Data-Driven Detection of Financially Motivated Cyber Attacks in the Energy Market

Ghadeer Obaid Alsharif

SUBMITTED IN FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

School of Computing Science
College of Science and Engineering
University of Glasgow

June, 2026

Abstract

The increasing integration of information and communication technologies into modern power systems has enhanced operational efficiency while simultaneously exposing electricity markets to sophisticated cyber threats. Among these, financially motivated False Data Injection Attacks (FDIAs) targeting Locational Marginal Prices (LMPs) pose significant risks to market integrity, economic fairness, and grid stability. Unlike operational disruptions, stealthy LMP manipulation aims to preserve statistical normality while inducing economically advantageous distortions, thereby evading conventional residual-based bad data detection mechanisms. Existing defense strategies remain limited, often relying on stationarity assumptions, trusted infrastructure, or low-dimensional settings that are inconsistent with real-world electricity markets. This thesis develops a comprehensive data-driven framework for detecting stealthy and economically motivated cyber-attacks in wholesale electricity markets. First, a physics-consistent synthetic benchmarking framework (SMLT) is constructed to systematically model and reproduce stealthy LMP manipulation scenarios under varying attack intensities and knowledge assumptions. The dataset enables quantitative characterization of statistical, temporal, and spatial signatures of manipulation, providing a controlled foundation for evaluating detection methods. Second, an incremental, unsupervised change-point detection framework is proposed for near–real-time monitoring of streaming LMP signals. The method balances detection sensitivity, false alarm stability, and bounded delay within short market settlement intervals. Third, to address the inherent non-stationarity of electricity markets, a drift-aware anomaly detection framework is introduced that integrates concept drift detection with adaptive anomaly scoring, enabling robust discrimination between natural regime shifts and adversarial manipulation. Finally, spatial dependencies among nodal prices are modeled using graph signal processing techniques, allowing physics-informed dimensionality reduction, scalable monitoring, and anomaly localization with quantifiable spatial accuracy. Extensive evaluation on both synthetic and real-world market datasets demonstrates that the proposed framework achieves improved detection robustness, reduced false positives, controlled detection delay, and enhanced scalability compared to existing baselines. By jointly addressing stealthiness, non-stationarity, real-time constraints, and spatial structure, this thesis advances the state of the art in energy market cybersecurity and establishes a principled foundation for adaptive, data-driven protection against financially motivated cyber-attacks in electricity markets.

Contents

List of Tables

List of Figures

Acknowledgements

It has been ten years since I began two long journeys at the same time: motherhood and higher education. Looking back, it is striking how much they mirror each other; sharing the same cycles of challenge and growth, joy and stress, hardship and reward. At times it felt overwhelming and almost impossible, yet in hindsight it has passed surprisingly quickly.

I would like to express my sincere gratitude to my supervisor, Dr. Christos Anagnostopoulos, for his unwavering guidance and support throughout my PhD journey. His kindness, encouragement, and consistently approachable manner made the process not only manageable but genuinely meaningful and, at times, truly enjoyable. I am also deeply grateful to my former supervisor, Dr. Angelos Marnerides, who was the first person I worked with at the University of Glasgow. Although he has since left the university, his early guidance and support have continued to shape and resonate throughout my academic journey. I would also like to thank my examiners, Dr. José María Maza Ortega and Dr. Michele Sevegnani, for their time, careful evaluation, and valuable feedback.

My appreciation also extends to my colleagues and friends at the University of Glasgow. In particular, I would like to thank my friends Dr. Tahani Aladwani and Dr. Nawal Alanazi for the many conversations and coffee breaks that often went far beyond research and PhD discussions, and for their constant encouragement, support, and friendship throughout this journey.

Most importantly, I would like to express my deepest gratitude to my husband, Sultan, for his patience, understanding, and unwavering support throughout this long journey. His constant encouragement and belief in me have been a steady source of strength through every stage of this work. My heartfelt love goes to my children, Safana, Naif, and Abdullelah. I am deeply sorry for the time I had to take away from them during this PhD journey, and I am endlessly grateful for their patience, love, and the joy they bring to my life each and every day.

I also extend my sincere thanks to my mother, sisters, and brothers for their endless prayers, love, and support, even from afar. Their presence in my life, despite the distance, has been a constant source of comfort and strength throughout this journey.

كل ما كنته وسأكونه ليس إلا بفضل ما عهدتماني به من عظيم
بركما وإحسانكما بي ودعائكما لي..

أدعوا الله أن يجعل أجر هذا العمل في ميزان حسناتكما..

إلى أمي الحبيبة:

وضحى بنت ناصر بن عبد المحسن

إلى الغائب الحاضر في روحي ووجداني

والدي: عبيد بن فهيد بن شاهين

أهدي هذه الرسالة

Chapter 1

Introduction

1.1 Motivation

The rapid growth of modern power systems, driven by the integration of information and communication technologies (ICT) into power infrastructure, has significantly enhanced control and monitoring capabilities, thereby reshaping the dynamics of the energy trading ecosystem. This transformation enables more efficient operations and streamlined energy management. However, increased reliance on ICT also introduces new vulnerabilities, creating opportunities for adversaries to exploit these systems through stealthy FDIAs. Such sophisticated attacks can bypass conventional security mechanisms, remain undetected, and manipulate critical system data. Consequently, attackers may influence energy trading outcomes while maintaining long-term stealth to maximize financial gains, posing serious risks to market stability.

The problem of Locational Marginal Price (LMP) manipulation via FDIAs has attracted growing research attention in recent years, with most studies focusing on the design of attack strategies rather than detection or mitigation mechanisms, as surveyed in [1,2]. Existing research has extensively developed attacks capable of manipulating LMPs under limited system knowledge and restricted access, while avoiding the triggering of traditional bad data detection schemes. In contrast, defense strategies against LMP manipulation remain comparatively underdeveloped. Current protection approaches largely emphasize securing selected meters or optimizing phasor measurement unit (PMU) placement, often under strong assumptions of trusted infrastructure. Although data-driven detection methods have been proposed, many rely on stationarity or independent and identically distributed (i.i.d.) assumptions, which are rarely satisfied in realistic electricity market environments.

Detecting Locational Marginal Price (LMP) manipulation presents a set of intertwined technical and practical challenges that substantially complicate the design of reliable detection frameworks. First, the inherently stealthy nature of market manipulation makes it particularly difficult to identify using conventional anomaly detection techniques. Adversaries typically craft FDIAs with small magnitudes that remain within operational tolerances, thereby avoiding significant

shifts in observable statistical properties. Because these perturbations are carefully calibrated to preserve the overall distributional characteristics of the data, traditional residual-based methods and distribution-sensitive anomaly detectors often fail to distinguish malicious manipulations from normal market fluctuations. Second, the non-stationary behavior of power systems and electricity markets further exacerbates the problem. Market data are continuously influenced by fluctuating demand patterns, renewable generation variability, network reconfigurations, regulatory interventions, and strategic bidding behavior. These factors induce concept drift, meaning that the underlying data-generating process evolves over time. As a result, detection models trained on historical data may become obsolete or exhibit degraded performance if not properly adapted, leading to increased false positives or false negatives. Third, electricity markets operate with high-dimensional data streams, incorporating nodal prices, load profiles, generation dispatch levels, congestion patterns, and network states, while market clearing occurs at short intervals (e.g., every 5–15 minutes). This combination of dimensionality and temporal granularity imposes significant computational and scalability constraints. Detection mechanisms must operate in near real time, process large volumes of correlated features, and maintain robustness without incurring prohibitive computational overhead. Finally, the scarcity of publicly available, high-quality datasets that include labeled instances of LMP manipulation further limits methodological development and benchmarking. Due to confidentiality, security concerns, and the rarity of confirmed attack events, researchers often rely on simulated environments, which may not fully capture the complexity and adversarial dynamics of real-world markets.

1.2 Limitations of Existing Detection Paradigms

Existing approaches for detecting LMP manipulation often rely on several assumptions that limit their effectiveness in realistic electricity market environments:

1. ***Dependence on model-based system monitoring:*** Many FDIA detection methods are designed within the context of power system state estimation, where anomalies are identified by validating measurements against physical system models such as power flow equations. Under this paradigm, detection relies on identifying inconsistencies between observed measurements and model-based predictions. However, electricity market monitoring operates at a different layer of the system. Market monitoring primarily focuses on economic signals such as prices, cleared quantities, and congestion patterns rather than direct validation of measurement consistency. As a result, detection frameworks built around model-based system monitoring may fail to identify manipulation that manifests primarily through market outcomes while remaining consistent with physical system models.

2. ***Trusted measurement infrastructure:*** Many protection-oriented approaches rely on the belief that certain measurement devices, particularly PMUs, can be trusted and are immune to compromise. Under this perspective, protecting a subset of critical meters or deploying secure PMUs is considered sufficient to prevent manipulation of the system state. However, PMUs rely on GPS signals for time synchronization, which can be vulnerable to spoofing attacks. As a result, even protected measurement infrastructures may still be exposed to manipulation through compromised measurements. Moreover, market manipulation may also originate from other sources beyond physical measurements, such as malicious bidding behavior through market gateways or adversarial manipulation of forecasting inputs (e.g., load or renewable generation forecasts). These additional attack surfaces indicate that protecting physical measurement infrastructure alone is insufficient to prevent economically motivated manipulation in electricity markets.

3. ***Stationary and i.i.d. market data:*** A number of detection methods assume that the monitoring data are independent and identically distributed and generated under stable statistical conditions. In practice, electricity markets exhibit strong temporal dependencies and evolving operational regimes. Demand fluctuations, renewable generation variability, network reconfigurations, and strategic bidding behavior continuously influence the statistical properties of market signals. This limitation leads to two challenges. First, stealthy manipulation of LMPs may preserve the overall distributional characteristics of the data, making it difficult for parametric or distribution-dependent detection methods to distinguish manipulated prices from normal market fluctuations. Second, electricity markets evolve over time, introducing concept drift where the underlying data-generating process changes across different operational regimes. As a result, models trained on historical data may gradually lose reliability if they do not explicitly adapt to these long-term structural changes.

1.3 Thesis Statement

This thesis claims that Locational Marginal Price (LMP) manipulation through stealthy False Data Injection Attacks (FDIAs) cannot be reliably detected using conventional residual-based or distribution-dependent anomaly detection methods due to their reliance on stationarity and i.i.d. assumptions in inherently non-stationary, high-dimensional electricity market environments. To address this gap, the thesis develops an adaptive, data-driven detection framework that explicitly accounts for concept drift, temporal market dynamics, and computational scalability constraints, thereby enabling robust and near–real-time identification of economically motivated stealth attacks without assuming trusted infrastructure or complete system knowledge.

1.4 Research Questions

This thesis addresses the following research questions:

RQ1: What measurable statistical and spatial signatures characterize stealthy LMP manipulation, and how can these signatures be systematically reproduced to benchmark detection models?

RQ2: How can change-point detection methods enable reliable identification of stealthy LMP manipulation in streaming electricity market data?

RQ3: How can anomaly detection models remain reliable under non-stationary LMP signals, where natural regime shifts and adversarial manipulations coexist?

RQ4: How can spatial dependencies among LMP signals be exploited to improve scalable anomaly detection and spatial localization in large-scale electricity markets?

1.5 Contributions

This thesis makes significant contributions to the field of energy market cybersecurity, specifically in the data-driven detection of financially motivated cyber-attacks that manipulate Locational Marginal Prices (LMPs). The work advances the state of the art in market-level anomaly detection by addressing stealthiness, real-time monitoring, non-stationarity, scalability, and localization. The key contributions are summarized as follows:

1- Establishing a Benchmarking Framework for Stealthy LMP Manipulation

- Developed a physics-consistent synthetic data framework to systematically model stealthy FDIAs scenarios targeting LMP calculations under varying attack vectors and intensities.
- Quantitatively characterized the statistical, temporal, and spatial signatures of stealthy manipulation, including distributional preservation properties, coefficient-of-variation changes, and spatial propagation patterns across network buses.
- Constructed a reproducible benchmarking protocol that enables controlled evaluation of anomaly detection methods under distribution-preserving and economically motivated attack conditions.

2- Enabling Real-Time Detection through Incremental Change-Point Analysis

- Proposed an unsupervised, incremental detection framework tailored to streaming electricity market data for near–real-time identification of structural deviations in LMP signals.
- Designed a change-point–based monitoring mechanism that balances detection sensitivity, false alarm stability, and bounded detection delay in short market settlement intervals.

3- Enhancing Robustness under Non-Stationary Market Conditions

- Introduced a drift-aware unsupervised anomaly detection framework capable of distinguishing between natural regime shifts and adversarial manipulation in non-stationary LMP signals.
- Integrated concept drift detection mechanisms with adaptive anomaly scoring to mitigate performance degradation caused by evolving market dynamics.
- Validated the framework on both synthetic and real-world datasets, demonstrating improved detection stability, reduced false positives, and sustained performance across distributional shifts.

4- Leveraging Spatial Structure for Scalable Detection and Localization

- Modeled LMP signals as graph-structured representations over transmission networks, capturing spatial dependencies governed by physical grid constraints.
- Developed physics-informed node selection and dimensionality reduction techniques to improve computational scalability while preserving detection sensitivity.
- Proposed a localization mechanism that identifies manipulated regions with quantifiable spatial accuracy, enabling interpretable and system-level market monitoring.

These contributions collectively advance the state of the art in energy market cybersecurity by establishing a comprehensive, data-driven framework for detecting stealthy and financially motivated LMP manipulation. The thesis demonstrates that real-time monitoring, robustness to non-stationarity, and scalable spatial modeling can be achieved simultaneously through principled statistical design and graph-informed techniques. The proposed methodologies are grounded in established theoretical principles from statistical learning, change-point detection, and graph signal processing, and are validated through controlled simulation environments and real-world electricity market data. Emphasis is placed on reproducibility, rigorous benchmarking, and comprehensive experimental evaluation under realistic manipulation scenarios.

1.6 Framework Overview

This thesis develops a unified monitoring framework for detecting and analyzing anomalies in electricity market price signals. The framework consists of four interconnected components that progressively address the key challenges of real-time market monitoring. First, an incremental change-point detection module identifies structural deviations in streaming LMP signals, enabling near–real-time detection of abnormal market behavior. Second, a drift-aware

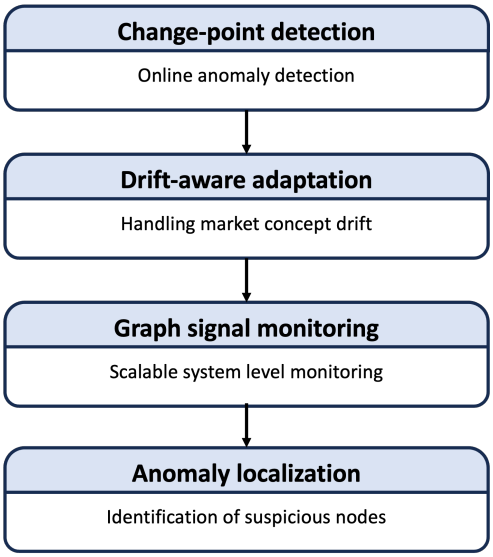


Figure 1.1: Conceptual overview of the proposed monitoring framework for electricity market anomaly detection and localization.

detection mechanism is introduced to maintain reliable performance under non-stationary market dynamics by distinguishing between natural regime shifts and adversarial manipulation. Third, graph signal monitoring leverages the spatial structure of electricity networks to model dependencies among nodal prices, enabling scalable system-level analysis through graph-based representations. Finally, an anomaly localization component identifies suspicious nodes within the network, providing spatial insight into potential manipulation events. Figure 1.1 shows the conceptual structure of the proposed monitoring framework and highlights how these components interact to support adaptive, scalable, and real-time anomaly detection in electricity markets.

1.7 Publications from This Research

The following publications have resulted from the research conducted during this PhD:

1. G. O. Alsharif, C. Anagnostopoulos, and A. K. Marnerides, "Energy Market Manipulation via False-Data Injection Attacks: A Review," *IEEE Access*, vol. 13, pp. 42559–42573, 2025. doi: 10.1109/ACCESS.2025.3548914.

2. G. O. Alsharif, C. Anagnostopoulos, and A. K. Marnerides, "Incremental Learning Detection of Distributed Financially Motivated Attacks in Energy Markets," in *Proceedings of the 2025 IEEE 45th International Conference on Distributed Computing Systems Workshops (ICDCSW)*, 2025, pp. 189–194. doi: 10.1109/ICDCSW63273.2025.00038.

3. G. O. Alsharif, C. Anagnostopoulos, and A. K. Marnerides, "Drift-aware Unsupervised Detection of Stealthy FDIA Towards Energy Market," in *Proc. IEEE Int. Conf. on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2025, pp. 1–7. doi:10.1109/SmartGridComm65349.2025.11204629.

4. G. O. Alsharif, C. Anagnostopoulos, A. K. Marnerides, and P. Mathaios, "SMLT: A Synthetic Dataset for Stealthy Manipulation of Energy Market via False Data Injection Attacks," Accepted at *ECML-PKDD Workshops*, 2025.

5. G. O. Alsharif, C. Anagnostopoulos, and A. K. Marnerides, "Physics-Informed Sampling for Adaptive Anomaly Detection of Graph-Smooth Signals in Energy Markets," Under Review in *IEEE Transactions on Power Systems*.

1.8 Thesis Outline

This thesis is organized in the following manner:

Chapter 2 provides background and a systematic literature review. It discusses smart grid infrastructure, energy market operations, and state estimation mechanisms, followed by a comprehensive survey of FDIAs and existing defense strategies. The chapter identifies key research gaps in market-level anomaly detection, handling non-stationarity, and structural monitoring.

Chapter 3 addresses RQ1 by establishing a benchmarking framework for stealthy LMP manipulation. It introduces the SMLT synthetic dataset and systematically characterizes the statistical, temporal, and spatial signatures of manipulation under varying attack scenarios. The chapter provides a controlled experimental foundation for evaluating detection mechanisms.

Chapter 4 addresses RQ2 by presenting an incremental, unsupervised change-point detection framework for real-time monitoring of streaming electricity market data. The proposed method enables near–real-time identification of structural deviations in LMP signals while maintaining controlled detection delay and false alarm rates.

Chapter 5 addresses RQ3 by introducing a drift-aware anomaly detection framework to handle the non-stationary nature of electricity markets. The chapter integrates concept drift detection mechanisms to distinguish between natural regime shifts and adversarial manipulation, improving robustness under evolving market conditions.

Chapter 6 addresses RQ4 by leveraging spatial dependencies among LMP signals by modeling them as graph-structured signals over the transmission network. It introduces physics-informed dimensionality reduction, graph-based concept drift detection using reduced smoothness, and anomaly localization mechanisms to enhance scalability and interpretability.

Chapter 7 concludes the thesis by summarizing the findings, discussing limitations, and outlining directions for future research in energy market cybersecurity and data-driven anomaly detection.

Chapter 2

Background and Literature Review

2.1 Introduction

Electrical Power Systems (EPSs) have undergone a significant transformation with a shift towards greener energy solutions, leading to the development of smart grids. This transition has resulted in the increased integration of Distributed Energy Resources (DERs), such as wind turbines and solar power, which alter both the operational and market dynamics of energy systems. In addition to these technological advancements, energy markets have evolved from simple supplier-retailer-customer models to competitive frameworks, where market participants actively bid and offer energy within dynamic pools. Locational Marginal Prices (LMPs) have emerged as a key mechanism for determining the market dynamics in modern energy systems. LMPs represent the cost of delivering electricity at different locations within a transmission grid, reflecting the marginal costs of generation, grid congestion, and transmission losses [1, 3].

