





Robust flexibility needs assessment with bid matching framework for distribution network operators

Md Umar Hashmi, *Member, IEEE* , Arpan Koirala, *Graduate Student Member IEEE* , Hakan Ergun, *Senior Member, IEEE* ,
and Dirk Van Hertem, *Senior Member, IEEE* 

Abstract

For market-based procurement of low-voltage flexibility, DSOs identify the amount of flexibility needed for resolving probable distribution network (DN) voltage and thermal congestion. A framework is required to avoid over or under-procurement of flexibility in the presence of uncertainty. To this end, we propose a scenario-based robust chance-constrained (CC) day-ahead flexibility needs assessment (FNA) framework. The CC level is analogous to the risk DSO is willing to take in flexibility planning. Multi-period optimal power flow is performed to calculate the amount of flexibility needed to avoid network issues. Future uncertainties are considered as multiple scenarios, which are utilized to solve the flexibility needs assessment optimal power flow (FNA-OPF) problem. The FNA tool calculates the temporal and locational ramp-up and ramp-down flexibility needs of the DN. We also propose a Pareto optimal mechanism for selecting CC level to reduce flexibility needs while reducing DN congestion. Zonal clustering of a DN feeder is performed using electrical distance as a measure, with spatial partitioning and silhouette coefficient. Further, a zonal and nodal bid matching algorithm is proposed, where alternate flexibility bids in the same zone are identified in case nodal matching is not exact, compared to the calculated FNA. The proposed bid matching algorithm assists the system operator in bid selection by avoiding rerunning power flow evaluations. The zonal and nodal bid matching leads to more than 43% additional reduction of probable network incidents compared

Corresponding author email: mdumar.hashmi@kuleuven.be

M.U. Hashmi, A. Koirala, H. Ergun, and D. Van Hertem are with KU Leuven, division Electa & EnergyVille, Genk, Belgium

This work is supported by the H2020 EUniversal project, grant ID: 864334 (<https://euniversal.eu/>) and the energy transition funds project BREGILAB organized by the FPS economy, S.M.E.s, Self-employed and Energy.

to naive nodal bid matching. Furthermore, we also calculate FNA for a real German DN with 646 nodes and 331 consumers.

Index Terms

Flexibility, Scenario generation, Optimal power flow (OPF), Distribution system operator (DSO), Zonal clustering, Bid matching, Chance constraint, Risk-based planning

CONTENTS

I	Introduction	6
I-A	Literature review	7
I-A2	Stochastic optimization based flexibility operation	8
I-A3	Zonal clustering and bid matching	9
I-B	Contributions and observations	9
II	Identifying zones of an LV DN	11
II-A	Distance measure for clustering	11
II-B	Spectral partitioning	12
II-C	Goodness of a cluster	12
III	Day-ahead flexibility needs assessment	13
III-A	Flexibility: definition	13
III-B	Building blocks for day-ahead FNA	14
III-C	Notation used	15
III-D	Optimization formulation	15
III-E	Robust FNA under uncertainty with risk	16
III-F	Pareto optimal tuning of chance constraint level	17
III-G	Pseudocode for calculating FNA	18
IV	Bid matching algorithm	20
V	Numerical case studies	23
V-A	Case study 1: Flexibility needs assessment	24
V-A1	Projecting CC levels on network issues	28
V-A2	Pareto optimal tuning of CC level	29
V-B	Case study 2: Marginal value of energy and power	31
V-C	Case study 3: Bid matching	32
V-D	Case study 4: FNA on a large DN	35
VI	Conclusion and future works	37
	References	41

Appendix		46
A	Multivariate distribution and Cholesky decomposition	46
B	Temporal and spatially correlated scenarios	47

Abbreviation

CC	Chance constraint
DA	Day-ahead
DG	Distributed Generation
DN	Distribution Network
DNI	Distribution Network Incidents
DSO	Distribution System Operator
ECDF	Empirical cumulative distribution function
FNA	Flexibility needs assessment
FSP	Flexibility service provider
LV	Low Voltage
OPF	Optimal power flow

I. INTRODUCTION

Distribution Networks (DN) are undergoing an abrupt increase in new loads, such as electric vehicles and distributed generations (DG), such as rooftop PV installations, fueled by new policies, e.g., fit for 55 [1]. The uncertainty in DN operation due to DG and loads with high simultaneity factors are increasing, and distribution system operators (DSOs) must plan flexible resources to ensure the reliable operation of DN [2]. ENTSO-e defines power system flexibility "as the ability of the power system to cope with variability and uncertainty in demand, generation and grid capacity while maintaining a satisfactory level of reliability at all times" [3]. Flexibility assists power networks to damp fluctuations, that could potentially increase the reliability of the DN [4], [5].

NOMENCLATURE

Operators

$\mathbf{I}_{(-\infty, z]}(y)$ Indicator operator with output equal to 1 if $y \leq z$

$F_{i,t,+}(z)$ ECDF for ramp down FNA for all scenarios

$F_{i,t,-}(z)$ ECDF for ramp up FNA for all scenarios ‘

E Set of edges in the DN graph

$i \in \{1, ..N\}$ Set of nodes in the DN

N_G Set of nodes with generator connected

N_L Set of nodes with load connected

$t \in \{1, .., T\}$ Time ‘

Variables and parameters

$(\Delta P_{i,t}^{\text{flex}+})^*, (\Delta P_{i,t}^{\text{flex}-})^*$ Ramp down and up FNA at time t , node i considering all scenarios

$\Delta P_{i,t}^{\text{flex}+}, \Delta P_{i,t}^{\text{flex}-}$ Ramp down and up flexibility power activated at time t , node i

$\Delta P_{i,t,\max}^{\text{flex}+}$ Ramp down flexibility power limit

Δt Sampling duration

ϵ_{cc} Chance constrain level

$\lambda^{\text{rampDown}}, \lambda^{\text{rampUp}}$ Ramp down and ramp up flexibility cost

$E_i^{\text{flex}+}, E_i^{\text{flex}-}$ Ramp down and up flexibiliy energy activated at time t , node i

$E_{i,\max}^{\text{flex}+}$ Ramp down flexibility energy limit

The challenge is that these new loads and generations in low voltage (LV) networks can lead to network congestion issues, which traditionally were solved using network upgrading. On the other hand, these loads and generation can operate with a level of coordination or flexibility to avoid congestion in the local network as well as at the system level. However, there is no ready-to-use solution currently available that the DN operators can use to utilize such flexible resources available in their network, remunerate the participants of the market as well as ensure the healthy operations of the network. A framework is needed for the quantification of flexibility needs while considering load and generation uncertainties. How should the DSO use the results of the calculation of the flexibility needs to choose the best flexibility bids? In this context, the goal of this paper is twofold. Firstly, to propose a flexibility need assessment (FNA) framework to quantify the amount of *temporal* and *locational* flexibility needed by the distribution system operator while considering uncertainty in load profiles and DER generation. Secondly, to present an algorithm for DN operators to select the bids in the flexibility market using the quantified FNA. This tool was part of a flexibility framework demo under the H2020 EUuniversal project [6]. The proposed FNA framework and bid matching algorithm would provide recommendations to the DSO to *plan* and *procure* flexibility in day-ahead (DA) setting for avoiding DN voltage and thermal congestion incidents.

A. Literature review

The review of the existing works is done under three headings: a) DN flexibility and their use case, b) stochastic optimization used in flexibility context, and c) bid matching algorithms.

Activation of flexibility in LV DN could aim at (a) mitigating LV DN issues [7] and/or (b) providing services to MV and HV transmission or DN [8]. The latter would require coordination in the operation of flexible resources, ensuring it does not create new problems [9]. In [10], it is shown that small electricity prosumers can utilize their flexible resources for providing grid services while making a profit in the DA market. Authors in [11] schedule flexible resources for DA energy dispatch, which minimizes the energy procurement cost.

Authors in [12] propose a computational method to calculate prosumer multi-period flexibility forecast. Using these feasible flexibility spaces, and prosumers can participate in the energy market [13]. In [14], a case study of Northern Germany is presented, where every MW of additional wind installation will require 0.7 MW of flexibility to avoid curtailment of renewable generation. Prior works [15]–[17] present market mechanisms for end-user flexible resources for solving LV network

issues. The market mechanism presented in [15] utilizes contractual flexibility by DSOs for the H2020 Interflex project. Authors in [16] propose an efficient design for flexibility markets. Different types of flexibility market designs are detailed in [18] and [19]. Recent works such as [20]–[23] discuss the need for flexibility provisioning for planning problems. In [20], FNA is utilized for the prioritization of future investments. Authors in [21], [23] quantify flexibility provisions based on the conditional value at risk. In [24], authors observe that locational and temporal flexibility mapping is crucial for avoiding future network issues. They propose a mapping tool for Northern Ireland via GIS analysis.

The procurement of flexible resources in the DA market can be costly. Under-procuring can lead to otherwise avoidable distribution network incidents (DNIs). Traditionally these DNIs are mitigated through unplanned load shedding and/or distributed renewable energy curtailment [25]. On the other hand, over-procuring of flexible resources may not be efficient due to high reservation costs [26]. FNA should consider the uncertainties in the network. In [27], an active DN with uncertainties are taken to demonstrate the role of flexibility in improving the network operation indices using robust stochastic optimization. In [28], a stochastic method to obtain a dispatch plan at the common point of coupling between TSO and DSO is presented, considering flexibility from photovoltaics as DG and energy storage.

Most prior works in the literature have focused on a specific aspect of needs assessment for DN. The key research gap we are filling in the paper is that we are providing a complete framework consisting of scenario generation, network clustering, optimization-based needs assessment, the consideration of risk metrics and bid matching algorithm for flexibility procurement.

2) *Stochastic optimization based flexibility operation:* Power system optimization under uncertainty itself has a broad range of concepts and is detailed in [29]. Deterministic optimal power flow (OPF) gives optimal results for a snapshot of all possible scenarios, making it not useable for flexibility procurement applications under uncertainties. A robust, viable solution is identified using robust optimization over a given uncertainty set. It is unclear how we can define such an uncertainty set for the FNA. Thus, we utilize generated scenarios for creating bounds on uncertainty. The flexibility needs values that provide a feasible solution for all scenarios will often lead to an over-procurement of flexible resources. These values of flexibility needs will be robust in the true sense [30]. Most recent work using stochastic optimization uses robust optimization [27],

[28]. However, the interest in DN-FNA is to avoid over-procurement, chance constraints (CC) optimization could be a preferred method. The CC can be implemented analytically, [31]–[33], or by using some linear reformulation [34]. Applying CC provides a risk-averse flexibility needs of the DN [35]. This is identified by using non-parametric empirical cumulative distribution functions (ECDF) for each node of the network and for each time instant. So that extreme tail events are considered in either direction when calculating DA flexibility, the ramp-up and ramp-down flexibility needs are quantified independently. The CC level is analogous to the risk DSO is willing to take in procuring flexible resources in DA flexibility market.

3) *Zonal clustering and bid matching*: DNI in LV DN are often a local problem in which flexible resources in the proximity respond to avoid these DNIs. To that end, we cluster DN onto *connected* zones. The proposed zonal bid selection algorithm finds alternate flexibility bids in a zone if the nodal matching of bids is not met entirely. The hypothesis is that the flexibility bids in the neighborhood of nodes where DNI is observed could effectively mitigate such DNIs. In the literature, power network clustering is widely utilized for the operation and planning of transmission and DNs [36]–[39]. In [36], [37], network clustering is used for transmission system operator-level resource dispatch and planning. Authors in [38] use zonal voltage regulation in the context of an LV DN with high penetration of DG. The proposed framework for zonal clustering partitions the DN in connected segments, a key challenge in other partitioning methods. These zones are used to aggregate nodal FNA. Flexibility needs for MV DN zones is identified in [40]. The resources offered in the flexibility market frequently fall short of precisely meeting needs for flexibility by the DN. To this end, we propose nodal and zonal bid matching for finding alternate flexibility bids for solving DNIs.

B. Contributions and observations

The key contributions of the paper are:

- *Framework for flexibility needs assessment*: Using the scenario generation¹ detailed in Appendix A, FNA-OPF is solved. The spatial and temporal robust flexibility needs of a DN are identified for avoiding all DNIs for all scenarios. For risk-averse FNA, a chance-constrained nodal FNA is calculated. We propose a Pareto optimal mechanism for selecting the level of CC used in deciding

¹The scenario generation takes as input PV size in kWp installed at a node, forecasted nodal load profiles, forecast of normalized PV generation, and the associated forecast errors for load and solar generation. The forecast error creates a boundary of uncertainty.

the DN FNA that considers DSO's risk of over and under-procurement of flexible resources, while reducing the unavoided DNIs. Further, the nodal FNA is aggregated in a zone to identify zonal FNA.

- *Algorithm for bid matching:* In the flexibility market, bids will be gathered by the flexibility service provider. The matching of FNA and flexibility bids at the nodal level will often lead to partial matching. In this case, DSO will not be able to run 100s of power flows for optimal bid selection. To solve this problem, we propose zonal bid matching preceded by nodal matching.

- *Observations from four numerical case studies:*

- The first case study shows the FNA output of a DN. A Pareto optimal mechanism is provided for tuning CC level. The selection of CC level is projected onto unavoided DNIs using power flow simulations. Nodal FNA is aggregated to form zonal FNA.

- The second case study quantifies the marginal value of flexibility based on energy or power needs. Energy and power needs are often valued differently in many energy markets. It is observed that power needs are twice as important as the energy needs.

- The third case study projects flexibility bid matching onto probable DNI. We observe that the proposed bid matching leads to an additional 6.5% (on average) probable voltage incident reduction and up to 28% reduction in thermal incidents for 20% of load and generation used as flexibility bids.

- In the last case study, we identified FNA for a real suburban German DN with 646 nodes and 331 consumers.

This paper is organized as follows. In Section II, zonal clustering of a DN based on electrical distance is provided. In Section III, the metrics used in FNA are quantified and the FNA-OPF problem is formulated. Section IV summarizes the bid-matching algorithm based on the nodal and zonal FNA. Section V presents the numerical case studies. Section VI concludes the paper.

II. IDENTIFYING ZONES OF AN LV DN

Identifying the zones of an LV DN will help the DSO in planning the flexibility needs of a network. Although the analytical division of a DN is expected to work well, however, due to the sheer number of DN feeders, it becomes crucial to have a standardized framework for dividing DN into zones based on electrical and/or geographical distances. For example, in the UK, there are more than 1.3 million LV feeders [41]. In this section, we develop a clustering framework to identify the best-suited LV DN zonal partition using electrical distance as a measure.

The proposed zone formation solves the following challenges: (i) the formation of connected zones requires an incidence matrix-based measure, therefore, we consider admittance as a measure. (ii) the admittance matrix cannot be used directly, therefore, spectral decomposition of a doubly stochastic matrix is used, and (iii) the selection of the appropriate number of zones is not known a priori. In order to apply unsupervised clustering techniques such as k -means, one should know how many clusters are needed. We use the silhouette score as a measure for identifying the best number of clusters. Although the different methods mentioned above have been well discussed in the literature, in this work, they have been, for the first time, combined to cluster the DN into zones in the context of FNA.

A. Distance measure for clustering

A DN can be represented in standard directed graph notation $G = (N, E, W)$, where N denotes the set of nodes such that $|N| = D$ (D denotes the total number of nodes), E denotes edges and W denotes edge weight matrix. We aim to divide N into p groups $\{M_1, M_2, \dots, M_p\}$ such that $M_i \subset N$ and $M_i \cap M_j = \emptyset, \forall i \neq j$. The edge weights can be considered a penalty for cutting that line in the graph while clustering. It also measures the connection strength of two nodes in a graph. A good measure would separate the weakly connected nodes and identify nodes that should be clustered together. A power network can be clustered based on static network parameters such as line impedance/admittance or dynamic parameters such as power flows, line losses, and voltage variations [37], [39], [42]. The electrical distance matrix is used for partitioning LV DNs in [36], [37]. Similar to these prior works, we use the line admittance matrix as the distance measure parameter. The weight $w_{ij} = Y_{ij} = 1/|R_{ij} + jX_{ij}|$, where Y_{ij} , R_{ij} and X_{ij} denotes admittance, resistance and impedance between nodes i and j and the edge weight $w : N \times N \rightarrow \mathbb{R}^{\geq 0}$ such that (a) $w_{ij} = w_{ji}, \forall i, j$, (b) $w_{ij} = 0$, if $(i, j) \notin E$, (c) $w_{ii} = 0, \forall i$.

