DYNAMIC ZONAL FLEXIBILITY MARKETS FOR CONGESTION MANAGEMENT: OPTIMISING DISTRIBUTION GRID CAPACITY

Victor S. Aasvær[1], Anders S. Ryssdal[1], Pedro Crespo del Granado[1], Asgeir Tomasgard[1], Selina Kerscher[1,2]*

[1]*Dept. of Industrial Economics and Technology Management, Norwegian University of Science and Technology, Trondheim, Norway*
[2]*Dept. of Electrical, Electronic, Communications and Systems Engineering, University of Oviedo, Gijón, Spain*
**UO286905@uniovi.es*

Keywords: ZONAL FLEXIBILITY MARKET, CONGESTION MANAGEMENT, SYSTEM COST

Abstract

The integration of volatile renewable energy sources into the electric grid and the corresponding rise in power demand due to electrification is creating challenges for congestion management (CM). CM aims at optimising the grid capacity utilisation and thereby enables hosting more distributed energy resources. Flexible resources have the potential to reduce the costs associated with CM, which raises the demand for an efficient framework for trading flexibility. The aggregated flexibility from stakeholders such as energy communities serves for valid bids into large-scale flexibility markets. This paper develops a market design aimed at improving CM by examining the concept of zonal flexibility. The design utilises zones that are dynamically allocated based on grid congestion after the day-ahead market. The proposed market design is compared to both the optimal utilisation of flexible resources and to conventional redispatch in order to evaluate its economic efficiency. Our findings indicate that it performs close to a nodal flexibility market. Future work needs to develop more refined partitioning algorithms to stabilise the performance across various scenarios.

1 Introduction

As the power system transitions to electrified economies and increases deployment of renewable energy technologies to mitigate climate change, it faces new challenges in managing congestion in the power grid. The power grid was designed for small increments on projected loads, and now is under strain due to rapidly accelerating electrification (e.g. transport and heat). In addition, the intermittent nature of renewable energy sources (RES) introduces new challenges for congestion management (CM), raising the risk of grid instability.

Managing congestion is critical to ensure reliable grid operations. Well-functioning and efficient CM is an enabler for hosting more distributed energy resources (DER) in the grid. One key approach to achieving this is through market-based flexibility procurement. Villar et al. list the participants in flexibility markets (FMs) as: transmission and distribution system operators (TSOs and DSOs), balancing responsible parties (BRPs), aggregators and retailers [1]. TSOs buy flexibility for balancing needs, while both TSOs and DSOs buy flexibility for CM and voltage control. BRPs buys flexibility to balance their own portfolio, while aggregators and retailers provide flexibility services [1] [2]. An effective use of flexibility requires a robust market framework for trading [3], which in turn will create a more resilient and responsive energy system that is better able to meet the challenges of the future and host more DER.

Centralised optimisation models are common to model flexibility markets [1]. Numerous authors have presented FMs that are cleared by minimising the operational costs of a particular participant [4] [5] [6] [7] [8]. Spiliotis et al. analyse the trade-off between grid expansion and flexibility dispatch by developing a long-term planning model which minimise the DSO's total cost [4]. In [5], a framework for trading prosumer flexibility used to minimise the DSO's cost of resolving network congestion is presented. Zhang et al. analyse a FM where a DSO procures flexibility from aggregators to relieve congestions in local grids [6]. Three trading setups representing different contractual arrangements between DSOs and aggregators are presented. The work is extended in [7], where two of the trading setups are analysed in more detail and a quantitative example is provided. In contrast, [8] presents a FM design where an aggregator operates the market and minimise the cost of dispatching its flexibility units.

Centralised optimisation models maximising social welfare have also been discussed in the literature [9] [10]. In [9], Jiang et al. argue for the need for a joint energy and flexibility market clearing to capture the interaction between the TSO and DSOs. The authors propose a bi-level optimization problem, where the upper-level problem maximises social welfare and the lower-level problem minimises the cost of supplying flexibility to the TS. In [10], a framework for trading flexibility consisting of a day-ahead (DA) and an intraday market cleared by an independent third party is presented. Both the DA and intraday markets are cleared with the objective of maximising social welfare.

The literature review highlights that many papers focus on minimising one of the participant's costs, and all papers consider models with a nodal grid resolution. The inclination to prioritise the minimisation of one participant's costs can be attributed to the geographical extent of the named studies. Many of these papers solely concentrate on the low-voltage DS, where DSOs and aggregators are the primary participants in the FM. Despite some mention of the need for coordination between the TSO and DSOs for the interaction between the transmission and distribution systems, these studies largely overlook the impact of flexibility dispatch on the TS. Moreover, only a handful of studies analyse FMs wherein both the TSO and DSOs function as procurers of flexibility.

While there is an extensive body of literature on zonal versus nodal pricing in DA markets, zonal FMs have not received much attention in the research. Ramos et al. discuss the concept, pointing out that nodal pricing has been criticised for its difficulty in explaining certain nodal prices [3]. However, there is a research gap on the design and modelling of zonal FM. The notion of dynamic zones has not yet been applied to FMs.

In this paper we provide an initial investigation into zonal FMs by proposing a novel FM design aimed at improving CM. A central idea behind the zonal market design is to guarantee liquidity and other advantages associated with the grouping of nodes into zones, while concurrently conveying accurate pricing signals, thus ensuring the optimal utilisation of flexibility. The grid congestion is addressed by the system operators, which participate in the market as flexibility procurers.

Theoretically, zonal markets are less economically efficient than nodal markets as they do not account for all system information. Furthermore, these markets may experience internal congestion after clearing, necessitating a traditional re-dispatch at a nodal level by the system operators. The implementation of a zonal FM thus incurs an alternative cost that this paper aims to assess.

Zonal FMs offer several potential benefits that warrant further investigation. Firstly, zonal markets may be more accessible to various actors. For example, aggregators that represent flexibility across multiple nodes may find it easier to participate in a zonal market with zonal-based bidding. Additionally, zonal market prices may be politically more acceptable to market actors, as price differences can be readily attributed to significant grid congestions.

A zonal FM may also differ from a nodal FM in terms of liquidity, price signal effects, market power, and gaming vulnerability. It is important to note that these benefits are uncertain and require further investigation. Advanced modelling beyond the scope of this paper is necessary to analyse these potential advantages. This paper solely focuses on short-term economic efficiency of zonal FMs, leaving the investigation of zonal FM benefits for future research.
The posed research questions are:

- How can zonal FMs contribute to facilitating more efficient CM?
- To what extent can dynamic zones limit the efficiency loss associated with zonal pricing in FMs?

To answer these questions, the zonal FM is modelled and bench-marked against a nodal FM and a case where flexibility is unavailable.

The main contributions are:

- A basic conceptualisation of zonal FMs.
- A novel market design for a multi seller, multi buyer FM connected to both the TS and the DS, designed to improve CM after DA clearing.
- An analysis of the role of dynamic zones in FMs.

Section 2 describes the applied methodology, while section 3 presents and discusses the results and several sensitivities. Finally, section 4 concludes the key findings.

2 Methodology

This section elaborates on the proposed market design and modelling, the zonal partitioning problem, the mathematical formulation of the study cases and data.

2.1 Flexibility market design and modelling

The proposed FM aims to enhance the management of intra-zonal congestion after day-ahead by providing system operators (SO) with more flexible resources. The market operates through discrete auctions held throughout the day with an hourly time resolution for bids, similar to both the DA and intraday markets. The FM is administered by an independent third party, with SO serving as flexibility procurers.

The core concept of the FM is to dynamically divide the grid into zones, based on the internal congestion arising from the DA clearing. This zoning strategy enables the SO to mitigate congestion through inter-zonal trading. It is worth noting that the optimal zone sizes may be strongly influenced by the degree and location of congestion. The proposed FM zones are distinct from the DA zones and are hence referred to as local flexibility zones (LFZs).

SO are engaged in the FM as flexibility procurers. Other considered participants are consumer aggregators, also representing energy communities, energy storage, demand-side industry, and intermittent power producers. This list aims to capture the expected behaviours and characteristics of the FM's participants, except for non-intermittent generation, which is assumed to provide only re-dispatch options. While CM is the primary focus of this paper, the market is also open to BRPs and prosumers seeking to re-balance their portfolios.

