Mapping Flexibility

Data Pipelines, Innovation Ecosystems, and Strategic Decision-making in Distribution Networks

Master's thesis in Computer science and engineering

William Ackemo

Axel Löfqvist

Department of Computer Science and Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2026

Master's thesis 2026

Mapping Flexibility

Data Pipelines, Innovation Ecosystems, and Strategic Decision-making in Distribution Networks

William Ackemo
Axel Löfqvist

UNIVERSITY OF GOTHENBURG

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Computer Science and Engineering
Chalmers University of Technology
University of Gothenburg
Gothenburg, Sweden 2026

Mapping Flexibility
Data Pipelines, Innovation Ecosystems, and Strategic Decision-making in Distribution Networks
William Ackemo
Axel Löfqvist

Supervisor: Vinh Ngo, Computer Science and Engineering
Supervisor: Ksenia Onufrey, Technology Management and Economics
Advisor: Joris van Rooij, Göteborg Energi
Examiner: Marina Papatriantafilou, Computer Science and Engineering

Master's Thesis 2026
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover:

Typeset in LaTeX
Gothenburg, Sweden 2026

Mapping Flexibility
Data Pipelines, Innovation Ecosystems, and Strategic Decision-making in Distribution Networks
William Ackemo
Axel Löfqvist
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg

Abstract

Electrification is increasing the pressure on electricity distribution grids, causing congestion and capacity problems, while grid infrastructure is characterized by long planning and construction lead times. Consequently, Distribution System Operators (DSOs) must make better use of the existing grid and its existing resources through *flexibility*, defined as ability to shift loads away from peaks and reduce congestion risks through different technologies. Thus, a DSO must increasingly rely on *flexibility market* mechanisms to manage capacity. A central challenge is to identify where, when, and how flexibility in the grid, and how it can be captured and deployed to solve grid related issues. One way to address this challenge is to analyze the large amounts of smart meter data generated by electricity consumption in the city, while also better understanding how the organization can utilize these resources.

This thesis examines how data-driven analysis and strategic decision-making jointly shape the deployment of flexibility through an *interdisciplinary* approach. The study combines advanced quantitative data analysis and clustering techniques with a qualitative organizational and *innovation ecosystem* study. It examines how interdependent actors jointly enable flexibility in the grid, while creating a data pipeline and identifying relevant use cases and trade-offs based on interviews with key stakeholders within the DSO organization and its surrounding external actor network. The quantitative analysis develops consumption baselines and applies clustering to identify recurring demand patterns among households and small to medium-sized enterprises. The qualitative analysis examines how these analytical outputs can be interpreted, acted upon, and embedded in organizational decision-making and flexibility market development.

The findings show that smart meter data can be used to show where, when, and how much flexibility exists. The study also finds that, within the regional flexibility market context, geographical clustering provides limited additional value, allowing analytical efforts to focus instead on clustering customers based on similarities in consumption behavior. Furthermore, the results highlight that clustering outputs must remain operationally actionable, where fewer, clearer, and more explainable customer groups are more useful for decision-making than highly granular segmentations.

These findings imply that flexibility deployment is both a technical analytics challenge, and an organizational and ecosystem challenge. By linking advanced data analytics and clustering methods to the practical needs of DSOs, *aggregators*, and other ecosystem actors, the study shows how smart meter data can support more resource-efficient, lower-risk, and strategically informed flexibility deployment in modern electricity distribution systems.

Keywords: smart meter data, flexibility markets, electricity distribution systems, DSOs, data-driven decision-making, energy system planning, qualitative interviews, organizational analysis, innovation ecosystem analysis, innovation ecosystems, dynamic capabilities, technology adoption, data pipeline design, consumption modeling, ARIMA, SARIMAX, clustering methods, Wards linkage, K-means, DBSCAN, load profiling, peak demand, demand-side flexibility, geographic clustering, computational efficiency, resource allocation, interdisciplinary.

Acknowledgements

We would like to express our sincere gratitude to our examiner Marina Papatriantafilou and supervisors Quang Vinh Ngo, Ksenia Onufrey for their guidance, constructive feedback, and continuous support throughout this thesis. Their expertise and insights have been integral in shaping the direction and quality of this work.

We extend our special thanks to our industry advisors at Göteborg Energi AB, Joris van Rooij, for providing us with this opportunity and for sharing their practical expertise and perspectives from the energy sector and Joris for his support and insightful comments. We also thank Göteborg Energi AB for granting access to the data, infrastructure, and domain knowledge that made this research possible.

Furthermore, we would like to thank the interviewees who contributed valuable organizational and ecosystem insights through talks, and discussions during the course of this thesis.

The grammar, spelling, and sentence structure of this document were reviewed using Overleaf's built in language and writing assistance tools. Artificial intelligence tools were additionally used to assist with the conversion of tables and figures in latex, code structuring and refinement, and improving the readability and clarity of selected text sections. All technical content, analysis, results, interpretations, and conclusions remain the responsibility of the authors.

William Ackemo, Gothenburg, 2026-07-06Axel Löfqvist, Gothenburg, 2026-07-06

Contents

List of Figures

List of Tables

1
Introduction

Sweden is experiencing increasing electricity demand driven by electrification of transport, industrial transformation, and broader decarbonization objectives [1]. At the same time, expansion of grid capacity and generation assets is characterized by long lead times and regulatory complexity [2]. As a result, Distribution System Operators (DSOs) face growing pressure to optimize utilization of existing infrastructure rather than relying solely on physical grid reinforcement.

One increasingly important response to these challenges is *flexibility*, the ability to shift or reduce electricity consumption in response to grid needs [3]. Flexibility is central to maintaining system reliability and managing local capacity constraints, but its value depends on more than its technical availability. It must also be understood in relation to the actors involved, like DSOs, flexibility providers such as large scale battery companies and industries, their incentives, organizational capabilities, and market arrangements that determine whether flexibility can actually be deployed [4].

Flexibility must also be inferred through analysis of consumption patterns, peak load dynamics, spatial variation, and external conditions such as weather and electricity prices [5]. The deployment of Advanced Metering Infrastructure (AMI) has thus created new possibilities for such analysis by generating high-resolution electricity consumption data across the distribution grid [6].

This thesis studies flexibility as both a data-analytical and a managerial challenge by combining smart-meter data analysis with a qualitative study of the organizations *innovation ecosystem* through interviews and managerial concepts. Innovation ecosystems describe how a system of actors align towards a common goal to achieve outcomes that one single actor could not achieve by itself [7].

1.1 Context of Study

In Sweden's Updated National Energy and Climate Plan, the ability to adjust electricity consumption by shifting or reducing load on the grid, flexibility, is highlighted as a necessary capability for maintaining system reliability in the coming decades [3]. This development is not unique to a single city or company, but reflects a structural transformation affecting distribution grids nationally and across Europe [8]. This thesis examines these dynamics in the empirical context of Gothenburg, Sweden, where concentrated load growth, adoption of electric vehicles, and industrial electrification make the challenges particularly visible [9].

Traditionally, grid planning has relied on assumptions of peak demand growth and long-term load forecasts [10]. However, increasing temporal variability in consumption patterns challenges static planning approaches [5]. Electricity demand now reflects not only structural consumption needs, but also behavioral factors such as electric vehicle charging schedules, distributed generation, price responsiveness, and weather-dependent heating loads. These dynamics directly influence when and where peak loads occur.

In this thesis flexibility is studied through the difference of maximum observed load and actual consumption at a given point in time. It reflects the extent to which demand could, in principle, be shifted without reducing total energy use [5]. At any given moment, the available flexibility corresponds to the gap between current consumption and technical or behavioral capacity limits [11]. Identifying this gap requires understanding both temporal variation (day, week, month, season) and spatial distribution across the grid.

1.2 Motivation

The increasing complexity of electricity consumption creates an increased need for more informed data approaches to planning of the distribution grid. As electrification accelerates, DSOs must determine whether capacity challenges should be addressed through infrastructure reinforcement, flexibility mechanisms, tariff design, or a combination of these tools. Thus, the DSO that organizes this system, needs both the capability to identify and operationalize data-streams and an active involvement with dependent actors to draw the most out of flexibility.

DSOs broadly have three categories of tools available when addressing capacity constraints [2], [4]. The first is physical infrastructure reinforcement, which provides long-term capacity expansion but involves substantial capital expenditure and long planning horizons [10]. The second is the activation of flexibility mechanisms, including structured flexibility procurement where consumers are compensated for temporary load reductions [4]. The third is tariff design and price-based incentives, which may influence consumption behavior over longer time scales if regulatory conditions permit.

The relevance of each tool depends on when, where, and how much flexibility exists within the grid [5]. If peak loads are structurally concentrated and largely insensitive to external signals, infrastructure reinforcement may be justified. If load peaks are temporally sharp and potentially shiftable, flexibility mechanisms or tariff-based incentives may provide more efficient solutions. The purpose of this thesis is therefore not to prescribe a single intervention, but to develop an evidence based structure that connects quantitative flexibility analysis to the strategic decision space of the DSO.

1.3 Interdisciplinary Opportunities

This thesis studies flexibility as both a data analytical and organizational challenge. It combines a quantitative analysis grounded in the Computer Science and Networks discipline with a qualitative analysis grounded in Management and Economics of Innovation. While the two parts are designed to be independently robust, their main value lies in how they inform each other when studying flexibility in the context of a DSO.

The interdisciplinary approach creates several opportunities. First, it allows the quantitative analysis to be guided by organizationally relevant questions, such as grid planning, flexibility procurement, and collaboration with external actors. Second, it gives the qualitative analysis an empirical foundation by linking strategic discussion to observed consumption patterns. Third, it makes it possible to evaluate flexibility both as a technical potential and as an organizational capability.

The interdisciplinary approach also has some inherent challenges. Quantitative and qualitative findings can emerge at different stages of the research process, making integration difficult. Interview insights may guide early analytical choices before all data patterns are known, while unexpected quantitative findings may appear after interviews have already been conducted. The interdisciplinary value of the thesis, therefore, depends on the creation of meaningful points of connection between the two analytical methods throughout the research process.

1.4 Flexibility as Data-Driven Grid Analysis

Unlike traditional grid planning approaches that often rely on static assumptions about peak demand growth, flexibility analysis requires a more dynamic understanding of electricity consumption. Peak loads are not only determined by the total level of demand, but also by when demand occurs, and whether it is sensitive to external conditions.

The quantitative component of this thesis focuses on constructing a scalable analytical framework capable of processing and structuring large volumes of smart meter data into operationally meaningful representations of electricity consumption behavior. Rather than analyzing individual consumption events in isolation, the objective is to identify broader behavioral patterns related to flexibility and peak formation within the distribution grid.

To achieve this, the study investigates how clustering techniques can be applied to aggregated smart meter behavior in order to create interpretable consumer groupings based on similarities in consumption and surge patterns. Since the available data is unlabeled, the objective is not to maximize predictive accuracy in a supervised learning setting. Instead, emphasis is placed on developing an effective data pipeline and evaluating clustering approaches capable of generating computationally feasible, balanced, and operationally actionable clusters. The analytical framework further integrates smart meter consumption data with contextual information such as geographical data, weather observations, and electricity price signals. However, insights obtained during the qualitative study indicated that geographical proximity between flexibility sources was of lower operational importance than behavioral similarity, since loads can frequently be rerouted within the distribution grid. Consequently, the clustering methodology prioritizes consumption characteristics and peak behavior over strict physical adjacency.

The purpose of the framework is therefore not only analytical description, but also organizational applicability. The resulting pipeline and clustering outputs are evaluated based on their ability to produce interpretable and operationally relevant representations of consumption behavior.

Answering these questions requires integrating high-resolution smart meter data with geographical information, weather observations, and electricity price signals.

1.5 Strategic Relevance for DSOs

This thesis also investigates how data-driven flexibility insights can influence the strategic direction of a DSO. Furthermore, the thesis examines how such insights can support the deployment of internal resources and capabilities in ways that strengthen the DSO's ability to respond to electrification and capacity challenges.

Strategic management literature has increasingly emphasized that firms must manage operating under uncertainty, institutional change, and technological complexity [4], [12]. These conditions are highly relevant to DSOs, which must respond to rapid developments in flexibility resources, regulatory frameworks, and market based solutions. The ability to interpret and act upon flexibility related data can therefore be understood as part of the DSO's broader strategic capability.

However, flexibility is not only an internal organizational issue. Although firm-level capabilities determine how a grid operator can govern and utilize flexibility markets, the success of such markets ultimately depends on the willingness and ability of external actors to participate. These actors may include Battery Energy Storage System (BESS) providers, Photovoltaic (PV) owners, and actors that aggregate flexibility resources [13], [14].

Thus, the *innovation ecosystem* involved in making flexibility options possible is also integral to understanding the potential of data findings for strategic deployment of internal resources [13]. The qualitative analysis therefore provides the strategic context needed to interpret how data-driven flexibility insights could support internal resource deployment and external ecosystem coordination [15].

1.6 Research Questions

This thesis is guided by three research questions that together address flexibility in the distribution grid from both a data driven and strategic perspective.

1. What clusters can be created to capture the consumption patterns of massive amounts of smart meter data to gain insights about the flexibility in the distribution grid?

2. Building on the flexibility insights derived from RQ1, how can a DSO position itself within the emerging flexibility ecosystem, considering the role of external actors, and internal resources and capabilities?

3. How can a data pipeline be constructed to effectively utilize the readily available smart meter- consumption, location data, weather and spot price data based on the organizational needs? And how can this data become useful for an organization considering their operations and broader ecosystem?

1.7 Contributions

This thesis demonstrates how data streams, containing the information of when and where events are happening and the relationship between them, can be translated into decision-relevant insights for a DSO by iteratively aligning quantitative analysis with qualitative understanding of organizational priorities and ecosystem constraints.

Furthermore, the thesis contributes by developing an evidence based analytical structure for understanding flexibility in the distribution grid. By combining high-resolution consumption data with geographical information, temperature variables, and electricity price signals, the thesis identifies when, where, and under what conditions flexibility may exist.

In addition, the thesis contributes by informing on the selection of tools for capacity management discussed in section 1.2. By identifying under what conditions flexibility may exist, and by understanding internal and external actors conditions that shape participation, the thesis helps clarify when different tools become relevant.

Finally, the thesis contributes to the broader discussion on how distribution grids can transition toward more flexibility oriented and data informed planning practices under increasing electrification. Thus, the thesis shows how DSOs can better utilize existing infrastructure while preparing for future demand growth.

1.8 Overview

The thesis is organized as follows. Chapter 2 presents the relevant background on electricity distribution grids, smart meter data, and strategic management perspectives relevant to DSOs. Chapter 3 defines the research problem and outlines the challenges involved with investigating flexibility in the distribution grid. Chapter 4 describes the data sources and methodological approach deployed. Chapter 5 describes the risks and ethical considerations involved in the research. Chapter 6 presents the quantitative and qualitative findings, and the interdisciplinary analysis. Chapter 7 concludes the thesis by summarizing the main findings, contributions, limitations, and directions for future research.

2
Background

This chapter introduces the foundational theory and concepts needed to analyze and interpret the empirical data and findings of this study. The first part of the background establishes the operational and related studies context, beginning with the CLUE framework for time series clustering that outlines relevant limitations and possibilities within a related study. Then it broadens the context to explain the role of a DSO in modern electricity grid markets and the technical requirements of flexibility. Furthermore, the technical foundation of high-dimensional data processing and modeling, and the managerial concepts required for strategic decision-making are established. These two foundations will work together ensuring that the subsequent analysis is grounded in both technical rigor and organizational relevance.

2.1 CLUE

The increasing availability of Advanced Metering Infrastructure (AMI) data has led to the development of methods aimed at extracting structure from high-dimensional electricity consumption time series. A notable contribution in this domain is the CLUE framework (Clustering-Based Load Understanding and Exploration) developed by Linus Magnusson and Rasmus Thorsson at Chalmers University of Technology and the University of Gothenburg [6].

Magnusson and Thorsson address the computational and methodological challenges of clustering large-scale temporal electricity data. Their toolchain integrates multiple clustering algorithms, with particular emphasis on IP.LSH.DBSCAN, enabling significant computational speedups compared to traditional DBSCAN approaches. The framework also incorporates automated parameter optimization, multiple distance metrics (including Euclidean, Angular, and Dynamic Time Warping) [16], and flexible representations using both raw time series and feature-based abstractions. Through a case study with Göteborg Energi, analyzing consumption data from approximately 7,500 customers, CLUE demonstrates the ability to identify meaningful consumption profiles and detect anomalies in high-dimensional load data [6]. The work contributes a scalable and interactive methodology for pattern discovery in electricity consumption time series.

While CLUE focuses primarily on clustering and exploratory load profiling, it highlights several limitations that are directly relevant to this thesis. In particular, Magnusson and Thorsson emphasize the need to evaluate clustering stability across multiple time periods and to develop techniques for tracking clusters over time. Such longitudinal analysis would enable the identification not only of static consumption patterns, but of structural changes and temporal persistence.

This temporal dimension is closely related to the objectives of the present work. Rather than clustering individual load profiles in isolation, this thesis examines how consumption patterns evolve across days, weeks, and seasons, and how such evolution relates to geographical distribution and external drivers such as weather and price signals. The emphasis is therefore less on algorithmic clustering performance and more on interpreting temporal and spatial variability in relation to distribution grid flexibility and planning decisions. In this sense, the present study complements the CLUE framework by extending the analytical focus from pattern identification within a fixed dataset toward decision-oriented interpretation of when, where, and under which conditions load variability and flexibility emerge in the distribution grid.

2.2 Distribution System Operator

A Distribution System Operator (DSO) is the entity responsible for operating, maintaining, and developing the low- and medium-voltage electricity distribution grid within a defined geographical area. The DSO ensures reliable electricity delivery from the transmission system to end users, including households, commercial consumers, and industry. Core responsibilities include maintaining power quality, ensuring supply reliability, managing grid capacity constraints, connecting new consumers and producers, and planning long-term infrastructure investments. Unlike electricity suppliers, DSOs operate the physical network infrastructure that enables electricity transport. In practice, DSOs must balance current operational reliability with long-term capacity planning under increasing electrification. In this thesis, the DSO context defines the decision environment in which smart meter data, weather variables, and price signals are analyzed. The objective is not market trading optimization, but improved understanding of load behavior to support grid planning and congestion management.

2.2.1 Price Signals and Local Load Behavior

Wholesale electricity markets generate price signals that increasingly influence consumption and production behaviour. As electricity prices day-ahead and intraday markets become more volatile, primarily due to weather dependent renewable generation and supply and demand imbalances, consumers and flexibility providers respond by adjusting their usage patterns.

Dynamic pricing contracts, demand response programs, and automated energy management systems enable households and businesses to shift consumption toward lower priced hours. Electric vehicle charging, battery storage operation, and flexible industrial processes could be more responsive to price variations. Similarly, distributed generators adjust output decisions in response to market incentives.

Although DSOs do not directly participate in wholesale electricity trading, these price induced behavioral adjustments affect local load profiles and congestion patterns. High price volatility can lead to synchronized demand shifts, potentially creating new local peaks rather than smoothing them. Conversely, effective price responsiveness can alleviate stress during critical periods.

Understanding how price signals interact with weather conditions and consumer behavior is therefore essential for anticipating grid constraints. Incorporating price variables into analytical models enhances the DSO's ability to interpret emerging load patterns and assess the implications of market driven flexibility for distribution grid stability.

2.2.2 Decision Environment of the DSO

The DSO operates within a complex sociotechnical ecosystem shaped by regulatory frameworks, market design, technological innovation, and evolving consumer behavior. While the DSO's mandate focuses on reliability and infrastructure management rather than trading or generation, its operational reality is increasingly influenced by external market outcomes and decentralized decision-making. This creates a decision environment characterized by heightened uncertainty and interdependence.

2.3 Quantitative Data Concepts

This section describes some of the core concepts which will be used for the querying, modeling and the environment the data is structured and located in.

2.3.1 Querying: Filters, Aggregations, and Joins

Querying refers to the process of retrieving and transforming data from relational tables. In structured data systems, this is typically performed using operations such as *filtering*, *aggregation*, and *joins* [17]. Filtering selects a subset of rows based on specified conditions, such as a time interval, a specific group of meters, or a particular geographical area. In the present study, filtering is applied during data ingestion to retain only meters belonging to the selected region and observations falling within the analysis window.

Aggregation summarizes granular observations into higher-level representations [17], for example computing hourly averages or daily peak values from 15-minute measurements. In this study, 15-minute smart meter measurements are aggregated into hourly consumption profiles for time-series modeling, while meter-level observations are aggregated to daily statistics during surge detection. Aggregation reduces data dimensionality while preserving the temporal patterns required for subsequent analysis.

Join operations combine records from multiple tables based on shared keys [17], such as linking a meter identifier in a measurement table to its geographical coordinates in a separate metadata table. These operations are required to integrate smart meter measurements with geographical metadata, weather observations, and electricity price data. In the implementation, *broadcast joins* are particularly important, where small reference tables, such as location metadata, are replicated across worker nodes to avoid costly network shuffles when joining against large consumption datasets.

2.3.2 Statistical Time Series Models

The objective of the time-series modeling stage is to estimate expected aggregate electricity consumption and thereby separate structural demand patterns from short-term deviations. Rather than being used primarily as forecasting tools, the models serve as baseline estimators against which residual consumption behavior can be evaluated.

Two model specifications are considered. The first is an Autoregressive Integrated Moving Average (ARIMA) model [18], which captures the underlying autoregressive structure of the aggregate consumption series. The second extends this framework through a Seasonal Autoregressive Integrated Moving Average model with Exogenous Variables (SARIMAX), incorporating both seasonal effects and external explanatory variables [19]. Comparing the two specifications enables assessment of the extent to which temperature, electricity prices, and time-of-day effects explain variation in aggregate demand beyond the autoregressive structure alone.

The fitted values from each model define the expected consumption baseline, while the residuals provide a measure of consumption behavior that cannot be explained by the modeled temporal and external drivers. These residuals subsequently form the basis for surge identification and behavioral analysis.

2.3.3 Datasets and Data Streams

A data stream refers to continuously generated data that arrives over time, such as smart meter readings recorded every 15 minutes or hourly weather updates. Streams are unbounded in nature and are can be processed incrementally as new observations arrive. A dataset, in contrast, is a stored and finite collection of data that can be queried repeatedly. The exloration done in this study will however explore a medium in between the two, with the 15 minute data points being batched into 4 hour segments before reaching the accessed database.

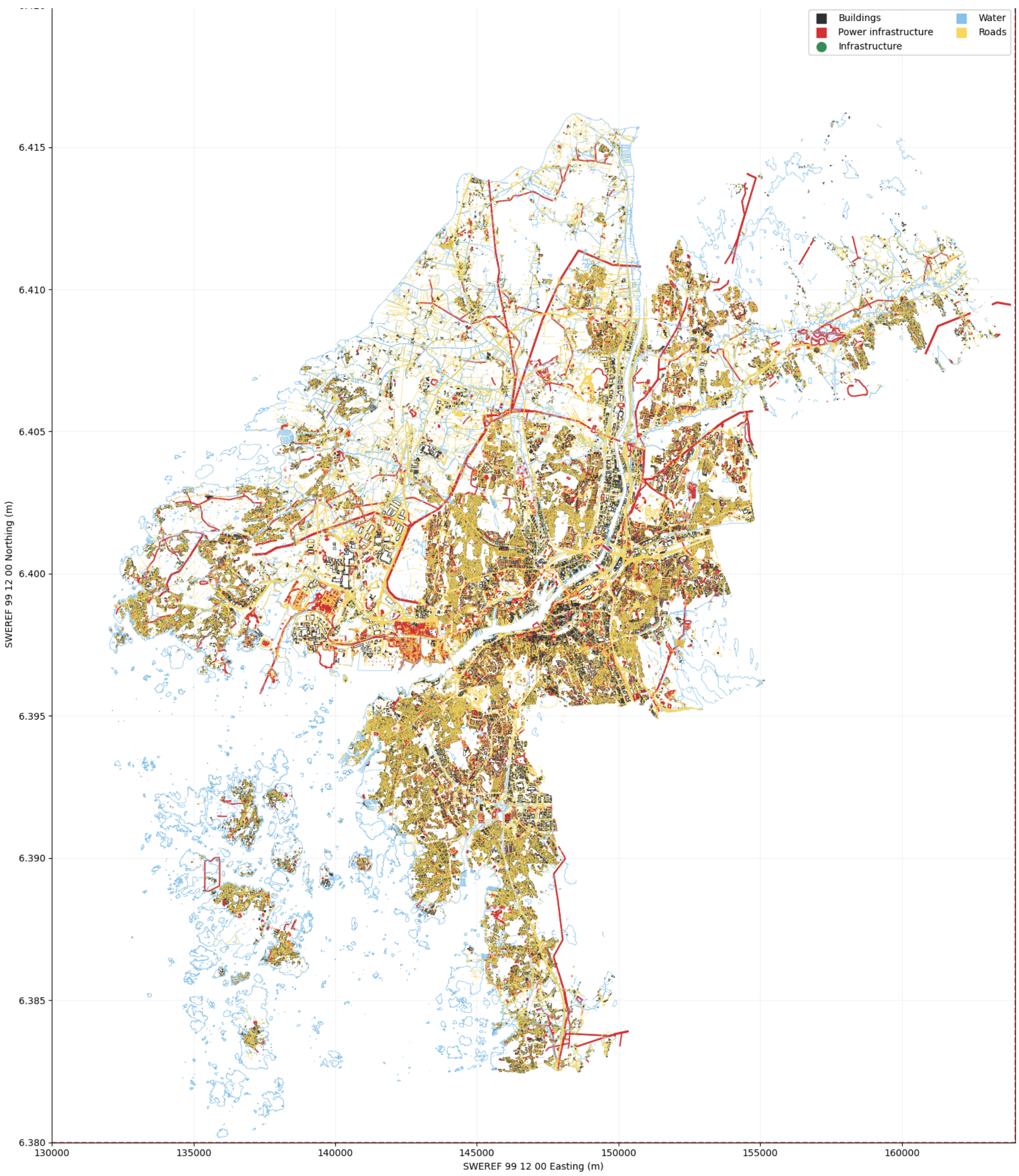

Figure 2.1: Base map over the Gothenburg city area with power infrastructure

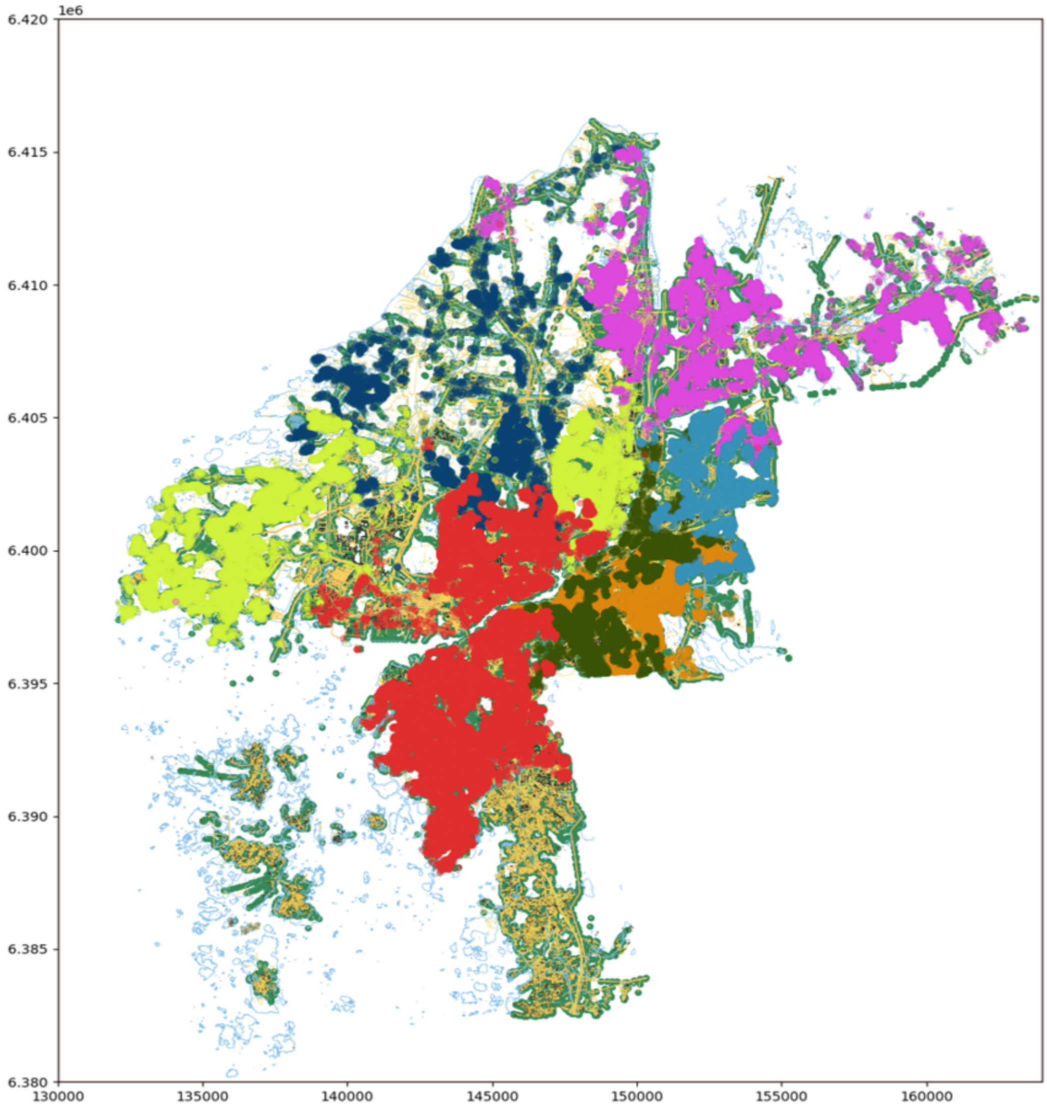

Figure 2.2: Placement of all smart meters with connected locational coordinates

Once streaming data is persisted in storage, it becomes part of a dataset that can be analyzed retrospectively using batch-oriented methods. In the context of this thesis, smart meter measurements originate as data streams but are periodically loaded into the database at 4 hour intervals, where they form structured datasets. The analytical work is performed on these persisted datasets rather than on live streams. This distinction clarifies that the focus is on historical pattern identification and aggregated analysis, rather than real-time control or streaming computation.

2.4 Managerial Concepts

This section outlines the theoretical foundations and managerial concepts used to interpret the empirical data collected in this thesis. The flexibility market can be characterized as a system of interdependent actors where value creation depends not only on the decisions of DSOs, but also on the capabilities and participation of external actors providing flexible resources.

To capture these dynamics, three theoretical perspectives are applied. First, the concept of an innovation ecosystem is used as an overarching lens to understand the broader system in which DSOs and flexibility providers interact. This perspective highlights how value creation emerges through interdependencies between multiple actors, including infrastructure providers, technology owners, regulators, and market platforms [13]. Second, the concept of dynamic capabilities is used to analyze how DSOs internally can respond to the evolving flexibility landscape. This perspective focuses on the organizational processes through which firms sense changes in their environment, seize opportunities, and transform internal structures in order to adapt to new market conditions [20]. Finally, the participation of flexibility providers is examined through innovation diffusion and technology adoption perspectives. These concepts help explain how external actors decide to adopt flexible technologies and engage in flexibility markets, as well as the barriers and incentives shaping their participation [14], [21]. Together, these perspectives provide a conceptual foundation that captures the flexibility market as an evolving ecosystem, the strategic capabilities of DSOs operating within it, and the adoption dynamics influencing participation among complementary actors.

2.4.1 Structure and Development of Innovation Ecosystems

Innovation ecosystems is a concept used to describe the co-dependence between different stakeholders, institutions, and actors that collectively influence value creation [12], [22]. Furthermore, innovation ecosystems places greater emphasis on joint value creation, knowledge flows, and collaborative innovation processes than common innovation system perspectives.

Innovation ecosystems are inherently dynamic and evolve over time as actors enter, exit, or shift roles in response to technological, market, and institutional changes [23], [24]. This dynamic nature implies that ecosystem boundaries are often fluid, especially during the emerging phase of the innovation ecosystem, and that the roles and relationships between actors are continuously reconfigured. As a result, ecosystem analysis must account not only for the structural composition of actors, but also for the processes through which coordination and adaptation occur.

A central implication is that the success of innovation ecosystems depends not only on a focal firm's internal capabilities but also on the alignment of multiple interdependent actors [13]. In addition, innovation success can be distinguished in *execution risk*, which relates to a firm's ability to deliver on its own activities, and the external *ecosystem risks*: *co-innovation risk* and *adoption chain risk* [13], [25]. Co-innovation risk arises when external actors fail to innovate related solutions, risking your innovation to not succeed. The adoption chain risk on the other hand emerges when the organization depends on external actors to adopt your innovation for your offering to provide value. This distinction is particularly relevant in emerging markets, where the timing, incentives, and readiness of multiple actors must align [24], [26].

From a strategic perspective, leveraging an innovation ecosystem can be divided into two related phases. First, the ecosystem must be mapped and understood by identifying key actors, their roles, resources and interdependencies [13]. These actors may include firms, suppliers, customers, regulators, research institutions, and platform providers. Second, firms must develop strategies that explicitly consider innovation ecosystem dependencies and risks, recognizing that innovation outcomes are often dependent on the coordinated actions of multiple actors rather than a single organization.

Furthermore, the ecosystem perspective enables strategic decision-making regarding where, when, and how to compete [13]. Rather than focusing solely on internal capabilities, firms must evaluate which ecosystem positions offer the greatest potential for value creation, influence, and long-term advantage. This includes decisions on whether to shape the ecosystem, adapt to existing structures, or leverage complementarities created by others.

Having established the strategic importance of ecosystem mapping, the next step is to examine the different roles actors may occupy within such systems. Within an innovation ecosystem, actors can assume different strategic roles within the system, such as *keystones*, *complementors*, *niche players*, or *orchestrators* [15], [27]. These roles are defined by the actor's position, the degree of interdependence with other actors and their influence over ecosystem structure and value flows. These forms of influence and interdependence are also reflected in an actor's network position. This includes the extent to which it is highly connected to other firms, acts as a broker or gatekeeper of interactions, or serves as a critical bridge for the functioning of the wider ecosystem [28].

A keystone actor occupies a central and stabilizing role by providing core assets such as platforms, standards, or shared infrastructure that enable other actors to interact and innovate [29]. Rather than capturing all value directly, keystones enhance the overall productivity, robustness, and diversity of the ecosystem. Complementors in turn, develop products or services that enhance the value of the keystone's offering, often focusing on differentiation or specialized capabilities.

Niche players operate in more specialized segments, contributing to ecosystem diversity through focused expertise or targeted solutions, while typically exerting limited influence over overall ecosystem direction [27]. Orchestrators, by contrast, actively coordinate ecosystem participants by aligning incentives, setting strategic direction, and facilitating collaboration among actors. In some ecosystems, the roles of keystone and orchestrator may be combined, while in others they remain distinct.

While these role typologies provide a useful analytical lens, empirical research suggests that ecosystem roles are often fluid and contested rather than fixed [24]. Actors may shift roles over time as technologies evolve or as their capabilities and strategic priorities change, towards for example more regulatory ones. Furthermore, ecosystems may involve asymmetric dependencies, where some actors exert disproportionate influence over standards, access, or value capture [15]. Such asymmetries creates potential tensions between value creation and value appropriation, particularly for smaller actors or new entrants that depend on the ecosystem but have limited ability to shape its rules or capture value.

Coordination is also important in aligning actors in the ecosystem. This is achieved through a combination of formal and informal governance mechanisms such as technological standards, modular interfaces, contractual agreements, and shared platforms that enable distributed innovation while maintaining system coherence [30]. The orchestrator role may also rely on incentive structures, rules, and relationship management to align actors with broader ecosystem objectives. The effectiveness of these coordination mechanisms is critical, as failures in alignment can limit or delay value creation [15], [26]. In this context, ecosystem success depends not only on the presence of capable actors, but also on the existence of mechanisms that facilitate collaboration, reduce uncertainty, and enable interoperability between different components of the system [30].

2.4.2 Dynamic Capabilities and Competitiveness of a Firm

To understand which internal resources and capabilities enable a DSO to navigate rapid changes and uncertainty in energy demand, as well as the tools it deploys within its ecosystem of actors, the concept of *dynamic capabilities* can be applied [20], [31]. As summarized in 2.3, dynamic capabilities refer to the organization's ability to integrate, build and reconfigure internal and external competencies in response to rapidly changing environments [32]. In the literature dynamic capabilities are divided into *sensing*, *seizing*, and *transforming* activities and serves as a foundation for the capabilities that need to be utilized, developed and integrated into the organization to be apt in facing challenges of uncertainty.

Sensing refers to the organizational processes through which firms systematically scan, explore and interpret changes in their technological, regulatory and market environments. It includes activities such as monitoring customer needs, engaging with regulators and ecosystem partners, investing in exploratory initiatives, and evaluating emerging technologies [20].

From an ecosystem perspective, sensing also becomes a leadership activity [31]. Rather than passively observing the environment, firms actively shape and structure their surrounding ecosystem by engaging with complementary actors, institutions, and regulatory bodies to in collaboration identify emerging opportunities [33]. In this view, sensing involves relational coordination and participation in collective learning processes that influence the direction of innovation within the ecosystem.

Seizing refers to the organizational processes through which identified opportunities are evaluated, selected, and translated into concrete strategic commitments [20]. This involves allocating financial and managerial resources, designing appropriate governance structures, and adapting or redefining elements of the business model to capture value. Seizing may include developing new value propositions, establishing partnerships, building new capabilities, and implementing mechanisms for revenue generation. Therefore seizing moves from recognition to execution by incorporating opportunity-driven initiatives into operational and strategic decision-making.

In an ecosystem context, seizing extends beyond internal decision making to encompass leadership in orchestrating complementary actors and aligning incentives across organizational boundaries [31], [33]. Firms must coordinate with partners, negotiate governance arrangements, and ensure value creation and capture mechanisms are viable at the ecosystem level. Seizing therefore entails committing not only firm specific resources but also relational and institutional investments that stabilize and legitimize new configurations of collaboration.

Transforming concerns the continuous renewal and reconfiguration of organizational structures and processes to sustain new strategic directions over time [20]. It involves realigning routines, incentives, and competencies to ensure that new initiatives become integrated into the core business. Transforming may require modifying legacy processes, redistributing decision rights, developing new skill sets, and reshaping organizational mindsets to support the evolving value creation logic. Through such a reconfiguration, the organization strengthens the link between input resources and long-term value creation while maintaining adaptability in the face of ongoing change.

In an ecosystem context, transforming extends beyond internal renewal to involve continuous reconfiguration of roles, relationships, and coordination mechanisms across the ecosystem [31]. This includes engaging in ongoing problem solving to address misalignments, resolve bottlenecks, and maintain the stability of the overall system as it evolves [34]. Ecosystem leaders play a central role in this process by adjusting governance structures, facilitating adaptation among partners, and ensuring that interdependencies remain functional over time and sustaining robustness of the overall system [31]. Transforming thus becomes a relational and systemic activity, where the focal firm not only adapts its own organization but also actively contributes to reshaping the configuration of the ecosystem to sustain collaboration and long-term value creation.

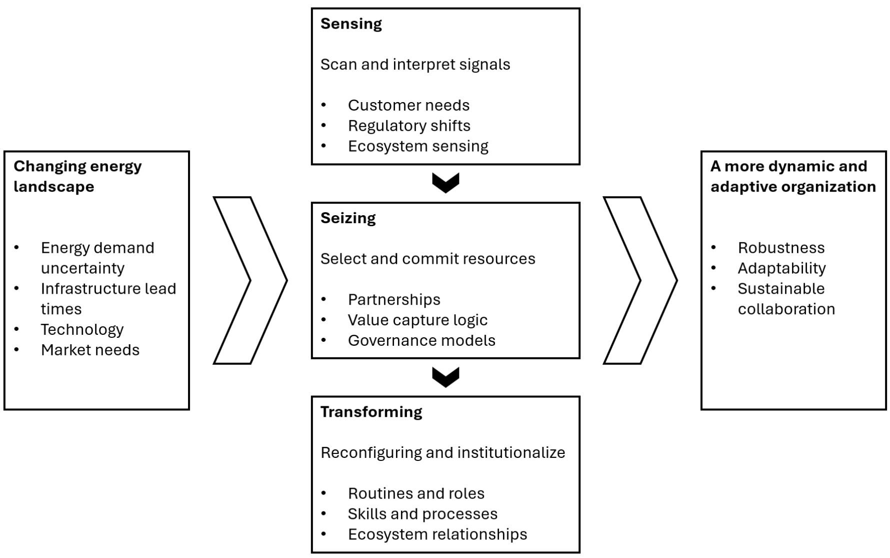

Figure 2.3: Overview of the dynamic capabilities concept and factors

2.4.3 Innovation Diffusion and Market Participation

The concept of *Innovation diffusion* provides a useful theoretical lens for analyzing flexibility market participation because it connects actor-level adoption decisions with the broader spread of new technologies, practices, and market arrangements. Innovation diffusion can be understood from several complementary perspectives. From a sociological perspective, diffusion emerges through the aggregation of many adoption processes, where the adoption of a new idea is shaped by information exchange through communication channels and interpersonal networks [14]. Innovation literature have also analyzed diffusion as a historical process, often represented through S-curves that describe how technologies spread gradually over time [35]. A market-oriented perspective further highlights how diffusion may be reinforced by mechanisms such as economies of scale, learning of scale, learning effects, network externalities, technological interrelatedness, and increasing informational returns [21].