B. Spectral partitioning

Spectral clustering is used for power network partitioning or creation of zones or network reduction in [37], [38]. The previously described weight matrix cannot be directly used, as the diagonal elements are all zeros. Authors in [37] use the normalized Laplacian matrix for spectral partitioning. Tutorial [43] transforms the weight matrix into a doubly stochastic matrix. A double stochastic matrix is a special type of Markov matrix where not only each row but also each column adds to 1. Spectral properties of doubly stochastic matrices have been detailed in [44]. For this transformed matrix, all eigenvalues are real and smaller than or equal to 1, with one eigenvalue exactly equal to 1. For identifying k partitions in a graph, the k highest eigenvalues and corresponding orthonormal eigenvectors are identified. The eigenvector matrix of the order $N \times k$ is used for DN partitioning, in effect reduces the dimensionality of the problem. k -means clustering is used to partition the spectral data points.

C. Goodness of a cluster

Previously, we detailed partitioning a given LV DN into k zones. In this subsection, we deal with identifying the best-suited values of k . In a real-world LV DN network partitioning problem, we may not know how many clusters we want. We use performance indices for measuring the goodness of a partition and, based on different values of k , identify the best value which fits our needs. In this work, the goodness of a cluster is measured using the *silhouette coefficient*, the maximum of the mean silhouette values of the members in the network clusters. The *silhouette coefficient* of a clustering scheme is a confidence indicator of association between the members of the clusters [38], [45]. Rousseeuw's interpretation based on the value of the *silhouette coefficient* [46] is used for zone selection, and the complete Algorithm is detailed in Appendix B.

III. DAY-AHEAD FLEXIBILITY NEEDS ASSESSMENT

Flexibility is referred to as resources that can be activated or deactivated, thus in effect increasing or decreasing the consumption based on the grid's needs. For a DN, nodal voltages, and line loadings should remain within operational bounds. Violation of these limits could damage appliances, and affect DN operation, or grid elements. Authors in [18] present two business models for operating electrical flexibility in the context of services provided to transmission and distribution system operators and assisting them to accommodate a greater amount of renewable energy. Depending on the application of flexibility, flexible resource needs can be measured in terms of response speed (or ramp rate), duration (or energy), and direction (ramp up and ramp down power) [47]. Next, the framework for calculating the flexibility needs that DSO procures in the energy market is detailed.

A. Flexibility: definition

Prior work, [48], used three parameters for defining flexibility: (a) ramp rate, (b) power, and (c) energy. The unit of these metrics used for describing flexible resources are watt per second, watt, and joule, respectively. The flexibility model used in this work utilizes the power and energy parameters. This model resembles the energy storage model used in [49]. The active power flexibility at node i and time t is given as

$$\Delta P_{i,t}^{\text{flex}} = \Delta P_{i,t}^{\text{flex}+} + \Delta P_{i,t}^{\text{flex}-}, \quad (1)$$

where $\Delta P_{i,t}^{\text{flex}+}$ and $\Delta P_{i,t}^{\text{flex}-}$ denotes ramp down and ramp up flexibility, respectively. Ramp-down flexibility decreases the nodal load and is analogous to load curtailment. Ramp-up flexibility increases the nodal load and is analogous to generation curtailment. Ramp-down flexibility can be provided by deactivating consumer loads such as HVAC, water heaters, pool pumps [50] etc. Ramp-up flexibility can be provided by solar generation curtailment, and activating electrical loads. Prosumer energy storage if not fully charged or discharged can provide both ramp-up and ramp-down flexibility [51]. Note these flexible resources have a limited ramp rate, power, and energy to be utilized as flexibility. As an example, the temperature of a water heater can be temporally constrained to be within some bounds of water temperature. If utilized as ramp-up flexibility to increase the nodal load, it can only be operated till the water temperature hits the upper bound (or the lower bound for ramping down). Thus, to respect technical device and user comfort constraints,

we use power and energy constraints to define flexible resources. The ramp-down and ramp-up flexibility energy for node i are given as

$$\begin{aligned} E_i^{\text{flex}+} &= \sum_t \Delta P_{i,t}^{\text{flex}+} \Delta t, \\ E_i^{\text{flex}-} &= \sum_t \Delta P_{i,t}^{\text{flex}-} \Delta t, \end{aligned} \quad (2)$$

with ramp-up and down power and energy constraints given as

$$\Delta P_{i,t}^{\text{flex}+} \in [0, \Delta P_{i,t,\text{max}}^{\text{flex}+}], \quad \forall t, \quad (3)$$

$$\Delta P_{i,t}^{\text{flex}-} \in [\Delta P_{i,t,\text{min}}^{\text{flex}-}, 0], \quad \forall t,$$

$$E_i^{\text{flex}+} \in [0, E_{i,\text{max}}^{\text{flex}+}], \quad (4)$$

$$E_i^{\text{flex}-} \in [E_{i,\text{min}}^{\text{flex}-}, 0].$$

Since ramp-up and ramp-down flexibility levels can compensate each other, therefore, they are dealt with separately to correctly assess the DN needs. Next, we describe the metrics used to quantify flexibility needs.

B. Building blocks for day-ahead FNA

The limits on ramp-up and down power and energy flexibility needs are infrastructure constraints that the optimization needs to consider. With these inputs and dimensions to the FNA-OPF problem, the following blocks are utilized:

- 1) *Objective function*: the optimization minimizes the cost of operation of flexible resources while ensuring network issues are avoided. Assuming an ample amount of flexible resources available, there are many solutions that can solve DN issues, some of which could have the same objective function value. In our prior work, [52] we observe the minimum amount of flexibility needed is identified using a flat activation cost. Thus, in this work, we utilize a flat penalty factor for penalizing the use of ramp-down and ramp-up flexibilities. The penalty factors are denoted as $\lambda^{\text{rampDown}}$, λ^{rampUp} . These factors are fixed for all nodes at all times.
- 2) *Flexibility power needs assessment metrics*: Ramp up ($\Delta P_{i,t}^{\text{flex}-}$) and ramp down ($\Delta P_{i,t}^{\text{flex}+}$) power needs for a node i , time t is calculated by solving FNA-OPF, given in (5).
- 3) *Flexibility energy needs assessment metrics*: Ramp up ($E_i^{\text{flex}-}$) and ramp down ($E_i^{\text{flex}+}$) energy needs of DN for node or zone i .

- 4) Using the DN clustering method presented in Section II, DN zones are created. Nodal flexibility needs are calculated using FNA-OPF aggregated over zones and henceforth referred to as zonal FNA.

C. Notation used

A power network consists of nodes (N), branches (E), generators, and loads. Each node $i \in N$ has two variables, i.e., voltage magnitude ($V_{i,t}$) and phase angle ($\theta_{i,t}$) for time i which depends on power injection and load magnitude. The branch admittance $(i, j) \in E$ together with the nodal voltage differences govern the power flow and losses. The subset of nodes with loads connected is denoted as $N_L \subset N$. The active and reactive power demand in these nodes are denoted as $P_{i,t}^d$ and $Q_{i,t}^d$. Nodes with generators connected are denoted as $N_G \subset N$, have active and reactive power generation denoted as $P_{i,t}^g$ and $Q_{i,t}^g$.

D. Optimization formulation

The flexibility needs assessment is performed by solving the optimal power flow problem which minimizes the cost of dispatching flexible resources. We solve a multi-period optimal power flow problem with penalties associated with generation and load curtailment (3). Further, multi-period coupled energy curtailment constraint is considered along with instantaneous curtailed power constraint (4). The optimization problem is denoted as flexibility needs assessment optimal power flow dispatch or FNA-OPF. FNA-OPF is an adaptation of the resource dispatch tool introduced in [53].

$$(\text{FNA-OPF}) \min_{\Delta P_{i,t}^{\text{flex}}} \sum_i \left(\lambda^{\text{rampDown}} E_i^{\text{flex}+} - \lambda^{\text{rampUp}} E_i^{\text{flex}-} \right) \quad (5a)$$

subject to, (3), (4)

$$V_{\min}^i \leq |V_{i,t}| \leq V_{\max}^i, \quad \forall i \in N, t \in \{1, \dots, T\}, \quad (5b)$$

$$(P_{i,t}^g) - (P_{i,t}^d - \Delta P_{i,t}^{\text{flex}}) + jQ_{i,t}^d = \sum_j s_{i,j,t}, \quad \forall i, j \in N, \quad (5c)$$

$$|s_{i,j,t}| \leq s_{i,j}^{\max}, \quad \forall i, j \in N, \quad (5d)$$

$$P_{i,t}^g \in [P_{\min,i}^g, P_{\max,i}^g], \quad \forall i \in N_G, \quad (5e)$$

$$s_{i,j,t} = \mathbf{Y}_{ij}^* V_{i,t} V_{j,t}^* - \mathbf{Y}_{ij} V_{i,t} V_{j,t}^*, \quad \forall (i, j) \in E \cup E^R, \quad (5f)$$

$$\angle(V_{i,t}V_{j,t}^*) \in [\theta_{i,j}^{\min}, \theta_{i,j}^{\max}], \quad \forall i, j \in N, \quad (5g)$$

where $s_{i,j,t}$ denote the apparent power flow from node i to j at time t . (5a) is the objective function of the FNA-OPF, i.e., the cumulative cost of ramp-up and ramp-down flexibility activations are minimized over the horizon of optimization. (5b), (5d) and (5g) denote the voltage constraint for nodes, thermal limit constraint, and phase angle constraints for branches, respectively. (5c) denotes the nodal balance of active and reactive power in the network. (5e) denotes the generator output power limits. (5f) denotes Ohm's law.

To avoid that both the ramp up and down flexibility described in (1) are non-zero simultaneously, they have been defined as $\Delta P_{i,t}^{\text{flex}+} \geq 0, \Delta P_{i,t}^{\text{flex}-} \leq 0$. As such, an increase in magnitude in either of them increases the objective function value in (5a). The FNA-OPF identifies the amount of temporal and nodal flexibility needed for solving the DNIs manifesting in the form of voltage and thermal violations. The generation cost is not included in FNA-OPF because this would affect the amount of flexibility activated, as in such a case there will be two components to the objective function. Quantifying the impact of the interaction between the generator cost and the cost of flexibility activated is not the goal of this work.

E. Robust FNA under uncertainty with risk

Robust optimization is a two-stage optimization problem where the inner level minimizes the objective function and the outer level maximizes the inner-optimization within an uncertainty set [30], [54]–[56]. For the robust FNA, we utilize scenarios for defining the uncertainty set. The inner *min*-problem solves the FNA-OPF for all generated scenarios. The outer level *max*-problem finds the robust flexibility levels for all time steps at a day-ahead level. The robust FNA problem is given as

$$\max_{s \in S} \min_{\Delta P_{i,t}^{\text{flex}}} \text{FNA-OPF}(s), \quad (6)$$

where $\Delta P_{i,t}^{\text{flex}}$ is the decision variable for the inner *min* problem. $s \in S$ represents the scenario of load and generation used for solving FNA-OPF and S , the set of all the scenarios, denotes

the uncertainty set considered. The ramp-down and ramp-up flexibility needs are modeled as an empirical cumulative distribution function (ECDF), given as

$$\begin{aligned} F_{i,t,+}(z) &= \frac{1}{S} \sum_{s=1}^S \mathbf{I}_{(-\infty, z]}(\Delta P_{i,t,s}^{\text{flex}+} \text{ for scenario } s), \forall i, t, \\ F_{i,t,-}(z) &= \frac{1}{S} \sum_{s=1}^S \mathbf{I}_{(-\infty, z]}(\Delta P_{i,t,s}^{\text{flex}-} \text{ for scenario } s), \forall i, t, \end{aligned} \quad (7)$$

where identity function is given as

$$\mathbf{I}_{(-\infty, z]}(y) = \begin{cases} 1, & \text{if } y \leq z, \\ 0, & \text{otherwise,} \end{cases} \quad (8)$$

ECDF provides a non-parametric model of the distribution of the ramp up and ramp down flexibility needs calculated for all scenarios without making an assumption on the underlying structure of the distribution, such as a normal or a beta distribution. Based on ECDF, the chance-constrained robust flexibility needs of the distribution network are identified below

$$\begin{aligned} (\Delta P_{i,t}^{\text{flex}+})^* &= F_{i,t,+}^{-1}(\epsilon_{cc}), \\ (\Delta P_{i,t}^{\text{flex}-})^* &= F_{i,t,-}^{-1}(1 - \epsilon_{cc}), \end{aligned} \quad (9)$$

where $\epsilon_{cc} \in [0, 1]$ denotes the chance constraint level.

The load profiles and their associated forecast error profiles are utilized by generating S scenarios. These scenarios are used as input for FNA-OPF. Note that to obtain a feasible solution in all the scenarios, we solve FNA-OPF where either load and/or generation can be curtailed. The activated flexibility is analyzed as shown in Fig. 1. We apply chance constraint analytically on ECDF distribution of $\Delta P_{i,t}^{\text{flex}+}$ and $\Delta P_{i,t}^{\text{flex}-} \forall S$. If the ϵ_{cc} is too low, then the risk for over-procurement is high. However, if the ϵ_{cc} is too high, then the risk of under-procurement is high, see (9). The CC level avoids over-procurement of flexibility. The process of selecting ϵ_{cc} is detailed next in Section III-F.

F. Pareto optimal tuning of chance constraint level

In multi-objective optimization, some of the objectives may be in conflict with others. In such a case, an optimal solution for one objective could lead to unacceptable outcomes for other objectives. In power systems, prior works such as [57]–[59] utilize multi-objective optimization. Authors in [59] use epsilon-constrained Pareto front for an electric vehicle scheduling problem of an aggregator

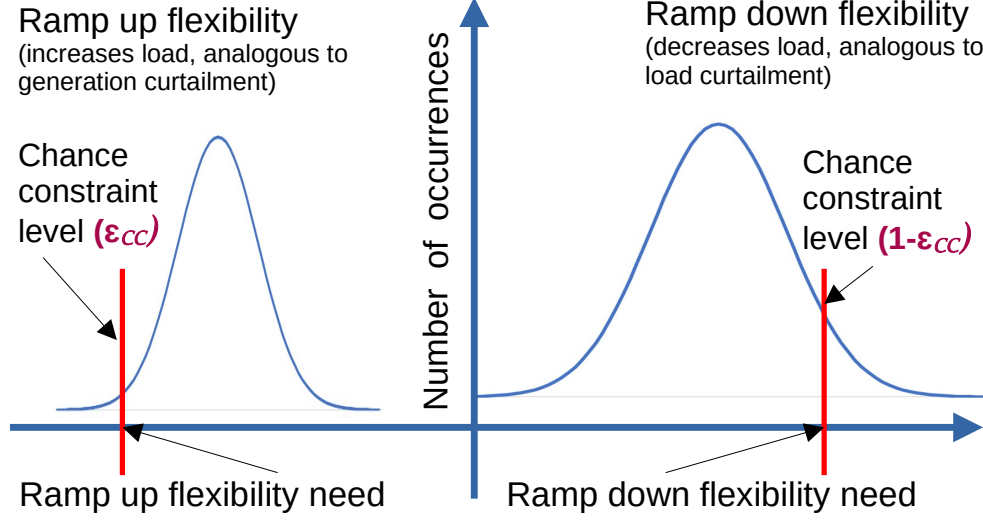


Fig. 1: FNA metrics for ramp up and down power and energy using chance constraint level denoted as ϵ_{cc} .

while considering DSO's goal to minimize losses, at the same time reducing the charging cost. Pareto optimality is frequently used to select between conflicting goals of a multi-objective optimization problem [60]. In our work, we use the Pareto optimal value of chance constraint to *reduce over-procurement of flexible resources* while *reducing the probable DN congestion incidents*. CC level is applied to the temporal and locational distribution of ramp-up and ramp-down power needs for the generated scenarios. These distributions often have a fat-tail, as shown in Fig. 1. Planning for the absolute worst case may lead to drastic over-design of flexibility needs.

Tuning CC level (ϵ_{cc}) for FNA: The selection of ϵ_{cc} is crucial for DSO to appropriately plan for day-ahead flexibilities. Both over and under-procurement is not efficient. We project ϵ_{cc} on mean DN voltage and thermal congestion over all the generated scenarios. Next, the Pareto optimal level of CC is utilized for reducing DN flexibility activation while reducing DNI. In Fig. 1 the robust ramp up and ramp down FNA is visualized, in which the extreme tail events are not considered. The knee point of the Pareto surface is identified using the algorithm proposed in [61]. In numerical results, we demonstrate the steps considered for calculating the ϵ_{cc} .

