which when multiplied by the hourly generation, $Q_{i,h}$ results in the spot market revenues.
- $CMP_{K,i}$ is the capacity market price profile perceived by generation unit *i*, which when multiplied by K_i results in the capacity market revenues.
- Q_i represents the full set of hourly generation values of generation unit *i*.

- $C_i(Q_i)$ 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:

$$Q_{i,h} \leq K_i \qquad \forall h \in H \ (\alpha_{i,h}) \tag{9}$$

This constraint is equivalent to the second constraint in the centralised problem. The first and third constraints found in the centralised problem are only present through the regulator's perspective and are therefore only translated through the spot and capacity market prices, respectively, in the objective function in this decentralised problem.

In order to obtain the first-order necessary conditions the Lagrangian function of this second problem, *L*, is formulated and its first partial derivatives with respect to the decision variables is computed.

$$L(Q_i,K_i,\alpha_{i,h}) = \sum_{h=1}^{H}\left[SMP_{Q,i,h}\cdot Q_{i,h}+\left(Q_{i,h}-K_i\right)\cdot\alpha_{i,h}\right]-C_i\left(Q_i\right)+CMP_{K,i}\cdot K_i - I_i(K_i) \tag{10}$$

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

$$\frac{\partial L}{\partial K_i} = CMP_{Ki} - \frac{\partial I(K)}{\partial K_i} - \sum_{h=1}^{H}\alpha_{i,h} = 0 \tag{11}$$

Equation 11 is very similar to equation 6, without the global constraint described by equation 4, which is only present in the centralised problem, but with the additional term $CMP_{K,i}$.

Unification of the centralised problem and the generator problem

Comparing equation 11 to equation 6 the following expression is obtained:

$$CMP_{K,i} = \frac{\partial RM(K,D)}{\partial K_i}\cdot\beta \tag{12}$$

This allows us to draw the conclusion that firm supply depends on the marginal contribution to the reliability metric *RM*, which is the same conclusion presented in chapter 3.

Consumer problem

The consumer problem is very similar to that of the generator, although it does not consider any investment and generation costs. The objective of each consumer, *j*, is to maximise their personal net benefit, PNB_j, considering their utility function, and both the spot market and capacity market prices:

$$\underset{D_{j,h}}{\text{Max}}\ PNB_j = U_j\left(D_j\right)-\sum_{h=1}^{H}\left[SMP_{D,j,h}\cdot D_{j,h} + CMP_{D,j,h}\cdot D_{j,h}\right] \tag{13}$$

Where:

- D_j represents the full set of hourly generation values of consumer *j*.
- $U_j(D_j)$ represents the utility that consumer *j* obtains by the use of electricity.
- $SMP_{D,j,h}$ is the hourly spot market price perceived by demand *j*, which when multiplied by the hourly demand, $D_{j,h}$, results in the spot market disbursements.
- $CMP_{D,j,h}$ is the capacity market price perceived by consumer *j*, which results in the capacity market disbursements when multiplied by $D_{j,h}$. In CRMs, these disbursements would be allocated as a lump sum either ex-ante, based on estimations of consumer demand, or ex-post, based on actual consumption. In this case, no constraints are considered.

In order to obtain the first-order necessary conditions the Lagrangian function of this problem, *L*, is formulated and its first partial derivatives with respect to the decision variables is computed.

$$L(D_{j,h}) = U_j\left(D_j\right) - \sum_{h=1}^{H}\left[SMP_{D,j,h}\cdot D_{j,h} + CMP_{D,j,h}\cdot D_{j,h}\right] \tag{14}$$

When computing the first partial derivative of this expression with respect to the decision variable $D_{j,h}$ the following optimality condition is obtained:

$$\frac{\partial L}{\partial D_{j,h}} = \frac{\partial U_j\left(D_j\right)}{\partial D_{j,h}} - CMP_{D,j,h} - SMP_{D,j,h} = 0 \tag{15}$$

Unification of the centralised problem and the consumer problem

Comparing equation 15 to equation 7 equation 16 is obtained, given that the variation in the utility of consumer *j* due to a marginal variation in its consumption is equivalent to the variation in the utility of demand as a whole (only the utility of consumer *j* will be affected):

$$CMP_{D,j,h} + SMP_{D,j,h} = \frac{\partial RM(K,D)}{\partial D_{j,h}}\cdot\beta - \pi_h \tag{16}$$

If the focus is placed on the capacity market disbursements of consumer *i*:

$$CMP_{D,j,h}\cdot D_{j,h} = \frac{\partial RM(K,D)}{\partial D_{j,h}}\cdot D_{j,h}\cdot\beta \tag{17}$$

This allows us to draw several conclusions:

1. The firm supply demanded by consumer *j* depends on its consumption during scarcity conditions.
2. Accordingly, cost allocation to consumer *j* will also depend on its consumption during instances in which a marginal variation of demand would cause a variation in the reliability metric.

