Quantification of Operational Flexibility from a Portfolio of Flexible Energy Resources

Digvijay Gusain*, Miloš Cvetković, Peter Palensky

EEMCS, Delft University Of Technology, Delft, The Netherlands

Abstract

Flexibility will be required as more renewable generation is integrated in the power system. Therefore, entities such as Flexibility Service Providers (FSPs) have emerged who can control a diverse set of resources in their portfolio to provide requested flexibility. An accurate and reliable assessment of the amount of dispatchable flexibility in FSP's portfolio considering demand response activation, ramp rate, and post-activation rebound effects, etc. is important. In case FSP portfolio cannot fulfill the flexibility request, it must be known beforehand to avoid imbalance penalties. In this paper, we propose three flexibility quantification metrics: Expected Unserved Flexible Energy (EUFE), Expected Duration of Insufficient Flexibility (EDIF), and Expected Flexibility Index (EFI). These metrics are calculated using scenario-based simulations. The three metrics provide the FSP with quantifiable as well as graphical information on the magnitude and the duration of insufficient flexibility well ahead of gate closure time. This helps the FSP to design appropriate demand response programs (DRPs) and formulate their operational policy. These metrics account for uncertainty in forecasts of flexibility requests by evaluating multiple scenarios by conducting an operational simulation. With a representative case study, it is shown how an FSP can use the metrics to design DRPs, and when needed, schedule intraday market participation.

Keywords: flexibility quantification, flexibility service provider, statistical methods, power systems

1. Introduction

The increased flexibility needs in power systems have created an attractive opportunity for actors such as aggregators to act as Flexibility Service Providers (FSPs). The FSPs find customers in entities such as Virtual Power Plants (VPPs), Balance Responsible Party (BRP), Transmission System Operators (TSO), Distribution System Operators (DSO) who require flexibility to alleviate problems such as congestion management, feeder voltage management, forecast error correction, uncertain net-load ramps management, etc. Of particular interest are the flexibility needs arising specifically from the forecasting of power generation of variable renewable energy sources (VRES) present at various levels in the grid. Since weather cannot be forecasted much ahead of time, the actual power generation varies from the forecasted power generation, leading to requirements for correction of forecasting errors.

The BRP holds the responsibility to manage its imbalance before the system operator (SO) takes over the responsibility of management and settlement of imbalances (typically 3-4 Program Time Units (PTU) before the time of delivery. 1 PTU = 15 mins). The penalties for causing system imbalances are already large, and expected to increase even further in the future [1]. As such, it is of interest to the BRP to arrange the required flexibility beforehand to avoid any imbalance penalties. Enabled by a growing communication infrastructure, the FSPs utilize advanced control and coordination techniques to control a pool of flexible resources which can provide valuable flexibility to the BRPs [2, 3]. Typically, these resources consist of distributed generators (DG), thermostatically controlled loads (TCLs) (refrigeration systems, thermal storage, buildings, etc.), power-to-gas (P2G) systems, etc. This flexibility from the FSP can be purchased by the BRP from short-term markets such as the intraday markets [4], be contracted by the BRP as a bilateral agreement, or some form of a demand response market (DRX) [5]. Alternatively, the FSP can operate in close cooperation with the BRP, by controlling the resources in the BRP portfolio to fulfill the BRP's flexibility needs. The advantage of this idea is that the internal imbalances caused by the activation of flexibility are known to the BRP, and are managed as part of the flexibility provision scheme [6]. Such a model is already used by energy suppliers in the Netherlands [7].

In either case, it is crucial for the FSP to accurately assess and quantify the flexibility of its portfolio in order to provide reliable service. Estimating the flexibility of a pool of resources, however, is a challenging task. A pool of

*Corresponding author

Email addresses: d.gusain@tudelft.nl (Digvijay Gusain), m.cvetkovic@tudelft.nl (Miloš Cvetković), p.palensky@tudelft.nl (Peter Palensky)

Nomenclature

FSP	Flexibility Service Provider
EUFE	Expected Unserved Flexible Energy (MWh)
EFI	Expected Flexibility Index
EDIF	Expected Durations of Insufficient Flexibility
DRP	Demand Response Program
BRP	Balance Responsible Party
VRES	Variable Renewable Energy Sources
PTU	Program Time Unit
FRP	Flexibility Request Party
IPP	Individual Power Profile
FR	Flexibility Request
UF	Unserved Flexibility
MES	Multi Energy System
SO	System Operator
VPP	Virtual Power Plant
i	Index of flexible resource
k	Index of PTU
t	Index of time in signal
N_i	Total number of flexible resources
N_t	Total number of samples in signal
Δ_t	Time resolution of a signal (min)
P	Electrical power (MW)
$\dot{Q}$	Thermal Power (MW_{th})
$\overline{P}$, $\underline{P}$	Lower and upper bounds
ΔP	Available power deviation (MW)
r	Ramp rate (MW/min)
$\uparrow$ $\downarrow$	Upward and downward directions
τ	Time duration (min)
ν	Flexibility set
π	FR signal (MW)
P^{π}_{UF}	UF signal corresponding to π (MW)

resources can have a diverse set of characteristics such as ramp rates, up/downtimes, power ratings, etc. If the characteristics of each of the resources are represented as a set of static time-invariant values such as power deviations, ramp rates, etc., then methods derived from set theory can be used to represent the aggregated flexibility. This is referred to as the *available flexibility*. The amount of flexibility that can be dispatched, however, can be different. It can be constrained by factors such as device inter-temporal constraints, network capacity constraints, the geospatial spread of the flexible resources on the network, and even the operational strategy employed. Furthermore, the uncertainty of the flexible load demand and the VRES generation can lend additional complexity to the problem of quantification. This much more constrained form of flexibility is referred to as the *operational flexibility*. In this paper, we propose three metrics to quantify operational flexibility.

A few studies [2, 8, 9] have quantified available flexibility in power systems, but not the operational flexibility. In [10], the authors formulate an optimal policy for energy-constrained battery systems. In this reference, flexibility is defined as the maximum time a set of battery systems can follow a power reference signal before exhausting its energy. However, the authors do not consider ramp constraints, the network constraints, and are limited to consideration of upward flexibility, i.e., the battery systems that can only discharge. In [11], the authors investigate the flexibility of a flexible resource from its location in the grid. The proposed methodology, however, does not indicate the aggregated flexibility from all such resources in the grid and does not take into account the operational strategy of the FSP. An interesting take on flexibility is seen in [12] where the authors propose a flexibility metric that does account for transmission limits. The flexibility of the system is quantified as a tuple of lower and upper bounds of the largest uncertainty under which the system can remain operationally feasible. The metric introduced in [12] is a deterministic one, derived from solving a robust optimization problem. Although robust optimization guarantees the feasibility and optimality of the solution even against the worst-case scenario, it does not provide an expected solution, which is more relevant for studies on the short-term markets, especially the intraday markets [13, 14]. In [15], the authors introduce the Normalized Flexibility Index as a metric to quantify flexibility in individual generating resources and the power system as a whole. The authors use deployable range, the summation of total power deviation, and the average ramp rate, to calculate the contribution of each resource to system flexibility. This approach is developed only for generation-side flexibility. Authors in [16] provide insights on the quantification of flexibility in the context of supply adequacy and reserve requirement in power systems. The quantification metric is derived from a process control paradigm, where three-dimensional polyhedra are used to represent available flexibility. The authors use this notion to assess the balancing reserves and flexibility in supply to meet the demand. However, the measure does not include flexibility from responsive demand or sector coupling. More

recently, authors in [17] have proposed two tools: Flexibility Solution Modulation Stack and Flexibility Solution Contribution Distribution to quantify flexibility from the perspective of the entity providing it, i.e. the FSP. Their focus in quantifying flexibility is limited to only power system components, and cannot take into account the flexibility from sector coupling. The authors in [18, 19, 20] have proposed flexibility envelopes. The primary purpose of the envelopes is for flexibility adequacy planning from a generation viewpoint. The authors account for sub-hourly dynamics by using dynamic models of the power system generators, providing flexibility to the power systems. The method proposed by the authors does not account for flexibility from sector coupling. Additionally, since the models are included within the economic dispatch framework that the authors propose, the models are simplistic linear time-invariant systems. This limits the representation of more complex, flexibility providing resources such as TCLs which have flexibility rebound effects. A rebound effect occurs when the changes in power consumption of a flexible resource from activation of flexibility at a particular time restrict its ability to do so again in subsequent time steps [21, 22].

To summarize, a useful flexibility metric must:

1. provides an assessment of portfolio-level aggregated flexibility considering operational characteristics of diverse resources in the portfolio,
2. account for resource and network constraints,
3. account for uncertainty in VRES power production, flexible load demand, and flexibility requests,
4. accounts for the applied strategy of control and coordination of resources, and finally,
5. is intuitive to understand.

In this paper, we propose three new metrics to quantify operational flexibility. The proposed flexibility metrics are the first, to the best of the authors' knowledge, to jointly satisfy all the above requirements. The key characteristic of the newly proposed metrics is in their ability to quantify the portfolio response to uncertain flexibility requests while accounting for the network, resource, and spatio-temporal constraints and the operational strategy employed to dispatch flexibility. We use a scenario-based simulation method to achieve this.

The paper is further divided into the following sections. Section 2 introduces the need for short-term flexibility planning to the FSP. Section 3 dives into the current flexibility quantification techniques and their limitations. It then introduces the proposed metrics. Section 4 provides an example case study on which we showcase the usefulness of the proposed metric. Section 5 discusses the research and provides while Section 6 provides the conclusion of the study.

2. Short Term Flexibility Planning

The increasing levels of sector integration between electricity, gas, heat, transportation, etc. have created ample opportunities for the FSP to include flexible demand in its portfolio. The demand response policies (DRP) designed by the FSP, therefore, play a crucial role in determining the amount of operational flexibility that can be extracted from its portfolio. The FSP needs short-term flexibility planning to design suitable demand response policies that maximize the flexibility from its portfolio. Just as the system operator (SO) requires metrics such as EUE, LOLE, etc. to evaluate and compare long term flexibility measures (such as network expansion, generation planning, etc.), the FSP also requires metrics to evaluate short term flexibility measures (such as designing DRPs, selecting an operational strategy, comparing the value of flexibility from different products, etc.).

Figure 1 shows a typical scenario-based flexibility evaluation procedure for an FSP. For a given flexibility request (FR) signal the FSP expects to receive from a Flexibility Requesting Party (FRP) and the designed DRP (χ), the FSP executes an operational simulation subject to the operational constraints of the resources in its portfolio. Such constraints can be the minimum duration for flexibility provision, the maximum number of flexibility activation requests received, etc. Since the FR signal is an uncertain quantity for the FSP to determine, the FSP executes the operational simulation for multiple scenarios, each with a different realization of the FR signal. The results of these simulations are then quantified into flexibility metrics which are used to compare different DRPs and select the most appropriate policy.