LMPs are essential to ensure efficient resource allocation and market operations. However, they are vulnerable to manipulation, particularly through False Data Injection Attacks (FDIA), which exploit weaknesses in the grid's Information and Communication Technology (ICT) infrastructure. FDIA involves injecting false data into the system to deceive state estimation processes, leading to inaccurate results and indirect manipulation of LMPs. This manipulation can have significant financial, operational, and stability-related consequences for the energy market. For instance, artificially inflated or deflated LMPs distort market price signals and may cause financial losses to legitimate market participants, including consumers. Furthermore, persistent FDIA can affect grid stability by disrupting dispatch decisions, exacerbating congestion, and increasing the likelihood of grid failures.

The significance of LMP manipulation and FDIA in energy systems is not merely theoretical, as evidenced by real-world instances. One notable example is the Western U.S. energy crisis of 2001, where manipulative practices led to severe shortages in California, resulting in blackouts and the collapse of major energy companies [4]. More recently, in 2022, the Federal Energy Regulatory Commission (FERC) imposed a $4 million penalty on a major energy company and

several individuals to engage in fraudulent trading activities within the PJM energy markets [5]. Additionally, in the UK, Ofgem fined companies for market manipulation, such as false reporting capacity, which resulted in £12.8 million in illicit profits [6].

Given the growing vulnerability of energy markets to FDIA and similar attacks, there is an urgent need for a better understanding of these risks and for the development of effective countermeasures. The existing literature has made substantial progress in studying the threats posed by FDIA, particularly in the context of power system operation and electricity market security. However, comprehensive and up-to-date reviews synthesizing these findings and examining the implications for LMP manipulation remain limited. To the best of our knowledge, only two review papers closely relate to this topic: [7] and [1]. Deng et al. [7] provide a survey of FDIAs on power system state estimation, focusing mainly on attack construction, operational impacts, and protection mechanisms at the grid level. Zhang et al. [1] review profit-oriented cyberattacks in electricity markets, emphasizing attack strategies and their economic implications. In contrast, this work focuses specifically on LMP manipulation through FDIA in energy markets and provides a data-centric perspective on the smart grid ecosystem, including a categorization of relevant data sources, an overview of defense mechanisms against LMP manipulation, and a discussion of market-level anomaly detection approaches. Table 2.1 summarizes the key differences between these surveys and the scope of the present work. Table 2.2 provides detailed information on the methodology employed to gather review materials.

Table 2.1: Comparison of existing FDIA survey papers with this work

Feature	**Deng et al. [7]**	**Zhang et al. [1]**	**This Work**
Main research scope	FDIAs on power system state estimation	Profit-driven cyberattacks in electricity markets	LMP manipulation through FDIA and systematic analysis of defense mechanisms
System level studied	Grid operation and state estimation	Electricity market operations	Electricity market operations
Focus on market attacks	Very limited	Strong focus on market attacks	Primary focus on LMP manipulation
Data-centric analysis of smart grid	Not included	Not included	Detailed categorization of smart grid data sources
Defense mechanisms	General protection methods for state estimation	Limited discussion	Systematic taxonomy of defenses against LMP manipulation
Detection methods	BDD and state estimation protection	Limited discussion	Market-level anomaly detection and monitoring approaches
Data-driven detection research gaps	Not discussed	Not discussed	Extensively discussed
Market-level resilience discussion	No	Limited	Explicit focus on market security and resilience

The main contributions of this review are as follows:

- A detailed background on both the energy system infrastructure and energy markets, with a particular focus on their inherent vulnerabilities to market manipulation and associated risks.
- This comprehensive systematic review provides an in-depth analysis of existing methods for LMP manipulation via FDIA and the corresponding defense mechanisms.
- Perform categorization to enhance the understanding of LMPs manipulation threat models.
- To the best of our knowledge, this is the first study to conduct a thorough survey of defenses against LMP manipulations.
- Outline of challenges and future directions for the energy industry to further investigate and address LMPs manipulation challenges.

The remainder of this chapter is organized as follows. Section 2.2 presents background information on smart grids, while Section 2.3 analyzes energy market operations. Sections 2.4 and 2.5 provide an in-depth examination of threat models and countermeasures for LMP manipulation. Section 2.6 discusses challenges and future research directions. Finally, Section 2.7 concludes the chapter.

Table 2.2: Methodological Overview of the Literature Review

Component	Description
Review Method	A structured narrative review supported by backward and forward snowballing [8].
Start Set Selection	**Search keywords:** False Data Injection Attack; Cybersecurity Attack; Energy Market; Locational Marginal Price; Market Manipulation; Attack Detection; Anomaly Detection. **Databases searched:** Google Scholar, IEEE Xplore, ACM Digital Library, ScienceDirect, and SpringerLink.
Backward and Forward Snowballing	References and citations of the selected start-set papers were systematically examined. Additional studies were included when they satisfied the predefined inclusion criteria.
Inclusion Criteria	• Directly related to FDIAs in energy markets. • Discusses the impact of FDIAs on locational marginal price. • Published as a peer-reviewed paper. • Published between 2011 and 2024.
Exclusion Criteria	• Focuses only on operational FDIAs without addressing financial consequences. • Focuses only on market fraudulent practices without considering FDIAs. • Duplicate publication. • Non-peer-reviewed publication.
Data Extraction Focus	**Threat models:** Scope, attack types, attack vectors, and attacker knowledge requirements. The review also considers whether the proposed models are theoretical or practically implementable. **Defense mechanisms:** Defense type, defense level, and analysis methods used in the proposed solutions. Defense mechanisms may also be categorized as prevention, detection, or mitigation approaches.

2.2 Smart Grid: An Overview

A Smart Grid (SG) represents a transformative approach to the generation, distribution, and consumption of electrical energy. By leveraging information and communication technologies (ICT), SGs significantly enhance the efficiency, reliability, and sustainability of the power grid infrastructure [9]. SGs extend conventional power systems by integrating advanced sensing, communication, automation, and data analytics into grid operation and planning. In transmission systems, measurements are continuously collected from substations, transmission lines, and generation assets and transmitted to control centres, where they support monitoring, state estimation, dispatch, and market operation. Control centres also send control commands and dispatch instructions back to controllable assets. The development of smart grids therefore increases the volume, speed, and diversity of operational data, enabling improved situational awareness, faster response to disturbances, and more efficient market operation. At the same time, this increased dependence on digital monitoring and communication expands the cyber-attack surface of power systems and electricity markets.

In this data-driven environment, contextual data collected from different points of the power system are processed to support monitoring, control, forecasting, and market-related decision-making [10]. Data and predictive analytics play a crucial role in defining SG functionality by enabling automatic processing of measurements, state inference and estimation, outage and fault detection, demand response, and voltage control and regulation [11]. However, this data-centric and analytics-driven intelligence integration in the SG introduces various cybersecurity challenges that significantly impact the stability and reliability of the SG [12]. Fig. 2.1 illustrates (a) the architecture of the SG, highlighting (b) the data flow and (c) the corresponding FDIA threats that originate from the lower levels, where contextual data are generated, and ascend to the upper levels, where operational and trading processes occur.

Thus, this section provides background on SG data sources, data acquisition, communication, and analytics, with particular emphasis on the transmission-level information flows that support state estimation, dispatch, and market operation. Understanding these components is essential for analysing the threat models of energy market manipulation discussed in later sections.

2.2.1 Contextual Data Sources

Contextual data from SGs are generated in real time at a high velocity and volume, fitting into the 5V framework of Big Data [10]. Extracting meaningful insights from SG data is crucial for SG applications, necessitating a deep understanding of its diverse sources, which encompass the entire electricity lifecycle from generation and transmission to distribution and consumption, as well as non-electric contextual data such as environmental conditions, demographic trends, and market dynamics. These types of data directly influence electricity-related data. Table 2.3 summarizes the main contextual data sources in SGs, categorizing them into four key groups:

Generation & Consumption Data, covering electricity generation from various sources and consumption patterns; Smart Grid Infrastructure Data, including distribution and transmission lines, substations, and related parameters; Market & Economic Data, involving electricity markets, prices, regulatory policies, and economic indicators; and External Data, encompassing weather conditions, local holidays, and social events impacting electricity consumption and generation behavior.

Table 2.3: Contextual Data Sources in Smart Grids

Category	Description	Representative Examples
Generation and Consumption	Electricity generation and demand-side usage data used to characterise power production, load behaviour, and consumption patterns.	Active power, reactive power, total energy consumption, peak demand, load profiles.
Grid Infrastructure	Operational and physical measurements describing the condition, performance, and reliability of smart grid assets.	Voltage magnitude, transformer loading, line losses, frequency deviations, fault records.
Market and Economic	Pricing, market-operation, and economic data that influence electricity trading, demand response, and regulatory decisions.	Spot prices, tariffs, locational marginal prices, carbon costs, bidding and auction data.
External Contextual Factors	Non-electrical data that may affect electricity generation, demand, market prices, or grid operating conditions.	Temperature, wind speed, solar irradiance, humidity, public holidays, extreme weather events.

2.2.2 Data Acquisition

The infrastructure for data acquisition in the SG was designed to gather comprehensive data from various sources across the grid. This involves deploying IoT devices such as smart meters, Remote Terminal Units (RTUs), and Phasor Measurement Units (PMUs) at strategic locations. Contextual data collected from these devices include electricity usage at residential and commercial properties and real-time (monitoring of) voltage, current, and frequency along transmission lines. Such data acquisition devices ensure a continuous and detailed flow of contextual information from end-user sites through distribution and transmission networks, providing robust knowledge for grid management and optimization.

2.2.3 Data Transmission

Data transmission enables the exchange of measurements, status information, and control-related data across the cyber layer of the power system. In transmission-level operation, data are exchanged between field devices, substations, control centres, and other operational entities. Different communication standards and protocols are used depending on the operational context. For example, IEC 61850 supports communication within substation automation environments, while IEC 60870-6/TASE.2 supports information exchange between control centres [13]. These communication paths allow measurements, topology information, equipment status, and operational data to be transferred to the control centre, where they support monitoring, state estimation, dispatch, and market-operation applications. Therefore, the integrity of transmitted data is critical: delayed, replayed, modified, or falsified data may propagate through state estimation and dispatch processes and ultimately affect congestion assessment LMP calculation.

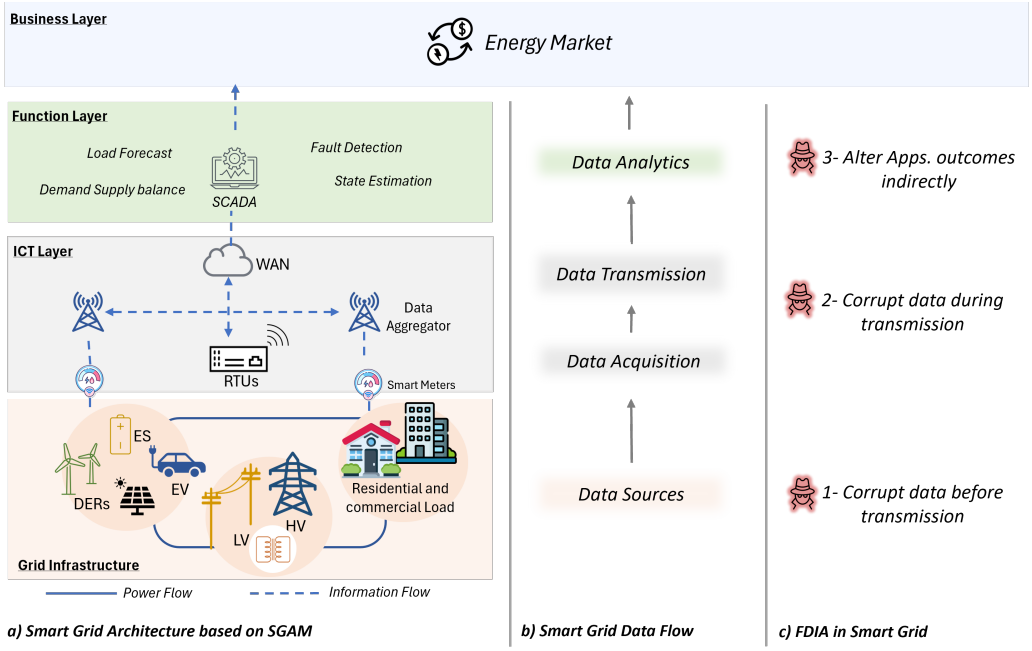


Figure 2.1: Overview of a) Smart Grid Architecture, b) Corresponding Data Flow, and c) General FDIA towards upper layers.

2.2.4 Data & Predictive Analytics

The most important steps in data processing and knowledge derivation within an SG are data analysis and inference. This enables the extraction of useful insights and patterns that support robust and scalable decision-making. Data and predictive analytics aims to identify hidden, potentially useful, and reusable information and trends in massive datasets and knowledge bases, transforming them into practical information and generalizable patterns. In [14], the

predictive and data analytics tasks in an SG were classified as event detection and identification analytics, contextual state and operational analytics, and customer analytics. Event analytics aid in diagnosing and detecting power system events such as faults, outages, and abnormal conditions, as well as identifying malicious attacks and electricity theft.

For instance, real-time fault detection Machine Learning (ML) algorithms quickly isolate and manage faults, minimizing downtime enhanced with self-healing properties. Contextual state and operational analytics involve tasks like state inference and estimation, SG topology identification, and energy forecasting. These analytics ensure efficient system operation, accurate load forecasting, and optimal resource dispatch (along with smart task offloading decision making). An example includes ML-based analytics to forecast energy demand ensuring balanced energy distribution and reducing operational costs. Customer analytics focus on 'understanding' and 'interpreting' consumer behavior and energy consumption patterns, enabling utilities to offer tailored demand response programs and improve customer classification and intelligent profiling. This involves analyzing smart meter data to identify long-/short-term consumption trends and optimize energy usage, benefiting both the utility and the customer through 'personalized' energy-saving recommendations and dynamic pricing strategies.

2.2.5 FDIA in a Smart Grid

In recent decades, the energy sector has been the target of numerous high-profile cyber-physical attacks, demonstrating the susceptibility of critical infrastructures to such threats. One of the earliest notable incidents occurred in 1982, when a cyberattack orchestrated by the CIA during the Cold War led to a devastating gas pipeline explosion in Siberia. The attackers successfully manipulated the pipeline's control systems, triggering catastrophic failures [15]. Recently, in 2015 Ukraine has experienced a significant attack on its power grid, causing a massive blackout that affected more than 225,000 customers. Hackers breached the Supervisory Control and Data Acquisition (SCADA) system and remotely activated circuit breakers, and left operators struggling to manually restore power [16]. These examples highlight escalating risks to the energy industry, with reports indicating a substantial increase in cyber-related disruptions. For instance, between 2011 and 2014, the United States recorded 362 power interruptions, underlining the increasing frequency and severity of these attacks [17].

The vulnerabilities of Smart Grids (SGs) are closely linked to the design and functionality of their data acquisition and transmission systems. These systems depend on a variety of Internet of Things (IoT) devices, including smart meters, RTUs, and PMUs, strategically deployed across the grid to collect real-time, detailed data [18]. These devices monitor key parameters such as electricity consumption, voltage, current, and frequency along transmission lines, enabling continuous oversight of grid operations. However, the extensive deployment of these devices introduces numerous points of vulnerabilities. If attackers gain access to these devices, they can manipulate the transmitted data, resulting in false readings that can disrupt the grid operations.

This vulnerability extends to higher levels of the system, where compromised data can have cascading effects. Moreover, the reliance on IoT devices makes the grid susceptible to weaknesses in their security protocols, such as inadequate encryption or authentication measures, which attackers can exploit to compromise the overall integrity of the system.

The consequences of cyberattacks on smart grids can be far-reaching, encompassing financial losses, power supply disruptions, and instability in grid operations. The effects on the grid stability can be particularly severe. Cyberattacks involving false data injections can lead to incorrect load shedding, unnecessary power generation rescheduling, or even large-scale blackouts [17]. A 2014 study demonstrated that spoofed GPS signals could activate protective measures, triggering load shedding and widespread instability [19]. Such attacks undermine the resilience of the power grid, potentially causing catastrophic failures if left unaddressed. Beyond the instability impact, the economic repercussions are also concerning, with attacks leading to inflated operational expenses, energy theft [20, 21], and market price manipulation [1], ultimately impacting both utility providers and their customers. For example, an FDIA targeting grid topology can obscure outages, causing operational confusion and significant financial losses [22]. In some instances, attackers have manipulated circuit breaker statuses, resulting in financial damages of up to $100,000 per day [23].

Although energy theft is beyond the scope of this study, understanding the key differences between energy theft and LMP manipulation is essential to develop effective countermeasures against these threats. Table 2.4 provides a summary of these distinctions, highlighting their respective objectives, targeted systems, and the nature of compromised data.

Table 2.4: Comparison of Financially Motivated Attacks in Smart Grids

Aspect	Energy Theft	LMP Manipulation
Objective	Manipulates the billing system to avoid or reduce payment.	Compromises market-relevant data to manipulate Locational Marginal Price (LMP) outcomes.
Attack Origin	Often initiated from compromised end-user smart meters or distribution-level devices.	Typically initiated through transmission-level measurements, Remote Terminal Units (RTUs), or related monitoring data.
Adversary	Typically consumers or prosumers seeking to avoid payment.	Typically a market participant seeking illegal financial gain.
Affected System	Affects the integrity of the distribution system.	Affects the integrity of the transmission system and wholesale market operation.
Affected Market	Primarily impacts the retail electricity market.	Primarily impacts the wholesale electricity market.
Consequences	Causes monetary losses to the distribution system operator or electricity supplier.	Causes financial losses to other market participants and may reduce market efficiency and fairness.

2.3 Energy Market

2.3.1 Market Evolution

The evolution of the electricity market has seen significant transitions in the reshaping of energy distribution and consumption. Initially, monopoly utilities managed the entire process, from generation to distribution within distinct regions. However, this centralization has faced efficiency and adaptability. The creation of power pools interconnected neighboring utilities through transmission networks, enabling cost-effective supply and enhanced reliability during high demand and emergencies. Despite these improvements, the full potential of interconnection was realized with the emergence of wholesale markets, which allowed real-time pricing and transactions, even if they initially overlooked transmission limitations. This has evolved into nodal markets, incorporating transmission constraints into pricing, leading to effective incentive structures and further innovation. LMPs have refined this approach by reflecting the marginal value of energy at specific times and locations, thus fostering efficient incentives and market optimization [3].

2.3.2 Market operation

Wholesale electricity markets are fundamental to the energy sector, enabling the exchange of electricity between producers and retailers. These markets are structured to guarantee efficient resource allocation, grid stability, and competitive pricing for consumers [3]. A successful wholesale market model is commonly structured around Locational Marginal Prices (LMPs), which correspond to bid-based, security-constrained economic dispatch [24]. Wholesale markets have two main settlements: day-ahead and real-time. The day-ahead market deals with scheduled energy based on forecasts of supply and demand, whereas the real-time market is required to correct deviations from these forecasts and maintain the balance between generation and demand during actual operation [25].

Day-ahead Market

The day-ahead market operates as a short-term forward market. In this market, load-serving entities and generators submit their anticipated demands and offer schedules before the operating day. The market operator clears the market using forecasted demand, generation offers, and network constraints. This process determines the scheduled generation and expected LMPs for each location in the system. In this sense, the day-ahead market mainly deals with planned energy schedules based on forecasts.

Real-time Market

The real-time market operates closer to actual delivery and is required because actual system conditions rarely match day-ahead forecasts exactly. Demand, renewable generation, outages, and congestion may differ from the expected values used in the day-ahead market. Therefore, the real-time market adjusts dispatch instructions to maintain the balance between generation and demand at each operating interval. This process is tightly coupled with the calculation of LMPs, which reflect the marginal cost of supplying electricity at specific locations and times.