Adoption decisions are influenced by perceived characteristics of the innovation, where relative advantage, compatibility, complexity, trialability and observability are central factors shaping adoption [14], [36]. These characteristics are relevant for flexibility markets, where participation may offer advantages such as additional revenue streams, improved asset utilization, or reduced grid congestion, but can also involve costs, technical uncertainty, measurement requirements, bidding procedures, and uncertainty regarding activation frequency [37]. The perceived compatibility of flexibility with the external actors operational routines and business models therefore becomes central for understanding adoption.

Adoption is also shaped by the characteristics of the adopting actors. These are distinguished between *innovators*, *early adopters*, *early majority*, and *laggards*, reflecting differences in risk tolerance, resources, legitimacy, and openness to change [14]. These actor types may help understand how maturing innovations and emerging markets can perceive increased challenges in gaining new adopters. One such challenge in the literature is *crossing the chasm* between early adopters and the early majority [38]. It is suggested that in contrast to the early adopters who want to support and experiment and be part of the development of the innovation, the early majority rather wants a more complete solution that is ready to be deployed from the start. Thus, attracting these early majority actors might require the implementation and maturity of the technology to be higher in order to take new steps in gaining adopters.

However, diffusion of technologies is not only determined by the perceived characteristics of the innovation or by the willingness of individual actors. Diffusion of technologies is also found to be shaped by system conditions that may enable or challenge adoption [21], [39]. In the literature these challenges are distinguished between actor-level challenges, such as limited knowledge, resources, and motivation, and system-level challenges, such as institutional uncertainty, infrastructure lock-in, weak market formation, and limited interaction between actors. Furthermore, studies of the drivers and barriers in flexibility markets add to this perspective by showing that participation may be constrained by barriers related to market design, administration, trust in relationships, technical prerequisites, and standardization barriers [40].

Applied to flexibility markets, these perspectives suggest that participation can be analyzed both as an adoption decision made by individual actors and as a part of a broader socio-technical diffusion process [21], [40]. Actors such as DSOs, aggregators, property owners, industrial firms, and end-users can be understood as potential adopters that evaluate whether participation aligns with their technical capabilities, organizational routines, resources, and strategic objectives. At the same time, the wider diffusion of flexibility market participation depends on how knowledge experience, trust, standards, and market expectations develop across the system.

2.4.4 Summary of Managerial Concepts

In this section, the theoretical lenses presented in this chapter are summarized. As shown in table 2.1 the summary follows the lenses, its purpose in the thesis, key concepts handled, and analytical relevance.

Theoretical lens	**Purpose in the thesis**	**Key concepts**	**Analytical relevance**
Innovation ecosystem	Used to understand the flexibility market as a system of interdependent actors whose activities must align for value creation to occur.	Actor interdependencies, complementary innovations and actors, ecosystem roles, coordination, governance, platforms, standards, and ecosystem risk.	Helps analyze how actor alignment, roles, and dependencies shape flexibility market development.
Dynamic capabilities	Used to analyze how DSOs respond to technological, regulatory, and market uncertainty in a changing energy landscape.	Sensing, seizing, transforming, resource reconfiguration, organizational adaptation, and ecosystem leadership.	Helps explain how DSOs identify opportunities, commit resources, and adapt internal structures over time.
Innovation diffusion and adoption	Used to explain how and why actors decide to adopt flexible technologies and participate in flexibility markets.	Relative advantage, compatibility, complexity, trialability, actor-level barriers, system-level barriers, and adoption incentives.	Helps analyze the drivers and barriers influencing flexibility providers' participation.

Table 2.1: Summary of theoretical lenses

3

Problem Definition

DSOs are currently navigating a period of profound structural uncertainty [41]. The drivers of electrification, industrial transformation, and changing consumption behavior create increasing pressure on local grid capacity. Because physical reinforcement projects involves extensive lead times and regulatory constraints, operators can no longer rely on expansion alone. Instead, they must now balance resource allocation between traditional infrastructure investments and alternative flexibility-based mechanisms under conditions of incomplete information.

At the same time, the data environment of DSOs has fundamentally changed [42], [43]. High-resolution smart meter data provides detailed observations of when and where electricity is consumed. Weather variables such as temperature and solar irradiance influence both demand and distributed generation. Electricity spot prices introduce an additional market-based signal that can affect consumption patterns. Despite this abundance of data, these streams are often analyzed separately and not systematically integrated into strategic decision-making processes.

3.1 Decision Context in a DSO Environment

DSOs face strategic trade-offs when managing grid capacity [44]. In areas experiencing recurring high load, the operator may choose to reinforce physical infrastructure. Alternatively, flexibility markets can be used to procure temporary reductions in demand from large consumers or aggregators [5]. Tariff-based price signals may incentivize behavioral changes if consumers respond to economic incentives. In more critical situations, contractual curtailment agreements can ensure security of supply.

Each of these mechanisms carries different cost structures, risk profiles, regulatory dependencies, and time horizons [4]. Infrastructure investments provide long-term certainty but require significant capital and long planning processes [10]. Flexibility markets offer adaptive and decentralized solutions but depend on sufficient actor participation and reliable activation [45]. Tariff-based solutions rely on assumptions regarding price elasticity and regulatory feasibility. The difficulty lies not only in choosing between these instruments, but in understanding where and when each mechanism is appropriate within the distribution grid. Such decisions require evidence regarding the drivers of peak load, the responsiveness of consumers to price signals, the spatial concentration of capacity constraints, and the temporal structure of demand volatility.

The opportunity is thus to study how flexibility can be utilized through the DSO's available tools and iteratively examine these together with domain experts at the DSO and with external flexibility providers to assess which patterns are decision-relevant, whether emerging insights are supported by data, and how these insights align with existing tools and planning processes. If certain geographical areas exhibit persistent structural overload independent of weather or price signals, infrastructure reinforcement may be justified. If load peaks correlate strongly with short-term price variation, future tariff design could become a helpful instrument, particularly under evolving regulatory conditions. If peak formation is concentrated among identifiable large consumers, targeted flexibility contracts may be more efficient than generalized measures.

The contribution of this thesis therefore lies in demonstrating how integrated data analysis, combined with iterative qualitative interpretation, can support DSOs in evaluating when to invest in hard infrastructure, when to rely on market-based flexibility mechanisms, and when to prepare for future regulatory opportunities such as more dynamic tariff structures. Rather than proposing a single optimal solution, the work develops a structured analytical and organizational framework that strengthens the evidence base for strategic grid governance under uncertainty.

3.2 Challenges

This section describes the challenges considered in our research. The challenges are divided into three parts, the challenges specific to the quantitative research process, the challenges specific to the qualitative research process, and the challenges specific to the interdisciplinary research process.

3.2.1 State of the Art

Smart meter clustering literature has increasingly focused on the challenges associated with large-scale, volatile, and uncertain consumption data [46]. As smart metering infrastructures have expanded, the analytical problem of flexibility detection and identification is increasingly operationalization as a clustering problem over high-dimensional consumption time series. This reduction is motivated by the absence of labeled data and by the need to infer structure in heterogeneous load profiles without predefined class boundaries. However, traditional time-series clustering approaches have encountered three recurring structural limitations: the computational burden imposed by large customer populations and high temporal granularity, the volatility of individual load profiles which causes interfering effects between distinct load features such as magnitude, overall trend, and short-duration spikes, and the uncertainty arising from the fact that the same customer may exhibit substantially different consumption patterns across different days [46].

A recurring observation in the literature is that these limitations stem from operating directly in the time domain, where metrics such as Euclidean distance between the smart meters temporal vector with values describing their flexibility are sensitive to slight temporal shifts, noise, and scale differences. This can result in either excessive numbers of clusters or large intra-cluster variance when applied to raw smart meter readings [46]. As a result, research has increasingly explored approaches that separate distinct behavioral features prior to clustering, reduce the effective sample size through staged representations, and model uncertainty probabilistically rather than eliminating it through temporal averaging.

Within this context, three classical clustering families have been identified as relevant to smart meter analytics, each addressing different structural requirements [46].

Connectivity-based clustering methods, such as hierarchical agglomerative clustering, capture the nested relational structure between consumers and do not require the number of clusters to be specified a priori. However, these methods require the construction of a full pairwise distance matrix, which becomes computationally prohibitive at the scale of smart meter deployments. Ward's minimum-variance linkage is one commonly applied variant in load profiling studies, valued for producing compact and interpretable groupings [47].

Centroid-based clustering methods, such as K-Means, are among the most widely adopted approaches due to their computational efficiency and scalability [46]. These methods partition the feature space around representative centroids, but their reliance on time-domain distance metrics makes them sensitive to the volatility and uncertainty characteristic of individual smart meter data. A spike or minor temporal shift in one load profile can disproportionately inflate inter-cluster distances, causing similar profiles to be assigned to different clusters [46]. MiniBatch K-Means extends the scalability of K-Means further by updating centroids using randomly sampled data subsets rather than the full dataset at each iteration, substantially reducing memory and computational overhead while preserving acceptable clustering quality [48].

Distribution-based clustering methods, of which Gaussian Mixture Models (GMMs) are the most prominent example in load profiling, characterise each cluster as a probability distribution rather than a deterministic centroid [46]. This is particularly suited to smart meter data, where day-to-day behavioural variability means that a single customer may legitimately belong to different clusters on different days. By assigning probabilistic cluster memberships rather than hard assignments, distribution-based methods preserve the uncertainty inherent in individual consumption behaviour rather than artificially eliminating it [46]. Density-based clustering methods, such as DBSCAN [16], offer a related advantage in that they can identify non-linear cluster structures and explicitly isolate outlier consumers without forcing all observations into predefined assignments. This is particularly relevant for flexibility analysis, where irregular consumption patterns may themselves carry operational significance. However, density-based methods remain sensitive to parameter selection and can produce fragmented or unstable structures across heterogeneous populations.

Taken together, the literature demonstrates that no single clustering family universally dominates across all smart meter applications. The suitability of a given approach depends not only on the statistical properties of the data, but on the operational and organisational context in which the resulting clusters must be used [46]. In particular, when clustering is intended to support high-level managerial decision-making such as flexibility procurement, consumer segmentation for aggregators, or strategic communication by a Distribution System Operator (DSO), the interpretability and granularity of the resulting clusters become binding constraints alongside mathematical performance. A method that produces thousands of micro-clusters may achieve high intra-cluster similarity, but renders the output operationally meaningless if no human or organisational process can act on it. The following section operationalises these constraints in the context of the present study.

3.2.2 Quantitative Study Challenges

A central challenge of the quantitative analysis lies in the fact that the available data is unlabeled. There is no pre-defined classification of which consumers represent flexibility resources, nor any established ground truth against which clustering performance can be evaluated in a traditional supervised learning sense. The objective is therefore not to optimise predictive accuracy, but to produce clusters that are sufficiently interpretable, computationally feasible, and operationally meaningful within the DSO context established by the qualitative study.

This constraint directly shapes which methodological families, as reviewed above, are appropriate for the present analysis. Density-based approaches such as DBSCAN, while capable of detecting non-linear behavioral structures and isolating outliers, are prone to producing highly fragmented cluster solutions under heterogeneous smart meter populations [16]. In a dataset where thousands of meters are observed at high temporal resolution, such methods risk generating outputs in which the majority of consumers are assigned to micro-clusters of one to five meters. While analytically precise, this granularity is incompatible with the use cases identified in the qualitative study: an aggregator seeking to identify actionable consumer segments, or a DSO seeking to design targeted flexibility initiatives, cannot operationalise a solution composed of thousands of near-identical behavioural fragments. Connectivity-based methods face analogous limitations from the opposite direction, as full distance matrix construction becomes computationally infeasible at the scale of this dataset [46].

Beyond method selection, the study is further simplified by the geographic scope of the analysis. Since all consumers reside within the same distribution region, spatial heterogeneity between regions does not need to be accounted for within the clustering framework. The clustering objective is therefore purely behavioral: to identify consumption patterns that are strategically relevant, rather than to explain geographic variation in demand. This reduces the dimensionality of the problem and strengthens the case for centroid-based and distribution-based approaches, which partition the behavioral feature space directly without requiring spatial covariates.

To address these constraints, the thesis defines flexibility primarily through surge behavior, defined as the deviation between infrequent peak consumption events and typical consumption patterns during comparable time periods. Rather than pursuing maximally granular or theoretically optimal clusters, the objective is to construct behaviorally coherent groupings that capture strategically relevant consumption characteristics while remaining interpretable to the managerial stakeholders identified in the qualitative study. The clustering framework therefore prioritises solutions that produce a manageable number of distinct, explainable segments capable of supporting the use cases identified in the qualitative study: assisting aggregators in identifying relevant consumer segments, supporting future flexibility procurement initiatives, and enabling the DSO to better understand and engage with strategically significant consumption behaviors.

From a computational perspective, the system must operate on high-volume streaming data within a fixed processing window. This imposes strict limitations on algorithmic complexity and necessitates staged aggregation. Raw 15-minute consumption data is therefore progressively transformed into reduced representations suitable for statistical modeling and clustering.

From an organizational perspective, outputs must be interpretable and usable in operational decision-making processes. This includes supporting communication with external actors such as aggregators, as well as internal planning and flexibility procurement strategies. As a result, model complexity is balanced against interpretability, and excessive segmentation is avoided when it reduces practical usability.

3.2.3 Qualitative Study Challenges

In this section, the challenges related to conducting a qualitative study are presented. How these challenges are managed is further expanded in chapter 4.3.

Conducting a case study on a DSO is appropriate for analyzing a real world context in an industrial city and at an energy company that is heavily investing in enabling electrification during turbulent times [4]. This context provides an opportunity for insights into real world problems and actions not otherwise achieved through purely theoretical research. Although a case study is limited in terms of external validity and generalization, they allow for an in-depth examination crucial for the goal of this thesis [49]. Thus, insights from the case study have to be objectively analyzed to determine the generalizability boundaries that could emerge in the research.

The main challenge with conducting qualitative research is thus the generalizability of the results in similar contexts. For example, it is uncertain whether perceptions from a DSO or external actors be generalized to a common understanding. Another challenge to be considered is if these findings will come to change over time as the flexibility landscape changes or as the studied organization transforms. The benefit of handling these challenges in the study is the increased relevance of the result to the organization and the inclusion of factors from different actors that risk being missed in pure theoretical research. Furthermore, this study can gain further insights that are simply not yes-or-no answers, but can show the complexity and nuance of the problem.

Another challenge is the scope and saturation of qualitative data from interviews that can be achieved in the limited time of the thesis. The data collected are dependent on multiple factors, such as interview design, interview process, respondent availability, political events that impact the respondent, and their willingness to give honest accounts of their operations. Thus, the data must be actively interpreted with these factors in mind when synthesizing the findings.

3.2.4 Interdisciplinary Challenges

A central challenge in this thesis concerns the sequencing and integration of findings across the quantitative and qualitative parts of the research. Although the study is designed as an iterative process, the two analytical streams do not necessarily develop at the same pace or produce findings at the same time. Quantitative analysis requires aggregation, feature construction, and validation before interpretable results can be produced. In contrast, qualitative insights from interviews and organizational engagement in the study may emerge earlier, later, or in parallel, depending on stakeholder availability, interview timing, and the evolving understanding of the problem context.

This creates a challenge for how findings are aligned during the research process. For example, early qualitative insights may guide the quantitative analysis before the full picture of the organization and external actors landscape has fully emerged. In contrast, data analysis can reveal unexpected patterns only after interviews have been conducted, limiting the possibility of immediately exploring these findings with relevant stakeholders.

Another issue concerns the extent to which stakeholders can or are able to discuss the implications of the research findings. Interviewees may be restricted by data sensitivity, commercial considerations, specific know-how, or uncertainty about future regulatory and market developments. They may also avoid speculating about what types of data or analytical outputs would be most useful before concrete results are available. As a result, stakeholder input must be interpreted in light of these limitations, and the research design must account for variation in access, disclosure, and perspective across respondents.

This sequencing challenge is particularly important because the value of the interdisciplinary approach depends on the interaction between the two perspectives. The challenge is therefore not only to conduct two separate analyzes, but to create meaningful points of connection between them throughout the research process.

3.3 Scope

This thesis is positioned as a decision-support study for a distribution system operator, focusing on how integrated Internet of Things data streams can be translated into decision-relevant insights. The scope is deliberately constrained to ensure analytical depth, interpretability, and practical relevance. The thesis does not aim to establish causal effects, optimize flexibility market design, or generalize results beyond contexts comparable to the studied distribution grid. Findings are interpreted as empirically grounded insights derived from observational data, intended to inform rather than replace existing planning judgment.

The research targets results in certain limitations of scope and depth of the thesis, particularly those related to resulting considerations of factors, number of interviewees, and data, influencing the system and organization. However, the research targets to contribute to the dynamics of a flexibility market in an energy grid company context.

3.4 Computational Efficiency

The focus of this thesis is not on optimizing computational performance or benchmarking execution speed of data processing pipelines. However since computational efficiency is a crucial aspect about all feasibility of data processing systems, this study will examine and reason about the time complexity for each implemented step into the pipeline. The computational approach prioritizes robustness, transparency, and practical feasibility within an operational analytics environment. Methods and data summaries are required to be computationally reasonable and compatible with recurring data updates, but execution time is evaluated only against basic operational constraints rather than theoretical optimality. As a result, solutions are designed to be "good enough" for decision-support purposes, rather than maximally efficient in a computational sense.

3.4.1 Ecosystem Perspective

The ecosystem perspective adopted in this thesis is intentionally limited to actors that materially influence flexibility-related planning and investment decisions at the distribution level. This includes stakeholders whose actions affect load behavior, net-load variability, or the feasibility of flexibility measures. The analysis does not aim to provide a comprehensive mapping of the energy ecosystem. Regulatory design, national market structures, and peripheral actors are considered only insofar as they directly constrain or enable distribution-level decision-making. This bounded ecosystem view ensures that the analysis remains focused on decision-relevant interactions rather than broad sectoral description.

3.4.2 Internal Organizational Perspective

The internal perspective is limited to understanding how analytical outputs can support existing planning and investment processes within a distribution system operator. The thesis does not seek to evaluate the broader organizational performance, governance structures, or internal incentives unrelated to flexibility and the flexibility market. Iterative interviews with selected organizational stakeholders are used to identify priority questions, interpret analytical results, and assess the practical relevance of proposed indicators.

3.4.3 Transparent and Explainable Modeling

A central constraint of this thesis is the requirement for transparency and explainability in all analytical methods. Models and indicators must allow inspection of assumptions, drivers, and relationships between variables. Black-box approaches that do not permit meaningful interpretation of results are explicitly excluded. Predictive performance is considered important, but never at the expense of interpretability and examinability. This emphasis reflects the role of the analysis as input to planning and investment decisions, where results must be explainable to both technical and non-technical stakeholders and withstand internal scrutiny, and not just "the AI told me to".

4

Methods

This chapter outlines the methodological choices conducted to investigate how the combination of quantitative and qualitative analysis can support a DSO's strategic grid governance.

4.1 Interdisciplinary Study

This study adopted an interdisciplinary approach combining innovation management and computer science engineering in a quantitative study and a qualitative study to gain further depth than one discipline alone would. Mixed methods studies are found to provide a more holistic and nuanced view, however these benefits also come with challenges involving their difference in nature, sequence and which approach that takes precedence and priority [49]. The quantitative and qualitative studies were conducted concurrently, by gathering insights about what is quantitatively possible and visible, simultaneously, as the empirical findings from interviews gave guidance and nuance to areas of interest in the flexibility topic. Both studies interacted iteratively and helped triangulate the evidence continuously, providing additional analytical depth.

In terms of sequencing the two studies, the research team took part in findings continuously by both participating in the interview study where guiding questions to the quantitative study could be made and during internal meetings where quantitative insights and possibilities were discussed to further frame the study and point out important stakeholders in terms of technical opportunity and needed understanding. During the analytical synthesis the results were discussed and put in relation to one another further pointing to relevant themes and discussion in the thesis.

4.1.1 Interdisciplinary Process

In this section, the interdisciplinary process is described in more detail. Our study had an initial exploratory phase in which the study was more broad in understanding the problem and system of the DSO. A middle phase, where the study was focused on flexibility and DSO tools. Lastly, a more validating phase where the relevance of our findings was discussed with stakeholders.

During the initial exploratory phase the quantitative work focused on understanding the smart meter data and assessing the possibilities of combining it with spatial data, weather data, and spot price data. At the same time, exploratory interviews were conducted to develop an initial understanding of the DSO's organizational context, relevant operational processes and potential stakeholders. This parallel exploration helped establish both technical possibilities and the organizational relevance of the study.

The early interaction between the two studies played an important role in shaping the research direction. Initial data experimentation clarified what types of flexibility-related patterns could be made and what analytical limitations existed. In parallel, the interviews provided contextual understanding of how such patterns could relate to the DSO's operations, decision-making processes, and stakeholder needs. Reflections after interviews were therefore used to reconsider the relevance of different data topics, while emerging quantitative insights helped inform new questions and topics that could be explored during interviews.

As the study progressed, the interdisciplinary process shifted from exploration and scoping toward gathering more findings and ultimately validation and analytical refinement. Quantitative findings were discussed in relation to interview insights, while respondent accounts were used to interpret the practical meaning and organizational relevance of the data-driven results. This iterative process helped ensure that the analysis did not rely only on what was technically observable in the data, but also considered how such observations could be understood and applied in practice. However, this stage was also challenging due to limited respondent availability and limited access to operational data that could validate the findings further.

4.2 Quantitative Study

The quantitative analysis is structured as an analytical study of electricity consumption behavior from high-frequency smart meter data. The objective is not predictive accuracy against labeled outcomes, but the construction of interpretable behavioral structures that can support identification of flexibility potential within the distribution grid.

A central challenge is that flexibility is not directly observable in the data and does not exist as a labeled variable. Consequently, flexibility must be evaluated through proxy constructs derived from deviations in consumption behavior, in particular surge-like events relative to expected load patterns.

4.2.1 Data-to-Insight Pipeline Architecture

The analytical workflow is structured as a sequential data-to-insight pipeline that progressively transforms high-frequency smart meter observations into interpretable behavioral representations of consumer flexibility. Rather than analyzing individual meters directly, the methodology first establishes regional consumption dynamics at the aggregate level before systematically de-aggregating the analysis to identify and characterize individual behavioral patterns.

The pipeline consists of five stages:

1. **Data loading and preparation:** Smart meter measurements, weather observations, electricity market data, and spatial information are integrated into a unified analytical dataset and prepared for distributed processing.

2. **Aggregate consumption construction:** Individual meter observations are aggregated into regional consumption profiles that capture the overall demand dynamics of the study area and provide the basis for system-level modeling.

3. **Regional baseline modelling:** Aggregate consumption patterns are modeled using statistical time-series techniques to estimate expected system behavior and identify the dominant temporal, seasonal, and external drivers of demand.

4. **Meter-level surge identification:** The aggregate analysis is complemented by a disaggregated examination of individual meter behavior, where statistically significant consumption surges are detected and quantified for each consumer.

5. **Behavioral clustering:** Meter-specific surge characteristics are transformed into behavioral feature representations and grouped into clusters of consumers exhibiting similar flexibility-related consumption patterns.

This architecture establishes a clear separation between system-level consumption modeling and consumer-level behavioral analysis. Aggregate modeling provides a regional understanding of demand dynamics, while the subsequent disaggregation enables the identification of individual flexibility signals that form the basis for behavioral clustering and strategic flexibility assessment.

Data Pipeline and Implementation

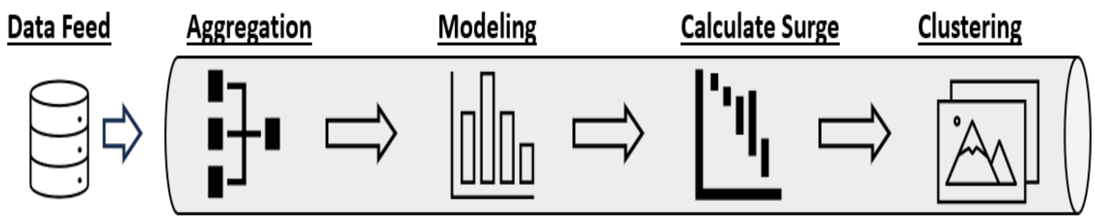

Figure 4.1: Overview of data pipeline functionality

The analytical pipeline, see figure 4.1 is implemented as a distributed data processing system designed to handle high-frequency smart meter data under near-real-time operational constraints. The system combines distributed computation via PySpark with in-memory statistical modeling in Python and Pandas, enabling scalable transformation of raw consumption data into structured behavioral representations. Two primary requirements constrain the design throughout: computational feasibility under large-scale streaming data ingestion, where new observations arrive at 15-minute intervals, and operational deployability, where all transformations must remain executable within bounded processing windows.

To satisfy these constraints, the methodology is implemented as a staged reduction process, in which raw data is progressively transformed into lower-dimensional representations prior to modeling and clustering. This ensures computational tractability while preserving temporal and behavioral structure.

4.2.2 Computational Complexity Summary and Notation

Table 4.1 summarizes the asymptotic complexity of each pipeline stage, together with the main bottleneck for each. The overall pipeline is dominated by the SARIMAX state-space estimation and the full Ward dendrogram construction.

Table 4.1: Computational complexity of each stage in the data pipeline.

Stage	**Complexity**
Data loading and ingestion	$\mathcal{O}(N \cdot T)$
Hourly aggregation	$\mathcal{O}(N \cdot T)$
ARIMA (grid search + fitting)	$\mathcal{O}(T \log T)$
SARIMAX	$\mathcal{O}(T^2)$
Daily meter aggregation (Spark)	$\mathcal{O}(N \cdot T)$
Surge flagging (broadcast join)	$\mathcal{O}(N \cdot T)$
Ward hierarchical clustering	$\mathcal{O}(n^2)$
DBSCAN (cosine)	$\mathcal{O}(n^2)$
Mini-batch k-means	$\mathcal{O}(b \cdot k \cdot i \cdot T)$

N = number of meters;
T = number of time points in aggregated series;
D = number of days;
n = number of meters with detected surges;
b = mini-batch size; k = number of clusters; i = clustering iterations.

4.2.3 Surge Definition and Behavioral Deviation Extraction

Surges are defined as statistically significant positive deviations from expected consumption behavior. Let ε_t denote the deviation at time t, measured as the difference between observed consumption and an expected baseline.

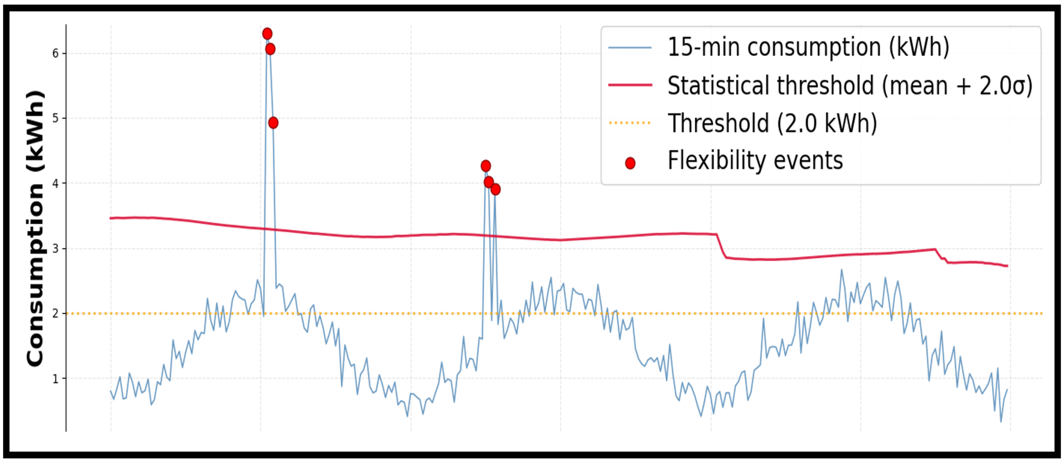

Figure 4.2: Illustrative example of a surge event representing a temporary increase above expected consumption.

A surge is identified when a deviation exceeds a threshold derived from the underlying distribution of observations:

$\varepsilon_t > \mu_\varepsilon + \lambda\sigma_\varepsilon$

where ε_t denotes the deviation at time t, μ_ε is the mean deviation over the reference period, σ_ε is the corresponding standard deviation, and λ is a threshold multiplier controlling detection sensitivity.

This formulation isolates the upper tail of the deviation distribution, capturing infrequent but potentially significant increases in consumption. To improve operational relevance, the statistical criterion is complemented by an absolute magnitude threshold, preventing minor fluctuations in low-consumption contexts from being classified as flexibility signals. The resulting surge events transform continuous consumption measurements into discrete behavioral observations that can subsequently be aggregated, analyses, and clustered.

The surge threshold combines a statistical criterion with a minimum energy criterion. The baseline consumption is denoted as a moving average including the surges like the red line seen in Figure 4.2. A threshold of two standard deviations was selected to limit detections arising from normal day-to-day consumption variability while retaining events that represent meaningful increases in demand. Under a normal distribution, only approximately 2.3% of observations exceed this threshold, providing a simple mechanism for filtering routine fluctuations. In addition, a minimum threshold of 2 kWh per 15-minute interval was imposed to ensure operational relevance. This value was chosen with electrification-driven flexibility in mind, as a typical 11 kW residential EV charger corresponds to approximately 2.75 kWh over a 15-minute period [50]. The threshold was set slightly lower to also capture lower-power charging technologies and other flexibility-related consumption events. Together, these criteria balance statistical significance with practical relevance while avoiding excessive sensitivity to normal consumption cycles.

Behavioral Clustering of Flexibility Signals

Surge events are aggregated at the meter level to construct behavioral profiles representing each consumers deviation patterns over time. These profiles form the input to clustering algorithms designed to identify groups of consumers with similar flexibility-related behavior.

Multiple clustering approaches are evaluated, including hierarchical clustering, centroid-based methods, and density-based clustering. The primary evaluation criteria are:

- Computational scalability under large datasets
- Interpretability of cluster structure
- Balance and stability of cluster sizes
- Operational usefulness for flexibility targeting

Unlike traditional clustering applications, the objective is not purely statistical separation, but the construction of meaningful behavioral segments that can support downstream operational use cases.

4.2.4 Data Sources and Integration

The pipeline integrates four data sources into a unified temporal structure:

- **High-frequency smart meter consumption data** (15-minute resolution), stored in a distributed table with fields `METER_ID`, `TIME_STAMP`, and `CONSUMPTION_KWH`.
- **Spatial metadata**, sourced from data structure mapping meter identifiers to geographical areas and grid-level regions.
- **Weather observations**, providing hourly air temperature aligned to UTC timestamps.
- **Electricity spot price data** at 15-minute resolution, expressed in EUR/MWh.

All datasets are temporally aligned and aggregated to a common hourly resolution to support consistent time-series modeling. This integration enables the analysis of consumption behavior as a function of both internal system dynamics and external drivers such as temperature and price signals.

Data Ingestion and Area Filtering

The ingestion stage performs three parallel loading operations before joining them into a single analytical structure. Location filtering is applied first, reading spatial metadata from the Parquet file, an efficient file format which compresses the data. Retaining only meters within the target geographic area. The resulting set of identifiers is broadcast across the cluster, enabling a broadcast join against the base consumption table containing the entries for each smart meter:

```
location_ids = (
```

```
    spark.read.parquet(LOCATION_FILE)
    .filter(F.col("area") == AREA_FILTER)
    .select("location_id")
    .distinct()
)

cons_15 = (
    spark.read.table("15min_parquet_file")
    .filter(
        (F.col("TIME_STAMP") >= START_DATE)
        & (F.col("TIME_STAMP") < END_DATE)
    )
    .join(broadcast(location_ids), on="location_id", how="inner")
    .dropna(subset=["kwh"])
    .cache()
)
```

Caching `cons_15` after the join is intentional: the same filtered consumption frame is reused in both the aggregate time-series modeling stage and the per-meter surge detection stage, and materialising it once avoids redundant re-execution of the shuffle-intensive join.

Hourly aggregation is then applied to produce the series used by the time-series models. Timestamps are truncated to hour boundaries via `F.date_trunc("hour", ...)`, after which a `groupBy()` and `agg()` reduce the 15-minute frame to hourly totals:

```
hourly_sdf = (
    cons_15
    .withColumn("hour", F.date_trunc("hour", "ts"))
    .groupBy("hour")
    .agg(F.sum("kwh").alias("kwh"))
)

hourly_pdf = (
    hourly_sdf
    .toPandas()
    .set_index("hour")
    .sort_index()
    .asfreq(pd.tseries.frequencies.to_offset("h"))
)
hourly_pdf["kwh"].interpolate(method="time", limit=3, inplace=True)
```

Missing intervals of up to three consecutive hours are filled through time-weighted linear interpolation. Longer gaps require explicit repair and are handled separately in the baseline modeling stage.

Weather and price series are loaded analogously. Temperature observations are grouped by hour using a custom truncation function and averaged within each window; price records originally at 15-minute resolution are resampled to hourly means using `.resample("h").mean()`. Both series are subsequently joined to `hourly_pdf` and forward- and backward-filled for short remaining gaps, producing the exogenous regressor matrix used by the SARIMAX model.

Data Integrity and Repair

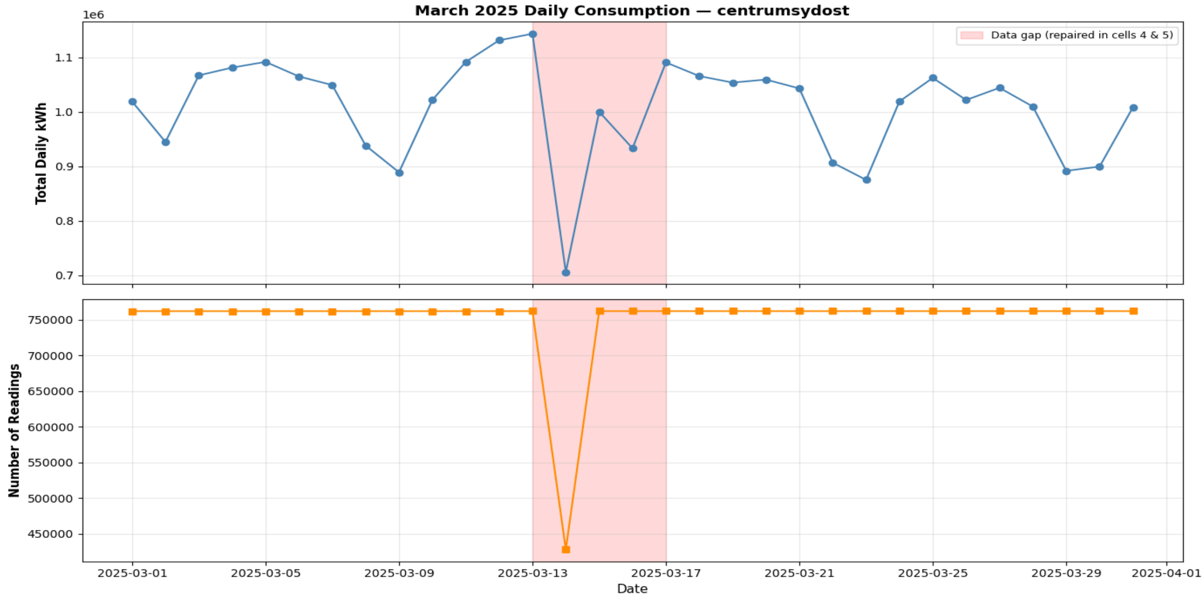

Figure 4.3: Reconstruction of consumption data

The consumption time series contains intermittent missing periods caused by data collection gaps. Simple linear interpolation is avoided for longer outages because it produces unrealistically flat stretches that distort seasonal decomposition and residual statistics, see 4.3. Instead, a pattern-preserving substitution approach is applied: temporally adjacent periods with similar seasonal characteristics serve as donors, preserving the intra-day load shape that is central to accurate baseline estimation.

For the identified five-day reporting gap, hourly patterns from the immediately following five days are mapped back onto the missing period:

```
donor_offset = pd.Timedelta(days=5)
donor_values = hourly_repaired.loc[gap_end : gap_end + donor_offset, "kwh"]
donor_values.index = donor_values.index - donor_offset
hourly_repaired.loc[gap_mask, "kwh"] = donor_values.reindex(
    hourly_repaired.loc[gap_mask].index
)
```

Donor days are selected from climatically similar periods to ensure that temperature-driven daily profiles remain consistent with the surrounding observations. This repair is applied exclusively to the aggregate hourly series used for time-series modeling; per-meter heatmaps and surge counts retain the original, unmodified data.

4.2.5 Baseline Consumption modeling

To separate structural consumption dynamics from anomalous deviations, a baseline expected consumption trajectory is estimated from the repaired aggregate series. Two model specifications are evaluated.

ARIMA An ARIMA(p, d, q) model is fitted to the univariate hourly aggregate series [51]. The specification is selected via grid search over the ranges $p \in \{0, 1, 2, 3\}$, $d \in \{0, 1\}$, $q \in \{0, 1, 2, 3\}$, with model selection based on the Akaike Information Criterion (AIC) [18], which penalises model complexity while rewarding goodness of fit. The optimal order is stored for reuse in the SARIMAX specification. The resulting fitted values define the aggregate baseline against which residuals are measured.

SARIMAX The ARIMA specification is extended with a seasonal component capturing the 24-hour daily cycle, as well as exogenous regressors including hourly temperature, spot price, and cyclical time-of-day features [19]:

$$\text{hour\_sin}_h = \sin\left(\frac{2\pi h}{24}\right), \quad \text{hour\_cos}_h = \cos\left(\frac{2\pi h}{24}\right) \tag{4.1}$$

The SARIMAX seasonal order is fixed at $(P, D, Q, s) = (1, 0, 1, 24)$, reflecting a single seasonal autoregressive and moving-average term with a 24-hour period [19]. Stationarity and invertibility constraints are relaxed to permit a wider search over the coefficient space:

```
model = SARIMAX(
    endog=y_train,
    exog=X_train[["avg_temp", eur_col, "hour_sin", "hour_cos"]],
    order=best_order,
    seasonal_order=(1, 0, 1, 24),
    enforce_stationarity=False,
    enforce_invertibility=False,
)
result = model.fit(maxiter=200, disp=False)
```

Incorporating external drivers allows the SARIMAX residuals to isolate consumption deviations that cannot be explained by temperature, price, or time-of-day patterns, thereby reducing spurious surge detections.

4.2.6 Per-Meter Surge Detection

While baseline modeling operates on aggregate consumption, flexibility analysis requires disaggregation to the individual meter level. The transition from system-level anomalies to consumer-level behavioral signals is achieved through a two-stage Spark aggregation.

Daily per-meter statistics. The cached 15-minute frame is grouped by meter identifier and calendar date, computing for each meter-day the count of intervals, the total consumption, and the sum of squared values needed for a one-pass variance estimate:

```
meter_day = (
    cons_15
    .withColumn("date", F.to_date("ts"))
    .groupBy("location_id", "date")
    .agg(
        F.count("kwh").alias("cnt"),
        F.sum("kwh").alias("sum_kwh"),
        F.sum(F.col("kwh") * F.col("kwh")).alias("sum_sq")
    )
)
```

This `groupBy().agg()` step reduces approximately three million 15-minute rows to around 30,000 meter-day rows, a 100-fold dimensionality reduction that makes the subsequent rolling window computationally feasible.