The proposed zonal FM is modelled using a four-step approach illustrated in Figure 1. In the first step, a load flow analysis based on the disaggregated DA volumes calculates the congestion on each line. In the second step, the configuration of the LFZs is determined based on the congestions. Steps 1 and 2 are performed independently for each hour. Section 2.2 presents the heuristic for obtaining the optimal zonal configuration.

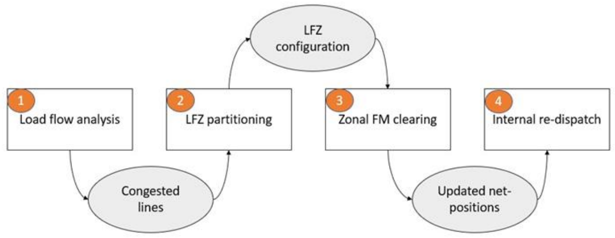

Fig. 1 Modelling approach for the Zonal FM case.

As step 3, the zonal market clearing problem is solved based on the LFZs from step 2. Congestion may still be present within the LFZs after the clearing since the zonal FM clearing ignores intrazonal constraints and the available flexibility may be insufficient to alleviate the interzonal congestion entirely. Step 4 addresses this issue by utilising conventional re-dispatch. The model in step 4 uses a nodal grid, and the starting net positions are equal to the net production obtained in step 3. The formulations for the models used in step 3 and 4 can be found in Section 2.3 and the nomenclature in Appendix B.

2.2 Zonal partitioning problem

The zonal partitioning problem has the dual goal of maximising the amount of congestion being considered in the market clearing while creating zones that are practical for the market participants, securing liquid markets. Having multiple objectives complicates the problem which is addressed by solely considering congestion, converting the other objectives to constraints or solution-requirements. As a result, the zonal partitioning problem solved for this paper concerns maximising the amount of congestion in inter-zonal lines, with requirements for a maximum number of zones and a minimum number of grid nodes per zone. This can be defined as a maximum k-cut problem, an NP-complete problem that involves partitioning a graph into k connected components [11]. As our graph- partitioning problem is NP-complete, it must be solved using a heuristic.

The zonal partitioning heuristic takes as input a set of nodes, lines and associated congestion volumes. The algorithm follows an iterative approach where one zone is divided into two at each iteration. The selection of the zone to be split is based on finding the best cut, which maximises the amount of congestion between the two resulting zones. The algorithm ends when there is no more internal congestion or when the desired number of zones has been reached. A more detailed description of the algorithm is presented in Appendix A.

An important aspect of the algorithm is the choice to only split zones in two. This is an advantage with regards to simplicity, but it may come at the cost of losing potential optimal solutions. Additionally, the heuristic may not find the optimal way of splitting a certain zone in two. There are thus several ways of losing optimal solutions in this heuristic, but it has proven to be both fast and easy to implement.

It is also important to note that maximising the inter-zonal congestion does not necessarily result in the most efficient zonal configuration. The performance of a zonal configuration is, for example, dependent on how the FM clearing will affect internal congestion, which is neither considered in the heuristic nor in the zonal model itself. Future research should therefore investigate the zonal FM performance of more sophisticated partitioning algorithms. One example is [12], who present a zone partitioning method using power transfer distribution factors (PTDFs) and spectral clustering.

2.3 Mathematical formulation of the study cases

To assess the efficacy of the zonal FM, two additional modelling cases are introduced: nodal FM and business-as-usual (BAU). The nodal FM case utilises a nodal grid approach to clear the flexibility market, providing a benchmark for the most efficient use of the flexibility resources. The BAU case also uses a nodal grid formulation, but it only considers the re-dispatch resources available outside the FM. This case corresponds to only running step 4 in the zonal FM case and represents the threshold beyond which a zonal FM will lead to increased costs to society, equivalent to a worst-case benchmark.

The nodal FM and the BAU case can each be modelled in a single step. The nodal FM model serves as a foundation since it shares most of its notation with the other models. The formulations for these models are presented subsequently.

2.3.1 Nodal FM model: The objective function of the nodal FM model minimises the total cost of flexibility and re-dispatch over a 24-hour period.

$$\min \sum_{t\in\mathcal{T}}\Big(\sum_{b\in\mathcal{B}^+} C_{tb}x_{tb}^+ + \sum_{b\in\mathcal{B}^-} C_{tb}x_{tb}^-\Big) \qquad (1)$$

where $\mathcal{B}^+$ and $\mathcal{B}^-$ are sets containing the up-regulatory and down-regulatory bids, respectively. In the model context, bids can either represent available flexibility or available redispatch. The power flow in each line is calculated using PTDFs.

$$P_{nt}^{DA} + \sum_{b \in \mathcal{B}_n^+} x_{tb}^+ - \sum_{b \in \mathcal{B}_n^-} x_{tb}^- = p_{nt} \qquad n \in \mathcal{N}, t \in \mathcal{T} \tag{2}$$

$$f_{lt} = \sum_{n \in \mathcal{N}} PTDF_{ln} \cdot p_{nt} \qquad l \in \mathcal{L}, t \in \mathcal{T} \tag{3}$$

$$-CAP_l \leq f_{lt} \leq CAP_l \qquad l \in \mathcal{L}, t \in \mathcal{T} \tag{4}$$

In Equation 2, net positions are calculated hourly at each node by adjusting the DA volumes with the bid volumes cleared up and down. The power flow is determined in accordance with Equation 3, while complying with the capacity constraints specified in Constraint 4.

The nodal FM model only has one stage, which means it considers both flexibility bids and available re-dispatch volumes simultaneously. This is a valid simplification when assuming that the SO have perfect information about the available re-dispatch, but it requires two distinct balancing constraints: one to accommodate adjustments made by flexible resources (Equation 5), and another to factor in conventional re-dispatch (Equation 6).

$$\sum_{b \in \mathcal{B}^+ \setminus \mathcal{B}^{RE}} x_{tb}^+ - \sum_{b \in \mathcal{B}^- \setminus \mathcal{B}^{RE}} x_{tb}^- = 0 \qquad t \in \mathcal{T} \tag{5}$$

$$\sum_{b \in \mathcal{B}^+ \cap \mathcal{B}^{RE}} x_{tb}^+ - \sum_{b \in \mathcal{B}^- \cap \mathcal{B}^{RE}} x_{tb}^- = 0 \qquad t \in \mathcal{T} \tag{6}$$

We introduce constraints ensuring the authenticity of the battery participants' behaviour.

$$\sigma_{it} = \sigma_{i(t-1)} + \sum_{b \in \mathcal{B}_i^-} x_{tb}^- \cdot \eta^- - \sum_{b \in \mathcal{B}_i^+} x_{tb}^+ / \eta^+ \qquad i \in \mathcal{I}^{Batt}, t \in \mathcal{T} \setminus \{1\} \tag{7}$$

$$\sigma_{i1} = S_i^{init} \qquad i \in \mathcal{I}^{Batt} \tag{8}$$

$$\sigma_{i|T|} = S_i^{init} \qquad i \in \mathcal{I}^{Batt} \tag{9}$$

$$0 \leq \sigma_{it} \leq S_i^{Cap} \qquad i \in \mathcal{I}^{Batt}, t \in \mathcal{T} \tag{10}$$

Constraint 7 calculates the state of charge for battery actors. By requiring the state of charge to be greater or equal to zero in Constraint 10, Equation 7 ensures that battery actors cannot sell more power than they have available. Constraints 8 and 9 fix the initial and final state of charge of batteries. This results in a more restrictive operation than would be the case in a practical setting, especially as each day is modelled independently.