The inclusion of operational simulation in Fig. 1 serves an important purpose in this quantification process. An important determinant of flexibility is the possible rebound effect [23, 24] that occurs after its activation. The authors believe that it is imperative that any flexibility quantification metric must be inclusive of this information. By executing an operational simulation, such effects are inherently captured through resource operational constraints. After settling in on a suitable DRP (coming out of redbox in Fig. 1), the FSP can determine if flexibility obtained from its portfolio is sufficient to fulfill the expected flexibility requests. If the portfolio flexibility is deemed insufficient, any additional flexibility can be procured either via the intraday market and demand response market or through bilateral contracts.

3. Flexibility Quantification

3.1. State of the art

Some efforts have been made to quantify and represent short-term flexibility from a portfolio using static resource characteristics. For example, in the case of thermostatically controlled loads, temperature bands around the nominal operating point are used to represent flexibility

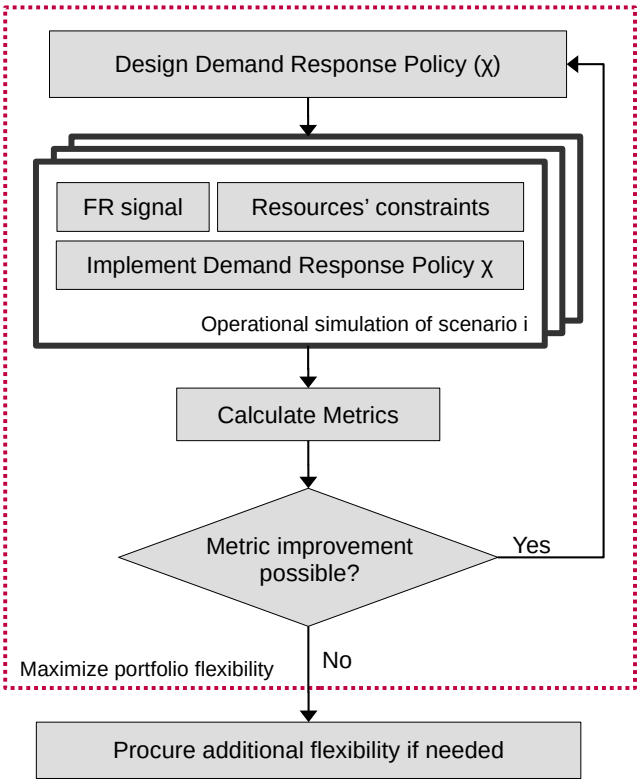

Figure 1: Planning for short term flexibility

[25]. Set notations have also been employed to represent the flexibility of a resource. In reference [26], the authors quantify flexibility from a single resource with a set of power, ramp rate, and energy deviations from the nominal operating point, and represent it as a polytope. When considering flexibility from a portfolio of resources, total flexibility is proposed to be calculated by applying set-theory methods. One such well-known and ubiquitously used method is the *Minkowski Summation* as proposed in [26]. However, as shown next, this is not an accurate representation of flexibility.

Consider a pool of N_i flexible resources. We denote the flexibility of resource i at a random PTU (Program Time Unit) k with the set $\nu_{i,k}^{\uparrow,\downarrow} = (\Delta P_{i,k}^{\uparrow,\downarrow},\ r_i^{\uparrow,\downarrow})$. Here, $\Delta P_{i,k}^{\uparrow}$ is the available power deviation $(\overline{P}_i - P_{i,k})$ in the upward direction and $\Delta P_{i,k}^{\downarrow}$ is the available power deviation $(P_{i,k} - \underline{P}_i)$ in the downward direction. $r_{i,k}^{\uparrow}$ and $r_{i,k}^{\downarrow}$ are the available ramp rate in upward and downward direction respectively. $P_{i,k}$ is the scheduled operating state at start of PTU k, whereas $\overline{P}_i$ and $\underline{P}_i$ are the maximum and minimum allowable power limits of resource i. The Minkowski Summation method for representing the aggregated flexibility available from this pool of N_i resources in PTU k is given by Eq. (1).

$$\begin{aligned}\nu_{ms,k}^{\uparrow} &= \Big(\sum_{i=1}^{N_i} \Delta P_{i,k}^{\uparrow}, \sum_{i=1}^{N_i} r_{i,k}^{\uparrow}\Big)\\ \nu_{ms,k}^{\downarrow} &= \Big(\sum_{i=1}^{N_i} \Delta P_{i,k}^{\downarrow}, \sum_{i=1}^{N_i} r_{i,k}^{\downarrow}\Big)\end{aligned} \tag{1}$$

For the sake of simplicity, we focus on the case of ramping up of a resource without losing generality and drop the $\uparrow,\downarrow$ notations. Then, $\nu_{ms,k}$ represents the aggregated flexibility (in the upward direction) at PTU k calculated using the Minkowski summation method (denoted as ms). However, since the Minkowski summation is a set summation method and is calculated for an interval of time, in this case for a PTU, it ignores the sub-hourly and sub-PTU resource inter-temporal dynamics important for flexibility assessment. To specifically account for these peculiarities in flexibility assessment, it is necessary to conduct an operational simulation. Consider in the same PTU k, discretized time series with $\Delta_t = 1$ min and $N_t = 15$, such that simulation time is $\Delta_t \cdot N_t = 15$ min. The simulated power profile of the resource within the PTU k can be represented by Eq. (2) as shown in Fig. 2.

$$\begin{aligned}&P_i[t] = \min(a'_{i,k}t + b_{i,k})\\ &\text{where}\quad a_{i,k} = [r_{i,k}\quad 0];\quad b_{i,k} = \begin{bmatrix}0\\ \Delta P_{i,k}\end{bmatrix}\end{aligned} \tag{2}$$

When a resource starts ramping at the start of PTU k at $t = 0$, depending on its ramp rate, the time to reach the required power level can be more or less than the PTU duration ($\tau_{PTU} = 15$ mins). We define *Time to Ramp*, denoted by $\tau_{i,k}$ (Eq. (3)), as the time taken by the resource i in PTU k to change its operating power by $\Delta P_{i,k}$ units.

$$\tau_{i,k} = \frac{\Delta P_{i,k}}{r_{i,k}} \tag{3}$$

If $\tau_{i,k}$ is less than the duration of a PTU, then the resource will ramp up and stay at the new power level for the remainder of the duration of the PTU as shown in Fig. 2. When flexibility from multiple resources is activated in PTU k, the aggregated continuous-time power profile of the portfolio will naturally be the sum of individual continuous-time power profiles (considering a copperplate network). We refer to this as an Individual Power Profile (IPP) Summation method. If all the resources have the same $\tau_{i,k}$, the ramp rate of the portfolio will be given by Eq. (1). Consequently, the time to ramp for the portfolio will be given by Eq. (4).

$$\tau_{ms,k} = \frac{\sum_{i=1}^{N_i} \Delta P_{i,k}}{\sum_{i=1}^{N_i} r_{i,k}} \tag{4}$$

In the case where at least two resources have different $\tau_{i,k}^{i}$, the ramp rate of the aggregated power profile will decrease at each $\tau_{i,k}$, making the aggregated ramp rate of the portfolio, a piecewise linear function. This decrease occurs because at each $\tau_{i,k}$, the resource i will reach the

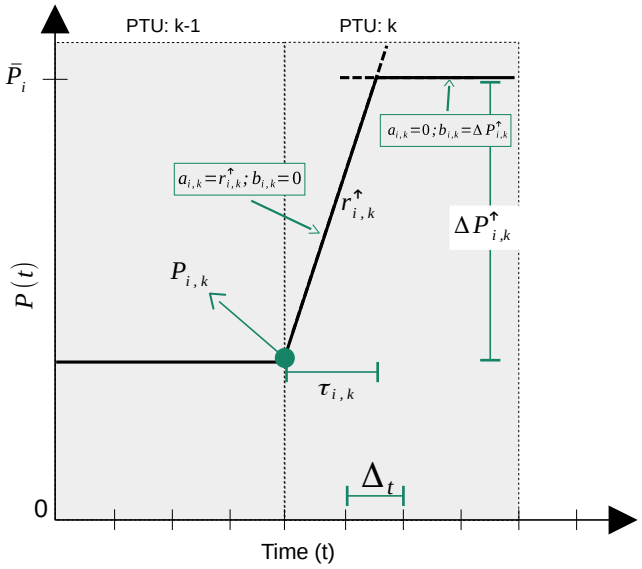

Figure 2: The sub-PTU power profile for resources with short time to ramp $\tau_{i,k}$.

specified power set-point and can no longer contribute to the aggregated ramp of the portfolio. Therefore, Eq. (5) gives the time to ramp for the portfolio.

$$\tau_{agg,k} = \min(\max\{\tau_{i,k} : \forall i \in N_i\}, \tau_{PTU}) \tag{5}$$

Although it seems logical to conclude that $\tau_{agg,k} = \tau_{ms,k}$, it can be proved that $\tau_{ms,k} \leq \tau_{agg,k}$. The basis for the proof lies in *Mediant inequality*.

Proposition 3.1 (Mediant Inequality)**.** *For positive real numbers a, b, c, d such that $a/c \leq b/d$ then the following holds:*

$$\frac{a}{c} \leq \frac{a}{c} \oplus \frac{b}{d} \leq \frac{b}{d}$$
$$\text{where} \quad \frac{a}{c} \oplus \frac{b}{d} = \frac{a+b}{c+d}$$

Corollary 3.1.1. *For any PTU k, let $\Xi_k = \{\tau_{i,k} : \forall i \in N_i\}$ be the set of time to ramp for each resource in the portfolio, then,*

$$\max(\Xi_k) \geq \frac{\sum_{i=1}^{N_i} \Delta P_{i,k}}{\sum_{i=1}^{N_i} r_{i,k}}$$

Proof. In the set Ξ_k, let j be the resource with the largest $\tau_{j,k} = \frac{\Delta P_j}{r_j} = \max(T_k)$. Therefore,

$$\frac{\Delta P_j}{r_j} \geq \frac{\Delta P_i}{r_i} \iff \Delta P_j \cdot r_i \geq \Delta P_i \cdot r_j \quad \forall i$$

Then, from Proposition 3.1, the following holds.

$$\frac{\Delta P_j}{r_j} \geq \frac{\Delta P_1 + \cdots + \Delta P_N}{r_1 + \cdots + r_N}$$

$$\iff \Delta P_j r_1 + \ldots \Delta P_j r_N \geq \Delta P_1 r_j + \ldots \Delta P_N r_j$$

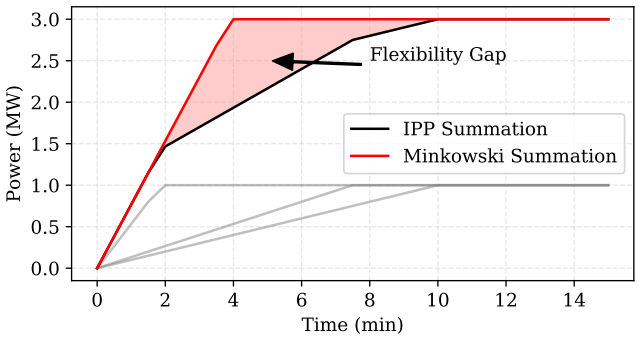

Figure 3: Aggregated portfolio power profile generated using Minkowski Summation method and IPP Summation method. The shaded region is the Flexibility Gap. The gray lines indicate IPP for R1, R2, and R3.