Sliding time window. Rolling statistics are computed over a centred seven-day window spanning ± 3 days around each observation, capturing the weekly behavioral cycle while adapting gradually to seasonal trends. The one-pass standard deviation formula $\sigma = \sqrt{(\Sigma x^2 - (\Sigma x)^2/n)\,/\,(n-1)}$ is applied directly to the aggregated sums, avoiding re-materialisation of the full 15-minute data:

```
w_day = (
    Window.partitionBy("location_id")
    .orderBy("date")
    .rowsBetween(-3, 3)
)

meter_day_thresh = (
    meter_day
    .withColumn("r_cnt",   F.sum("cnt").over(w_day))
    .withColumn("r_sum",   F.sum("sum_kwh").over(w_day))
    .withColumn("r_sumsq", F.sum("sum_sq").over(w_day))
    .withColumn("roll_mean", F.col("r_sum") / F.col("r_cnt"))
    .withColumn("roll_std", F.sqrt(
        (F.col("r_sumsq") - F.col("r_sum") * F.col("r_sum")
         / F.col("r_cnt"))
        / F.greatest(F.col("r_cnt") - 1, F.lit(1))
    ))
    .withColumn(
        "threshold",
        F.col("roll_mean") + SURGE_SIGMA * F.col("roll_std")
    )
)
```

The choice of a seven-day window balances several competing requirements. A window of seven days provides approximately 672 15-minute intervals per meter, yielding stable estimates of the local mean and standard deviation. The centred specification uses both past and future days, making the rolling statistics more representative of genuine local behavior than a purely backward-looking window. Finally, the weekly span captures the structural difference between weekday and weekend consumption profiles, which is particularly important for meters with strong business-hours periodicity.

Surge flagging via broadcast join. The threshold frame, now containing one row per meter-day, is broadcast to all executors and joined back to the full 15-minute data. A dual condition is applied: the statistical condition ensures the deviation is significant relative to local variability, while the absolute condition filters trivial fluctuations in low-consumption meters:

```
surges = (
    cons_15
    .withColumn("date", F.to_date("ts"))
    .join(broadcast(meter_day_thresh),
          on=["location_id", "date"], how="inner")
    .withColumn(
        "excess_kwh",
        F.greatest(F.col("kwh") - F.col("roll_mean"), F.lit(0.0))
    )
    .withColumn(
        "is_surge",
        (F.col("kwh") > F.col("threshold"))
        & (F.col("excess_kwh") > SURGE_ABS_KWH)
    )
    .groupBy("location_id", "date")
    .agg(
        F.sum(F.col("is_surge").cast("int")).alias("surge_count"),
        F.sum(
            F.when(F.col("is_surge"), F.col("excess_kwh"))
            .otherwise(0.0)
        ).alias("excess_kwh")
    )
)
```

The resulting `daily_pdf` frame carries one row per meter-day and two target columns: `surge_count`, recording the number of 15-minute surge intervals on that date, and `excess_kwh`, recording the total consumption above the rolling mean during those intervals. The formal surge condition at meter level i and time t is:

$$x_{i,t} > \mu_{7d}(t) + \lambda\,\sigma_{7d}(t) \quad \text{and} \quad x_{i,t} - \mu_{7d}(t) > \delta \tag{4.2}$$

where $\mu_{7d}(t)$ and $\sigma_{7d}(t)$ are the seven-day rolling mean and standard deviation, $\lambda = 2$ is the statistical multiplier, and $\delta = 2.0kWh$ is the absolute minimum excess (`SURGE_ABS_KWH`).

The dual-threshold design reflects a deliberate methodological choice. The statistical condition alone can be triggered by high relative variation in meters with very low baseline consumption, where even a fraction of a kilowatt-hour of excess registers as a multi-sigma event. The absolute condition prevents such structurally insignificant deviations from being classified as flexibility signals, ensuring that the surge inventory represents operationally meaningful behavior.

Behavioral Representation and Feature Construction

Each meter is represented as a fixed-dimensional behavioral vector for clustering. Two pivot tables are constructed from `daily_pdf`, transforming the long-format meter-day frame into wide-format matrices in which each row corresponds to a meter and each column to a calendar date:

```
pivot_surge = daily_pdf.pivot(
    index="location_id",
    columns="date",
    values="surge_count"
).fillna(0)

pivot_flex = daily_pdf.pivot(
    index="location_id",
    columns="date",
    values="excess_kwh"
).fillna(0)
```

Only meters that recorded at least one surge throughout the analysis period are included. Meters with no surge activity exhibit structurally stable consumption profiles and are excluded from the clustering step, as they do not contribute observable flexibility signals. The surge-count matrix `pivot_surge` serves as the primary clustering input, with the flexibility potential matrix `pivot_flex` used for post-clustering quantification of excess energy by cluster.

Looking to the behavioral clustering of flexibility patterns, Ward's minimum-variance hierarchical clustering is applied to the surge-count matrix. The algorithm merges pairs of clusters iteratively, selecting at each step the merge that minimises the increase in total within-cluster variance:

$$\Delta(C_i, C_j) = \frac{|C_i|\,|C_j|}{|C_i| + |C_j|}\,\|\mu_i - \mu_j\|^2 \tag{4.3}$$

Ward's method is preferred over centroid-based alternatives for two reasons. First, it minimises within-cluster variance directly, producing compact and interpretable behavioral groups rather than optimising a geometric distance to a predefined centroid. Second, it generates a full dendrogram that makes the hierarchical structure explicit and supports data-driven selection of the cluster count.

Full Ward's clustering has $\mathcal{O}(n^2)$ time complexity and $\mathcal{O}(n^2)$ space complexity [52], which becomes infeasible when the number of meters with surge activity exceeds the threshold `WARD_MAX_METERS` = 5,000. A hybrid approximation is applied in this case: a random sample of `WARD_MAX_METERS` meters is drawn with a fixed seed, the full dendrogram is computed on the sample, and k cluster centroids are derived by cutting the dendrogram. Remaining meters are then assigned to the nearest centroid by Euclidean distance:

```
centroids = [
    data_sample[sample_labels == c].mean(axis=0)
    for c in range(1, k + 1)
]
dists_to_c = cdist(all_data, centroids, metric="euclidean")
all_labels  = dists_to_c.argmin(axis=1) + 1
```

This hybrid strategy reduces the effective complexity of the assignment phase to $\mathcal{O}(n \times k \times D)$, where D is the number of calendar days in the analysis period, making the approach scalable to datasets substantially larger than the full-dendogram limit.

DBSCAN on the other hand is implemented using a brute-force nearest neighbour search with a cosine distance metric, resulting in a time complexity of $\mathcal{O}(N^2)$ rather than the $\mathcal{O}(N \log N)$ achievable with tree-based indexing structures [16]. This approach was selected following initial testing on the regional smart meter datasets, where the brute-force pairwise distance table demonstrated faster practical performance at the relevant dataset scales.

Compared to partition methods such as k-means [53], DBSCAN is particularly well-suited to identifying noise and arbitrarily shaped clusters, making it advantageous for detecting low-density or irregular groupings that deviate from the dominant consumption patterns [16]. Accordingly, the DBSCAN clustering stage is oriented towards surfacing smaller outlier subgroups rather than capturing the broad structural segmentation of the load profiles.

4.2.7 K-Means Clustering

In addition to hierarchical clustering, centroid-based clustering was evaluated using K-Means. Whereas Wards linkage constructs a hierarchical dendrogram through iterative variance-minimising merges [47], K-Means partitions the behavioral feature space into a fixed number of non-overlapping clusters by minimising within-cluster distances to cluster centroids [48].

The algorithm operates on the same surge-count representation defined in Section 4.2.6, where each meter is represented as a fixed-dimensional vector of daily surge activity.

The K-Means objective function is defined as:

$$\min_{C} \sum_{i=1}^{k} \sum_{x \in C_i} \|x - \mu_i\|^2 \tag{4.4}$$

where C_i denotes cluster i and μ_i its centroid.

Unlike Wards linkage, which produces a hierarchical structure, K-Means directly optimizes a flat partition of the feature space [53]. This substantially improves computational scalability, particularly for larger datasets.

MiniBatch approximation.

For datasets exceeding 3,000 meters with surge activity, MiniBatchKMeans was used instead of standard K-Means. Mini-batch optimization reduces memory consumption and computational cost by updating centroids using randomly sampled subsets of the data rather than the full dataset during each iteration [48].

The resulting complexity is approximately:

$$\mathcal{O}(b \cdot k \cdot i \cdot T) \tag{4.5}$$

where b denotes the mini-batch size, k the number of clusters, T the number of time points, and i the number of optimization iterations.

Compared to full K-Means, the MiniBatch approximation slightly increases within-cluster variance but enables scalable clustering under operational constraints associated with near-real-time processing.

4.2.8 Comparative Evaluation of Clustering Methods

The clustering paradigms evaluated in this study represent different trade-offs between scalability, interpretability, and structural flexibility.

Wards hierarchical clustering produced compact and interpretable behavioral groups while preserving the hierarchical structure of meter relationships through the dendrogram representation. This improved interpretability and enabled data-driven selection of cluster granularity. However, the quadratic scaling of dendrogram construction limited computational feasibility for very large datasets.

K-Means substantially improved scalability and execution speed, particularly when combined with MiniBatch optimization. The method produced relatively balanced cluster sizes and performed efficiently under distributed operational constraints. However, the centroid-based formulation assumes approximately spherical cluster geometry and is more sensitive to centroid initialization and outliers.

DBSCAN provided the greatest structural flexibility by identifying arbitrarily shaped density regions and explicitly isolating outlier meters. Nevertheless, parameter sensitivity and unstable cluster formation reduced operational consistency across geographic regions.

Overall, Wards linkage was selected as the primary clustering framework due to its balance between interpretability, cluster stability, and operational relevance. K-Means was retained as a computational benchmark and scalability reference, while DBSCAN served primarily as a robustness and outlier-detection comparison.

4.2.9 Computational Complexity and Pipeline Bottlenecks

While most stages of the pipeline scale linearly with the volume of meter observations, $\mathcal{O}(N \cdot T)$, practical performance is primarily influenced by data movement and aggregation operations within the distributed Spark environment. Data ingestion, hourly aggregation, daily meter aggregation, and surge detection are all dominated by distributed scans, joins, and group-by operations rather than computational complexity alone.

The most computationally expensive analytical components are the baseline consumption models. ARIMA model selection scales as $\mathcal{O}(T \log T)$ due to the grid search over candidate specifications, while SARIMAX scales approximately as $\mathcal{O}(T^2)$ because of state-space estimation. These models are therefore executed as offline batch analyses rather than as components of a real-time operational workflow.

The primary scalability constraint of the clustering stage is Ward's hierarchical clustering, which exhibits $\mathcal{O}(n^2)$ time complexity and $\mathcal{O}(n^2)$ memory complexity. For larger populations of surge-active meters, the hybrid approximation described in Section 4.2.6 is used to bound computational cost while preserving the interpretability advantages of hierarchical clustering. Alternative methods such as DBSCAN ($\mathcal{O}(n^2)$) and MiniBatch K-Means ($\mathcal{O}(b \cdot k \cdot i \cdot D)$) provide computationally lighter benchmarks but were not selected as the primary clustering framework.

Overall, computational performance is governed by three factors: distributed aggregation over high-frequency consumption data, offline time-series model estimation, and hierarchical clustering of surge-active meters.

4.3 Qualitative Study

This section presents the methodological decisions made to examine how flexibility is understood, organized, and developed around a DSO and its wider flexibility market ecosystem.

4.3.1 Methodological Approach

The managerial part of the study was based on a qualitative research approach. The approach is appropriate for investigating a real world problem with multiple factors such as technical, organizational, regulatory and market related factors [49]. The research design also allowed perspectives outside the organization to be investigated, making this approach suitable for incorporating insights from flexibility providers and enablers from the wider system relevant to a DSO. The design also allowed the study to capture and reflect on the perception and experiences of stakeholders rather than solely on quantitative comparisons, allowing for a deeper understanding of how the surrounding environment affects decisions around flexibility.

The study was designed as a case study focusing on Göteborg Energi's strategic engagement within the flexibility market. The case study method is an effective method to generate deep contextual insights [54]. Although case studies may be limited in terms of external validity and generalization, they allow for a thorough examination of a specific context and its challenges [49]. The chosen case gave important context of the activities and challenges a prominent actor in the flexibility space is faced with, in an industrial city under electrification.

To further structure the case study, an *abductive approach* was deployed, which is suitable to match and guide evidence for further exploration [55]. The abductive approach consists of *systematically combining*, confronting empirical evidence with theory, and letting the match of data direct or redirect further research, and enables the researcher to triangulate and reveal evidence not found otherwise. The systematic combining also supported the iterative development of our qualitative and quantitative research. The findings from the respective research aspect deepened our understanding of specific emerging themes found, which led to further investigation.

4.3.2 Interview Study

To capture stakeholder perspectives in the qualitative research design, an interview study was conducted. The interviews were semi-structured to allow the different perceptions, perspectives and emerging themes of the respondents to be elaborated on to provide more detailed information [56]. The initial interviews were exploratory to provide early insights to inform and guide the abductive approach in iteratively designing the theoretical foundation and focus of the study [55]. The exploratory interviews were designed with broader questions to gain initial insight into the organization and its possibilities and related stakeholders.

The semi-structured interviews were designed to capture evidence to the central managerial concepts from the ecosystem lens. However, the interviews were also adapted to fit the interviewee's expertise and context, which meant that internal and external interviews were different in conceptual theme and foundation. Whilst the internal interviews at the DSO were focused on the ecosystem perspectives from a dynamic capability point of view, the external stakeholders were examined through an ecosystem perspective more focused on their challenges and enablers in adopting and contributing through flexibility resources. The two interview foundations were thus developed in parallel to one another, enabling the findings from both sides to improve and guide relevant themes to be discussed. New questions were also added and removed before and during interviews where appropriate to increase relevance and exchange of information [56].

The interview participants were chosen by a *snowball sampling* method [49]. This was performed by letting the first exploratory interviews with internal flexibility market actors at the DSO set the frame for topic relevant stakeholders. This was then expanded upon if additional stakeholders emerged in the study, so that the ecosystem actors could be effectively mapped. However, due to the scope of the study in terms of time, this constraint also affected the sample size. The collected data were further validated by letting the respondents review their contribution and, where necessary, allow them to refine their statements in case they were unsure at the time of the interview.

The interview study resulted in 14 interviews that were conducted continuously during the data gathering. The interviews were either conducted physically or online due to convenience of the participants and all interviews were recorded on a separate device and transcribed with consent. The interview's duration was approximately an hour but could vary due to schedule and relevance and exhaustion of topic and some interviews were conducted with two or more respondents at the same time due to schedule and availability. In the majority of the interviews, the full research team was present to increase the exchange between the two disciplines of the study. Where relevant questions for the quantitative study emerged during the interview, room was allowed in the interview session to help guide the interdisciplinary area of the study.

Respondent	Organization	Role/Area	Time
R1	DSO	Business Relations	1h
R2	DSO	Technical Expert	30 min
R3	DSO	Department Manager	1h
B1	External	CEO Aggregator	1h
B2	External	Technical Expert Aggregator	1h
B3	External	Technical Expert	1h
B4	External	Technical Expert	1h
B5	External	Business Development	1h
R4	DSO	Department Manager	30 min
R5	DSO	Business Development	1h
B6	External	Strategy Manager	1h
R6	DSO	Strategy Manager	1h
B7	External	BESS product manager	30 min
E1	External	Expert from Academia	1h
R7	DSO	Flexibility Manager	1h
B8	External	Regulatory Expert	1h
B9	External	Regulatory Expert	1h
B10	External	Regulatory Expert	1h

Table 4.2: Overview of interview respondents

4.3.3 Qualitative Data Analysis

The gathered data was analyzed inductively through systematically coding and conceptualizing the empirical findings [49]. The respondents experiences and perceptions were reviewed and coded into first order concepts reflecting the empirical data as transparently as possible. These concepts were then compared and grouped into broader theoretical second order themes. Subsequently, these themes were aggregated into thematic dimensions, forming the structure under which narrative the empirical findings and discussion was created.

In the secondary data collection scientific articles were used to support the abductive approach in guiding the findings [49], [55]. The analysis involved systematically comparing empirical findings with previous studies to form conceptual themes, and to establish an understanding of the empirical data supported by previously studied organizations [49]. Literature was chosen based on their comparative relevance to the empirical evidence found and therefore iteratively developed during the data collection.

5

Risk Analysis and Ethical Considerations

In this section, the risks of handling data and ethical considerations of the research process are addressed.

5.1 Handling of Smart Meter Data

Smart meter data, consumption, production, geographical location, weather, time, is of essence to the thesis, the generalization and quality of insights that can be deduced. Each data point represents one perspective of a specific facility and their operations, and the aim of the thesis is not to evaluate or put value on or harm individual actors, but rather what findings the research of aggregated and clustered data points can produce for future action [49]. Thus, careful considerations must be made not to point out specific actors and to give an objective view of what can be done with data and not to put specific data points and the actor behind the data in a negative light.

5.2 Handling of Interview Data

The internal actors of Goteborgs Energi and the external stakeholders that are subject of being interviewed for the thesis are an important factor in the result and process of the thesis. Therefore, the integrity of stakeholders, anonymity, confidentiality, and the use of data and informed consent is of the essence for the research study [49]. To maintain the integrity of the interviewees and the organization, they receive in good time information about the purpose of the study to make an informed decision if they will participate. Consent also includes being informed about how the interview is recorded, where the data will be placed, and for how long the data will be kept before removal. The interviewee were also be able to review their contribution before publication.

6

Evaluation and Results

This section first provides an aggregate overview of the identified flexibility potential across all regions and clusters, establishing a baseline for the scale and distribution of surge-related behaviour. It then examines the extent to which this potential can be concentrated through targeted selection of high-activity clusters, with a focus on the trade-off between coverage and meter population. Following this, the exploratory clustering approaches are evaluated, comparing Wards linkage, DBSCAN, and k-means in terms of cluster structure, stability, and suitability for identifying operationally meaningful groupings within the data. The analysis then shifts to a regional breakdown, where cluster-level characteristics are interpreted in context and related back to insights from the managerial interviews, particularly concerning practical applicability, operational constraints, and the relative importance of behavioral versus geographical segmentation.

6.1 Clustering Benchmark and Method Selection

To identify recurring flexibility behaviors among smart meters, multiple clustering approaches were evaluated and benchmarked.

The objective was not solely to maximize cluster separation, but rather to produce operationally actionable groupings capable of supporting flexibility analysis and demand-side management strategies.

Three clustering approaches were evaluated:

1. Ward hierarchical clustering
2. K-Means clustering
3. DBSCAN using both Euclidean and cosine distance metrics

The benchmark compared computational performance, cluster structure, scalability, and behavioral interpretability.

6.2 Data Pipeline

The computational performance is evaluated on a dataset covering the Central Gothenburg distribution region. The dataset consists of 37,049 smart meters observed at 15-minute resolution over a one-year period, corresponding to approximately 4 × 15-minute intervals per day. Of these, 21,638 meters exhibit at least one detected surge event and are therefore included in the downstream behavioral analysis.

Table 6.1 reports execution times for each major stage of the data pipeline. The results reflect a full end-to-end batch execution including time-series modeling components as well as the operationally relevant continuous processing path.

To distinguish between offline analytical modeling and deployable near-real-time analytics, two total runtime configurations are reported. The first includes all components of the pipeline, including ARIMA grid search and SARIMAX estimation. The second excludes these components since that for continuous analytics these larger cyclic models aren't needed for the core of flexibility analysis, but rather to examine the larger aggregate relationships between the data.

Table 6.1: Execution time of data pipeline components with starting input of all the 15 min entries for all the smart meters for the given time period of 01-02-2025 to 01-02-2026, ran on a python jupyter notebook environment (Central Gothenburg smart meter set).

Pipeline stage	**Runtime (s)**
Data loading & ingestion	19.3
Hourly aggregation	580.6
ARIMA (grid search + fit)	99.2
SARIMAX	584.9
Daily meter aggregation	50.5
Surge flagging	115.4
Total (full pipeline)	**1450.0 s**
Total (continuous analytics, excluding ARIMA + SARIMAX)	**765.9 s**

The results indicate that the dominant computational cost is concentrated in time-series modeling, particularly SARIMAX estimation and ARIMA grid search. When these components are excluded, the remaining pipeline exhibits substantially lower runtime and is dominated by aggregation and surge detection stages.

From an operational perspective, this separation is critical. The full pipeline is appropriate for offline analytical model calibration and structural consumption analysis, whereas the reduced configuration represents a feasible baseline for continuous or near-real-time analytics under 15-minute data ingestion constraints.

6.3 Clustering Techniques

Benchmarking of clustering techniques within the pipeline described in Section 6.1, ran within a python jupyter notebook environment.

Table 6.2: Clustering benchmark comparison on Central Gothenburg Region

Method	Time Complexity	Runtime
Ward Linkage	$\mathcal{O}(n^2)$	139 s
K-Means	$\mathcal{O}(nkt)$	50 s
DBSCAN (Cosine)	$O(n^2)$	49 s

6.4 Cross-Regional Overview

This section presents the comparative baseline across all study regions, including data coverage, aggregate load structure, and regional differences in temporal peak behavior with clusters created through Ward's method, to conduct these analytics on larger tangible, and balanced clusters to bundle and reason about the available flexibility according to the focus areas denoted in the interviews.

Table 6.3: Surge statistics by region and clusters derived from Ward's linkage

Region / Cluster	Meters	Meters >1 Surge	Surges	%-Surges	Excess [kWh]	Power [MW]
Total (all regions)	196122	135608	15083159	0.32%	97167249	11,09
Targeted (No C1s)	37597	37597	13073698	0,99%	90175666	10,29
South-East (Total)	31745	19652	1704261	0.25%	12331420.5	1.41
C1	15260	15260	197979	0.04%	702388.9	0.08
C2	1522	1522	254137	0.48%	1804381.1	0.21
C3	1188	1188	404226	0.97%	2958865.2	0.34
C4	1019	1019	576944	1.62%	3695984.2	0.42
C5	663	663	270975	1.17%	3169801.1	0.36
North (Total)	7197	6146	1024551	0.48%	5561948.3	0.63
C1	3061	3061	85544	0.08%	318586.3	0.04
C2	1738	1738	326783	0.54%	1337981.3	0.15
C3	1347	1347	612224	1.30%	3905380.7	0.45
West (Total)	16040	13843	2623028	0.54%	15396204.6	1.76
C1	6931	6931	238499	0.10%	837003.8	0.10
C2	3560	3560	777180	0.62%	3708231.1	0.42
C3	3352	3352	1607349	1.37%	10850969.7	1.24
North-East (Total)	17552	14666	1843537	0.36%	10994662.2	1.26
C1	10549	10549	376067	0.10%	1224224.1	0.14
C2	2445	2445	625412	0.73%	3225278.4	0.37
C3	1672	1672	842058	1.44%	6545159.7	0.75
East (Total)	12557	9445	1007369	0.30%	5886114.6	0.67
C1	6114	6114	92172	0.04%	258940.8	0.03
C2	1443	1443	147135	0.29%	512777.3	0.06
C3	892	892	251118	0.80%	1358115.5	0.16
C4	845	845	444093	1.50%	2746180.1	0.31
C5	151	151	72851	1.38%	1010100.9	0.12
South (Total)	73982	50218	5367645	0.31%	35774072.2	4.08
C1	37482	37482	764575	0.06%	2490509.2	0.28
C2	7055	7055	1739114	0.70%	10227732.2	1.17
C3	5681	5681	2863956	1.44%	23055830.8	2.63
Central (Total)	37049	21638	1512768	0.20%	11222826.2	1.28
C1	18614	18614	254625	0.04%	1159929.8	0.13
C2	1221	1221	686103	1.60%	4102528.7	0.47
C3	975	975	274515	0.80%	2120676.1	0.24
C4	828	828	297525	1.03%	3839691.6	0.44

Table 6.3 presents the aggregated surge statistics for each geographical region and their corresponding clusters. The column *Meters* refers to the subset of all available

smart meters belonging to a given geographical area. Each region was analyzed independently before clustering was applied to the corresponding consumption profiles.

The column *Meters >1 Surge* represents the number of meters for which at least one surge event was detected during the analyzed period. As part of the preprocessing pipeline, all meters without any detected surge occurrences were removed from the surge analysis in order to focus exclusively on meters exhibiting flexibility-related behavior.

The column *Surges* represents the absolute number of detected surge events occurring between 1 February 2025 and 1 February 2026. A surge event corresponds to a consumption value significantly exceeding the local baseline behavior as determined by the surge detection methodology.

The metric *%-Surges* represents the proportion of all possible measurement events classified as surges. This value is computed as:

$$\text{\%-Surges} = \frac{\text{Surges}}{\text{Meters} \times 35040}$$

where 35040 corresponds to the total number of 15-minute measurement intervals within one year:

$$4 \times 24 \times 365 = 35040$$

Consequently, the metric describes how frequently surge behavior occurs relative to the total number of possible observations for the analyzed meter population.

The column *Excess [kWh]* represents the cumulative excess energy consumption above the calculated 7-day moving average baseline. More specifically, for each detected surge event, the difference between the observed consumption peak and the corresponding moving average baseline was accumulated. This value was interpreted as a proxy for potentially flexible or shiftable load, since the excess demand may represent consumption that could theoretically be temporally redistributed.

Finally, the metric *Power [per hour] [MW]* estimates the average hourly flexible power available to a flexibility market operating on 1-hour trading windows. The value is calculated by distributing the total yearly excess energy across all hours of the year:

$$\text{Power [MW]} = \frac{\text{Excess [kWh]}}{365 \times 24 \times 1000}$$

It should be noted that no temporal weighting or simultaneity analysis has yet been applied. The calculation therefore assumes that all detected excess consumption contributes equally regardless of when the surge occurs. Temporal alignment, simultaneity effects, and cluster-specific behavioral patterns will be further analyzed in the subsequent individual cluster analyses.

To further evaluate the practical implications of the clustering approach, an aggregated comparison was performed between the complete set of detected surge meters and a targeted subset excluding the least active clusters (all C1 clusters across regions). The results are presented in Table 6.4.

Table 6.4: Comparison between total and targeted flexibility potential

Scenario	**Meters >1 Surge**	**Excess [kWh]**	**Power [MW]**
Total (all regions)	135608	97167249	11.09
Targeted (No C1s)	37597	90175666	10.29

The results demonstrate a highly concentrated distribution of flexibility potential across the analyzed meter population. By excluding the least active clusters (C1), approximately 92.7% of the total detected flexibility potential remains, despite only targeting 27.7% of the meters exhibiting surge behavior.

This finding indicates that a relatively small subset of highly active meters accounts for the overwhelming majority of the theoretically flexible load. Consequently, large-scale flexibility programmes may not necessarily require broad participation across the entire smart meter population. Instead, significantly higher efficiency may be achieved through targeted engagement strategies focused on the most surge-intensive behavioral clusters.

The comparison also highlights a substantial increase in flexibility density. For the complete population, the average hourly flexible power available per surge-active meter was approximately:

$$\frac{11.09 \text{ MW}}{135608} \approx 0.08 \text{ kW per meter}$$

whereas for the targeted subset excluding all C1 clusters, the corresponding value increased to:

$$\frac{10.29 \text{ MW}}{37597} \approx 0.27 \text{ kW per meter}$$

However, when analyzing the average hourly flexibility contribution only across the actively targeted surge windows, the effective average flexibility availability increases substantially. The complete dataset corresponds to an average tradable flexibility contribution of approximately 0.08 kW per participating surge event, while the targeted subset reaches approximately 0.27 kW per participating surge event.

This further reinforces the conclusion that the clustering methodology successfully isolates behaviorally distinct groups with substantially different flexibility characteristics. Depending on the operational objectives of the flexibility market, different cluster combinations may therefore be prioritized. Certain clusters exhibit highly concentrated but infrequent surges, while others display more persistent and sustained excess consumption patterns, potentially making them more suitable for specific demand response strategies, trading horizons, or grid balancing applications.

6.5 Regional Aggregated and De-aggregated Cluster Analysis

This section delves deep into the region of Gothenburg South East one of the 7 regions and examines the regions aggregate results both in terms of the autoregressive findings from ARIMA, the aggregate behaviors found when accounting for the external factors such as weather and electricity pricing from SARIMAX, the comparison between them regarding how the surge patterns differ. Meaning the first 5 figures show how all the meters in the Gothenburg South East Region behave together. The later figures showcases the breakdown of clusters emerging, their differences in surge behaviors and what this equates to in terms of power. Lastly how these clusters emerge and are represented geographically. The entire cluster breakdown is done with Ward's linkage to showcase the dendrogram and thus how similar the cluster are to eachother as well as Ward's innate feature of producing balanced cluster so that insights can be gained from groupings with a tangile amout of smart meters [47].

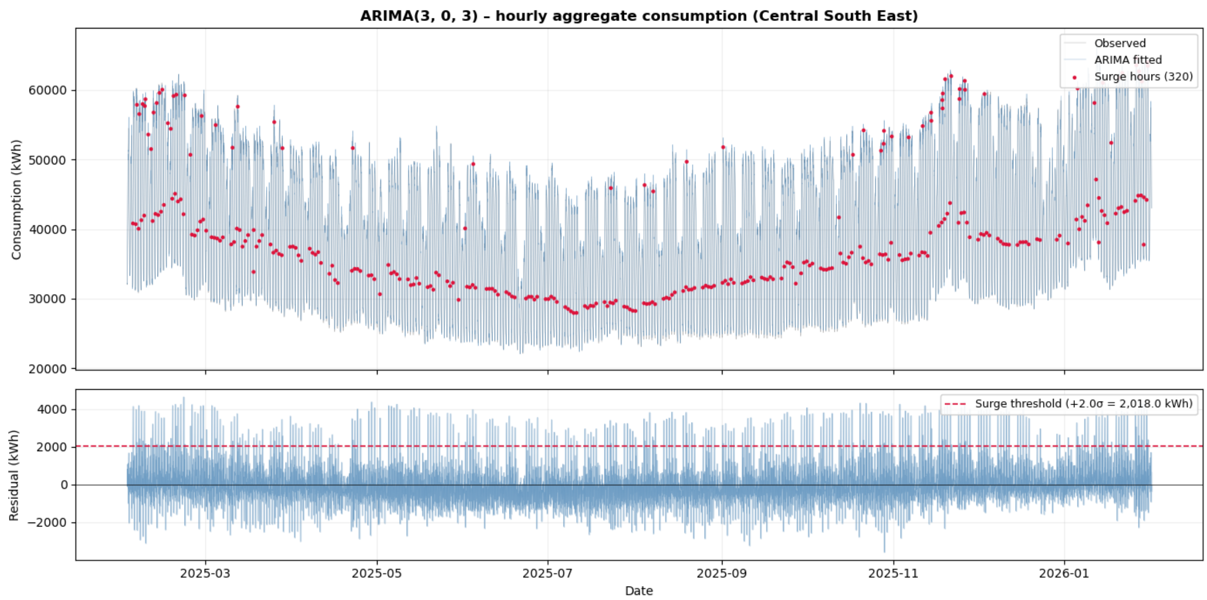

Figure 6.1: ARIMA

An ARIMA(p, d, q) model is fitted to the hourly aggregated time series. The specification is selected via grid search over $p \in \{0, 1, 2, 3\}$, $d \in \{0, 1\}$, and $q \in \{0, 1, 2, 3\}$, with model selection based on the Akaike Information Criterion (AIC). The resulting optimal configuration is ARIMA$(3, 0, 3)$, capturing short-term autoregressive and moving-average dependence without differencing.

For the ARIMA model, the residual series exhibits comparatively lower variance, reflecting that most short-term autocorrelation structure is captured by the autoregressive and moving-average terms, leaving primarily idiosyncratic noise.

It is observed that the model produces more frequent and structurally consistent surge patterns, with peaks occurring at relatively stable consumption levels across the time series, see Figure 6.1, whereas the SARIMAX model exhibits less regular and more dispersed surge behavior, see Figure 6.2.

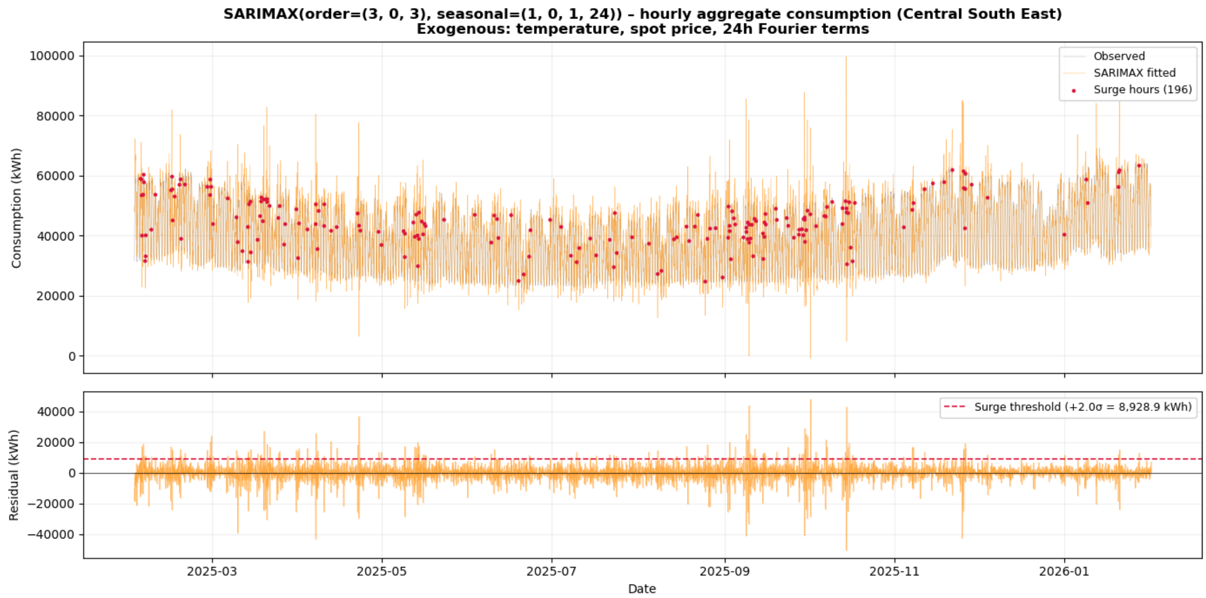

Figure 6.2: SARIMAX

SARIMAX$(p, d, q) \times (P, D, Q, s)$ model incorporating seasonality and exogenous variables. The seasonal component is fixed at $(P, D, Q, s) = (1, 0, 1, 24)$ to capture daily periodicity in the hourly series. The final specification is SARIMAX$(3, 0, 3) \times (1, 0, 1, 24)$ with exogenous regressors including temperature, electricity spot price, and cyclical time-of-day features.

For the SARIMAX model, the residuals display higher volatility, see Figure 6.2, which is primarily attributed to the inclusion of exogenous variables such as the electricity spot price, see Figure 6.4, whose pronounced short-term fluctuations are propagated into the error structure. Possibly due to the lag between consumption behavior, and the difference in response to the pricing signals leads to additional features instead of reducing variability on the aggregate level. This observation can be seen in the surges occuring in the aggregate SARIMAX model, they are fewer and occur in greater variance across the consumption band of the aggregated region.

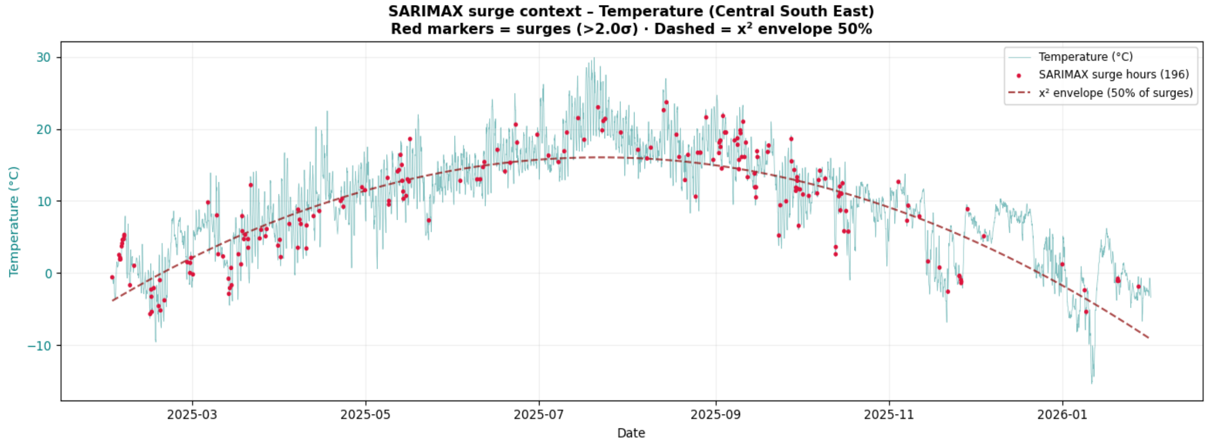

Figure 6.3: Surges displayed over weather data

The regions surge behavior related to the temperature changes occur mainly during the period where the weather starts to shift from cold to warm (01-02-2025 to 10-05-2025) and the return (01-09-2025 to 15-10-2025). With many of the aggregate surges appearing alongside the drops in temperature or the lowest temperature measurements, meanwhile almost no surgest occur at the very warmest measurements.

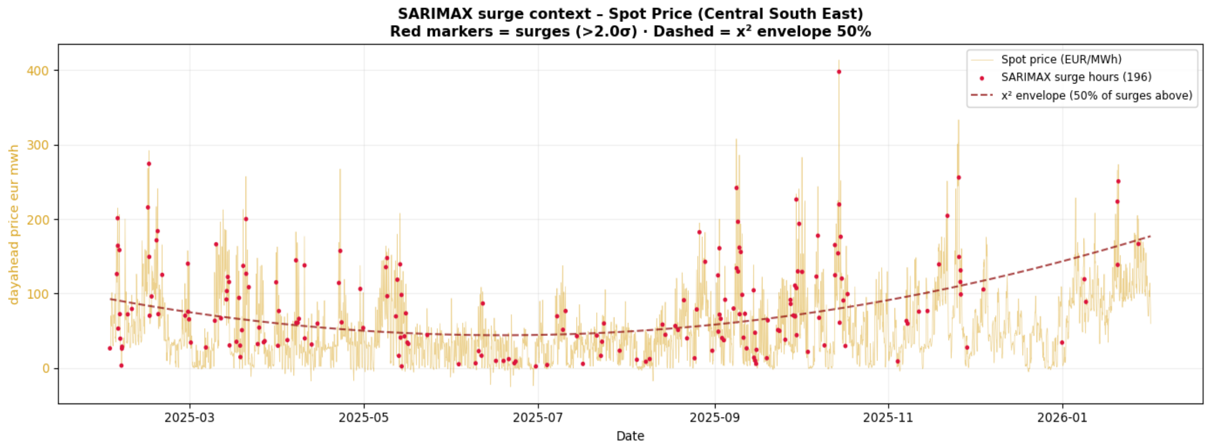

Figure 6.4: Surges displayed over EPEX SPOT price data

The regions surge dependency of the EPEX SPOT price is next to non-existent, with aggregate surges occurring regardless of pricing signals, the thing to note is that at each sudden spike in the electricity price, there often is a surge occurring. But since these prices are set ahead of time in the day-ahead market, these surges most likely relate to wider capacity constraints from weather shifts.

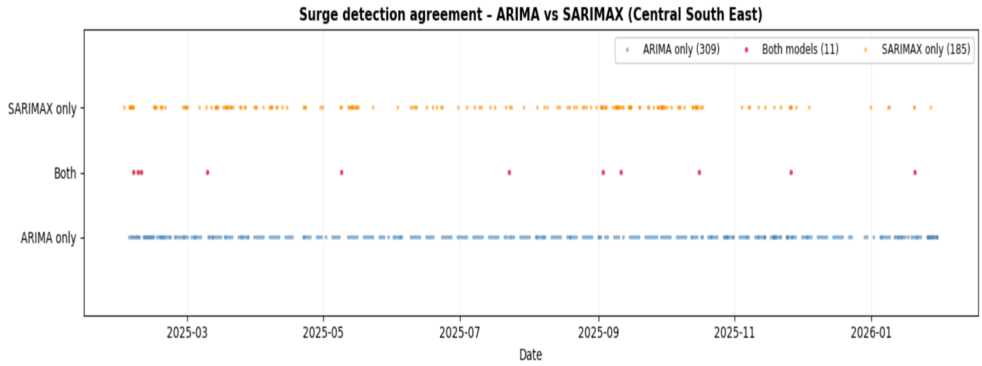

Figure 6.5: ARIMA v. SARIMAX Surge Occurrences

The stark difference in aggregate surge behavior becomes more apparent when comparing the two auto-regressive models, see Figure 6.5. With the consistent surge appearances of the ARIMA model evenly spread throughout the year and in comparison the SARIMAX surges occur mainly in relation to the spikes in electricity price occurring.