The aggregators' behaviour is subject to constraints. Constraint 11 stipulates that within a six-hour interval, an aggregator cannot curtail more than a fraction λ of the consumption scheduled in the DA clearing, $\mathcal{K}_i$. Similarly, constraint 12 limits the amount of extra power consumption during the day. Both constraints are meant to reflect that aggregators represent end-users and, despite being motivated by price signals, are unlikely to adjust their average consumption significantly.

$$\sum_{\tau=t}^{t+5} \Big(\sum_{b \in \mathcal{B}_i^+} x_{\tau b}^+ - \sum_{b \in \mathcal{B}_i^-} x_{\tau b}^- \Big) \leq \mathcal{K}_i \cdot 6\lambda \qquad i \in \mathcal{I}^{Agg}, t \in \mathcal{T} \setminus \{|T|-5, \ldots, |T|\} \tag{11}$$

$$\sum_{t \in \mathcal{T}} \Big(\sum_{b \in \mathcal{B}_i^-} x_{tb}^- - \sum_{b \in \mathcal{B}_i^+} x_{tb}^+ \Big) \leq \mathcal{K}_i \cdot 24\zeta \qquad i \in \mathcal{I}^{Agg} \tag{12}$$

Constraints 13 and 14 ensure that volumes cleared up and down are non-negative and are kept below the maximum volume of the bid.

$$0 \leq x_{tb}^+ \leq V_{tb}^+ \qquad t \in \mathcal{T}, b \in \mathcal{B}^+ \tag{13}$$

$$0 \leq x_{tb}^- \leq V_{tb}^- \qquad t \in \mathcal{T}, b \in \mathcal{B}^- \tag{14}$$

2.3.2 Zonal FM model: This subsection describes the model used in step 3 of the zonal FM case. Constraints 5-14 are identical to those in the nodal FM model, with the exception that re-dispatch bids are not included. Thus, equation 6 is absent and Constraint 13 and 14 only apply to flexibility bids, $b \in \mathcal{B} \setminus \mathcal{B}^{RE}$. Apart from this, the differences between the two case models lie in the objective function and the power flow constraints which are expounded upon below.

Since this case is a zonal market and the zones are dynamic, we introduce a set $\mathcal{Z}_t$ representing the zones in time period t. The net positions are calculated in Equation 15 similarly to Constraint 2 in the Nodal FM model, but the nodes are now aggregated into their respective zones. Hourly line flows for critical branches, $l \in \mathcal{L}_t$, are calculated using zonal PTDFS (ZPTDFs) indexed by time period, line, and zone. ZPTDFs are derived from PTDFs, and a detailed description of this derivation is presented in Appendix C.

Unlike the nodal FM model, the zonal FM model does not constrain line flows below the line capacity. This is because conventional re-dispatch is not available in the model clearing, and inadequate flexibility combined with hard flow constraints may lead to an infeasible solution. Hence, Constraint 17 introduces a variables y_{lt} which represents congestion on a critical branch and is penalised in the objective function. These variables are forced to be non-negative in Constraint 18.

$$P_{zt}^{DA} + \sum_{n\in\mathcal{N}_{zt}} \Big(\sum_{b\in\mathcal{B}_n^+\setminus\mathcal{B}^{RE}} x_{tb}^+ - \sum_{b\in\mathcal{B}_n^-\setminus\mathcal{B}^{RE}} x_{tb}^- \Big) = p_{zt} \qquad t\in\mathcal{T}, z\in\mathcal{Z}_t \tag{15}$$

$$f_{lt} = \sum_{z\in\mathcal{Z}_t} ZPTDF_{lzt}\cdot p_{zt} \qquad t\in\mathcal{T}, l\in\mathcal{L}_t \tag{16}$$

$$-CAP_l - y_{lt} \le f_{lt} \le CAP_l + y_{lt} \qquad t\in\mathcal{T}, l\in\mathcal{L} \tag{17}$$

$$0 \le y_{lt} \qquad t\in\mathcal{T}, l\in\mathcal{L} \tag{18}$$

The objective function for the zonal FM case is given by Equation 19. The two first terms correspond to the objective function of the nodal FM case, while the last term penalises congestions in critical branches, where γ represents the unit penalty of congestion.

$$\min \sum_{t\in\mathcal{T}} \Big(\sum_{b\in\mathcal{B}^+\setminus\mathcal{B}^{RE}} C_{tb}x_{tb}^+ + \sum_{b\in\mathcal{B}^-\setminus\mathcal{B}^{RE}} C_{tb}x_{tb}^- \Big) + \gamma \sum_{t\in\mathcal{T}} \sum_{l\in\mathcal{L}_t} y_{lt} \tag{19}$$

2.3.3 Zonal FM model: In the BAU case, where only conventional re-dispatch volumes are included, the model formulation closely resembles the model for the nodal FM case. The objective function remains the same with the exception that we sum over $\mathcal{B}^+\cap\mathcal{B}^{RE}$ and $\mathcal{B}^-\cap\mathcal{B}^{Re}$ since flexible resources are not available. Constraints 2-4, 6 and 13-14 are also present, while the remaining constraints are irrelevant.

Besides being the focus of the BAU case, a re-dispatch model must also be run in the zonal FM case after the FM clearing to guarantee alleviation of all congestions. This corresponds to step 4 in Figure 2.1. We then run the same re-dispatch model described in the previous paragraph, except that net positions must account for adjustments cleared in the zonal FM. Equation 2 is therefore replaced by Equation 20:

$$P_{nt}^{DA} + \left(\sum_{b\in\mathcal{B}_n^+\setminus\mathcal{B}^{RE}} X_{tb}^+ - \sum_{b\in\mathcal{B}_n^-\setminus\mathcal{B}^{RE}} X_{tb}^- \right) + \left(\sum_{b\in\mathcal{B}_n^+\cap\mathcal{B}^{RE}} x_{tb}^+ - \sum_{b\in\mathcal{B}_n^-\cap\mathcal{B}^{RE}} x_{tb}^- \right) = p_{nt} \qquad n\in\mathcal{N}, t\in\mathcal{T} \tag{20}$$

where X_{tb}^+ and X_{tb}^- are the volumes cleared up and down in the zonal FM, respectively.

2.4 Study data

The following subsections contain details on data resources from the transmission system, the DA market, the distribution system, the market participants and bids, the available redispatch, and further implementation details.

2.4.1 Transmission system: The TS used in modelling is based on real-life data for the Nordic TS from 2012 (see [13] for the full data set). Comprising a total of 446 nodes and 770 lines, the data set spans across Norway, Sweden, Denmark, and Finland, with voltage levels ranging between 110 kV and 420 kV. Each of these nodes is located within one of the 20 areas, displayed as N1-N11, S1-S6, D1-D2 and FI in Figure 2a. Additionally, the data set includes information about generator capacities and load volumes for the nodes, as well as capacities and reactances for the lines.

a

b

Fig. 2 (a) Areas in the TS data set [13], (b) Flexibility area with red TS nodes from the reference dataset and artificially constructed green DS nodes.

2.4.2 DA market data: The DA volumes for production and load is gathered from historic Nord Pool data, using the interval from 17th of November to 30th November 2022 [14]. The Nord Pool data is aggregated for each DA zone, so it needs to be disaggregated onto the nodes in the TS. The following principles are used in the disaggregation process:

- Each of the 20 areas in the TS data set described above is assigned to a designated DA zone in accordance with their respective geographic locations. As a result, each individual node within the TS data set is also associated with a specific DA zone.
- Utilising the load volumes and installed capacities belonging to the TS data set, nodal weights are computed for both production and consumption. These weights represent the proportion of the aggregate DA zonal volumes that are assigned to a particular node.
- As the load volumes and installed capacities from the TS data set are only indicative of a specific moment in time, their derived nodal disaggregation weights remain constant. To ensure more dynamic congestion patterns, stochastic perturbations are therefore introduced to the disaggregated nodal production and load volumes.

2.4.3 Distribution system: To investigate the effects of the proposed FM, we study a sub-part of the NO1 DA zone, encompassing the metropolitan Oslo area (hereafter referred to as the flexibility area (FA)). When modelling, this area is represented both by a high voltage TS and a medium voltage DS (see Figure 2b). Additionally, we include the entire Nordic TS in the modelling to ensure realistic system flows, and to observe whether the FM has any adverse effects on the surrounding TS.