BIBLIOGRAPHY

ACER, Agency for the Cooperation of Energy Regulators, 2025. Security of EU electricity supply. 2025 Monitoring report.

ACER, Agency for the Cooperation of Energy Regulators, 2024a. Progress of EU electricity wholesale market integration: 2024 Market Monitoring Report.

ACER, Agency for the Cooperation of Energy Regulators, 2024b. Security of EU electricity supply: 2024 Monitoring Report.

ACER, Agency for the Cooperation of Energy Regulators, 2024c. DECISION No 06/2024 OF THE EUROPEAN UNION AGENCY FOR THE COOPERATION OF ENERGY REGULATORS of 2 May 2024 on the European Resource Adequacy Assessment for 2023.

ACER, Agency for the Cooperation of Energy Regulators, 2023a. DECISION No 04/2023 OF THE EUROPEAN UNION AGENCY FOR THE COOPERATION OF ENERGY REGULATORS of 27 February 2023 on the European Resource Adequacy Assessment for 2022.

ACER, Agency for the Cooperation of Energy Regulators, 2023b. Security of EU electricity supply: 2023 Monitoring Report.

ACER, Agency for the Cooperation of Energy Regulators, 2022a. DECISION No 02/2022 OF THE EUROPEAN UNION AGENCY FOR THE COOPERATION OF ENERGY REGULATORS of February 2022 on the European Resource Adequacy Assessment for 2021.

ACER, Agency for the Cooperation of Energy Regulators, 2022b. Security of EU electricity supply in 2021: Report on Member States approaches to assess and ensure adequacy.

ACER, Agency for the Cooperation of Energy Regulators, 2020a. Methodology for the European resource adequacy assessment in accordance with Article 23 of Regulation (EU) 2019/943 of the European Parliament and of the Council of 5 June 2019 on the internal market for electricity. ACER Decision on the ERAA methodology: Annex 1.

ACER, Agency for the Cooperation of Energy Regulators, 2020b. 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.

AEEGSI, Autorità per l'Energia Elettrica il Gas e il Sistema Idrico. 2017. Mercato italiano della capacità - ultimi parametri tecnico-economici. Consultation document 592/2017/R/EEL.

AEMC, Australian Energy Market Commission, Reliability Panel, 2022. 2022 Review of the Reliability Standard and Settings. Report.

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

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

ARERA, Autorità di Regolazione per Energia Reti e Ambiente, 2024. Criteri e condizioni per la disciplina del sistema di remunerazione della disponibilità di capacità produttiva di energia elettrica, ai sensi dell'articolo 2 del decreto legislativo 19 dicembre 2003, n. 379.

Awara, S., Murphy, C., Schleifer, A., Stephen, G., Denholm, P., 2023. Impacts of hybridization and forecast errors on the probabilistic capacity credit of batteries. Next Energy, Volume 1, Issue 2, Article 100021.

Bastos, J. P., Cunha, G., Barroso, L. A., Aquino, T., 2018. Reliability mechanism design: An economic approach to enhance adequate remuneration and enable efficient expansion. Energy, Volume 158, Pages 1150-1159.

Batlle, C., Mastropietro, P., Rodilla, P., Pérez-Arriaga I.J., 2022. The System Adequacy Problem: Lessons Learned from Both Sides of the Atlantic. In: Capacity Mechanisms in the EU Energy Markets: Law, Policy, and Economics - Second Edition. Oxford University Press. ISBN 978-0-19-874925-7.

Batlle, C., Rodilla, P., Mastropietro, P., 2021. Markets for Efficient Decarbonization: Revisiting Market Regulation and Design. IEEE Power and Energy Magazine, Volume 19, Issue 1, Pages 20-28, Jan.-Feb. 2021.

Batlle, C., Rodilla, P., 2010. A critical assessment of the different approaches aimed to secure electricity generation supply. Energy Policy, Volume 38, Issue 11, Pages 7169-7179.

Batlle, C., and Pérez-Arriaga, I. J., 2008. Design criteria for implementing a capacity mechanism in deregulated electricity markets. Utilities Policy, Volume 16, Issue 3, Pages 184-193.

Batlle, C., Vázquez, C., Rivier, M. & Pérez-Arriaga, I. J., 2007. Enhancing power supply adequacy in Spain: Migrating from capacity payments reliability options. Energy Policy, Volume 35, Issue 9, Pages 4545-4554.