The cost of operation in the real-time market involves determining a solution that optimally allocates power generation G_i to meet the expected load D_i, while ensuring compliance with physical security constraints at each bus i. A simplified DC optimal power flow formulation can be written as follows:

$$\min \sum_{i=1}^{N_b} c_i(G_i) \tag{2.1}$$

subject to:

$$\sum_{i=1}^{N_b} (G_i - D_i) = 0 \tag{2.2}$$

$$P_i^{\min} \leq G_i \leq P_i^{\max}, \quad \forall i \in N_b \tag{2.3}$$

$$\sum_{i=1}^{N_b} GSF_{l-i}(G_i - D_i) \leq F_l^{\max}, \quad \forall l \in L \tag{2.4}$$

$$\sum_{i=1}^{N_b} GSF_{l-i}(G_i - D_i) \geq F_l^{\min}, \quad \forall l \in L \tag{2.5}$$

where $P_i^{\min}$ and $P_i^{\max}$ represent the lower and upper generation capacities at bus i, respectively. The term GSF_{l-i} denotes the Generation Shift Factor between line l and bus i, while $F_l^{\max}$ and $F_l^{\min}$ are the upper and lower transmission limits of line l, respectively. The LMP at bus i is then estimated as:

$$LMP_i = \lambda + \sum_{l \in L} GSF_{l-i}\mu_l \tag{2.6}$$

where λ is the Lagrange multiplier associated with the power-balance constraint, and μ_l is the Lagrange multiplier associated with the transmission constraint of line l. This formulation shows that LMPs are affected by both the system-wide energy balance and local congestion conditions.

This formulation is a simplified DC representation of market clearing. It neglects reactive power, voltage-magnitude constraints, losses, unit commitment constraints, reserve requirements, and detailed market-specific rules. Nevertheless, it captures the main relationship between network constraints, congestion, and nodal prices, and is therefore useful for explaining how false or manipulated system data can influence LMPs.

2.3.3 State Estimation and its Role in Market Operation

State Estimation (SE) is a fundamental operation of various power system applications. It involves estimating the system state variables, such as bus voltage magnitudes and bus voltage angles, from measurements collected across the power system. These measurements are commonly obtained through SCADA and related monitoring infrastructure, including active and reactive bus power injections, branch power flows, and bus voltages [26]. SE plays a key role in determining the LMPs for each node by providing the system operator with the necessary input to calculate the real-time operational state of the system. Through SE, the system operator makes informed technical and economic decisions, including congestion management, optimal power flow, and procurement of ancillary services [26].

In the context of LMP calculation, inaccuracies in state estimation, or deliberate misrepresentation of system states, can affect the system's perceived congestion levels and subsequently the calculated LMPs. If a market participant can influence the state estimation process, for example by providing false data, it may be possible to artificially create or hide congestion at specific locations, leading to distorted LMPs.

In a power system with N_b buses, the sensor measurements obtained through SCADA can be denoted as:

$$z = h(x, g) + \varepsilon \tag{2.7}$$

where $h(\cdot, \cdot)$ represents the nonlinear relationship between the measurement vector z, the system state x, and the grid topology g, while ε denotes measurement noise with covariance matrix R. SE aims to infer a system state $\hat{x}$ that best aligns with the measurements z. Given z, h, and R, the system state can be estimated using the Weighted Least Squares (WLS) method:

$$\hat{x} = \arg\min_{x}(z - h(x))^{\top} R^{-1}(z - h(x)). \tag{2.8}$$

The minimisation is commonly solved using iterative approximation methods, such as Newton–Raphson [26]. Erroneous data arising from faulty sensors or topological errors can be detected using methods such as the Largest Normalised Residual (LNR) method. According to this method, the estimated state variables $\hat{x}$ are regarded as valid only when the norm of the measurement residuals falls below a predefined threshold γ:

$$LNR = \|z - h(\hat{x})\|_2 \leq \gamma. \tag{2.9}$$

This normal SE process provides the background for the state estimation attacks discussed later in Section 2.4.1.

2.3.4 Financial Trading in Electricity Markets & Transmission Rights

Financial trading in electricity markets involves leveraging price differentials and market fluctuations to profit from buying and selling electricity [27]. System operators facilitate this process by opening a market gateway where third parties, known as 'traders', engage in financial transactions to enhance market liquidity. These traders were neither generators nor consumers of electricity; they participated solely in financial transactions without handling actual electricity. The primary objective of a trader is to profit by purchasing electricity in the DA market and selling it in the RT market, or vice versa. It is crucial that any amount of electricity bought in the DA market is exactly sold back in the RT market [28]. Financial trading raise concerns for market operators, as traders may engage in market manipulation to maximize their profits [29].

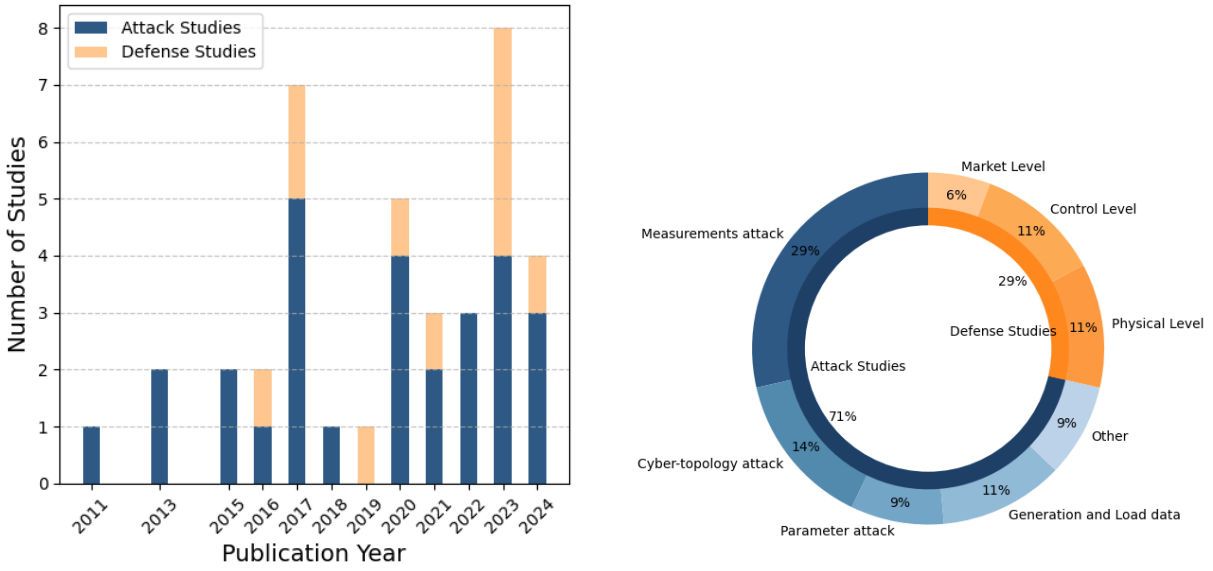
Financial Transmission Rights (FTRs) are financial instruments that enable market participants to hedge against unforeseen changes in nodal prices and associated congestion charges [30]. These rights entitle the holder to a stream of revenue based on the price differences between specified nodes within the market. FTRs are crucial for managing price risk and ensuring price certainty for those involved in the production, distribution, and consumption of electricity. By securing FTRs, market participants efficiently forecast and stabilize their costs, thereby enhancing market efficiency and promoting investment in grid infrastructure. However, owing to the heavy dependence of Security-Constrained Economic Dispatch (SCED) on transmission constraints, FTR holders can engage in LMP manipulation by reducing transmission capacity to create fictitious congestion for their benefit [31]. Further details are provided in Section 2.4.

2.4 LMPs Manipulation through FDIAs

LMPs manipulation refers to unauthorized actions that compromise various underlying measurement units, such as RTUs and communication channels, to mislead LMPs in the energy market [32]. These stealthy activities are designed to avoid detection, bypass BDD, and persist over an extended period, thereby ensuring long-term gains for attackers [1]. They carried out FDIA and Man-in-the-Middle (MITM) attacks or their combination by malicious market participants. Although attacks on the center are unlikely because of strong security measures [33], an adversary is capable of manipulating the LMP by altering the system data reported to the control centre.

LMPs manipulation can be initiated by any market participant seeking to maximize benefits. The main actors in the energy market, include generators, demand-side participants, and financial traders. There are four primary scenarios for potential LMP manipulation: (i) a generator attempting to increase the LMP at their connected bus to boost revenue, (ii) a demand-side participant aiming to decrease the LMP at their connected bus to reduce costs, (iii) a financial trader striving to maximize the price difference between the selling and purchase buses to optimize their financial gains, and (iv) any of the aforementioned participants with FTR profiting from the price differences across the two locations.

There has been a noticeable increase in the literature addressing the problem of LMP manipulation through FDIA, with more than 70% of the reviewed studies focusing on this issue, as shown in Fig. 2.2a. These attacks primarily target system data from various sources, including measurements, parameters, and topology, to mislead the state estimation. Such attacks constitute 80% of the analysed threat models, as depicted in Fig. 2.2b. The classification adopted in this section follows the structured narrative review methodology summarised in Table 2.2, where the selected studies are organised according to the manipulated input or attack vector. Other attack vectors, such as generation and load data alteration attacks and dispatch signal attacks, have



(a) Number of publications per year.

(b) Classification of attack and defense studies.

Figure 2.2: Overview of the energy market cyber-security literature: publication trends and study classification.

also been found in the literature. This section provides a classification of existing threat models based on their attack vectors, along with a theoretical explanation of the attack models and a state-of-the-art review of LMP manipulation threat models. Table 2.5 summarizes the technical details, categorized by attack type, attack vectors, and the knowledge required by attackers.

2.4.1 State Estimation Attack

Measurement attacks aim to deceive SE processes and bypass BDD mechanisms, thereby significantly impacting dispatch operations and real-time LMPs. Manipulating measurements yields a one-time benefit, requiring repeated actions for subsequent LMP calculations [34]. Parameter and topology attacks target parameter processors and topology processors modules, allowing indirect modification of network parameters and topology datasets without accessing the highly secure control center. These attacks manipulate LMP calculations, ensuring long-term advantages for the adversaries [34]. A visual representation of these attacks is shown in Fig. 2.3.

Measurement Attack

Adversaries aim to execute FDIA by deceiving SE $\hat{x}$ in (2.7) into a compromised estimated system state $\hat{x}_a$ such that

$$\hat{x}_a = \hat{x} + c. \tag{2.10}$$

Factor c denotes the variation in the state of the targeted buses in the EPS. This manipulation was achieved by altering the measurements received at the control center to

$$z_a = z + a, \tag{2.11}$$

where a denotes the injected attack vector. To avoid detection through the BDD, the attack vector is defined as

$$a = \mathrm{h}(\hat{x} + c) - \mathrm{h}(\hat{x}). \tag{2.12}$$

In the scenario where the FDIA is constructed with complete knowledge of $h(.)$ and g, the norm of the residual in 2.9 remains unaltered [55], that is,

$$\begin{aligned} \mathrm{LNR}_a &= ||z_a - \mathrm{h}(\hat{x}_a)|| = ||z + a - \mathrm{h}(x) - \mathrm{h}(\hat{x} + c))|| \\ &= ||z - \hat{z}|| = \mathrm{LNR}. \end{aligned} \tag{2.13}$$

Therefore, the manipulated measurements bypass the BDD without detection.

In scenarios with partial or no-prior knowledge, the attacker faces a trade-off between maximizing profit and remaining undetected by minimizing the value of γ. In the worst case where $\gamma = 0$, the attack becomes undetectable [56]. The concept of data integrity attacks in the energy market was first introduced in [29]. However, their approach assumed that the attacker had access to complete knowledge of the network topology and the optimal state. In realistic scenarios, it is impractical for an adversary to possess such information, given that EPSs data are vast, highly secure, and subject to dynamic changes owing to network topology adjustments in normal and contingency situations.

To address this limitation, subsequent research [37, 41, 43, 45, 47] developed FDIA models with partial knowledge of the SG. In [37], the authors proposed using linear independent component analysis (ICA) to infer the transmission line congestion status, thus, reducing the number of measurements that could be manipulated. In [43] and [47], attackers exploited historical data on the probability distributions of loads at different nodes across an SG. They employed Monte Carlo simulations to generate a range of expected power flow scenarios for the transmission lines. Zhang et al. [45] proposed an LMP manipulation without raising suspicion from market operators. This attack was designed to evade detection during the SE phase, while maintaining a normal LMP patterns. Mengis et al. [41] propose a worst-case robustness FDIA, where the attacker is capable of achieving profit while minimizing the value of γ in (2.9) to nearly zero.

Without prior knowledge of the SG topology and parameters, Tan et al. [39] designed online stealthy attacks that exploit online real-time streaming of measurements z. In their experiments, they achieved a large LMP deviation by implementing FDIA for data streaming in two measurements units.

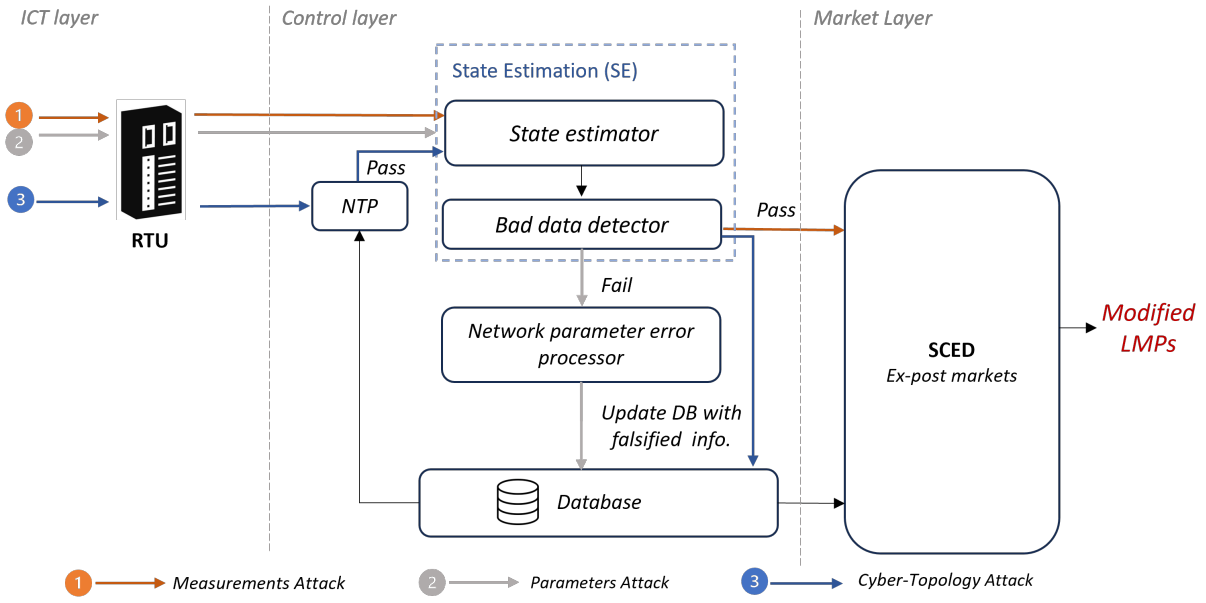


Figure 2.3: State Estimation Attack Paths through MITM and FDIA on RTU.

Parameter Attack

The control center maintains a database of network parameters, including transformer tap settings, transmission line ratings, and branch admittances. However, these parameters may become inaccurate and obsolete over time owing to a lack of regular updates. To address this issue, the system operator initiates Network Parameter Error Processing (NPEP) to identify and correct errors in network parameters [57, 58].

The NPEP process comprises three main steps: (i) error identification, (ii) parameter estimation, and (iii) parameter updating. Given the actual SG parameters g in (2.7) and incorrect network parameters as $\tilde{g}$, in the presence of parameter errors, we can express the measurement model as follows:

$$z = \mathrm{h}(x, \tilde{g}) + \varepsilon = \mathrm{h}(x, g) + [\mathrm{h}(x, \tilde{g}) - \mathrm{h}(x, g)] + \varepsilon. \tag{2.14}$$

If $\tilde{g}$ is sufficiently large, the LNR value surpasses the threshold γ in (2.9), thereby triggering a BDD alert.

Similar to the AC state estimation, NPEP estimate parameter $\hat{g}$ as a WLS problem with the aim of minimizing the objective:

$$\hat{g} = \arg\min_{g} (z - \mathrm{h}(x, g))^{\top} \mathbf{R}^{-1} (z - \mathrm{h}(x, g)). \tag{2.15}$$

In practice, processing measurements collected over multiple time frames significantly improves the accuracy of network parameter estimation [59]. Hence, any erroneous parameters stored in the database can be effectively rectified by updating them using more accurate estimates.

Attackers exploit vulnerabilities within the NPEP to indirectly manipulate the network parameter dataset, without requiring access to a securely protected control center [58]. This requires fewer capabilities, and can be executed offline. This leads to the manipulation of LMP calculations, thereby securing long-term benefits for the adversary. Well-designed parameter attacks pose challenges in detection due to their independence from the system operation points [34].

A study [38] examined the economic implications of Transmission Line Rating (TLR) attacks within two-settlement electricity markets, exploring how nodal prices in real-time markets can be manipulated through TLR attacks. In addition, the potential for cyber attackers to manipulate branch reactance and deceive LMP was explored in [34] and [33]. In [44], targeted transformer tap ratios and phase-shift angles were used to minimize the number of targeted measurements and introduce distortions into the outcomes of SE.

Cyber-Topology Attack

Cyber-topology attacks within EPSs focus on manipulating the network topology estimate. At the core of this manipulation is the Network Topology Processor (NTP), a critical element within the SCADA system that is instrumental in discerning connectivity statuses across the power network. NTP compiles data from various switching devices, such as circuit breakers, responsible for the connection or disconnection of the EPS components [40]. The network topology was represented by an incidence matrix,

$$\mathbf{A}(q, p) \in [1, 0, -1]^{N_b \times L}, \tag{2.16}$$

where N_b is the number of buses and L is the number of transmission lines in an EPS. A transmission line connecting buses q and p is represented by 1 in the matrix, signifying the origin at bus q, and -1 denotes the termination at bus q. If there is no line connecting the two buses, it is denoted as zero. In a cyber-topology attack, an attacker can manipulate the cyber information sent to the control center, thereby altering the incidence matrix in (2.7). This, consequently affect the LMP calculations in SCED. For a successful cyber-topology attack, the attacker must also manipulate a set of measurements and bypass both the BDD and NTP to deceive the LMPs [40].

The impact of cyber-topology attacks on SCED solutions was studied in [40] and [23]. Undetectable cyber-topology attacks, as described in [35], conducted a MITM attack on the connection between the RTUs and the control center, affecting electricity prices in the wholesale market. In [46], a coordinated cyber-physical attack was proposed to manipulate an LMP. Reinforcement Learning was employed to assess the real-time electricity market vulnerability under cyber-topology attacks, where the attacker can modify the solution to the optimal-power-flow problem, thereby influencing LMPs with limited information about the grid topology structure [50].

2.4.2 Dispatch Signal

In addition to falsifying underlying measurements, malicious generators can further corrupt the dispatch signal by targeting the connection between the Automatic Generation Control (AGC) system and the generator unit [48]. AGC serves as a control mechanism used to adjust the output of generation resources in real-time, ensuring a balance between supply and demand while regulating the system frequency within acceptable limits. The attacker manipulates the signal to the generator to produce more energy, thereby generating illegal revenue at the expense of other generators in the same area [60].