While information about the TS is often openly available, DSOs do not commonly publish detailed information about the DS. We therefore design a DS specifically for this paper. The following principles were used when designing the DS:

- Each TS node is assumed to be a net production node and each DS node is assumed to be a net consumption node.
- Three example DS radial grids are designed and one of these is attached to each TS node. Before attaching these radials, they are scaled and rotated to fit the terrain.
- The disaggregated DA load calculated as described above is allocated to the DS nodes.

For a detailed description of the DS design, see Appendix D.

2.4.4 Market participants: The flexibility market participants, like the grid system, are assumed to remain constant throughout the project. They are thus generated and allocated to nodes before running cases and sensitivities. Attempting to best represent a real system, we only allocate them to DS nodes, with each participant limited to one node for simplicity.

When a participant is created and allocated to a DS node, the participant is also defined with a size based on the expected node load, and a cost scaling factor that is randomly chosen on a uniform distribution between 0.5 and 1.5. These properties are later used when designing market bids of the participant and in model constraints. For a detailed description of how market participants are allocated to nodes, see Appendix D.

2.4.5 Market bids: Available flexibility and re-dispatch volumes are represented by bids in the models. Each bid is characterised by its cost $[€/MWh]$, a maximum volume $[MWh]$, a node in the grid, an implementation hour and a direction. Flexibility bids are also associated with a market participant. The FM bids are important to represent the behaviour of the four participant types presented in section 2.1. We therefore try to adapt the bid frequency, volumes and cost based on how the participants are expected to behave in practice.

The direction of a bid says whether the bid will increase or decrease the net production in the node. An "UP" bid will in this context mean a bid that increases the net production, while a "DOWN" bid is defined as decreasing the net production in the node. These definitions are fitting from a market clearing perspective but are less intuitive when looking at the practical implications for various market participants. For example, an UP bid from an aggregator or industry actor represents an option to reduce their power consumption.

Bids are first generated deterministically based on a method described in Appendix D and cloned for each hour. Then, noise is introduced so that bids are differing in volumes and cost between hours, as well as removing bids from certain hours. This process is based on random variables and will give different results for each day and hour. The aim is to show how the FM clearing may vary between hours depending on available bids, and to show the solution's resilience to noise.

2.4.6 Available redispatch: Bids that represent conventional re-dispatch are not co-generated with flexibility bids. These participants can be regarded as supply-side actors as they are the most common providers of re-dispatch volumes. In the Norwegian power system, which is largely dominated by hydro power, re-dispatch costs are heavily influenced by water values, which are complex to calculate and beyond the scope of this work. To approximate these costs, we use the average price premium obtained from the balancing markets for the same period as we collect DA data [15]. Given that balancing operations are closer to real-time and offer fewer options for planning than re-dispatch, we argue that the balancing market price premium can be considered an upper bound for the cost of conventional re-dispatch.

2.4.7 Implementation details: The models are solved on a computer running on Windows 10, with Intel Core i7-10700 CPU and 16 GB of RAM. They are implemented in Python using the Pyomo package [16] and Gurobi solver with default settings [17]. Solving all three cases for a period of 14 days takes 4453 s.

Please see our repository on GitHub for more insight into the code implementation: LocalFlex_public.

3 Results

This section presents and discusses the model results focusing specifically on system cost and traded volumes.

3.1 System cost

The system cost is minimised for all three cases, which under the assumptions of rigid demand equals a maximisation of social welfare.

Figure 3 represents the daily FA costs for all cases and demonstrates that the cost of the zonal FM case closely tracks the cost of the nodal FM case throughout most of the simulated days. Focusing solely on the costs incurred in the FA, the nodal FM approach incurs costs of 5.946 M€, the zonal FM approach incurs costs of 6.071 M€, and the BAU strategy results in costs of 6.590 M€. This represents a relative cost reduction of 9.8% and 7.9% for the nodal FM and zonal FM approaches, respectively.

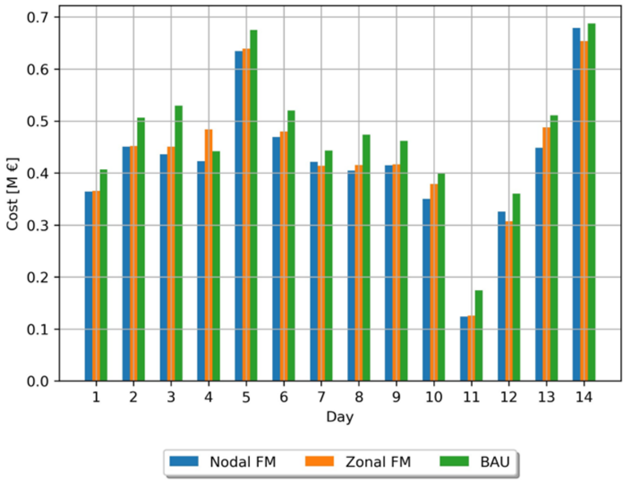

Fig. 3 Daily FA costs for all cases which sum up to 5.946 M€, 6.071 M€, and 6.590 M€ for nodal FM, zonal FM and BAU, respectively.

As mentioned in section 2.4, we include a model of TS for the whole Nordic region to ensure realistic line flows and study the effects of activating flexibility inside the FA. The implications are apparent in Figure 3, which reveal that the zonal FM may in fact be cheaper than the nodal FM, and more expensive than the BAU case for individual days. This is because the models may transfer costs into or out of the FA depending on the most optimal outcome for the overall system. Notably, when flexibility is accessible within the FA, the FA ends to bear a higher share of the total costs. This is evident when looking at the total system cost, which is found to be 124.843 M€, 124.897 M€, and 125.816 M€ for the nodal FM, zonal FM, and BAU cases, respectively. Relative to the FA BAU cost, the nodal and zonal cases then exhibit a performance increase of 14.8% and 13.9%, respectively. This result is notably superior to isolating the FA area costs.

The findings in Figure 3 align with the expectations outlined in Section 2.3. Specifically, the BAU case is identified as the costliest option, while the nodal FM approach is consistently the most economical alternative. The superiority of the nodal FM approach over the BAU case is unsurprising from a theoretical perspective. Additionally, it is well-established that zonal markets cannot theoretically outperform nodal markets in terms of cost efficiency. At best, the zonal method can emulate the nodal approach and clear the FM efficiently, but this is improbable due to less available grid information. Nevertheless, the results indicate that the zonal FM performance is significantly closer to the nodal than the BAU case across all simulated days.

3.2 Volumes

In Table 1, the total dispatch of flexibility units and redispatch for the whole system is presented for all three cases. The same numbers are presented for the FA in Table 1.

Table 1 Total system volume adjustments in GWh.

Type	Nodal FM	Zonal FM	BAU
Flexibility	27.312	27.161	0
Redispatch	2,495.678	2,496.858	2,516.318
Total	2,522.99	2,524.02	2,516.318

Table 2 Total flexibility area volume adjustments in GWh.

Type	Nodal FM	Zonal FM	BAU
Flexibility	27.312	27.161	0
Redispatch	117.735	120.354	131.811
Total	145.047	147.515	131.811

The total volume of flexibility adjustments in the nodal FM case amounts to 27,312 MWh, while that of the zonal FM case is 27,161 MWh. A slightly larger number of flexible resources is therefore utilised in the Nodal FM case. For both the total system and the FA, the sum of flexibility and conventional re-dispatch is lowest for the BAU case. This result is unsurprising, as the uniform cost of conventional re-dispatch effectively makes the BAU case minimise activation volumes.

For the FA, the volume of conventional re-dispatch is 2,619 MWh higher in the Zonal FM than in the Nodal FM case, which is a major reason for the cost differences presented in section 3.1. The larger volume of activated flexibility in the nodal FM case is one reason, but even more important is the effectiveness of the activated flexibility volumes. The ZPTDFs used in the zonal FM case are similar for all nodes in a zone, which translates to an absence of nodal information. The activated volumes of flexibility in the zonal FM case are therefore less effective in alleviating the congestion, necessitating a greater conventional re-dispatch volume in the next stage. The Zonal FM may also activate flexibility resources that must be countered by re-dispatch when all grid constraints are revealed.