Béres, R., Van der Wel, A., Fattahi, A., Van der Broek, M., 2024. The impact of national policies on Europe-wide power system transition towards net-zero 2050. Energy, Volume 310, Article 133216.

Bidwell, M., 2005. Reliability Options: A Market-Oriented Approach to Long-Term Adequacy. The Electricity Journal, Volume 18, Issue 5, Pages 11-25.

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

BMWK, Federal Ministry for Economic Affairs and Climate Action, 2024. Electricity market design of the future: Options for a secure, affordable and sustainable electricity system.

BOG, Belgian Official Gazette, 2021. Program Law 2021-12-27/01.

Borenstein, S., 2002. The Trouble With Electricity Markets: Understanding California's Restructuring Disaster. Journal of Economic Perspectives, Volume 16, Number 1, Pages 191-211.

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.

Brito-Pereira, P., Bruninx, K., De Vries, L., Mastropietro, P., Rodilla, P. Future-proofed resource adequacy metrics: a model-based assessment of multi-metric vs. composite-metric reliability standards. Sustainable Energy, Grids and Networks, Volume 44, Pages 101957-1 - 101957-15.

Brito-Pereira, P., Rodilla, P., Mastropietro, P., Batlle, C., 2022a. Self-fulfilling or self-destroying prophecy? The relevance of de-rating factors in modern capacity mechanisms. Applied Energy, Volume 314, Article 118939.

Brito-Pereira, P., Mastropietro, P., Rodilla, P., Barroso, L. A., Batlle, C., 2022b. Adjusting the aim of capacity mechanisms: Future-proof reliability metrics and firm supply calculations. Energy Policy, Volume 164, Article 112891.

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, Volume 38, Pages173 188.

Busby, J. B., Baker, K., Bazilian, M. D., Gilbert, A. Q., Grubert, E., Rai, V., Rhodes, J. D., Shidore, S., Smith, C. A., Webber, M. E., 2021. Cascading Risks: Understanding the 2021 Winter Blackout in Texas. Energy Research & Social Science, Volume 77, Article 102106.

Bushnell, J., Flagg, M., Mansur, E., 2017. Capacity Markets at a Crossroads. Energy Institute at Haas, Working Paper 278.

Byers, C., Levin, T., Botterud, A., 2018. Capacity market design and renewable energy: Performance incentives, qualifying capacity, and demand curves. The Electricity Journal, Volume 31, Issue 1, Pages 65-74.

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

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

Caramanis, M., 1982. Investment Decisions and Long-Term Planning under Electricity Spot Pricing. IEEE Transaction on Power Apparatus and Systems, Volume PAS 101, Issue 12, Pages 4640-4648.

Caramanis, M. C., Bohn R. E., Schweppe, F. C., 1982. Optimal Spot Pricing: Practice and Theory. IEEE Transactions on Power Apparatus and Systems, Volume PAS 101, Issue 9, Pages 3234-3245.

CNMC, Comisión Nacional de los Mercados y la Competencia, 2020. Circular 3/2020, de 15 de enero, de la Comisión Nacional de los Mercados y la Competencia, por la que se establece la metodología para el cálculo de los peajes de transporte y distribución de electricidad.

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.

CPUC, California Public Utilities Commission, 2009. Baselines for Retail Demand Response Programs. Technical presentation.

Cramton, P., Ockenfels, A., Stoft, S., 2013. Capacity Market Fundamentals. Economics of Energy & Environmental Policy, Volume 2, Issue 2, Pages 27-46.

Cramton, P., Stoft., S., 2006. The Convergence of Market Designs for Adequate Generating Capacity with Special Attention to the CAISO's Resource Adequacy Problem. Center for Energy and Environmental Policy Research, 06-007 WP.

Cramton, P., Stoft, S., 2005. A Capacity Market that Makes Sense. The Electricity Journal, Volume 18, Issue 7, Pages 43-54.

Creti, A., Fabra, N., 2007. Supply security and short-run capacity markets for electricity. Energy Economics, Volume 29, Issue 2, Pages 259-276.

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.

Dent, C. J., Sanchez, N., Shevni, A., Smith, J. Q., Wilson, A. L., Yu, X., 2024. Resource adequacy and capacity Procurement: Metrics and decision support analysis. Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy. Volume 239, Issue 1, Pages 199-209.

DESNZ, Department for Energy Security & Net Zero, 2023. Exploring Reliability Standard Metrics in a Net Zero Transition. DESNZ research paper No. 2023/050.