G. Pseudocode for calculating FNA

The inputs and outputs of the proposed FNA framework are shown in Fig. 2.

The pseudocode for identifying nodal and zonal FNA is detailed in Algorithm 1.

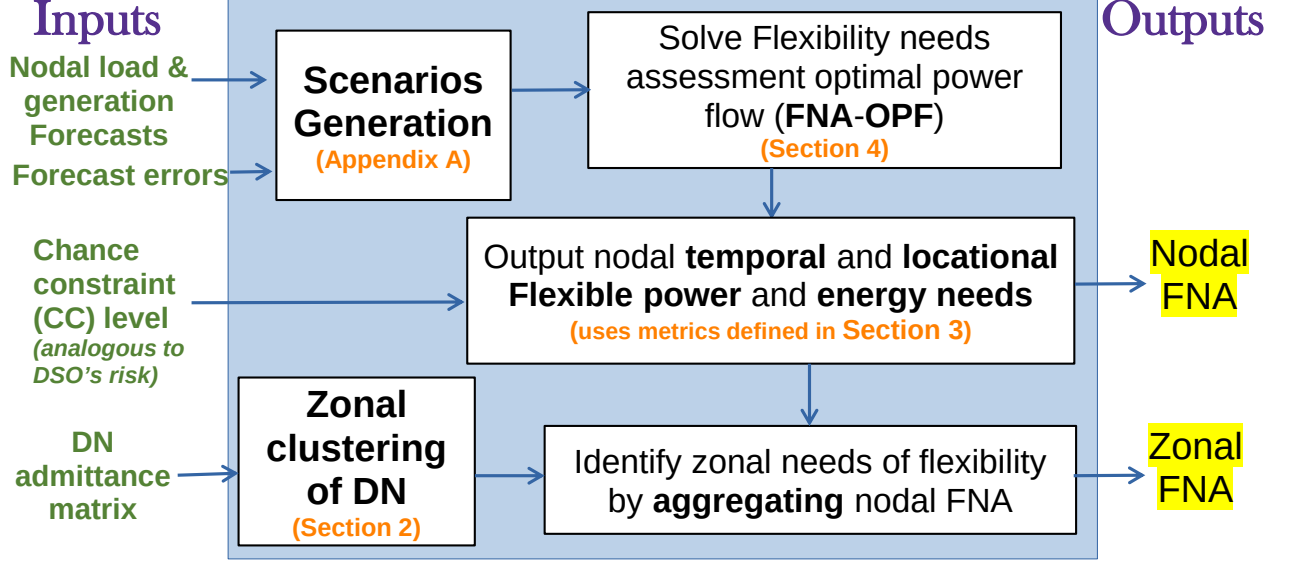


Fig. 2: The inputs and the outputs of scenario-based robust chance-constrained FNA under uncertainty.

Algorithm 1 Pseudocode for calculating nodal & zonal FNA.

Inputs: Nodal load forecast, forecast errors

- 1: Generate scenarios using forecast profiles and forecast errors as detailed in Appendix A,
 - 2: Solve FNA-OPF using (5) and store the $\Delta P_{i,t}^{\text{flex}}$ for each $s \in S$,
 - 3: The robust FNA can be calculated by solving (6),
 - 4: Calculate ECDF functions $F_{i,t,+}, F_{i,t,-}$ for identified FNA $\Delta P_{i,t,s}^{\text{flex+}}, \Delta P_{i,t,s}^{\text{flex-}}, \forall s \in S$ using (7).
 - 5: Set the value of ϵ_{cc} by comparing unavaoided DNIs with the amount of flexibility needed. Pareto optimal point is calculated using [61],
 - 6: Using (9), calculate risk based FNA. The calculated FNA considers the probable uncertainties and risk level in form of ϵ_{cc} .
 - 7: Using Algorithm 2 and Sec. II, calculate the zones of the DN,
 - 8: Aggregate nodal FNA in Step 6 in zones for identifying zonal FNA.
-

The FNA calculation using Algo. 1 does not assume the location of flexible resources. Although we assume that flexibility is derived at nodes where there is a load or generation source connected. For all the nodes with no load and generation connected, flexibility (ramp up and ramp down) is assumed to be zero. Ramp-down flexibility (derived from load curtailment) is available only at nodes with loads (positive consumption), while ramp-up flexibility is available not only at nodes with DG (negative consumption) but also nodes where nodal load can be increased by activating flexible loads.

IV. BID MATCHING ALGORITHM

The operation of flexibility markets will require utilities to accept bids offered in the bidding platform of the flexibility market. In the proposed bid matching, the flexibility bids in time and location are matched with FNA calculated using Algo. 1. Matching nodal FNA will often lead to insufficient resources, as one-to-one matching of nodal bids and nodal needs may not be met. In such a case, DSO's may not have enough time to rerun the detailed evaluation for selecting appropriate flexibility bids. We propose a bid selection algorithm that first matches nodal bids and then finds alternate zonal bids for residual flexibility needs which have not been satisfied in the nodal matching. This hierarchical bid selection algorithm can also be used for finding alternate flexibility options for solving DNIs.

The bid matching algorithm is shown as a flow chart in Fig. 3. The DNI are localized network incidents of voltage and thermal violations. With high probability, a voltage and thermal congestion incident can be solved by flexible resources in the neighborhood. Using this hypothesis, we propose a nodal bid with zonal flexibility matching algorithm which can be used by DSOs for flexibility procurement. The zonal matching takes into consideration the unmatched FNA after nodal matching. Partial nodal matching could be caused due to scarcity of flexibility bids in the market. Zonal matching of unmatched nodal FNAs would provide alternative flexibility resources which could resolve network issues. The inputs for the bid-matching algorithm are (a) nodal FNA, (b) zonal FNA, and (c) flexibility bids aggregated by Flexibility Service Provider or FSP.

The proposed bid matching leads to three scenarios,

- (C1) Flexibility needs fully met,
- (C2) Flexibility needs approximately met, and
- (C3) Flexibility needs partially not met.

In case of nodal residual flexibility is not identified in the zones, then in such a case, flexibility needs are partially not met. If zonal matching can meet flexibility needs, it is called approximate matching of flexibility. And if flexibility needs are entirely met by nodal bid matching, it is referred to as exact matching.

In numerical simulations, we will quantify the impact of the proposed bid matching on the additional DNI that DSO's could avoid. The performance metrics used for evaluation are:

- Additional contribution due to zonal matching is given as

$$\text{Additional match} = \text{Nodal residual} - \text{Zonal residual}. \quad (10)$$

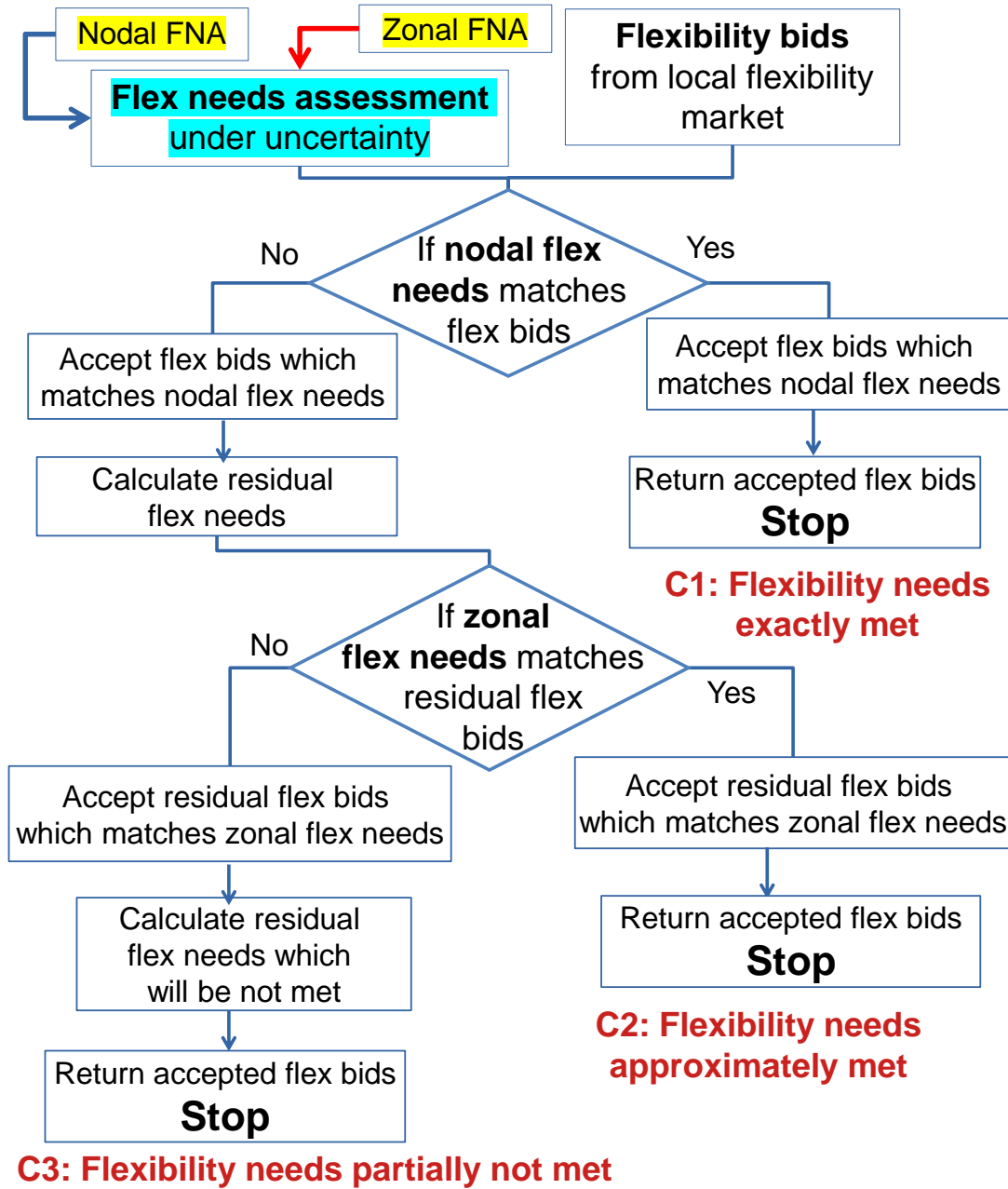


Fig. 3: Flexibility bid selector flowchart.

In (10), nodal residual refers to the unmet FNA after nodal bid matching and zonal residual refer to the unmet FNA after nodal plus zonal bid matching.

- Reduction in additional network incidents reduction (ANIR) due to zonal bid matching is given as

$$ANIR = \frac{DNI \text{ post nodal match} - DNI \text{ post zonal match}}{DNI \text{ without flexibility}}. \quad (11)$$

(11) denotes the metric used for DNI reduction attributed due to zonal bid matching post nodal bids are matched. Note, in (11) that we use DNI without flexibility in the denominator to avoid scaling issues caused to division with zero or numbers that are close to zero.

V. NUMERICAL CASE STUDIES

The DN considered is an adaptation of one of the Spanish LV feeders described in [62]. The DN consists of 76 nodes and 75 branches connecting the nodes. 52 loads are connected in 28 DN nodes, implying many nodes have more than one load connected. The network diagram is shown in Fig. 20 (in Appendix B). The network used in this work along with load profiles can be found on GitHub [63]. The numerical simulations are performed using PowerModels.jl in Julia / JuMP [64]. As the FNA-OPF problem is non-convex, the IPOPT solver [65] is used to solve (5). The feasibility of the solution of FNA-OPF is verified using power flow calculations. All simulations are performed on a personal computer with 32 GB RAM, i7-8665U CPU at 1.9GHz processor.

The load and PV distributions are selected so that the network issues are visible for significant periods. The total load of the DN seen at the substation over a day is 2.719 MWh. The total installed PV is equal to 338 kW peak, with a cumulative PV generation over a day of 1.709 MWh. Based on the above load profiles, referred to as nominal nodal profiles, 1000 scenarios are generated using Cholesky decomposition (detailed in Appendix A) and these scenarios are shown in Fig. 4. For the scenario generation, the forecast error for nodal load is assumed to be 30% and for PV generation is assumed to be 40%. Note from Fig. 4, that the load perturbations are reduced during aggregation at the substation, however, PV generation due to the high simultaneity factor leads to significantly high fluctuations during the day. For the nominal load profile, the self-sufficiency is 62.85% which means a substantial part of the consumer load is met using PV generation. The voltage set-point at the slack bus is set at 1.01 per unit.

Due to the high amount of PV installations and high peak loads during the late evening, we observe DN voltage and thermal constraint violations. For the 1000 scenarios generated, we observe that only 57.03% of optimal power flow calculations (without flexibility) are feasible for all time steps using an iteration limit of 3000. We observe under voltage (voltage < 0.95) for 6.165% of the samples in all scenarios, and for 3.275% of samples we observe over-voltage (voltage > 1.05). Thermal overload is calculated by performing OPF with relaxed network constraints, as OPF output for infeasible cases is not reliable. Power flow calculations are used for all non-feasible OPF solutions. We observe thermal over-loads for 0.11% of samples over all samples, i.e., (number of scenarios) \times (number of nodes) \times (number of time steps in 1 scenario). The step-by-step procedure for calculating nodal and zonal FNA is detailed in Algorithm 1.

The mean silhouette score for the 76-node network based on the best cluster evaluation described

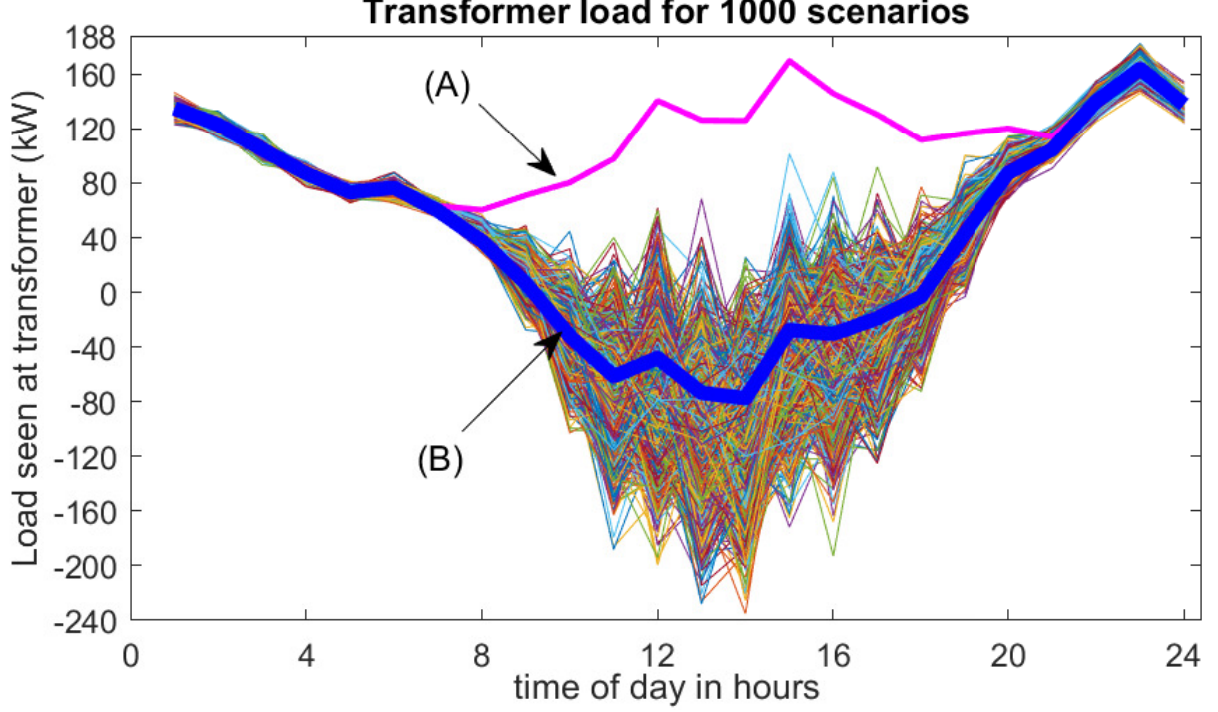


Fig. 4: Net load as seen at transformer with 1000 scenarios. (A) shows the nominal aggregate load profile without PV generation, (B) shows the nominal aggregate load profile with PV generation.

is shown in Fig. 5. The mean silhouette score for 12 clusters is 0.8487. The high silhouette score implies the nodes in a cluster are strongly related to each other and weakly related to other nodes in the other clusters. Fig. 20 shows the zones of the DN. Note that the naming of zones is arbitrarily selected based on k -means centroid. The stability of DN clustering is ensured by performing 100 Monte Carlo simulations.