Lastly, the LFZs fail to capture all congestion within the FA. This is demonstrated in Figure 4, which illustrates the total FA congestion following the DA clearing and the extent of congestion captured by the LFZs. The remaining congestion is entirely disregarded in the zonal clearing and must be addressed through conventional re-dispatch.

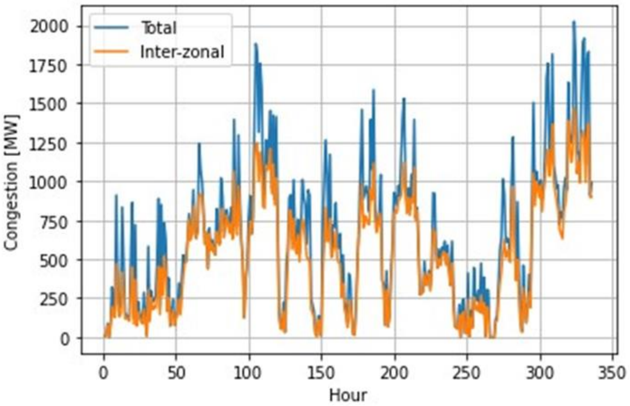

Fig. 4 Total FA congestion and congestion on lines connecting LFZs.

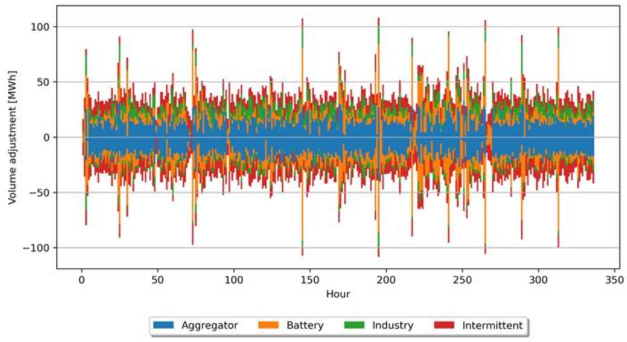

Fig. 5 Volume adjustments by flexibility type for the zonal FM case.

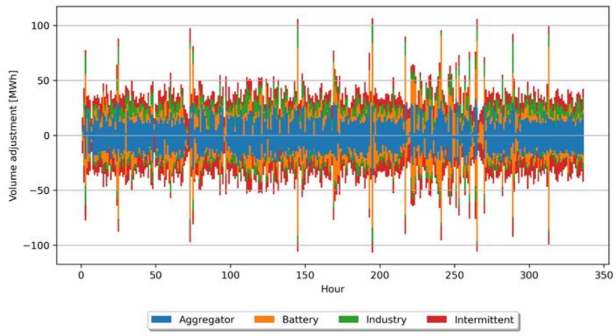

Fig. 6 Volume adjustments by flexibility type for the Nodal FM case.

Figures 5 and 6 provide an overview of the hourly volumes of activated flexibility, categorised by flexibility type, for both the zonal FM and nodal FM cases. The graphs demonstrate significant deviations across individual hours, which may be attributed to noise introduced for bid availability, cost, and volume. Additionally, certain isolated hours show significant surges in volume when compared to the average volume. These spikes are caused by batteries satisfying their initial and final states of charge each day, and thus represent a weakness in the model that nevertheless is similar for both cases. Overall, the dispatch of flexibility volumes in the two cases exhibit notable similarities in total amount and market participant shares. Despite the mentioned similarities, Figure 7 shows that the exact bids activated differ between the two cases. More precisely, it shows that only about 2/3 of the activated flexibility is activated in both cases. In the zonal FM model, the zonal PTDFs means that the cheapest bids within the zones are accepted first. In contrast, the clearing in the nodal FM may prefer more expensive bids in nodes where high PTDFs reduces the required volume adjustments. Since nodal differences lead to diverging clearing preferences between the nodal FM and zonal FM cases, it suggests that using ZPTDFs results in an efficiency loss. This is also the case when all congestion is captured in inter-zonal lines.

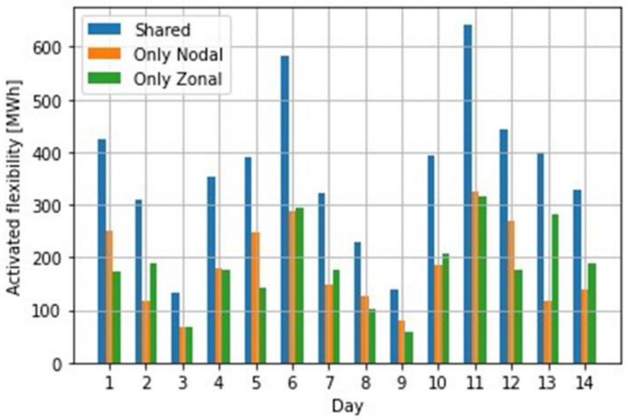

Fig. 7 Activated flexibility that is shared or unique to either the Zonal FM or Nodal FM case, for each day.

3.3 Discussion and sensitivities

To account for uncertainties in the input data, this section highlights sensitivities on several parameters, and discusses weaknesses in the modelling framework.

The determination of the optimal number of LFZs is a critical consideration in the implementation of the zonal FM. As outlined in section 2.2, the heuristic for identifying LFZs relies on a specified number of zones, and thus cannot effectively ascertain the optimal number of LFZs. The primary rationale for adopting a zonal FM is that aggregating nodes into zones may have various benefits. Most of these benefits are dependent on the number of participants and the amount of flexibility volumes, so less aggregation may be needed when these numbers rise. However, the optimal LFZ configuration is also heavily influenced by congestion levels, congestion distribution and distribution of flexibility resources, complicating the optimal LFZ size further.

A sensitivity was run to test the impact of LFZ sizes, although the results are apparent because the models focus on economic efficiency. Specifying three LFZs instead of five resulted in a cost increase of 86,421 € over the first five days of the time horizon, for instance. This increase was primarily due to the zonal FM clearing neglecting a considerable amount of congestion, which leads to an unfavourable dispatch of flexibility. Similarly, higher numbers of smaller zones would mean that the zonal method approaches the nodal method, increasing economic efficiency. A method to determine the zonal configuration should consider the impact on economic efficiency and weigh it against the benefits of aggregating to certain zone sizes, neither of which are considered by this paper's partitioning algorithm.

As FMs are only an emerging concept and yet to be implemented in large scale, the volumes that will be available in such markets are challenging to estimate. In Table 3, the FA cost is presented for all three cases with a scaling of available flexibility volume by 0.5, 2, 4, 10, and 15.

As anticipated, the cost of both the nodal FM and the zonal FM cases decrease as more flexibility becomes available. The absolute cost difference between the zonal FM and the nodal FM cases remains relatively stable up to a ten-fold increase in flexibility volume. However, for the scaling of 15, the absolute cost differences increase significantly due to a rise in internal congestion from the zonal FM, which has to be addressed through re-dispatch after the zonal FM clearing.

Internal congestion initially occurs when flexibility volumes are scaled by a factor greater 10. It has already been argued that higher volumes allow for smaller zone sizes, which would bring more grid information and limit the internal congestion problem. Nevertheless, different congestion patterns or capacities on internal grid lines could arguably make the zonal FM clearing create internal congestion with much lower flexibility volumes as well. Further assessment is required to determine the prevalence of such conditions and to evaluate the effectiveness of improved grid partitioning algorithms in mitigating the problem.

Table 3 Flexibility area cost in M€ for different scaling of available flexibility volume.

Scaling	Nodal FM	Zonal FM	BAU
0.5	2.444	2.439	2.560
2	2.195	2.313	2.560
4	2.083	2.200	2.560
10	1.683	1.805	2.560
15	1.426	1.639	2.560

Similar to the previous discussion regarding available flexibility volumes, the assignment of costs to both flexibility and re-dispatch resources is subject to a high degree of uncertainty. Modifying one or both of these costs would significantly influence the system costs associated with the three cases, presented in section 3.1, but would not notably affect the relative performance of the zonal FM when compared to both the nodal FM case and the BAU case. Decreasing cost of re-dispatch would narrow the differences between the Nodal FM case and the BAU case, for instance, but would also bring the zonal FM case closer to the nodal FM case in terms of performance.