□

Since the $\tau_{ms,k} \neq \tau_{agg,k}$, there exists a difference between the continuous-time aggregated power profile generated using the Minkowski Summation method and that generated using the IPP Summation method. We refer to this gap as the *Flexibility Gap*.

To illustrate the flexibility gap, let us consider a small example. A portfolio of three flexible resources (R1, R2, R3), at a randomly chosen PTU k, is selected. Each of the three resources can increase their power consumption by 1 MW with ramp rates of 1.5, 2, 8 MW/PTU. The flexibility set ν of the portfolio is given by {(1, 1.5), (1, 2), (1, 8)}. Following Eq. (1), the total ramp, and power deviation available from the portfolio in this PTU is $\nu_{ms,k} = [(3, 11.5)]$. We will drop the PTU notation (k) henceforth in this example since we consider a single PTU. We take $\Delta_t = 1$ min and $N_t = 15$, such that total simulation time is $\Delta_t \cdot N_t = 15$ min.

Next, consider a flexibility request $\nu_{req} = [(2.5, 11.5)]$, implying a flexibility request for $\Delta P = 2.5$ MW at 11.5 MW/PTU ramp rate. When all the resources in the portfolio are activated simultaneously to fulfill the request, the ramp rate of the aggregated profile trajectory is $\sum r_i = 11.5$ MW/PTU. Among the three resources, the one with the lowest τ_i is R3 ($\tau_3 = 0.125$ PTU $= 2$ min). When ΔP reaches 1 MW, R3 will stop contributing to the total ramp at $t = 2$. With only R1 and R2 now contributing, the total ramp decreases to $\sum r_i = 3.5$ MW/PTU. This process will continue as R2 ($t = 7.5$ min), and then finally R1 reach their respective $\Delta P = 1$ MW ($t = 10$ min). The continuous-time power profile of the aggregated portfolio (in red) and individual resources (in gray) within the PTU is shown in Fig. 3. It is seen that the total portfolio ramp changes at each τ_i. In the same figure, the black line represents the power profile generated using the Minkowski Summation method. It can be seen here that $\tau_{agg} = \tau_1 \geq \tau_{ms}$, where τ_1 is the time to ramp for R1. The shaded region between the two curves is the flexibility gap.

3.2. Inspiration from System Adequacy Metrics

In this section, we seek to be inspired by the metrics used by system operators (SO) to ensure adequacy in power systems. These will serve as an inspiration for the metrics this paper proposes. The SO uses the power system adequacy metrics to make appropriate investment and policy decisions such that system reliability is maintained. The Loss of Load Probability (LOLE) is an example of such an adequacy metric and it describes the expected number of hours when the load will not be served by a given generation capacity. This information helps the SO to determine if the amount of generation is adequate to serve the forecasted inflexible load. Similarly, the Expected Unserved Energy (EUE) metric, uses the forecasted data for a given time horizon (typically, a year), and calculates the expected amount of energy that would not be served with the given generation portfolio. The SO uses these metrics extensively in generation planning, network expansion planning, etc.

Since, the power system adequacy metrics (such as LOLE, EUE, etc.) were created for a system that was based on the concept of "generation must always follow demand", this concept is no longer "always" applicable to the renewable-rich power system. This limits their applicability in future power system planning. To overcome this challenge, several authors suggested modifications to existing adequacy metrics that can assist the traditional generation adequacy studies with VRES. The Insufficient Ramp Resource Expectation (IRRE) metric proposed in [27] is a power system flexibility metric that takes into account the operational schedules of generators in the portfolio and provides the expected number of observations when the system will face ramping shortages arising from the variations in netload. Another metric introduced in [28] is called Periods of Flexibility Deficit (PFD). The PFD measures the frequency of flexibility shortfalls of specified time duration for each hour. The flexibility needs are derived from netload ramps, similar to those in IRRE. The calculation of IRRE and PFD follows that of LOLE and helps the SO in generation planning considering challenges from VRES. These metrics quantify the system inflexibility based on which generation planning decisions (such as investments in a more high-ramping generation like gas turbines) can be formulated.

Due to this lack of metrics to quantify short-term flexibility, many studies employ power system adequacy metrics in some form to the short-term flexibility planning. Reference [29] incorporates the Expected Energy Not Served Supplied (EENS) metric with its Conditional Value at Risk (CVaR) to determine optimal sizing and location of the DER units in the distribution system using a risk-based optimization framework. In [30], the authors employ Expected Unserved Energy (EUE) and Loss of Load Probability (LOLP) to formulate a reliability constrained unit commitment problem. Other adequacy metrics such as SAIDI and SAIFI [31, 32], etc. are also employed to conduct operational studies to evaluate the impact of flexibility from DER and flexible loads in a microgrid. While the metrics IRRE and PFD overcome some challenges posed by VRES, they are still primarily focused on generation planning in the long term.

The disadvantage of using power system adequacy metrics for short-term flexibility studies stems directly from their calculation methodology, and consequently, their interpretation. Firstly, these metrics are calculated by aggregation of yearly data, which implies that sub-hourly dynamics are unaccounted for. This is important since, in operational time scales, such dynamics can constrain the amount of flexibility [33]. Secondly, the impacts stemming from activation of flexibility, such as previously mentioned flexibility rebound effects cannot be captured through these metrics. Thirdly, since these metrics are designed to evaluate the power system's ability to meet demand in the long term [34], they do not consider any network-related constraints or uncertainty in demand and VRES generation.

Despite the disadvantages, the methodology used for calculating these metrics is noteworthy because it condenses complex relationships between various variables (such as those defined in AC and DC power flow equations) available as time-series data (yearly load, generation profiles) into a single numerical value that conveys useful information to the SO about adequacy of generation in the power system. To the best of the authors' knowledge, metrics specifically addressing the requirements for short-term flexibility quantification highlighted in Section 1 do not exist. Few metrics that do quantify short term flexibility are furthermore shown to be inaccurate (Section 3.1).

In Section 3.3, we therefore propose three new flexibility quantification metrics. These metrics take inspiration from the methodology used in the calculation of system adequacy metrics to address the complexities introduced by requirements listed in Section 1.

3.3. Proposed Metrics

Revisiting the requirements for a useful flexibility metric from Section 1, we propose three flexibility quantification metrics derived from scenario-based simulations. The proposed metrics provide a comprehensive overview of the operational flexibility available to the FSP from its portfolio. However, to understand and derive the metrics, it is first important to define the following terms:

- Flexibility Request (FR) signal: The FR signal (denoted by π) is a time series of power setpoints that the FSP receives from the FRP.
- Unserved Flexibility (UF) signal: The UF signal (denoted by P^{π}_{UF}) is determined after running an operational simulation using the FR signal. The FSP aims to follow the FR signal as closely as possible by controlling and coordinating the resources in its portfolio. Therefore, any non-zero values of P^{π}_{UF} will imply the FSP is unable to fulfill some part of the flexibility request. Technically, if P_{sh} is the power

signal that represents the power shifted by FSP to fulfill the FR signal, then, we can define P^{π}_{UF} as UF signal given by Eq. (6).

$$P^{\pi}_{UF} = |\pi - P_{sh}| \tag{6}$$

We can now define the three metrics:

3.3.1. Expected Unserved Flexible Energy (EUFE)

The energy in the UF signal (P^{π}_{UF}) after serving a flexibility request π is the unserved flexible energy (UFE). However, the FR signal is an uncertain quantity for the FSP to determine. To address uncertainty in the flexibility evaluation, the FSP considers a range of FR signals ($\pi_1, \pi_2 \ldots \pi_n \in \Pi$) to obtain the expected value of the UFE. This expected value is termed as Expected Unserved Flexible Energy (EUFE) and calculated using Eq. (7). The EUFE metric is similar to the calculation of the power system reliability metric Expected Unserved Energy.

$$\text{EUFE} = \frac{1}{N_\pi} \cdot \sum_{\pi \in \Pi} \frac{\Delta_t}{60} \cdot \sum_{t=1}^{N_t} P^{\pi}_{UF}[t] \tag{7}$$

Here, P^{π}_{UF} is the UF signal in MW, Δ_t is the time resolution in min and N_t is the total time steps in the UF and FR signals, and N_π is the total number of FR signal scenarios.

3.3.2. Expected Durations of Insufficient Flexibility (EDIF)

The Expected Duration of Insufficient Flexibility (EDIF) is a graphical measure of flexibility. It aggregates the UF signal calculated after analyzing all the N_π UF signals to generate an overview of time periods of insufficient flexibility in the portfolio and its direction and presents it as a heatmap.

3.3.3. Expected Flexibility Index (EFI)

Expected Flexibility Index or EFI provides a more general and intuitive indication of the flexibility of a portfolio. EFI is a unitless measure of the operational flexibility of the portfolio. It is the expected value of $\hat{F}_\pi(0)$ calculated over the set of FR signals ($\pi_1, \pi_2 \ldots \pi_n \in \Pi$). Here, $\hat{F}_\pi$ is the Empirical Cumulative Distribution Function (ECDF) (shown in Appendix A) of the UF signal P^{π}_{UF}. $\hat{F}_\pi(0)$ denotes the fraction of the UF signal π that is zero. The FSP's ability to fulfill the FR signal is indicated by $\hat{F}_\pi(0)$, which is the fraction of the total time steps where the value of P^{π}_{UF} is 0. The FSP is completely flexible to fully service the FR signal π when $\hat{F}_\pi(0) = 1$. Conversely, the FSP is completely incapable of servicing any FR signal when $\hat{F}_\pi(0) = 0$. While no particular value is a reference for EFI, the FSP should desire a value as close to 1 as possible.

The FSP uses the EFI metric as part of the flexibility assessment of its portfolio (as shown in Fig. 1). The FSP evaluates the EUFE and EFI of its portfolio and by combining these two metrics with the EDIF heatmap, increases flexibility within its portfolio by implementing appropriate DR actions and operational policies. When internal flexibility cannot be increased any further (i.e., when EFI cannot be increased through DR actions), it uses the EUFE and EDIF heatmap to calculate the amount of additional energy that would need to be traded on the intraday market. The FSP's evaluation of the flexibility metrics also depends on the operational policy it applies. If needed, the FSP can change it and use another policy (such as maximizing available storage level at end of the day, minimizing expected imbalance costs, minimize degradation of storage, etc.) as well. This gives the FSP the freedom to not only evaluate flexibility from resources but also compare the impact of a particular operational policy on the flexibility obtained from its portfolio.