A. Case study 1: Flexibility needs assessment

The objective of this case study is to evaluate the steps for calculating scenario-based robust FNA using risk tuning. The 76-bus Spanish DN is utilized. The methodology used is detailed in Algorithm 1. The key outcome of this numerical case study is the DA nodal and zonal FNA for the DN.

FNA-OPF is executed for unbounded power and energy flexibilities. Thus, we identify the minimum amount of flexibility needed to avoid all network issues in all scenarios. FNA based on a chance constraint level of 5% is shown in Fig. 6. Parametric and non-parametric distributions are used for identifying the ϵ_{cc} and $1 - \epsilon_{cc}$ quartiles of ramp down and ramp up flexibility needs. The

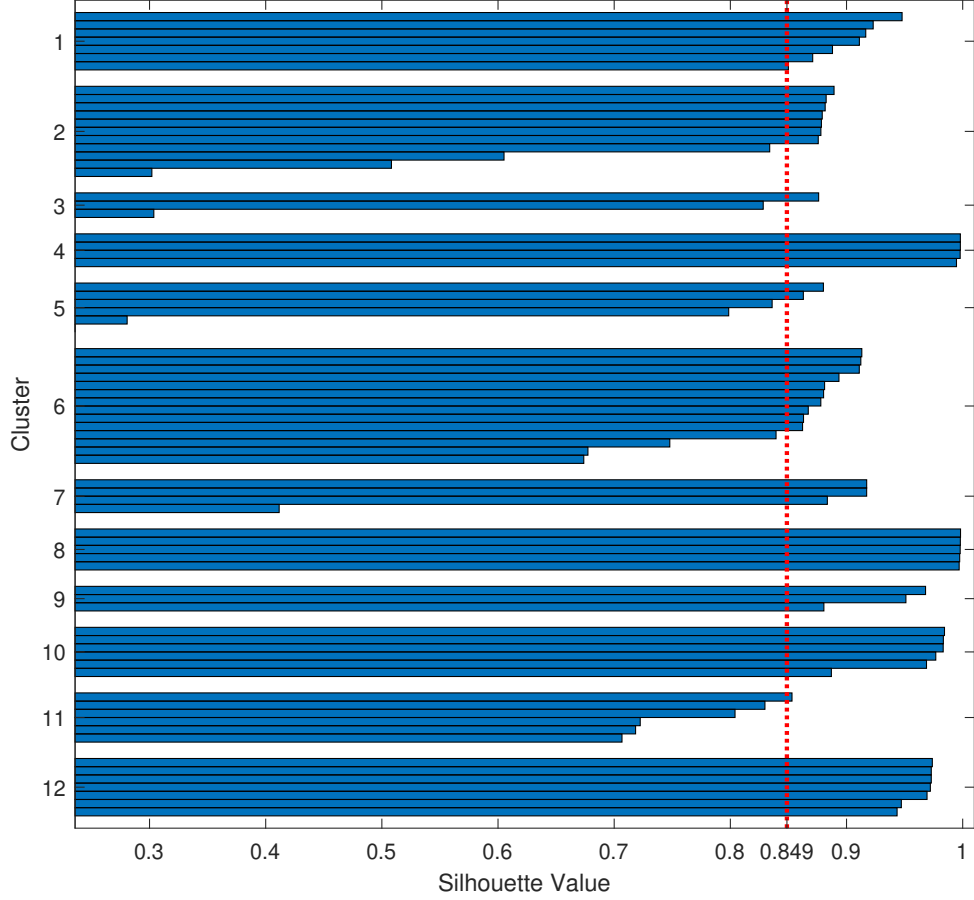


Fig. 5: Silhouette score for 76 node network with 12 zones.

parametric distribution of flexibility needs is fitted using a normal distribution. Note from Fig. 6 that the quality of the approximation depends on the distribution of the flexibility needs for all scenarios. We utilize a non-parametrized ECDF for applying CC as a normal distribution fit, shown in Fig. 6(a). From the numerical results, we observed that the error between FNA calculated by applying chance constraint level on the inverse ECDF and on an approximate Gaussian fit is less than 1% for most points, however, few instances had an error exceeding 5%. Since the distribution of FNA of a DN is not known, ECDF is used for all subsequent results. Note for ramp-up flexibility, the CC level is $(1 - \epsilon_{cc})$ level because of the sign reversal.

Fig. 7 shows the nodal ramp-up and ramp-down flexibility power needs of the DN. The nodal power needs are aggregated into zones of the DN. Zonal FNA is shown in Fig. 8. It can be observed that the majority of generation and load flexibility is needed in zones 1, 3, 4, 5, 6, and 11. Note that all these zones are located at the end of the DN feeder. Implementing a flexibility market can help

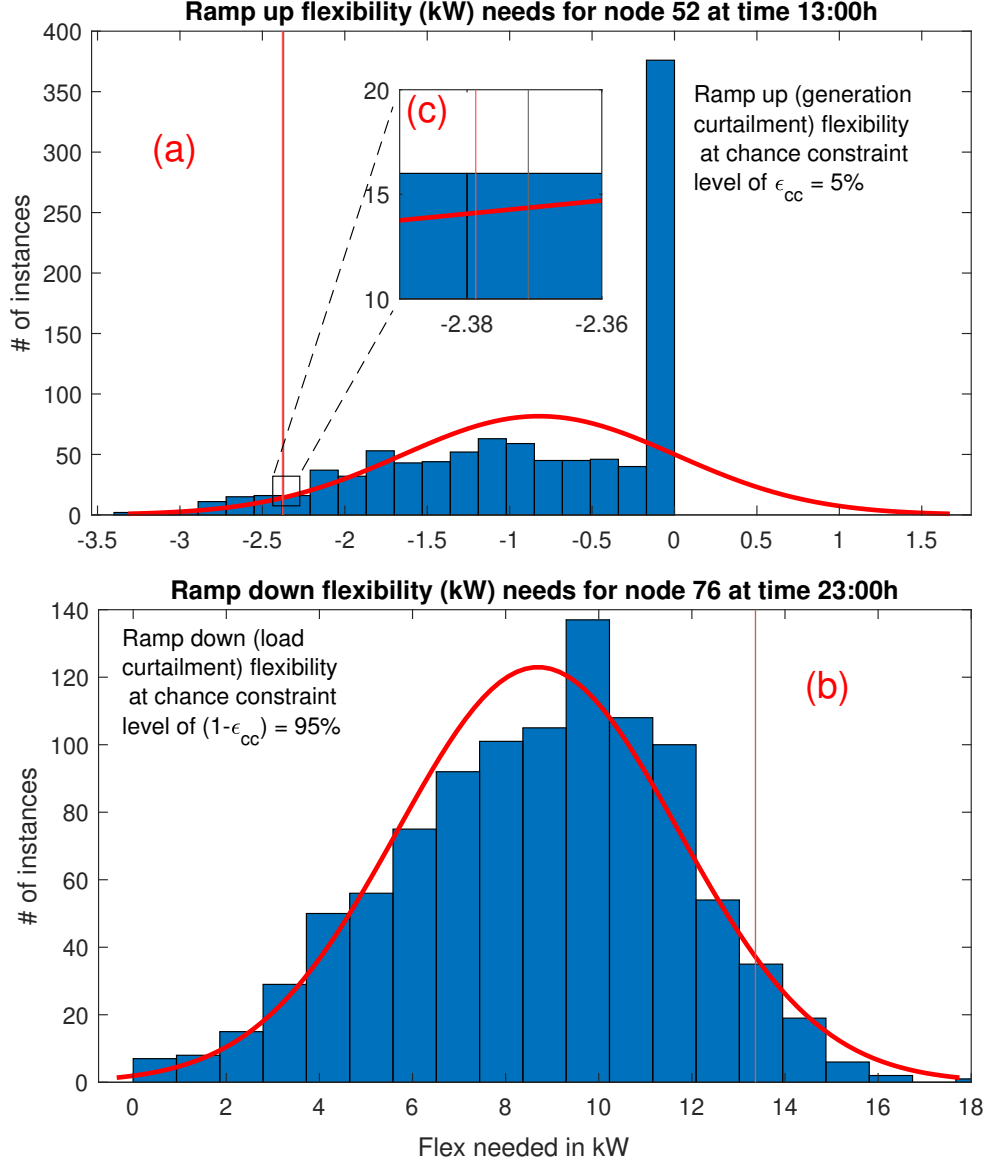


Fig. 6: FNA calculation based on normal distribution fit and CC quartiles. The error between inverse ECDF and quantile of approximate normal distribution fit is 0.33% and 0.02% respectively.

reduce prosumers' spatial inequality. The participation of flexible resources in such a market can create an additional revenue stream for such prosumers and can alleviate the locational disparity, also observed in [66].

From Table I, it can be observed that zonal FNA leads to a reduction in standard deviation (STD) compared to nodal FNA. The mean and STD denote the parameters for the normal distribution fitted to the histogram of ramp-up and down FNA. We observe that in this numerical case study that the

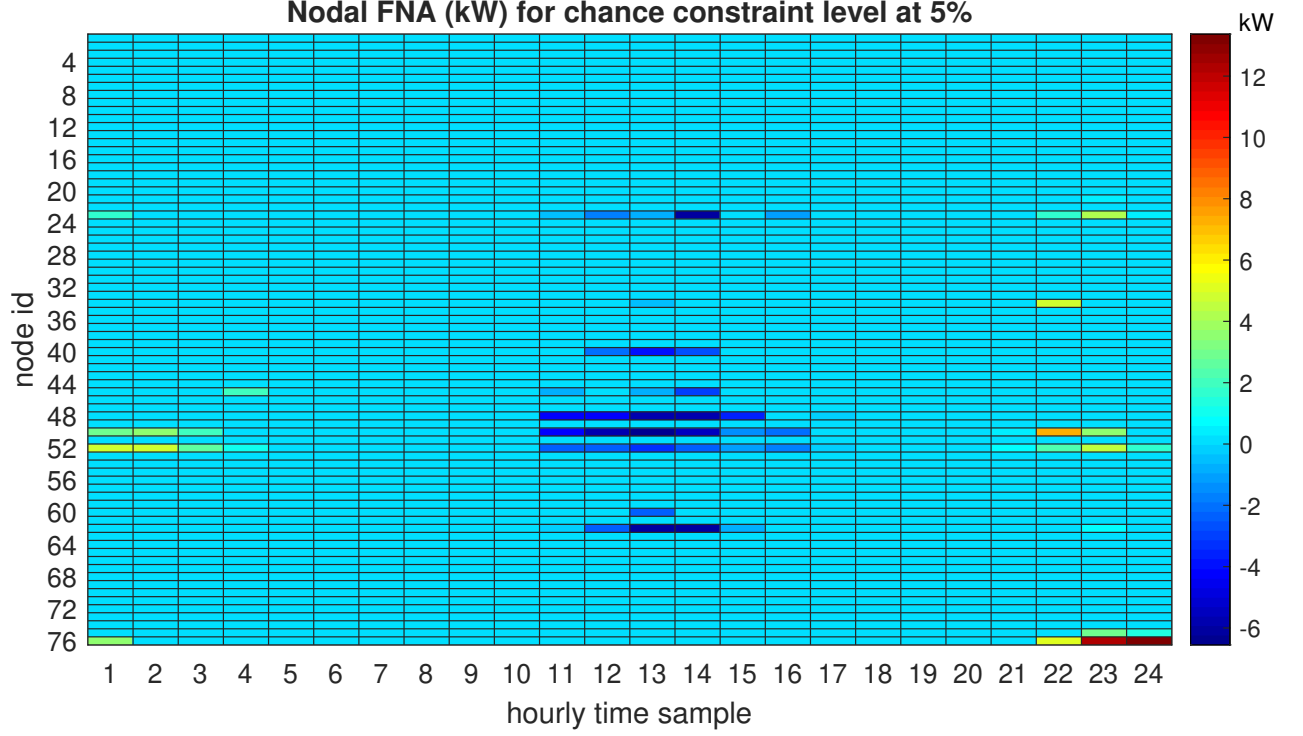


Fig. 7: Nodal flexibility power needs with $\epsilon_{cc} = 5\%$.

zonal flexibility needs have an STD which is 43.48% and 57.85% less for ramp up and down *nodal* FNA. Thus, zonal FNA is more immune to uncertainty. The aggregation of uncertain parameters leads to the reduction in the variance of uncertainty is a well-known concept [67]. The advantages of flexibility aggregation are also observed in [68]. Numerically, it is observed that the variance of the day-ahead FNA can be reduced by zonal aggregation of nodal FNA. This implies aggregation makes it more predictable for the DSO to plan flexibility in a day-ahead setting. The zonal FNA can help DSOs to identify alternative flexible resources which could probably solve DNIs.

TABLE I: Comparing aggregate nodal and zonal FNA variation for DA.

Perf. index	Nodal flex need (kWh)		Zonal flex need (kWh)	
	Ramp Up	Ramp Down	Ramp Up	Ramp Down
Mean	- 250.39	328.7	- 250.39	328.7
STD	328.5	256.6	185.8	108.2
STD reduction	-	-	43.48%	57.85%

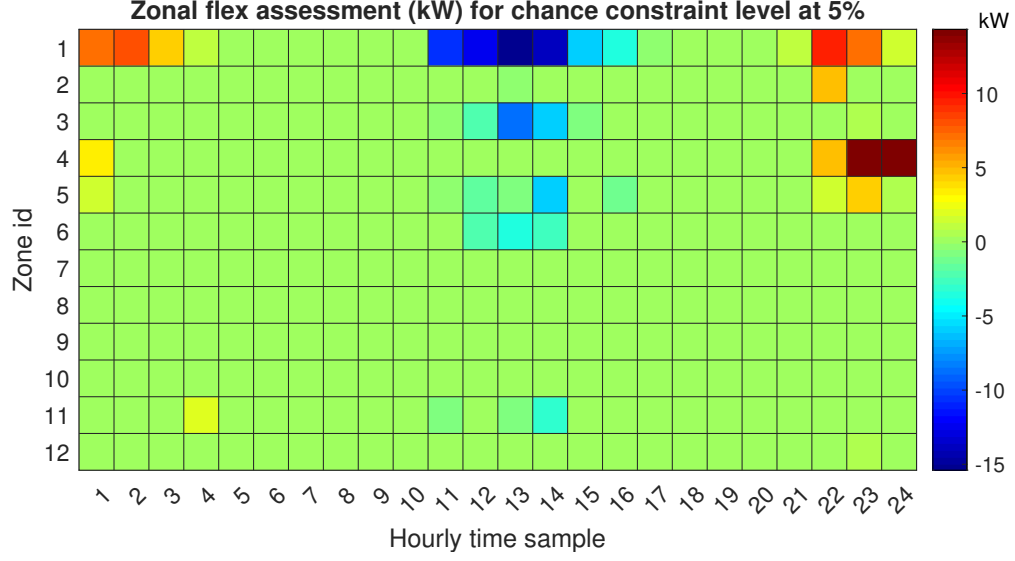


Fig. 8: Zonal flexibility power needs (kW) with $\epsilon_{cc} = 5\%$.

Table I shows that a significant amount of ramp-up and ramp-down flexibility is needed for DN, with DNIs observed for less than 10% of the time. In our numerical example, mean ramp-up and down needs are around 9.2% and 12.1% of the time of day, respectively. With growing DG integration and load uncertainty, the need for flexibility planning will be crucial for maintaining the reliability of DN.

Fig. 9 shows the impact of the chance constraint level on the flexibility energy needs. The heavy tail distribution can be avoided by a small value of CC. Choosing $\epsilon_{cc} = 5\%$ can reduce ramp-up energy needs by 68% and ramp-down energy needs by 40%, respectively. The CC level can also be used for identifying the criticality of a resource in ensuring DN reliability, as shown in the next section.

The computation time for 1 scenario of FNA-OPF for the 76-node network is on average 7 seconds. Therefore, nodal and zonal FNA are calculated according to Algo. 1 for 1000 scenarios takes around 2 hours.