De Vries, L., Sánchez Jiménez, I., 2022. Market Signals as Adequacy Indicators for Future Flexible Power Systems. Oxford Open Energy, Volume 1, Pages 1-5.

De Vries, L., Heijnen, P., 2008. The impact of electricity market design upon investment under uncertainty: The effectiveness of capacity mechanisms. Utilities Policy, Volume 16, Issue 3, Pages 215-227.

EC, European Commission, 2024. Regulation (EU) 2024/1747 of the European Parliament and of the Council of 13 June 2024 amending regulations (EU) 2019/942 and (EU) 2019/943 as regards improving the Union's electricity market design.

EC, European Commission, 2021. Commission Decision of 27.8.2021 on THE AID SCHEME SA.54915 – 2020/C (ex 2019/N) Belgium – Capacity remuneration mechanism.

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, 2018. State Aid SA.42011 (2017/N) – Italy – Italian Capacity Mechanism.

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.

EC, European Commission, 2016c. Commission Decision of 8.11.2016 ON STATE AID SCHEME SA.39621 2015/C (ex 2015/NN).

EIA, U.S. Energy Information Administration, 2019. Assumptions to the Annual Energy Outlook 2019. Electricity Market Module.

EIRGRID, 2016. I-SEM Capacity Market: Methodology for the Calculation of the Capacity Requirement and De-rating Factors. SEM-16-082a. Appendix 1.

Elia, 2024a. Adequacy and Flexibility Study for Belgium: 2024-2034.

Elia, 2024b. Capacity Remuneration Mechanism (CRM) functioning rules (Version 4): Proposal from Elia.

Elia, 2024c. Preparation of the CRM Y-1, Y-2 and Y-4 auctions with Delivery Periods 2026-2027, 2027-2028 and 2029-2030: Report of the transmission system operator containing the information to determine the volume to be contracted and proposals for other parameters.

Elia, 2021. Baseline Methodology Assessment. Final report.

Elia, 2019a. CRM Design Note: Derating factors.

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

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

Elkouk, A., Pokhrel, Y., Satoh, Y., Bouchaou, L., 2022. Implications of changes in climate and human development on 21st-century global drought risk. Journal of Environmental Management, Volume 317, Article 115378.

EMRS, Electricity Market Reform Settlement Limited, 2024a. G12 - CM Supplier Capacity Market Demand Forecast, EMRS Guidance.

EMRS, Electricity Market Reform Settlement Limited, 2024b. G15 - CM Supplier Payments, EMRS Guidance.

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

Ensslin, C., Milligan, M., Holttinen, H. O'Malley, M., Keane, A., 2008. Current Methods to Calculate Capacity Credit of Wind Power, IEA Collaboration. 2008 IEEE Power and Energy Society General Meeting – Conversion and Delivery of Electrical Energy in the 21st Century, 20-24 July 2008, Pittsburgh, PA, USA.

ENTSO-E, European Network of Transmission System Operators for Electricity, 2025. European Resource Adequacy Assessment 2024 Edition: ACER's approved and amended version (August 2025). Executive Report.

ENTSO-E, European Network of Transmission System Operators for Electricity, 2024. European Resource Adequacy Assessment 2023 Edition: ACER's approved and amended version. Executive Report.

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

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

EPE, Empresa de Pesquisa Energética, 2020a. Relatório Final do Grupo Temático: Critérios de Garantia de Suprimento, Portaria No. 187/2019 (in Portuguese).

EPE, Empresa de Pesquisa Energética, 2020b. Critério de Garantia de Suprimento (in Portuguese). Infographic. https://shorturl.at/oKGrn. Accessed on 07/04/2025.

EPRI, Electric Power Research Institute, 2024. Metrics and Criteria Insights from Case Studies and Recommendations and Considerations for Future Practice. Technical Report 3002030638. Principal investigators: G. de Mijolla, A. Tuohy.

EPRI, Electric Power Research Institute, 2022. Resource Adequacy for a Decarbonized Future: A Summary of Existing and Proposed Resource Adequacy Metrics. Technical report.

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

ERAWA, Economic Regulation Authority Western Australia, 2024. Procedure Change Proposal: Benchmark Reserve Capacity Price. Working document EEPC_2024_01.

ERCOT, Electric Reliability Council of Texas, 2021. Review of February 2021 Extreme Cold Weather Event – ERCOT Presentation. Technical presentation.

ESIG, Energy Systems Integration Group, 2025. Gaps, Barriers, and Solutions to Demand Response Participation in Wholesale Markets. Technical Report.