It must be noted here that both the FR signal and the UF signal (input to the operational simulation and the output generated from it), are time series characterized by signal time resolution Δ_t and the total number of time steps N_t. The resolution of this time series depends on the user-selected simulation and the dynamics of interest. For example, in a typical optimal dispatch simulation with an interest in sub-hourly dynamics, the Δ_t can be in minutes (1 min, 5 min, 15 min, etc.), while in closer to real-time dynamic simulation with predictive control and an interest in transient dynamics, Δ_t can be in seconds. In either case, once these signals are known, the proposed metrics can be calculated.

4. Example Study

Consider a virtual power plant (VPP) that aggregates a lot of distributed wind generators in a region. The VPP optimizes its power exchange for each hour of day D with the external grid by executing a day-ahead scheduling program on D-1. The VPP is its own BRP, once the schedule is confirmed by the TSO, it is solely responsible for any imbalance occurring in its area. Since the VPP is a collection of distributed VRES, forecasting errors, and hence imbalances are inevitable. It, therefore, requires flexibility to avoid unforeseen costs arising from the resulting imbalance. In this example, we only consider the flexibility needs arising from forecasting errors, although the proposed metrics can be applied to evaluate portfolio flexibility to participate in other value chains as well, such as congestion management in distribution grids. The VPP contracts an FSP to obtain this flexibility. The FSP can use multiple electrical and thermal loads connected to the local electric and heat network and exploit their flexibility. The FSP, therefore, must evaluate its portfolio's flexibility to properly assist the VPP to manage its imbalance.

4.1. FSP's Portfolio Description

In this example, the portfolio of the FSP is a multi-energy system (as shown in Fig. 4). The choice of selecting the FSP portfolio as a multi-energy case study as

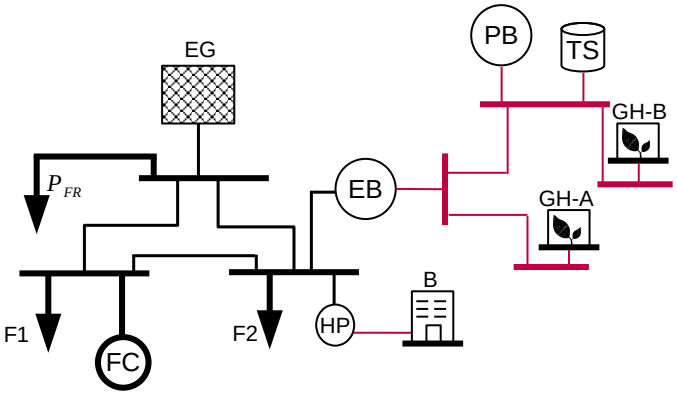

Figure 4: Proposed multi-energy system setup representing the FSP's portfolio. The red area represents the heat network while the black part represents the electricity network.

opposed to a pure power system one is natural. Given the prevalence of P2X resources and consequently, tighter integration between electricity and other energy sectors, the MES must be analyzed as a whole. Ignoring the effect of one energy sector in an MES to analyze the P2X resources will result in sub-optimal results [35]. The electric network interfaces to the higher voltage level via the main substation (here represented as an external grid (EG)). There are two additional substations (SS-A and SS-B) servicing two feeders (F1 and F2). These substations are connected to EG via underground cables. Additionally, SS-A connects a small-scale fuel cell (FC) to the electrical grid while SS-B is connected to a large electrical boiler (EB). The EB interfaces the electrical grid to the local heat grid. The heat grid is responsible for supplying necessary heat to two nearby greenhouses (GH-A and GH-B). For emergencies, the heat grid also contains a peak boiler (PB) operated using natural gas. A thermal storage (TS) tank is available as a buffer and to reduce the dependence on GB. A building load (B) is connected to SS-B via a heat pump (HP). The FSP controls F1, F2, FC, EB. Additionally, it controls the temperatures inside the building (θ_B) and the greenhouses (θ_{GH}). The building temperature is controlled by modulating the HP power, while the temperatures inside GH-A, GH-B are controlled by modifying the thermal power demand of greenhouses. To meet the thermal power demand, the FSP controls the EB and GB thermal power outputs. Table 1 lists the upper and lower limits of power and ramp rates of flexible resources in the given portfolio.

4.2. FR Signals

The FSP receives the FR signal from the VPP to self-balance its renewable-rich portfolio throughout the day. To prepare for it, the FSP needs to assess the internal flexibility of the portfolio. Since the FR signal is an uncertain quantity for the FSP to determine, to consider this uncertainty, the FSP generates possible scenarios of the FR signals. In this case, we assume that the VPP has a

Table 1: Flexibility parameters for resources

Resource	Parameter	Value	Unit
EB	$\underline{P}_{EB}$, $\overline{P}_{EB}$	0, 1	MW
F_i	$\underline{P}_{F,i}$, $\overline{P}_{F,i}$	0, 3	MW
F_i	$r^{\downarrow}_{F,i}$, $r^{\uparrow}_{F,i}$	-3, 3	MW/PTU
FC	$\underline{P}_{FC}$, $\overline{P}_{FC}$	0, 3	MW
FC	$r^{\downarrow}_{FC}$, $r^{\uparrow}_{FC}$	-3, 3	MW/PTU
HP	$\underline{P}_{HP}$, $\overline{P}_{HP}$	0, 2	MW
HP	$r^{\downarrow}_{HP}$, $r^{\uparrow}_{HP}$	-2, 2	MW/PTU
B	$\underline{\theta}_B$, $\overline{\theta}_B$	24, 25	$^{\circ}C$
GH_i	$\underline{\theta}_{GH}$, $\overline{\theta}_{GH}$	33, 34	$^{\circ}C$
PB	$\underline{\dot{Q}}_{PB}$, $\overline{\dot{Q}}_{PB}$	0, 0.5	MW_{th}
TS	$\underline{\dot{Q}}_{TS}$, $\overline{\dot{Q}}_{TS}$	-0.5, 0.5	MW_{th}
TS	$\underline{\dot{S}}_{TS}$, $\overline{\dot{S}}_{TS}$	0, 0.5	MW_{th}

significant amount of wind power in its portfolio. Due to inherent uncertainty in weather, there almost always exists a difference between the day-ahead committed power from the wind generator and the actual power generated, which gives rise to wind power forecasting errors. In this case, we assume that the VPP portfolio is dominated by wind power, and therefore, the VPP's flexibility requests comprise mainly of these wind power forecasting errors. It is shown in [36] that wind power forecasting errors can be modeled using a Gaussian distribution. The Gaussian model is parameterized by $\mathcal{N}(0, 0.1)$. We consider power balancing at the point of common coupling of VPP and the grid and assume, with no loss of generality, that network constraints are respected. A dummy power system load represents the FR signals at the point of common coupling.

4.3. Operational Simulation

The FSP looks to maximize its ability to service any flexibility request (π) it receives from the VPP. In other words, the FSP's operational policy X is to minimize the UF signal and is given by Eq. (8).

$$X = \min \mid P^{\pi}_{UF} \mid \tag{8}$$

Since the direction of imbalance, P^{π}_{UF}, is not as relevant as the magnitude (imbalance in whichever direction it occurs, is always penalized), we minimize the absolute value of P^{π}_{UF}. In this example study, we set $\Delta_t = 15$ min and $N_t = 96$ to represent an optimal re-dispatch for the entire day. We are interested in understanding the impacts of thermal temperature range on flexibility and flexibility from sector coupling. Since the dynamics in the heat sector are slower than those in the electricity sector, this choice of time resolution is justified. It must be noted, however, that it is also possible to consider finer resolution signals with $\Delta_t = 5$ min or $\Delta_t = 1$ min to quantify flexibility using the proposed method and metrics when resources such as

electrolyzers, battery systems are included where the sub-PTU dynamics can be more relevant in the quantification process.

The objective in Eq. (8) is subject to various equality and inequality constraints in Appendix B. These constraints form the original DRP employed by the FSP. The resulting MILP problem is modeled in Python using pyomo [37] environment and solved using the Gurobi solver. The simulations are carried out on a dual-core Intel i7-10510U CPU @ 1.80GHz running Ubuntu 20.04.

4.4. Overview of Case Studies

In total, we investigate five cases. These are divided into three subsections for easier understanding. The first subsection Section 4.5 is the *Flexibility Assessment using Minkowski Summation*. We perform Minkowski summation as the baseline and show the difference between this assessment and the one obtained using the proposed EFI and EUFE metrics, and the EDIF heatmap (calculated using scenario-based simulation approach) is highlighted. The second subsection Section 4.6 is the *Flexibility Assessment for Designing DRP*. In this subsection, we investigate two cases (C2, C3), built upon the results from C1. Herein, we show the use of proposed metrics and heatmap to assist the FSP in tuning their DRPs. The third subsection Section 4.7 is the *Flexibility Assessment for Intraday Market Participation*. In this subsection, two more cases (C4 and C5) are investigated which further build upon the results from C3. The idea here is to show the use of the proposed metrics and heatmap in assisting the VPP+FSP entity plan for flexibility shortages using the intraday market. The cases are explained in further detail in the following subsections.

4.5. Flexibility Assessment Using Minkowski Summation

This method explicitly takes into account only the ramp rates (in terms of MW/min) of the electrical assets of the FSP —F1, F2, FC, EB, and HP. In other words, the flexibility of the resources connected to the heat grid (e.g. TS, ...) and the operational inter-dependencies resulting from sector coupling are not captured by the method.

As discussed previously, Minkowski Summation for flexibility assessment is not based on simulation, but on set summation. The FSP using this method for flexibility assessment does so by calculating the aggregated power deviation and ramp rate available from the resources' planned operational state (the D-1 optimal power schedule) for each required interval of time and at each required time step using Eq. (1). This gives upper and lower bounds on the power excursions that its portfolio can take in each interval of time at each time step. For our MES case, the Minkowski evaluated bounds are shown in Fig. 5.