1) Projecting CC levels on network issues: The CC levels are analogous to the risk DSOs are willing to take while planning for DN flexibility in a day-ahead setting. It will be useful to understand how the CC levels project to avoid probable network issues. We perform traditional OPF calculations with different levels of CC for FNA. The network voltage and thermal congestions are analyzed and listed in Table II where the mean values of probabilities over 1000 scenarios are

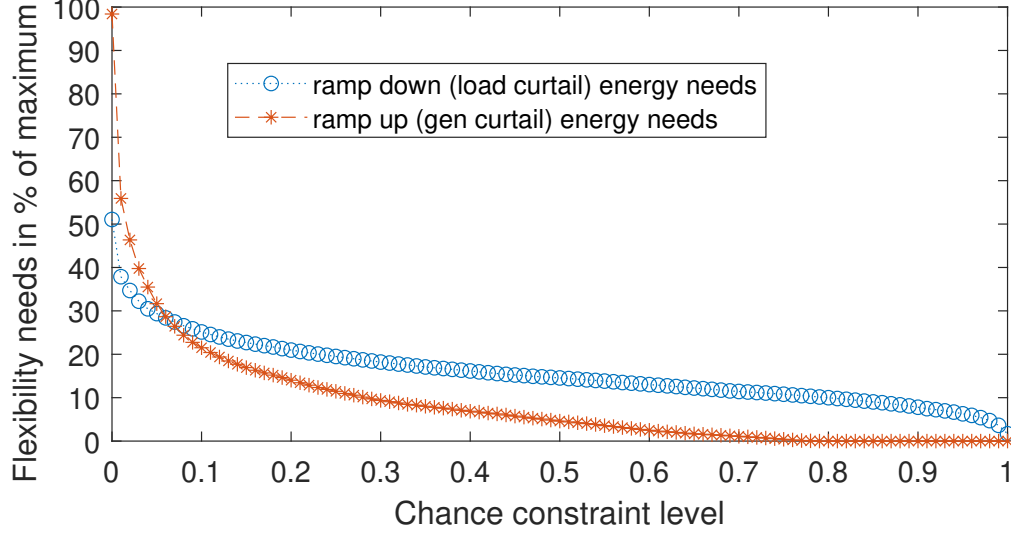


Fig. 9: Flexibility needed with chance constraint level ϵ_{cc} .

provided. Note for the $\epsilon_{cc} = 5\%$, the mean under-voltage probability reduced from 6.165% for no flexibility to 0.076%, a reduction of 98.77%. The mean over-voltage probability is reduced by 84.2% and the thermal over-load probability is reduced by 77.9%. The mean hours of congestion (thermal and/or voltage) for $\epsilon_{cc} = 5\%$ is reduced by 93.23%. As expected, a robust FNA with $\epsilon_{cc} = 0$ will eliminate all DN issues. From Table II we observe that CC levels cannot be directly projected on probable DN issues and will be governed by DN layout and the nodal distribution of demand and DG. Next, we discuss the mechanism for selecting the appropriate CC level for FNA.

2) *Pareto optimal tuning of CC level:* The FNA of a DN is governed by network layout and nodal load profiles. A Pareto front is built for the two conflicting goals of reducing the amount of flexibility needed, and the percentage of time in which the DN is congested. As the chance constraint level approached zero, the FNA increases and the percentage of time DN is congested reduces. However, planning so much flexibility would be very expensive. The desired solution is to procure less amount of flexibility while ensuring DN congestion instances are also low. The Pareto optimal solution is found for these conflicting goals of the DSO using the knee point calculation framework proposed in [61]. Fig. 10 shows that the Pareto optimal value of the CC-level for the considered DN is 5%.

TABLE II: Projecting CC levels on probabilities of network congestion.

	CC level %	Under voltage %	Over voltage %	Thermal over load %	% of hours with congestion
No flexibility	-	6.165	3.275	0.110	46.36
With	0	0	0	0	0
robust	1	0.066	0.102	0.007	1.108
CC	5	0.076	0.518	0.024	3.138
FNA	10	0.203	0.915	0.044	6.196
planning	20	0.544	1.440	0.082	12.16
	30	0.996	1.891	0.117	17.63
	40	2.079	2.563	0.153	26.43

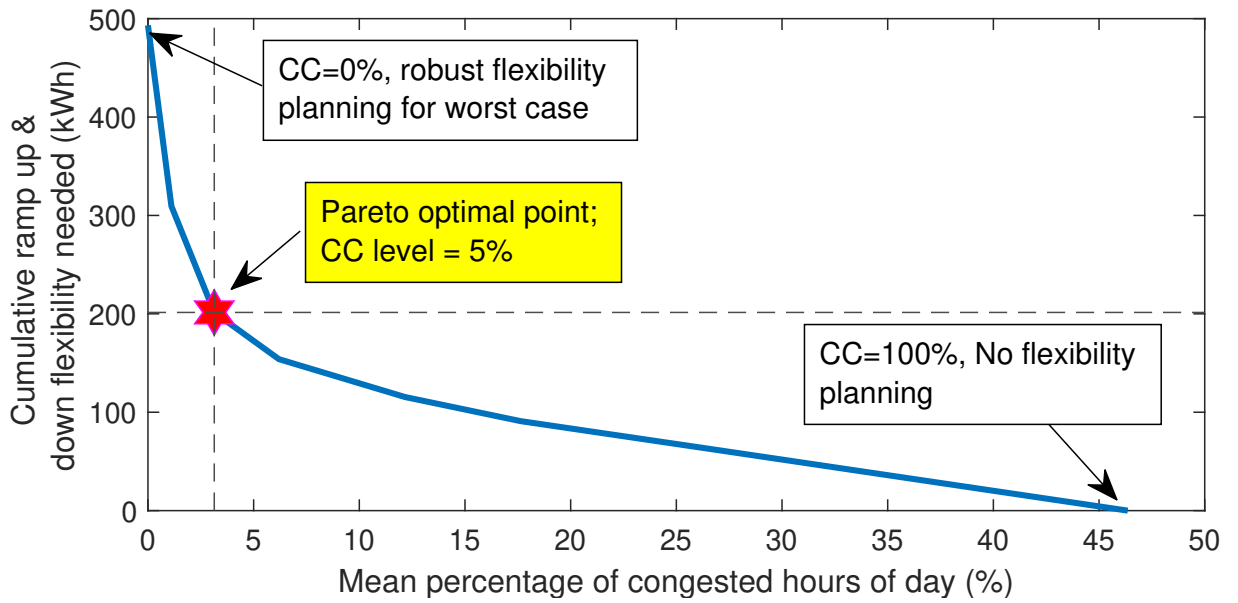


Fig. 10: Tuning CC level using Pareto optimality.

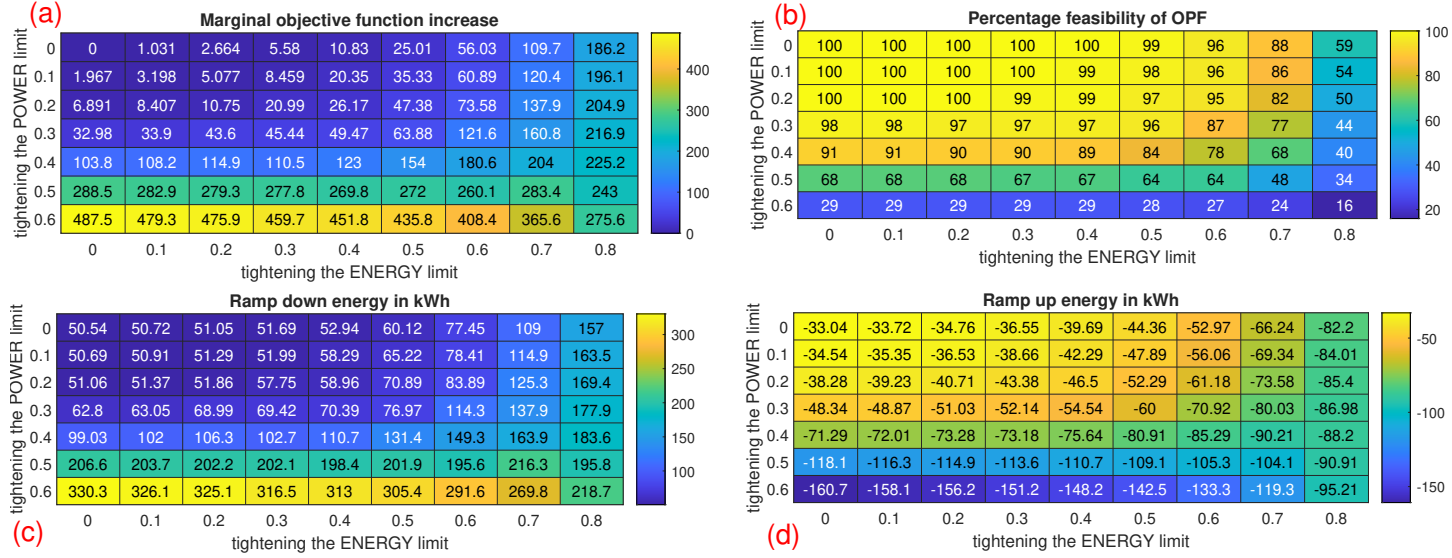


Fig. 11: DN flexibility needs with flexibility energy and power bound tightening; (0,0) denotes unconstrained solution of FNA-OPF.

B. Case study 2: Marginal value of energy and power

The objective of the second case study is to assess the marginal value of energy and power limits of flexibility. This is motivated by the power and energy attributes of flexible resources being valued differently in many energy markets. The 76-bus Spanish DN is utilized. The methodology used is based on power and energy constraint tightening in fixed steps of 10% for solving FNA-OPF. The key outcome of this numerical case study is the quantification of power as an attribute is roughly twice more valuable for DSOs compared to energy for flexible resources.

Flexible resources are valued differently in different types of markets [69]. The goal of this case study is to quantify the marginal impact of energy and power constraints on flexible resources. The ramp-up and down energy and power constraints are tightened in 10% steps, to evaluate: (a) marginal increase in the objective function value of FNA-OPF, (b) calculate the percentage feasibility of 100 scenarios, (c) cumulative ramp-up and (d) cumulative ramp-down energy needs. These metrics are shown in Fig. 11. The key observations from this study are:

- The marginal cost for ramp-up and down power is more than two times that of the energy. This is observed in Fig. 11(a) where the marginal increase in the objective function is more than 487 % compared to the unconstrained objective function value for FNA-OPF. This marginal

increase is due to a reduction in flexibility power ranges by 60%. However, for up to 80% reduction in energy constraint over a day, the marginal increase in the objective function is around 186%. Similar trends can be observed for other combinations of simultaneous tightening of power and energy constraints.

- The percentage of feasible FNA-OPF calculations is shown in Fig. 11(b). It is expected that the feasibility will reduce with tighter flexibility power and energy constraints. Note that we limit the levels of power and energy bounds tightening to 0.6 and 0.8 respectively, as the majority of points are infeasible for any further tightening. The solutions with extremely high levels of tightening are unreliable, as in such a case either there is no solution possible or a maximum number of iterations is reached. In both these cases, the solution for FNA-OPF cannot be used.
- Fig. 11(b) also indicates the combinatorial nature of the flexibility needs assessment problem. We can observe that there are many combinations of feasible solutions that can solve the same network congestion. This motivates the zonal plus nodal bid selection framework presented in Section IV.
- The ramp-up and ramp-down energy are shown in Figs. 11(c) and (d). It is observed that as power and energy constraints are tightened, a larger amount of flexibility is needed somewhere else in the network to mitigate and solve congestion. As the distance from the congestion location increases, the efficiency of the procured flexibility to solve congestion decreases. The FNA approximately doubles for 70% tightening of the energy constraint, while the FNA quadruples for 50% tightening of the power constraint.

C. Case study 3: Bid matching

The objective of this case study is to perform the evaluation of the bid matching algorithm using the 76-bus Spanish DN. The methodology used is the nodal and zonal bid matching detailed in Section IV. It is observed that the zonal and nodal bid matching leads to more than 43% additional reduction of probable DNIs.

In this case study, we assess the impact of nodal and zonal bid matching with only nodal bid matching on the probable network incidents avoided. The starting point of the bid matching is the nodal and zonal FNA calculated in Case Study 1 in Sec. V-A. The DN FNA is matched with flexibility bids available in the flexibility market. As the goal of this paper is not to develop a flexibility market, we assume flexibility bids available as a fraction proportional to the load (intake) and distributed generation (offtake) connected at the point of common coupling.

The aggregated temporal FNA is shown in Fig. 12. The blue-shaded region shows the ramp-down FNA, and the red-shaded region shows the ramp-up FNA of the DN. Note that although thermal overloads due to reverse power flow might not be happening, the over-voltage incidents are observed, in line with observations made in [70]. Thus, we observe FNA in ramp-up flexibility is higher although the reverse power flow seen at the substation is low, see Fig. 4.

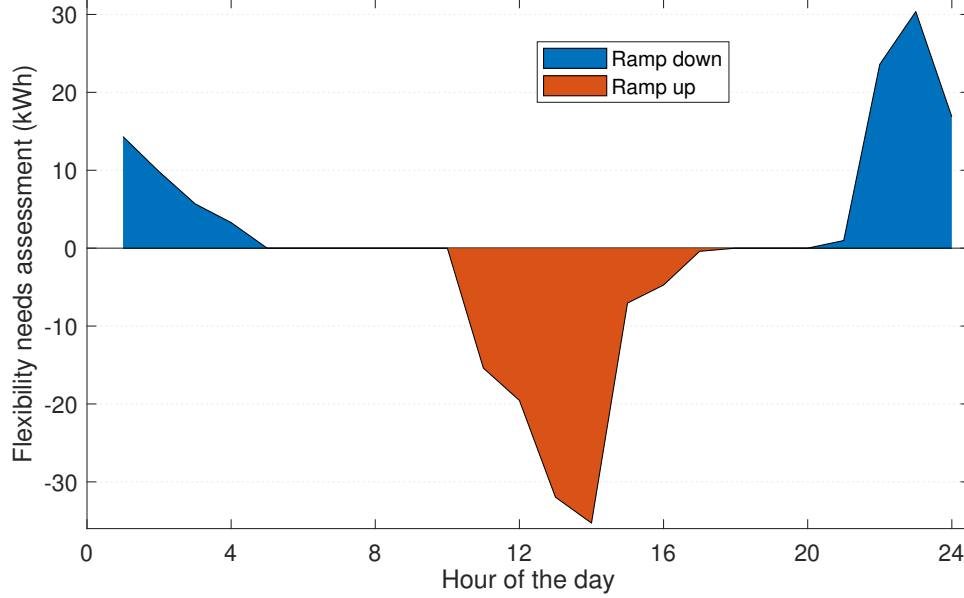


Fig. 12: Temporal FNA calculated in Case-study 1.

Using the scenario generation tool described in Appendix A, we generate 1000 scenarios. To produce proportional bids, the mean ramp-up and ramp-down profiles of these scenarios are used. The bids refer to the bids in the flexibility market. In Fig. 13, we show the additional matching achieved for ramp-up and ramp-down flexibility due to the zonal matching proposed in Sec. IV. Observe that as the fraction of load/DG generation as flexibility increases, the matching improves. With a very large amount of flexibility bids, the nodal residual would be very close to zero. Clearly, zonal matching improves the bid selection in case of nodal bid scarcity. Note, in Fig. 13, the flexibility bid level exceeds 100% as this fraction is calculated over the mean levels of nodal load profiles. From Fig. 13, note that additional matching increases up to 19 kW ramp down for 80% average load as bids and 17 kW ramp up for 55% average generation as bids.