ESIG, Energy Systems Integration Group, 2024. New Resource Adequacy Criteria for the Energy Transition - Modernizing Reliability Requirements. Report by the ESIG Resource Adequacy Task Force.

ESIG, Energy Systems Integration Group, 2023. Ensuring Efficient Reliability - New Design Principles for Capacity Accreditation. Report of the Redefining Resource Adequacy Task Force.

ESIG, Energy Systems Integration Group , 2021. Redefining Resource Adequacy for Modern Power Systems: A Report of the Redefining Resource Adequacy Task Force.

Estache, A., and Rodriguez-Pardina, M., 1996. Regulatory Lessons from Argentina's Power Concessions. Public Policy for the Private sector. The World Bank Group, Note No. 92.

Fabra, N., 2018. A primer on capacity mechanisms. Energy Economics, Volume 75, Pages 323-335.

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, Volume 24, Issue 2, Pages 541-551.

Finon, D., Meunier, G., Pignon, V., 2008. The social efficiency of long-term capacity reserve mechanisms. Utilities Policy, Volume 16, Issue 3, Pages 202-214.

Finon, D., Pignon, V., 2008. Electricity and long-term capacity adequacy: The quest for regulatory mechanism compatible with electricity market. Utilities Policy, Volume 16, Issue 3, Pages 143-158.

Fraunholz, C., Miskiw, K. K., Kraft, E., Fichtner, W., Weber, C., 2023. On the Role of Risk Aversion and Market Design in Capacity Expansion Planning. The Energy Journal, Volume 44, Issue 3, Pages 111-138.

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

Freire-Barceló, T., Martín-Martínez, F., Sánchez-Miralles, Á., Rivier, M., Gómez, T., Huclin, S., Chaves, J. P., Ramos, A., 2022. Storage and demand response contribution to firm capacity: Analysis of the Spanish electricity system. Energy Reports, Volume 8, Pages 10546-10560.

Garver, L., 1966. Effective Load Carrying Capability of Generation Units. IEEE Transactions on Power Apparatus and Systems, Volume 85, Issue 8, Pages 910 919.

Goldberg, M. L., Kennedy, G., 2013. Measurement and Verification for Demand Response. Prepared for the National Forum on the National Action Plan on Demand Response: Measurement and Verification Working Group

Gollmer, R., Nowak, M., Römisch, W., & Schultz, R., 2000. Unit commitment in power generation - A basic model and some extensions. Annals of Operations Research, Volume 96, Pages 167-189.

Gorman, W. The quest to quantify the value of lost load: A critical review of the economics of power outages. The Electricity Journal, Volume 35, Issue 8, Article 107187.

Herrero, I., Rodilla, P., Batlle, C., 2015. Electricity market-clearing prices and investment incentives: The role of pricing rules. Energy Economics, Volume 47, Pages 42-51.

Hogan, W., 2013. Electricity Scarcity Pricing Through Operating Reserves. Economics of Energy & Environmental Policy, Volume 2, Issue 2, Pages 65-86.

Hogan, W., 2005. On an "Energy Only" Electricity Market Design for Resource Adequacy.

Hohl, C., Lo Prete, C., 2025. Capacity markets vs "energy only" markets with improved scarcity pricing under increased wind penetration. Renewable Energy Focus, Volume 55, Article 100738.

Hormaza Mejia, A., Brouwer, J., Copp, D. A., 2023. Performance and dynamics of California offshore wind alongside Western US onshore wind and solar power. Renewable Energy Focus, Volume 47, Article 100490.

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.

IPCC, Intergovernmental Panel on Climate Change, 2023. Climate Change 2023: Synthesis Report. Contribution of Working Groups I,II, and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.

Isaza Cuervo, F., Botero Botero, S., 2016. Wind power reliability valuation in a Hydro-Dominated power market: The Colombian case. Renewable and Sustainable Energy Reviews, Volume 57, Pages 1359-1372.

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

Joskow, P. L., 2008. Capacity payments in imperfect electricity markets: Need and design. Utilities Policy, Issue 16, Pages 159-170.

Joskoy, P. L., Tirole, J., 2007. Reliability and competitive electricity markets. RAND Journal of Economics, Volume 38, Number 1, Pages 60-84.

Joskow, P.L., 2006. Competitive Electricity Markets and Investments in New Generating Capacity, Center for Energy and Environmental Policy Research, 06-009 WP.

Keppler, J. H., 2017. Rationales for capacity remuneration mechanisms: Security of supply externalities and asymmetric investment incentives. Energy Policy, Volume 105, Pages 562-570.

Komorowska, A., Kaszynski, P., Kaminski, J., 2023. Where does the capacity market money go? Lessons learned from Poland. Energy Policy, Volume 173, Article 113419.