4 Conclusion

This paper investigates the economic efficiency of flexibility markets with dynamic zones. Our findings suggest that the presented zonal FM performs close to the nodal FM in terms of cost. The cost differences can be predominantly attributed to the lack of nodal information in the zonal FM case, resulting in less efficient use of flexibility and more need for re-dispatch after the market clearing. Sensitivity analyses reveal that the performance is contingent on the amount of congestion, available flexibility, and the zonal configuration. A refined zonal partitioning algorithm that accounts for these factors may further stabilise and improve performance compared to the presented results. This paper highlights that a zonal FM can reach an economic efficiency close to the theoretical optimum. By decreasing the cost of congestion management, the integration of more distributed energy resources becomes more viable from a system's perspective.

An improved zonal partitioning method can serve to investigate the robustness of zonal FMs and economic performance to different system states. Future research should therefore focus on developing a more sophisticated zonal partitioning algorithm for FMs for determining the optimal number of zones. Further investigation should examine the comprehensive value of implementing a zonal FM design and weigh the trade-offs associated with economic efficiency losses.

5 Acknowledgements

This paper was prepared as a part of the PowerDig project (Digitalization of short-term resource allocation in power markets), funded by the Research Council of Norway through the ENERGIX program (p-nr: 320789). We are also grateful to industrial partners Statnett and Statkraft supporting the PowerDig project. This paper was developed from a master's thesis [18].

6 References

[1] J. Villar, R. Bessa und M. Matos, „Flexibility products and markets: Literature review," Electric Power Systems Research, Bd. 154, p. 329–340, 2018.

[2] P. Arboleya, M. Kippke und S. Kerscher, „Flexibility management in the low-voltage distribution grid as a tool in the process of decarbonization through electrification," Energy Reports, pp. 248-256, 2022.

[3] A. Ramos, C. D. Jonghe, V. Gómez und R. Belmans, „Realizing the smart grid's potential: Defining local markets for flexibility," Utilities Policy, Bd. 40, pp. 26-35, 2016.

[4] K. Spiliotis, A. I. R. Gutierrez und R. Belmans, „Demand flexibility versus physical network expansions in distribution grids," Applied Energy, Bd. 182, pp. 613-624, 2016.

[5] S. S. Torbaghan, N. Blaauwbroek, P. Nguyen und M. Gibescu, „Local market framework for exploiting flexibility from the end users," 2016 13th International Conference on the European Energy Market (EEM), 2016.

[6] C. Zhang, Y. Ding, J. Østergaard, H. W. Bindner, N. C. Nordentoft, L. H. Hansen, P. Brath und P. D. Cajar, „A flex-market design for flexibility services through DERs," in IEEE PES ISGT Europe 2013, 2013.

[7] C. Zhang, Y. Ding, N. C. Nordentoft, P. Pinson und J. Østergaard, „Flech: A Danish market solution for DSO congestion management through der flexibility services," Journal of Modern Power Systems and Clean Energy, Bd. 2, p. 126–133, 2014.

[8] P. Olivella-Rosell, P. Lloret-Gallego, Í. Munné-Collado, R. Villafafila-Robles, A. Sumper, S. Ø. Ottessen, J. Rajasekharan und B. A. Bremdal, „Local Flexibility Market Design for Aggregators Providing Multiple Flexibility Services at Distribution Network Level," Energies, Bd. 11, 2018.

[9] T. Jiang, C. Wu, R. Zhang, X. Li, H. Chen und G. Li, „Flexibility Clearing in Joint Energy and Flexibility Markets Considering TSO-DSO Coordination," IEEE Transactions on Smart Grid, pp. 1-1, 2022.

[10] S. S. Torbaghan, N. Blaauwbroek, D. Kuiken, M. Gibescu, M. Hajighasemi, P. Nguyen, G. J. M. Smit, M. Roggenkamp und J. Hurink, „A market-based framework for demand side flexibility scheduling and dispatching," Sustainable Energy, Grids and Networks, Bd. 14, pp. 47-61, 2018.

[11] O. Goldschmidt und D. S. Hochbaum, „A Polynomial Algorithm for the k-Cut Problem for Fixed k," Mathematics of Operations Research, Bd. 19, p. 24–37, 1994.

[12] Y. Hu, P. Xun, W. Kang, P. Zhu, Y. Xiong und W. Shi, „Power System Zone Partitioning Based on Transmission Congestion Identification Using an

Improved Spectral Clustering Algorithm,“ Electronics, Bd. 10, 2021.

[13] H. Farahmand, S. Jaehnert, T. Aigner und D. Huertas-Hernando, Task 16.3 Nordic Hydro Power Generation flexibility and transmission capacity expansion to support the integration of Northern European wind power production: 2020 and 2030 case studies, TWENTIES, 2013.

[14] Nord Pool, Day-ahead volumes, Nord Pool.

[15] Nord Pool, Regulating prices, Nord Pool.

[16] W. E. Hart, J.-P. Watson und D. L. Woodruff, „Pyomo: modeling and solving mathematical programs in Python,“ Mathematical Programming Computation, Bd. 3, p. 219–260, 2011.

[17] Gurobi Optimization, LLC, Gurobi Optimizer Reference Manual, 2023.

[18] V. S. Aasvær und A. S. Ryssdal, „ntnuopen,“ 2023. [Online]. Available: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/3122235.

[19] S. H. Low, „Convex Relaxation of Optimal Power Flow—Part I: Formulations and Equivalence,“ IEEE Transactions on Control of Network Systems, Bd. 1, pp. 15-27, 2014.

[20] D. Schönheit, R. Weinhold und C. Dierstein, „The impact of different strategies for generation shift keys (GSKs) on the flow-based market coupling domain: A model-based analysis of Central Western Europe,“ Applied Energy, Bd. 258, p. 114067, 2020.

[21] NVE, NVE.

[22] SINTEF Energi, Planbok, SINTEF Energi.

[23] J. E. Trohjell und I. H. Vognild, 1994.

Appendix A: Zonal partitioning heuristic

To make the heuristic easier to understand, we divided it into two parts: an outer and an inner algorithm, called layers. The outer layer is shown in pseudo code below (Algorithm 1) and can be thought of as performing the more high-level loops of the heuristic.

Algorithm 1 Outer layer: Optimal partitioning of a network

```
repeat
    Index of best zone to split, i,:= 0
    Value of best split, v,:= 0
    The zones created by best split, s,:= {}
    for all zones z_j in Z do
        if z has congestion and |N_z| ≥ k · 2 then        ▷ k is min. number of nodes per zone
            Split zone z_j in two parts, z_j1 and z_j2, and calculate value of cut v_j
            if v_j > v then
                i = j
                v = v_j
                s = [z_j1,z_j2]
            end if
        end if
    end for
    Remove z_i from Z
    Add s[0] and s[1] to Z
until |Z| = n                                            ▷ n is number of zones to be partitioned
```

The outer layer start with a set Z containing one zone, which represents the entire area to be partitioned. As can be seen in Algorithm 1, the heuristic tries to split each congested zone with a sufficient number of nodes into two smaller zones. Only the split option with the highest cut value is chosen for each iteration, and the heuristic ends when the set of zones contains the desired number of partitions. The way of splitting a zone is described by the inner layer shown in Algorithm 2.