As seen by Figure 6.6 plenty of smart meters clustering into *C1* have next to no surge activity, even after excluding the meters with no activity. With *C2* still having little to no consistent peak patterns throughout the cluster, except for a faint week-cycle pattern visible, with two consolidated surges occurring late December 2025, which can be best seen in Figure 6.8. Continuing on cluster *C3* and looking to the dendogram

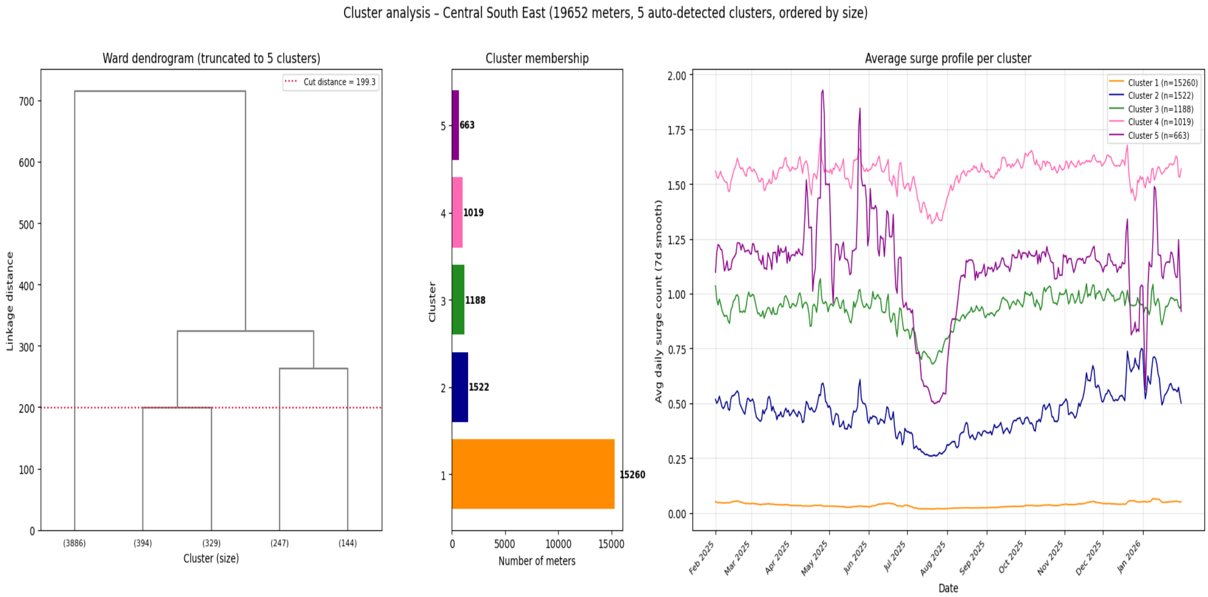

Figure 6.7: Cluster analytics, covering Ward's dendrogram of how closely related the clusters are, the amout of smart meters adherent to each cluster, and the average surge profile of the clusters

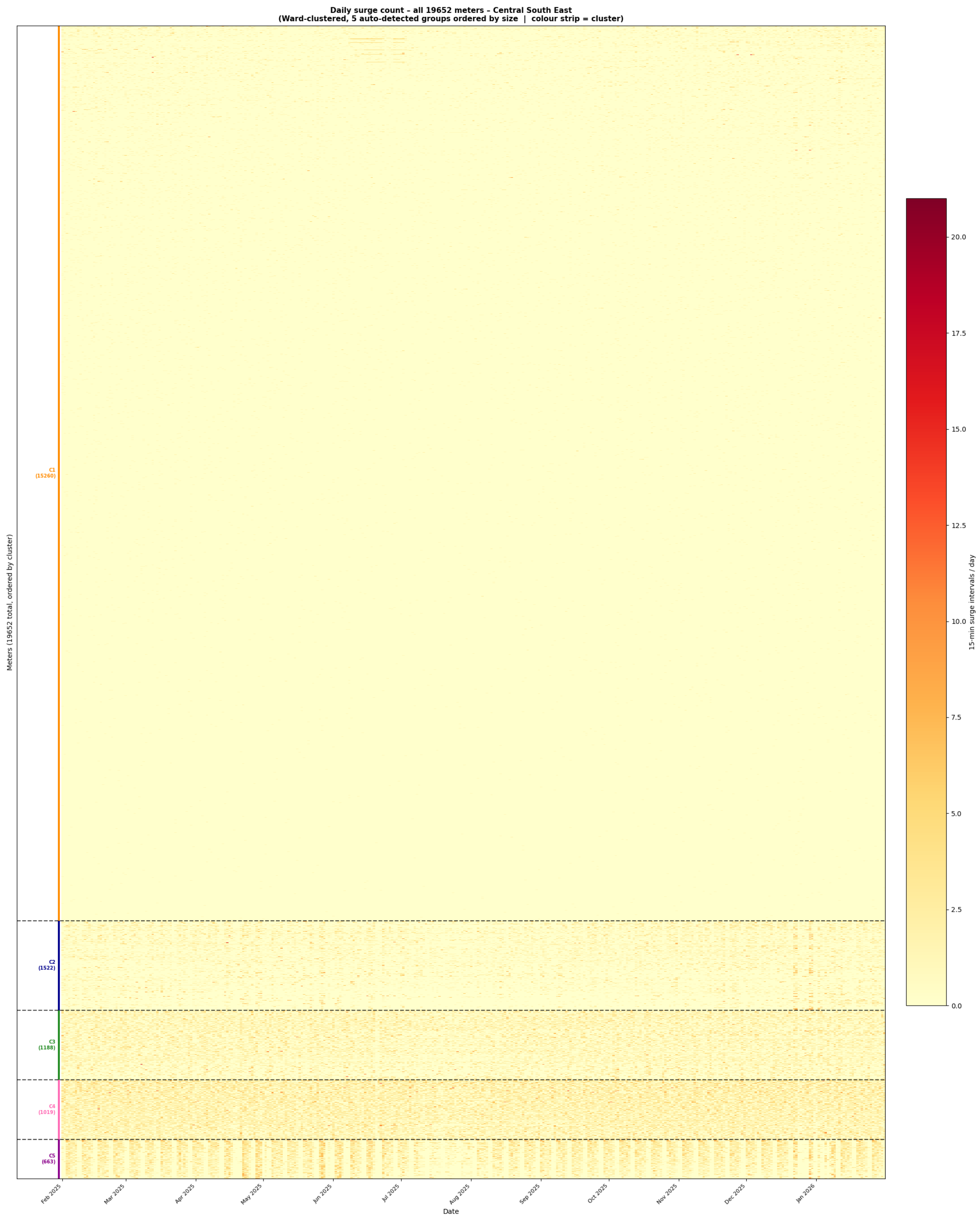

Figure 6.6: Heatmap showcasing cluster capture of meter surge data

Looking to the analysis of the clusters, starting with the "Average surge profile per Cluster", see Figure 6.7. It can be clearly observed that the most populated cluster *C1* is also the one with the lowest activity. The other 4 clusters follow a clear seasonal dip in surges occurring around the summer months of July and August, with varying effects of volatility. Noting that the clusters of individual smart meters experiences relatively small changes in surge occurrences during the colder month, while cluster *C3* and *4* showcases a similar muted behavioral pattern, in contrast to cluster *C5*, with its extremely volitive profile which was examined further both geographically and in terms of available flexibility [kWh], see Figure 6.9.

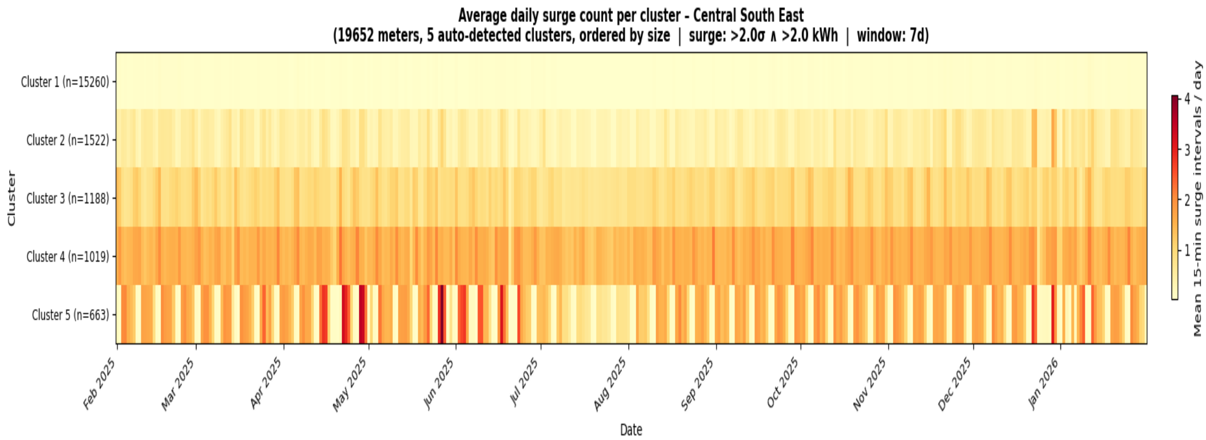

Figure 6.8: Cluster heatmap showcasing surge occurrence

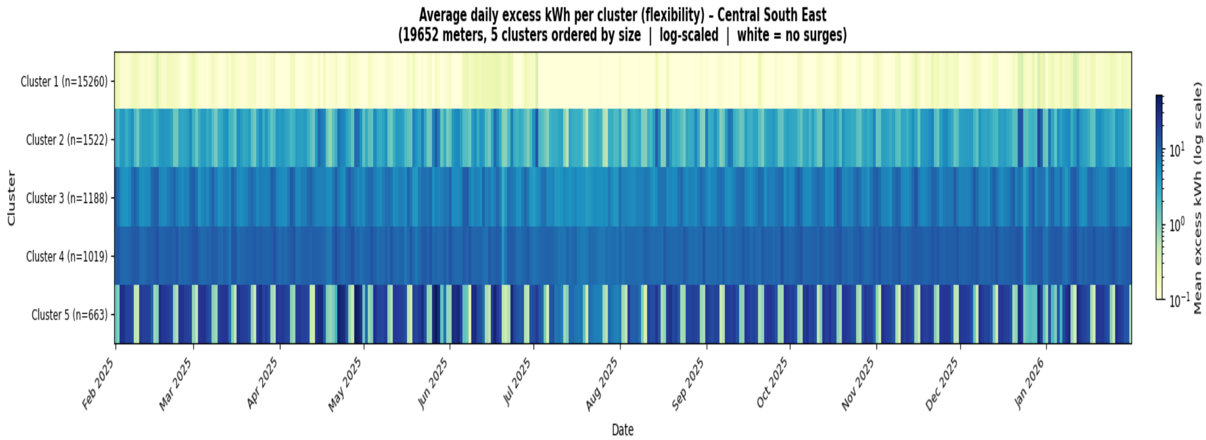

Figure 6.9: Cluster heatmap showcasing excess kWh from surges

Through deeper analysis of geographical distribution of smart meters and their adherent clusters *(not included here to due privacy concerns)*. It can be observed that the clusters *C1*, *C3* and *C4* with the more muted and stable surge profile are mostly found in residential areas, meanwhile the volatile cluster *C5* with clear week-day / week-end cycles are almost exclusively found in industrial and shopping areas.

6.6 Qualitative Empirical Findings

This chapter presents the results aimed at answering RQ2, with empirical evidence from the interviews with internal DSO respondents, external flexibility providers and aggregators, an academic expert, and regulatory and policy respondents. As presented in figure 6.10, the empirical findings were analytically synthesized into first-order concepts, second-order themes and aggregate dimensions using inductive analysis [49].

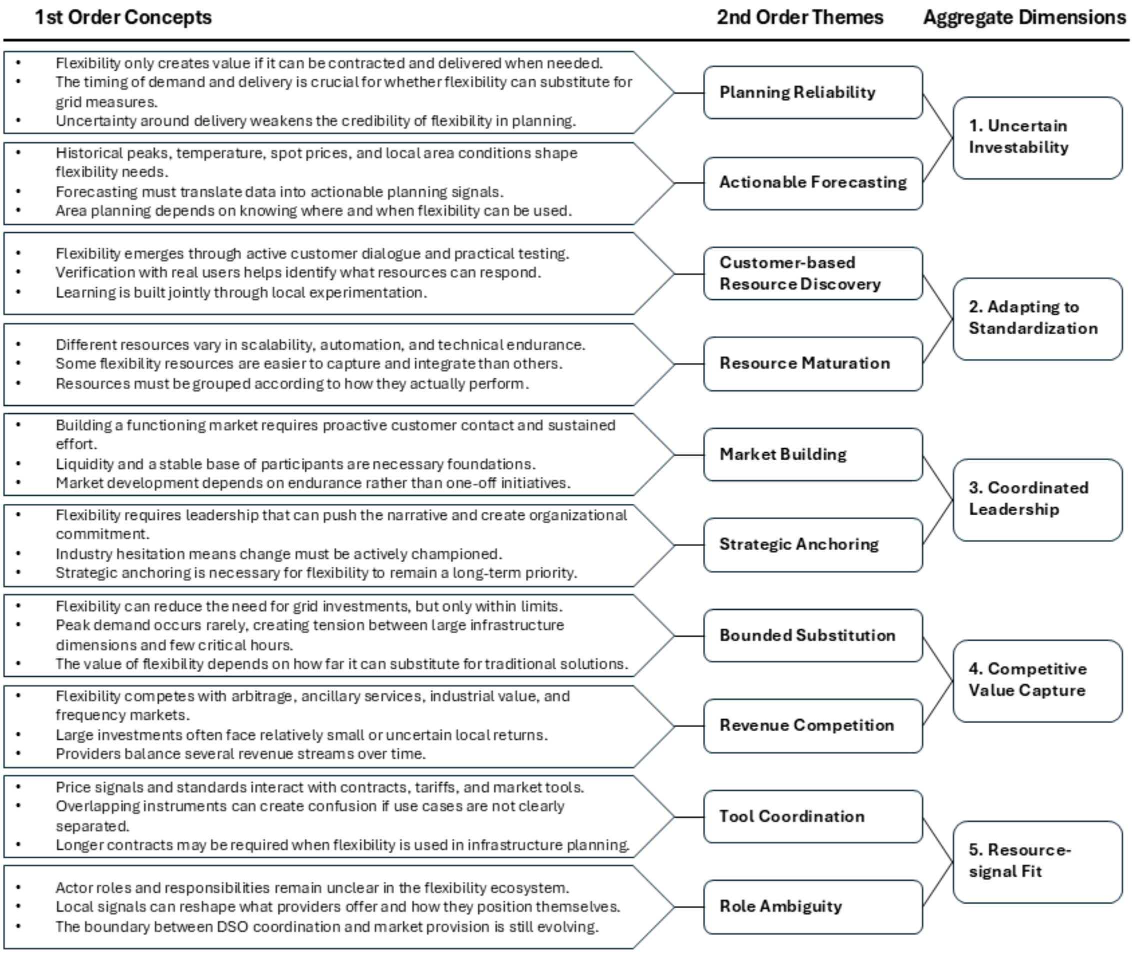

Figure 6.10: Overview of first order concepts, second order themes and aggregate dimensions

These aggregate dimensions are presented in logical sequence, first explaining how flexibility cannot be planned, procured, invested in, or regulated unless actors perceive it as reliable, visible, and controllable. Secondly, once uncertainty is acknowledged, the question becomes how actors learn what works and turn that learning into scalable structures, through experimentation and standardization. Thirdly, experimentation and standardization has shown to not happen automatically, someone must lead and coordinate market actors, customers, and regulators. Fourth, because even if flexibility is technically available and coordinated, providers will only supply if the business case is attractive. Fifth, the DSO tools, contracts, bidding, and contracts must be aligned with how different resources actually respond.

6.6.1 Information Uncertainty and Measurability Shape Planning Legitimacy and Investability

The internal and external interviews strongly agree that uncertainty is one of the central barriers to flexibility. Internal respondents worry about whether flexibility can be trusted as a system tool. External respondents worry about whether uncertainty can be reduced enough for customers, investors, and project developers to commit resources. Both perspectives point to the same conclusion, that without measurability and better visibility, flexibility remains difficult to scale.

When discussing what flexibility can do for a DSO, internal respondents broadly agreed that there is a large potential in flexibility, but it was repeatedly emphasized that the practical challenge lies in matching theoretical concepts with real operating conditions. As R4 emphasized,

> *The devil is in the details* [R4]

particularly when flexibility is expected to replace or delay grid investments. For flexibility to become usable in planning, respondents described that the relevant resource must be available at the right time, in the right place, and with certainty that the resource will be available to support actual decisions. R4 therefore framed flexibility not primarily as a concept problem, but as a problem of decision usefulness and operational precision. R4 further adds

> *If we look deeper into the grid, prognoses and forecasts become less important, in this world, it matters more if the customer chooses to build on that or that street, depending on what they are asking for* [R4]

illustrating the issue of substituting or delaying a grid investment with flexibility when it is operationalized in the real world.

Respondent R6 argued that a key long-term challenge is to enable

> *The traditional grid planner to see and trust a flexible resource instead of a network investment.* [R6]

indicating that flexibility is not only competing with infrastructure economically, but also institutionally, whilst adding that the market needs to work really well for this to be the case. Similar accounts were also made by R7 who exemplified the challenges

> *How do we get flexibility to become part of the infrastructure? How can we dare to connect customers thanks to the flexibility we have? There we still have some journey to make* [R7]

pointing toward uncertainty challenges in the flexibility technology, in terms of knowing that the flexibility will exist the coming years at a location. Their account suggests that these solutions would need more conditional or contracted flexibility rather than market based participation. R6 further relates this issue to their core operations, and how they in the past handled uncertainty in situations:

> *We have sometime in history learned that we don't have to have to fully dimension everything. It will probably be somewhat the same with this, gradually* [...] *but during the construction phase it will probably be a quite interesting dynamic between demand of agreements, belief in the market, and the where you are positioned in the development* [R6]

Forecasting was repeatedly described as central to this credibility by several respondents. R3 described short-term forecasts as critical both for trading on the local flexibility platform and for future dynamic tariff models, where price signals would reflect predicted congestion in the network. But as R5 remarked

> *The current forecasting tool still has limitations* [R5]

emphasizing that data accuracy could impact timing of flexibility demand. In their account, the anticipated peak can move, disappear, or emerge elsewhere between the forecast and the actual delivery hour. This makes activation at two days earlier inherently uncertain and highlights the importance of better forecasting and delivery if the market is to become more targeted and reliable.

R5 and R7 explained that one of the most important development areas lies not simply in collecting more data, but in improving the capability to convert multiple data types into actionable signals. As R5 remarked

> *We need to see how and where flexibility exist* [...] *to include spot prices into the prognosis* [...] *and temperature data relation* [...] *but also if a user in an area has potential for flexibility* [R5].

highlighting the different types of data needed but also how it ultimately needs to be translated into something useful like identifying potential users. R7 also added that forecasts must account for how flexibility is traded, load patterns, network switching that can move loads between areas, significant grid users, large batteries, and production plans.

Several respondents also emphasized that measurability concerns not only prediction but also delivery quality. As R5 explained, market maturity should not be understood only in terms of contracted capacity, but in terms of reliability and endurance. The challenge is therefore not just to onboard megawatts but to know how much flexibility can actually be delivered, for how long, and under what conditions. R3 described a similar insight after the winter season, where the organization had learned that some customers were unable or unwilling to reduce load precisely when system stress was highest. In their words, some customers effectively said

> [...] *when it is cold, then I cannot be flexible* [R3]

even though that is when the peak emerges.

R1's interview contributed a more practical and relational perspective to this dimension. Rather than describing measurability as an abstract system property, the respondent illustrated how flexibility is often made measurable through joint testing and interpretation. Historical data are used to identify unusual spikes, customers are asked what happens operationally at those times, and then specific machines or processes are tested to determine what actually can be shifted. In that sense, measurability is not fully given beforehand, but constructed through interaction between grid knowledge, meter data, and customer understanding.

From the external side, the same pattern appears, but with a more explicit investment logic. B4 explains that local flexibility demand is difficult to include in a battery business case because the need is uncertain in the short term, while the investment horizon is long:

> *It is very hard to know how many megawatts the local DSO will need the next 2 to 3 years. But when we build a case for a battery it is projected for 15, 20, 25 years, so it is very hard to forecast already on a national level.* [B4]

B6 expresses the same challenge in more financial terms. The respondent argues that external investment depends on being able to reduce uncertainty and quantify the expected demand:

> *If you can at least say that for ten years it will be this much and this much, it becomes much easier. We also need to bring in capital. It is not a bank that invests. It will have to be some venture capital* [...] *that way it becomes a much more expensive project.* [...] *We have to minimize the uncertainty on our side so that the projects become investable.* [...] *We need some indication. It does not have to be perfect numbers, but there has to be something that simplifies the calculation, so that one can quantify the risk or quantify the loss in opportunity cost.* [B6]

B6's account suggests that flexibility demand must be expressed in quantifiable terms of how much, how often, and for how long, if external actors are expected to invest against it. B5 exemplifies this issue by describing how a lack of transparency can lead project developers to spend time and money on flexibility projects that are never implemented:

> *There is much less transparency in Sweden about where and how much flexibility is needed from local DSOs* [...] *so there are now many project developers who plan and spend a lot of time and money to bring projects forward but only a fraction gets implemented in the end* [B5]

B7 adds a more customer-facing version of the same issue. According to B7, rapid changes in market conditions make it difficult to provide customers with reliable financial forecasts:

> *Things are changing rapidly in the market, capacity, tariffs, electricity prices, and which services are most attractive at any given moment. One*

of our biggest challenges in the sales process is to provide a credible forecast for future financial performance. [B7]

B7s account suggests that even when customers understand what a battery does in principle, they still hesitate because future savings and revenues remain uncertain. B1 reinforces this by showing that uncertainty is also rooted in operational frictions such as incomplete data, lack of confidence in metering, and the cost of installing energy metering:

Half of our costs go to installing energy metering [...] *this causes a lot of sales friction* [B1]

The account of B1 further adds that external actors often proceed with imperfect information simply because requesting more data can make them lose the customer.

E1 added that data sharing between DSO's and external actors is a broader system-level challenge. More data could improve coordination, but actors may regard commercial or operational data as sensitive. The issue is therefore not simply collecting more data, but deciding what data can be shared, how, and under what conditions. B8, B9 and B10 extended this issue by showing that measurability is also shaped by regulatory reporting and the still emerging methods used to describe flexibility needs. B10 emphasized that the first reporting plans and FNA-related assessments showed differences in how network companies interpreted and assessed their flexibility needs. This suggests that flexibility is not only difficult to measure operationally, but also difficult to translate into comparable planning categories across DSO's. It was remarked as maturing and expected to improve as companies get more accustomed to the methods. B9 further anchored the FNA and reporting context, where flexibility needs and inputs can help develop a national understanding of flexibility demand.

Taken together, the results show that uncertainty limits flexibility both as a planning instrument and as an investment object. As summarized in 6.11, DSOs and flexibility providers share an interest in long-term commitment without excessive risk, but both depend on more targetable factors to reduce uncertainty. Internally, respondents describe flexibility as promising but difficult to rely on unless it can be forecasted, located, measured, and verified with sufficient precision. The central issue is therefore not whether flexibility has theoretical value, but whether grid planners can trust it enough to use it as an alternative or complement to traditional network investments. Externally, the same uncertainty appears as a barrier to investment and customer adoption. Project developers, aggregators, and customers need clearer indications of where flexibility is needed, how much demand can be expected, for how long, and under what conditions revenues or savings can be realized. The section therefore suggests that measurability and visibility are not merely technical improvements, but key conditions for making flexibility legitimate in grid planning and investable in the market.

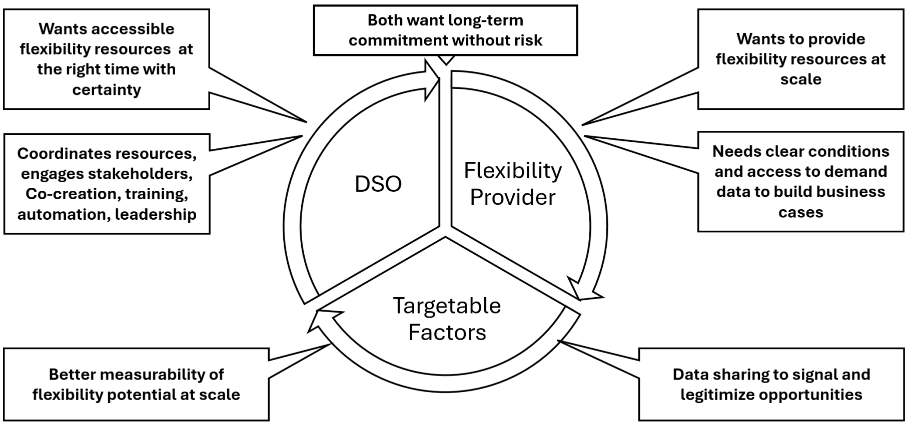

Figure 6.11: Overview of risks and commonalities in planning and providing flexibility

6.6.2 Flexibility Develops Through Experimentation and Standardization

A second finding is that flexibility is developed through a combination of experimentation and standardization. Respondents did not describe these as opposing logics. Rather, the material suggests that experimentation is needed to discover what kinds of resources, arrangements, and behaviors are actually viable, while standardization is needed to make the emerging solutions scalable and easier to coordinate across actors and contexts. External respondents largely agree, but add that experimentation becomes costly if it is not gradually translated into clearer and more repeatable structures.

This experimental logic was especially visible in the interviews with R6, R3, and R1. R6 described local flexibility as something that does not emerge spontaneously, but has to be created, recruited, and nurtured over time. As remarked by R6 regarding building the flexibility market

> *The market doesn't build itself. The timing might not even be right at this moment. Maybe in a few years, when the conditions are better, but we won't know whether that future will exist unless we work on it* [R6]

which highlights the need to be proactive in building value in flexibility. The argument was not that the market already works perfectly, but that it will never become useful unless actors invest in its development before all conditions are in place and have the endurance to do so.

Similar perspectives were also pointed out by other respondents, who described the local flexibility market not as a ready-made solution, but as a tool that is still being developed and tested. R3 framed previous work with flexibility as a proactive investment in future capability rather than as an immediate solution to existing capacity problems:

> *We have never in past runs procured flexibility in order to solve a capacity problem. Rather, the years we have spent working with Effekthandel Väst and flexibility should be seen as a proactive investment in a tool that we may need, much like proactively reinforcing the electricity grid when we know that a district will grow and that capacity will be needed later.* [R3]

The external interviews confirm this early-stage logic. B1 shows that providers are still learning what actually creates value in practice, for example by testing how tariff savings, spot-price response, and local flexibility revenues stack together. Respondent B6 adds their view on testing with the DSO

> *It is very beneficial for the DSO to test different solutions with an actor and we can quickly become guinea pigs, but only to a point.* [B6]

suggesting that experimentation is acceptable as long as it is commercially defensible but cannot continue indefinitely at the expense of returns.

R1s interview illustrated what such experimentation looks like in practice. The respondent described how flexibility often begins with curiosity, site visits, and trying out whether a machine, lighting system, charging pattern, or industrial process can be shifted. The respondents account showed that many resources do not enter the market through a standardized product logic from the outset, but through local discovery and testing. For instance, flexibility in tram depots or street lighting was identified by asking practical questions such as what runs in the morning, what can be dimmed, and what happens if something is delayed. Respondent E1 adds that the wider system is in an in-between phase where actors are still learning how to integrate storage and flexibility into electricity systems that were not designed around them.

Respondent R1 also highlighted some institutional differences regarding experimentation and standardization, explaining the need for action to progress:

> *Some want to have everything ready, they want the solution, have to test everything in secrecy before taking a small step, and that is not how the real world works* [R1]

At the same time, several respondents described the need for more standardization. R5 emphasized that standardized products for local flexibility markets have become an important development, particularly because they reduce complexity for flexibility service providers participating across multiple markets. Standardization, in their account lowers entry barriers, makes markets easier to understand, and helps new actors engage without facing a completely different structure in each local market.

Both sides further emphasize that continued scaling requires more standardization. R5 highlights the value of standardized local flexibility products, while B7 stresses that national guidelines for battery projects are still lacking and that scaling distributed batteries requires more harmonized approval and installation processes from the DSO. B2 and B3 reinforce this point from a project-development angle, describing battery deployment as partly dependent on permitting and approval processes:

> *It was a bit trial and error. It has changed a bit the last couple of years but we still see these problems, perhaps on larger facilities, to get the building permit* [B2]

R2s interview reinforces this transition from exploratory growth to more selective market maturity. While the respondent acknowledged the importance of building up a large market in terms of Megawatts, it was also argued that future value lies less in simply increasing the number of participants and more in attracting participants that are scalable, automated, and operationally reliable. This indicates that experimentation is gradually giving way to a more selective logic of market refinement, where not all resources have equal usefulness for the DSO.

R4s perspective further suggests that experimentation alone is insufficient if flexibility is to affect infrastructure choices. From a planning perspective, resources must eventually become robust enough to support more formal decision processes. Thus, experimentation is necessary to identify and validate possibilities, but standardization is necessary if flexibility is to become a stable part of system planning.

Another perspective was also highlighted by R7 who explained that standardized products are being developed partly in anticipation of EU regulation, with the ambition that similar products should be used across Swedish local flexibility markets. B8, B9 and B10 confirmed this from a regulatory perspective, describing the forthcoming EU network code for demand response and flexibility as an effort to remove barriers, clarify roles and harmonized participation across Europe. In B8's account, standardization is not introduced in one step. Rather, regulation moves through several layers, EU level framework guidelines, drafts by European transmission system operators and DSO organizations, review by regulatory authorities, negotiation by the Commission, and later national implementation. This means that standardization is both a legal process and a coordination process where European direction must translated into Swedish conditions.

B10 also presented regulatory sandboxes as a possible future tool that is under implementation. This tool is meant as a forum for actors to have a structured dialogue and interaction to test and experiment with new solutions and possibly but not mainly with regulatory schemes other than currently in place. B8 similarly described their innovation center as ways of providing regulatory guidance without issuing binding advance decisions. The combined regulatory perspective suggests that flexibility develops through both standardizing regulatory initiatives but also benefit from testing and experimenting jointly with the dependent actors to come with solutions.

The combined finding is that local flexibility is moving from a pilot logic toward a more mature market logic. However, as shown in 6.12, this creates tension, while more standardization can increase reliability and scalability, it may also reduce the openness that enabled early experimentation and smaller actors to participate.

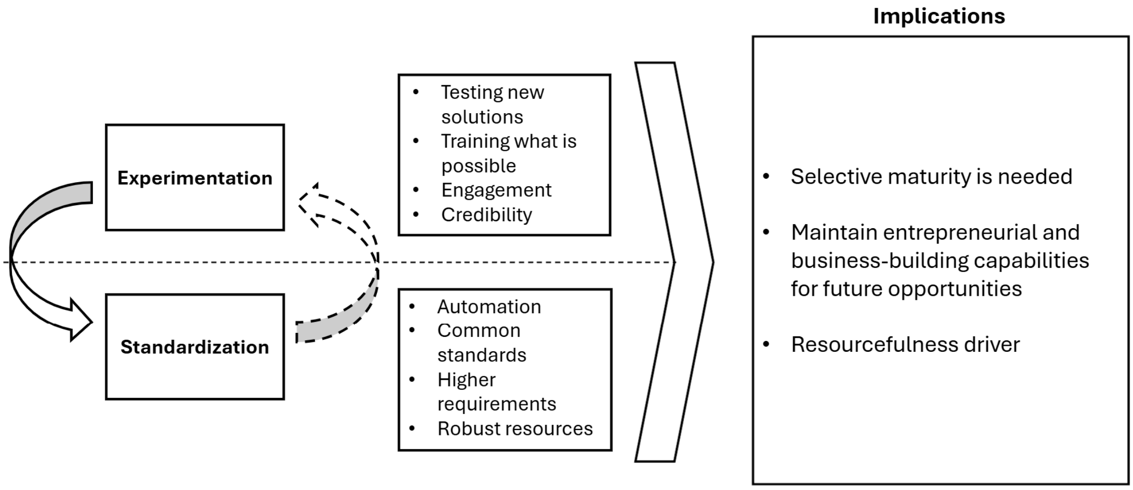

Figure 6.12: Overview of the tension and relationship between standardization and experimentation

6.6.3 Leadership and Orchestration Hold the Ecosystem Together

The third dimension concerns the role of leadership and orchestration. Across interviews, flexibility was not portrayed as something that naturally falls into place once market conditions or technologies exist. Rather, it emerged as something that requires proactive prioritization, internal anchoring, and coordination among actors with different roles, incentives, and capabilities. Externally, this is largely confirmed, but external respondents also show that orchestration includes credibility-building, matchmaking, intermediary actors and clearer signal-setting from system actors.

R3 argued that one reason the DSO has remained active in flexibility while other actors have reduced or paused their efforts is that leadership has consistently believed in the issue and prioritized it over time. In their view, flexibility is difficult precisely because it requires changed routines, systems, and mental models. R3 described leadership commitment as decisive for sustaining this kind of change over time:

> *It is a lot about leadership to lead change. We have had a leadership and a CEO that has believed in this path for 5-6 years now, and has dared to invest resources into it and not everyone else has dared to do that. It has been a decisive factor what type of prioritization the leadership has and what you dare to invest* [R3]

R6 expressed a similar understanding, although from another perspective. The organization had reached at least two moments where continuing the flexibility journey was questioned. Nevertheless, the organization decided to continue, partly because it believed that customers could help operate the system more efficiently, and partly because a municipal actor of the DSOs size had a responsibility to move the issue forward. The reasoning frames flexibility as a strategic and collective project rather than a narrow technical solution.

The external interviews validate this internal picture. B1 describes the DSO as an actor that helped them get started with their business already from the beginning:

> *They are stars that helped with credibility from day one* [B1]

B2 similarly experienced the DSO as having a clear strategic direction, even though all details of the market design were not yet settled:

> *My view of the DSO is that they want to get the market going and less important that it be absolutely correct at all times. It has been a bit of learning by doing from both sides* [B2]

B6s account also acknowledges that the DSO was actively trying to develop solutions, even if the overall market remained difficult:

> *They are doing good things, they show that they actually want to come with solutions* [B6]

The external perspective confirms that the DSO is experienced not just as a buyer of flexibility, but as a market-enabling actor that can influence its surroundings. Furthermore, external respondents ask for a more structured next step. B5 suggests that a more explicit tendering model, where a system actor says that a battery is needed at a specific location and developers bid to provide it, would reduce development costs and make investment cases clearer. B6 similarly suggests that actors need to feel welcome and need clearer conditions under which their participation becomes worthwhile. These views challenge the internal findings in a constructive way, since experimentation and market-building may have been the right first step, larger-scale external actors now appear to want more explicit coordination and clearer locational signals.

The leadership role was also visible in how respondents described internal collaboration. R3 noted that strategy, planning, and customer-related functions sit close together, interact frequently, and operate in a relatively innovative environment where ideas are tested and refined iteratively. Rather than following a fully predefined long-term roadmap, the respondent described a pattern of short-term goals, continuous exploration, and mutual adjustment between functions.

R5s account showed that coordination also takes place on the operational market. The respondent described a successful flexibility market as one characterized by several equivalent actors, diverse resource types, and enough liquidity, easy accessed flexibility resources, to avoid excessive dependence on one dominant participant. This suggests that the DSOs role is not simply to buy flexibility, but to actively shape the market composition and enable conditions under which different providers can participate efficiently.

The external respondents also described their dependencies and system actors. B1 depends on demand side stakeholders and data intermediaries to sell and gain access to operate, B3 depended on consultants who introduced the opportunity of BESS, and B7 points to aggregators, installation partners, DSOs, and local policy makers as critical project actors.

R7 also exemplified regulatory factors as a coordinated process driven by general national and EU direction but also influenced by the DSO's and TSO's involved in the system together bringing forward suggestions that institutions can interpret into regulation and policy. R7 also adds that by leading proactive solutions, it can help the DSO leverage and shape policy to an extent, by being able to show others how they achieved a solution.

Taken together this aggregate dimension show that flexibility markets has depended on active leadership and coordination to compensate for lacking market and technical maturity. Internally leadership has been important for sustaining commitment, allocating resources, and keeping flexibility on the agenda despite uncertainty and organizational change. Externally, respondents see leadership as a source of increased relevance, market enablement, and more structured conditions for participation. The chapter therefore suggest that orchestration happens through multiple actors in the ecosystem, the DSO, customers, aggregators, consultants, regulators, and policy actors. However, the central responsibility undeniably falls on the DSO to lead, align, and shape successful conditions for flexibility.

6.6.4 Competitive Revenue Streams Shape the Supply of Flexibility

The fourth finding is that the supply of flexibility is shaped by the relative attractiveness of different revenue opportunities. Respondents repeatedly pointed out that local flexibility does not compete only with internal inertia or technological barriers. It also competes with alternative markets, with other uses of the same resource, and with more conventional infrastructure solutions.

R2 explained that batteries, especially in the local flexibility markets, compete with ancillary services markets and arbitrage. R5 adds that there is also a limited visibility into the decision-making of the external stakeholders models.

B3 further noted how the initial value offer can change. The ancillary services markets previously offered particularly strong incentives because they paid well while consuming very little energy, while local flexibility became more attractive only as those revenue conditions changed.

The external interviews strengthen this pattern considerably. B1 explains that they aim to commit to as many revenue streams they can:

> *We commit to any type of flexibility we can, operations needs, spot price adjustments, flexibility market if there are incentives, and then we choose and stack everything we can to make cash for the client* [B1]

B3 similarly describes the battery acquisition as a specific opportunity that emerged for the business

> [...] *to have many legs to stand on and for the business to be more robust* [B3]

remarking that the battery investment was an interesting proposition at the time for the specific facility. However, the investment was not particularly lucrative compared to other alternatives in hindsight. B4 and B5 argue that local flexibility markets may not be the main reason for building their business case but still could contribute if they were in Gothenburg:

> *The local flexibility market is too small and too weak to facilitate a battery investment with only that as support* [...] *but if we were to invest for other reasons, then we would most likely be part of the local flexibility market as well* [B5]

R7 confirmed this from the DSO side, explaining that the DSO does not expect actors to build their entire business case around the local flexibility market. Instead, the market can add an additional revenue stream when an actor already has a resource or business model in place.

B6 explains from the perspective that there is no intrinsic preference for one market over another, the question is rather where the strongest commercial return is. These accounts show that the flexibility supply is determined by portfolio logic rather than by any commitment to one market category. Similarly, B7 argues that the strongest battery case begins with local site value, such as power tariff management or charging support, and that local flexibility market participation should be treated as *"a small upside to the whole"* [B7]. B7s account explain that a more robust path for the customer to contribute with flexibility may be one where the asset already makes sense for the customer locally, and where the flexibility market is layered on top rather than serving as the main justification.

R3 also discussed this perspective noting that flexibility is not a core business for the network operator, while it may well be part of the core business for firms that sell batteries, optimization services, or flexibility-related technologies. The respondent also pointed out that the current regulation still gives clearer returns on grid investment than on flexibility, meaning that the incentive structure for many network operators does not yet strongly support flexible alternatives.

B1 nuanced this picture by pointing out depending on which type of stakeholder they talked to, different incentives occur. Larger organizations like supermarkets might have more of an sustainability incentive, and see value in saving money in their budget. Compared to smaller actors who were more interested in how they could increase their revenue streams.

R6 argues the same customer pattern from a future prospects perspective, where external incentives may beneficially align to provide more flexibility without the DSO needing to invest large resources. Their example was the vehicle-to-grid opportunity, where the market moves to electric vehicles, and a large amount of flexibility would be made available without the DSO needing to build that market themselves, just the technical solutions around it.

R1 added further nuance describing that participation is not always driven by revenue maximization alone from their customers. Some actors participate partly because they want to contribute to the city, learn, or be part of a local development effort. However, the interviewee also made clear that such motivations coexist with business logic rather than replacing it. Battery actors, in particular, increasingly assess flexibility participation as part of a broader business case involving security, optimization, and possible future income.

The combined finding is therefore, firstly, both internal and external respondents agree that flexibility resources are shaped by competitive revenue streams and rarely follow local flexibility demand alone. Second, external material clarifies that flexibility will not attract investment at scale unless local value, stacked revenues, or both make the asset sufficiently robust against uncertainty and market volatility. Lastly, flexibility resources can be made available through external incentives not directly related to flexibility market interaction, such as vehicle-to-grid or to provide clearer business cases for projects.