Figure 5 can be interpreted as follows: the FSP is assured that there is enough flexibility in the portfolio to fulfill any received FR signal it receives at any point in time. In the rest of this section, we will use the metrics

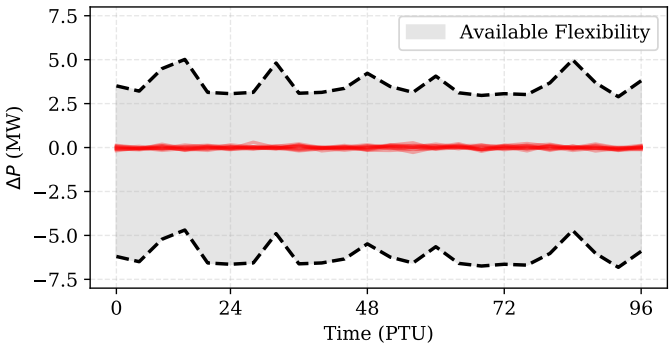

Figure 5: Assessing flexibility using Minkowski Summation. The red lines indicate the FR signal scenario set. The black dashed lines indicate the upper and lower levels of allowable power excursion calculated using the Minkowski method. The available flexibility bound is calculated in each PTU for the next PTU.

proposed in this paper to show that this characterization is invalid. A major limitation here in addition to previously mentioned issues is that the assessment from this method does not inform the FSP of changes occurring in the event of the activation of flexibility in any resource. In such an event, the resource's and consequently, the portfolio's operational state changes. The available power deviation and ramp rate from the resource change for a subsequent time step, directly impacting the flexibility available in subsequent periods. Therefore, a re-assessment of portfolio flexibility is needed every time flexibility in a resource is activated.

To compare with the method proposed in this paper, flexibility is evaluated by conducting a simulation-based assessment and then calculating the proposed metrics for this case. In this baseline case C1, the operational constraints given by Eqs. (B.1) to (B.13) form the DRP. The temperature in the office building is to be maintained between $24°C$ and $25°C$. For the greenhouses, the temperature must be maintained between $33°C$ and $34°C$. The total load on feeders F1 and F2 cannot be ramped by more than 3MW in either direction per PTU. The fuel cell generator power rating is 3MW, which limits its flexible operation. The values for the parameters are summarized in Table 1.

For C1, the operational simulation results in a EUFE value of 1.858 MWh and the EFI value of 0.906. Although the EFI is more than 0.9, a high flexibility index, a closer look at FR and UF signals in the worst-case scenario in Fig. 7 shows that while the FR signal is absorbed in most parts of the day ($P^{\pi}_{UF} = 0$), the UF signal still has large and frequent variations in the remaining time steps. This indicated that the DRP merely aggregates and shifts the imbalances in the FR signal to certain times during the day. This is supported by the calculation of imbalance energy in the FR signal set was 1.92 MWh, whereas the EUFE calculated after C1 simulation is 1.858 MWh. Therefore, the current DRP is highly ineffective and allows the FSP to reduce the imbalance energy by 3% only.

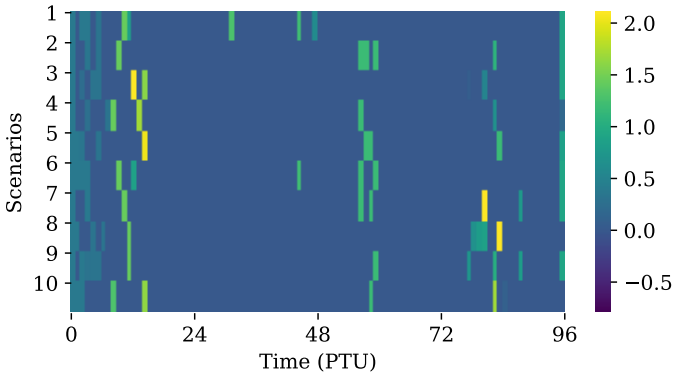

Figure 6: EDIF heatmap for C1. As is visible, there exist significant durations of insufficient flexibility early in the morning, few in the daytime, and then again at late night. This is true for almost all the scenarios.

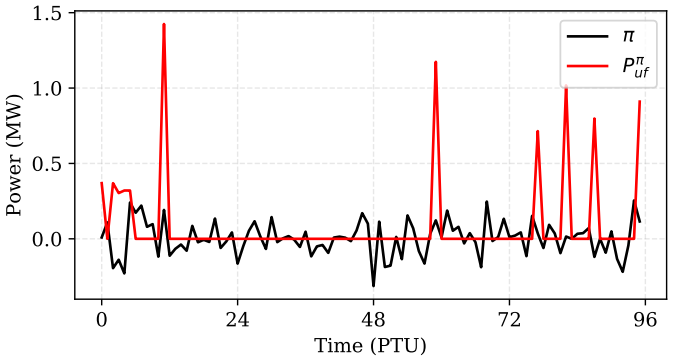

Figure 7: The FR signal realization from $\mathbf{\Pi}$ and the corresponding P^{π}_{UF} signal in C1 which has the worst UFE. While the FR signal is absorbed in most parts of the day ($P^{\pi}_{UF} = 0$), the UF signal still has large and frequent variations in the remaining time steps, indicating that the policy merely aggregated and shifted the imbalances to certain times in the day.

In Fig. 6, the expected times (PTUs) where the FSP's portfolio flexibility is insufficient to service the FR signal scenario set $\mathbf{\Pi}$ is shown. The streaks on the heatmap reflect the magnitude of P^{π}_{UF} for each scenario. The ECDF chart corresponding to the signals in Fig. 7 representing the flexibility index is shown in Fig. 8.

4.6. Flexibility Assessment for Designing DRP

It is visible from the calculated metrics and the EDIF heatmap that the baseline DRP on the portfolio is highly incapable to service the FR signal set, despite the assessment made using the Minkowski Summation method. The largest magnitudes in the set of P^{π}_{UF} can be seen particularly in the early mornings, midday, and then late-night periods. The FSP will want to reduce the EUFE of its portfolio while also improving the EFI. An obvious source of additional flexibility is the thermal flexibility from the thermal loads. An increased range for temperature variability will proportionally allow increased power variations (Eq. (B.4)), and therefore, create greater flexibility in the portfolio. By modifying constraints in C1, two new case

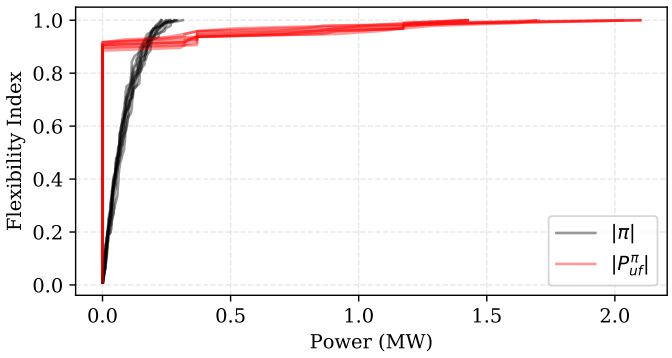

Figure 8: The ECDF of the FR signals and UF signals obtained from scenario simulations for C1 provide us with the expected flexibility index. This also shows that relatively larger power variations compared to FR signals exist, but only for few time steps.

studies are designed: C2 and C3, representing two options for DRP tuning.

- C2: The temperature bands are increased by $\pm 1°C$ across all three thermal loads (B, GH-A, GH-B). Since the modification is time-invariant, this is referred to as additional non-targeted DR.
- C3: The office building temperature bands are relaxed only for time intervals with low expected occupancy. The temperature limits in the building are set to $22 - 27°C$ between 0000h-0800h, and 1900h-0000h; to $24 - 25°C$ between 0800h-1900h. An additional constraint is added which ensures that the building temperature is equal to $24°C$ at 0800h (typical start of work hour). The temperature bands for greenhouses are kept unchanged from C1. The reason for this choice of DRP stems from the knowledge that crop yields are sensitive to ambient temperature variations and therefore maintaining the greenhouse temperature will have greater priority than maintaining thermal comfort in the office temperature. Since this DR strategy has time variance, it is referred to as additional targeted DR.

The operational simulation of C2 with updated temperature ranges results in a EUFE of 0.746 MWh. This shows an incredible increase in flexibility of the portfolio by $\approx 60\%$ compared to C1. The EFI of 0.934 corroborates the result that increasing temperature bands by ± 1 degrees on all thermal loads can provide significant flexibility to the FSP. This is also visible with the EDIF heatmap in Fig. 9 where the streaks are reduced in intensity and frequency of occurrence across all scenarios. It can be observed that for the same PTUs that had a high magnitude of unserved flexibility in C1, the corresponding values for C2 are lower, some even reduced to zero. Although the magnitude and frequency of occurrence of unserved flexibility are reduced to some extent, it exists prominently in most of the scenarios.

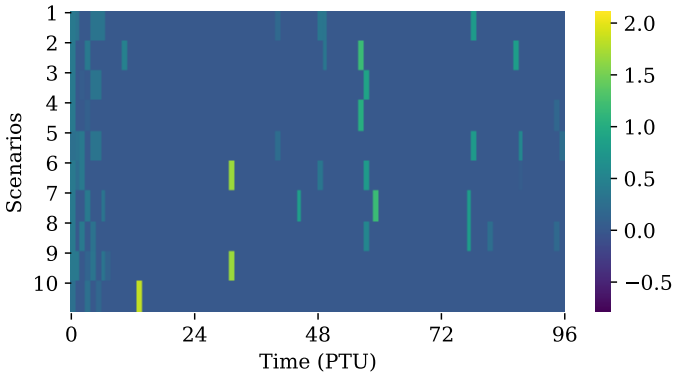

Figure 9: EDIF heatmap for C2. Increasing temperature bands on all thermal loads leads to significant reductions in durations of insufficient flexibility compared to C1.

With C3, the operational simulation resulted in a EUFE value of 0.608 MWh whereas the EFI is calculated to 0.948, which suggests that C3 offers an improvement over C2 and is, therefore, a more effective DRP. The decrease in EUFE values compared to C1 and C2 can be explained by the fact that the building heating load is directly connected to the electricity grid, whereas the greenhouses are connected to the heat grid. While the heat storage tank along with the peak boiler and the electric boiler serves the loads in the heat grid, the heat load for the building is supplied only from the electricity grid. Thus, the flexibility of the building thermal load is directly available to the power system, whereas flexibility from greenhouses is indirectly available, depending also on the operating state of resources in the heat network. The EDIF heatmap for C3 shown in Fig. 10 shows that the values of the unserved flexibility signal. Although there is not a significant difference, there seem to be lesser times of insufficient flexibility across all scenarios compared to C2.

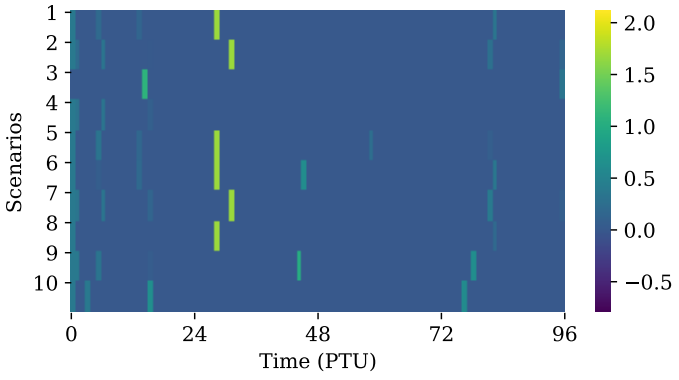

Figure 10: EDIF heatmap for C3. Targeted demand response from office building load by modifying temperature bands provides slightly more flexibility to the system compared to C2.