Algorithm 2 Inner layer: optimal splitting of a zone

```
Create list L, containing all congested lines in z

Value of best split, v,= 0
The zones created by best split, s,= {}
for all lines l_j in L do
    The cut lines for iteration j, C,:= []
    Cut iteration, i,:= 1
    loop
        Cut line in the middle of the shortest path between l_j[0] and l_j[1]
        Add the cut line c_i to set C
        if l_j[0] and l_j[1] are disconnected and zones are feasible then
            Break
        else if l_j[0] and l_j[1] are disconnected and resulting zones are infeasible then
            loop
                Remove c_i from C and reinstate it in graph
                Perform new cut c_i on the shortest path, this time on the line furthest from the infeasible zone
                if l_j[0] and l_j[1] are connected then
```

Break
else if $l_j[0]$ and $l_j[1]$ are disconnected and resulting zones are feasible **then**
Break x2
else if $l_j[0]$ and $l_j[1]$ are disconnected and resulting zones are infeasible **then**
Remove c_i from C and reinstate it in graph
$i- = 1$
end if
end loop
end if
end loop
if Value of cuts in C are $\geq v$ **then**
v = value of cuts in C
The two zones created are put into s
end if
end for

The inner layer contains slightly more complicated logic than the outer layer. The main idea is that the algorithm cuts a line in the middle of the shortest path between two nodes, $l[0]$ and $l[1]$. This is repeated until the two nodes are split into two disconnected areas. The complexity arises from protective measures trying to adapt infeasible solutions without throwing away the solution. A zone is infeasible when it has too few nodes. In that case, the algorithm will redo its last cut and instead choose a line on the shortest path closer to the feasible zone. If the solution is still infeasible, the algorithm will throw away the current cut c_i completely and try to redo cut c_{i-1}. If the algorithm returns to step $i = 1$ and still gets an infeasible solution, line j is discarded completely. The inner layer may in theory return no feasible solutions to the outer layer, and if this happens for all zones z_j, the heuristic will stop with $< n$ zones.

Appendix B: Modelling assumptions and complete model formulations

This appendix includes an overview of the most central assumptions made with regards to the market modelling of the three cases, as well as the nomenclature formulations of the models.

Perfect competition: In a market with perfect competition, there is no market power, and the market participants will bid their true marginal cost (this assumption is not valid for hydro power owners who will bid their water values, however, in this paper, hydro power units are included in the conventional re-dispatch category and not as a part of the flexibility units). This assumption simplifies the modelling process and is common for centralised optimisation models. However, it is debatable if this assumption is valid for FMs. The geographical scope of FMs could limit the number of potential market participants, which may be a problem in terms of market power. For the smaller lines in the DS, some actors may also get pricing power because they alone can cause or relieve congestion in the grid. Nevertheless, accounting for market power is not in this paper's scope, nor would it bring any valuable insight without significantly better data on the cost of flexibility and the behaviour of FM participants.

Completely inelastic demand: Since the system operators are responsible for the grid operation, they must make sure all line constraints are satisfied. We therefore assume the total amount of flexibility and conventional re-dispatch demanded by system operators is fixed and determined by adjustments needed to satisfy line constraints. Therefore, the problem of maximising social welfare reduces to the problem of minimising costs.

No changes to cross-border exchanges: The Nordic power system has interconnectors to other countries, such as Germany and the UK. Including the flow on these interconnectors complicates the data pre-processing, and changes to these flows are unlikely to alter the results of the FM performance. We therefore assume that the FMs and re-dispatch do not affect the transmission across interconnectors. Nevertheless, interconnectors were taken into account when determining the DA volumes.

DC power flow: We assume DC power flow. DC flow relies on the assumption that resistance is significantly smaller than reactance in the grid. This is a normal and valid assumption for the TS, but it is not always the case for medium voltage DS. For our DS specifically, these values are not satisfying the DC power flow assumptions, and AC power flow modelling would be needed to accurately model flows in the grid. However, including AC flow on these lines would require an extension to non-convex programming. Although there exist convex relaxation techniques to reduce this problem to a convex one [18], the extension would likely increase the complexity and run time of the models significantly. Additionally, we will argue that the load flow modelling accuracy does not reduce the validity of this paper's results, as we aim to provide an initial investigation into zonal FMs and all cases are using the same load flow equations. Therefore, improving load flow by using AC power flow is beyond the scope of this project and we assume DC power flow for the whole system.

Each day is independent: The models solve for the 24 hours cleared in the DA market and consider interdependencies in this time interval, but do not take into account the state of the system from the previous day. The 24-hour scope corresponds with how the FM would be cleared in practice, as its purpose is to adjust volumes after DA. However, the FM participants' behaviour would be dependent on their trading in previous days, and in some cases also their expectations for the future. These interdependencies are disregarded for the sake of simplicity and computational time.

System operators are not modelled explicitly: System operators are important FM participants in this paper, especially as they are responsible for procuring the flexibility that will alleviate congestion. When doing so, they must counter-trade in another zone or node to keep the market balance. This step can be skipped when modelling by only describing the volumes adjusted up and down. Assuming the system operators do not have any transaction costs connected to participating in the market, the models will then account for all the costs to society.

Congestion outside of the FA is also considered: We model the whole TS described in section 2.4.2, including the net positions after DA and the resulting congestions. All lines in the system are thus subject to capacity constraints. The reason is that the flexibility area is not an isolated system but is interdependent with the surrounding grid. Thus, solving solely for the flexibility area will not provide an accurate description of how the FM will affect the grid in practice.

Nodal Formulation
Sets and indices
N : Set of nodes, $n \in \mathrm{N}$
B^+ : Set of UP bids, $b \in \mathrm{B}^+$
B^- : Set of DOWN bids, $b \in \mathrm{B}^-$
B^{RE} : Set of bids representing redispatch
$\mathrm{L} \subset N \times N$: Set of lines, $l \in \mathrm{L}$
I^{Agg} : Set of aggregator market participants, $i \in \mathrm{I}^{Agg}$
I^{Batt} : Set of battery market participants, $i \in \mathrm{I}^{Batt}$
T : Set of time periods, $t \in \mathrm{T}$
Parameters
P_{nt}^{DA} : Net production in node n and period t from DA clearing
V_{tb} : Available volume of bid b in period t
C_{tb} : Cost of bid b in period t
$PTDF_{ln}$: PTDF for line l, node n
CAP_l : Capacity of line l
S_i^{init} : Initial battery storage for participant i, $i \in \mathrm{I}^{\mathrm{B}}$ S_i^{Cap} : Battery storage capacity for participant i, $i \in \mathrm{I}^{\mathrm{B}}$ η^+: Efficiency of battery discharging η^-: Efficiency of battery charging
Variables x^+_{tb} : Volume cleared up of bid b in period t x^-_{tb} : Volume cleared down of bid b in period t f_{lt} : Flow on line l in time period t p_{nt} : Net production in node n in period t σ_{it} : State of charge of participant i's battery in period t, $i \in \mathrm{I}^{\mathrm{B}}$

Zonal Formulation
The notation that differs or is added after the nodal model is presented below, while the zonal model is formulated in its entirety.

Sets and indices
Z_t : Set of zones in period t, $z \in Z_t$
L_t : Set of lines between zones in time period t, $l \in L_t$
Parameters
P_{zt}^{DA} : Net production in zone *z*, period *t* from DA clearing $ZPTDF_{lzt}$: ZPTDF for line *l* and zone *z* in time period *t* γ: Unit penalty of congestion
Variables p_{zt} : Net production in zone *z* in period *t* y_{lt} : Congestion on line $l \in L_t$ in time period *t*.

BAU
The BAU model can be seen as the nodal model without FM-specific constraints, as it only includes redispatch bids and does not take into account any constraints related to FM participants.

Appendix C: Zonal Power Transfer Distribution Factor

This appendix describes how PTDFs are transformed into ZPTDFs. The transformation requires generation shift keys (GSKs), which describe how a change in net position in a zone is distributed among the generating units in the zone [19]. As determining the GSKs exactly requires information about production and load, which is unavailable prior to the market clearing, several approaches to estimate GSKs exist. The choice of which generating units to include and how to assign the weight with which each unit contributes to the change in net position is called a GSKs strategy. [19] lists four GSKs strategies, where one of them uses the scheduled power output to determine the GSKs. Since the adjustments in the FM are made relative to DA volumes, we use the DA volumes to approximate the market output from the FM clearing and construct GSKs based on them.