6.6.5 Tariffs and Markets Must Be Aligned With How Resources Actually Respond

The final aggregate dimension concerns the relationship between tariffs, contracts, and market-based flexibility mechanisms. The respondents did not frame these as mutually exclusive solutions. Rather, they consistently described them as different instruments that may suit different resources, time horizons, and system needs. Across both groups, the repeated message is that current instruments do not yet align well enough with how flexible resources actually behave.

R3 argued that the local flexibility market is unlikely to remain in its current form indefinitely, and that tariff-based price signals could likely take over part of the flexibility market role. At the same time, the respondent maintained that certain local or contractual mechanisms will still be needed where more specific commitments are required. In their view, broad groups of customers are more likely to respond to price signals than to engage actively in a market platform, while more mature actors with stronger flexibility capabilities may instead be governed through contracts or more targeted market arrangements.

The external interviews make this problem more concrete. B1 argues that tariffs are a blunt tool because they punish a peak even if that peak does not coincide with when the network actually needs help. B2 and B3 show how recharging a battery between flexibility events can create an expense that *"eats up the profit"* [B2] from local flexibility participation. B6 similarly criticizes effect fees as poor system signals for fast-response assets. These accounts all point in the same direction: existing tariff logic often undermines rather than supports flexibility provision.

Several external stakeholders also perceive that different resources require different market logic. B1 stresses that demand side flexibility cannot be treated as a battery because

> *There is food in the store and we cannot provide flexibility at the cost of food going bad* [B1]

highlighting the local conditions that can limit the possibility of participation.

B7 notes that vehicle-to-grid might become useful in a large system context, but is not necessarily the right solution for peak management on a specific site. B6 generalizes this by arguing that different assets need different signals and that local flexibility markets *"cannot live in isolation"*[B6] from other market structures if participation is to remain economically sensible.

R5s interview strongly supports this interpretation. The respondent emphasized that different resource types have different operating logics. Batteries can often respond quickly, but are shaped by multiple competing markets. Backup generation is enduring but costly and undesirable to activate early. Operational resources such as cooling systems or industrial processes can provide flexibility only within the limits of their core activity. This means that no single steering instrument can be assumed to fit all resource types equally well. Instead, tariffs, contracts, and market products must be aligned with the technical and business characteristics of each resource and geography.

R2 expressed a similar view when discussing household and small-scale aggregated resources. They saw substantial flexibility potential in households, home batteries, and charging infrastructure, but also suggested that tariff-based signals may be especially relevant further down in the network, where broad, distributed behaviors matter more than explicit market participation. This points to a differentiated logic in which tariffs are better suited to wide-reaching behavioral steering, while local markets and contracts may be more appropriate for explicit and location specific relevant capacity needs.

R4 also reinforced the need for coordination across instruments. It was emphasized that grid expansion, flexibility markets, and tariffs should not be seen as alternatives in a binary sense, but *"all of them are needed"* [R4] to address peak-related capacity issues. In their account, the task is not to find one superior mechanism, but to combine several tools in a way that avoids overbuilding the grid for a small number of extreme hours.

R1s account shows what misalignment looks like in practice. The respondent described how some battery actors had charged between two peak periods in ways that created local problems, even though they were participating in the flexibility market. Because the organization was still in a learning-oriented phase, there were no strict sanctions in place. This illustrates that tariffs, contracts, and market rules cannot simply be layered on top of each other without careful coordination, they need to reflect how resources actually behave under real conditions.

Taken together, the findings suggest that the future development of flexibility is not primarily about choosing between tariffs and markets. Rather, it is about designing combinations of instruments that fit the response logic of different resource types and system needs.

6.6.6 Summary of the Empirical Pattern

Taken together, the internal and external interviews show that flexibility is understood as both a necessary future capability and an immature present-day coordination problem. Internal respondents mainly frame flexibility as a challenge for the DSO: how to build markets, create trust, make flexibility measurable, and eventually integrate it into planning and operation. External respondents largely agree with that ambition, but add that flexibility must also become investable, commercially understandable, and compatible with the economic and operational logic of the assets expected to provide it. More specifically, the findings suggest that the DSO is being pushed toward a broader role than historically has been the case. Instead of only building and operating network assets, the DSO must also identify potential flexibility resources, engage customers, validate delivery capability, coordinate with aggregators and other markets, interpret uncertain forecasts, and decide when flexibility is credible enough to influence network investment or operations. Several respondents therefore portray flexibility not only as a market or technical issue, but as a transformation of what the DSO needs to be able to do.

Strategic respondents emphasized long-term commitment, market building, and proactive investment in tools before their full value is visible. Planning-oriented respondents stressed the need for measurability, predictability, and risk management before flexibility can be treated as a credible substitute for infrastructure. Market-facing respondents emphasized supplier quality, liquidity, and the need to make local flexibility competitive relative to other revenue opportunities. Customer-facing respondents highlighted the importance of dialogue, testing, and practical problem-solving in identifying and activating flexibility in the first place.

Rather than pointing to one single obstacle, the empirical material suggests that flexibility is constrained by a combination of information uncertainty, immature coordination mechanisms, competing value streams, and the organizational challenge of integrating new tools into established planning and operational logics. At the same time, the interviews show that DSO has already moved beyond a purely conceptual interest in flexibility. The organization is actively building, testing, and selectively professionalizing a local flexibility ecosystem, even if many key questions remain open.

6.7 Interdisciplinary Analysis

The qualitative findings reveal why flexibility is difficult to scale, actors lack certainty about where flexibility exists, how reliable it is, how it can be effectively captured and how it should be governed to be an attractive solution. The quantitative analysis responds to this challenge by exploring how smart meter data, aggregation, surge detection, clustering and contextual variables can support more informed decision-making in distribution networks. Together, the two analyses show how data can help transform flexibility from an uncertain and fragmented opportunity into a more visible, measurable, and actionable resource.

The interdisciplinary analysis is introduced in two sections. The first section examines how data can support the operational challenges found in the qualitative results by both identifying possibilities and capturing these through the proposed surge detection and clustering methods. The second section then considers how these insights can contribute to operational transformation by supporting scalable data pipelines, ecosystem coordination, verification, and the design of the DSO tools.

6.7.1 Data as a Foundation for Sensing and Seizing Flexibility

The qualitative findings repeatedly highlighted that one of the central challenges facing the DSO is not necessarily a lack of flexibility resources, but rather limited visibility regarding where flexibility exists, how much can be expected, and which actors are most relevant to engage. Internal respondents emphasized the need to identify flexibility potential [R5, R7], understand its geographical distribution [R1, R4, R7], and develop more efficient methods for targeting relevant customers and flexibility providers [R1, R2].

The quantitative analysis directly addresses these challenges. Through surge detection and clustering, the study provides a systematic overview of both the magnitude and distribution of flexibility potential across the network. Rather than treating the smart meter population as a homogeneous resource pool, the clustering approach identifies distinct consumption profiles with substantially different flexibility characteristics. This allows flexibility potential to be quantified not only at an aggregate level, but also in terms of behavioral patterns and cluster-specific load profiles. The findings demonstrate that flexibility is highly concentrated. By targeting only 27.7% of the surge-active meters approximately 92.7% of the estimated flexibility potential remains accessible, see table 6.3. From an organizational perspective, this substantially improves the ability to prioritize resources and focus engagement efforts on the most relevant customer segments. Instead of relying on broad market participation, the DSO can identify clusters exhibiting the highest concentration of surge behavior and direct flexibility initiatives accordingly.

The analysis also provides a foundation for ecosystem coordination. Several respondents described how aggregators, project developers, and internal stakeholders often expend significant resources identifying suitable flexibility opportunities [B1, B5, B6]. The cluster-level results create a standardized representation of where flexibility is located, how it behaves over time, and which types of consumption profiles dominate different regions. Such information can reduce uncertainty for external actors while enabling the DSO to communicate system needs more effectively without providing preferential treatment to individual market participants. Furthermore, the identified load profiles establish an analytical baseline for future flexibility opportunities. Emerging resources such as vehicle-to-grid solutions may be evaluated against the observed surge patterns and consumption characteristics, in a more advanced setting than the current exploratory approach only incorporating the use into the existing consumption clusters. The threshold-based methodology developed in this study therefore not only measures existing flexibility, but also creates a framework through which future technologies can be assessed and compared against current system needs. Taken together, the quantitative analysis transforms large volumes of smart meter data into organizationally actionable information. By making flexibility visible, measurable, and segmentable, the analysis strengthens the DSO's ability to sense opportunities and size potential interventions before significant resources are committed.

6.7.2 From Flexibility Identification to Operational Integration

While identifying flexibility potential is a necessary prerequisite, the qualitative findings suggest that the greater challenge lies in transforming identified flexibility into a reliable operational resource. Respondents repeatedly emphasized the need for standardization, automation, accountability, and governance mechanisms capable of integrating flexibility into established planning and operational processes. The quantitative analysis contributes to this transformation by providing a repeatable and scalable methodology for monitoring flexibility resources. Rather than relying solely on individual pilot projects, customer dialogues, or manual assessments, the developed pipeline continuously classifies and evaluates consumption behavior using a common analytical framework. This creates a shared reference point through which flexibility providers, aggregators, and internal stakeholders can discuss opportunities using comparable metrics and definitions.

The clustering results are particularly relevant in this context, see Section 6.5. The identified clusters do not merely represent statistical groupings, but potentially reflect different businesses or household patterns in preferences and operational characteristics. Some clusters exhibit frequent and recurring surge behavior, while others display rarer but more concentrated events. Such differences suggest that flexibility resources may require different governance mechanisms, contractual structures, and incentive designs. Additionally, the analysis provides support for the qualitative finding that tariffs, contracts, and market mechanisms cannot be designed as one-size-fits-all solutions. The findings support the development of moving into automated flexibility solutions. By finding recurring patterns and similar consumption behaviors, the analysis creates opportunities to move away from manual customer acquisition toward more standardized approaches. This would allow aggregators and other market actors may target behavioral segments rather than individual customers, while the DSO can evaluate flexibility portfolios based on observable consumption characteristics rather than solely on self-reported capabilities.

Accountability came to be an important interdisciplinary contribution. Several interviewees highlighted the need to verify whether contracted flexibility is actually delivered. The analytical framework developed in this study provides a foundation for such verification by establishing measurable baseline behavior and enabling deviations from expected consumption patterns to be observed over time. This increases transparency and may strengthen confidence in flexibility as a planning and operational resource.

Finally, the findings provide insights relevant to the long-term evolution of flexibility markets and tariff structures. Previous organizational initiatives aimed at broader tariff-based transformations have struggled to achieve the desired impact. The results suggest that improved visibility into consumption behavior and flexibility characteristics may help future instrument design become more closely aligned with how different resource types actually respond. In this sense, the analytical capability developed in the study does not only support current flexibility markets, but also contributes to the future design of tariffs, contracts, and emerging flexibility mechanisms.

Viewed through the dynamic capabilities lens, the interdisciplinary contribution extends beyond identifying flexibility resources [20]. The combination of organizational understanding and data-driven analysis enables the DSO to progress from sensing opportunities, through seizing their potential impact, toward transforming the organization to work with flexibility in a more standardized, accountable, and operationally integrated way.

7

Discussion

In this chapter the results and implication of the proposed methods are discussed. The discussion follows five sections. First, a general discussion of proposed methods leading up to answering the research questions are discussed. Second, clustering evaluation is discussed in regards to RQ1. Third, the managerial implications of the study is discussed in regards to RQ2. Fourth, the interdisciplinary method is discussed in regards to RQ3. Lastly, a conclusion summarizing the findings and learnings.

7.1 Clustering Evaluation RQ1

The results examined in Section 6.5 examining Gothenburg South East and the other regions denoted in the Appendix demonstrate that clusters emerging from Ward's hierarchal clustering of the de-aggregated 15 min smart meter data table produce behaviorally coherent and understandable groups which would allow for operational value. The 5 clusters emerging each had distinct temporal signature and surge behaviors 6.7 that maps plausibly onto the underlying consumer groups. Cluster C1, the most populated group, consolidates meters with negligible surge activity even after the exclusion of entirely inactive meters, indicating that the majority of consumption in the region is structurally stable and contributes little observable flexibility. Clusters C3 and C4 exhibit muted but seasonally structured profiles, with a consistent suppression of surge activity during the summer months of July and August, and are geographically concentrated in residential areas. This seasonal pattern aligns with the heating-dominated demand structure of Swedish residential consumers. Cluster C5 stands apart with a highly volatile profile and pronounced weekday / weekend cycling, and is located almost exclusively in industrial and commercial zones. This behavioral signature is consistent with the structured operating hours of these consumers and represents the most operationally actionable flexibility signal, as its periodicity is predictable and the excess energy volumes are substantial.

How ever there are limitations that limit the extent conclusions can be drawn. Incorporating time-of-day surge distributions for instance, by aggregating surge counts into morning, afternoon, and evening bins before clustering would allow the method to distinguish demand response profiles that are relevant to different grid balancing services. Additionally, only meters with at least one surge event are admitted to the clustering step. While this is a deliberate choice to ensure that the feature space contains observable signal, it also means that the analytical output says nothing about the large population of structurally stable consumers.

An important improvement direction concerns the threshold parameters λ and δ that govern surge classification. The current values, $\lambda = 2$ standard deviations and $\delta = 2\,\text{kWh}$, are fixed globally across all meters and regions. For industrial meters with high baseline consumption, a 1 kWh absolute threshold may be so small relative to normal variability that it passes nearly all statistical exceedances, inflating surge counts for this consumer class. Conversely, for small residential meters, a two-sigma threshold may be too stringent and suppress genuine behavioral events. Adaptive thresholds calibrated to meter-level baseline magnitude. for example, expressing δ as a fraction of each meter's median consumption would produce a more uniform detection sensitivity across the consumer population and likely sharpen the cluster boundaries between residential and commercial archetypes.

7.2 Managerial Implications RQ2

In this discussion, the empirical findings are connected to the theoretical concepts presented in the background. The findings show that positioning a DSO within an emerging flexibility ecosystem is not only a question of creating a local flexibility market. Rather, it concerns how the DSO manages a set of tensions that arise when flexibility moves from experimentation toward more structured use in grid planning and operation. Across the empirical material, flexibility appears as a necessary future capability, but also as an immature coordination problem. Internal respondents describe the need to build markets, create trust, validate resources, and integrate flexibility in an automated way. External respondents largely support this ambition, but emphasize that flexibility must also become investable, commercially understandable, and compatible with operational logic of the resources expected to provide it.

Innovation ecosystem theory helps explain why flexibility depends on interdependent actors whose activities, incentives, and resources must align [13]. Dynamic capabilities theory helps explain how the DSO must sense opportunities, seize them through concrete initiatives, and transform internal routines under uncertainty [20]. Innovation diffusion and adoption theory helps explain why external actors do, or do not, participate in flexibility markets [14]. However, the empirical findings also nuance these theoretical concept categories. They reveal a number of managerial tensions that the DSO must actively navigate.

The following sections therefore discuss flexibility development through five central tensions. First, the tension between shared system value and fragmented actor incentives in an emerging innovation ecosystem. Second, the tension between DSO leadership and excessive ecosystem responsibility. Third, the tension between experimentation and organizational integration. Fourth, the tension between learning openness and scalable standardization. Fifth, the tension between technical flexibility potential and commercial adoption logic. Together, these tensions show that the DSOs managerial challenge is not simply to procure flexibility, but to orchestrate an ecosystem in which flexibility becomes measurable, reliable, commercially attractive, and institutionally supported.

7.2.1 Flexibility as an Emerging Innovation Ecosystem

This section discusses how innovation ecosystem theory helps explain the interdependence between actors in the flexibility market. The findings show that local flexibility development depends on a broad set of actors whose incentives, capabilities, and decision logics do not automatically align. The DSO depends on customers, aggregators, technology providers, battery developers, consultants, installers, regulators, market platforms, and internal functions to identify, activate, and scale flexibility. This confirms that flexibility is not created by the DSO alone. Instead, it emerges through the coordinated activities of several actors across the electricity system and surrounding market environment.

This finding aligns strongly with innovation ecosystem theory, which describes value creation as dependent on interdependencies, complementarities, actor roles, coordination, governance, standards, and ecosystem risk [26], [57]. In line with this theory, the flexibility market can be understood as an ecosystem in which the DSOs ability to create value depends on whether other actors are willing and able to contribute complementary resources. The DSO may identify a congestion problem, but it cannot solve it through flexibility unless external actors have suitable resources, understand the opportunity, trust the market conditions, and find participation worthwhile.

However, the findings also show an important tension. While flexibility creates system value for the DSO by reducing congestion, improving grid utilization, and potentially complementing infrastructure investment, this does not mean that external actors value flexibility in the same way. For many providers, local flexibility is only one possible value stream among several. Batteries, aggregators, and demand-side resources may instead prioritize ancillary services, arbitrage, tariff savings, operational needs, or local site value when these alternatives are more attractive. This means that the innovation ecosystem is not held together by one shared value proposition, but by overlapping and shifting interests creating an ecosystem risk for the DSO [13]. In the literature, this relates to the adoption chain risk and the risk of shifting roles, where the flexibility actor risk not adopting flexibility and the role an actor takes in the system are uncertain and changing [24], [25]. These two challenges can be seen as opposing in this context, since invested resources in coordinating actors could be wasted if conditions change.

This creates a managerial challenge for the DSO. The DSO cannot assume that technical flexibility potential will automatically become available just because it is valuable for the grid or through previously robust incentives. Instead, it must actively understand and manage the fragmented incentives of external actors [13]. Managerially, this means that the DSO should map not only where flexibility is needed, but also why different actors would choose to provide it. The DSO should therefore clarify locational needs, communicate the value of participation in actor-specific terms, and design market conditions that fit the business logic of different providers. In this sense, ecosystem orchestration is not only about coordinating actors, but about continuously maintaining enough alignment between system value and actor-specific incentives for flexibility to remain useful [33]. However, this needed increased orchestration could also create a risk of wasting resources when conditions change, such as regulation changing the conditions of flexibility tools.

7.2.2 The Changing Role of the DSO

The findings indicate that the DSO's role is changing. The DSO is no longer only a grid owner, system planner, or buyer of flexibility. Instead, it increasingly acts as a market-enabling and ecosystem-shaping actor [24], [33]. Respondents describe how the DSO created credibility, facilitates learning, engages customers, translates system needs, supports experimentation, and contributes to the development of regulation and processes. External actors also describe the DSO as important for making flexibility relevant and credible, especially when starting up participation.

This finding connects closely to the innovation ecosystem concepts of keystone and orchestrator roles [27], [29]. The DSO resembles a keystone actor because it controls or influences central ecosystem assets, including grid knowledge, network data, customer relationships, local system needs, and parts of the market design. At the same time, it resembles an orchestrator because it coordinates actors, reduces uncertainty, enables experimentation, and shapes the conditions for participation. The findings therefore align with the theoretical view that central ecosystem actors can influence value creation not only through their own resources, but by shaping how other actors interact and contribute [13], [25], [33].

The tension is that the more the DSO acts as an ecosystem leader, the more it risks becoming responsible for solving uncertainty that is actually distributed across the ecosystem. Internal respondents describe the DSO as proactive and committed, while external respondents ask for clearer locational signals, more predictable investment conditions, and more standardized processes. This shows that the DSOs early leadership has helped initiate the ecosystem, but also that the next phase requires a clearer distribution of responsibility. If the DSO provides too little structure, external actors may hesitate to invest. If the DSO absorbs too much uncertainty, the ecosystem may become overly dependent on the DSO, and the DSO risks investing to many resources into uncertain opportunities.

For DSO managers, the implication is that leadership should shift from broad market activation toward structured orchestration [13]. The DSO should continue to provide direction, credibility, and system knowledge, but should avoid becoming the sole actor responsible for making every business case work. More specifically, the DSO should clarify what information it can provide, what risks external actors must carry, and where regulation need to support coordination. This includes providing clearer locational signals, more transparent flexibility needs, and more predictable participation processes. The DSOs role should therefore be to reduce unnecessary uncertainty, not to remove all uncertainty from the ecosystem.

7.2.3 Dynamic Capabilities in Flexibility Market Development

In this section the flexibility development is discussed as an organizational capability challenge for the DSO. The findings show that flexibility development has required the DSO to work under uncertainty and gradually build new organizational capabilities. Flexibility is not yet a stable product or a fully mature market mechanism. Instead, it has required practical testing, customer dialogue, market development, data analysis, regulatory engagement, and internal learning. Several respondents emphasize that leadership and long-term commitment have been important because the value of flexibility is not always visible from the beginning.

This aligns well with dynamic capabilities theory [32]. Sensing is visible in how the DSO identifies flexibility opportunities through customer dialogue, data analysis, testing, and exploration of different resource types. Seizing is visible in how the DSO develops tools, contracts, market platforms, co-learning processes, and potentially new tariff structures. Transforming is visible in attempts to integrate flexibility into planning routines, forecasting tools, market design, customer engagement, and regulatory discussions. The empirical material therefore supports the theoretical argument that dynamic capabilities are needed when organizations face technological, regulatory, and market uncertainty, such as increased pressure on grid capacity and more volatile production profiles of renewable generation technologies [20], [31].

However, the findings also show that the transformation dimension is not yet complete. The DSO has clearly sensed and seized flexibility opportunities, but flexibility is not yet fully integrated into ordinary planning and operational decision-making. Planning-oriented respondents emphasize that flexibility must become more measurable, predictable, and reliable before it can be used as a credible alternative to infrastructure investment. This creates a tension between experimentation and organizational integration. Experimentation has been necessary to learn what flexibility is and where it can be found, but integration requires more formal routines, decision criteria, and reliable standards. This adds a mechanism to the relationship between sensing, seizing and transforming in innovation ecosystems [20], [31]. The findings suggest that the tools and routines developed to sense and seize opportunities do not automatically lead to transformation. Instead, they can create different directional forces. On one hand, tools that identify flexibility potential, enable experimentation and collaboration can move the organization toward transformation. On the other hand, if these tools remain disconnected from the core organization, they may reinforce an experimental mode rather than enable full organizational integration.

Managerially, this means that the DSO should not treat flexibility only as a market or procurement activity. It should treat flexibility as an organizational capability that must be institutionalized. The DSO should develop routines for identifying suitable resources, validating delivery performance, forecasting availability, comparing flexibility with grid investments, and deciding when flexibility is reliable enough to influence planning or operations. The key implication is that pilot activity should increasingly be translated into repeatable organizational processes. Without this transformation, flexibility risks remaining an experimental side activity rather than becoming part of the DSOs long-term toolbox.

7.2.4 From Experimentation to Standardization

In this section the evident paradox between experimentation and standardization is discussed.The findings show that local flexibility has developed through experimentation, but that scaling now requires more standardization and automation. In the early phase, broad exploration helped the DSO identify resources, build relationships, test market arrangements, and learn how different actors respond. This experimentation was necessary because flexibility potential is often not known in advance. It must be discovered through interaction with customers and providers, practical testing, and local learning.

This finding aligns with innovation diffusion theory, particularly the importance of trialability [14]. Actors are more likely to engage with an innovation when they can test it, learn from it, and participate without committing to excessive risk from the beginning [37]. At the same time, the findings also align with innovation ecosystem theory, which emphasizes the importance of standards, modular interfaces, governance mechanisms, and coordination structures for scaling [30]. External actors describe fragmented rules, uncertain processes, and local variation as barriers to investment and national expansion. Internal respondents also describe a shift toward more standardized products, automation, higher delivery requirements, and clearer resource suitability.

The central tension is that the openness that enabled early learning can become a barrier to scaling. Too much experimentation may create uncertainty for actors that need predictable conditions before investing. However, too much standardization may exclude smaller actors, emerging technologies, or locally specific resources before their potential has been fully understood. The challenge is therefore not experimentation versus standardization, but how to combine learning openness with scalable predictability.

For DSO managers, the implication is that standardization should be selective. Core market elements should become clearer and more predictable, including product definitions, baseline requirements, delivery expectations, communication routines, and locational signals. At the same time, the DSO should preserve controlled experimentation for new resource types, new customer segments, and uncertain use cases. This would allow the flexibility ecosystem to mature without losing the learning capacity that enabled its development in the first place. Scaling flexibility therefore requires moving from open-ended experimentation toward structured experimentation within a more standardized market framework. Conceptually, this highlights a paradoxical arrangement in which low entry barriers and trialability support early adoption, while standards and common requirements may also be perceived as lowering entry barriers by making participation more predictable and scalable [14]. In the case of flexibility markets, this could reflect the diversity among flexibility providers in the ecosystem, considering there are both actors that benefit greatly from trial and those who want clearer standards and better access to scale.

7.2.5 Adoption and Participation Among Flexibility Providers

This section connects external actor participation to innovation diffusion and adoption theory [14], [21]. The findings indicate that technical flexibility potential does not automatically lead to market participation. External actors evaluate flexibility in relation to their own business models, operational constraints, investment logic, and alternative revenue opportunities. Batteries, aggregators, and flexible customers compare local flexibility revenues with ancillary services, arbitrage, tariff savings, operational value, sustainability benefits, and other priorities. For some actors, local flexibility is an additional upside rather than the main reason for investing in a flexible resource.

This aligns with innovation diffusion and adoption theory. Participation depends on whether flexibility offers relative advantage, fits existing operations, is understandable, involves manageable complexity, and can be tested without excessive risk. The findings particularly confirm the importance of relative advantage and compatibility [36], [37]. Flexibility providers are more likely to participate when the market complements an existing asset, business model, or site need. Social motives, sustainability ambitions, and local engagement may support participation, but they rarely replace the need for a credible commercial logic. However, examples have also been made about future vehicle-to-grid opportunities as well as previous notions of adoption due to temporary heightened incentives, suggest that participation can be a momentary opportunity that needs to be captured and proactively prepared for.

The tension is that the DSO may see flexibility potential from a system perspective, while providers evaluate the same potential through a commercial and operational lens. A resource that is technically valuable to the grid may not be commercially available if the provider can earn more elsewhere, faces operational disruption, or finds the market too complex. This means that adoption is shaped not only by whether flexibility is possible, but by whether participation is worthwhile compared to alternative uses of the same resource.

For DSO managers, the implication is that flexibility markets must be designed around adoption logic, not only system needs. The DSO should reduce administrative complexity, clarify expected revenues and requirements, and increase transparency around activation needs. Participation can be encouraged by lowering friction and making value easier to access. One such evident opportunity is through public tendering of battery parks and large projects. Another evident opportunity discussed in the interviews is through translating available data from smart metering to locational signals that aggregators and project developers can utilize to scale more flexibility and make clearer business cases.

7.2.6 Theoretical Contribution and Practical Implications

Taken together, the findings contribute to theory by showing how innovation ecosystems, dynamic capabilities, and adoption theory interact in a regulated infrastructure context. Flexibility is an ecosystem challenge because value depends on several interdependent actors whose incentives are only partly aligned. It is a dynamic capabilities challenge because the DSO must sense, seize, and transform under uncertainty while integrating flexibility into established planning and operational routines. It is an adoption challenge because external actors will only participate when flexibility fits their incentives, resources, and business models.

The findings also refine the theoretical perspectives. Innovation ecosystem theory is supported, but the empirical material shows that the flexibility ecosystem is not yet mature or fully aligned [24]. It is held together by overlapping interests that can shift as markets, regulation, and technologies develop. Dynamic capabilities theory is supported, but the findings show that sensing and seizing flexibility opportunities is easier than transforming them into ordinary planning and operational routines [31], [32]. Adoption theory is also supported, but the findings show that participation is shaped not only by the characteristics of flexibility as an innovation, but also by competing revenue streams and the operational behavior of different resources [14], [35].

The main managerial implication is that the DSO should position itself as an active orchestrator of flexibility, but not as the only actor responsible for making the ecosystem work. The DSO should use its grid knowledge, customer relationships, and system understanding to reduce uncertainty and create clearer participation conditions. However, it should avoid designing flexibility only from the DSO perspective or absorbing all uncertainty on behalf of external actors. Flexibility products, tariffs, and contracts must reflect how different providers create value and how their resources respond in practice.

7.3 Interdisciplinary Implications RQ3

This study answers RQ3 by demonstrating how a data pipeline can transform readily available smart meter, location, weather, and spot price data into organizationally useful information. The pipeline combines data ingestion, aggregation, forecasting, surge detection, and clustering to identify flexibility behavior at both individual and aggregated levels. By integrating technical data processing with organizational requirements, the pipeline moves beyond data collection and enables systematic analysis of where flexibility exists, how it behaves, and how it can potentially be utilized.

The results further demonstrate that such a pipeline can be implemented at operational scale. The full pipeline processed over 196,000 smart meters and a full year of 15-minute measurements, while the continuous analytics configuration completed within operationally feasible execution times. This indicates that the methodology can support recurring analysis rather than remaining a purely experimental framework.

The inclusion of weather and spot price data also provides contextual understanding of consumption behavior. The SARIMAX analysis showed that flexibility-related consumption patterns are influenced by external conditions, particularly seasonal temperature changes, while electricity price signals exhibited weaker explanatory power at the aggregated smart meter level. This illustrates how combining consumption and contextual data can improve interpretation of flexibility behavior and support more informed decision-making.

The findings show that the primary value of the pipeline lies in increasing visibility of flexibility resources. Through surge detection and clustering, flexibility becomes measurable and segmentable rather than hidden within large volumes of consumption data. The results demonstrate that flexibility potential is concentrated among specific consumption profiles, enabling more efficient targeting of customers, flexibility providers, and geographical areas. Notably, approximately 92.7% of the identified flexibility potential remained accessible by focusing on only 27.7% of surge-active meters, highlighting the value of behavioral segmentation for resource prioritization.

Beyond internal decision-making, the pipeline also creates value within the broader flexibility ecosystem. The generated insights provide a common analytical basis for communication between DSOs, aggregators, project developers, and other stakeholders. By combining behavioral clustering with geographical information, the analysis makes it possible to identify not only which consumption profiles exhibit flexibility, but also where they are located within the network. This supports more targeted flexibility initiatives and improves coordination between actors seeking flexibility resources.

The study therefore shows that data becomes useful when technical analysis is translated into organizational capability. Internally, the pipeline supports sensing and evaluating flexibility opportunities, while externally it contributes to greater coordination and transparency across the ecosystem. The main contribution is not the data itself, but the ability to transform large-scale consumption data into actionable information that supports planning, engagement, and future flexibility development.

7.4 Future Studies

The goal of this section is to clarify the limitations of the study and identify how future research can build on its findings. Although the study demonstrates how readily available data can be structured and analyzed to identify flexibility potential, it does not validate whether the identified flexibility can be activated in real operational settings. Furthermore, the study is based on one specific interpretation of flexibility and a limited set of stakeholder interviews, meaning that the findings should be viewed as an exploratory foundation rather than a complete operational solution. The pipeline identifies flexibility consumption patterns, but further research is needed to examine under which parameters clusters are made useful and scalable to capture more flexibility.

Future studies should therefore test the pipeline in live operational contexts. One important direction is to investigate how and under which circumstance data can be shared for desired outcomes within the ecosystem. The research highlight several opportunities for strategic data sharing to signal and align actors, but doing so in an ethical and impact-full way still remains to be studied. Another direction is modifying the pipeline and parameters for practical use cases such as specific aggregator customer profiles or internal customer relations functions, or for the bidding, to further drive standardization within flexibility.

Overall, this study should be viewed as an exploratory foundation rather than a complete operational solution. It shows how a data pipeline can be constructed and how its outputs can become useful for a DSO and its broader ecosystem, but further research is needed to validate, govern, and integrate these insights into daily operations.

7.5 Conclusion

This thesis has explored how flexibility resources can be identified, analyzed, and contextualized through the combination of large-scale data analytics and organizational perspectives. The results demonstrate that meaningful flexibility insights emerge when technical analysis and ecosystem understanding are considered together rather than independently.

The proposed methodology transforms large volumes of smart meter, weather, price, and location data into actionable information regarding flexibility behavior and resource availability. Through historical consumption, surge detection, clustering, and geographical analysis, the study demonstrates that flexibility potential is highly concentrated among specific behavioral groups, enabling more targeted and efficient flexibility initiatives. Furthermore, the computational efficiency of the analytical pipeline allows hundreds of thousands of smart meters and millions of measurements to be processed within operationally feasible timeframes, making large-scale recurring analysis practical.

From an organizational perspective, the findings show that flexibility development extends beyond technical resource identification and depends on the coordination between DSOs, aggregators, customers, regulators, and other ecosystem actors. The study highlights how data-driven insights can strengthen ecosystem orchestration, while organizational requirements provide direction for analytical development.

The primary contribution of this work is therefore the demonstration of how analytical and organizational perspectives reinforce one another. Analytical methods become more valuable when aligned with operational and strategic needs, while organizational decision-making becomes stronger when supported by scalable and evidence-based insights. As electricity systems continue to evolve, the combination of scalable data-driven analytics and organizational capability development will be central for turning flexibility from a theoretical resource into a practical tool for distribution grid management

Bibliography

[1] RISE Research Institutes of Sweden, *How sweden can meet increasing electricity demand*, Accessed: 2026-03-23, n.d. [Online]. Available: https://www.ri.se/en/energy-and-electrification/power-production/story/how-sweden-can-meet-increasing-electricity-demand.

[2] Nordic Transmission System Operators, "Nordic grid development perspective 2025," Nordic TSOs, Jun. 2025, Accessed: 2026-03-23.

[3] Ministry of Climate and Enterprise. "Swedens final updated national energy and climate plan (necp) 20212030," Accessed: Feb. 1, 2026. [Online]. Available: https://commission.europa.eu/document/download/26d2c93e-641d-489f-a160-a7052fde58bb_en?filename=SE_FINAL%20UPDATED%20NECP%202021-2030%20%28English%29.pdf.

[4] E. Andersson, "Local capacity resources and the need for business models for bess usage/balancing," Third Conference of Swedish Electricity Storage and Balancing Centre; Head of Strategy at Göteborg Energi Elnät (DSO), Oct. 2025.

[5] C. Flygare, A. Wallberg, E. Jonasson, V. Castellucci, and R. Waters, "Correlation as a method to assess electricity users contributions to grid peak loads: A case study," *Energy*, vol. 288, 2024. DOI: 10.1016/j.energy.2023.129805. [Online]. Available: https://doi.org/10.1016/j.energy.2023.129805.

[6] L. Magnusson and R. Thorsson, "Clue – clustering-based load understanding and exploration," English, Student essay, University of Gothenburg, 2025. [Online]. Available: https://hdl.handle.net/2077/89783.

[7] M. Thelen, V. Hornung-Prähauser, M. Lassnig, and G. Pressmair, "Emergence of new ecosystems for innovative e-mobility services: Exploring business model patterns for vehicle-to-grid technology," in *RD Management Conference "Responsible and Responsive Innovation for a Better Future" in Seville, Spain on 19 June 2023*, Jun. 2023.

[8] R. Cadenovic, S. Vitiello, J. Landeka, and D. Pozo Camara, "The role of electricity distribution systems in assessing flexibility needs: Flexibility needs assessments according to the energy market design reform," Publications Office of the European Union, Tech. Rep., 2024. DOI: 10.2760/6973071. [Online]. Available: https://data.europa.eu/doi/10.2760/6973071.

[9] Göteborg Energi Nät AB, "Elektrifieringsrapport nr 1 2025: Göteborgs elektrifiering," Göteborg Energi, Gothenburg, Sweden, Tech. Rep., Jun. 2025, Accessed: 2026-03-23. [Online]. Available: https://www.goteborgenergi.se/om-oss/vad-vi-gor/hallbarhet/elektrifieringsrapporten.

[10] J. Torriti, R. Hanna, B. Anderson, G. Yeboah, and A. Druckman, "Peak residential electricity demand and social practices: Deriving flexibility and greenhouse gas intensities from time use and locational data," *Indoor and Built Environment*, vol. 24, no. 7, pp. 891–912, 2015. DOI: 10.1177/1420326X15600776. [Online]. Available: https://doi.org/10.1177/1420326X15600776.

[11] M. R. Bailey, D. P. Brown, B. C. Shaffer, and F. A. Wolak, "Take the load off: Time and technology as determinants of electricity demand response," National Bureau of Economic Research, Working Paper 33755–33824, 2025, pp. 1–68.

[12] R. M. Grant, *Contemporary Strategy Analysis*, 12th ed. Wiley, 2025.

[13] R. Adner, "Match your innovation strategy to your innovation ecosystem," *Harvard Business Review*, 2006.

[14] E. M. Rogers, *Diffusion of Innovations*, 5th ed. Free Press, 2003.

[15] M. G. Jacobides, C. Cennamo, and A. Gawer, "Towards a theory of ecosystems," *Strategic Management Journal*, vol. 39, pp. 2255–2276, 2018. DOI: 10.1002/smj.2904.

[16] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD'96)*, AAAI Press, 1996, pp. 226–231.

[17] M. Armbrust et al., "Spark sql: Relational data processing in spark," in *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, ser. SIGMOD '15, Melbourne, Victoria, Australia: Association for Computing Machinery, 2015, pp. 1383–1394, ISBN: 9781450327589. DOI: 10.1145/2723372.2742797. [Online]. Available: https://doi.org/10.1145/2723372.2742797.

[18] T. Ozaki, "On the order determination of arima models," *Journal of the Royal Statistical Society Series C: Applied Statistics*, vol. 26, no. 3, pp. 290–301, Nov. 1977, ISSN: 0035-9254. DOI: 10.2307/2346970. eprint: https://academic.oup.com/jrsssc/article-pdf/26/3/290/48620221/jrsssc_26_3_290.pdf. [Online]. Available: https://doi.org/10.2307/2346970.

[19] F. R. Alharbi and D. Csala, "A seasonal autoregressive integrated moving average with exogenous factors (sarimax) forecasting model-based time series approach," *Inventions*, vol. 7, no. 4, 2022, ISSN: 2411-5134. DOI: 10.3390/inventions7040094. [Online]. Available: https://www.mdpi.com/2411-5134/7/4/94.

[20] D. J. Teece, "Business models and dynamic capabilities," *Long Range Planning*, vol. 51, no. 1, pp. 40–49, 2018, ISSN: 0024-6301. DOI: 10.1016/j.lrp.2017.06.007.

[21] I. Mignon and A. Bergek, "System- and actor-level challenges for diffusion of renewable electricity technologies: An international comparison," *Journal of Cleaner Production*, 2016. DOI: 10.1016/j.jclepro.2015.09.048.

[22] O. Granstrand and M. Holgersson, "Innovation ecosystems: A conceptual review and a new definition," *Technovation*, 2020. DOI: 10.1016/j.technovation.2019.102098.

[23] L. A. d. V. Gomes, A. L. F. Facin, M. S. Salerno, and R. K. Ikenami, "Unpacking the innovation ecosystem construct: Evolution, gaps and trends," *Technological Forecasting and Social Change*, vol. 136, pp. 30–48, 2018. DOI: 10.1016/j.techfore.2016.11.009.

[24] O. Dedehayir, S. J. Mäkinen, and J. R. Ortt, "Innovation ecosystems as structures: Actor roles, timing of their entrance, and interactions," *Technological Forecasting and Social Change*, vol. 183, 2022. DOI: 10.1016/j.techfore.2022.121875.

[25] R. Adner, *The Wide Lens: What Successful Innovators See That Others Miss.* Penguin, 2012.

[26] R. Adner, "Ecosystem as structure: An actionable construct for strategy," *Journal of Management*, vol. 43, no. 1, pp. 39–58, 2017. DOI: 10.1177/0149206316678451.

[27] M. Iansiti and R. Levien, "Strategy as ecology," *Harvard Business Review*, vol. 82, no. 3, pp. 68–78, 2004.

[28] B. Iyer, C.-H. Lee, and N. Venkatraman, "Managing in a small world ecosystem: Lessons from the software sector," *California Management Review*, vol. 48, no. 3, pp. 28–47, 2006. DOI: 10.2307/41166348. [Online]. Available: https://doi.org/10.2307/41166348.

[29] M. Iansiti and R. Levien, "Strategy as ecology," *Harvard Business Review*, 2004.

[30] E. Autio and L. Thomas, "Innovation ecosystems: Implications for innovation management?" In *The Oxford Handbook of Innovation Management*, M. Dodgson, D. M. Gann, and N. Phillips, Eds., Oxford University Press, 2014, pp. 204–228.

[31] N. J. Foss, J. Schmidt, and D. J. Teece, "Ecosystem leadership as a dynamic capability," *Long Range Planning*, vol. 56, no. 1, 2023. DOI: 10.1016/j.lrp.2022.102270. [Online]. Available: https://doi.org/10.1016/j.lrp.2022.102270.