The temperature of the office building and the greenhouses is also shown in Fig. 11. The larger temperature variation range in the early morning and night periods allows the building load to be more flexible with temperature variations, thus providing increased flexibility to the FSP.

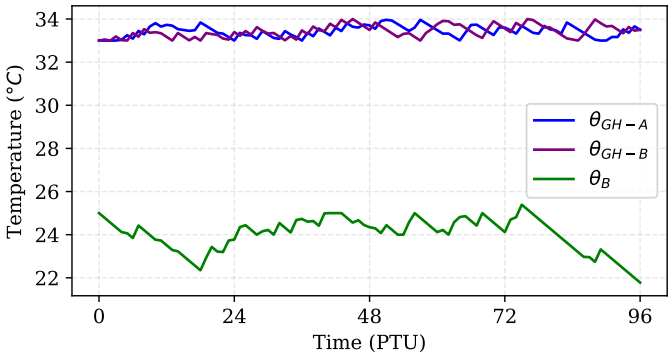

Figure 11: The temperatures for greenhouses and office building in FSP portfolio.

As shown in these cases, the metrics EUFE, EFI, and EDIF heatmap helps the FSP to compare flexibility from different DRPs and resources in a quantitative manner. In the previous cases, it can be seen from an improved EFI and EUFE values, that targeted DR actions will have a greater impact on the operational flexibility of the FSP portfolio.

4.7. Flexibility Assessment for Intraday Market Participation

Although a lot more combinations of DR actions can be evaluated, such as targeted temperature band variations for greenhouses, however, for the sake of brevity, we assume that this is the maximum flexibility that the FSP can obtain from its portfolio. This implies that the EUFE of 0.608 MWh (value from C3) still needs to be procured from external sources and therefore, this information is relayed by the FSP to the VPP. In this example, the additional flexibility is procured by the VPP from the intraday market. Two strategies (and consequently, two new cases) for the VPP are envisioned:

- C4: VPP is a risk-taking entity and buys energy equivalent to average UFE in each PTU in C3.
- C5: VPP is a risk-averse entity and buys energy equivalent to UFE in the worst performing scenario in C3.

In C3, the EFI value was 0.948, which implies that on average, the portfolio was able to service flexibility requests for 94.8% of the 96 PTUs (leaving ≈ 6 PTUs to be filled) evaluated for each scenario. Therefore, on average, 0.608 MWh of energy needs to be procured for 6 PTUs per scenario. The location of these PTUs in the day can be approximated by analyzing the EDIF heatmap. Upon consideration of the EDIF heatmap for C3, the six periods of insufficient flexibility are PTUs 12, 28, 31, 45, 76, 78. Therefore, 0.608 MWh/6 = 0.101 MWh of energy needs to be procured for the 6 PTUs.

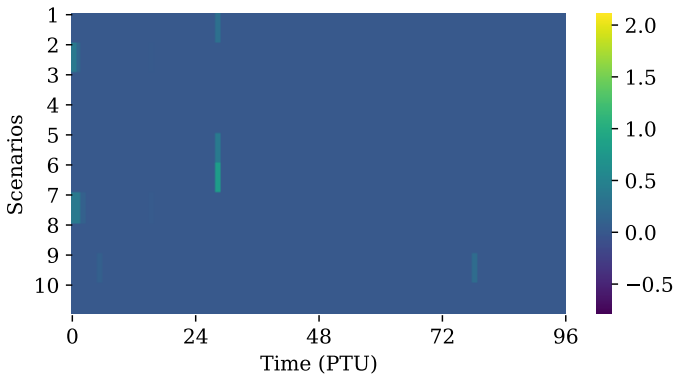

Figure 12: EDIF heatmap for C4. The energy to be procured from the intraday market is calculated based on the EFI and EUFE values, along with EDIF heatmap from C3.

In C4, we specify the power bought from the intraday market ($P_{ID}(k)$) at PTU k in Eq. (B.2) as 0.101 MWh/0.25h = 0.405 MW for each of the 6 PTUs listed. On executing the operational simulation, the EUFE value is calculated to 0.077 MWh, while the EFI is 0.987. The corresponding EDIF heatmap is shown in Fig. 12. Despite the expectation that procuring the calculated amount of energy will result in ideal EFI (=1) and EUFE (=0) values, the obtained values can be reasonably explained.

Investigating the EDIF heatmap for C3, it can be seen that inflexibility only occurs in a few scenarios at isolated PTUs and each of them is of different magnitudes. In C4, we calculated the power to be procured from the intraday market using expected values (EUFE and EFI). This masks a magnitude of insufficient flexibility in each PTU and scenario. For a risk-taking VPP, this could be an acceptable solution, since the EDIF heatmap suggests that most of the inflexibility is measured for some scenarios and only at a very small number of PTUs. An EFI close to 1 (0.987) reaffirms that the small inflexibility is likely to occur in one or two PTUs in a tiny number of scenarios.

In C5, the FSP is a risk-averse entity, which means that it would analyze the worst-case scenario and plan its market participation accordingly. To achieve this, more attention needs to be paid to the EDIF heatmap from C3. The scenario with the highest unserved flexible energy is selected, and intraday market participation is derived by analyzing this scenario. For C3, this is scenario 7. The unserved flexible energy in scenario 7 is 0.82 MWh and is seen in PTU 1, 2, 7, 16, 32, 82, and 96. This indicates an insufficient energy equivalent to 0.82 MWh/6 = 0.1171 MWh for each PTU which needs to be procured from the intraday market. The value of P_{ID} for these PTUs is set to 0.1171 MWh/0.25h = 0.4686 MW in Eq. (B.2). The operational simulation with the updated values of P_{ID} results in EUFE of 0, and EFI of 1. The EDIF heatmap for C5 is shown in Fig. 13, where it can be seen that there are no streaks, implying in all scenarios, all flexibility requirements are met with the proposed intraday market participation. The FR signal and UF signal for the evaluated scenario 7 is shown in Fig. 14. Figure 15 represents the ECDF chart for signals in C5. The results from the cases are summarized in Table 2 and plotted in Fig. 16.

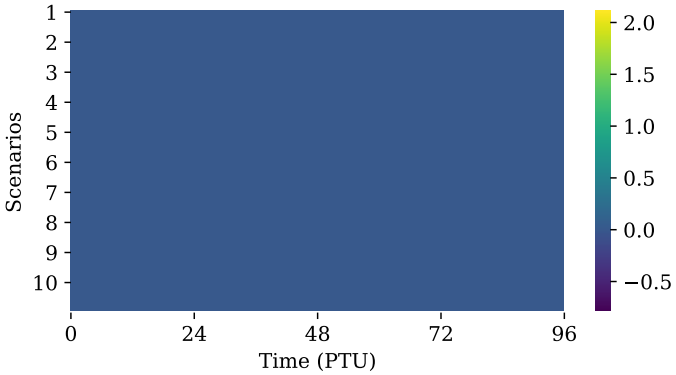

Figure 13: EDIF heatmap for C5. No streaks indicate that the expected procured energy from intraday market will be sufficient to guarantee that any flexibility request from VPP can be accommodated.

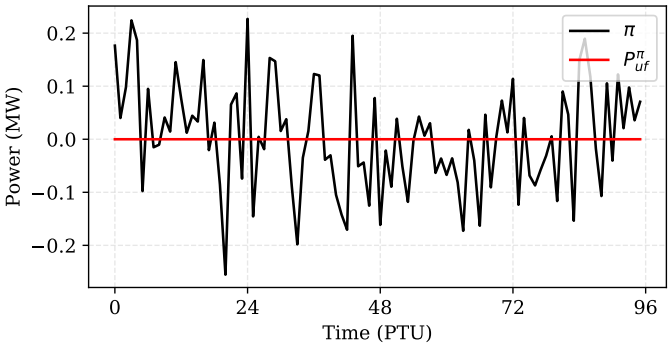

Figure 14: π^7 and P^7_{UF}. The unserved flexibility is zero, which implies system is able to fulfill the FR request.

4.8. Economic Analysis

While EUFE indicates the volume of the flexible energy, the value for FSP in using the proposed method to design DRPs can be best motivated by analyzing the unforeseen costs it would incur due to inaccurate flexibility assessment. In the UK, according to ref. [38], the average imbalance price in 2019 was 57.06 €/MWh. The Minkowski summation method for flexibility quantification suggested the portfolio flexibility was sufficient to fulfill every FR signal request at each time step. However, using the simulation-based approach, we determined that the EUFE is 1.858 MWh. For the FSP, this directly translates to a yearly cost of € 38696. With tuning of DRP, as shown with C2 and C3, the FSP can reduce the EUFE to 0.746 MWh and 0.608 MWh respectively. Consequently, its unforeseen costs decrease to € 15536 and € 12662 respectively. When FSP engages in timely intraday or day-ahead market participation to correct these expected insufficiencies in the flexibility calculated in C3, then as shown in

Table 2: Flexibility metric values for different cases

Case	Characteristics	EUFE	EFI	Cost (€)
C1	Baseline	1.858	0.906	38696
C2	Additional non-targeted DR	0.746	0.934	15536
C3	Additional targeted DR	0.608	0.948	12662
C4	IDM with risk-neutral VPP	0.077	0.987	1603
C5	IDM with risk-averse VPP	0.0	1.00	0

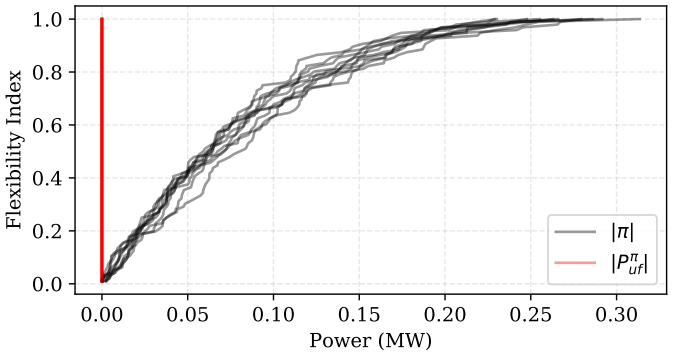

Figure 15: The ECDF for the signals in Fig. 14 shows the flexibility index is 1, implying a totally flexible system.