2.4.3 Generation & Load Data

The integrity of the generation and load data is crucial for ensuring the fairness and efficiency of the energy markets. Falsifying such data potentially enables market participants to manipulate market outcomes for financial gain. In [36], the authors explore FDIAs in real-time power markets using look-ahead dispatch models and propose strategies to manipulate generator ramp constraints for financial arbitrage through capacity withholding. In [42], they investigated potential market manipulation by aggregators of distributed generation in renewable energy-integrated power systems, focusing on strategically curtailing generation to maximize profit. In [51], Load Altering attacks were introduced, which created artificial high-loading conditions to sell energy at inflated prices. In [52], Load Redistribution attacks were demonstrated by manipulating the line flow and load measurements in the SG to influence the energy market.

Highlights: This section classifies threat models related to energy market operations based on scope, attack type, attack vector, and the knowledge required to launch such attacks. Notably, most inputs to the LMP calculations can be manipulated without prior knowledge of the current state of the grid. In more sophisticated attacks, adversaries can adjust LMPs to align with expected patterns, complicating the detection of these manipulated prices. Additionally, while there is substantial research investigating SE FDIA, more studies are needed to explore FDIA in dispatch signals, generation data, and load data. More insights are provided in Section 2.6

Table 2.5: Current Threat Models for Energy Market Manipulation

Publication	Year	Attack Type	Attack Vector	Attackers' Knowledge
[29]	2011	Measurements attack	Set of measurements	Complete
[35]	2013	Cyber-topology attack	Breaker status	Complete
[36]	2013	Generation data	Ramp-rate	Complete
[37]	2015	Measurements attack	Voltage angle	Partial
[38]	2015	Parameter attack	Transmission line rating	Complete
[39]	2016	Measurements attack	Set of measurements	No prior knowledge
[40]	2017	Cyber-topology attack	Breaker status	Complete
[23]	2017	Cyber-topology attack	Breaker status	Complete
[41]	2017	Measurements attack	Set of measurements	Partial
[42]	2017	Generation data	Output power	No prior knowledge
[43]	2018	Measurements attack	Power flow	Partial
[34]	2020	Parameter attack	Branch reactance	Complete
[44]	2020	Parameter attack	Phase angle & transformer tap ratios	Partial
[45]	2021	Measurements attack	Power flow	Partial
[46]	2022	Cyber-topology attack	Breaker status	Partial
[47]	2022	Measurements attack	Power flow	Partial
[48]	2022	Dispatch signal attack	Dispatch signal	Complete
[49]	2023	Measurements attack	Set of measurements	Partial
[50]	2023	Cyber-topology attack	Breaker status	No prior knowledge
[51]	2023	Load altering attack	Load data	Partial
[52]	2024	Load redistribution attack	Load data	Partial
[53]	2024	Measurements attack	Set of measurements	Complete
[54]	2024	Measurements attack	Set of measurements	Partial

2.5 LMPs Manipulation Defense Mechanisms

Limited research has been conducted on the detection of LMP manipulation. Most existing studies primarily focus on securing the SE process to ensure the reliability of higher-level applications, including economic decisions made at the market level. Consequently, current defense mechanisms are predominantly oriented towards either protecting the SE process or implementing detection techniques to identify anomalies in the data reported to the control center. This section provides an overview of the current state of research on LMP manipulation defense, categorizing the work into two main areas: protection and detection. Table 2.6 summarizes the defense strategies, detailing the publication year, defense level, and methods employed in each study.

2.5.1 Protection Strategies

In protection strategies, specific measurement units are identified as potential targets for malicious activities. Robust security measures are implemented to protect these selected units, including the optimized placement of the PMUs. PMUs provide precise measurements of electrical waveforms (voltage and current) in real-time with a high sampling rate and use GPS signals to synchronize their measurements across different locations. This ensures consistent and coherent data analysis [65]. Therefore, this feature makes PMUs exhibit resilience against data integrity attacks in comparison to RTUs, challenging attackers to manipulate measurements without raising suspicion of detection.

The study in [61] proposed a least-budget Defense strategy to protect pre-selected meters from the FDIA. The defense strategy is formulated as a mixed integer nonlinear programming problem, which is solved using Benders decomposition based on the most vulnerable system meters. In [62], a three-tier defender-attacker-operator methodology was proposed to simplify the process of identifying the minimum set of meters to protect to prevent market manipulation. In [65] and [66], the placement of PMUs was optimized to defend against data integrity attacks in power markets. Different strategies were used to identify the most vulnerable bus/branch and to determine the optimal number of PMUs to be deployed.

However, protection-based mechanism have certain limitations. First, it relies on the impractical assumption that PMUs are immune to attack. Indeed, PMUs are susceptible to GPS spoofing attacks, which can significantly affect the accuracy of their measurements [68]. Moreover, securing a subset of measurements reduces the level of data redundancy, thereby affecting SE accuracy [64].

2.5.2 Anomaly Detection

In [58], the authors designed a detection strategy to identify malicious modifications of critical network parameters by analyzing the minimal protection set at a single branch. The algorithm proposed in [63] aims to detect cyberattacks by analyzing the inconsistencies between specified network parameters and historical data of power injections and voltage phasors. In [39], a detection algorithm using a shift factor matrix as a trustworthy baseline identified attacked real-time measurements. Improved SE approach that allow real-time detection of FDIA was presented in [64]. This method leverages Power Transmission Distribution Factors (PTDF) to calculate variations in load and power flow, aiming to detect indicators LMPs manipulation. In [45], a model using market-level data was proposed to differentiate between safe and risky periods for LMPs manipulation detection. These studies relied on statistical models for attack detection, with limitations in capturing complex and nonlinear system relationships, especially in the presence of sophisticated attacks [66]. Moreover, these techniques depend on utilizing historical data, limiting their ability to identify cyber-attacks when manipulated data are consistent with past measurements [64].

Highlights: This section explores the current defense mechanisms to protect energy market operations from manipulation. Protection strategies secure critical measurement units such as PMUs, whereas anomaly detection methods analyze data to identify signs of manipulation. However, research on defenses against market manipulation remains less developed than the proposed threat models. This underscores the ongoing need for advanced anomaly detection models that can utilize multiple data sources to detect manipulated prices and identify the origin of attacks.

Table 2.6: Current Energy Market Defense Research

Publication	Year	Category	Defense Level	Method	Description
[39]	2016	Anomaly Detection	Control Level	Model-based	Use real-time measurements and baseline shift factor matrices to identify attacked measurements
[61]	2017	Protection using PMU	Physical Level	Game-theory	Define critical meters using Benders decomposition algorithm
[62]	2017	Protection using PMU	Physical Level	Model-based	Define critical meters using Three-tier defender-attacker-operator algorithm
[63]	2019	Anomaly Detection	Control Level	Model-based	Detection of the inconsistencies between network parameters and the historical measurements data
[64]	2020	Anomaly Detection	Control Level	Model-based	Detection of the variations in load and power flow
[45]	2021	Anomaly Detection	Market Level	Model-based	Differentiate safe and risky periods for detecting LMPs anomalies
[65]	2023	Protection using PMU	Physical Level	Optimization	Define critical buses/branches using Binary Harris Hawk optimization
[66]	2023	Protection using PMU	Physical Level	Optimization	Define critical buses/branches using a proposed algorithm to increase system observability
[58]	2023	Protection	Control Level	Model-based	Define minimal measurement set to detect changes in system parameters by analyzing both observability and Jacobian matrix
[67]	2024	Anomaly Detection	Market Level	Data-driven	Market-level integrated detection against cyber attacks in real-time market operations by self-supervised learning

2.6 Challenges & Future Directions

2.6.1 Potential LMP Manipulation Threats

Given the foundational role of LMP and SCED in market operations, inputs into SCED require careful scrutiny because of their potential vulnerabilities, which could be exploited to manipulate market prices and disrupt dispatch settlements. The threat posed by SE has been thoroughly examined in the literature (as discussed in Section 2.4). However, other types of attacks, such as those involving load altering attacks [51] and manipulations in load and generation forecasts [69, 70], remain largely unexplored in the context of their impact on market operations. Furthermore, the study of individual behaviors exhibited by DER owners, particularly concerning their bidding strategies and capacity withholding [36], represents an ongoing research frontier within the energy market context. The study and understanding of these less-explored attack vectors are important for fortifying the resilience of market mechanisms against potential manipulation and ensuring the reliability of EPSs. Furthermore, SG data analytics relies on a variety of machine learning applications to optimize their operations [71]. These applications encompass various functionalities, including resource scheduling, demand response forecasting, energy trading, and grid balancing. ML algorithms analyze historical data to make predictions, enabling efficient allocation of distributed energy resources and enhancing decision-making processes [72]. However, these ML applications are susceptible to adversarial ML attacks that have not yet been thoroughly investigated, particularly in the context of energy market manipulation. Adversarial ML involves manipulating input data to deceive ML models, potentially leading to incorrect predictions or decisions [73, 74]. For example, adversarial attacks can compromise the accuracy of energy demand forecasts, disrupt resource scheduling, and manipulate trading strategies. Therefore, additional studies are necessary to explore the impact of adversarial ML on energy trading and develop ML models capable of detecting such attacks.

2.6.2 LMPs Manipulation Detection

As shown in Fig.2.2b, there is a notable shortage of studies addressing LMP manipulation Defense, particularly in the area of LMP manipulation detection. Detection is a critical component of the market security. It provides a proactive mechanism to identify and respond to malicious activities before they escalate into significant financial or operational disruptions. Although protection mechanisms have been widely studied, they are inherently limited. These protection-centric approaches often involve reducing measurement redundancy [64], which can compromise the accuracy of state estimation (SE) results. Furthermore, protection-based mechanisms alone

are insufficient for ensuring comprehensive security in market operations. A determined attacker can exploit vulnerabilities in these defenses, finding ways to corrupt or bypass them altogether. This highlights the importance of detection, as it complements protection by addressing scenarios in which safeguards fail.

A promising direction to addressing this gap is the development of data-driven detection models. These models leverage advanced machine learning and statistical techniques to analyze vast amounts of market and system-level data, identifying anomalous patterns indicative of manipulation. Unlike traditional model-based detection approaches that rely on predefined energy system models, data-driven models dynamically learn from historical and real-time data, improving their adaptability to emerging threats. Model-based approaches, although effective in capturing known attack scenarios, may struggle with novel or evolving manipulation techniques. In contrast, data-driven methods can detect subtle anomalies by leveraging complex correlations between system variables and market behaviors, thereby enhancing the detection accuracy. Moreover, incorporating market-level data is crucial for strengthening detection capabilities. Market data, such as nodal prices and transaction patterns, provides a broader perspective on system behavior beyond physical measurements. Time-series analysis of market data, when integrated with system-level data, allows for the identification of normal and abnormal market conditions, enabling the localization of attack origins and the recognition of their spatial effects across different node prices. Despite its potential, the application of market-level data for LMP manipulation detection remains largely unexplored, creating a significant research gap that warrants further investigation and development to enhance cybersecurity in the energy market .

Furthermore, the effectiveness of the LMP manipulation detection can be significantly enhanced by deploying real-time detection models. Given the dynamic nature of energy markets, where price fluctuations and system conditions evolve rapidly, real-time detection is essential to provide timely alerts and mitigation strategies. Traditional post-event analysis methods, which are valuable for forensic investigations, lack the immediacy required to prevent market disruptions and financial loss. Real-time detection models, driven by streaming data analytics and edge computing, enable the continuous monitoring of market activities, allowing for instant anomaly detection and rapid response. Various techniques, including adaptive thresholding, online learning, and event-driven architectures, help quickly identify suspicious activities and activate automated countermeasures to mitigate threats before they spread. Moreover, integrating real-time detection systems with market operator control frameworks can provide actionable intelligence, ensuring a more resilient and secure electricity market.

2.6.3 Utilizing Multiple Data Sources

Several studies have demonstrated that an increased number of data sources significantly enhances the accuracy of market manipulation detection [21]. We argue that integrating diverse datasets, including weather, measurement, and market data, can effectively yield robust detection results in the context of financially motivated attacks. The combination of these varied data types not only enriches the analytical framework but also contributes to a more comprehensive understanding of the complex patterns associated with LMP manipulation. Hence, this point opens an important research direction that could contribute significantly to the detection of financially motivated attacks in the energy market.

2.6.4 Need for LMPs Anomaly Detection Dataset

Developing an LMP anomaly detection dataset is crucial for advancing the research on energy market integrity and resilience. Current datasets focus on normal historical LMPs, which, while useful for understanding market behavior, lack the support of effective anomaly detection. Additionally, existing anomaly detection datasets primarily focus on system measurements, potentially missing crucial market dynamics. To bridge these gaps, future datasets should encompass a diverse range of scenarios and market conditions, including normal operations and anomalous behaviors such as price spikes or prolonged deviations. This broader scope should integrate contextual factors such as weather data, demand patterns, and market events alongside LMP values to enhance dataset robustness and provide a comprehensive context for interpreting anomalies. Annotated instances of known manipulation or market disturbances are essential for training robust detection models that support both supervised learning approaches and unsupervised anomaly detection methods. Granularity in temporal and spatial dimensions is critical for the effective detection and localization of anomalies across different scales within the energy grid. This approach ensures that anomalies can be promptly and accurately identified, contributing to enhanced market transparency and resilience against manipulation.

2.7 Conclusions

The increasing integration of digital technologies in energy markets has introduced new vulnerabilities, making them susceptible to sophisticated cyberattacks, particularly FDIAs. Our review highlights a diverse range of attack vectors that can manipulate LMPs and disrupt market stability and financial fairness. While existing defenses primarily focus on securing state estimation and anomaly detection, they remain limited in addressing the broader market-level impact of FDIAs. Our study highlights the need for a more holistic approach that incorporates market-level data analytics, multi-source information fusion, and advanced machine learning techniques for real-time detection. Additionally, adversarial attacks on predictive models and

market forecasting remain largely unexplored, presenting a critical research gap. Addressing these challenges requires the development of robust anomaly detection datasets and the integration of cross-disciplinary insights to enhance energy market resilience, ensuring both security and market integrity in an increasingly digitalized landscape.

Chapter 3

Benchmarking Stealthy Market Manipulation via Synthetic Data

3.1 Introduction

Building on the market and threat background presented in Chapter 2, this chapter addresses a fundamental bottleneck in electricity market cybersecurity research: the absence of standardized, labeled datasets that model stealthy Locational Marginal Price (LMP) manipulation through False Data Injection Attacks (FDIAs). In deregulated electricity markets, LMPs are computed through state estimation and optimal power flow procedures, and directly determine financial settlements. Because LMPs reflect congestion and network conditions inferred from real-time measurements, adversaries can manipulate prices by injecting carefully crafted false data, thereby creating artificial congestion or altering dispatch outcomes for financial gain.

Although prior work has extensively analyzed attack strategies and their financial impact, comparatively less attention has been given to reproducible benchmarking for detection research [75]. Existing anomaly detection studies often rely on simplified attack models, limited simulation horizons, or case-specific configurations that hinder systematic comparison and generalization. Moreover, publicly available electricity market benchmarks typically lack labeled adversarial scenarios and long-horizon time-series data, making it difficult to rigorously evaluate data-driven detection methods under controlled yet realistic manipulation conditions.

To address this gap, we introduce the **S**tealthy **M**anipulated **L**MP **T**imeseries (`SMLT`) dataset[1], a reproducible, open-source framework for modeling stealthy FDIAs targeting electricity market operations. The dataset simulates diverse manipulation vectors affecting transmission ratings, system parameters, topology, and demand profiles, while preserving realistic operational constraints through Matpower-based optimal power flow simulations. Time-series data are generated over extended horizons with ground-truth anomaly labels, enabling systematic evaluation of AI-based anomaly detection algorithms.

[1]The dataset and code are available at: `https://github.com/Ghadeer101/SMLT_Dataset/`

The main contributions of this chapter are as follows:

1. We introduce SMLT, an open-source dataset designed to model stealthy FDIA-driven LMP manipulation using labelled time-series market data. Unlike general synthetic power-system benchmarks, which mainly support power-flow, dispatch, or planning studies, SMLT focuses on adversarial market manipulation scenarios and provides ground-truth labels for anomaly-detection evaluation.

2. We incorporate eight manipulation scenarios identified in the literature, covering diverse attack vectors including transmission limits, network parameters, topology alterations, and load perturbations.

3. We generate hourly-resolution time-series data spanning up to 20 weeks, capturing both short- and long-term system variability, with ground-truth labels provided for each timestep.

4. We develop an open-source FDIA simulation framework built on Matpower, enabling reproducibility and extensibility of attack scenarios.

5. We provide an in-depth spatio-temporal characterization of manipulated LMP behavior and demonstrate a case study on anomaly detection using the GEM model.

The remainder of this chapter is organized as follows. Section 3.2 reviews related work and existing benchmarking limitations. Section 3.3 details the construction of the `SMLT` dataset, including the benchmark system, attack modeling framework, and data generation process. Section 3.4 presents empirical analyses of statistical and spatial manipulation effects. Section 3.5 provides an anomaly detection case study. Lessons learned are discussed in Section 3.6, and Section 3.7 concludes the chapter with limitations and future directions.

Table 3.1: List of Notations for Chapter 3

Symbol	Description
$\mathcal{P}$	Power system case (network model used for simulation)
$\mathbf{a}_{\text{nom}}$	Nominal attack vector describing the manipulated parameter
$\mathbf{a}_t$	Attack vector applied at time step t
α	Attack magnitude scaling factor
$\mathcal{T}_a$	Set of time indices during which the attack occurs
t	Time index (hourly timestep)
$\mathcal{S}$	Generated dataset containing LMPs, timestamps, and labels
$\mathbf{l}_t$	Load vector for all buses at time t
y_t	Binary attack label at time t (1 = attack, 0 = normal)
λ_t	Locational Marginal Price (LMP) vector at time t

Symbol	Description
$\mathrm{LMP}_n^{(i)}$	Normal (baseline) LMP at bus i
$\mathrm{LMP}_m^{(i)}$	Manipulated LMP at bus i under attack
$\Delta\mathrm{LMP}^{(i)}$	LMP deviation between normal and manipulated states at bus i
p	Nominal system parameter value (e.g., load, line rating)
b	Injected bias introduced by the attacker
q_1, q_3	First and third quartiles of a data distribution
$T_{\min}, T_{\max}$	Lower and upper bounds for outlier filtering
c	Quartile multiplier used in outlier detection (set to 1.5)
μ	Mean of an LMP time series
σ	Standard deviation of an LMP time series
CV	Coefficient of Variation ($CV = \sigma/\mu$)
$V^{(i)}$	Visibility metric measuring distributional change at bus i
$D^{(i)}$	Detectability metric capturing temporal statistical variation
$\Delta CV_T^{(i)}$	Average temporal change in CV over T time windows
z_{ij}	Complex impedance of the transmission line between buses i and j
r_{ij}	Resistance of the transmission line between buses i and j
x_{ij}	Reactance of the transmission line between buses i and j
$w(i, j)$	Edge weight representing impedance magnitude between buses i and j
$d(i, j)$	Electrical distance between buses i and j
$\mathcal{P}(i, j)$	Set of all paths between buses i and j
B_j	Set of buses grouped in electrical distance bin j
$\bar{d}_j$	Average electrical distance in bin j
$\bar{y}_j$	Average LMP disturbance in bin j
S	Spreadability metric measuring spatial propagation of attack impact
N_{bins}	Number of electrical distance bins used for spatial analysis
k, h	GEM model parameters controlling neighborhood size and entropy estimation

3.2 Related Work

A wide range of benchmark datasets has supported the development of power system modeling and electricity market studies. The IEEE standard test systems, such as the IEEE 30-bus system and the Reliability Test System (RTS), have long served as the foundation for simulation and optimization tasks [76, 77], though their economic data is outdated and non-transparent. To enable more realistic market studies, researchers have developed economically enhanced test systems, such as the PJM 5-bus and IEEE 30-bus with market benchmarks [78], hydrothermal

dispatch models [79], and the reduced WECC model incorporating demand response and energy storage [80]. Additional systems reflect regional structures, including the Chilean grid test case [81] and a renewable-integrated IEEE 118-bus system designed to study market impacts under high penetration of clean energy [82].