[32] D. J. Teece, G. Pisano, and A. Shuen, "Dynamic capabilities and strategic management," *Strategic Management Journal*, vol. 18, no. 7, pp. 509–533, 1997, Accessed: 2026-03-18. [Online]. Available: http://www.jstor.org/stable/3088148.

[33] E. Niesten, L. Striukova, and E. Bacha, "Leadership cognition as a driver of ecosystem strategies in sustainable business ecosystems," *M@n@gement*, vol. 28, no. 5, pp. 89–110, 2025. DOI: 10.37725/mgmt.2025.11468. [Online]. Available: https://doi.org/10.37725/mgmt.2025.11468.

[34] D. P. Hannah and K. M. Eisenhardt, "How firms navigate cooperation and competition in nascent ecosystems," *Strategic Management Journal*, vol. 39, no. 12, pp. 3163–3192, 2018. DOI: 10.1002/smj.2750.

[35] A. Grübler, "Time for a change: On the patterns of diffusion of innovation," *Daedalus*, vol. 125, no. 3, pp. 19–42, 1996, ISSN: 00115266. Accessed: Mar. 2, 2026. [Online]. Available: http://www.jstor.org/stable/20027369.

[36] L. G. Tornatzky and K. J. Klein, "Innovation characteristics and innovation adoption-implementation: A meta-analysis of findings," *IEEE Transactions on*

Engineering Management, vol. EM-29, no. 1, pp. 28–45, 1982. DOI: 10.1109/TEM.1982.6447463.

[37] P. Ponnaganti, R. Sinha, J. R. Pillai, and B. Bak-Jensen, “Flexibility provisions through local energy communities: A review,” *Next Energy*, vol. 1, no. 2, p. 100 022, 2023, ISSN: 2949-821X. DOI: 10.1016/j.nxener.2023.100022.

[38] G. A. Moore, *Crossing the Chasm: Marketing and Selling Disruptive Products to Mainstream Customers*, 3rd ed. HarperBusiness, an imprint of HarperCollinsPublishers, 2014.

[39] S. O. Negro, F. Alkemade, and M. P. Hekkert, “Why does renewable energy diffuse so slowly? a review of innovation system problems,” *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3836–3846, 2012. DOI: 10.1016/j.rser.2012.03.043.

[40] J. Palm, A.-R. Kojonsaari, I. Öhrlund, N. Fowler, and C. Bartusch, “Drivers and barriers to participation in sweden’s local flexibility markets for electricity,” *Utilities Policy*, vol. 82, 2023. DOI: 10.1016/j.jup.2023.101580. [Online]. Available: https://doi.org/10.1016/j.jup.2023.101580.

[41] Kungl. Ingenjörsvetenskapsakademien (IVA). “Iva-seminarium: Elnätets utmaningar i västra götaland,” IVA Väst, Accessed: Mar. 23, 2026. [Online]. Available: https://www.iva.se/det-iva-gor/evenemang/elnatets-utmaningar-i-vastra-gotaland/.

[42] Göteborg Energi. “Elmätarbyte.” accessed 2026-03-23. [Online]. Available: https://www.goteborgenergi.se/kundservice/elmatarbyte.

[43] A. Bolin, “Economic implications of smart meter deployment in enhancing energy efficiency and demand flexibility in the residential smart economy,” *Energy & Buildings*, vol. 349, 2025. DOI: 10.1016/j.enbuild.2025.116496. [Online]. Available: https://doi.org/10.1016/j.enbuild.2025.116496.

[44] Eurelectric. “What is flexibility in the power sector?” Eurelectric - Powering People, accessed 2026-03-23. [Online]. Available: https://www.eurelectric.org/in-detail/what-is-flexibility-in-the-power-sector/.

[45] International Energy Agency. “Flexibility electricity 2026 analysis.” IEA, accessed 2026-03-23. [Online]. Available: https://www.iea.org/reports/electricity-2026/flexibility.

[46] R. Li, F. Li, and N. D. Smith, “Multi-resolution load profile clustering for smart metering data,” *IEEE Transactions on Power Systems*, vol. 31, no. 6, pp. 4473–4482, 2016. DOI: 10.1109/TPWRS.2016.2536781.

[47] F. Murtagh and P. Legendre, “Wards hierarchical agglomerative clustering method: Which algorithms implement wards criterion?” *Journal of Classification*, vol. 31, no. 3, pp. 274–295, Oct. 2014, ISSN: 1432-1343. DOI: 10.1007/s00357-014-9161-z. [Online]. Available: http://dx.doi.org/10.1007/s00357-014-9161-z.

[48] D. Sculley, “Web-scale k-means clustering,” in *Proceedings of the 19th International Conference on World Wide Web*, ser. WWW ’10, Raleigh, North Carolina, USA: Association for Computing Machinery, 2010, pp. 1177–1178, ISBN: 9781605587998. DOI: 10.1145/1772690.1772862. [Online]. Available: https://doi.org/10.1145/1772690.1772862.

[49] E. Bell, A. Bryman, and B. Harley, *Business Research Methods*, 6th ed. Oxford University Press, 2022.

[50] E. Hartvigsson, N. Jakobsson, M. Taljegard, and M. Odenberger, "Comparison and analysis of gps measured electric vehicle charging demand: The case of western sweden and seattle," *Frontiers in Energy Research*, vol. Volume 9 - 2021, 2021, ISSN: 2296-598X. DOI: 10.3389/fenrg.2021.730242. [Online]. Available: https://www.frontiersin.org/journals/energy-research/articles/10.3389/fenrg.2021.730242.

[51] Statsmodels, *Statsmodels.tsa.arima.model.arima*, https://www.statsmodels.org/stable/generated/statsmodels.tsa.arima.model.ARIMA.html, Accessed: 2026-05-04, 2026.

[52] D. Müllner, *Modern hierarchical, agglomerative clustering algorithms*, 2011. arXiv: 1109.2378 [stat.ML]. [Online]. Available: https://arxiv.org/abs/1109.2378.

[53] J. B. McQueen, "Some methods of classification and analysis of multivariate observations," in *Proc. of 5th Berkeley Symposium on Math. Stat. and Prob.*, 1967, pp. 281–297.

[54] E. W. K. Tsang, "Generalizing from research findings: The merits of case studies," *International Journal of Management Reviews*, vol. 16, no. 4, pp. 369–383, 2014. DOI: 10.1111/ijmr.12024.

[55] A. Dubois and L.-E. Gadde, "Systematic combining: An abductive approach to case research," *Journal of Business Research*, 2002. DOI: 10.1016/S0148-2963(00)00195-8.

[56] S. Kvale, S. Brinkmann, and S.-E. Torhell, *Den kvalitativa forskningsintervjun*, 3rd ed. Studentlitteratur, 2014, Översättning av Sven-Erik Torhell.

[57] A. Gawer and R. Henderson, "Platform owner entry and innovation in complementary markets: Evidence from intel," *Journal of Economics & Management Strategy*, vol. 16, no. 1, pp. 1–34, 2007. DOI: 10.1111/j.1530-9134.2007.00130.x.