Next, we perform power flows for all 1000 scenarios for the whole day and identify the network congestion incidents caused due to voltage and thermal limit violations. In Fig. 14, we show these

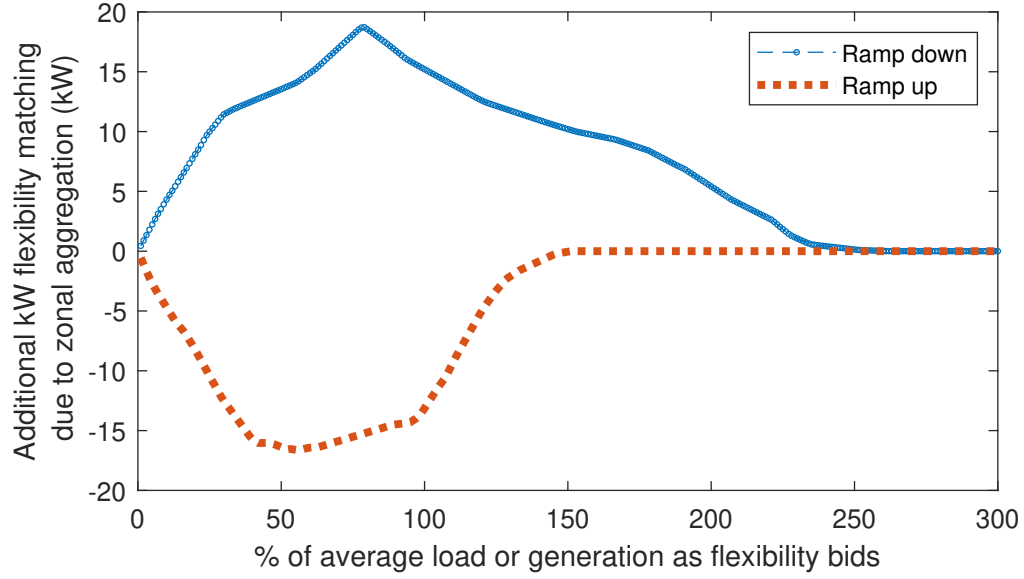


Fig. 13: Additional contribution of zonal matching given by (10).

network incidents for the following cases:

- no flexibility case,
- with flexibility needs exactly met, i.e., FNA,
- zonal and nodal bid matching for {20%, 40%, 60%, 80%, 100%} of load and generation as flexibility bids.

Note that as the amount of bids increases, the bid matching improves and leads to a reduction in network incidents.

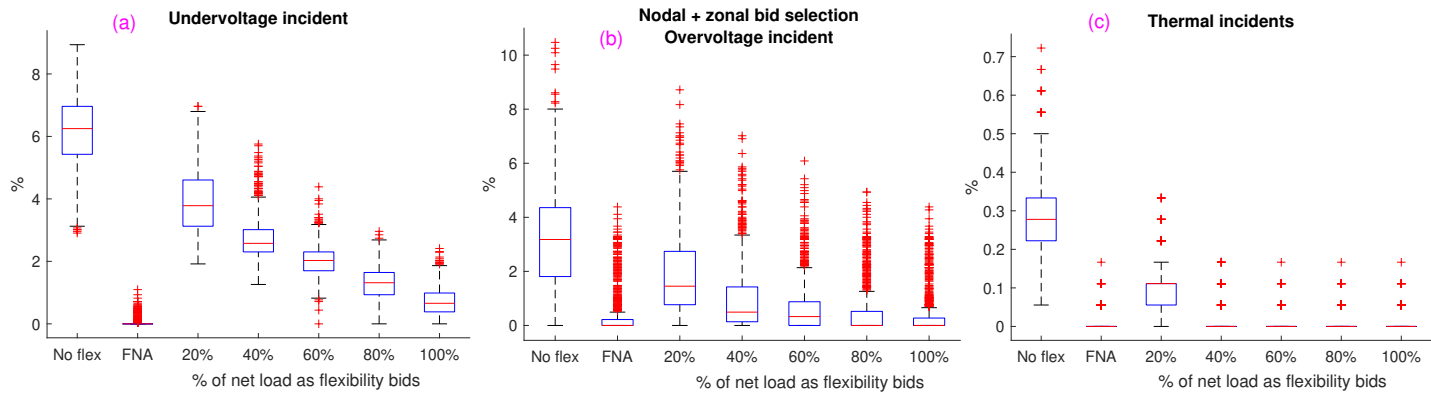


Fig. 14: Matched flexibility bids when projected on DNI: under-voltage, over-voltage and thermal overloads.

Next, we assess the marginal improvement in network congestion incident reduction due to zonal

bid matching of nodal residual. For this, we perform power flows using nodal matched bids denoted as C1 and then rerun power flows for all scenarios using nodal and zonal bid matching denoted as C2 in Fig. 3. This metric is shown in Fig. 15. Observe that zonal matching leads to more than 6.5% (on average) additional voltage incident reduction for 20% load and generation as bid. For thermal incident reduction, zonal bid matching leads to a 28% (on average) reduction. Clearly, zonal bid matching provides a rational manner of selecting bids based on the flexibility needs assessment.

D. Case study 4: FNA on a large DN

The objective of this case study is to perform FNA for a real demonstration DN in Germany for the EUniversal project. In this case study, we apply the FNA framework for a larger DN. This numerical case study considers a real suburban German DN used for EUniversal demo [71]–[73]. The DN consists of 646 nodes with 331 loads connected to it. Algorithm 1 is utilized as the methodology for calculating FNA. The key outcome of this numerical case study is the DA nodal and zonal FNA of the DN.

The details of a digital twin for synthetic data generation, node reduction, and network parsing for this demo network can be found here [73]. Fig. 16 shows the DN with 7 clusters identified using the DN clustering framework detailed in Section II. The load profiles are selected from a pool of historical load profiles based on prosumer meta-data such as annual kWh, and PV size installed. Based on historical data, the load profile forecast error is assumed to be 30% and PV generation forecast error of 40% is used for scenario generation. Based on these scenarios, the FNA is calculated. The temporal flexibility needs are shown in Fig. 17. Note from Fig. 17 that the ramp-up flexibility needs of this DN are both zero, as the PV penetration, in this case, was very low. The total installed PV is 61 kWp and the annual kWh served exceeds 923 MWh.

Fig. 18 shows the zonal FNA for the DN. Observe that most flexibilities needed are in zone 1, 3, and 4; all these zones are end-of feeders. More than 80% (30.04, 30.45, 19.33% respectively) of the flexibility needed is located in these three zones.

The computation time for 1 scenario is on average equal to 58 seconds. Thus, for calculation of nodal and zonal FNA detailed in Algo. 1 using 100 scenarios takes 1.6 hours. Since the proposed FNA is used in DA, therefore, the computation time is not a constraint at these levels, however, further work is required to improve the computation time by parallelization to adapt FNA calculations for the real-time application.

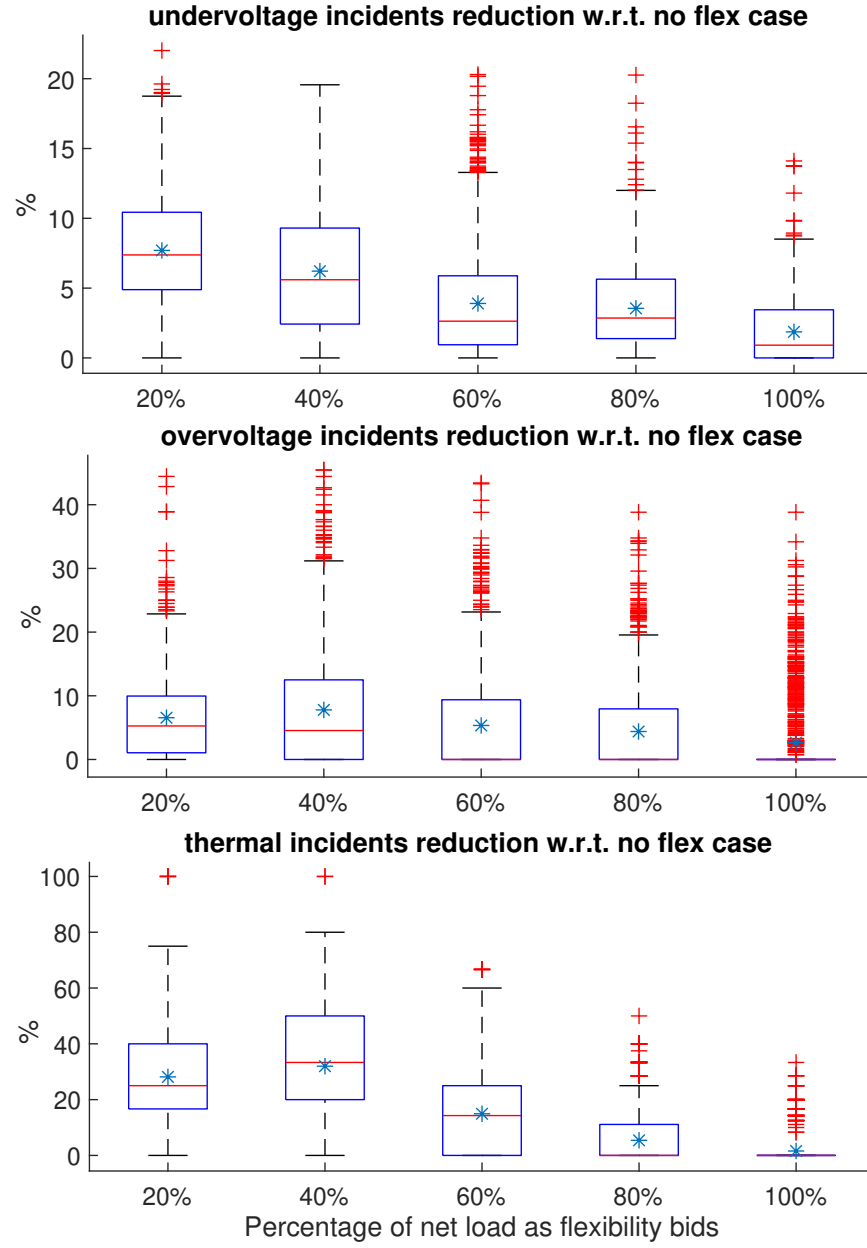


Fig. 15: Benefit of zonal matching on the network incidents calculated using (11). The incident reductions shown are with respect to the case, with no flexibility. The blue star shows the mean value of incident reduction for the given flexibility bid level.



Fig. 16: German DN with clusters marked [73].

VI. CONCLUSION AND FUTURE WORKS

Power system flexibility will be crucial for the reliable operation of a distribution network (DN) for ensuring a large-scale integration of distributed generation and new loads such as electric vehicles. An optimization-based framework for calculating the day-ahead flexibility needs assessment (FNA) of a DN is presented. The system operator uses the FNA calculated for selecting flexibility bids in the flexibility market. Thus, we provide DSOs with qualitative recommendations in the form of FNA and bid selection for flexibility procurement.

The FNA tool considers future uncertainties in the form of PV generation and demand scenarios. The flexibility needs assessment optimal power flow problem (FNA-OPF) identifies the temporal and locational (nodal) ramp-up and ramp-down power and energy needs of a DN, that minimizes the cost of dispatching flexible resources. Based on the distribution of FNA, a chance constraint-based

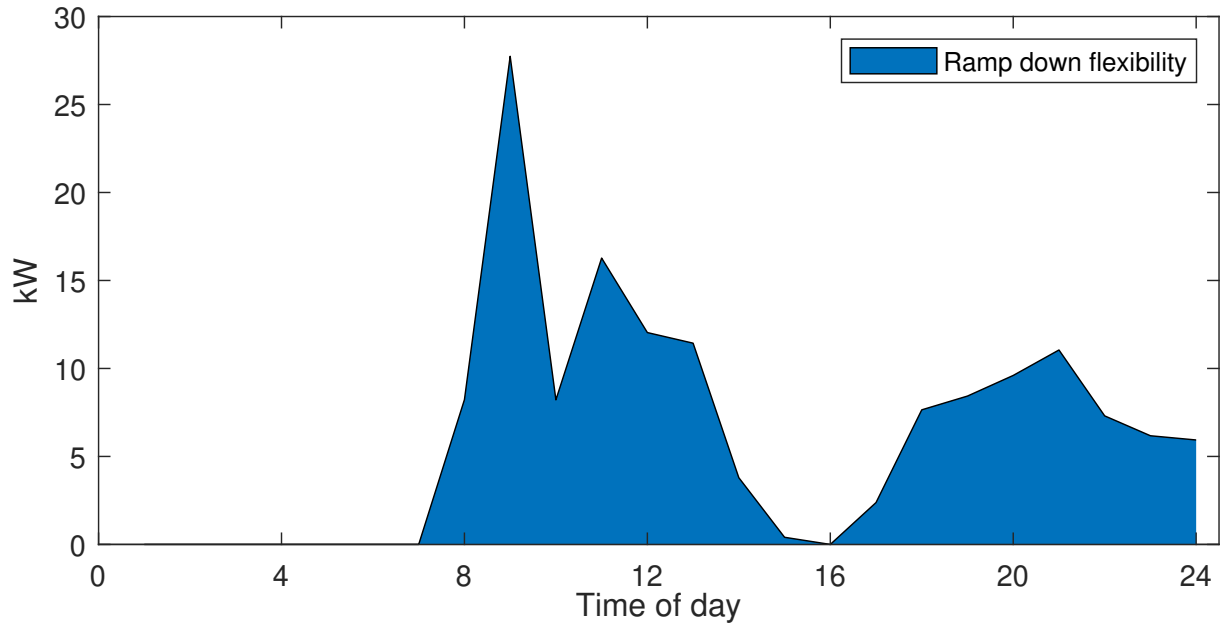


Fig. 17: Day-ahead temporal flexibility needs for German DN.

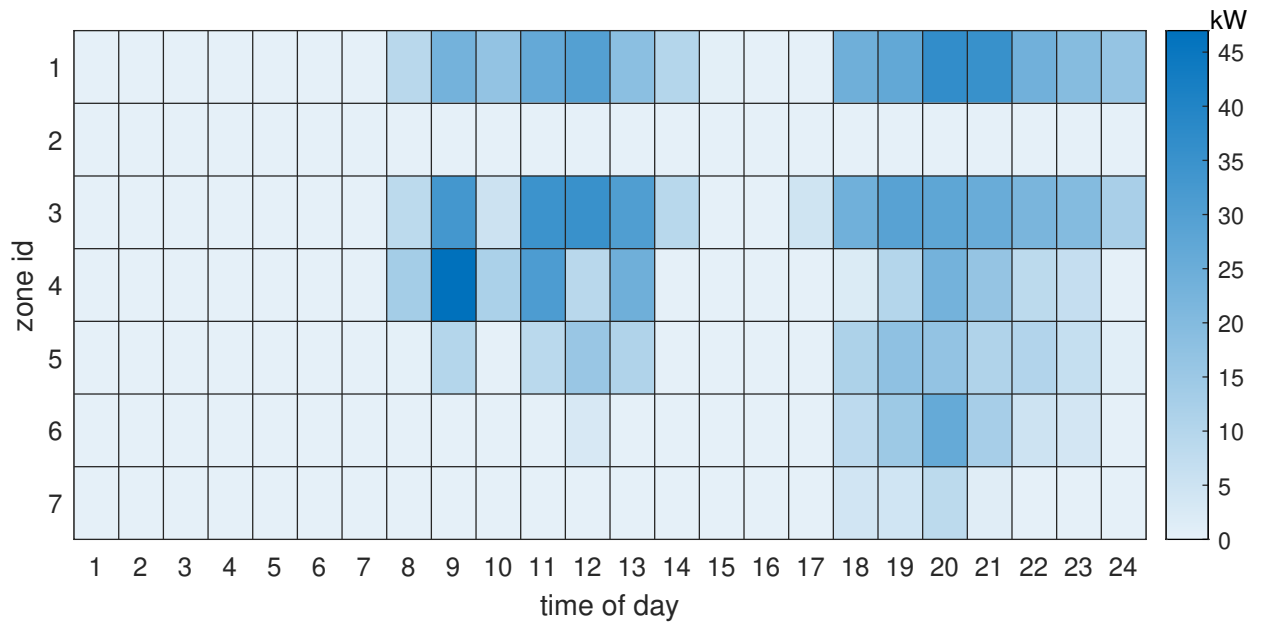


Fig. 18: Temporal flexibility needs are calculated in a DA horizon.

robust FNA is identified to avoid under or over-procurement of flexible resources. The appropriate number of zones is determined based on an electrical distance as a measure, spatial partitioning, and silhouette coefficient. The nodal FNA is aggregated in these zones for calculating zonal FNA. We propose zonal and nodal bid matching algorithm. The purpose of the zonal analysis is to assist DSOs in finding alternative resources in the same zone in case the nodal FNA identified are not available in the flexibility market.