More recently, synthetic grid models from Texas A&M University have enabled large-scale analysis with statistically grounded network structures, load profiles, and economic data [83, 84]. These datasets, along with the WECC/NPCC dataset developed for electricity market modelling and analysis, provide a strong foundation for market simulation, price analysis, and policy evaluation.

However, despite their strengths, none of these benchmarks incorporate adversarial scenarios such as FDIAs, nor do they provide labeled time-series data suitable for training or evaluating AI models in the context of electricity market cybersecurity. The lack of realistic, annotated datasets capturing LMP manipulation remains a critical obstacle to developing and benchmarking anomaly detection algorithms in this domain.

3.3 SMLT Dataset Construction

3.3.1 Baseline Benchmark

The development of SMLT is grounded in a well-established benchmark for electricity market studies provided in [76], with a particular focus on the NPCC test case. This benchmark was selected because it meets key requirements: (1) an open-source system model, and (2) time-series load data. It provides hourly load profiles spanning a full operational year, enabling the design of FDIA strategies that evolve over time. In addition, the openly available system model supports the systematic manipulation of diverse attack vectors to simulate the effects of FDIAs on LMPs. These features ensure that our experiments are conducted under realistic and reproducible conditions, enabling detailed analysis of cyber-physical impacts on electricity markets. Table 3.2 summarises the main characteristics of the NPCC test system used in this chapter.

3.3.2 Threat Model

The SMLT dataset assumes a strategic adversary capable of manipulating selected system parameters in order to influence LMPs while remaining operationally stealthy. The attacker may alter parameters such as transmission limits, generator constraints, topology information, or load injections. To ensure practical realism, all perturbations are constrained to remain within feasible operating ranges. Specifically, manipulated values must satisfy

$$0.2 \cdot p \leq p + b \leq 3 \cdot p,$$

Table 3.2: Summary of the NPCC test system used for SMLT dataset generation.

Property	Value
Test system	NPCC test system
Number of buses/nodes	140
Number of generators	48
Number of load buses	76
Number of transmission branches/lines	233
Bus types	94 PQ buses, 45 PV buses, 1 slack bus
Base power	100 MVA
Voltage levels	13.8–500 kV
Total active load	30,349.76 MW
Total generation capacity	47,323.23 MW

where p denotes the nominal parameter value and b represents the injected bias. These constraints prevent unrealistic parameter changes and reduce the likelihood of triggering basic validation mechanisms. A formal formulation of the FDIA model and its interaction with state estimation and electricity market operations is presented in Chapter 4.

3.3.3 FDIA Scenarios

The attack scenarios incorporated in the SMLT dataset are adapted from existing literature and follow the threat model described in the previous subsection. Each scenario targets a different layer of the power system, including physical parameters, network topology, generator constraints, and load injections. All attacks aim to manipulate LMPs while remaining operationally feasible and bypassing standard Bad Data Detection (BDD) mechanisms. In what follows, we briefly describe the scenarios under consideration:

S1: Transmission Line Rating (TLR) Attack.

This attack falsifies line thermal limits to simulate congestion by reducing the rated capacity of selected transmission lines, as provided in [38]. The injected false data causes the system to overestimate congestion, leading to inflated LMPs at nearby buses. It is a cyber-physical attack that manipulates network constraints without altering load or topology data.

S2: Critical Parameter Attack.

In this scenario, the attacker alters static network parameters such as line reactance or susceptance values stored in EMS databases. These changes affect injection shift factors (ISFs) and propagate pricing distortions throughout the network. Since these parameters are weakly observable, the attack evades standard detection methods. The scenario is based on the formulation in [85].

S3: Cyber-Topology Attack.

This attack involves falsifying breaker status data to change the perceived grid topology while leaving physical infrastructure untouched. By simulating line openings or closures, the attacker misleads the state estimator and alters the SCED formulation. The resulting dispatch errors cause artificial congestion and LMP shifts. This scenario is drawn from the topology attack framework in [86].

S4: Ramp-Induced Data (RID) Attack.

The RID scenario manipulates generator telemetry used to enforce ramp constraints in look-ahead dispatch. By injecting a false generation value at the initial time step, the attacker narrows the feasible ramping window, causing higher-cost generators to be dispatched. This indirectly inflates LMPs and supports profitable arbitrage. The attack is inspired by the method in [87].

S5: Load-Altering Attack (LAA).

This scenario simulates abrupt load injections at specific buses using compromised demand-side infrastructure such as smart meters or automated devices. These changes are timed to avoid triggering protection mechanisms while still shifting the real-time dispatch and LMP outcomes. The attack emphasizes market exploitation over grid disruption and follows the model proposed in [88].

S6: Aggregator-Based Strategic Curtailment.

Unlike traditional FDIA, this scenario does not involve false data but rather a physical strategy where a renewable aggregator curtails output to induce price manipulation. By anticipating how reduced net injection alters real-time pricing, the attacker exploits economic signals to increase revenue. This scenario reflects subtle but realistic manipulation strategies, as discussed in [89].

3.3.4 Dataset Generation

Building on the previously described attack scenarios, we construct eight distinct test cases by varying key parameters such as attack magnitude, duration, and target buses. Figure 3.1 illustrates the end-to-end dataset generation framework. At each time step, the process begins by loading hourly demand profiles and system case data as input. The system may then be subjected to an FDIA according to a predefined case. The attack logic, described in Algorithm 1, determines whether and how false data are applied. The corresponding LMPs are then computed using the OPF solver provided by MATPOWER. Specifically, the `runopf` function is used, which solves the AC-OPF problem by default. The resulting outputs, consisting of LMP values for each bus, are labelled, timestamped, and cleaned to produce the final dataset.

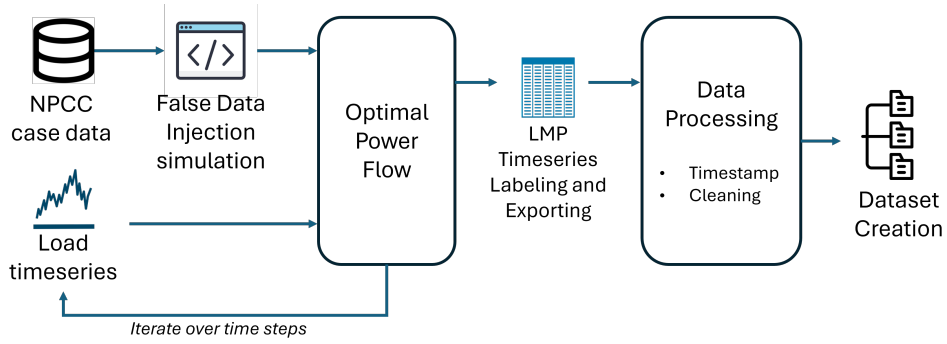


Figure 3.1: Overview of the SMLT dataset development framework.

Algorithm 1 FDIA Generalized Process

Input: Power system case $\mathcal{P}$, nominal attack vector $\mathbf{a}_{\text{nom}}$, attack magnitude α, attack time indices $\mathcal{T}_a$.
Output: $\mathcal{S}$: Time series of LMPs, labels, and timestamps
function LMPAttackSimulation($\mathcal{P}, \mathbf{a}_{\text{nom}}, \alpha, \mathcal{T}_a$)
    Initialize $\mathcal{S} \leftarrow \emptyset$
    for each timestep t **do**
        $\mathbf{l}_t \leftarrow$ GetLoadProfile($\mathcal{P}, t$) ▷ Load vector for all buses at time t
        if $t \in \mathcal{T}_a$ **then**
            $\mathbf{a}_t \leftarrow \alpha \cdot \mathbf{a}_{\text{nom}}$, $y_t \leftarrow 1$ ▷ FDIA present
        else
            $\mathbf{a}_t \leftarrow \mathbf{a}_{\text{nom}}$, $y_t \leftarrow 0$ ▷ No FDIA
        end if
        $\lambda_t \leftarrow$ RunOPF($\mathcal{P}, \mathbf{l}_t, \mathbf{a}_t$)
        $\mathcal{S} \leftarrow \mathcal{S} \cup \{(\lambda_t, y_t, t)\}$
    end for
    return $\mathcal{S}$
end function

Table 3.3 summarizes the FDIA cases included in the SMLT dataset. Each test case specifies details of the attack scenario, including the attack vector $\mathbf{a}_{\text{nom}}$, attack magnitude α, targeted bus, and attack duration; whether it spans a full week or targets only peak-hour injections. Each case is evaluated along two key dimensions: (1) stealthiness: whether the attack bypasses standard bad data detection (BDD); and (2) effectiveness: whether it successfully manipulates LMPs after solving the OPF. It is denoted as ΔLMP in Table 3.3, quantifies the difference between the normal and manipulated LMP at bus i:

$$\Delta\text{LMP}^{(i)} = \text{LMP}_n^{(i)} - \text{LMP}_m^{(i)} \tag{3.1}$$

Both state estimation and OPF are implemented using Matpower 8.0, ensuring compatibility with industry-standard power system analysis tools.

Table 3.3: Summary of attack cases and their validation outcomes

	Scenario	$\mathbf{a}_{\text{nom}}$	α (p.u.)	Target Bus	**Duration**	**BDD**	ΔLMP	**Profit**(per week)
Case 1	S1	$L_{\text{rate},109}$	0.14	Bus 115	Week	Pass	0.19$	127.1 $/MWh
Case 2	S1	$L_{\text{rate},109}$	0.2	Bus 115	Week	Pass	1.7$	1146 $/MWh
Case 3	S2	R_{181} X_{181}	2 1.5	Bus 128	Week	Pass	2.47$	416.37 $/MWh
Case 4	S3	$L_{\text{breaker},109}$	-	Bus 115	Week	Pass	−3.24$	-545.8$/MWh
Case 5	S4	$G_{\text{pmax},13}$ $G_{\text{ramp},13}$	0.2	Bus 50	Week	Pass	3.18$	534.5$/MWh
Case 6	S5	P_{115}	1.2	Bus 115	Peak hours	Pass	0.93$	630.95$/MWh
Case 7	S1, S5	P_{115}, $L_{\text{rate},109}$	1.2 0.2	Bus 115	Peak hours	Pass	1.74$	293.2$/MWh
Case 8	S6	$G_{\text{pmax},15,16,19,20}$	0.02	Bus 56	Peak hours	Pass	1.15$	193.5$/MWh

Dataset Reproducibility. The `SMLT` dataset and the complete generation pipeline are publicly available in the GitHub repository at `https://github.com/Ghadeer101/SMLT_Dataset`. The repository includes the MATLAB simulation script used to generate FDIA attack scenarios, dataset metadata, preprocessing utilities, and analysis scripts used in the experiments. Dataset generation is deterministic given the specified attack parameters and hourly demand profiles. The parameter ranges used to construct the eight attack scenarios are summarized in Table 3.3. While `SMLT` provides a controlled benchmark for studying stealthy market manipulation, it remains a synthetic dataset and therefore cannot capture the full operational complexity and regulatory constraints present in real-world electricity markets.

3.3.5 Data Preprocessing

To mitigate the effect of extreme LMP values and reduce the influence of spurious data artifacts, we apply a column-wise outlier filtering method based on quartile statistics [90]. For each bus-level LMP time series, we compute the first (q_1) and third (q_3) quartiles and define the lower and upper bounds for acceptable values as $T_{\min} = q_1 - c \cdot (q_3 - q_1)$ and $T_{\max} = q_3 + c \cdot (q_3 - q_1)$, where c is a user-defined multiplier set to 1.5 in our experiments. Any values outside this range are considered outliers and are replaced with the median of the corresponding column. This replacement is performed independently for each LMP series, preserving the statistical structure of the data while mitigating distortion from anomalous spikes.

3.4 Empirical Observations

In this section, we present a preliminary set of empirical observations based the developed dataset. By analyzing the data from various perspectives, we aim to uncover key statistical and spatiotemporal patterns that characterize both normal operations and attack scenarios. We seek to answer the following questions: **RQ1:** What is the impact of stealthy FDIAs on the

distribution of LMP data? **RQ2:** How does the impact of an attack propagate throughout the system, and to what extent can the source of the attack be identified? The empirical findings presented here are not intended as an exhaustive evaluation, but rather as an initial exploration aimed at highlighting patterns, generating hypotheses, and informing future approaches to LMP anomaly and attack detection.

3.4.1 The impact of stealthy FDIAs on the distribution of LMP data

To answer **RQ1**, we introduce two evaluation metrics: *visibility* and *detectability*. These metrics are designed to capture statistical and temporal deviations induced by an attack, providing a quantitative foundation for identifying anomalous patterns in the system.

Visibility measures the extent to which the statistical distribution of the LMP shifts in response to an attack over a single day, represented by the Coefficient of Variations before (normal) and after the (manipulated) attack, $CV_n = \frac{\sigma_n}{\mu_n}$ and $CV_m = \frac{\sigma_m}{\mu_m}$, respectively, at the targeted bus i, defined as:

$$V^{(i)} = \frac{\left| CV_n^{(i)} - CV_m^{(i)} \right|}{CV_n^{(i)}}. \tag{3.2}$$

A larger $V^{(i)}$ indicates a more significant shift from the baseline, suggesting that the attack has a more visible impact on the system's behavior.

Detectability, on the other hand, quantifies the extent to which temporal variations in the statistical profile of the LMP at the targeted bus i can be observed during an attack. It is measured by tracking changes in the CV across consecutive time steps:

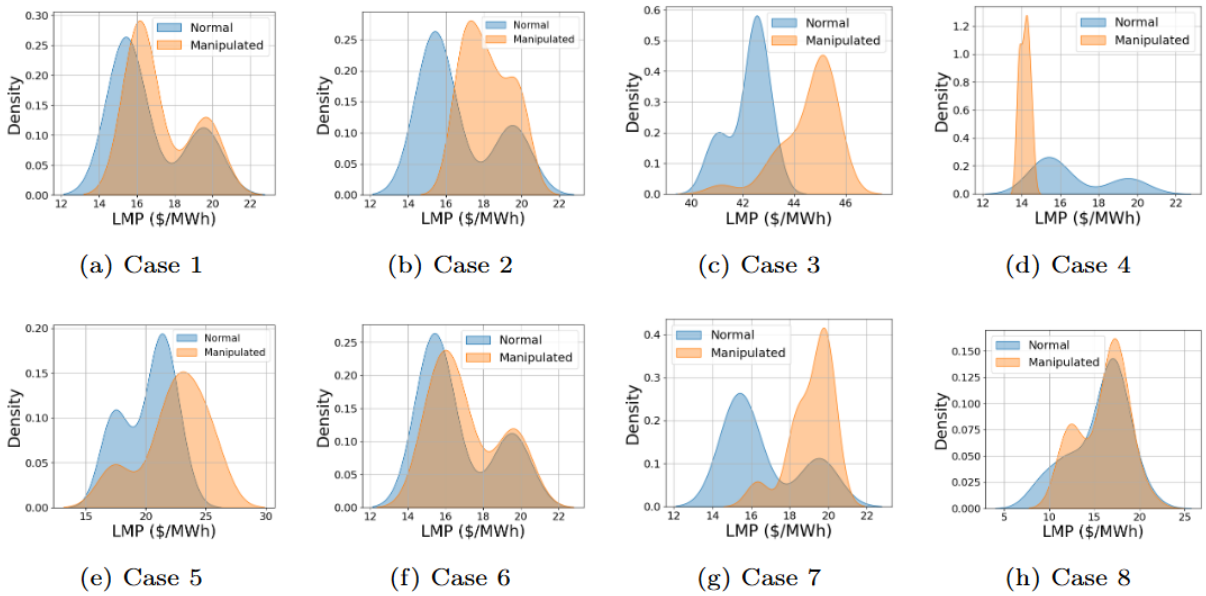
$$\Delta CV_T^{(i)} = \frac{1}{T} \sum_{t=1}^{T} \left| \frac{\sigma_t^{(i)}}{\mu_t^{(i)}} - \frac{\sigma_{t-1}^{(i)}}{\mu_{t-1}^{(i)}} \right|, \tag{3.3}$$

where $\mu_t^{(i)}$ and $\sigma_t^{(i)}$ represent the mean and standard deviation of the LMP at bus i, calculated over a distinct, non-overlapping 24-timestep window, respectively. Detectability is then calculated as the absolute difference in this metric between the normal (n) and manipulated (m) time series:

$$D^{(i)} = \left| \Delta CV_{T(n)}^{(i)} - \Delta CV_{T(m)}^{(i)} \right| \tag{3.4}$$

A higher $D^{(i)}$ indicates that the attack induces more significant fluctuations in the temporal behavior of the LMP, making it more detectable.

Figure 3.2 illustrates that the distribution of normal LMP values deviates significantly from a standard Gaussian-like profile and does not exhibit characteristics of independent and identically distributed (i.i.d.) data. This observation challenges a common assumption in many conventional anomaly detection models, which often rely on stationary or i.i.d. baselines. Moreover, as shown



(a) Case 1 (b) Case 2 (c) Case 3 (d) Case 4

(e) Case 5 (f) Case 6 (g) Case 7 (h) Case 8

Figure 3.2: Data distribution of LMPs at the targeted bus under normal and manipulated conditions for the eight test cases (a–h). Each subplot presents a KDE-based comparison over a one-day period.

in Figure 3.3, the distribution of daily LMPs exhibits continuous and nontrivial changes over time, further complicating the construction of a stable statistical model for normal behavior. This dynamic nature of the LMP data implies that stealthy attacks can be effectively masked within the natural variability of the market, making their detection substantially more difficult. For example, in Cases 3 and 7 (see Figure 3.2), the attacks induce noticeable shifts in the LMP distribution, reflected by high visibility scores (denoted as '+' in Table 3.4). However, due to the continuous evolution of the underlying LMP distribution (see Figure 3.3), these cases exhibit low detectability ('−' in Table 3.4), illustrating the challenge of distinguishing attack-induced changes from normal market dynamics.

Another important observation is that normal LMP time series exhibit long-term non-stationarity (see Figure 3.3), meaning that key statistical properties change over time. This poses significant challenges for power system modeling, as most forecasting and anomaly detection techniques rely on the assumption of stationarity. When applied to non-stationary data, such models are vulnerable to reduced accuracy and performance degradation, a phenomenon commonly referred to as concept drift. For example, a model trained during stable market conditions may perform poorly during periods of stress or in the presence of changing dynamics introduced by renewable generation. Furthermore, non-stationarity complicates anomaly detection: an evolving baseline can result in false positives when normal shifts are misclassified as attacks, or missed detections when actual attacks mimic expected variation. Addressing these issues requires adaptive modeling techniques that account for temporal variation, including seasonality adjustments, rolling model updates, and regime-switching frameworks.