A

Appendix 1

A.1 Northern Hisingen

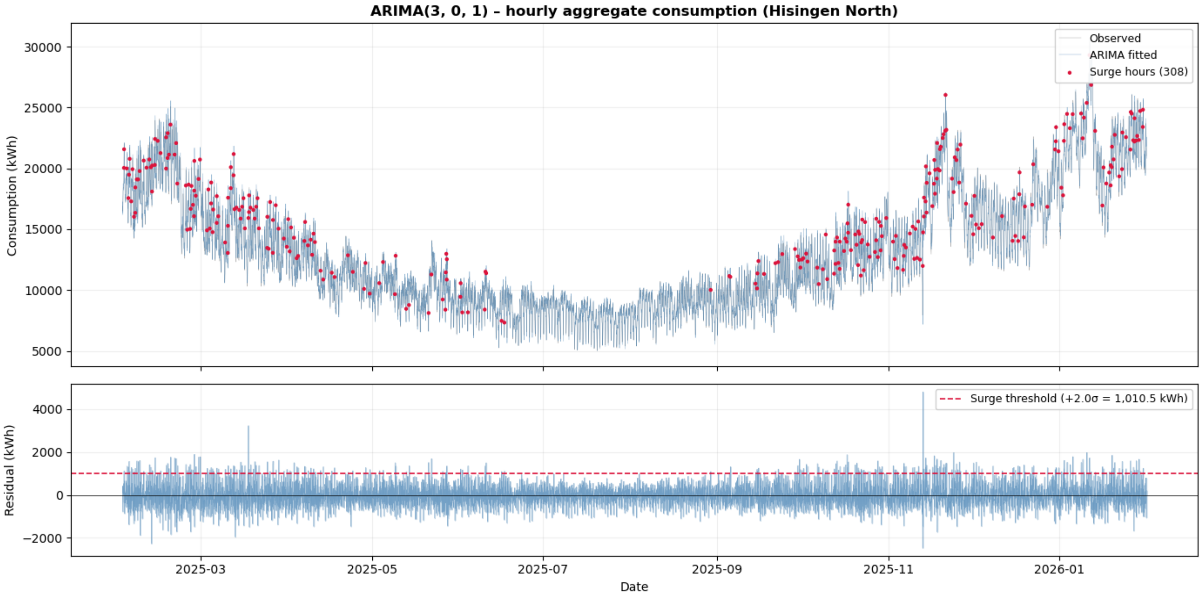

Figure A.1: ARIMA

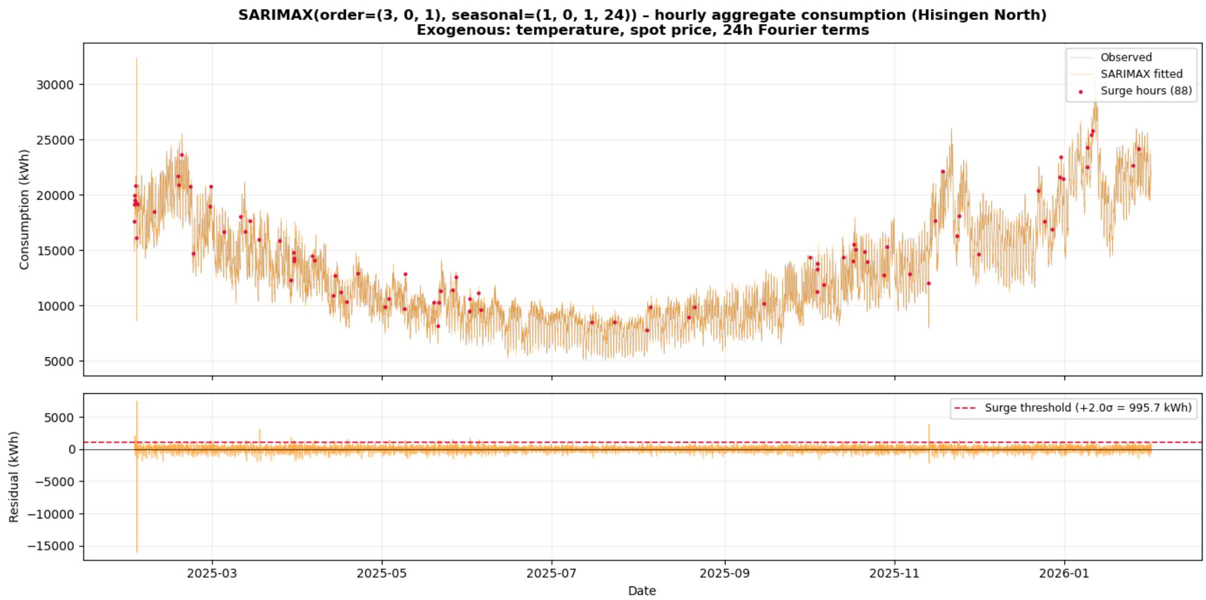

Figure A.2: SARIMAX

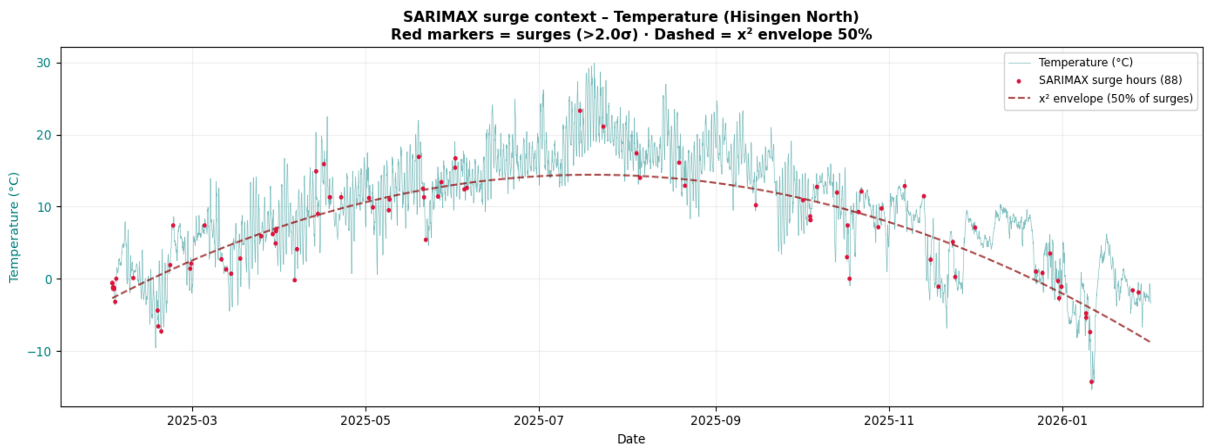

Figure A.3: Surges displayed over weather data

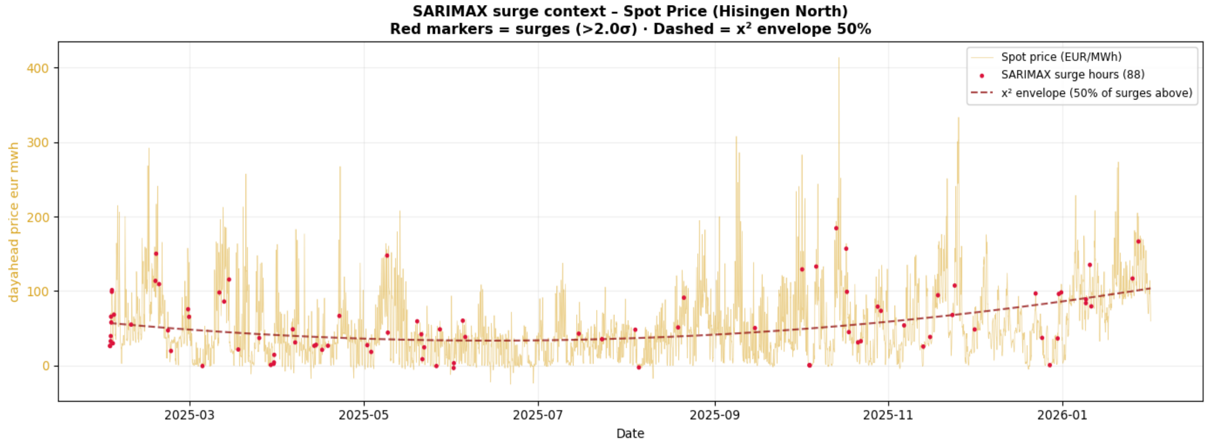

Figure A.4: Surges displayed over EPEX SPOT price data

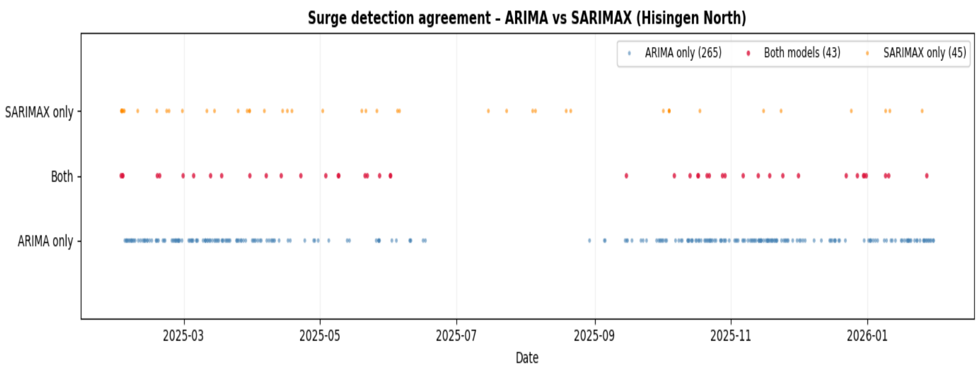

Figure A.5: ARIMA v. SARIMAX

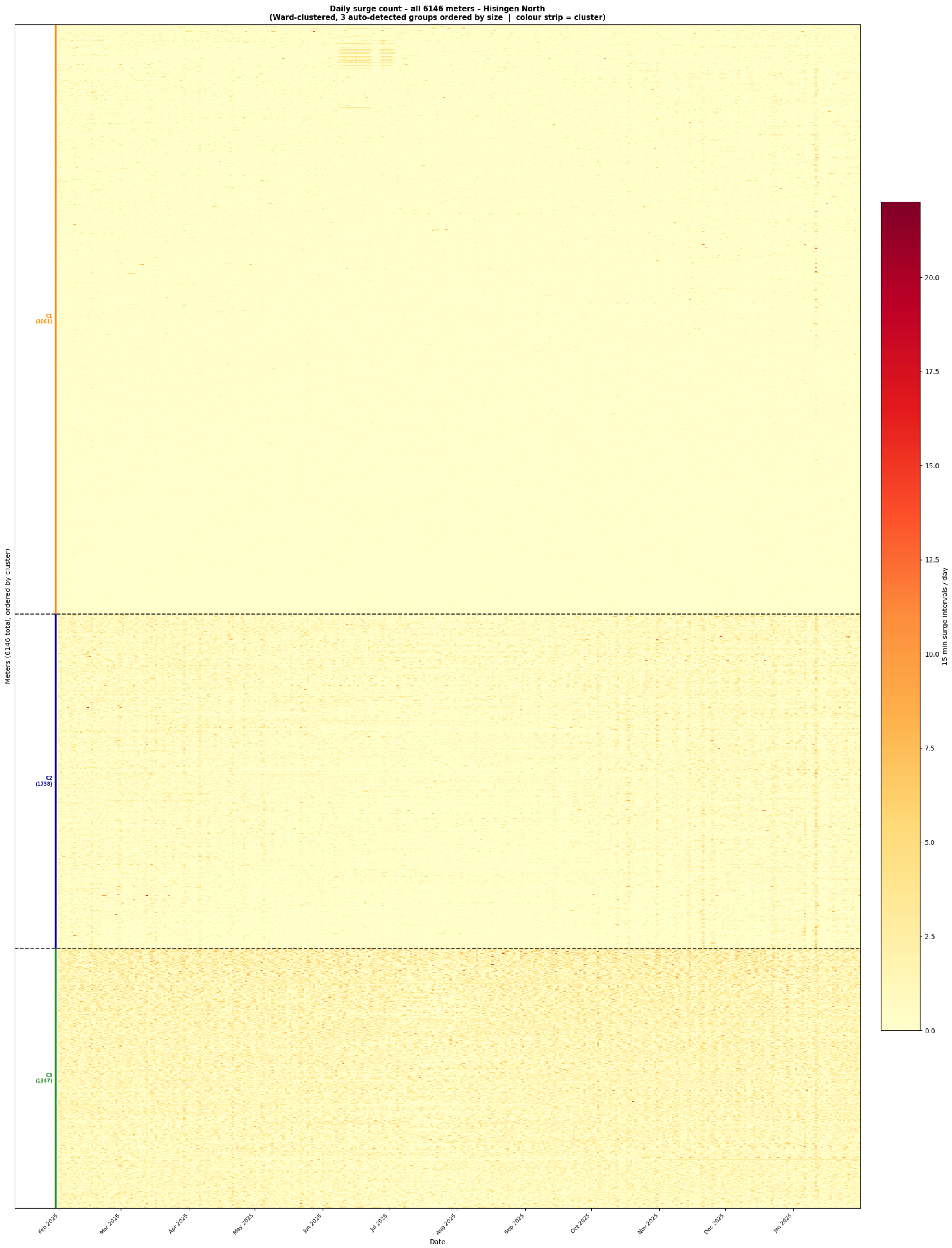

Figure A.6: Heatmap showcasing cluster capture of meter surge data

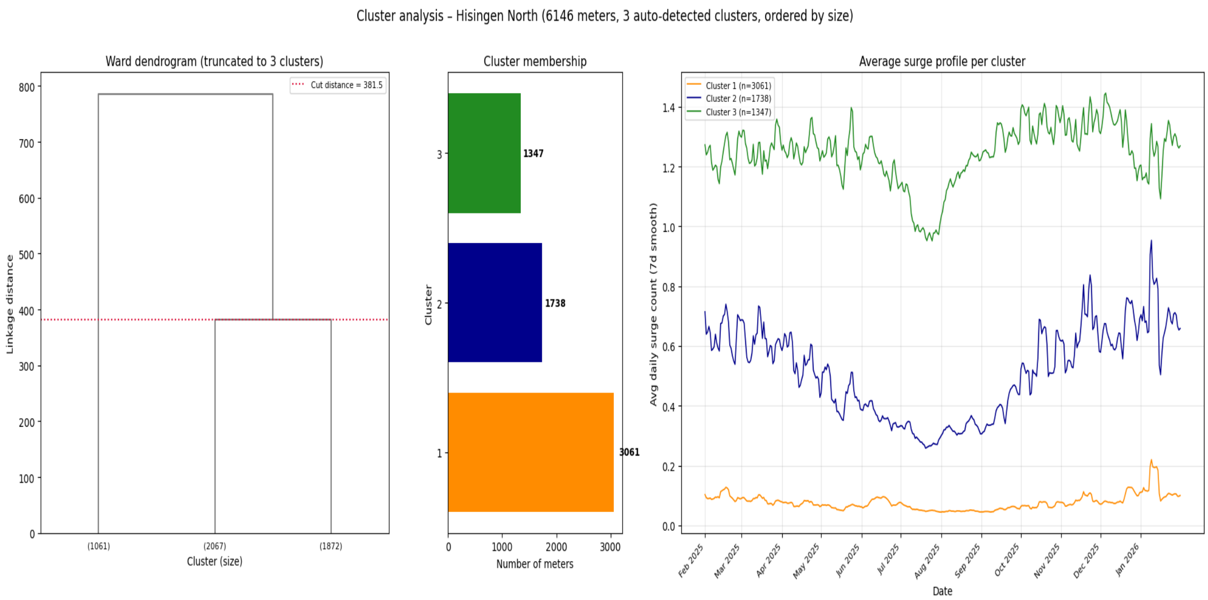

Figure A.7: Cluster analytics

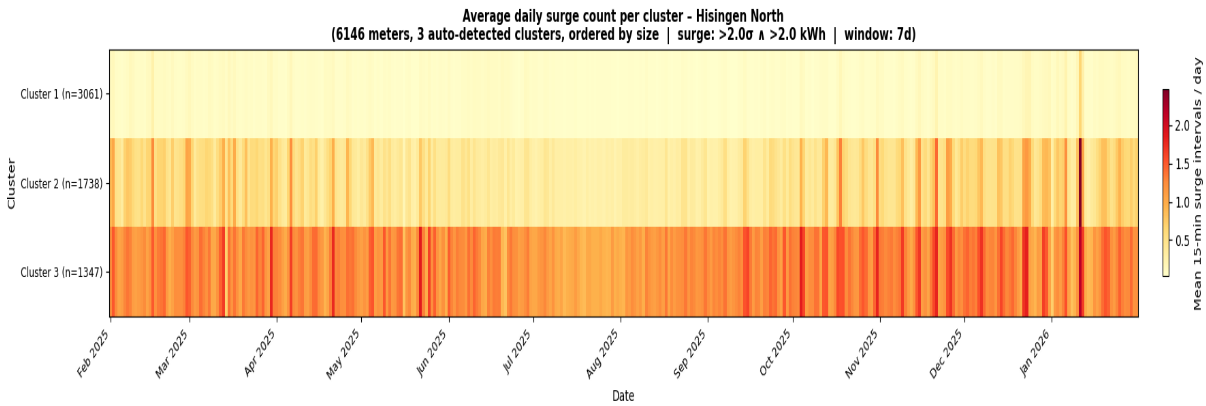

Figure A.8: Cluster heatmap showcasing surge occurrence

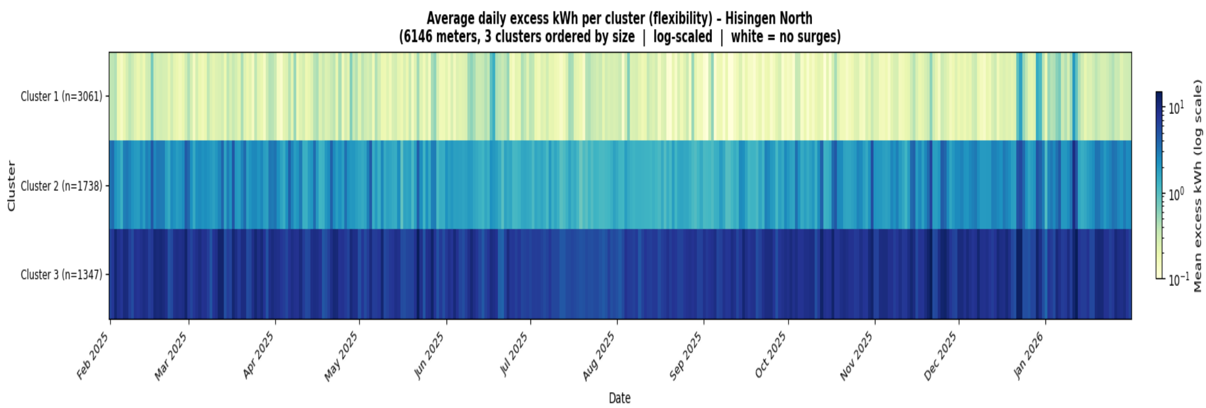

Figure A.9: Cluster heatmap showcasing excess kWh from surges

A.2 Western Hisingen

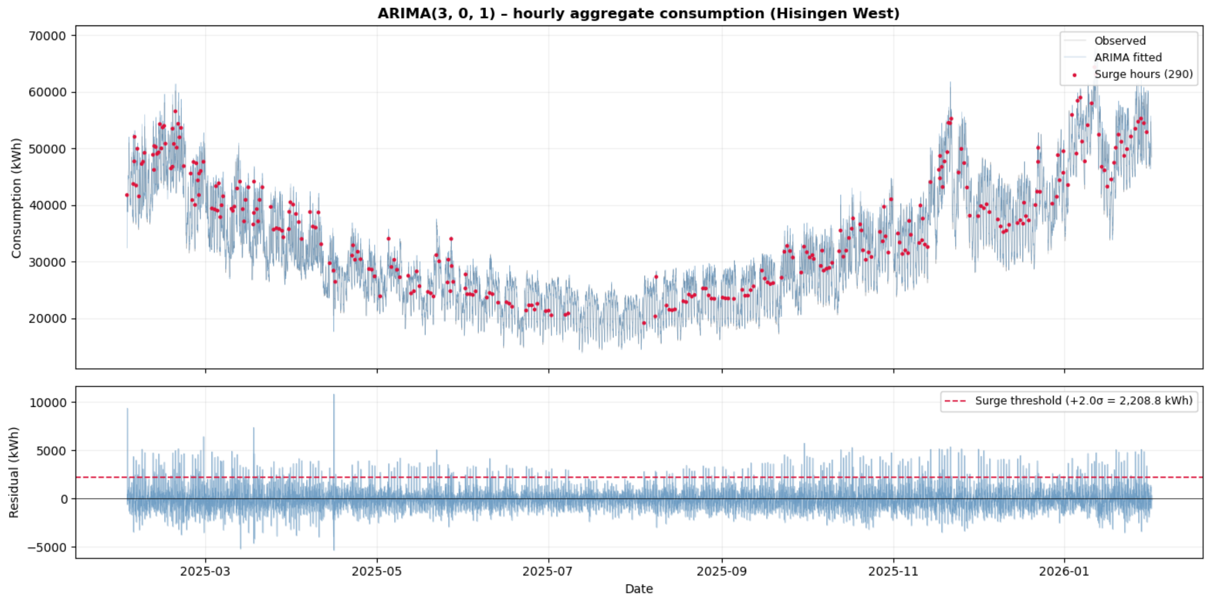

Figure A.10: ARIMA

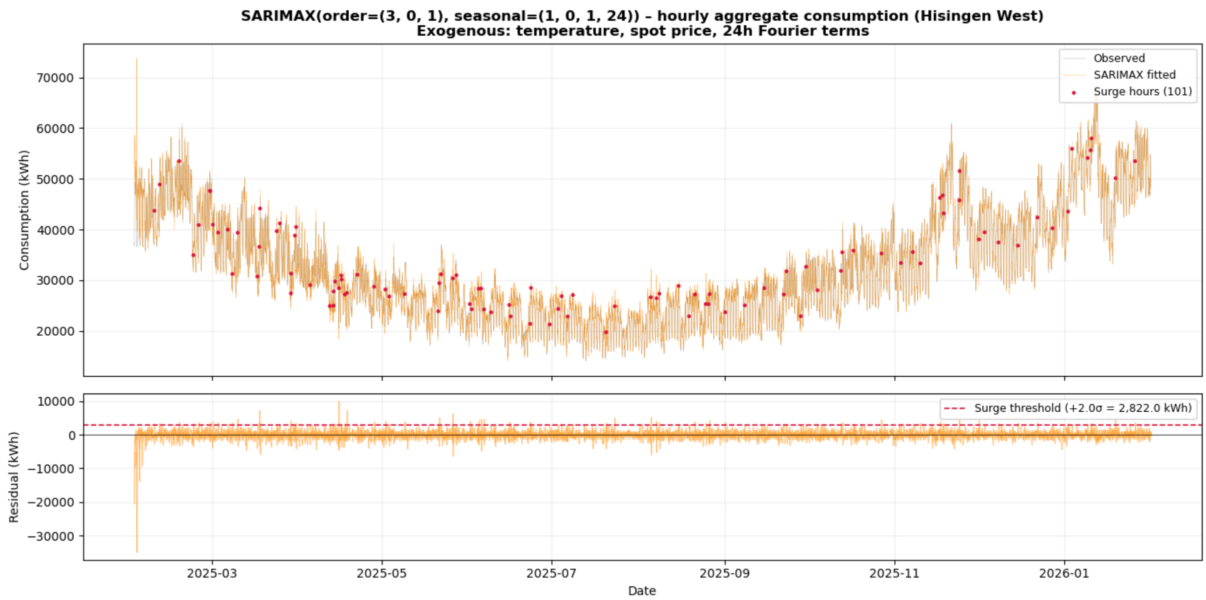

Figure A.11: SARIMAX

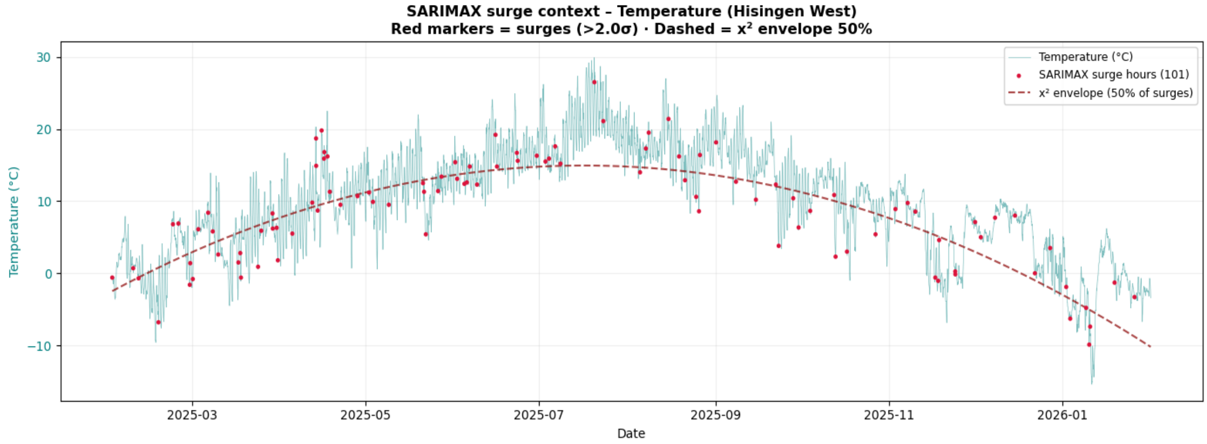

Figure A.12: Surges displayed over weather data

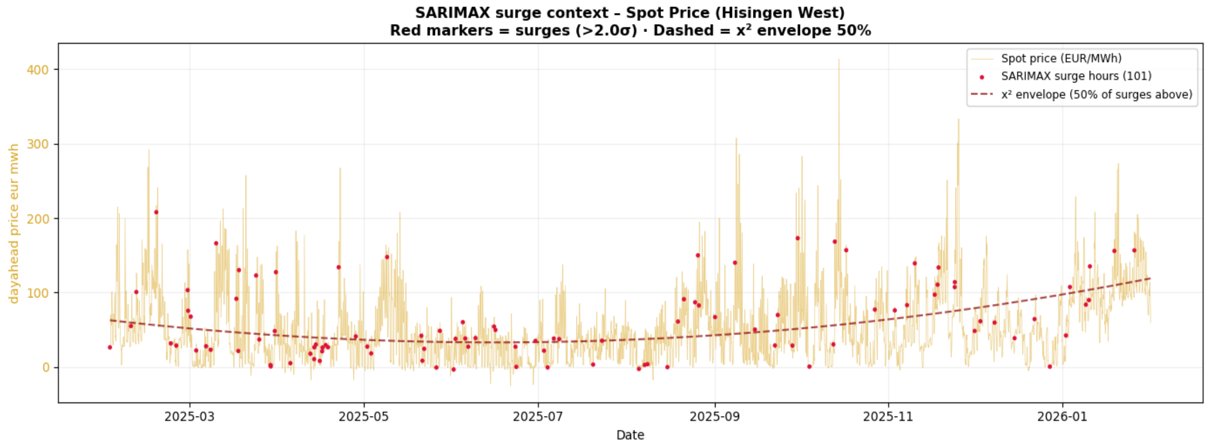

Figure A.13: Surges displayed over EPEX SPOT price data

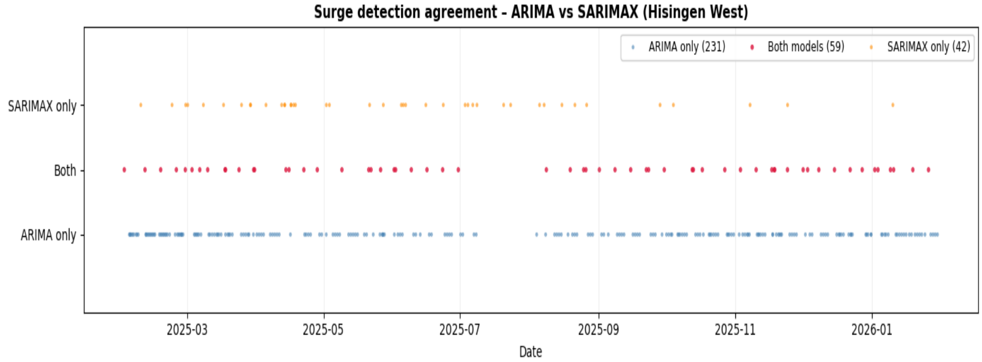

Figure A.14: ARIMA v. SARIMAX

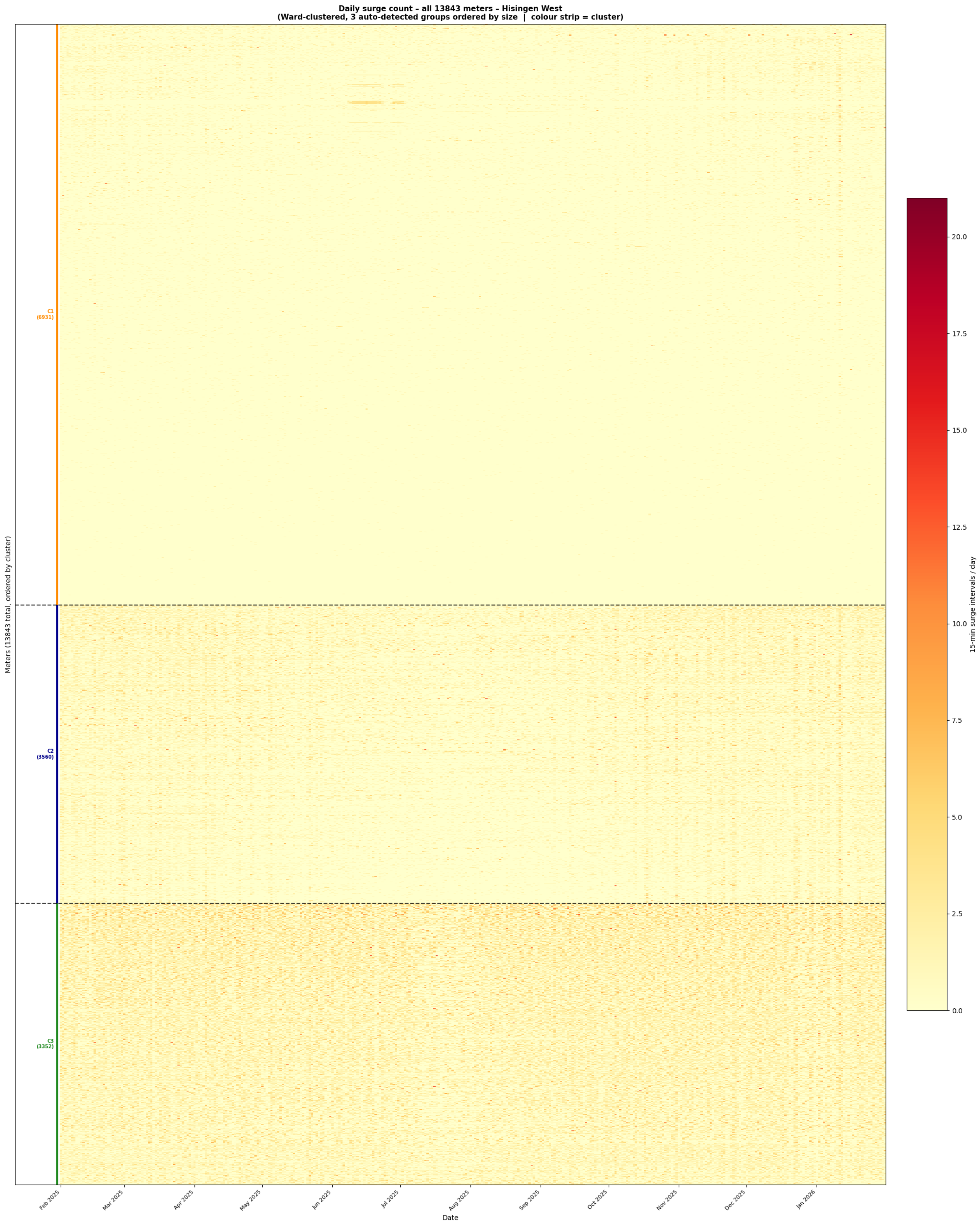

Figure A.15: Heatmap showcasing cluster capture of meter surge data

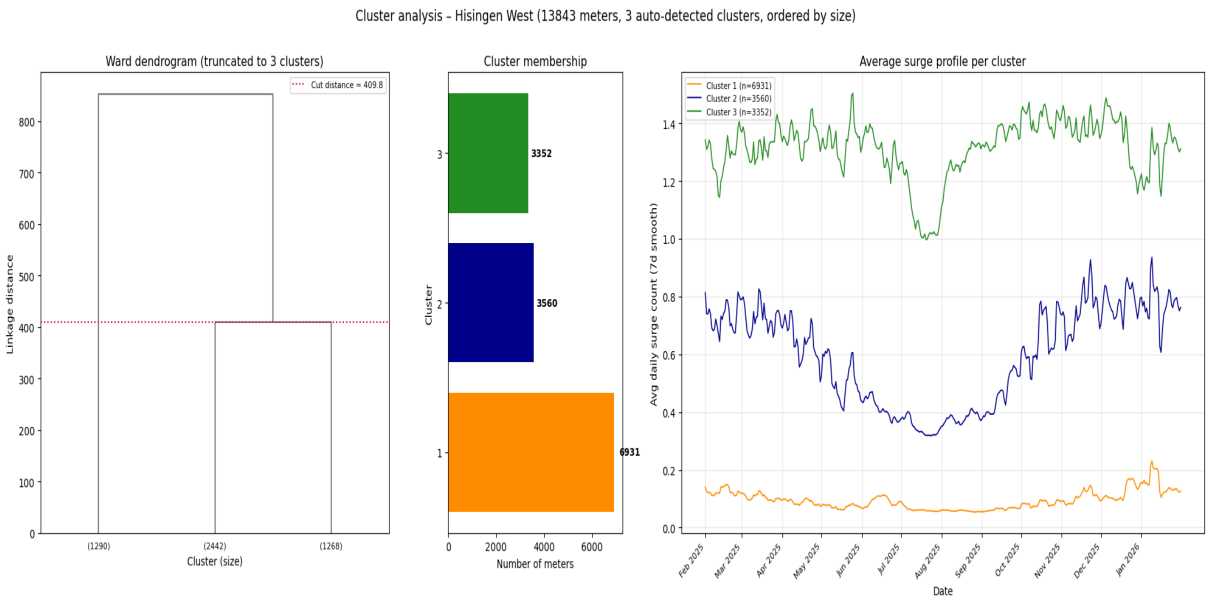

Figure A.16: Cluster analytics

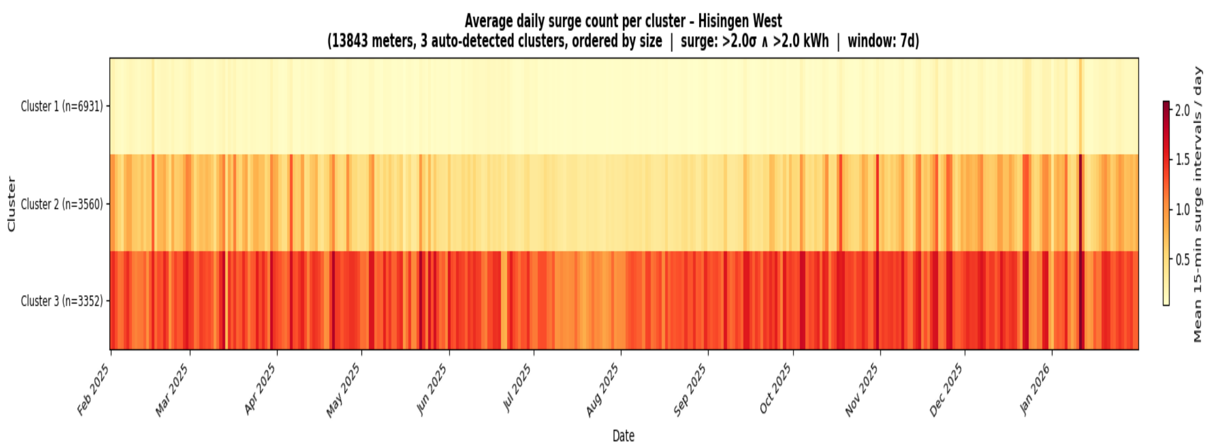

Figure A.17: Cluster heatmap showcasing surge occurrence

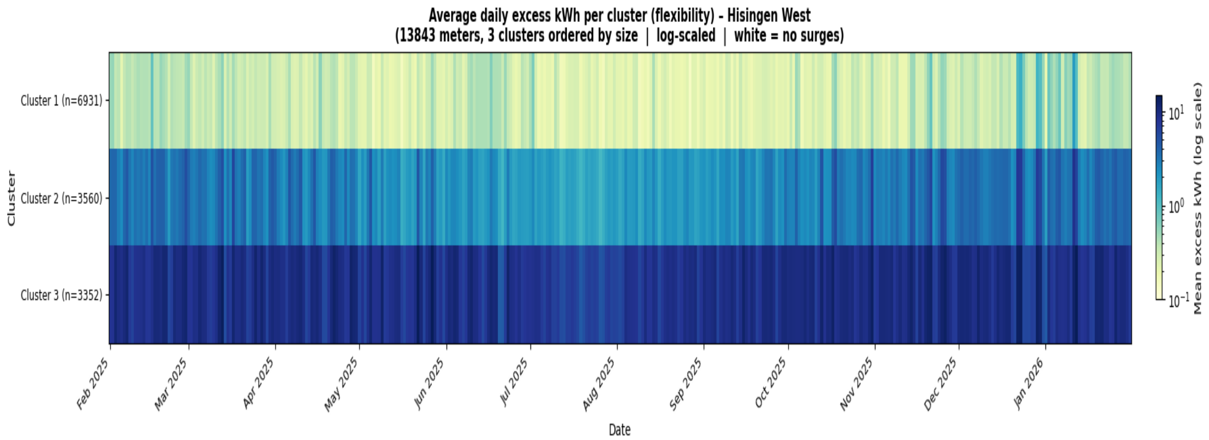

Figure A.18: Cluster heatmap showcasing excess kWh from surges

A.3 Northeast

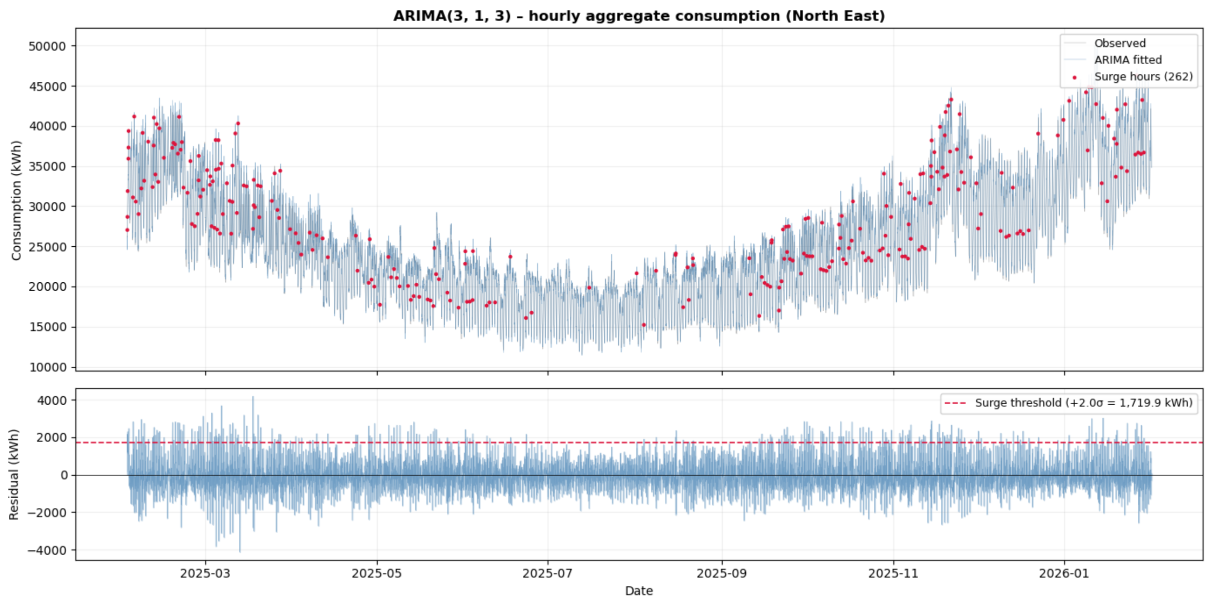

Figure A.19: ARIMA

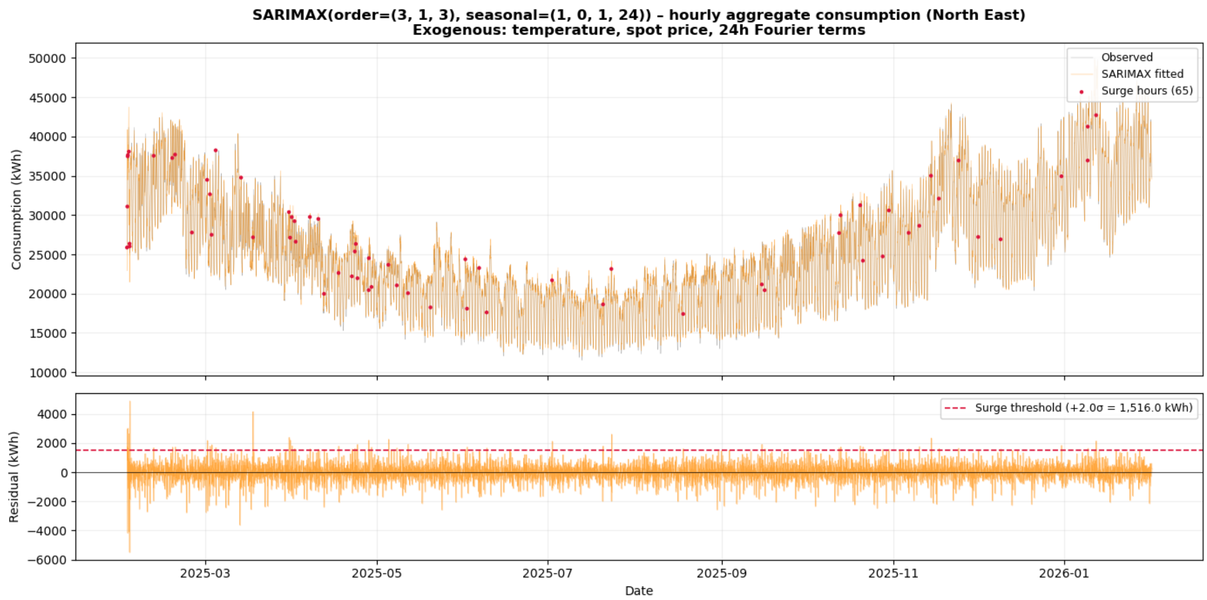

Figure A.20: SARIMAX

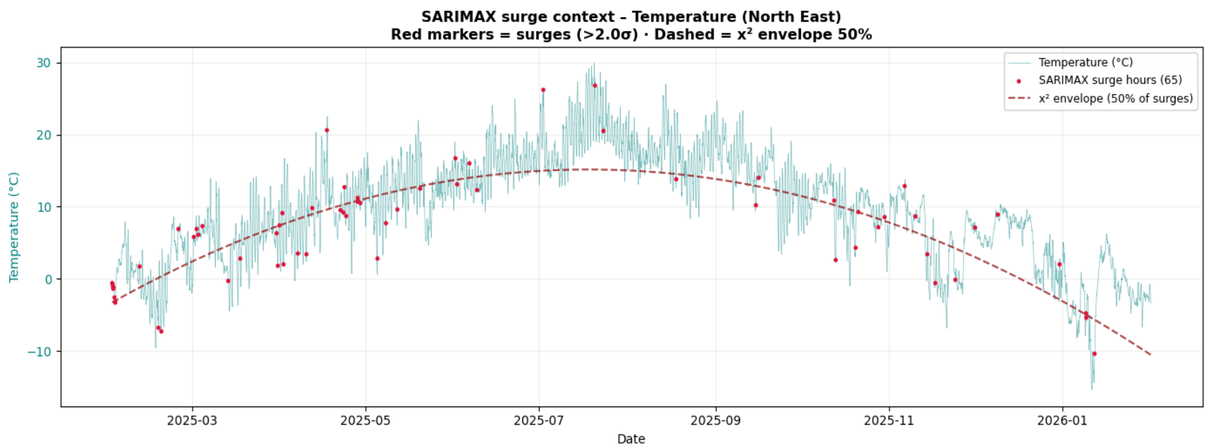

Figure A.21: Surges displayed over weather data

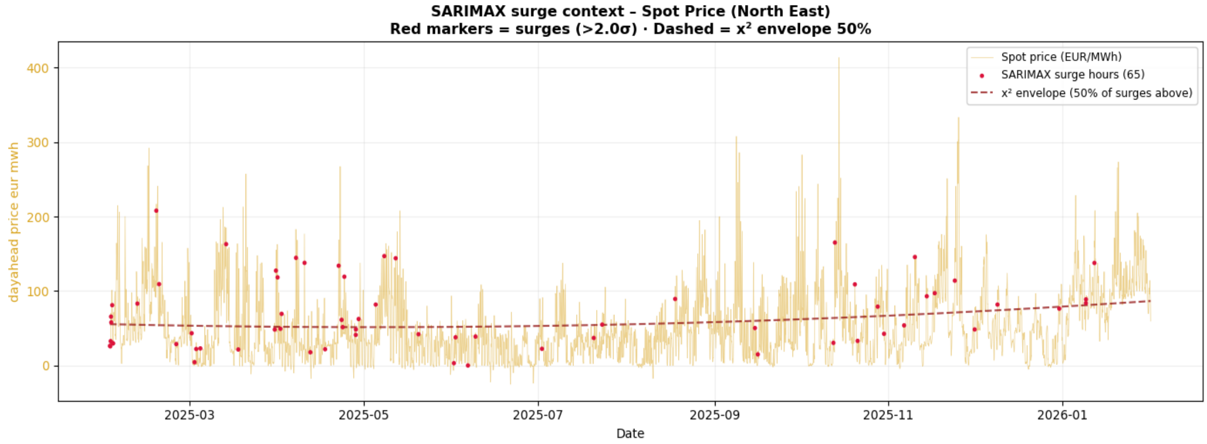

Figure A.22: Surges displayed over EPEX SPOT price data

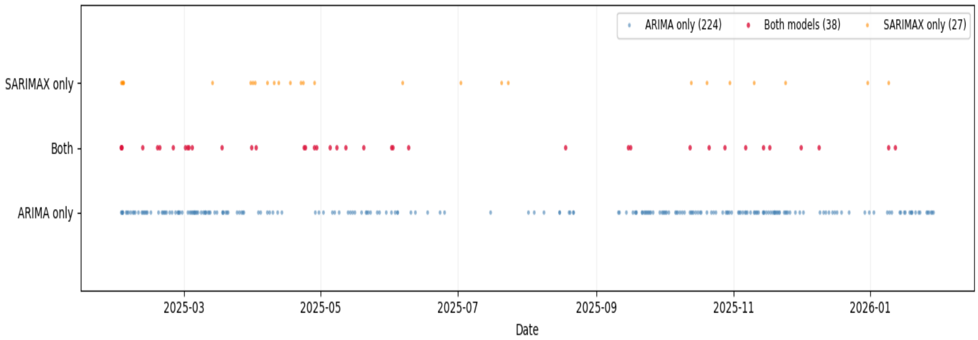

Figure A.23: ARIMA v. SARIMAX

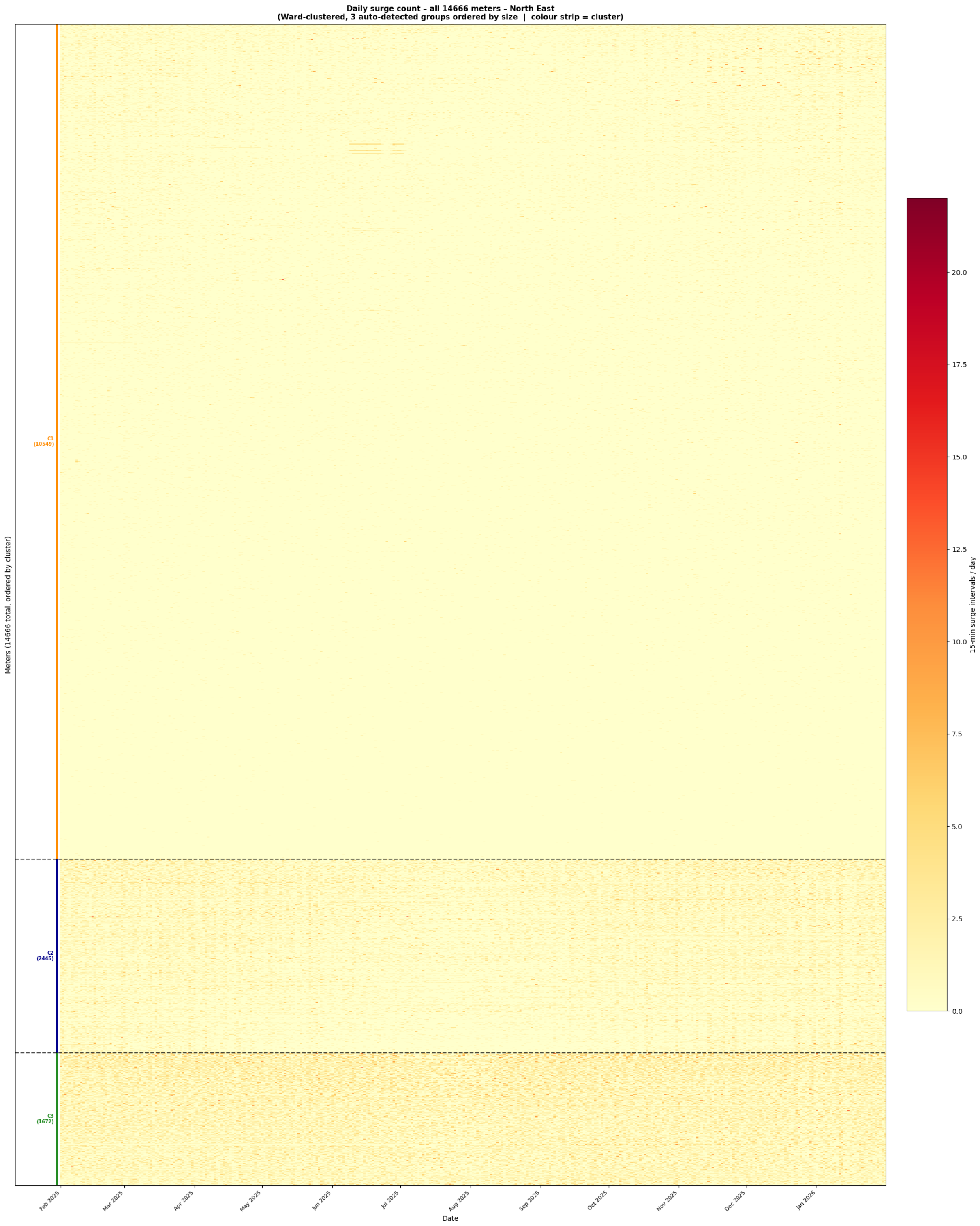

Figure A.24: Heatmap showcasing cluster capture of meter surge data

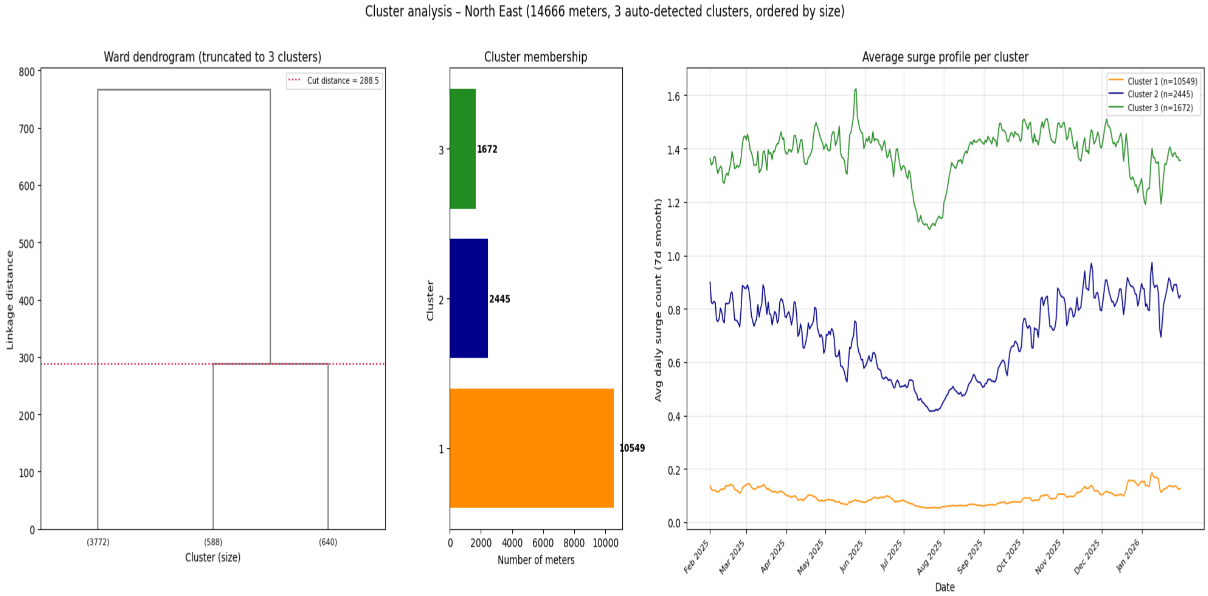

Figure A.25: Cluster analytics

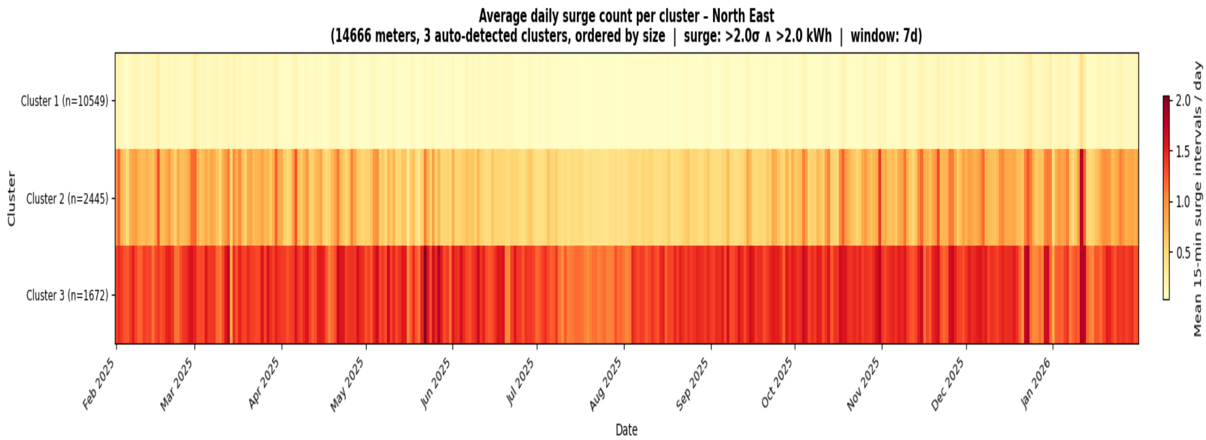

Figure A.26: Cluster heatmap showcasing surge occurrence

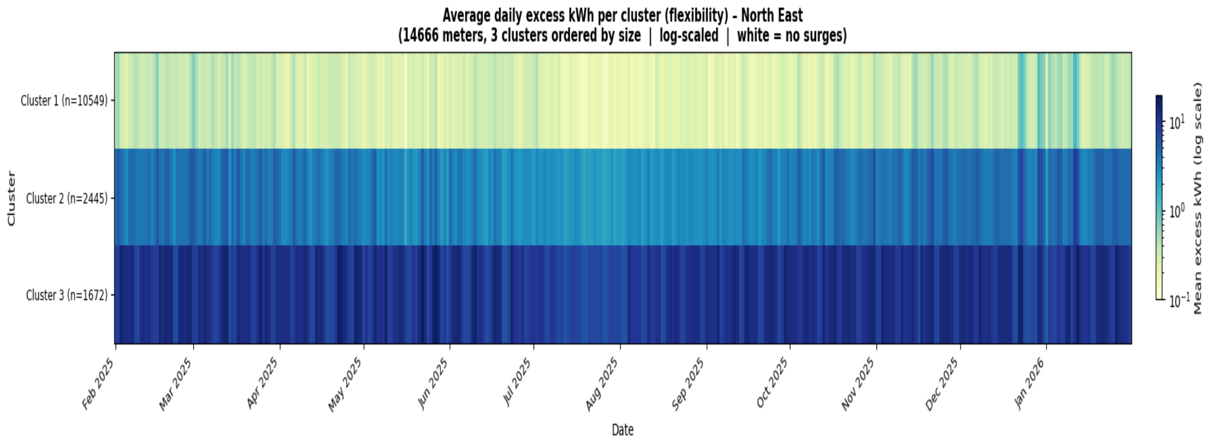

Figure A.27: Cluster heatmap showcasing excess kWh from surges