Kozlova, M., Huhta, K., Lohrmann, A., 2023. The interface between support schemes for renewable energy and security of supply: Reviewing capacity mechanisms and support schemes for renewable energy in Europe. Energy Policy, Volume 181, Article 113707.

Kozlova, M., Overland, I., 2022. Combining capacity mechanisms and renewable energy support: A review of the international experience. Renewable and Sustainable Energy Reviews, Volume 155, Article 111878.

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

Larsen, P. H., Lawson, M., LaCommare, K. H., Eto, J. H., 2020. Severe weather, utility spending, and the long-term reliability of the U.S. power system. Energy, Volume 198, Article 117387.

Lazard, 2023. Lazard's Levelized Cost of Energy Analysis - Version 16.0.

Lisnianski, A., Elmakias, D., Laredo, D., Ben Haim, H., 2012. A Multi-state Markov Model for a Short-term Reliability Analysis of a Power Generating Unit. Reliability Engineering & System Safety, Volume 98, Issue 1, Pages 1-6.

Lu, Z., Li, H., Qiao, Y., 2018. Probabilistic Flexibility Evaluation for Power System Planning Considering Its Association With Renewable Power Curtailment. IEEE Transactions on Power Systems, Volume 33, Number 3, Pages 3285-3295.

Lynch, M. A., Nolan, S., Devine, M. T., O'Malley, M., 2019. The impacts of demand response participation in capacity markets. Applied Energy, Volume 250, Pages 444-451.

Ma, J., Liu, Y., Zhang, S., Wu, L., Yang, Z., 2023. Comprehensive probabilistic assessment on capacity adequacy and flexibility of generation resources. International Journal of Electrical Power and Energy Systems, Volume 145, Article 108677.

Machado, R. H. S., Rego, E. E., Udaeta, M. E. M., Nascimento, V. T., 2022. Estimating the adequacy revenue considering long-term reliability in a renewable power system. Energy, Volume 243, Article 123022.

Mancarella, P., 2022. Considerations in Support of the 2022 Reliability Standards and Settings Review. Briefing Note for the Australian Energy Market Commission.

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, Volume 139, Article 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, Volume 112, Pages 253-262.

Mastropietro, P., Fontini, F., Rodilla, P., Batlle, C., 2018. The Italian capacity remuneration mechanism: Critical review and open questions. Energy Policy, Issue 123, Pages 659-669.

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

Mastropietro, P., Herrero, I., Rodilla, P., Batlle, C., 2016. A Model-Based Analysis on the Impact of Explicit Penalty Schemes in Capacity Mechanisms. Applied Energy, Volume 168, Pages 406-417.

Mastropietro, P., Batlle, C., Barroso, L.A., Rodilla, P., 2014. Electricity auctions in South America: Towards convergence of system adequacy and RES-E support. Renewable and Sustainable Energy Reviews, Issue 40, Pages 375-385.

Maurer, L.T.A., and Barroso, L.A., 2011. Electricity Auctions: An Overview of Efficient Practices. The World Bank.

Mertens, T., Bruninx, K., Duerinck, J., Delarue, E., 2021. Capacity credit of storage in long-term planning models and capacity markets. Electric Power Systems Research, Volume 194, Article 107070.

Mills, A. D., Rodriguez, P., 2020. A simple and fast algorithm for estimating the capacity credit of solar and storage. Energy, Volume 210, Article 118587.

MISO, Midcontinent Independent System Operator, 2024. Exploration of Alternative Resource Adequacy Criteria. Resource Adequacy Subcommittee (RASC-2024-4).

MITERD, Ministerio para la Transición Ecológica y el Reto Demográfico, 2024. Proyecto de Orden por la que se crea un mercado de capacidad en el sistema eléctrico peninsular español. Working document. Can be found at https://shorturl.at/xdbgV. Accessed on the 27/03/2025. In Spanish.

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.

Monitoring Analytics, 2023. PJM State of the Market – 2022. Technical report.

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, Volume 38, Issue 10, Pages 5758-5769.

Muaddi, S., Singh, C., 2022. Investigating capacity credit sensitivity to reliability metrics and computational methodologies. Applied Energy, Volume 325, Article 119825.

National Grid ESO, National Grid Electricity System Operator, 2023. Electricity Capacity Report.

National Grid, 2019a. De-rating Factor Methodology for Renewables Participation in the Capacity Market: Consultation Response Summary.

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

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

National Grid, 2015. National Grid EMR Electricity Capacity Report June 2015.