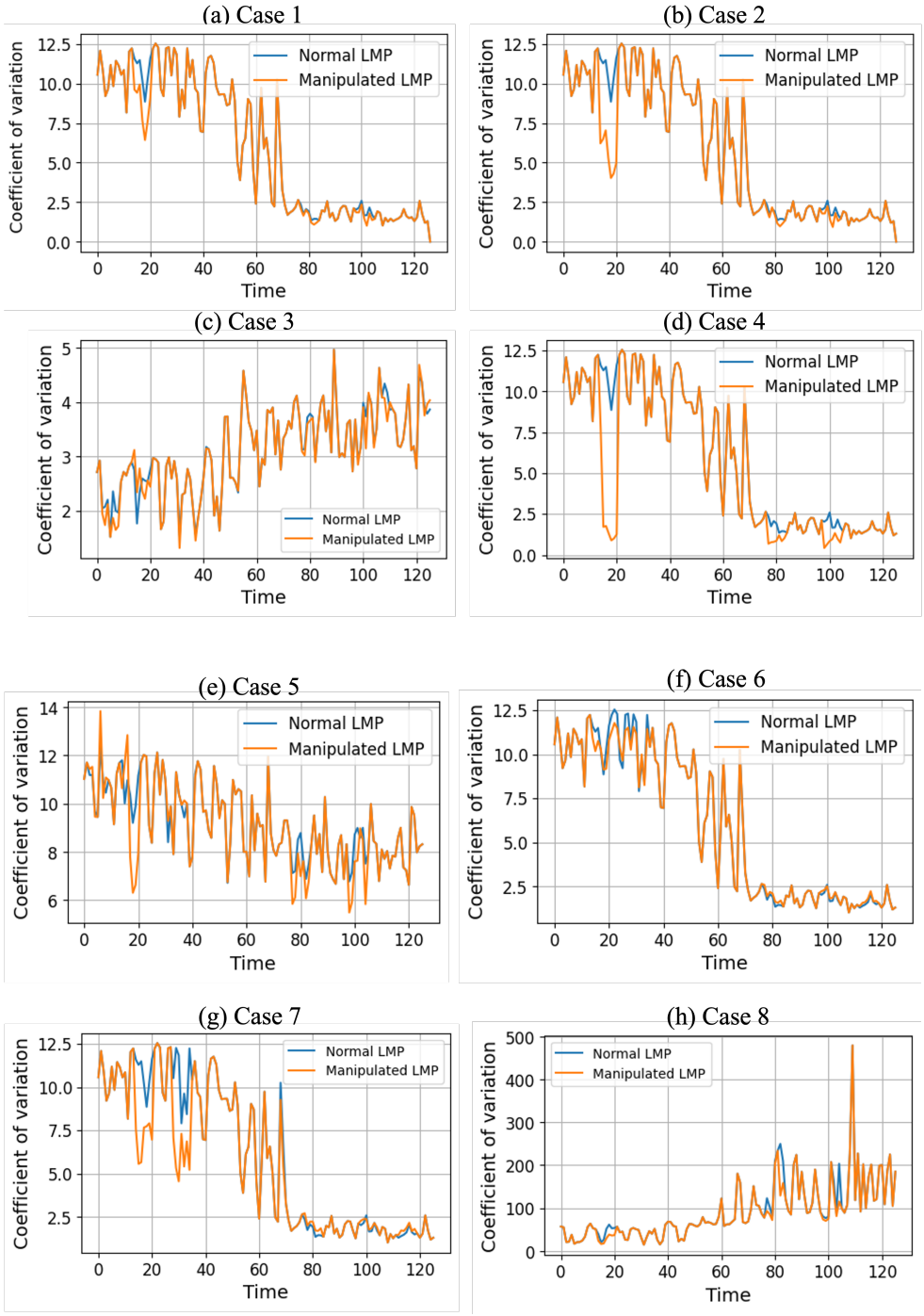


Figure 3.3: Temporal evolution of the CV for daily LMP under normal and manipulated conditions across eight attack cases.

3.4.2 The impact of stealthy FDIA throughout the system

To analyze the spatial impact of LMP manipulation, we adopt an impedance-based method consistent with the approach in [91] to quantify electrical proximity between buses. The power system is modeled as an undirected, weighted graph, where nodes represent buses and edges represent transmission lines. Each edge connecting bus i to bus j is weighted by the magnitude of its impedance:

$$w(i, j) = |z_{ij}| = \sqrt{r_{ij}^2 + x_{ij}^2}, \tag{3.5}$$

where r_{ij} and x_{ij} are the resistance and reactance of the transmission line between buses i and j. To quantify proximity from a target bus, we compute the shortest weighted path through the network using Dijkstra's algorithm. Mathematically, the electrical distance between buses i and j is defined as:

$$d(i, j) = \min_{p \in \mathcal{P}(i,j)} \sum_{l=1}^{k} w_l, \tag{3.6}$$

where $\mathcal{P}(i, j)$ is the set of all paths between buses i and j, k is the number of edges in path P, and w_{e_l} is the impedance magnitude of the l-th edge along the path. The resulting distance vector provides a physically grounded and topologically consistent measure of electrical proximity that is independent of system operating conditions and remains fixed for a given network topology. To reduce noise and reveal spatial trends, we group buses into bins based on their electrical distance from the target. For each bin B_j, we compute the average electrical distance $\bar{d}_j = \frac{1}{|B_j|} \sum_{i \in B_j} d_i$ and the average LMP disturbance $\bar{y}_j = \frac{1}{|B_j|} \sum_{i \in B_j} |\Delta\text{LMP}^{(i)}|$, resulting in a set of smoothed representative points $\{(\bar{d}_j, \bar{y}_j)\}_{j=1}^{N_{\text{bins}}}$ that characterize how the impact of the manipulation decays with electrical distance.

A metric is used to assess the spatial effects of an attack:

Spreadability is a quantitative measure indicating how widely the impact of a cyberattack on $\Delta\text{LMP}^{(i)}$ propagates across the power grid, as a function of electrical distance from the targeted bus. Higher spreadability indicates a broader impact that is sustained over longer distances within the network. It is measured by the Area Under the Curve (AUC) of the $\Delta\text{LMP}^{(i)}$ versus electrical distance plot and computed as follows:

$$S = \sum_{b=1}^{N_{\text{bins}}-1} \frac{\bar{y}_b + \bar{y}_{b+1}}{2} \cdot (\bar{d}_{b+1} - \bar{d}_b), \tag{3.7}$$

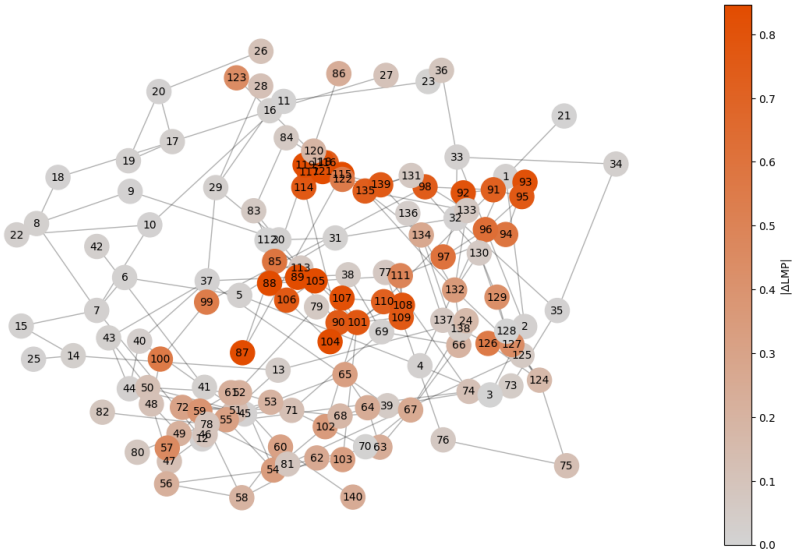


Figure 3.4: Bus-level heatmap of $|\Delta\text{LMP}^{(i)}|$ across the system.

where $\bar{d}_b$ and $\bar{y}_b$ are the average electrical distance and LMP disturbance in bin b, respectively. This metric provides an interpretable scalar that summarizes the spatial extent of attack influence across the network.

The results presented in Table 3.4, Fig. 3.4, and Fig. 3.5 indicate that the impact of LMP manipulation decayes as the distance from the targeted bus increases. In particular, buses located nearest to the attacked bus experience the most significant price deviations, while the effect fades progressively with greater separation. The rate at which this impact declines varies depending on the specific characteristics of each case. Notably, Cases 5 and 7 demonstrate a higher level of spreadability, suggesting that attacks involving generator data manipulation have a broader influence on the network. Conversely, cases with lower spreadability, such as Cases 1, 3, and 6, tend to be undetectable. This decreases the likelihood of detection, making these attacks harder to identify and more challenging to address, as will be shown in the next section.

Table 3.4: Qualitative and quantitative evaluation of attack cases across four metrics. Signs (−−, -, +, ++) provide a simplified view of metric magnitudes based on a quantile-based classification.

Case	Visibility(%)		Detectability		Spreadability	
	Value	Qual.	Value	Qual.	Value	Qual.
Case 1	16.67%	−−	0.0005	-	0.1650	-
Case 2	42.32%	+	0.0010	+	0.3668	+
Case 3	33.36%	+	0.0003	−−	0.0267	−−
Case 4	84.71%	++	0.0014	+	0.2275	+
Case 5	16.69%	-	0.0014	+	0.9137	++
Case 6	9.73%	−−	0.0005	−−	0.2077	-
Case 7	50.54%	++	0.0008	-	0.5780	++
Case 8	18.86%	-	0.0135	++	0.0465	−−

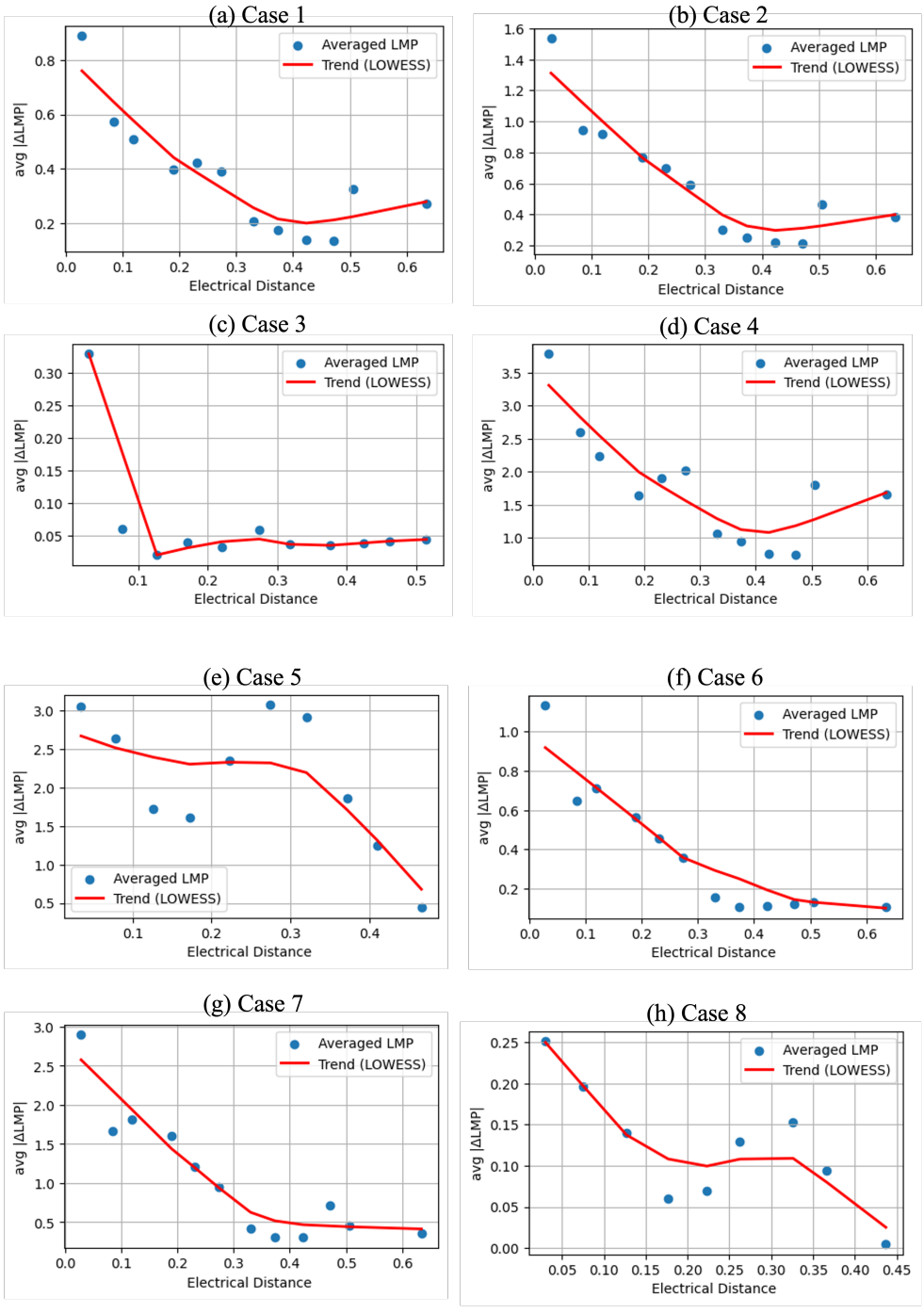


Figure 3.5: Average $\Delta\text{LMP}^{(i)}$ vs. electrical distance from the targeted bus across eight attack cases.

3.5 Case Study

This section presents an anomaly detection case study using our dataset, highlighting the performance of the Geometric Entropy Minimization (GEM) model [92], a multivariate, nonparametric, and unsupervised method shown to perform well in detecting anomalies in high-dimensional data streams. GEM was selected as a strong baseline due to its ability to capture subtle multivariate deviations in temporal data without relying on prior knowledge of attack signatures. Its sensitivity to distributional shifts makes it particularly well-suited for detecting stealthy FDIAs in electricity market settings. GEM estimates baseline statistics offline and applies them in real-time, offering practical advantages for deployment. In our experiments, GEM is configured with parameters $k = 2$, $h = 5$, and a sliding window size of 336. Table 3.5 summarizes its performance across eight test cases, evaluated both before and after a known drift in the data distribution.

The results presented in Table 3.5 demonstrate that, overall, GEM's detection performance before concept drift aligns closely with the detectability metric in most cases, indicating that the model generally captures how statistically observable each attack is. In particular, high $\Delta\text{LMP}^{(i)}$ values tend to correspond to higher detection rates, as larger economic disruptions are more likely to produce detectable anomalies in the system. However, this relationship is not absolute. For example, Case 3 causes a large price shift yet remains nearly undetected. This reflects the fact that detectability depends not just on economic impact but on how the attack manifests across GEM's multivariate inputs. In Case 3, the low spreadability likely caused the effect to remain localized and thus masked in the model's broader statistical view. Additionally, attacks during peak hours show reduced detectability before drift, as the high natural variability in system behavior during those periods makes it difficult for change point-based methods like GEM to distinguish subtle anomalies from normal fluctuations. After drift, as shown in Fig 3.6, GEM's performance degrades significantly. This is because GEM is not inherently adaptive to temporal changes in the LMP distribution. When concept drift occurs, such as due to seasonal dynamics, evolving load patterns, or topology changes, the model continues to rely on outdated baseline statistics, leading to a surge in false positives and reduced overall reliability.

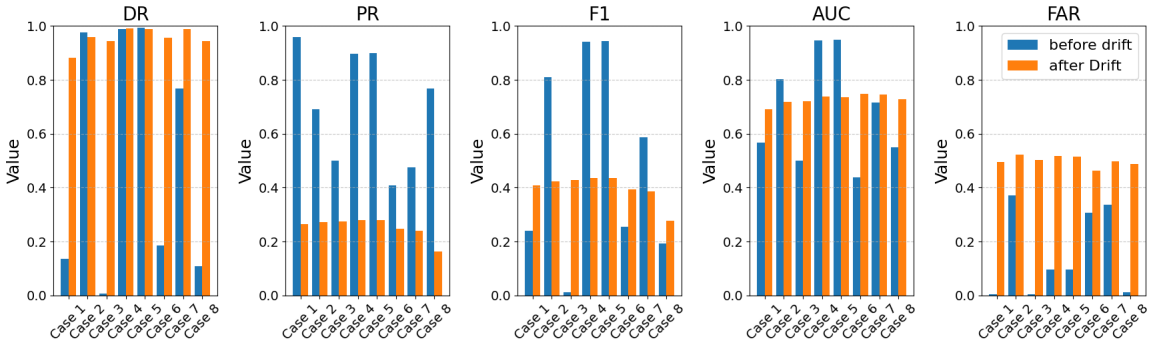


Figure 3.6: Performance of the GEM model before and after drift

Table 3.5: GEM model performance before and after drift

Case	GEM before drift					GEM after drift				
	DR	PR	F1	AUC	FAR	DR	PR	F1	AUC	FAR
1	0.137	0.958	0.240	0.566	0.005	0.881	0.265	0.408	0.692	0.496
2	0.976	0.692	0.810	0.802	0.372	0.958	0.272	0.423	0.718	0.523
3	0.006	0.500	0.012	0.500	0.005	0.943	0.276	0.428	0.721	0.502
4	0.988	0.897	0.941	0.946	0.097	0.991	0.280	0.437	0.737	0.517
5	0.994	0.898	0.944	0.949	0.097	0.988	0.280	0.437	0.736	0.515
6	0.185	0.409	0.254	0.438	0.308	0.956	0.247	0.393	0.747	0.463
7	0.769	0.476	0.588	0.715	0.338	0.989	0.240	0.387	0.746	0.497
8	0.110	0.769	0.192	0.549	0.011	0.945	0.163	0.278	0.728	0.489

3.6 Lessons Learned

Our empirical analysis and case study using the `SMLT` dataset reveal several key insights into the behavior of stealthy FDIA scenarios and the broader challenges of monitoring electricity markets. We hope these lessons, together with the dataset introduced in this work, will support the research community in advancing future efforts toward more secure, reliable, and fair electricity market operations.

1. **Importance of market-level data.** LMP data can serve as an effective signal for monitoring grid behavior and detecting FDIA attacks, particularly those aimed at financial gain. While such attacks may bypass traditional security mechanisms like BDD, they often produce noticeable distortions at the market level, making LMPs a valuable source for identifying anomalies in system operation.

2. **FDIA impacts are rarely isolated.** Attacks on a specific bus or node tend to propagate throughout the system, especially when they affect shift factors or dispatch outcomes. While the attacked bus often exhibits the most significant price distortion, neighboring and electrically close buses are also impacted, highlighting the system-wide footprint of even localized attacks.

3. **Importance of attack duration and timing.** Short-duration attacks, particularly those occurring during peak hours, are more difficult to detect due to the natural volatility of LMPs during those periods. Moreover, attacks launched during times of distributional drift or operational changes can be masked by ongoing fluctuations, making detection even more challenging.

4. **Importance of drift-aware models.** Our observations confirm that LMP distributions are non-stationary and evolve continuously over time. This highlights the need for adaptive, temporally aware detection models that can accommodate regime shifts, seasonal variability, and concept drift in both market dynamics and system behavior.

3.7 Conclusion

This chapter introduced SMLT, a synthetic benchmarking dataset designed to systematically evaluate stealthy FDIAs targeting Locational Marginal Prices (LMPs). By combining physics-consistent power system modelling, market-clearing mechanisms, and literature-grounded manipulation strategies, the dataset provides a reproducible and extensible platform for studying adversarial behaviour in electricity markets. The spatio-temporal analyses demonstrate that even bounded and carefully crafted attacks can induce subtle yet structured distortions in LMP distributions, motivating the need for robust anomaly detection frameworks.

Limitations and generalisability. Several limitations should be acknowledged. SMLT remains a synthetic dataset generated through simulation. Although it is constructed using a realistic NPCC-based market benchmark and OPF-based LMP calculation, it may still contain biases induced by the modelling assumptions of the benchmark system, the selected load profiles, the OPF formulation, and the predefined attack scenarios. In particular, the simulation does not fully capture all real market rules, operator interventions, reserve products, strategic bidding behaviour, contingency events, or unplanned outages. Therefore, models trained only on SMLT may risk learning simulation-specific patterns rather than general indicators of LMP manipulation.

To reduce overfitting to synthetic artefacts, the detection experiments are designed to evaluate general attack effects rather than memorised attack signatures. Models are trained on normal operating data and evaluated on unseen manipulated intervals, attack magnitudes, and operating conditions. The dataset also includes multiple attack vectors, attack durations, and target buses to reduce dependence on a single manipulation pattern. Nevertheless, validation on additional real-world market data remains necessary to further assess generalisability.