Four numerical case studies show the efficacy of our proposed FNA and bid matching. In the first case study, nodal and zonal FNA is identified for a Spanish DN. It is observed that zonal FNA is more immune to future uncertainties, and DSO's can utilize this feature along with nodal FNA for flexibility planning and operation. Pareto optimal tuning of chance constraint is performed that reduces flexibility procurement while ensuring the unavoided DN incidents are also reduced. The second case study illustrates the impact of the flexibility bounds for power and energy on the flexibility needs and the marginal cost of additional flexibility. We observe that the marginal impact of tighter power flexibility constraints on activation cost is more than twice compared of energy flexibility constraints. Bid matching algorithm is evaluated in the third case study, we observe that nodal with zonal bid matching compared to naive nodal bid matching leads to a reduction of additional 15% and 28% DN incidents associated with voltage violations and thermal overloads, respectively assuming 20% of the load or generation as flexibility bids. In the fourth case study, we implement the proposed FNA framework for a real German suburban DN.

In future directions of work, we would extend the single-phase equivalent formulation for FNA calculation to three-phase unbalanced DNs while improving the computational time. We will also analyze the impact of power and energy flexibility constraints for determining the best suitable location and size of flexibility sources in DNs. Further work is needed to quantify the impact of reactive power flexibility on active power FNA.

ACKNOWLEDGEMENT

This work is supported by H2020 EUniversal project (<https://euniversal.eu/>) (Grant agreement ID: 864334). We would like to thank David Brummund and Maik Staudt from Mitnetz Strom for providing detailed inverter operational rules in Eastern Germany.

REFERENCES

- [1] Council of the EU, 27 June 2022. Fit for 55. [Online]. Available: <https://ec.europa.eu/energy/en/topics/renewable-energy/renewable-energy-directive>
- [2] M. A. F. Ghazvini, G. Lipari, M. Pau, F. Ponci, A. Monti, J. Soares, R. Castro, and Z. Vale, "Congestion management in active distribution networks through demand response implementation," Sustainable Energy, Grids and Networks, vol. 17, p. 100185, 2019.
- [3] (2022) Entso-e 2021 technology factsheets. [Online]. Available: <https://tinyurl.com/4anawj89>
- [4] H. Nosair and F. Bouffard, "Flexibility envelopes for power system operational planning," IEEE Transactions on Sustainable Energy, vol. 6, no. 3, pp. 800–809, 2015.
- [5] I. F. Abdin and E. Zio, "An integrated framework for operational flexibility assessment in multi-period power system planning with renewable energy production," Applied energy, vol. 222, pp. 898–914, 2018.
- [6] The EUniversal project. [Online]. Available: <https://euniversal.eu/>
- [7] P. Arboleya, M. A. Kippke, and S. Kersch, "Flexibility management in the low-voltage distribution grid as a tool in the process of decarbonization through electrification," Energy Reports, vol. 8, pp. 248–256, 2022.
- [8] D. A. Contreras, S. Müller, and K. Rudion, "Congestion management using aggregated flexibility at the tso-dso interface," in 2021 IEEE Madrid PowerTech. IEEE, 2021, pp. 1–6.
- [9] A. G. Givisiez, K. Petrou, and L. F. Ochoa, "A review on tso-dso coordination models and solution techniques," Electric Power Systems Research, vol. 189, p. 106659, 2020.
- [10] X. Ayón, J. K. Gruber, B. P. Hayes, J. Usaola, and M. Prodanović, "An optimal day-ahead load scheduling approach based on the flexibility of aggregate demands," Applied Energy, vol. 198, pp. 1–11, 2017.
- [11] E. Hooshmand, A. Rabiee, S. Jalilzadeh, and A. Soroudi, "Optimal flexibility coordination for energy procurement in distribution networks," IET Renewable Power Generation, no. 6, pp. 1191–1203, 2021.
- [12] R. Pinto, R. J. Bessa, and M. A. Matos, "Multi-period flexibility forecast for low voltage prosumers," Energy, vol. 141, pp. 2251–2263, 2017.
- [13] M. Z. Liu, L. N. Ochoa, S. Riaz, P. Mancarella, T. Ting, J. San, and J. Theunissen, "Grid and market services from the edge: Using operating envelopes to unlock network-aware bottom-up flexibility," IEEE Power and Energy Magazine, vol. 19, no. 4, pp. 52–62, 2021.
- [14] A. C. Meißner, A. Dreher, K. Knorr, M. Vogt, H. Zarif, L. Jürgens, and M. Grasenack, "A co-simulation of flexibility market based congestion management in northern germany," in 2019 16th International Conference on the European Energy Market (EEM). IEEE, 2019, pp. 1–6.
- [15] R. Fonteijn, T. Van Cuijk, P. H. Nguyen, J. Morren, and J. G. Slootweg, "Flexibility for congestion management: A demonstration of a multi-mechanism approach," in 2018 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe). IEEE, 2018, pp. 1–6.
- [16] G. Tsousoglou, J. S. Giraldo, P. Pinson, and N. G. Paterakis, "Mechanism design for fair and efficient dso flexibility markets," IEEE transactions on smart grid, vol. 12, no. 3, pp. 2249–2260, 2021.
- [17] S. S. Torbaghan, N. Blaauwbroek, P. Nguyen, and M. Gibescu, "Local market framework for exploiting flexibility from the end users," in 2016 13th International Conference on the European Energy Market (EEM). IEEE, 2016, pp. 1–6.
- [18] K. Coninx, G. Deconinck, and T. Holvoet, "Who gets my flex? an evolutionary game theory analysis of flexibility market dynamics," Applied energy, vol. 218, pp. 104–113, 2018.

- [19] S. S. Torbaghan, N. Blaauwbroek, D. Kuiken, M. Gibescu, M. Hajighasemi, P. Nguyen, G. J. Smit, M. Roggenkamp, and J. Hurink, "A market-based framework for demand side flexibility scheduling and dispatching," Sustainable Energy, Grids and Networks, vol. 14, pp. 47–61, 2018.
- [20] S. Davidov and J. Curk, "Novel framework for investment prioritisation based on flexibility needs assessment," IET Generation, Transmission & Distribution, vol. 14, no. 25, pp. 6189–6194, 2020.
- [21] A. Laur, J. Nieto-Martin, D. W. Bunn, and A. Vicente-Pastor, "Optimal procurement of flexibility services within electricity distribution networks," European Journal of Operational Research, vol. 285, no. 1, pp. 34–47, 2020.
- [22] E. Hillberg, A. Zegers, B. Herndler, S. Wong, J. Pompee, J.-Y. Bourmaud, S. Lehnhoff, G. Migliavacca, K. Uhlen, I. Oleinikova et al., "Flexibility needs in the future power system," 2019.
- [23] X. Yang, C. Xu, H. He, W. Yao, J. Wen, and Y. Zhang, "Flexibility provisions in active distribution networks with uncertainties," IEEE Transactions on Sustainable Energy, vol. 12, no. 1, pp. 553–567, 2020.
- [24] O. Agbonaye, P. Keatley, Y. Huang, O. O. Ademulegun, and N. Hewitt, "Mapping demand flexibility: A spatio-temporal assessment of flexibility needs, opportunities and response potential," Applied Energy, vol. 295, p. 117015, 2021.
- [25] A. Haque, M. Rahman, P. Nguyen, and F. Blik, "Smart curtailment for congestion management in lv distribution network," in 2016 IEEE Power and Energy Society General Meeting (PESGM). IEEE, 2016, pp. 1–5.
- [26] G. Celli, F. Pilo, G. Pisano, S. Ruggeri, and G. G. Soma, "Risk-oriented planning for flexibility-based distribution system development," Sustainable Energy, Grids and Networks, vol. 30, p. 100594, 2022.
- [27] H. Kiani, K. Hesami, A. Azarhooshang, S. Pirouzi, and S. Safaee, "Adaptive robust operation of the active distribution network including renewable and flexible sources," Sustainable Energy, Grids and Networks, vol. 26, p. 100476, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352467721000473>
- [28] M. Rayati, M. Bozorg, M. Carpita, and R. Cherkaoui, "Stochastic optimization and markov chain-based scenario generation for exploiting the underlying flexibilities of an active distribution network," Sustainable Energy, Grids and Networks, vol. 34, p. 100999, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352467723000073>
- [29] L. A. Roald, D. Pozo, A. Papavasiliou, D. K. Molzahn, J. Kazempour, and A. Conejo, "Power systems optimization under uncertainty: A review of methods and applications," Electric Power Systems Research, vol. 214, p. 108725, 2023.
- [30] A. Ben-Tal and A. Nemirovski, "Robust optimization—methodology and applications," Mathematical programming, vol. 92, no. 3, pp. 453–480, 2002.
- [31] T. L. Santos, V. M. Cunha, and A. Mesbah, "Stochastic model predictive control with adaptive chance constraints based on empirical cumulative distributions," IFAC-PapersOnLine, vol. 53, no. 2, pp. 11 257–11 263, 2020.
- [32] K. Tagawa and S. Miyanaga, "Weighted empirical distribution based approach to chance constrained optimization problems using differential evolution," in 2017 IEEE Congress on Evolutionary Computation (CEC). IEEE, 2017, pp. 97–104.
- [33] P. Petsagkourakis, I. O. Sandoval, E. Bradford, F. Galvanin, D. Zhang, and E. A. del Rio-Chanona, "Chance constrained policy optimization for process control and optimization," Journal of Process Control, vol. 111, pp. 35–45, 2022.
- [34] T. Van Acker, F. Geth, A. Koirala, and H. Ergun, "General polynomial chaos in the current–voltage formulation of the optimal power flow problem," Electric Power Systems Research, vol. 211, p. 108472, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0378779622006022>
- [35] H. Zhang, Z. Hu, E. Munsing, S. J. Moura, and Y. Song, "Data-driven chance-constrained regulation capacity offering for distributed energy resources," IEEE Transactions on Smart Grid, vol. 10, no. 3, pp. 2713–2725, 2018.
- [36] S. Blumsack, P. Hines, M. Patel, C. Barrows, and E. C. Sanchez, "Defining power network zones from measures of electrical distance," in 2009 IEEE Power & Energy Society General Meeting. IEEE, 2009, pp. 1–8.

- [37] R. J. Sánchez-García, M. Fennelly, S. Norris, N. Wright, G. Niblo, J. Brodzki, and J. W. Bialek, "Hierarchical spectral clustering of power grids," IEEE Transactions on Power Systems, vol. 29, no. 5, pp. 2229–2237, 2014.
- [38] J. Ding, Q. Zhang, S. Hu, Q. Wang, and Q. Ye, "Clusters partition and zonal voltage regulation for distribution networks with high penetration of pvs," IET Generation, Transmission & Distribution, vol. 12, no. 22, pp. 6041–6051, 2018.
- [39] M. Zhang, Z. Miao, and L. Fan, "Power grid partitioning: Static and dynamic approaches," in 2018 North American Power Symposium (NAPS). IEEE, 2018, pp. 1–6.
- [40] F. Retorta, C. Gouveia, G. Sampaio, R. Bessa, and J. Villar, "Local flexibility need estimation based on distribution grid segmentation," in 2022 18th International Conference on the European Energy Market (EEM). IEEE, 2022, pp. 1–6.
- [41] V. Rigoni, L. F. Ochoa, G. Chicco, A. Navarro-Espinosa, and T. Gozel, "Representative residential lv feeders: A case study for the north west of england," IEEE Transactions on Power Systems, vol. 31, no. 1, pp. 348–360, 2015.
- [42] E. Cotilla-Sanchez, P. D. Hines, C. Barrows, S. Blumsack, and M. Patel, "Multi-attribute partitioning of power networks based on electrical distance," IEEE Transactions on Power Systems, vol. 28, no. 4, pp. 4979–4987, 2013.
- [43] J. P. Hespanha, "An efficient matlab algorithm for graph partitioning," Santa Barbara, CA, USA: University of California, 2004.
- [44] B. Mourad, "On a spectral property of doubly stochastic matrices and its application to their inverse eigenvalue problem," Linear algebra and its applications, vol. 436, no. 9, pp. 3400–3412, 2012.
- [45] F. Scarlatache, G. Grigoraş, G. Chicco, and G. Cârţină, "Using k-means clustering method in determination of the optimal placement of distributed generation sources in electrical distribution systems," in 2012 13th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM). IEEE, 2012, pp. 953–958.
- [46] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," Journal of computational and applied mathematics, vol. 20, pp. 53–65, 1987.
- [47] H. Tang and S. Wang, "Energy flexibility quantification of grid-responsive buildings: Energy flexibility index and assessment of their effectiveness for applications," Energy, vol. 221, p. 119756, 2021.
- [48] A. Ulbig and G. Andersson, "Analyzing operational flexibility of electric power systems," International Journal of Electrical Power & Energy Systems, vol. 72, pp. 155–164, 2015.
- [49] M. U. Hashmi, A. Mukhopadhyay, A. Bušić, and J. Elias, "Storage optimal control under net metering policies," arXiv preprint arXiv:2002.01524, 2020.
- [50] Y. Chen, M. U. Hashmi, J. Mathias, A. Bušić, and S. Meyn, "Distributed control design for balancing the grid using flexible loads," in Energy Markets and Responsive Grids. Springer, 2018, pp. 383–411.
- [51] A. Bušić, M. U. Hashmi, and S. Meyn, "Distributed control of a fleet of batteries," in 2017 American Control Conference (ACC). IEEE, 2017, pp. 3406–3411.
- [52] M. U. Hashmi, A. Koirala, H. Ergun, and D. Van Hertem, "Perspectives on distribution network flexible and curtailable resource activation and needs assessment," IEEE Transactions on Industry Applications, 2023.
- [53] —, "Flexible and curtailable resource activation in a distribution network using nodal sensitivities," in 2021 International Conference on Smart Energy Systems and Technologies (SEST). IEEE, 2021, pp. 1–6.
- [54] A. J. Conejo and X. Wu, "Robust optimization in power systems: a tutorial overview," Optimization and Engineering, pp. 1–23, 2021.
- [55] L. Baringo and A. J. Conejo, "Offering strategy via robust optimization," IEEE Transactions on Power Systems, vol. 26, no. 3, pp. 1418–1425, 2010.
- [56] D. Bertsimas and A. Thiele, "Robust and data-driven optimization: modern decision making under uncertainty," in Models, methods, and applications for innovative decision making. INFORMS, 2006, pp. 95–122.

- [57] E. E. Elattar, A. M. Shaheen, A. M. Elsayed, and R. A. El-Sehiemy, "Optimal power flow with emerged technologies of voltage source converter stations in meshed power systems," *IEEE Access*, vol. 8, pp. 166 963–166 979, 2020.
- [58] Y. Li and Y. Li, "Two-step many-objective optimal power flow based on knee point-driven evolutionary algorithm," *Processes*, vol. 6, no. 12, p. 250, 2018.
- [59] K. Knezović, A. Soroudi, A. Keane, and M. Marinelli, "Robust multi-objective pq scheduling for electric vehicles in flexible unbalanced distribution grids," *IET Generation, Transmission & Distribution*, vol. 11, no. 16, pp. 4031–4040, 2017.
- [60] K. Deb and H. Gupta, "Searching for robust pareto-optimal solutions in multi-objective optimization," in *International conference on evolutionary multi-criterion optimization*. Springer, 2005, pp. 150–164.
- [61] D. Kaplan. Knee point MATLAB Central File Exchange. [Online]. Available: <https://www.mathworks.com/matlabcentral/fileexchange/35094-knee-point>
- [62] A. Koirala, L. Suárez-Ramón, B. Mohamed, and P. Arbolea, "Non-synthetic european low voltage test system," *International Journal of Electrical Power & Energy Systems*, vol. 118, p. 105712, 2020.
- [63] "Network and load data Github," Online, <https://github.com/umar-hashmi/FNADData>.
- [64] C. Coffrin, R. Bent, K. Sundar, Y. Ng, and M. Lubin, "Powermodels. jl: An open-source framework for exploring power flow formulations," in *2018 Power Systems Computation Conference (PSCC)*. IEEE, 2018, pp. 1–8.
- [65] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.
- [66] M. U. Hashmi, D. Deka, A. Bušić, and D. Van Hertem, "Can locational disparity of prosumer energy optimization due to inverter rules be limited?" *IEEE Transactions on Power Systems*, 2022.
- [67] J. Miettinen, H. Holttinen, and B.-M. Hodge, "Simulating wind power forecast error distributions for spatially aggregated wind power plants," *Wind Energy*, vol. 23, no. 1, pp. 45–62, 2020.
- [68] S. S. Torbaghan, G. Suryanarayana, H. Höschle, R. D'hulst, F. Geth, C. Caerts, and D. Van Hertem, "Optimal flexibility dispatch problem using second-order cone relaxation of ac power flows," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 98–108, 2019.
- [69] M. U. Hashmi, W. Labidi, A. Bušić, S.-E. Elayoubi, and T. Chahed, "Long-term revenue estimation for battery performing arbitrage and ancillary services," in *IEEE Int. Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. IEEE, 2018, pp. 1–7.
- [70] G. Küpper, F. Promel, F. Benothman, B. Czarnecki, R. Magulski, and L. M. Carvalho. Deliverable: D3.3 system-level assessment framework for the quantification of available flexibility for enabling new grid services. [Online]. Available: <https://tinyurl.com/3sttrkd4>
- [71] G. S. Sampaio, F. Bockemühl, D. Brummund, K. Sinitsyna, M. Staudt, G. Milzer, M. Kaffash, C. Dumont, A. Debray, P. Crucifix, K. Vanthournout, R. D'hulst, M. Findura, M. U. Hashmi, and H. Ergun. Deliverable: D8.1 german demonstrator — demonstration of congestion management using market driven utilisation of flexibility options in a lv grid. [Online]. Available: <https://tinyurl.com/22ehrh6x>
- [72] G. de Almeida Terça, A. Delnooz, A. Sanjab, K. Kessels, and M. U. Hashmi. Deliverable: D5.2 methodology for dynamic distribution grid tariffs. [Online]. Available: <https://tinyurl.com/bdd35kvm>
- [73] M. U. Hashmi, D. Brummund, R. Lundholm, A. Koirala, and D. Van Hertem, "Consensus based phase connectivity identification for distribution network with limited observability," *arXiv preprint arXiv:2301.03938*, 2023.
- [74] A. Chakraborty, "Generating multivariate correlated samples," *Computational Statistics*, vol. 21, no. 1, pp. 103–119, 2006.
- [75] C. KI Williams, *Gaussian processes for machine learning*. Taylor & Francis Group, 2006.