NERC, North American Electric Reliability Corporation, 2024. Polar Vortex Review. Report.

NERC, North American Electric Reliability Corporation, 2021. 2021 Long-Term Reliability Assessment.

NERC, North American Electric Reliability Corporation, 2018a. ERO Reliability Assessment Process Document.

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

NERC, North American Electric Reliability Corporation, 2016. Probabilistic Assessment: Technical Guideline Document.

Neuhoff, K., de Vries, L., 2004. Insufficient incentives for investment in electricity generations. Utilities Policy, Issue 12, Pages 253-267.

Newbery, D., 2016. Missing money and missing markets: Reliability, capacity auctions and interconnectors. Energy Policy, Issue 94, Pages 401-410.

Nexant, 2017. California ISO Baseline Accuracy Work Group Proposal. Working document.

Nolan, S., Neu, O., O'Malley, M., 2017. Capacity Value Estimation of a Load-Shifting Resource Using a Coupled Building and Power System Model. Applied Energy, Volume 192, Pages 71-82.

Nolan, S., O'Malley, M., Hummon, M., Kiliccote, S., Ma, O., 2014. A methodology for estimating the capacity value of demand response. 2014 IEEE PES General Meeting | Conference & Exposition, National Harbor, MD, USA, Pages 1-5.

NPCC, Northwest Power and Conservation Council, 2019. Pacific Northwest Power Supply Adequacy Assessment for 2024.

NREL, National Renewable Energy Laboratory, 2018. U.S. Solar Photovoltaic System Cost Benchmark: Q1 2018.

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 07/04/2025.

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

ODI, Overseas Development Institute, 2016. Rethinking Power Markets: Capacity Mechanisms and Decarbonisation. Report elaborated by Laurie van der Burg and Shelagh Whitley.

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

Oren, S., 2005. Generation Adequacy via Call Options Obligations: Safe Passage to the Promised Land. The Electricity Journal. Volume 18, Issue 9, Pages 28-42.

Peluchon, B., 2019. Market design and the cost of capital for generation capacity investment. Working paper #41.

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

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

Petitet, M., Unel, B., Felder, F. A., 2024. Making Electricity Capacity Markets Resilient to Extreme Weather Events. Economics of Energy & Environmental Policy, Volume 12, Number 2.

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, Volume 103, Pages 30-46.

Pfeifenberger, J. P., Spees, K., Carden, K., Wintermantel, N., 2013. Resource Adequacy Requirements: Reliability and Economic Implications. The Brattle Group and Astrape Consulting technical report for the Federal Energy Regulatory Commission (FERC).

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

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

Pototschnig, A., Godfried, M., 2013. Capacity mechanisms and the EU internal electricity market: The regulator's view: ACER's report on capacity mechanisms. Technical report.

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., Mastropietro, P., Brito-Pereira, P., 2023. The Challenge of Integrating Demand Response in Capacity Remuneration Mechanisms: Providing a Comprehensive Theoretical Framework. IEEE Power and Energy Magazine, Volume 21, Number 4, Pages 67-71.

Rodilla, P., Mastropietro, P., Gómez, T., Agurto, R., Skerk, C., Batlle, C., 2018. Developing the Design of the Chilean Power Market to Allow for Efficient Deployment of Renewables. Oxford Energy Forum, Issue 114, Pages 19-21.

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, Volume 51, Pages 365-382.

Rodilla, P., Batlle, C., 2012. Security of supply at the generation level: Problem analysis. Energy Policy, Volume 40, Pages 177-185.

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

Saldarriaga-C., C. A., Salazar, H., 2016. Security of the Colombian energy supply: the need for liquefied natural gas regasification terminals for power and natural gas sectors. Energy, Volume 100, Pages 349-362.

Santos Borba, P. C., Sousa Júnior, W. C., Shadman M., Pfenninger, S., 2023. Enhancing drought resilience and energy security through complementing hydro by offshore wind power – The case of Brazil. Energy Conversion and Management, Volume 277, Article 116616.

Schweppe, F. C, Caramanis, M. C., Tabors, R. D., Bohn, R. E., 1988. Spot Pricing of Electricity. The Kluwer International Series in Engineering and Computer Science, 1988, Springer. ISBN 978-1-4612-8950-0.

SEMC, Single Electricity Market Committee, 2024. SEM Tariffs and Charges to apply from 1 October 2024 – 30 September 2025. Decision Paper SEM-24-056

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.

Sengupta, M., Xie, Y., Lopez, A., Habte, A., Maclaurin, G., & Shelby, J., 2018. The National Solar Radiation Data Base (NSRDB). Renewable and Sustainable Energy Reviews, Volume 89, Pages 51-60.