The current threat vector does not cover all possible forms of market manipulation. In particular, it does not explicitly model coordinated multi-attacker strategies, adaptive attacks that respond to the detector, adversarial machine-learning attacks, strategic bidding manipulation without physical data corruption, forecast manipulation, or sophisticated attacks that combine topology, parameter, measurement, load, and market-bid manipulation.

Future versions of the dataset could be extended by incorporating emerging and composite attacks. This could include coordinated manipulation across multiple buses, adaptive attacks with changing magnitude and duration, attacks on forecast inputs, combined cyber-market strategies, and richer multi-layer data sources such as SCADA measurements, PMU streams, renewable forecasts, demand forecasts, and bidding data. These extensions would improve the dataset's ability to support broader benchmarking of data-driven market-monitoring methods.

Despite these limitations, SMLT provides the necessary benchmarking infrastructure for developing and evaluating data-driven detection methods throughout the remainder of this thesis. The dataset enables controlled experimentation under realistic yet reproducible manipulation scenarios, forming the empirical foundation for the adaptive detection frameworks introduced in the following chapters.

Chapter 4

Incremental Prediction-Error Monitoring for Real-Time Market Anomaly Detection

4.1 Introduction

Building on the benchmarking framework introduced in Chapter 3, this chapter addresses the problem of detecting stealthy Locational Marginal Price (LMP) manipulation in streaming electricity market data. Financially motivated False Data Injection Attacks (FDIAs) target measurement streams transmitted from remote terminal units (RTUs) to the control center, with the objective of influencing state estimation outcomes and consequently altering LMP calculations [1,38,62]. These attacks are typically crafted to remain stealthy by bypassing conventional bad data detection mechanisms and preserving statistical plausibility. As a result, manipulated LMP trajectories may closely resemble legitimate market fluctuations, making detection at the market level particularly challenging.

Contribution. Existing defense mechanisms in energy markets can broadly be categorized into (i) protection-based strategies and (ii) detection-based approaches. Protection mechanisms often rely on securing selected measurements or optimizing PMU placement under strong trust assumptions [93, 94]. However, such approaches reduce redundancy, assume partial infrastructure immunity, and may not scale under realistic adversarial conditions. Detection mechanisms, on the other hand, typically operate at the grid level using physical system measurements and parameter-consistency checks [68, 95, 96]. These methods are primarily designed to detect anomalies in underlying power system states rather than market-level distortions reflected in LMP time series.

In deregulated markets, financially motivated manipulation ultimately manifests as deviations in LMP trajectories. LMPs are inherently temporal signals with strong autocorrelation and evolving market dynamics. Therefore, effective detection requires models that operate directly on streaming LMP data, adapt incrementally over time, and distinguish abrupt adversarial shifts from normal market variability.

Contribution. To address this challenge, we propose an incremental, unsupervised detection framework based on cumulative prediction error analysis. A Long Short-Term Memory (LSTM) model is trained to capture normal LMP temporal dynamics, and structural deviations are identified by monitoring accumulated prediction errors through a change-point detection mechanism. By aggregating prediction discrepancies over time, the proposed approach enhances sensitivity to subtle but persistent manipulation patterns while controlling false alarm rates. We benchmark the framework against two baselines: (i) a standard prediction-error thresholding model and (ii) the cumulative sum (CuSum) method. Experimental results demonstrate improved adaptability and reduced detection delay across varying attack intensities.

Although deep learning models such as Transformers and spatio-temporal Graph Neural Networks (GNNs) are powerful tools for forecasting and representation learning, they are not the primary methodological direction of this thesis. The focus of the thesis is on data-driven but interpretable monitoring frameworks for stealthy LMP manipulation, where the main requirements are unsupervised operation, limited reliance on labelled attack data, controlled false alarm rates, bounded detection delay, robustness to concept drift, and scalability to large electricity-market networks. These requirements are difficult to satisfy using purely deep-learning-based detectors, which often require large representative training datasets, extensive hyperparameter tuning, and may provide limited interpretability for market operators. Therefore, deep learning is used only in a limited role in this chapter, as a forecasting component for constructing prediction residuals, while the detection decision is handled through change-point monitoring. The later chapters follow the same principle by prioritising nonparametric, drift-aware, and graph-informed detection mechanisms over fully supervised deep-learning architectures.

The remainder of this chapter is organized as follows. Section 4.2 formulates the problem and outlines the detection setting. Section 4.3 details the proposed incremental framework. Section 4.4 presents experimental evaluation and comparative analysis. Section 4.5 concludes the chapter and discusses limitations.

Table 4.1: Notation used in Chapter 4

Symbol	Description
LMP_i	Locational Marginal Price at bus i
λ	Lagrangian multiplier associated with the power balance constraint
μ_l	Lagrangian multiplier associated with transmission constraint l
GSF_{l-i}	Generation Shift Factor between line l and bus i
$\hat{x}$	Estimated state vector of the power system
θ_i	Voltage angle at bus i
z	Measurement vector collected from the power system
$h(\cdot)$	Nonlinear measurement function

Symbol	Description
e	Measurement noise vector
σ^2	Variance of measurement noise
a	Attack vector injected into the measurement data
z_a	Compromised measurement vector after attack
$\hat{x}_a$	Manipulated state estimate after attack
c	State deviation caused by the attack
τ	Time at which an attack or change occurs
T	Stopping time of the detection algorithm
$\mathcal{F}_\tau$	Information set (filtration) available up to time τ
$J(T)$	Worst-case expected detection delay
α	Lower bound on expected run length before false alarm
p_0	Probability density function of normal LMP data
p_1	Probability density function of attacked LMP data
y_t	Observed LMP value at time step t
$\hat{y}_t$	One-step-ahead LMP forecast at time step t
$f(\cdot)$	LSTM forecasting function
w	Look-back window length used for forecasting
e_t	Forecasting residual at time step t
uc_t	Upper cumulative sum statistic at time step t
lc_t	Lower cumulative sum statistic at time step t
K	Reference value controlling CuSum sensitivity
λ^+	Upper detection threshold
λ^-	Lower detection threshold
μ_a	Magnitude of the attack-induced residual mean shift
H	Number of hidden units in the LSTM model
D	Input dimension of the LSTM model
N	Number of monitored nodes or buses
t	Time index of the LMP sequence

4.2 Problem Formulation

Consider a distributed power system where the LMP at a distributed node i is determined by the following relationship:

$$LMP_i = \lambda + \sum_{l=1} GSF_{l-i} \cdot \mu_l, \tag{4.1}$$

where λ is the Lagrangian multiplier for the equality constraint ensuring power balance, and μ_l represents the Lagrangian multiplier for the l-th transmission constraint. Accurate computation of λ and μ_l relies on precise state estimation (SE) of the power system, which is fundamental to maintaining system stability and market efficiency.

We consider a DC load-flow model in which voltage magnitudes are assumed fixed and reactive power and losses are neglected. Therefore, the state vector is represented only by bus voltage angles, i.e.,

$$\hat{x} = [\theta_1, \ldots, \theta_N]^T, \tag{4.2}$$

where $\theta_1, \ldots, \theta_N$ are the bus voltage angles to be estimated. Let z represent the vector of real-time measurements collected from the power system, including power flows and injections. The relationship between measurements and the state vector can be expressed as:

$$z = h(\hat{x}) + e, \tag{4.3}$$

where $h(\cdot)$ is the nonlinear measurement function and $e \sim \mathcal{N}(0, \sigma^2 I)$ represents measurement noise. The state estimation process solves for $\hat{x}$ such that the measurements are best matched to the model.

4.2.1 FDIA in the Energy Market

However, this reliance on real-time measurements exposes the SE process to FDIAs, in which an adversary intentionally manipulates the measurements to compromise the estimated state $\hat{x}$. Suppose the malicious attack occurs by introducing an attack vector a into the measurement data. The compromised measurements are then given by:

$$z_a = z + a, \tag{4.4}$$

resulting in a manipulated system state:

$$\hat{x}_a = \hat{x} + c, \tag{4.5}$$

where c represents the deviation introduced in the state estimation. To evade Bad Data Detection (BDD) mechanisms, the attack vector a is carefully crafted as:

$$a = h(\hat{x} + c) - h(\hat{x}), \tag{4.6}$$

ensuring that the residual norm remains unchanged. Specifically, the norm of the residual before and after the attack is:

$$\|z_a - h(\hat{x}_a)\| = \|z - h(\hat{x})\|. \tag{4.7}$$

This property allows the attack to bypass the BDD mechanism, rendering it undetectable through conventional residual-based methods. Consequently, the adversary can manipulate critical operational decisions, including LMP calculations, leading to potential economic and operational disruptions. The goal of this study is to investigate the vulnerabilities of the SE process to FDIAs and to propose robust detection mechanisms to mitigate their impact.

Our objective is to identify manipulated LMPs as quickly as possible following their occurrence at τ, as any alteration in the underlying measurements will manifest as anomalies at the market level. Change point detection, leveraging statistical differences between pre- and post-change data, offers an effective framework for achieving this goal.

4.2.2 Change Point Detection

In change point detection, data is observed in a sequence, and a stopping rule is used to identify when a change occurs. Sequential methods focus on minimizing the delay in detecting that change. One common way to measure performance is through the worst-case average detection delay, a concept introduced by Lorden [97]:

$$J(T) = \sup_{\tau} \operatorname{ess\,sup}_{\mathcal{F}_\tau} \mathbb{E}_\tau \left[(T-\tau)^+ \mid \mathcal{F}_\tau \right], \tag{4.8}$$

where $J(T)$ measures the detection delay in the worst-case scenario, considering all observations up to the change-point τ. The term $\mathcal{F}_\tau$ refers to the accumulated information available up to that point. While minimizing $J(T)$ helps ensure that anomalies are detected quickly, it is also crucial to manage the running length during normal conditions to prevent false alarms from occurring too frequently.

To balance these objectives, the sequential change detection problem is formulated as:

$$\inf_{T} J(T) \quad \text{subject to} \quad \mathbb{E}_\infty[T] \geq \alpha, \tag{4.9}$$

where $\mathbb{E}_\infty[T]$ represents the expected stopping time when no change occurs ($\tau = \infty$). The parameter α is a threshold that ensures a sufficiently long average running time under normal conditions.

For the LMP manipulation problem, let the pre and post attack probability density functions (PDFs) of the LMP data stream be denoted by p_0 and p_1, respectively. When these PDFs are fully specified, the quickest detection problem can be optimally solved using the well-known CuSum test. However, the characteristics of the time series are influenced by dynamic system conditions, making it challenging to precisely specify the PDFs p_0 and p_1. These complexities arise from variations in grid topology, demand-supply fluctuations, and market mechanisms, which introduce significant uncertainty into the distributional assumptions.

To address these challenges, this chapter adopts a prediction-based framework that models the expected LMP trajectory without relying on explicit pre- and post-change distributional assumptions. Attacks are then detected indirectly through persistent deviations between observed and predicted LMP values.

4.3 Methodology

4.3.1 Framework Overview

The proposed detection framework consists of two sequential stages. First, a forecasting model is trained on benign LMP observations to generate one-step-ahead predictions of the market signal. Second, the discrepancy between the observed and predicted values is transformed into a residual sequence and monitored online using a two-sided cumulative sum (CuSum) detector. This decomposition separates normal temporal prediction from anomaly decision-making. The forecasting stage provides a data-driven estimate of expected LMP behavior, while the sequential detection stage accumulates residual evidence over time to identify persistent deviations. An alarm is triggered when the upper or lower detection statistic crosses its corresponding decision threshold. Fig. 4.1 illustrates the overall structure of the proposed framework.

4.3.2 LSTM-Based LMP Forecasting

The first stage of the proposed framework aims to learn the normal temporal behavior of the LMP series and generate one-step-ahead forecasts. This forecasting stage is important because raw LMP trajectories may exhibit substantial short-term variability even in the absence of attacks. Directly detecting anomalies from the observed price sequence can therefore lead to a large number of false alarms. To reduce this effect, the detector operates on forecasting residuals rather than on the raw LMP signal itself. In this way, the prediction model serves as a baseline estimator of expected market behavior, while deviations from that baseline are passed to the anomaly detection stage.

To model the temporal dependence of LMP values, we employ a Long Short-Term Memory (LSTM) network [98]. LSTM networks are widely used for time-series forecasting because they can capture nonlinear sequential dependencies and preserve information over longer horizons through their gated memory structure. This makes them suitable for LMP data, which typically exhibit temporal correlation, local trends, and fluctuations driven by changing operating conditions.

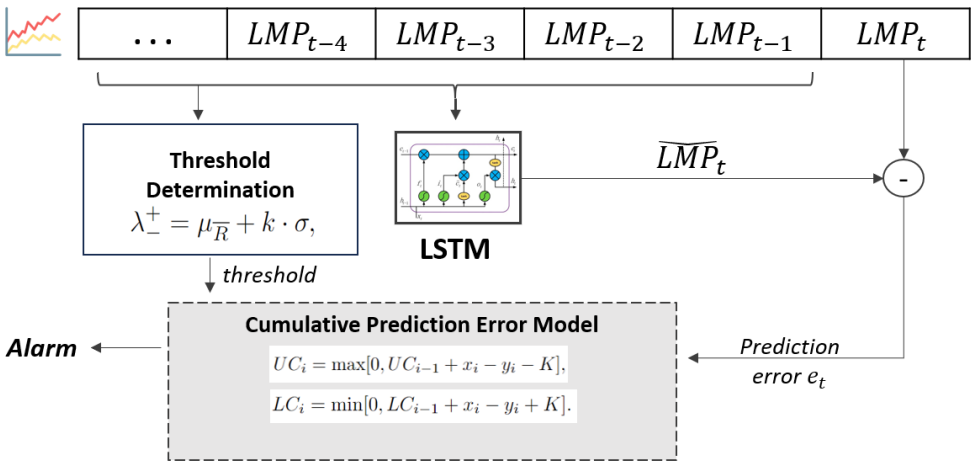


Figure 4.1: Overview of the proposed residual-based sequential detection framework for LMP manipulation.

In the general formulation, the forecasting task can be applied to any monitored LMP signal or to a selected subset of market nodes. Let y_t denote the observed LMP at a monitored node at time t. Given a look-back window of length w, the forecasting model learns a mapping from historical observations to a one-step-ahead prediction:

$$\hat{y}_t = f(y_{t-w}, y_{t-w+1}, \ldots, y_{t-1}), \tag{4.10}$$

where $f(\cdot)$ denotes the LSTM predictor and $\hat{y}_t$ is the one-step-ahead forecast. The prediction residual is then computed as

$$e_t = y_t - \hat{y}_t, \tag{4.11}$$

where e_t measures the discrepancy between the observed and forecasted LMP at time t. Under normal operating conditions, these residuals are expected to remain relatively small and approximately centred around zero. In contrast, persistent manipulation of the market signal may induce sustained deviations in the residual sequence. Therefore, the residual stream provides a more informative input for sequential anomaly detection than the raw LMP series itself.

4.3.3 Residual-Based Sequential Detection

The second stage of the proposed framework monitors the forecasting residuals in order to detect persistent deviations from normal LMP behavior. While isolated prediction errors may occur because of normal market variability or imperfect forecasting, attack-induced manipulations are expected to produce sustained shifts in the residual sequence. To distinguish between these two cases, we employ a sequential change-detection mechanism based on the cumulative sum (CuSum) procedure.

Let e_t denote the residual at time step t obtained from the forecasting stage. The detector maintains two cumulative statistics: an upper statistic for positive deviations and a lower statistic for negative deviations. These statistics are updated recursively as

$$uc_t = \max\left(0,\ uc_{t-1} + e_t - K\right), \tag{4.12}$$

$$lc_t = \min\left(0,\ lc_{t-1} + e_t + K\right), \tag{4.13}$$

where K is a reference value that controls the sensitivity of the detector. The parameter K determines how much deviation is treated as normal fluctuation before cumulative evidence begins to build. Small values of K make the detector more sensitive to weak changes, whereas larger values reduce sensitivity to minor variations and noise.

The upper statistic uc_t is designed to accumulate evidence of persistent positive shifts in the residual sequence, while the lower statistic lc_t captures persistent negative shifts. The reset operations in the CuSum recursions ensure that short-lived fluctuations do not accumulate indefinitely. As a result, the detector responds primarily to sustained deviations rather than to isolated forecasting errors.

A change is declared when either cumulative statistic crosses its corresponding decision threshold. Specifically, an alarm is raised when

$$uc_t \geq \lambda^+ \quad \text{or} \quad lc_t \leq \lambda^-, \tag{4.14}$$

where λ^+ and λ^- denote the upper and lower alarm thresholds, respectively. In this formulation, positive residual shifts may indicate an upward manipulation of LMP values, whereas negative residual shifts may indicate downward distortion or abnormal suppression of expected prices.

Algorithm 2 summarizes the online detection procedure. At each time step, the observed LMP is first compared with the corresponding one-step-ahead forecast to obtain the residual. The two CuSum statistics are then updated recursively, and an alarm is triggered once sufficient evidence has accumulated in either direction. Since each update requires only a small number of arithmetic operations, the procedure is suitable for real-time deployment in streaming market environments.

4.3.4 Detection Delay and False-Alarm Trade-off

The proposed detector exhibits the standard trade-off of sequential cumulative monitoring schemes. Under normal operating conditions, the residuals are expected to remain approximately centered around zero. As a result, the CuSum statistics tend to reset frequently, which helps limit false alarms. After an attack occurs, if the residuals experience a persistent mean

Algorithm 2 Residual-Based Sequential Detection

Require: Observed LMP sequence $\{y_t\}_{t=1}^{T}$, one-step-ahead forecasts $\{\hat{y}_t\}_{t=1}^{T}$, upper threshold λ^+, lower threshold λ^-, reference value K

Ensure: Alarm time indicating anomalous LMP behavior

Initialize $uc_0 = 0$ and $lc_0 = 0$
for $t = 1$ **to** T **do**
    $e_t = y_t - \hat{y}_t$
    $uc_t = \max(0, uc_{t-1} + e_t - K)$
    $lc_t = \min(0, lc_{t-1} + e_t + K)$
    if $uc_t \geq \lambda^+$ **or** $lc_t \leq \lambda^-$ **then**
        return Anomalous LMP detected at time t
    end if
end for

shift, then the corresponding CuSum statistic acquires a positive drift[1] on average and moves progressively toward the alarm threshold. In that case, the detection delay is approximately inversely related to the size of the post-change drift and directly related to the threshold. Therefore, larger thresholds improve robustness to false alarms but increase delay, whereas stronger attacks lead to faster detection.

4.3.5 Complexity Analysis

The CuSum detector requires constant computation per time step, resulting in a total time complexity of $O(t)$ over a sequence of length t for a single node. Each iteration involves only basic arithmetic operations and comparisons, resulting in negligible computational overhead. However, the overall runtime of the proposed framework is dominated by the LSTM forecasting stage. For an LSTM with H hidden units and input dimension D, the forward-pass computational complexity per time step is approximately $O(H(H+D))$, due to the matrix–vector multiplications and gate operations within the recurrent layer. In our implementation ($H = 24$, $D = 1$), this simplifies to $O(H^2)$ operations per prediction. Consequently, the end-to-end computational complexity for a single node over t time steps is $O(tH^2)$. Extending the framework independently to N nodes results in a total complexity of $O(NtH^2)$. Therefore, although the CuSum detector itself is computationally lightweight, the scalability of the overall system is primarily constrained by the forecasting model.

[1] In this chapter, "drift" refers to the directional movement of the cumulative detection statistic caused by a residual mean shift. It is not used here to denote concept drift in the data distribution, which is formally introduced and addressed in Chapter 5.