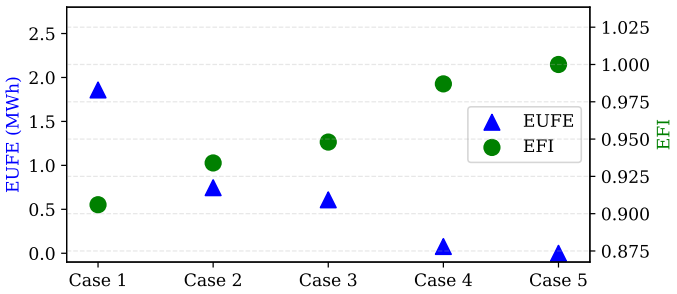

Figure 16: Comparison of EUFE and EFI for all 5 cases.

C4 and C5, the EUFE and consequently its costs are further reduced to € 1603 and € 0 respectively, a negligible amount compared to C1. The economic results are summarized in Fig. 17 and tabulated alongside the technical results in Table 2. This analysis shows that the proposed method offers the FSP a significant economic benefit, even in highly volatile balancing markets.

5. Discussion

Flexibility is an unavoidable necessity of the future power system and FSPs will play a major role in ensuring its availability. The ability to quantify flexibility is therefore necessary for the FSP to operate. In this research, we focus on accurate quantification of short-term operational flexibility considering peculiarities arising from activation of flexibility, particularly those from demand response.

A unique advantage of the proposed metrics is their usefulness in comparing resource flexibility that are defined in unique ways. For example, in C2 and C3, the FSP can compare the value of flexibility provided by an altering temperature bands for electricity grid-connected building compared to altering temperature bands for all thermal loads, two of which are connected to heat grid and therefore coupled to electricity grid indirectly. Such a comparison can also be conducted for comparing flexibility from resources with vastly different operational characteristics, such heat pump operation in a heat grid and a battery system in the electric grid. The introduced metrics allow FSPs to quickly assess contributions of such varied forms of flexibility intuitively and therefore design appropriate DRPs. This is shown in this paper by comparing flexibility derived from time-dependent and time-independent temperature bands for thermal loads.

In C4 and C5, it is seen how the three metrics are useful for deciding participation in intraday markets. It is a well known fact that prices of electricity closer to gate closure time can be highly unpredictable, and in many cases, unusually high. The earlier a trade is made, the lower is the uncertainty in price. Since the EUFE, EDIF heatmap are calculated offline, the information derived from these metrics can be used to procure any additional flexibility required well ahead of time.

However, there are a few limitations to the proposed method. The first one is computational complexity. The metrics are computed using scenario-based simulations, which involve executing FSP's operational policy subject to individual resource dynamics and constraints. Firstly, we assume that the model of the system is an accurate representation of a real physical system. Any simulation-based approach suffers from modelling assumptions and inaccuracies when compared to reality. Secondly, if addi-

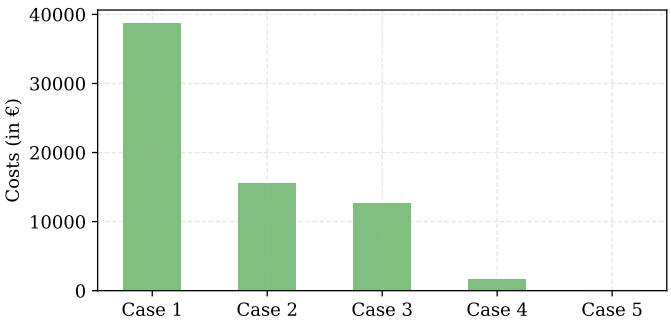

Figure 17: Comparison of unforeseen imbalance penalties for the FSP when using Minkowski Summation method (C1) and proposed method for flexibility assessment (C2, C3, C4, and C5).

tional complexity in models is included for more realistic representation (for example, including network constraints such as heat network transport delays, electric network voltage violations, etc.) the simulation problem can be hard to solve and becomes time-consuming. To overcome this, scenario simulations must be well formulated. In this paper, we considered simple resource dynamics and network constraints to show the applicability and usefulness of the metrics. The non-linearity in the objective formulation was reformulated to create a standard MILP problem.

A second limitation of the method is that the quantification process depends on the accuracy of scenarios of the FR signals. While the forecasted signals will never be 100% accurate and will introduce uncertainties in the results, using advanced scenario generation and reduction techniques can partially offset these limitations.

An interesting point of discussion in the presented approach is the choice of the objective function for the operational simulation. We proposed to minimize the absolute value of the UF signal, which means, we aim to minimize the absolute amount of energy in the UF signal, ignoring the direction component of the UF signal. This is a deliberate choice made by the authors for two reasons. One, the imbalance costs are associated with the magnitude of imbalance. The type of payment to be made (either from system to BRP or BRP to system) depends on BRP's imbalance relative to overall power system imbalance (which is a highly uncertain quantity). If the magnitude of imbalance is minimized, the payment (in either direction) can automatically be minimized. Two, consideration of direction appends additional complexity to the quantification process. Although it has not been studied here, the direction of the UF signal can be included in the objective formulation as required without affecting the quantification process.

In the future, the research will focus on defining appropriate values of EFI, EDIF, and EUFE for a given portfolio and objective. Additional work also needs to be done to evaluate different operational strategies and objectives for an FSP with FR signal scenarios generated using sophisticated state-of-art methods.

6. Conclusions

The paper focused on the problem of quantification of operational flexibility from a portfolio of flexible resources. Firstly, the concept of available flexibility and operational flexibility in power systems was clearly defined and differentiated. Next, existing methods to quantify flexibility were shown to be inadequate. We then proposed three new metrics: EUFE, EDIF, and EFI to measure portfolio flexibility. The metrics were calculated using scenario-based simulations and are shown to assist the FSPs to design appropriate DRP, compare flexibility from various resources, formulate operational policies, and identify potential insufficiency in flexibility provision from a portfolio. A case study consisting of a multi energy system was used to show how the proposed metrics can be used to enhance the flexibility servicing capability of the FSP. A short economic assessment was also conducted to highlight the value of the proposed metrics to the FSP.

Appendix A. Empirical Cumulative Distribution Function

ECDF is a common non-parametric estimator used in exploratory data analysis. Traditionally, it is used to determine the underlying CDF of a dataset.

For a given dataset, $\{x_1, ..., x_n\}$, the Empirical Cumulative Distribution Function (ECDF) is given mathematically by Eq. (A.1).

$$\hat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} I\{x_i \leq x\} \tag{A.1}$$

where $I\{\cdot\}$ is the indicator function. It has two possible values:

$$I(x_i \leq x) = \begin{cases} 1 & \text{when } x_i \leq x \\ 0 & \text{when } x_i > x \end{cases} \tag{A.2}$$

Appendix B. Optimization Problem

This section describes the constraints for optimization problem described in Section 4.3. These constraints form the baseline DRP formulated in C1. The equality constraints are:

$$\dot{Q}_{GH} + \dot{Q}_B = \dot{Q}_{TES} + \dot{Q}_{EB} + \dot{Q}_{PB} \tag{B.1}$$

$$P_{Fi} + P_{EB} + P_{HP} + \pi_i = P_{FC} + P_{EG} + P_{ID} \tag{B.2}$$

$$S_{TS}(t+1) = S_{TS}(t) - \dot{Q}_{TS}(t) - \dot{Q}_l(t) \tag{B.3}$$

$$\theta(t+1) = \theta(t) + \frac{\dot{Q}(t) - \dot{Q}_l(t)}{\dot{m} \cdot c} \tag{B.4}$$

$$\dot{Q}_{EB} = P_{EB} \cdot \eta_{EB} \tag{B.5}$$

$$\dot{Q}_B = \dot{Q}_{HP} = P_{HP} \cdot COP_{HP} \tag{B.6}$$

$$\sum_{t=0}^{N_t} P_{Fi}^{DA} - \sum_{t=0}^{N_t} P_{Fi} = 0 \tag{B.7}$$

$$\sum_{t=0}^{N_t} \tilde{\dot{Q}}_{PB}^{DA} - \sum_{t=0}^{N_t} \dot{Q}_{PB} = 0 \tag{B.8}$$

Equations (B.1) and (B.2) represent the thermal and electrical power balance constraints respectively. Constraints in Eqs. (B.3) and (B.4) represent the resource dynamics for thermal energy storage and general temperature evolution dynamics for building and greenhouses. η_{EB} represents the electrical boiler efficiency, while COP_{HP} represents the coefficient of performance of the HP. Constraints in Eqs. (B.5) and (B.6) relate the electrical and thermal powers for electric boiler and heat pump. Here, $\dot{Q}_B$ represents

the power consumption of the building. Equation (B.7) forms the load shifting constraint in electrical loads in feeders F1 and F2. It ensures that no load shedding occurs by operating the loads flexibly. Equation (B.8) defines flexibility in operating PB. Since the amount of gas to operate the PB was brought on D-1 in the gas market, its operation is fuel constrained. Assuming a linear relationship between fuel used, power output, and hours of operation of the PB, Eq. (B.8) ensures that the rescheduling of the gas-based PB uses the same amount of gas as contracted in the day-ahead market on D-1.

In addition to the equality constraints, Eq. (8) is also subject to the following inequality constraints:

$$\underline{\theta}_{GH} \leq \theta_{GH} \leq \overline{\theta}_{GH} \tag{B.9}$$

$$\underline{\theta}_{B} \leq \theta_{B} \leq \overline{\theta}_{B} \tag{B.10}$$

$$\underline{P}_{Fi} \leq P_{Fi} \leq \overline{P}_{Fi} \tag{B.11}$$

$$r^{\downarrow}_{Fi} \leq P_{Fi}(t+1) - P_{Fi}(t) \leq r^{\uparrow}_{Fi} \tag{B.12}$$

$$0 \leq P_{FC} \leq \overline{P}_{FC} \tag{B.13}$$

Equations (B.9) and (B.10) represent the upper and lower bounds on thermal load (GH-A, GH-B, B) temperatures, Eqs. (B.11) and (B.12) represents the bounds on power and ramp rates on feeders F1 and F2, and finally, Eq. (B.13) describes the bounds on the power output of FC. The *abs*() in Eq. (8) introduces non-linearity in the optimization problem. This can however, easily be linearized by introducing a slack variable in the problem formulation. The reformulated problem then becomes a Mixed Integer Linear Programming (MILP) problem which is easily solvable.

Acknowledgment

The authors would like to thank Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) for the financial support of the project Heat and Power Systems at Industrial Sites and Harbors (HaPSISH).

References

[1] M. Derks, Annual Market Update 2018 (2018).
URL https://www.tennet.eu/fileadmin/user_upload/Company/Publications/Technical_Publications/Dutch/Annual_Market_Update_2018_-_Final.pdf

[2] H. Hao, B. M. Sanandaji, K. Poolla, T. L. Vincent, Aggregate Flexibility of Thermostatically Controlled Loads, IEEE Transactions on Power Systems 30 (1) (2015) 189–198, conference Name: IEEE Transactions on Power Systems. doi:10.1109/TPWRS.2014.2328865.