A.4 East

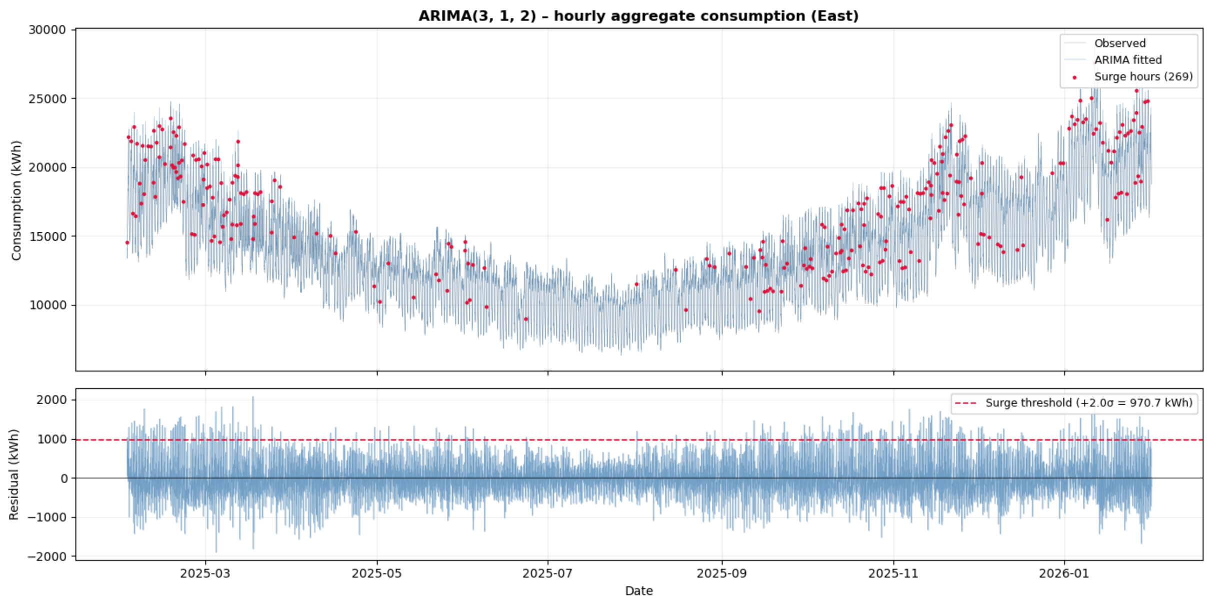

Figure A.28: ARIMA

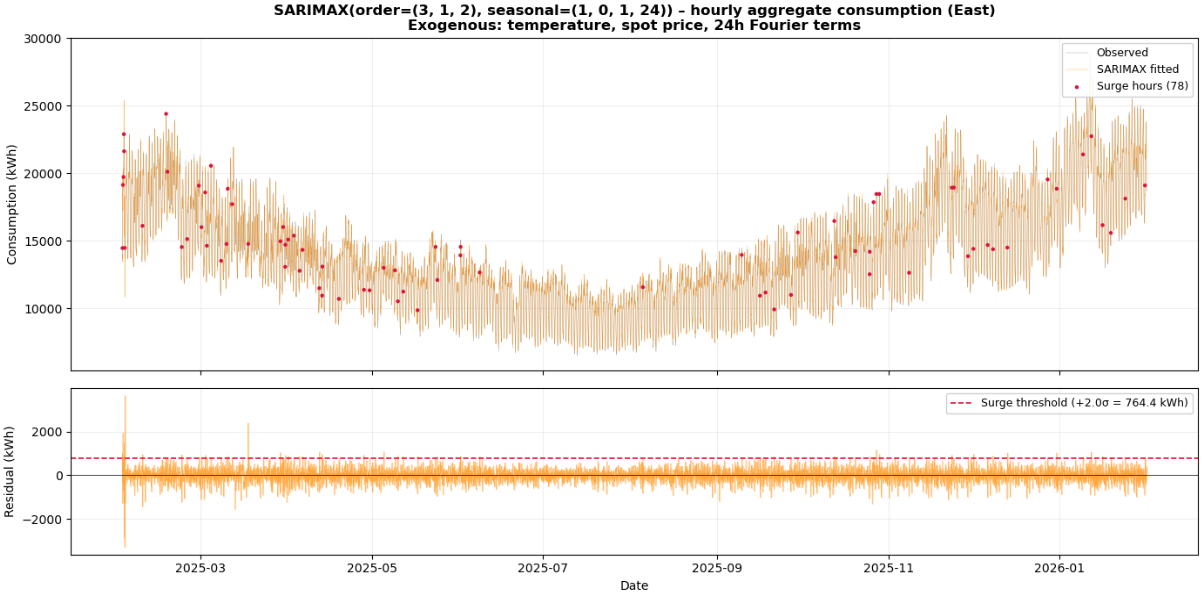

Figure A.29: SARIMAX

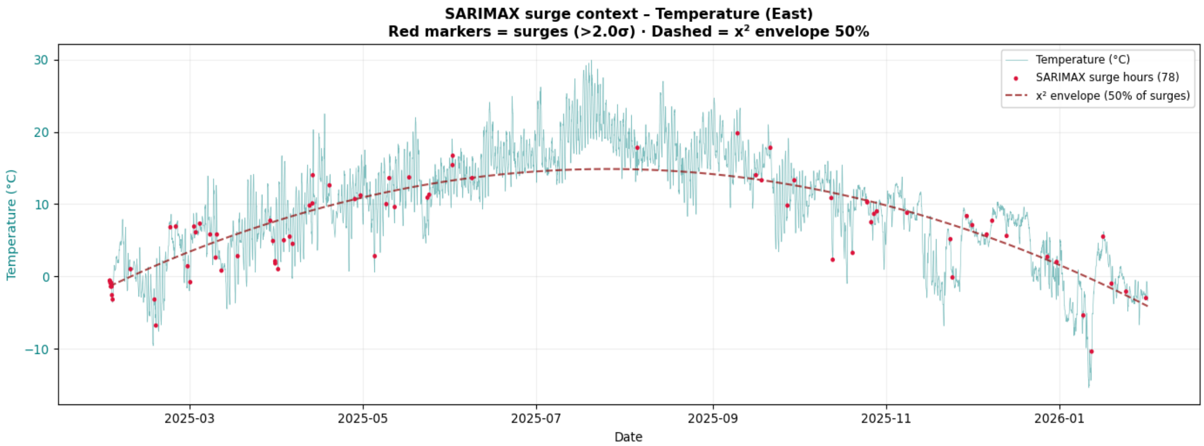

Figure A.30: Surges displayed over weather data

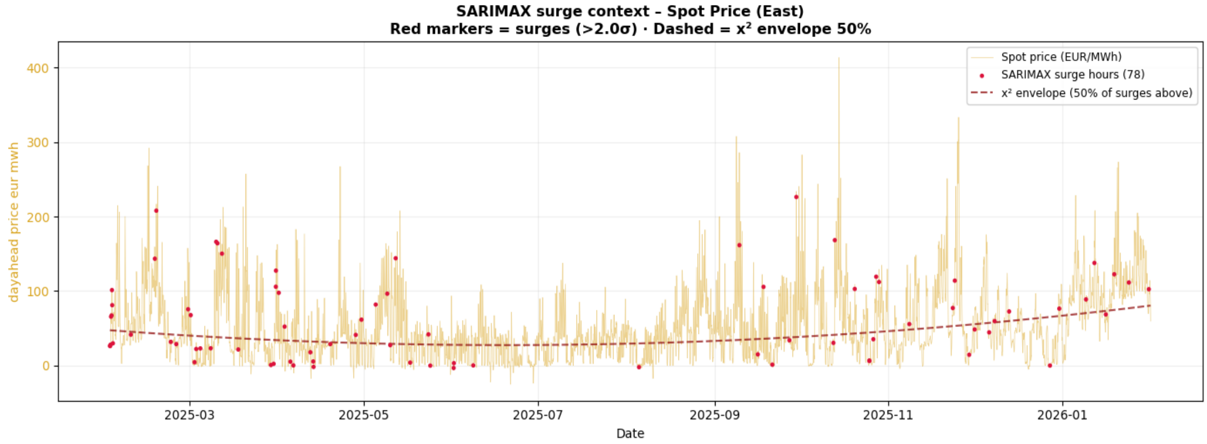

Figure A.31: Surges displayed over EPEX SPOT price data

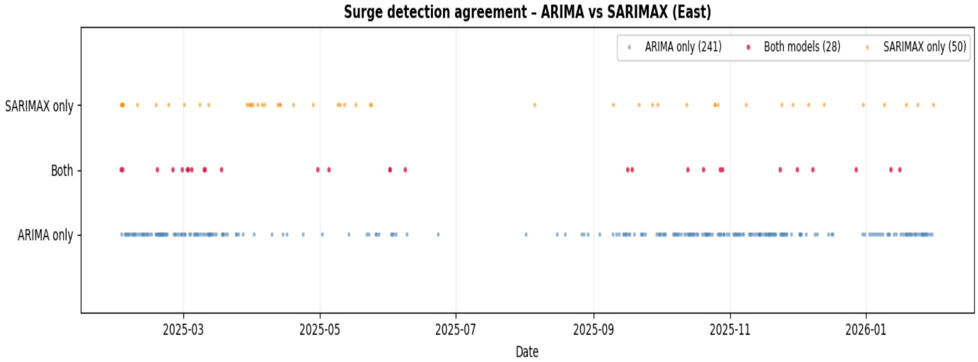

Figure A.32: ARIMA v. SARIMAX

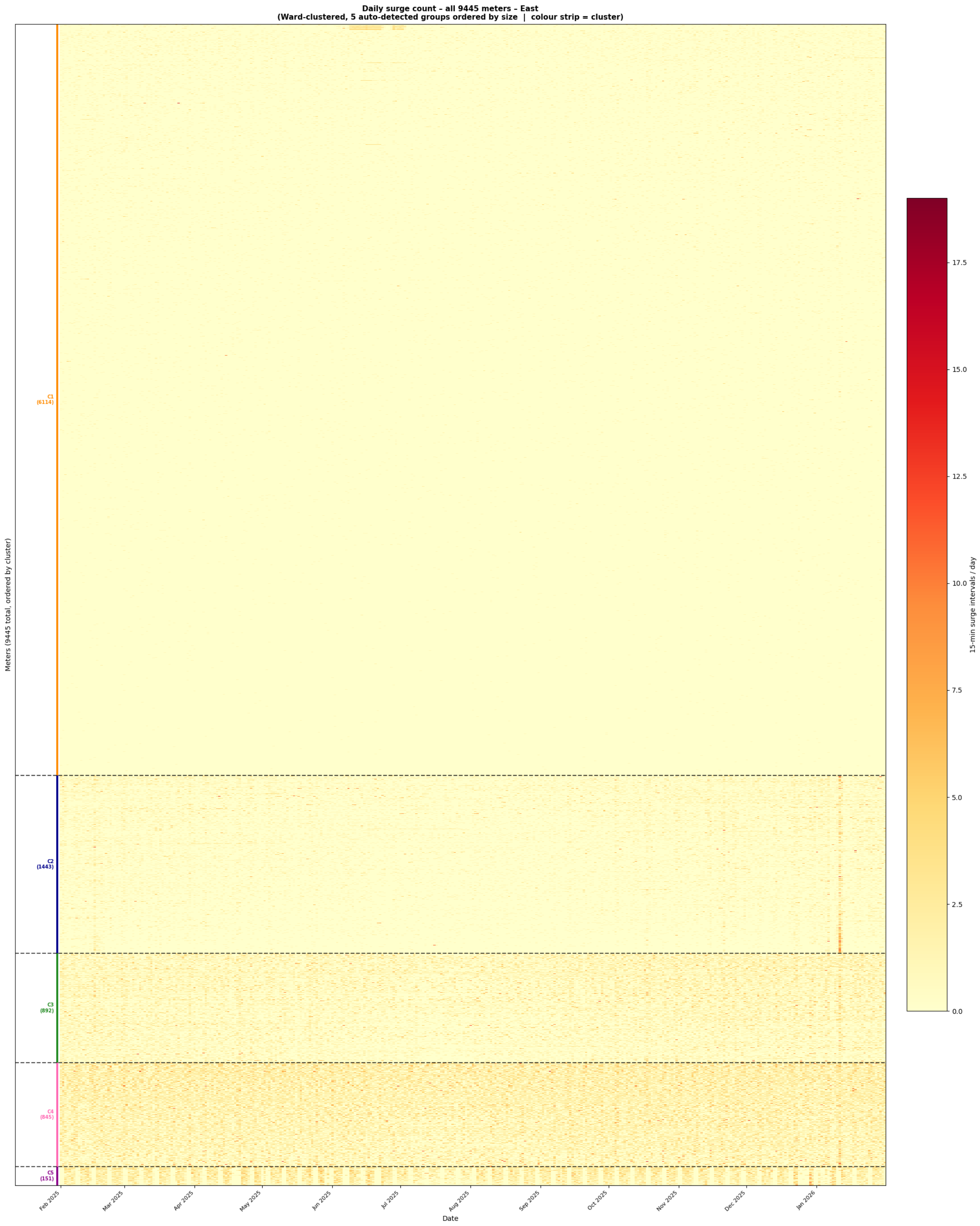

Figure A.33: Heatmap showcasing cluster capture of meter surge data

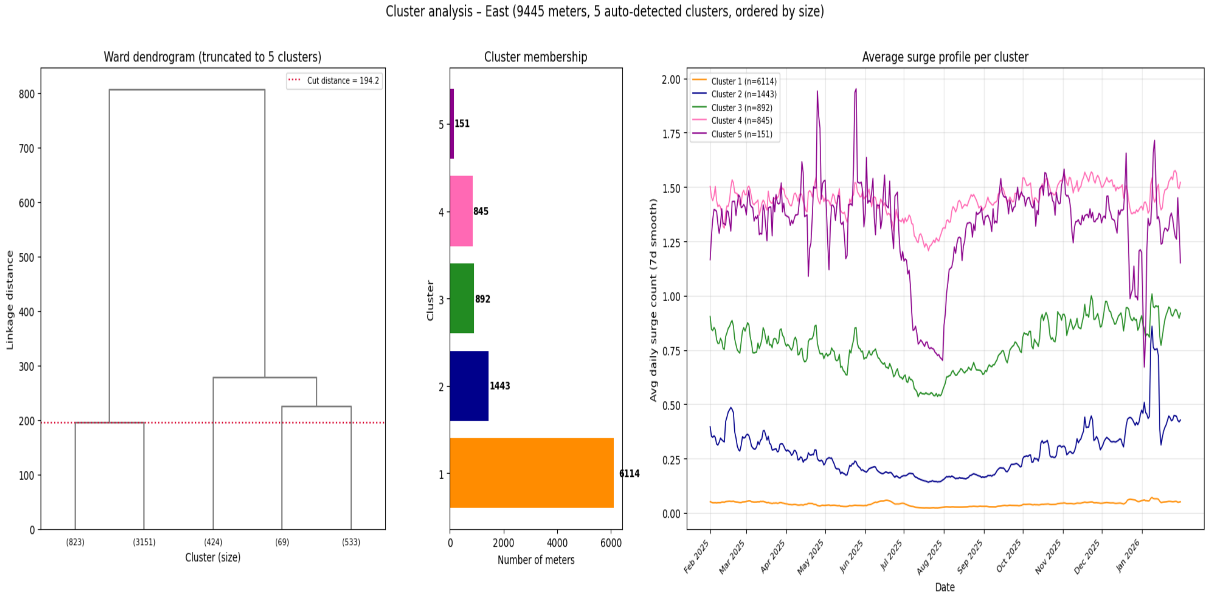

Figure A.34: Cluster analytics

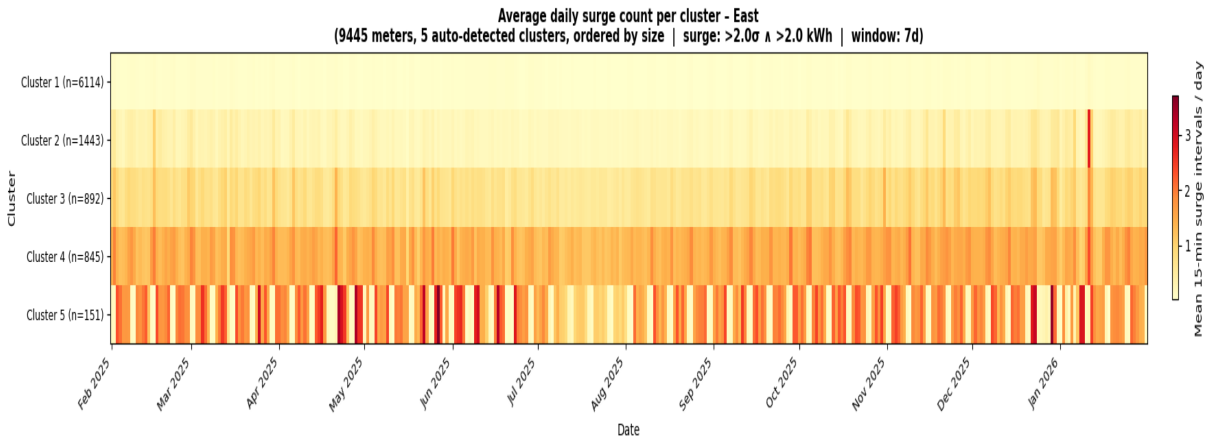

Figure A.35: Cluster heatmap showcasing surge occurrence

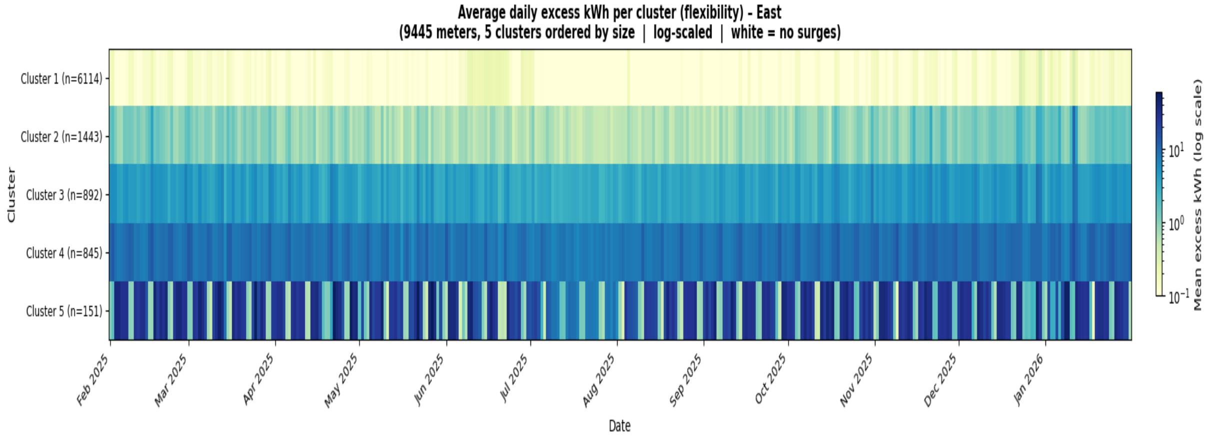

Figure A.36: Cluster heatmap showcasing excess kWh from surges

A.5 Southern Gothenburg

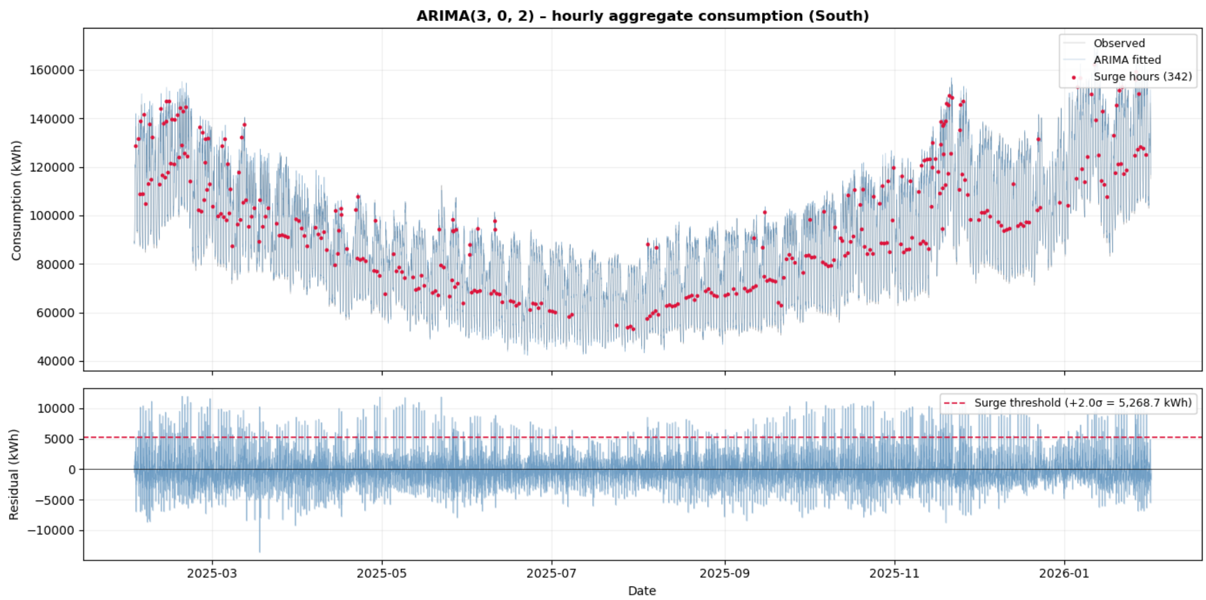

Figure A.37: ARIMA

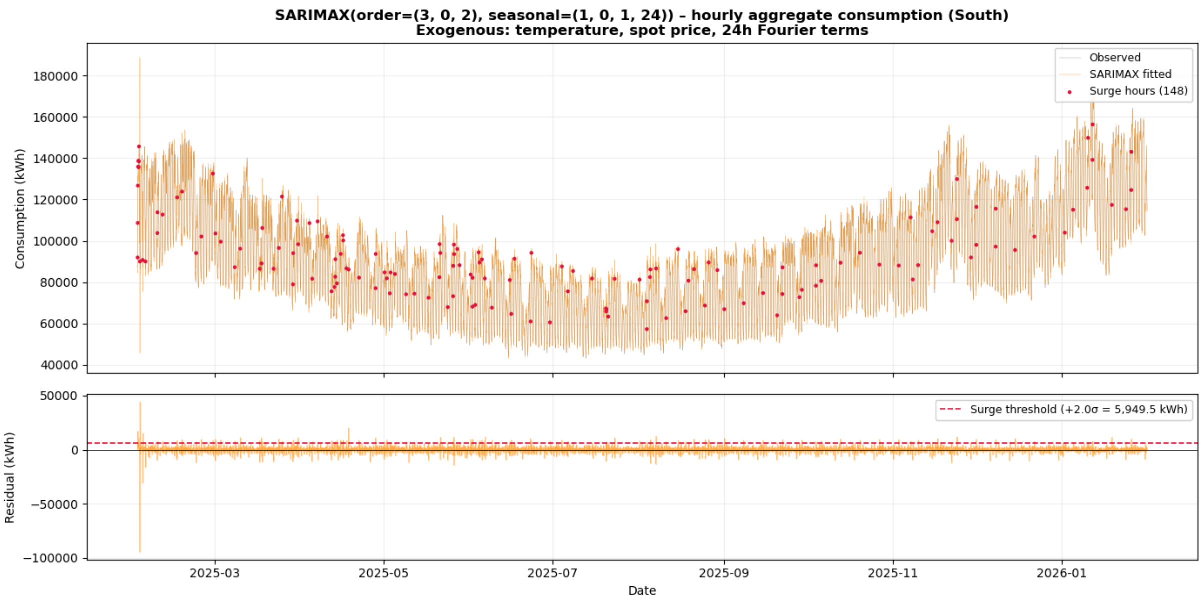

Figure A.38: SARIMAX

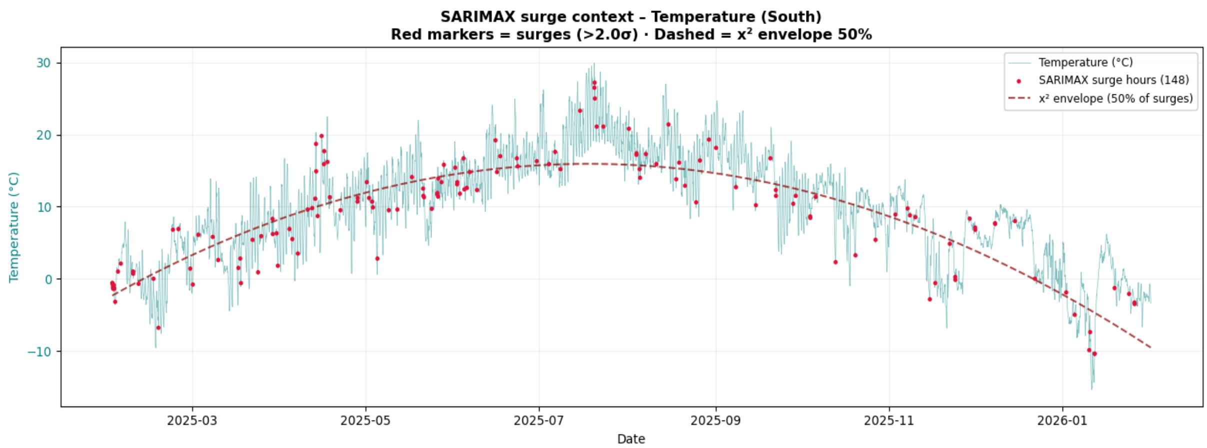

Figure A.39: Surges displayed over weather data

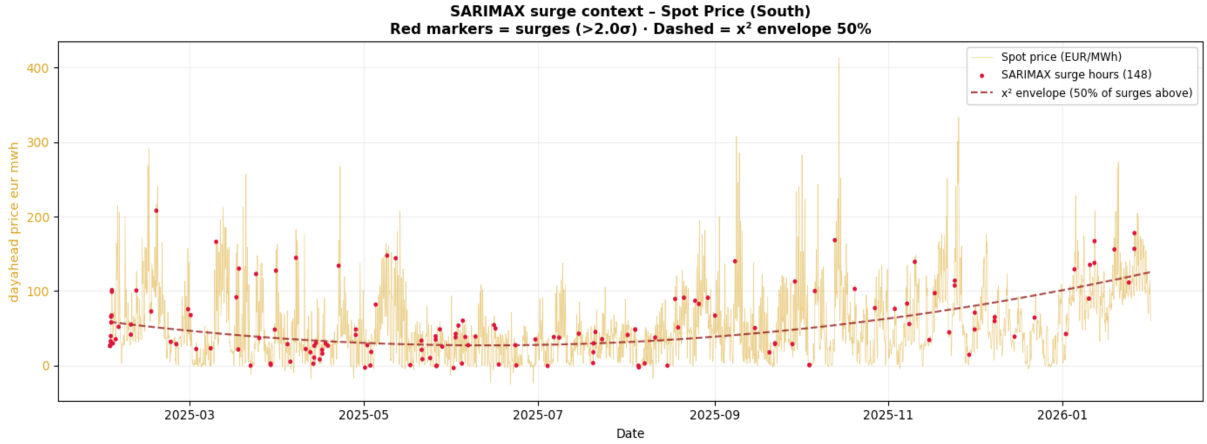

Figure A.40: Surges displayed over EPEX SPOT price data

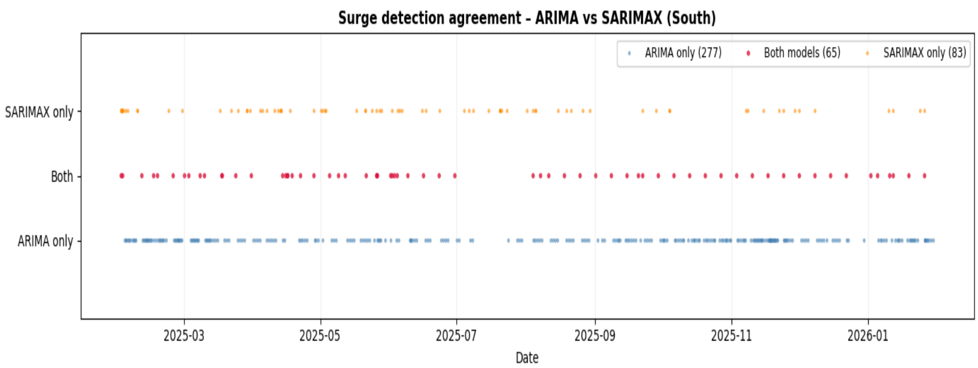

Figure A.41: ARIMA v. SARIMAX

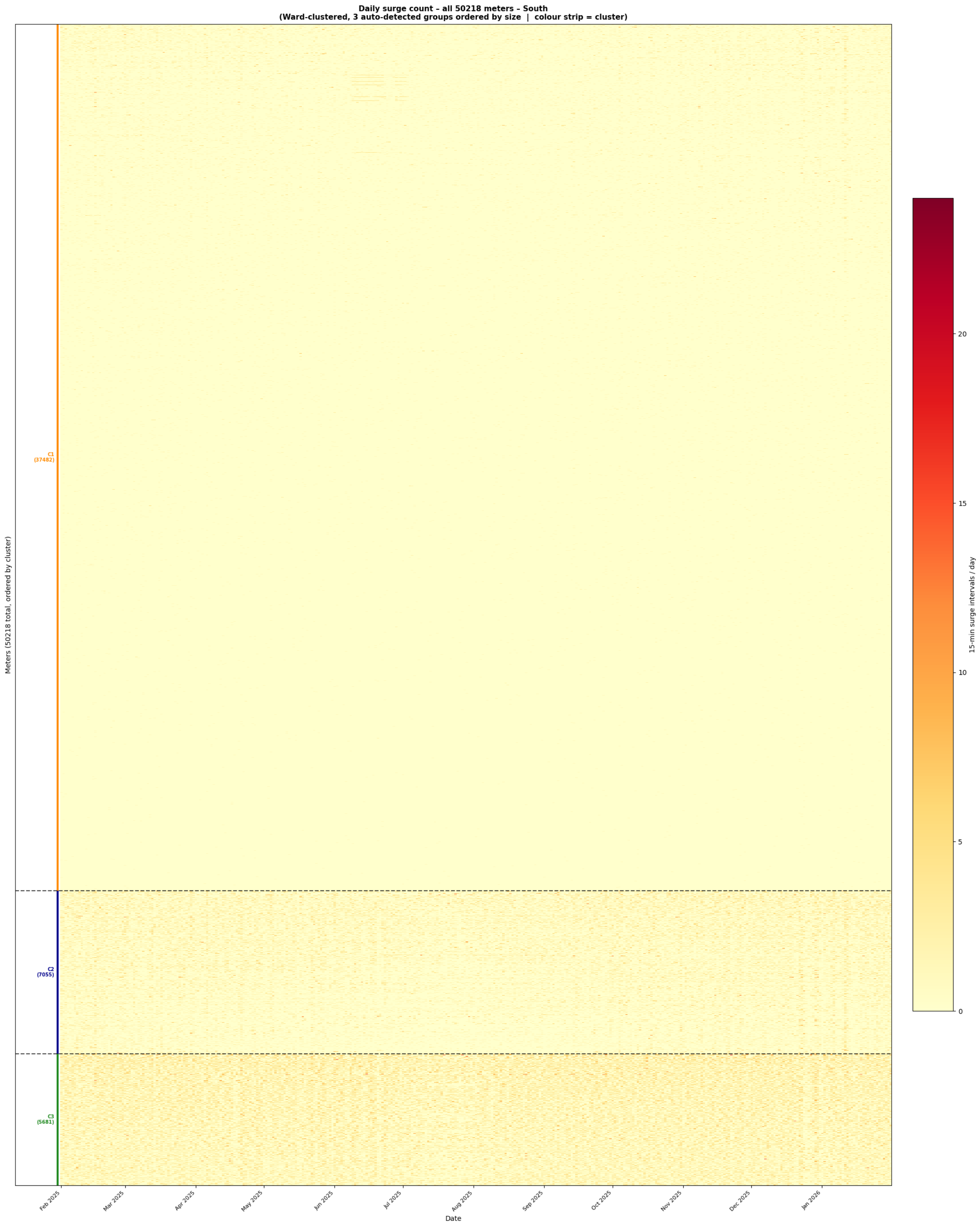

Figure A.42: Heatmap showcasing cluster capture of meter surge data

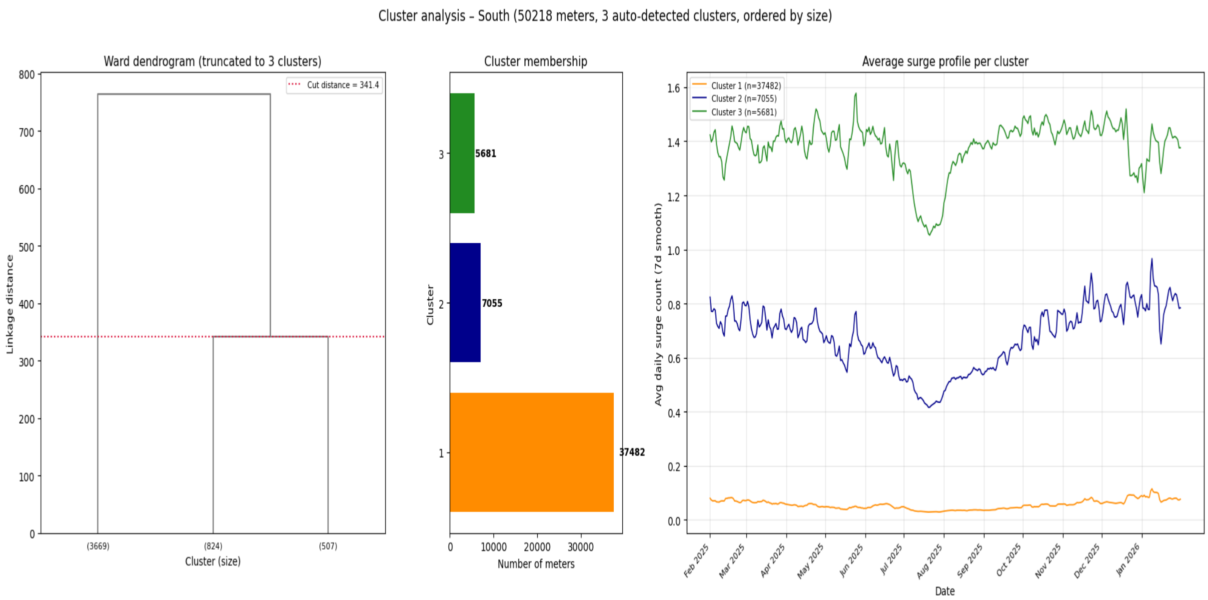

Figure A.43: Cluster analytics

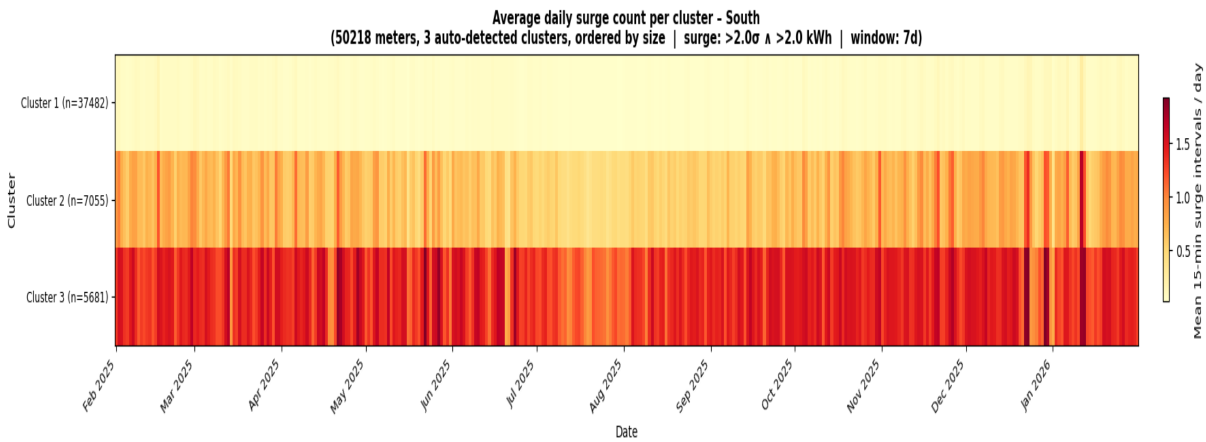

Figure A.44: Cluster heatmap showcasing surge occurrence

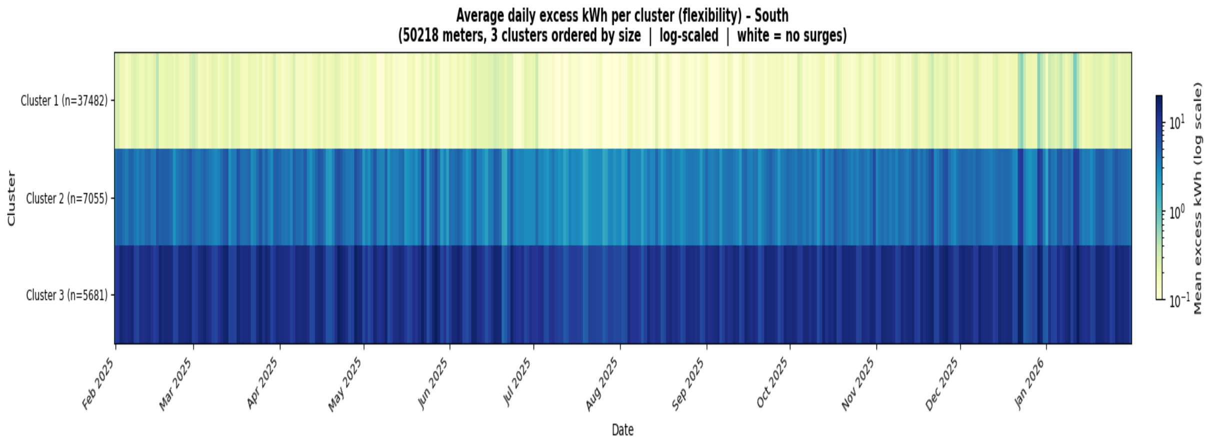

Figure A.45: Cluster heatmap showcasing excess kWh from surges

A.6 Central Gothenburg

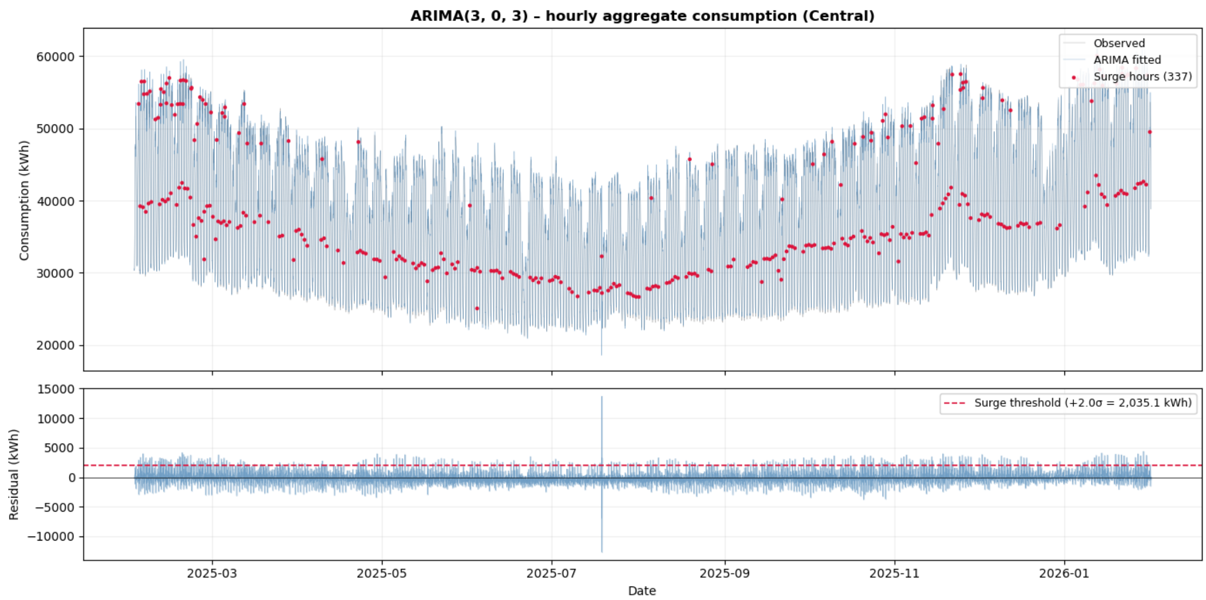

Figure A.46: ARIMA

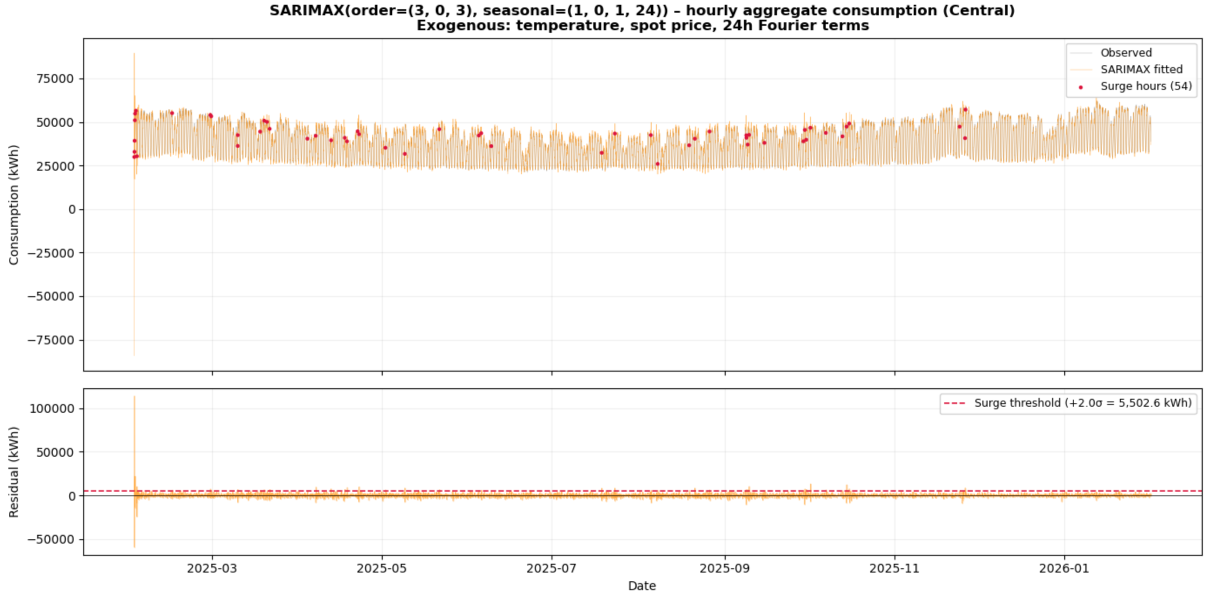

Figure A.47: SARIMAX

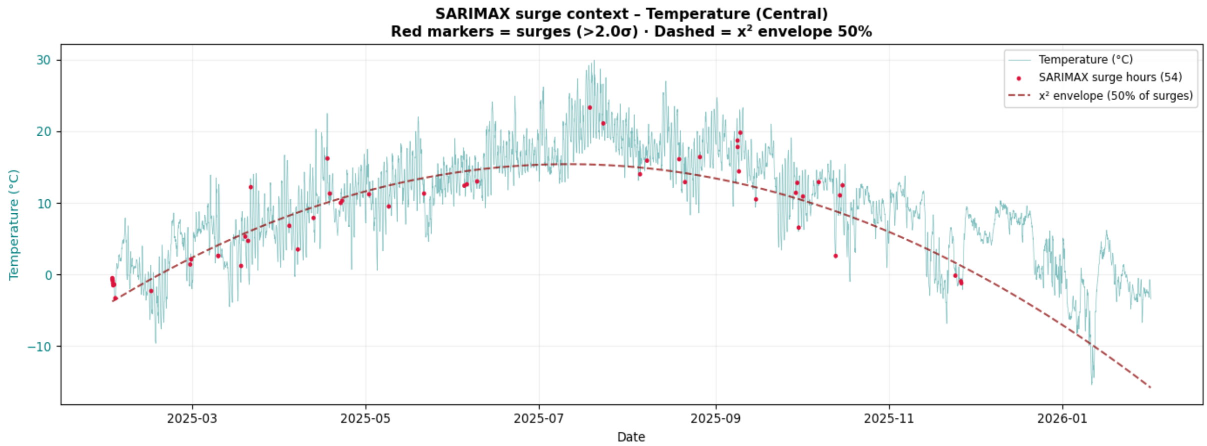

Figure A.48: Surges displayed over weather data

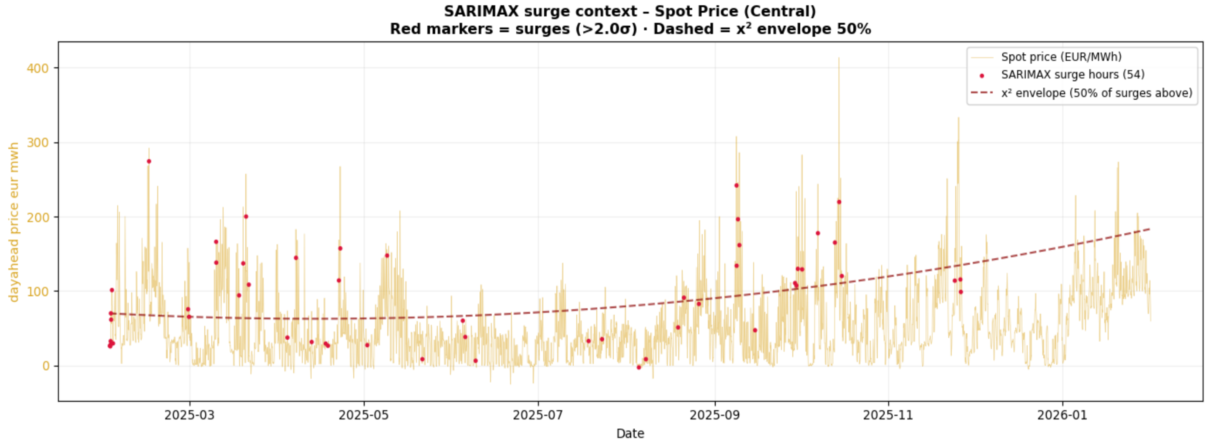

Figure A.49: Surges displayed over EPEX SPOT price data

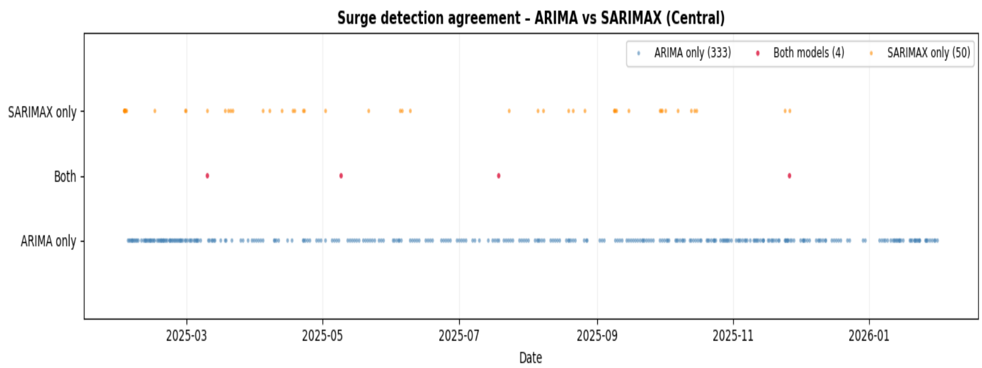

Figure A.50: ARIMA v. SARIMAX

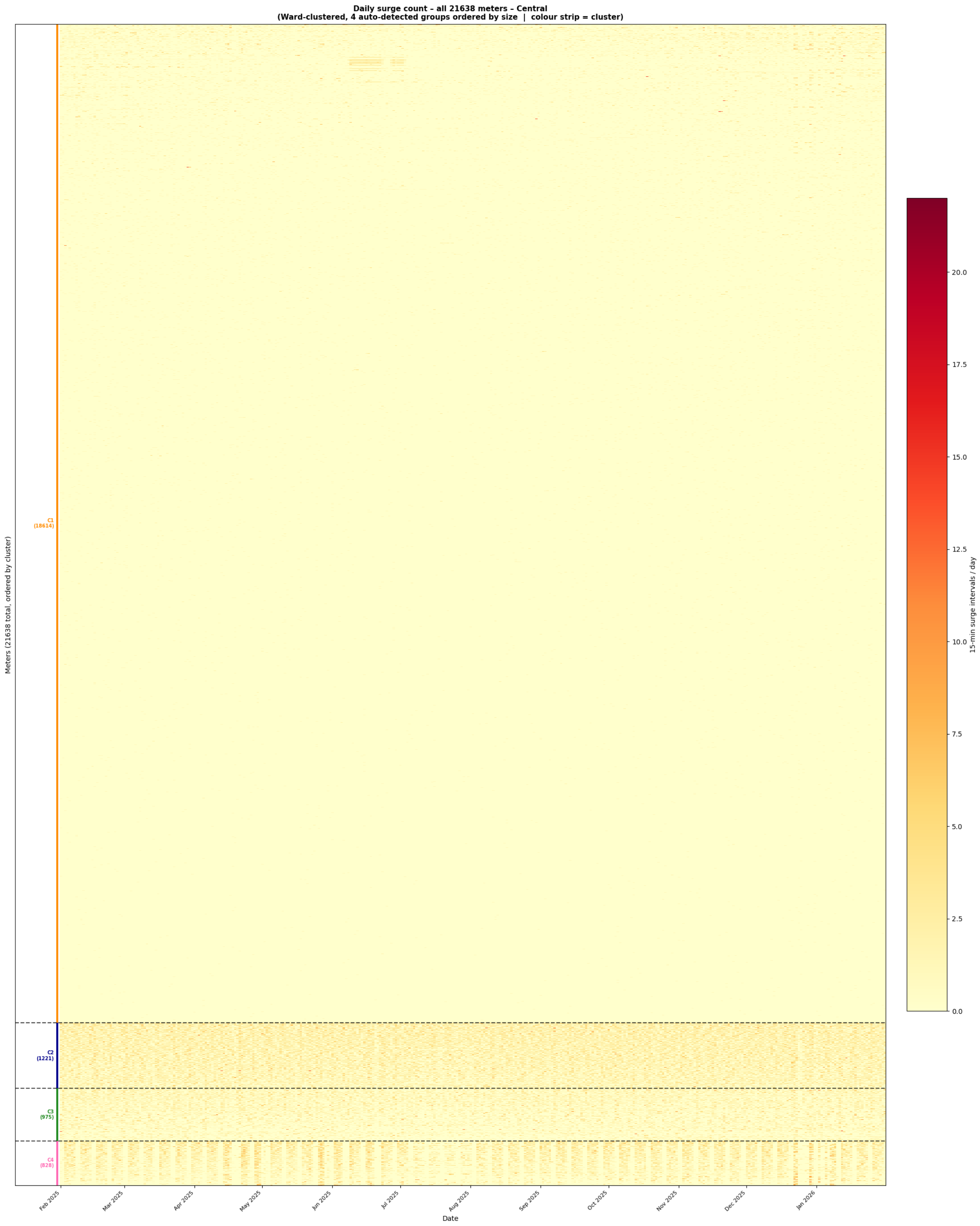

Figure A.51: Heatmap showcasing cluster capture of meter surge data

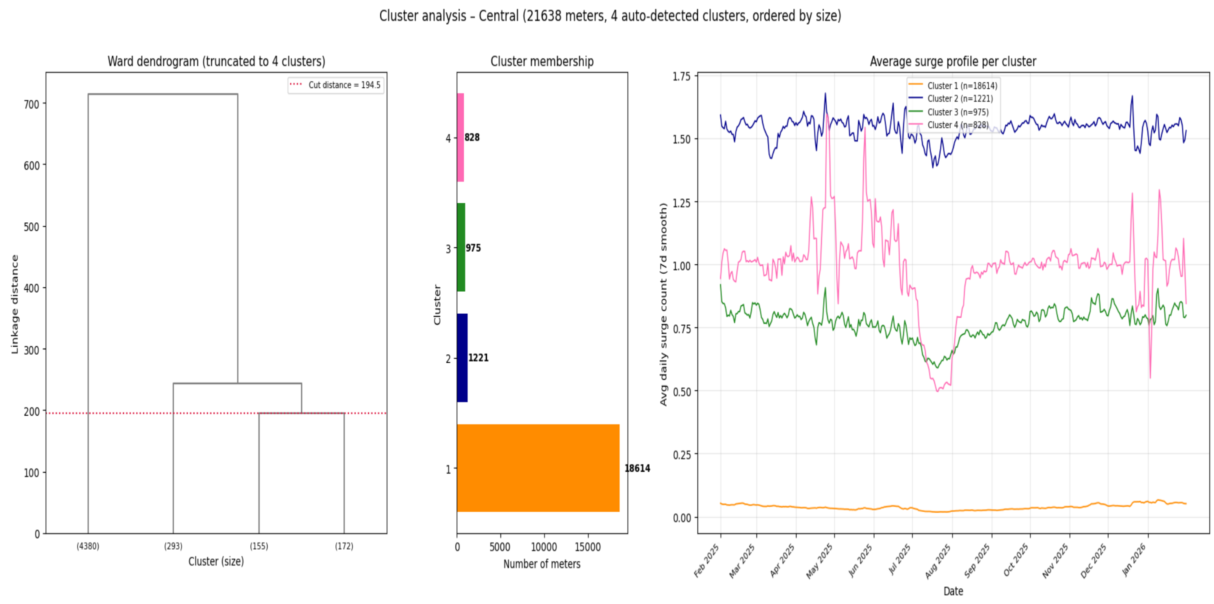

Figure A.52: Cluster analytics

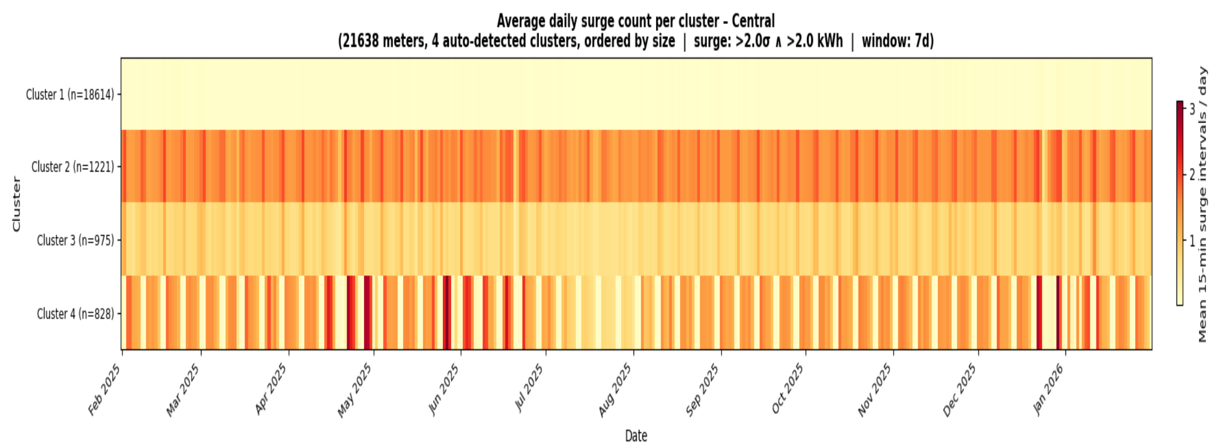

Figure A.53: Cluster heatmap showcasing surge occurrence

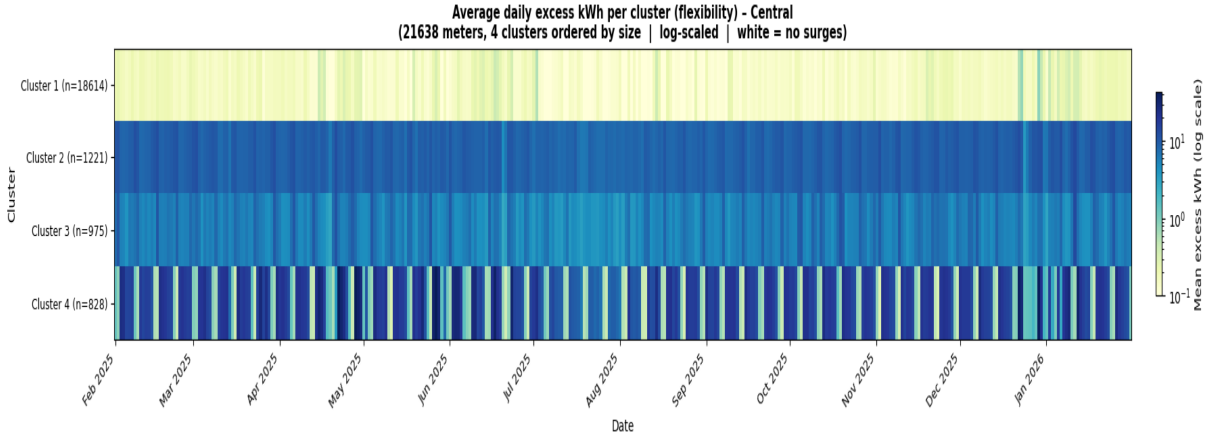

Figure A.54: Cluster heatmap showcasing excess kWh from surges