Let P_{nt}^{DA} be the net DA position at node *n* in time period *t*, N the set of nodes, N_z the set of nodes in zone *z*, T the set of time periods and Z_t the set of zones in time period *t*. The GSK for node *n*, zone *z* in time period *t* is then given by equation C1.

$$GSK_{nzt} = \frac{P_{nt}^{DA}}{\sum_{n \in \mathcal{N}_z} P_{nt}^{DA}} \qquad n \in \mathcal{N}, t \in \mathcal{T}, z \in \mathcal{Z}_t \tag{C1}$$

Furthermore, let $PTDF_{ln}$ be the PTDF for line *l* and node *n* and L be the set of lines. The ZPTDF for line *l*, zone *z* in time period *t* is then given by:

$$ZPTDF_{lzt} = \sum_{n \in \mathcal{N}_z} GSK_{nzt} \cdot PTDF_{ln} \qquad l \in \mathcal{L}, t \in \mathcal{T}, z \in \mathcal{Z}_t \tag{C2}$$

Appendix D: Data

DS design: The FA is a subpart of the NO1 ELSPOT zone and it is the only part of the grid where a DS needs to be constructed. Since the Norwegian DS has a radial structure, we base this construction on three example radials designed after NVE Atlas data [20], seen in figure D1.

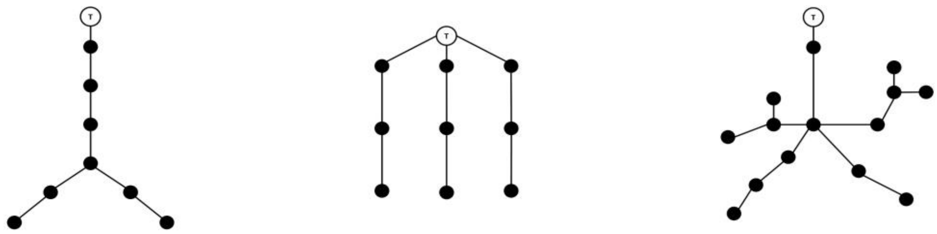
Figure D1: Example radials used to construct the DS grid.

For each TS node with a load, one of these radials are attached, scaled, and rotated, and the resulting grid can be seen in Figure D2. In the figure, red nodes represent TS nodes and green nodes represent DS nodes. Each radial grid is assigned a voltage, either 50 kV or 132 kV, based on the load in their respective TS node. Additionally, the example radials all have load distribution factors for their nodes, which are used to redistribute the load previously being allocated to the TS node.

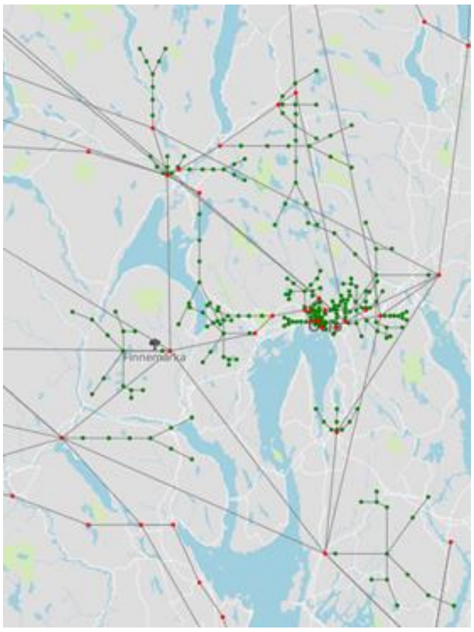

Figure D2: Flexibility area with red TS nodes from the reference dataset and artificially constructed green DS nodes.

The line properties are determined next. Based on data for a 72.5 kV overhead line from [21], the resistance and reactance values are set to 0.122 and 0.379 Ohm/km respectively. The line capacities are based on [22], who state that distribution lines in Norway with a voltage level of 66 kV typically have capacities between 50 MW and 125 MW, while lines with a voltage level of 132 kV typically have capacities in the range of 100 MW to 250 MW. To distribute capacities in the DS, we perform a load flow analysis on the DA volumes from 12.00-13.00 on 17th of November. We divide the set of DS lines into two sets: one with a voltage level of 50 kV and one with a voltage level of 132 kV. For each set of lines, we then partition the lines into ten quantiles based on their load flow. Additionally, the two capacity ranges are divided into 10 intervals each. The quantile to which a line belongs determines its capacity, where each line in a quantile is assigned the upper limit of the corresponding capacity interval. For instance, a 132 kV line that falls in the quantile with the highest flow will have a capacity of 250 MW, while a 50 kV line in the same quantile will have a capacity of 125 MW. This approach aims to evenly distribute the possible capacities while considering that some lines in the radial network may consistently experience higher load flows.

Market participants: The participants were allocated to nodes randomly based on the following probability:

$$p_{nt} = \begin{cases} 0 & L_n < d_t \\ b_t + a_t \cdot L_n & L_n \leq c_t \\ b_t + a_t \cdot c_t & L_n > c_t \end{cases} \qquad \text{D1}$$

Where p_{nt} is the probability of adding a participant of type t to node n, and L_n is the expected load [MW] in the node, using the reference dataset. a_t, b_t, c_t and d_t are parameters assigned to the various participant types to create an appropriate relationship between the probability and the expected load. Their values are found in Table D1 below along with the resulting number of participants.

Our use of Equation D1 and the weights in Table D1 to allocate participants is intended to increase the transparency of our assumptions and approach. The d_t values show that aggregators and industry are assumed to be present only in nodes with a certain amount of load. Similarly, c_t is used to define a threshold beyond which an increase in expected node load does not increase the probability of allocating a participant of type t. Lastly, a_t and b_t define the slope and intercept of the probability function, respectively.

Table D1 Participant types, their stochastic parameters and total number of market participants.

Participant type	a	b	c	d	Max probability	Count total
Aggregator	2.5 %	5 %	30 MW	1 MW	80 %	41
Battery	0.3 %	5 %	50 MW	0 MW	20 %	19
Industry	2 %	0 %	40 MW	3 MW	80 %	34
Intermittent	2.5 %	5 %	10 MW	0 MW	30 %	41

Market bids: There is a lack of reliable data on flexibility market participants, especially the costs of providing flexibility, and the existing sources have inconsistent assumptions compared to this project. The bid costs are therefore determined based on assumptions made specifically for this project. Similar to the generation of participants, the method used for generating bids is intended to be transparent. Instead of probabilities, however, it uses the participant type, size, and cost scaling to construct bids for each participant. Given a set of participants, we deterministically construct a set with bids assigned to the respective participant including a cost, volume, direction, and hour. Table D2 below presents the input used for this method, where the cost is scaled with the "cost scaling" attribute for each participant, and the bid size is given as a percentage of the participant's size.

Table D2 Rules to determine the bids of a FM participant.

Participant type	Number of bids		Bid cost [€/MWh]		Bid size [MWh]	
	UP	DOWN	UP	DOWN	UP	DOWN
Aggregator	3	2	2,8,15	5,15	5%,10%,25%	10%,25%
Battery	2	2	2,5	2,5	20%,80%	20%,80%
Industry	1	1	10	10	30%	10%
Intermittent	1	1	-10	-10	15%	15%

The Table D2 columns with the number of bids, indicate that aggregators and battery actors can place multiple UP or DOWN bids in a specific hour. This is reflected in the corresponding columns for bid cost and bid size, where each entry refers to a separate bid. Multiple bids are used to show the actors' increasing cost of providing flexibility when the volumes increase. For example, the aggregators will demand more money when cutting load beyond 5% of their size, as the cost then jumps from 2 €/MWh to 8 €/MWh, and batteries will offer a smaller percentage of their capacity at a lower price. The table reveals several other assumptions; for example, how it is easier for industry to decrease the power consumption than to increase it. Finally, the table indicate that intermittent power producers are not regarded as providers of flexibility, but rather as BRPs that need to procure flexibility due to incorrect forecasting in the DA market.

Further elaboration is needed regarding the costs presented in Table D2. It is first important to reiterate that the costs are not based on any external sources, as they were judged to be too uncertain to bring any value. However, there are also some assumptions behind these numbers that are more fitting to present here, and that relate to how costs are used in the FM models. When determining what an actor will charge for flexibility, we assume that the cost is closely related to the DA market. If an aggregator has expenses of 2 €/MWh power provided to the grid, we assume that it charges a DA price premium of 2 €, as we have perfect competition. Thus, the cost and the price charged in the FM are not the same, but are linked by the DA price, and it is the cost that is used when determining the bids. For modelling purposes, it is sufficient to only consider the costs as they appear to the participant and disregard the actual market price.