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

Shaner, M.R., Davis, S.J., Lewis, N.S., Caldeira, K., 2018 Geophysical constraints on the reliability of solar and wind power in the United States. Energy & Environmental Science, Volume 11, Issue 4, Pages 914-925.

Ssengonzi, J., Johnson, J. X., DeCarolis, J. F., 2022. An efficient method to estimate renewable energy capacity credit at increasing regional grid penetration levels. Renewable and Sustainable Energy Transition, Volume 2, Article 10033.

Stephen, G., Tindemans, S. H., Fazio, J., Dent, C., Figueroa Acevedo, A., Bagen, B., Crawford, A., Klaube, A., Logan, D., Burke, D., 2021. Clarifying the Interpretation and Use of the

LOLE Resource Adequacy Metric. 2022 17th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS).

Sutanto, D., Outhred, H. R., Lee, Y., 1989. Probabilistic power system production cost and reliability calculation by the Z-transform method. IEEE Transactions on Energy Conversion, Volume 4, Number 4., Pages 559-566.

Tapetado, P., Usaola, J., 2019. Capacity credits of wind and solar generation: The Spanish case. Renewable Energy, Volume 143, Pages 164-175.

Tarufelli, B., Eldridge, B., Somani, A., 2022. Capacity Markets for Transactive Energy Systems. Technical report prepared for the U.S. Department of Energy.

Terna, 2021. Mercato della Capacità: allegato 1 alla Disposizioni tecniche di funzionamento n.2. Regulations approved with Decree of 28 October 2021.

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'.

Tillmans, M., Schöttler, J., Praktinknjo, A., 2026. A review of probabilistic resource adequacy assessments in power systems: Methods, applications, and future challenges. Energy Policy, Volume 209, Part A, Article 114924.

Tong, D., Farnham, D.J., Duan, L., Zhang, Q., Lewis, N.S., Caldeira, K., Davis, S.J., 2021. Geophysical constraints on the reliability of solar and wind power worldwide. Nature Communications, 12, Article 6146.

Uryasev, S., 2000. Conditional value-at-risk: optimization algorithms and applications. Proceedings of the IEEE/IAFE/INFORMS 2000 Conference on Computational Intelligence for Financial Engineering (CIFEr) (Cat. No.00TH8520), New York, NY, USA, 2000, Pages 49-57.

Vazquez, C., Hallack, M., Vazquez, M., 2017. Price computation in electricity auctions with complex rules: An analysis of investment signals. Applied Energy, Volume 105, Pages 550-561.

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

Von der Fehr, N. H. M., Harbord, D., 1994. Capacity Investment and Long-Run Efficiency in Market-Based Electricity Industries. Conference on: Competition in the Electricity Supply Industry: Experience from Europe and the United States, Copenhagen, September 23-24, 1994.

Wang, R., Wang, S., Geng, G., Jian, Q., 2024. Multi-time-scale capacity credit assessment of renewables and energy storage considering complex operational time series. Applied Energy, Volume 355, Article 122382.

Wang, H., Shen, Z., Wang, Y., Wang, Y., Hu, J., Lan, Z., Jiang, X., Yu, H., Yang, L., 2023. Typical Capacity Remuneration Mechanisms and Enlightenment for an Imperfect Competition Electricity Market. 2023 IEEE 6th International Electrical and Energy Conference (CIEEC), Hefei, China, 2023, Pages 3508-3513.

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

Zachary, S., Wilson, A., Dent, C., 2022. The integration of variable generation and storage into electricity capacity markets. The Energy Journal, Volume 43, Number 4.

Zamani-Gargari, M., Kalavani, F., Abapour, M., Mohammadi-Ivatloo, B., 2018. Reliability assessment of generating systems containing wind power and air separation unit with cryogenic energy storage. Journal of Energy Storage, Volume 16, Pages 116-124.

Zeng, B., Wei, X., Sun, B., Qiu, F., Zhang, J., Quan, X., 2020. Assessing capacity credit of demand response in smart distribution grids with behavior-driven modeling framework. International Journal of Electrical Power & Energy Systems, Volume 118, Article 105745.

Zhou, Y., Mancarella, P., Mutale, J., 2016. Framework for capacity credit assessment of electrical energy storage and demand response. IET Generation, Transmission & Distribution, Volume 10, Number 9, Pages 2267-2276.

Zhou, Y., Mancarella, P., Mutale, J., 2015. Modelling and assessment of the contribution of demand response and electrical energy storage to adequacy of supply. Sustainable Energy, Grids and Networks, Volume 3, Pages 12-23.