[3] A. Lewandowska-Bernat, U. Desideri, Opportunities of Power-to-Gas technology, Energy Procedia 105 (2017) 4569–4574. doi:10.1016/j.egypro.2017.03.982.
URL https://linkinghub.elsevier.com/retrieve/pii/S1876610217310974

[4] J. P. Chaves-Ávila, R. A. Hakvoort, A. Ramos, Short-term strategies for Dutch wind power producers to reduce imbalance costs, Energy Policy 52 (2013) 573–582. doi:10.1016/j.enpol.2012.10.011.
URL http://www.sciencedirect.com/science/article/pii/S0301421512008695

[5] H. T. Nguyen, L. B. Le, Z. Wang, A Bidding Strategy for Virtual Power Plants With the Intraday Demand Response Exchange Market Using the Stochastic Programming, IEEE Transactions on Industry Applications 54 (4) (2018) 3044–3055. doi:10.1109/TIA.2018.2828379.

[6] A. van der Veen, M. van der Laan, H. de Heer, E. Klaassen, W. van den Reek, USEF Flexibility Value Chain, Tech. rep. (2018).
URL https://www.usef.energy/app/uploads/2018/11/USEF-White-paper-Flexibility-Value-Chain-2018-version-1.0_Oct18.pdf

[7] Peeeks - Eneco.
URL https://www.eneco.nl/over-ons/wat-we-doen/duurzame-bronnen/peeeks/

[8] L. Zhao, W. Zhang, H. Hao, K. Kalsi, A Geometric Approach to Aggregate Flexibility Modeling of Thermostatically Controlled Loads, IEEE Transactions on Power Systems 32 (6) (2017) 4721–4731, arXiv: 1608.04422. doi:10.1109/TPWRS.2017.2674699.
URL http://arxiv.org/abs/1608.04422

[9] E. Polymeneas, S. Meliopoulos, Aggregate modeling of distribution systems for multi-period OPF, in: 2016 Power Systems Computation Conference (PSCC), 2016, pp. 1–8. doi:10.1109/PSCC.2016.7540987.

[10] M. P. Evans, S. H. Tindemans, D. Angeli, A Graphical Measure of Aggregate Flexibility for Energy-Constrained Distributed Resources, IEEE Transactions on Smart Grid 11 (1) (2020) 106–117. doi:10.1109/TSG.2019.2918058.

[11] M. A. Bucher, S. Delikaraoglou, K. Heussen, P. Pinson, G. Andersson, On quantification of flexibility in power systems, in: 2015 IEEE Eindhoven PowerTech, IEEE, Eindhoven, Netherlands, 2015, pp. 1–6. doi:10.1109/PTC.2015.7232514.
URL http://ieeexplore.ieee.org/document/7232514/

[12] J. Zhao, T. Zheng, E. Litvinov, A Unified Framework for Defining and Measuring Flexibility in Power System, IEEE Transactions on Power Systems 31 (1) (2016) 339–347. doi:10.1109/TPWRS.2015.2390038.

[13] Z. Chen, L. Wu, Y. Fu, Real-Time Price-Based Demand Response Management for Residential Appliances via Stochastic Optimization and Robust Optimization, IEEE Transactions on Smart Grid 3 (4) (2012) 1822–1831. doi:10.1109/TSG.2012.2212729.

[14] M. Wei, J. Zhong, Optimal bidding strategy for demand response aggregator in day-ahead markets via stochastic programming and robust optimization, in: 2015 12th International Conference on the European Energy Market (EEM), 2015, pp. 1–5, iSSN: 2165-4093. doi:10.1109/EEM.2015.7216732.

[15] J. Ma, V. Silva, R. Belhomme, D. S. Kirschen, L. F. Ochoa, Evaluating and Planning Flexibility in Sustainable Power Systems, IEEE Transactions on Sustainable Energy 4 (1) (2013) 200–209. doi:10.1109/TSTE.2012.2212471.
URL http://ieeexplore.ieee.org/document/6313967/

[16] N. Menemenlis, M. Huneault, A. Robitaille, Thoughts on power system flexibility quantification for the short-term horizon, in: 2011 IEEE Power and Energy Society General Meeting, 2011, pp. 1–8, iSSN: 1944-9925. doi:10.1109/PES.2011.6039617.

[17] T. Heggarty, J.-Y. Bourmaud, R. Girard, G. Kariniotakis, Quantifying power system flexibility provision, Applied Energy 279 (2020) 115852. doi:10.1016/j.apenergy.2020.115852.
URL https://www.sciencedirect.com/science/article/pii/S030626192031326X

[18] H. Nosair, F. Bouffard, Flexibility envelopes for power system operational planning, IEEE Transactions on Sustainable Energy 6 (3) (2015) 800–809, publisher: IEEE.

[19] H. Nosair, F. Bouffard, Economic dispatch under uncertainty:

The probabilistic envelopes approach, IEEE Transactions on Power Systems 32 (3) (2016) 1701–1710, publisher: IEEE.

[20] H. Nosair, F. Bouffard, Reconstructing operating reserve: Flexibility for sustainable power systems, IEEE Transactions on Sustainable Energy 6 (4) (2015) 1624–1637, publisher: IEEE.

[21] P. Lütolf, M. Scherer, O. Mégel, M. Geidl, E. Vrettos, Rebound effects of demand-response management for frequency restoration, in: 2018 IEEE International Energy Conference (ENERGYCON), 2018, pp. 1–6. doi:10.1109/ENERGYCON.2018.8398849.

[22] D. Gusain, M. Cvetković, P. Palensky, Energy Flexibility Analysis using FMUWorld, in: 2019 IEEE Milan PowerTech, 2019, pp. 1–6. doi:10.1109/PTC.2019.8810433.

[23] D. Gusain, M. Cvetković, P. Palensky, Energy Flexibility Analysis using FMUWorld, in: 2019 IEEE Milan PowerTech, 2019, pp. 1–6. doi:10.1109/PTC.2019.8810433.

[24] P. Palensky, D. Dietrich, Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads, IEEE Transactions on Industrial Informatics 7 (3) (2011) 381–388. doi:10.1109/TII.2011.2158841.
URL http://ieeexplore.ieee.org/document/5930335/

[25] R. Yin, E. C. Kara, Y. Li, N. DeForest, K. Wang, T. Yong, M. Stadler, Quantifying flexibility of commercial and residential loads for demand response using setpoint changes, Applied Energy 177 (2016) 149–164. doi:10.1016/j.apenergy.2016.05.090.
URL http://www.sciencedirect.com/science/article/pii/S0306261916306870

[26] A. Ulbig, G. Andersson, Analyzing operational flexibility of electric power systems, International Journal of Electrical Power & Energy Systems 72 (2015) 155–164. doi:10.1016/j.ijepes.2015.02.028.
URL http://www.sciencedirect.com/science/article/pii/S0142061515001118

[27] E. Lannoye, D. Flynn, M. O'Malley, Evaluation of Power System Flexibility, IEEE Transactions on Power Systems 27 (2) (2012) 922–931. doi:10.1109/TPWRS.2011.2177280.
URL http://ieeexplore.ieee.org/document/6125228/

[28] E. Lannoye, P. Daly, A. Tuohy, D. Flynn, M. O'Malley, Assessing Power System Flexibility for Variable Renewable Integration: A Flexibility Metric for Long-Term System Planning, CIGRE Science and Engineering Journal 3 (2015) 26–39, accepted: 2016-12-09T15:47:30Z Publisher: CIGRE.
URL https://researchrepository.ucd.ie/handle/10197/8201

[29] R. Mena, M. Hennebel, Y.-F. Li, C. Ruiz, E. Zio, A risk-based simulation and multi-objective optimization framework for the integration of distributed renewable generation and storage, Renewable and Sustainable Energy Reviews 37 (2014) 778–793. doi:10.1016/j.rser.2014.05.046.
URL https://linkinghub.elsevier.com/retrieve/pii/S1364032114003712

[30] S. Goleijani, T. Ghanbarzadeh, F. Sadeghi Nikoo, M. Parsa Moghaddam, Reliability constrained unit commitment in smart grid environment, Electric Power Systems Research 97 (2013) 100–108. doi:10.1016/j.epsr.2012.12.011.
URL http://www.sciencedirect.com/science/article/pii/S037877961200363X

[31] S. Talari, M. Haghifam, The impact of load and distributed energy resources management on microgrid reliability, in: 22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013), Institution of Engineering and Technology, Stockholm, Sweden, 2013, pp. 1035–1035. doi:10.1049/cp.2013.1045.
URL https://digital-library.theiet.org/content/conferences/10.1049/cp.2013.1045

[32] I. Bae, J. Kim, Reliability Evaluation of Customers in a Microgrid, IEEE Transactions on Power Systems 23 (3) (2008) 1416–1422. doi:10.1109/TPWRS.2008.926710.

[33] C. M. Correa-Posada, G. Morales-Espana, P. Duenas, P. Sanchez-Martin, Dynamic Ramping Model Including Intraperiod Ramp-Rate Changes in Unit Commitment, IEEE Transactions on Sustainable Energy 8 (1) (2017) 43–50. doi:10.1109/TSTE.2016.2578302.
URL http://ieeexplore.ieee.org/document/7487051/

[34] European Commission. Directorate General for Mobility and Transport., Ref4e., Mercados., E bridge., Identification of appropriate generation and system adequacy standards for the internal electricity market: final report., Publications Office, LU, 2014.
URL https://data.europa.eu/doi/10.2832/089498

[35] B. Leitner, E. Widl, W. Gawlik, R. Hofmann, A method for technical assessment of power-to-heat use cases to couple local district heating and electrical distribution grids, Energy 182 (2019) 729–738. doi:10.1016/j.energy.2019.06.016.
URL http://www.sciencedirect.com/science/article/pii/S0360544219311399

[36] B.-M. Hodge, D. Lew, M. Milligan, H. Holttinen, S. Sillanpaa, E. Gómez-Lázaro, R. Scharff, L. Soder, X. G. Larsén, G. Giebel, Wind power forecasting error distributions: An international comparison, Tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States) (2012).

[37] W. E. Hart, C. D. Laird, J.-P. Watson, D. L. Woodruff, G. A. Hackebeil, B. L. Nicholson, J. D. Siirola, Pyomo — Optimization Modeling in Python, 2nd Edition, Springer Optimization and Its Applications, Springer International Publishing, 2017. doi:10.1007/978-3-319-58821-6.
URL https://www.springer.com/gp/book/9783319588193

[38] D. Kirli, M. Parzen, A. Kiprakis, Impact of the COVID-19 Lockdown on the Electricity System of Great Britain: A Study on Energy Demand, Generation, Pricing and Grid Stability, Energies 14 (3) (2021) 635, number: 3 Publisher: Multidisciplinary Digital Publishing Institute. doi:10.3390/en14030635.
URL https://www.mdpi.com/1996-1073/14/3/635