- [76] J. Orduz, “Sampling from a multivariate normal distribution,” blog, March, 2019 [Online]. [Online]. Available: <https://tinyurl.com/bcchm6c9>
- [77] T. Muschinski, G. J. Mayr, T. Simon, N. Umlauf, and A. Zeileis, “Cholesky-based multivariate gaussian regression,” Econometrics and Statistics, 2022.

APPENDIX

The focus of the present paper is to consider probable future uncertainties to provide FNA. This tool is developed as part of the EUniversal project to be implemented by the distribution system operator Mitnetz Storm in Germany [71]. Due to German data limitations, only historical forecast errors and point forecasts are available for scenario generation. This has led us to opt for a simple scenario generation model with data limitation. The generated scenarios using multivariate Gaussian distribution and Cholesky decomposition considers (a) the correlation of data points, and (b) parameter distribution. The Cholesky factor is utilized to form unique scenarios emulating different uncertainties. This structure of scenario generation can be used for historical data and point forecasts with known forecast errors. In the present case, large historical data is unavailable, instead, a point forecast along with its associated forecast error is known. In such a case, a multivariate Gaussian distribution can be formed by using point time-series forecast as the mean value and the variance of the distribution is proportional to forecast error [74].

A. Multivariate distribution and Cholesky decomposition

The multivariate Gaussian distribution is a generalization of the one-dimensional normal distribution to more than one dimension. For our case, X_i denotes the load profile of consumers connected at time i . The multivariate normal distribution of a S -dimensional random vector $\mathbb{X} = (X_1, X_2, \dots, X_S)^T$ can be written as $\mathbb{X} \sim N(\mu, \Sigma)$, with S -dimensional mean vector is given as $\mu =$

$$\mathbf{E}[\mathbb{X}] = \begin{bmatrix} \mathbf{E}[\mathbb{X}_1], \dots, \mathbf{E}[\mathbb{X}_S] \end{bmatrix}^T, \text{ and the covariance matrix is given as } \Sigma = \begin{bmatrix} \text{Cov}(\mathbb{X}_1, \mathbb{X}_1) & \dots & \text{Cov}(\mathbb{X}_1, \mathbb{X}_S) \\ \vdots & \ddots & \vdots \\ \text{Cov}(\mathbb{X}_S, \mathbb{X}_1) & \dots & \text{Cov}(\mathbb{X}_S, \mathbb{X}_S) \end{bmatrix},$$

where $\text{Cov}(\mathbb{X}_i, \mathbb{X}_j)$ (also denoted as $\Sigma_{i,j}$) denote the covariance between \mathbb{X}_i and \mathbb{X}_j and is given as

$$\Sigma_{i,j} = \mathbf{E}[(\mathbb{X}_i - \mu_i)(\mathbb{X}_j - \mu_j)], \quad i, j \in \{1, \dots, S\}. \quad (12)$$

Each component X_i has a distribution $N(\mu_i, \sigma_i^2)$ with $\Sigma_{ii} = \sigma_i^2$. The covariance matrix is a square matrix. To qualify as a covariance matrix, Σ must be symmetric ($\Sigma = \Sigma^T$) and positive semi-definite ($x^T \Sigma x \geq 0$).

In order to ensure that Cholesky decomposition does not fail, a small error is added to the covariance matrix for numerical reasons, e.g. $\Sigma = \Sigma + \epsilon I_S$, where I_S denotes identity matrix of order S , ϵ denotes small error value. This ensures that eigenvalues of Σ do not decay rapidly, which

stabilizes the decomposition. Due to the small magnitude of ϵ , it has inconsequential effects on the samples while ensuring numerical stability [75], [76].

Cholesky decomposition is used to calculate the lower triangular matrix, L , for the covariance matrix such that $\Sigma = LL^T$. L is also called the Cholesky factor. A lower triangular matrix is convenient because it reduces the calculation of a scenario as $\mu + LZ$ where $Z \sim N(0, I)$ to

$$\begin{aligned} X_1 &= \mu_1 + l_{11}z_1, \\ X_2 &= \mu_2 + l_{21}z_1 + l_{22}z_2, \dots \\ X_S &= \mu_S + l_{S1}z_1 + l_{S2}z_2 + \dots + l_{SS}z_S \end{aligned}$$

where z_1, z_2, \dots, z_S are independent and identically distributed random variables. The Cholesky factor is utilized to form unique scenarios [77] and is calculated using eigenvalue factorization. Since Σ is symmetric and positive semi-definite, therefore, its eigenvalues $\lambda_1, \dots, \lambda_S$ are non-negative. The covariance matrix can be denoted as $\Sigma = V \Lambda V^T$, where V is an orthogonal matrix, i.e. $VV^T = I$, Λ is a diagonal matrix with eigenvalues as diagonal entries. In this case, $L = V\Lambda^{0.5}$.

Consider J scenarios are generated for a time series with S intervals. The scenario matrix is given as

$$S_J = \begin{bmatrix} \mu_1 & \dots & \mu_1 \\ \mu_2 & \dots & \mu_2 \\ \vdots & \dots & \vdots \\ \mu_S & \dots & \mu_S \end{bmatrix}_{S,J} + \begin{bmatrix} l_{11} & \dots & 0 \\ l_{21} & \dots & 0 \\ \vdots & \ddots & \vdots \\ l_{S1} & \dots & l_{SS} \end{bmatrix} \text{rnd}(S, J), \quad (13)$$

where $\text{rnd}(S, J) \sim N(0, I)$ denotes uniformly distributed random numbers. (13) is used to generate J number of unique forecast profiles.

B. Temporal and spatially correlated scenarios

The proposed scenario generation tool considers the point forecast as the mean shown in (13). The covariance matrix is a function of forecast error calculated based on historical data. The forecast error is assumed to be 1.96 times the standard deviation of the normal distribution. This would ensure that more than 95% of the incidents are covered. The extreme tail events are ignored in the scenario generation. The scenario generation does not consider the spatial correlation. However, the generated scenarios do not ignore spatial correlation entirely. The net load of a node consists of solar generation as DG and load. Since the consumption of electricity is assumed positive, therefore,

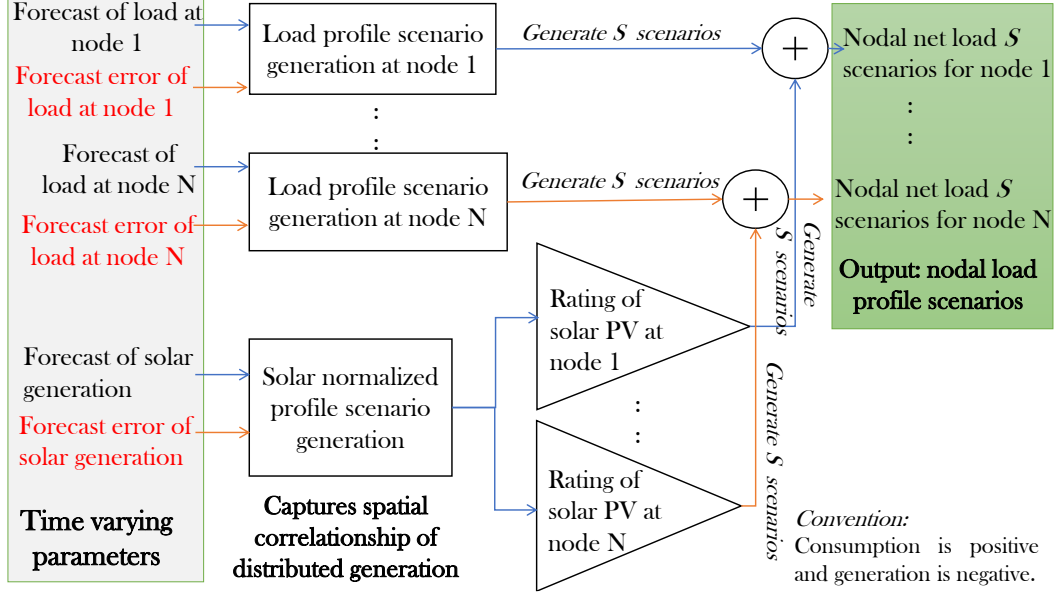


Fig. 19: Scenario generation decoupling load and generation to consider temporal and spatial correlations.

the load is strictly non-negative, and DG is a negative load. The load profile and PV generation are sampled separately depending on load forecast error and its point forecast, which are assumed to be known. On the other hand, DG scenarios are calculated based on their point forecast and associated forecast errors. The DG forecast scenarios are normalized in per kW installed capacity. These scenarios are assumed to be the same for all DGs in the feeder, assuming all DGs are installed in geographical proximity with the same solar irradiance. The average spread of 160 Spanish feeders, [62], is a square of the size of 158 meters. Therefore, this assumption would be valid for most DNs. Fig. 19 shows the input and output of scenario generation.

The silhouette value of a member $j \in M_x$ is calculated as, $s_j = \frac{b_j - a_j}{\max\{b_j, a_j\}}$, where a_i denotes the mean distance between member j and all other members in the group M_x and b_j denotes the mean distance between node j and all other nodes not in the group M_x . The mean silhouette value of a group M_x is given as $S_x = \frac{1}{|M_x|} \sum_{i=1}^{|M_x|} s_i$, where $|M_x|$ denotes the number of members in the group M_x . The maximum of the mean silhouette scores of a k partitioned network, i.e., *silhouette coefficient*, is given as $SC_T^k = \max(S_x, \forall x = 1, 2, \dots, k)$. For a given network, vary k to maximize the value of SC_T^k . Note SC_T^k may be high for a small number of clusters which may not suit the application, therefore, the selection of best k which increases SC_T^k , is a design problem.

The algorithm for selecting the best number of zones that maximizes the *silhouette coefficient* is

detailed in Algorithm 2. The output of the algorithm is the ideal number of zones for our flexibility assessment application.

Algorithm 2 Zones of distribution network.

Inputs: Network details, $SC_T^k = []$,

- 1: Calculate admittance matrix, Y , for the network,
 - 2: Make the electrical distance matrix double stochastic,
 - 3: Set value of $k = 2$ clusters.
 - 4: Calculate k the largest eigenvalues and eigenvectors.
 - 5: Use k -means clustering and calculate SC_T^k and concatenate,
 - 6: Increment k till $k \leq N$ and Go to Step 4,
 - 7: Analyze SC_T^k vector to select the best suited number of zones. The best-suited number of zones depends on (a) how many clusters will meet the required application for which clusters are formed and (b) the silhouette coefficient of the cluster. For example, if the *silhouette coefficient* is very high for very few numbers of clusters, but it does not meet our purpose, then we may opt for a *silhouette coefficient*.
-

The network diagram and the zones are shown in Fig. 20.

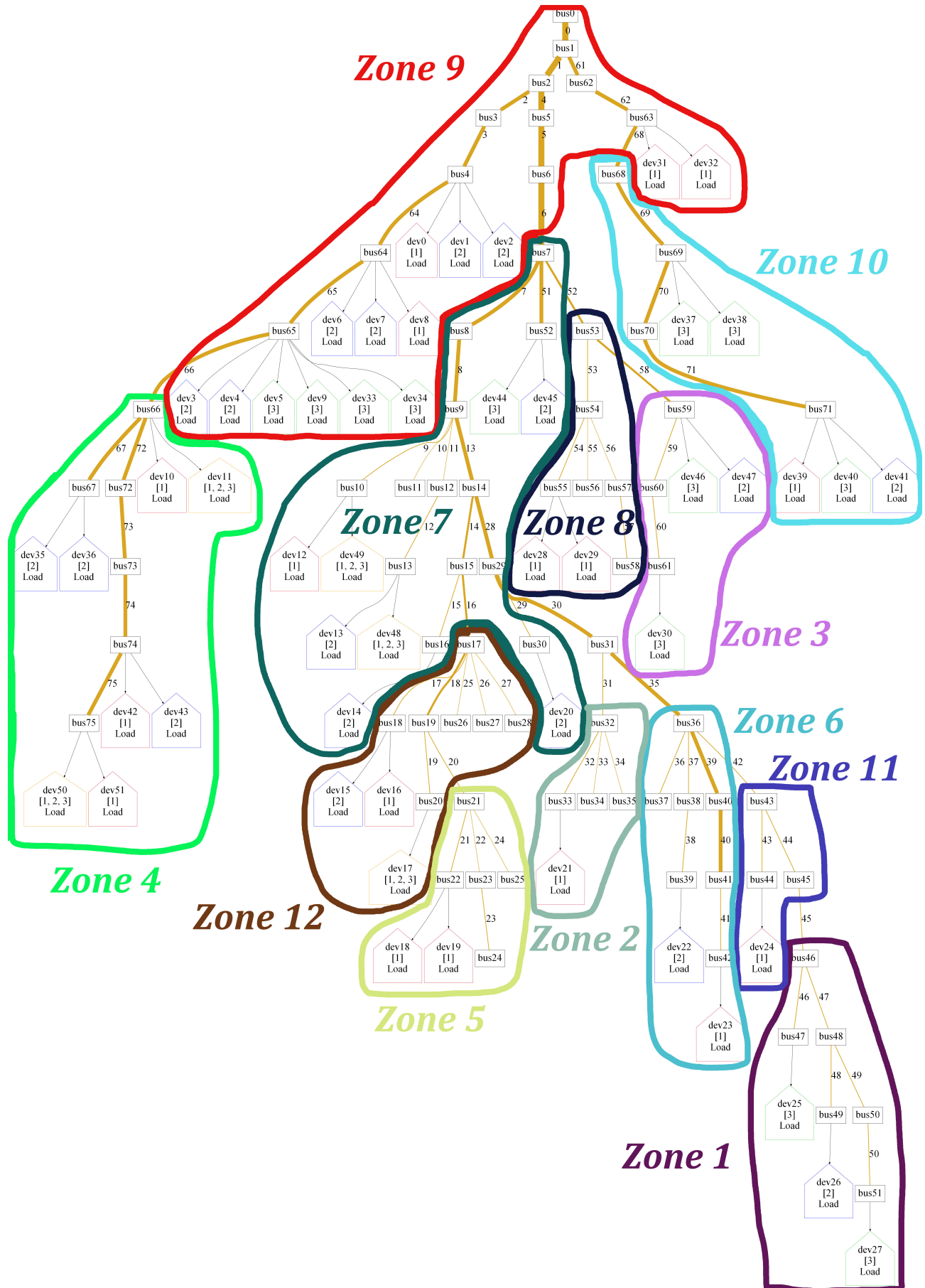


Fig. 20: Network diagram with 12 zones indicated.