4.4 Experimental Evaluation

4.4.1 Experimental Setup

The experiments are conducted on the NPCC test system using the `MATPOWER` simulation environment. The attack scenarios and corresponding LMP time series used in this chapter are derived from the `SMLT` dataset introduced in Chapter 3. In particular, we focus on the two transmission line rating (TLR) manipulation cases (Case 1 and Case 2), which were shown to generate stealthy yet economically meaningful LMP distortions.

Due to the spatial variability of LMP dynamics across the network, we focus on a representative location (Bus 115) where the attack impact is most pronounced. The resulting time series contains 8,784 hourly LMP observations, which are used for training and evaluation of the prediction and detection models. The dataset is partitioned into 80% for training, 10% for validation, and 10% for testing.

In the experimental implementation, the forecasting task is applied to the hourly LMP sequence at Bus 115. The look-back window is set to $w = 23$, so that the previous 23 hourly observations are used to predict the LMP for the next hour. The LSTM is implemented in a stateless configuration with a single hidden layer of 24 units.

To evaluate anomaly detection performance, manipulated LMP signals are generated by simulating TLR attacks within the OPF framework. In this scenario, an attacker associated with a generator at Bus 115 manipulates the reported transmission line rating r to create artificial congestion. This congestion alters the dispatch solution and leads to increased LMP values at nearby buses. The attack magnitude is varied from $0.1r$ to $0.3r$ to simulate different levels of manipulation intensity. Lower attack magnitudes correspond to more stealthy attacks that generate smaller but still profitable price distortions. Following common assumptions in parameter-based attacks, once the manipulated parameter is introduced, it persists for approximately one week before the system parameters are corrected [33].

The CuSum detection threshold is calibrated using the most recent 168 hours of normal (attack-free) LMP observations prior to activating the detector. The reference parameter K, which controls the sensitivity of the CuSum statistic, is evaluated over the range 0.1 to 1.0. Empirical tuning shows that $K = 0.2$ provides the best balance between detection rate and false alarm rate. The resulting decision thresholds are $|\lambda^-| = \lambda^+ = 3.5038$. Since these thresholds are derived solely from normal LMP behavior, they are applied consistently when evaluating both attack scenarios.

4.4.2 Baselines

Two baseline methods are used to evaluate the effectiveness of the proposed approach: (1) the Prediction Error Anomaly Detection method based on LSTM forecasting [99], and (2) the classical CuSum sequential change detection algorithm [100]. These baselines represent two complementary families of anomaly detection techniques commonly used in time-series monitoring. The prediction-error approach detects anomalies through deviations between predicted and observed values, while CuSum provides a statistically grounded framework for detecting distributional changes in streaming data.

The proposed method integrates elements from both paradigms by combining prediction-based monitoring with sequential statistical detection. Therefore, comparing against these two baselines provides a meaningful and representative benchmark.

To ensure a fair comparison, all models are trained and evaluated using the same dataset partitions and experimental protocol. Hyperparameters for each method are selected using validation data prior to testing, and no baseline is given additional information about the attack timing or structure. This ensures that performance differences reflect the capabilities of the detection methods rather than differences in experimental configuration.

4.4.3 Metrics

To evaluate the performance of our detector, we use the following key metrics:

$$\text{Recall (Detection Rate)} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$

where TP is the number of true positives and FN is the number of false negatives.

$$\text{False Alarm Rate (FAR)} = \frac{\text{FP}}{\text{FP} + \text{TN}},$$

where FP is the number of false positives and TN is the number of true negatives.

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{DR}}{\text{Precision} + \text{DR}},$$

where Precision is defined as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}.$$

$$\text{Detection Delay} = \frac{\sum_{i=1}^{n} (\text{Detection Time}_i - \text{Attack Time}_i)}{n},$$

where Detection Time_i is the time or step when the i-th attack is detected, Attack Time_i is the time or step when the i-th attack actually occurs, and n is the total number of attacks.

4.4.4 Impact of Attack Intensity on LMP Change

In this experiment, we explore the impact of varying attack intensities on LMP changes and the total financial loss over the course of one week. The first column represents the normal LMP (without any attack, where the attack intensity = 0), which is 15.6 $/MWh. Notably, with an attack intensity of 0.13, the attacker causes a modest increase of only 0.05 $/MWh in the average timestep LMP, which translates to a potential gain of up to 90$/MWh over a week. Furthermore, at an attack intensity of 0.2, the attacker increases the normal price by 0.89$/MWh, yielding potential benefits of up to 287$/MWh per week. These changes are subtle and difficult to distinguish from normal price fluctuations. This finding is consistent with the results in [38], which state that attack vectors with intensities ranging from 0.1 to 0.2 are particularly challenging to detect.

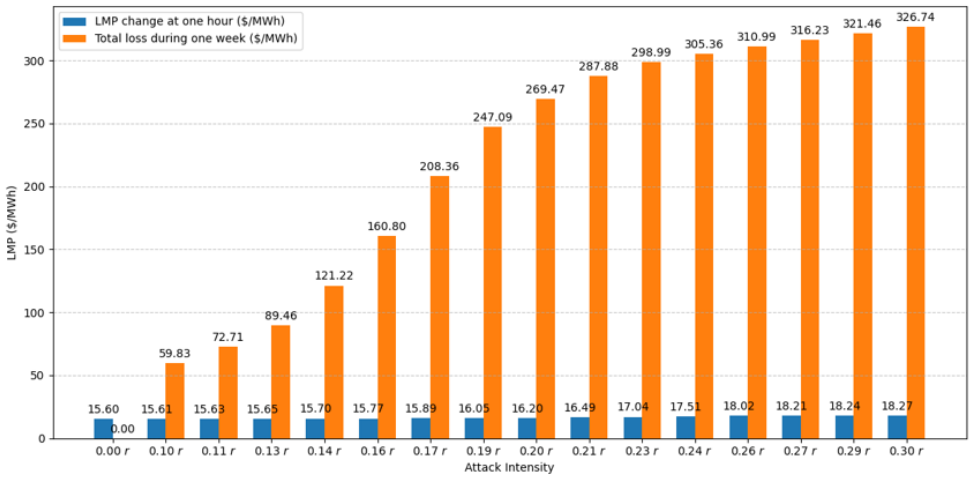


Figure 4.2: Impact of attack intensity on LMP Change ($/MWh) and total loss ($) during a week.

4.4.5 Detection Performance over Different Attack Intensity

Table 4.2 presents the detection performance results. For higher attack intensities, the CuSum model achieves a 93% DR, while the Prediction Error model achieves 63% DR. However, both models exhibit high FAR of 45% and 29%, respectively. For lower attack intensities, both baseline models fail to distinguish between normal and manipulated LMPs. In contrast, the proposed model consistently outperforms the baselines in detecting both high and stealthy attack intensities, achieving DRs of 98% and 92%, respectively, with a significantly lower FAR ranging from 8% to 7% across all cases. This superiority is further reflected in the F1 score, where our model achieves 84%, compared to approximately 47% for both baselines. These results align with our findings, where the manipulated LMPs are more easily detected using our model, especially for stealthy attacks, as shown in Fig 4.3.

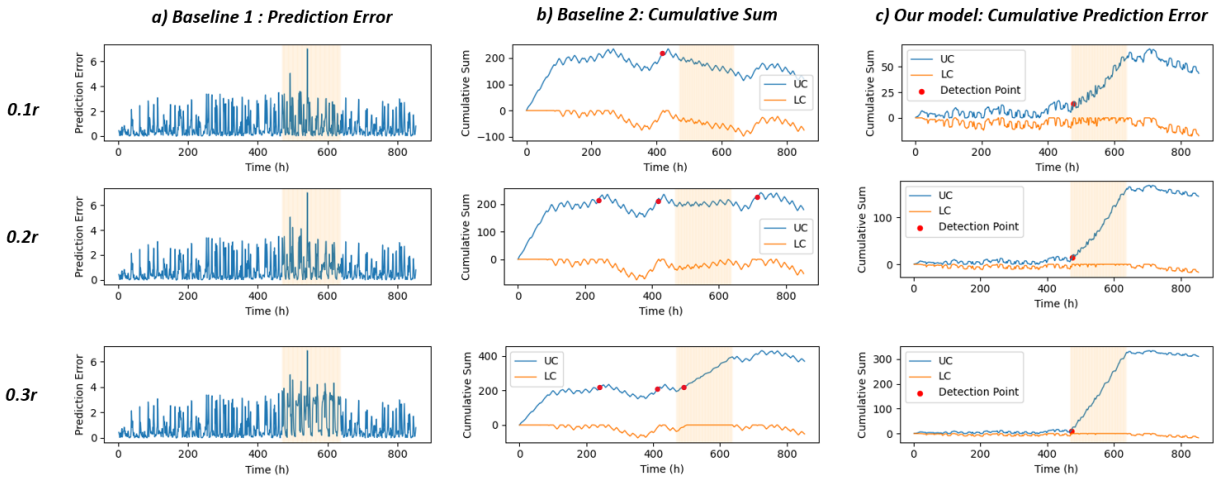
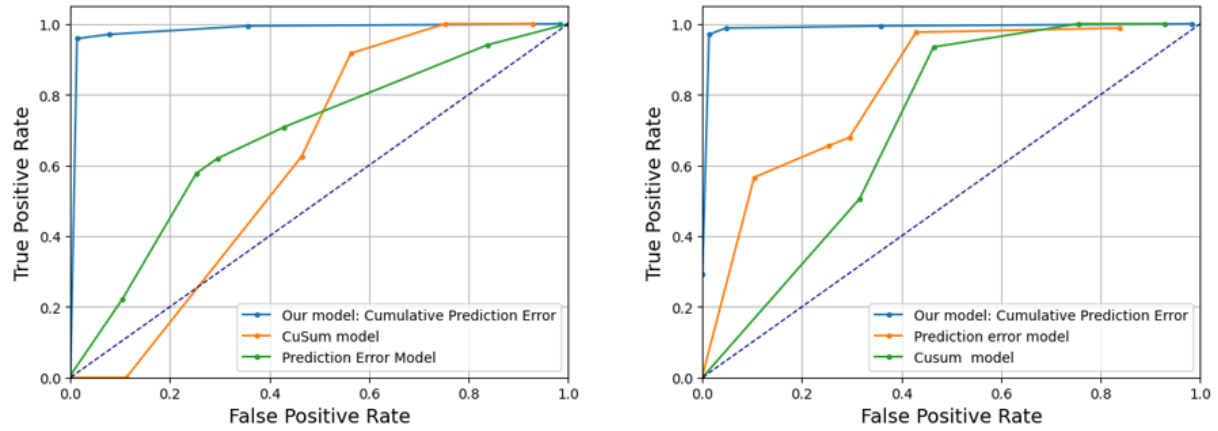


Figure 4.3: Visibility of manipulated LMP across different models at various attack intensities.

4.4.6 Detection Performance over Different Thresholds

We evaluate the performance of all models across different thresholds at two selected attack intensities: 0.15r and 0.3r. Fig.4.4 presents the ROC curves for each model. While the baseline models perform better at higher attack intensities, our model consistently outperforms them, particularly in detecting both high and stealthy attacks. This indicates that the proposed approach remains effective in accurately identifying manipulated LMPs, even in the presence of more subtle attack scenarios.



(a) Attack intensity 0.15r

(b) Attack intensity 0.30r

Figure 4.4: ROC curves comparing the detection performance of the proposed model and the baseline model at two different attack intensities.

4.4.7 Detection Delay

We analyze our model's detection delay under varying attack intensities. The detection delay is calculated as:

$$\text{Detection Delay} = \frac{1}{n}\sum_{i=1}^{n}(\text{Detection Time}_i - \text{Attack Time}_i),$$

where Detection Time_i represents the timestep when the i-th attack is detected, Attack Time_i denotes the timestep when the i-th attack begins, and n is the total number of attacks.

The results, shown in Table 4.2, indicate that the model takes longer to detect attacks at lower intensities. For instance, at an attack intensity of $0.1r$, the detection delay is 6 timesteps. However, as the attack intensity increases, the manipulated LMP becomes more easily distinguishable in the residual sequence, leading to a significant reduction in detection delay. At higher attack intensities, the detection delay is reduced to as few as 2 timesteps. This result is consistent with the behavior of sequential cumulative detectors: as attack intensity increases, the induced residual shift becomes larger, so the decision statistic crosses the threshold faster. Moreover, the false alarm rate remains relatively stable across attack intensities, indicating that improved responsiveness is not obtained at the expense of normal-condition stability.

Table 4.2: Detection performance of our model over different attack intensities compared with two baseline models.

Intensity	**Baseline 1: Prediction Error**				**Baseline 2: CuSum**				**Proposed Model**				
	Recall	Precision	F1	FAR	Recall	Precision	F1	FAR	Recall	Precision	F1	FAR	Delay
0.10r	0.435	0.266	0.330	0.293	0.000	0.000	0.000	0.146	0.917	0.733	0.815	0.081	6
0.12r	0.464	0.279	0.348	0.295	0.107	0.079	0.091	0.147	0.970	0.748	0.845	0.080	5
0.14r	0.566	0.320	0.409	0.295	0.464	0.294	0.429	0.464	0.970	0.751	0.847	0.078	5
0.17r	0.643	0.348	0.452	0.295	0.792	0.294	0.429	0.464	0.976	0.756	0.852	0.077	4
0.20r	0.655	0.353	0.458	0.295	0.905	0.323	0.476	0.464	0.976	0.756	0.852	0.077	4
0.22r	0.685	0.363	0.479	0.295	0.935	0.330	0.488	0.464	0.982	0.764	0.859	0.074	3
0.24r	0.673	0.359	0.474	0.295	0.935	0.330	0.488	0.464	0.982	0.764	0.859	0.074	3
0.27r	0.661	0.355	0.462	0.295	0.935	0.330	0.488	0.464	0.988	0.765	0.862	0.074	2
0.30r	0.679	0.361	0.471	0.295	0.935	0.330	0.488	0.464	0.988	0.765	0.862	0.074	2

4.5 Conclusions

This chapter presented an incremental, unsupervised framework for detecting financially motivated manipulation in streaming Locational Marginal Price (LMP) data. By combining LSTM-based temporal prediction with cumulative error monitoring through a change-point detection mechanism, the proposed approach enables near–real-time identification of structural deviations in market signals. Experimental evaluation demonstrates improved detection performance and false alarm stability compared to standard prediction-error thresholding and CuSum baselines, particularly under low- and medium-intensity stealthy attacks.

A key strength of the framework lies in its reliance solely on market-level LMP time series, allowing detection without requiring direct access to system-level measurements. However, several limitations remain. First, the current formulation is univariate and models each LMP signal independently, without explicitly incorporating cross-location dependencies. Second, the detection mechanism assumes relatively stable baseline dynamics and does not explicitly account for gradual distributional shifts caused by seasonal patterns, load variability, or evolving market regimes. Third, threshold selection remains fixed, which may limit robustness under prolonged non-stationary conditions.

These limitations motivate the next chapter, which addresses the challenge of detecting manipulation under non-stationary market behavior. Specifically, Chapter 5 develops a drift-aware anomaly detection framework capable of distinguishing natural regime shifts from adversarial manipulation in evolving LMP signals.

Chapter 5

Adaptive Anomaly Detection in Non-Stationary Market Signals

5.1 Introduction

This chapter addresses a critical practical challenge in electricity market monitoring: non-stationarity of Locational Marginal Price (LMP) signals. While real-time change-point detection can identify abrupt structural deviations, evolving market dynamics may gradually alter the statistical properties of LMPs, complicating the distinction between legitimate regime shifts and adversarial manipulation.

Detecting LMP manipulation in this setting presents two fundamental challenges. **CH1:** Stealthy attacks are deliberately designed to induce small, incremental deviations that remain statistically subtle. Even minor price perturbations (e.g., sub-dollar shifts) can generate substantial financial gains over sustained periods while preserving overall distributional characteristics. Such bounded manipulations are difficult to detect using fixed-threshold or stationary assumptions. **CH2:** Electricity markets are inherently non-stationary. Renewable generation variability, seasonal demand patterns, regulatory changes, and operational constraints cause the statistical properties of LMPs, such as mean, variance, and temporal dependence—to drift over time. Consequently, anomaly detection models trained under one regime may become outdated, leading to increased false alarms or missed detections when natural distributional shifts occur.

Related Work. Existing defense mechanisms largely rely on protection strategies or model-based detection methods grounded in predefined system representations [61–63, 66]. Although effective under specific assumptions, these approaches struggle under evolving market conditions and are particularly vulnerable to stealthy attacks that mimic legitimate variability. Furthermore, many existing anomaly detection models implicitly assume stationarity, limiting their robustness in realistic market environments where concept drift is unavoidable.

Contribution. To address these challenges, we propose a drift-aware, data-driven framework for detecting stealthy manipulation in multivariate LMP time series. The proposed methodology integrates two complementary components: (1) an online anomaly detection mechanism based on Geometric Entropy Minimization (GEM), which incrementally identifies deviations from the evolving baseline distribution; and (2) a window-based concept drift detection module that monitors distributional changes and triggers model adaptation when significant regime shifts are detected.

By explicitly modeling concept drift alongside anomaly detection, the framework aims to distinguish between gradual structural evolution of market behavior and abrupt adversarial manipulation.

The main contributions of this chapter are as follows:

1. A drift-aware, unsupervised framework for detecting manipulation in multivariate LMP time series.

2. A nonparametric change-point detection formulation that operates without restrictive distributional assumptions.

3. An adaptive anomaly detection mechanism capable of separating adversarial manipulation from natural regime shifts.

4. Comprehensive evaluation on both synthetic (SMLT) and real-world datasets under varying attack intensities.

The remainder of this chapter is organized as follows. Section 5.2 formulates the detection problem under non-stationary conditions. Section 5.3 presents the proposed drift-aware framework. Section 5.5 evaluates performance and robustness under diverse scenarios. Section 5.6 concludes the chapter and discusses implications for scalable market monitoring.

Table 5.1: Notation used in Chapter 5

Symbol	**Description**
t	Discrete time index
x_t	Multivariate LMP observation at time t
p	Dimensionality of the data stream (number of LMP nodes)
t_1	Change-point time at which anomaly or drift begins
f_{x_0}	Pre-change probability density function of x_t
f_{x_1}	Post-change probability density function of x_t
T	Stopping time at which a change is declared
$J(T)$	Worst-case average detection delay
$\mathcal{F}_t$	Information history available up to time t

Symbol	Description
β	Lower bound on the average false alarm period
ℓ_t	Log-likelihood ratio at time t
g_t	Sequential decision statistic for change detection
h	Detection threshold for anomaly declaration
X	Nominal historical dataset used to construct the baseline
S_1	Reference subset of nominal data
S_2	Secondary nominal subset used for baseline scoring
N	Total number of samples in the nominal dataset
N_1	Number of samples in S_1
N_2	Number of samples in S_2
x_j	Sample point in the nominal subset S_2
x_{test}	Incoming test point to be evaluated
k	Number of nearest neighbors used in distance computation
$e_j(i)$	Manhattan distance from x_j to its i-th nearest neighbor in S_1
d_j	Sum of k-NN Manhattan distances for point x_j
d_t	Sum of k-NN Manhattan distances for incoming point x_t with respect to S_1
α	Significance level used in GEM-based outlier testing
$\mathbb{I}\{\cdot\}$	Indicator function
$\hat{p}_t$	Empirical probability estimate for incoming point x_t
s_t	Log-score associated with $\hat{p}_t$
λ	Decay factor in the anomaly accumulation statistic
S_1'	Sliding/adaptive reference set built from recent normal samples
d_t'	Sum of k-NN Manhattan distances for x_t with respect to S_1'
y	Instantaneous drift measure, defined from $\lvert d_t - d_t' \rvert$
z	Accumulated drift statistic
H	Drift detection threshold
P	Probability set associated with the reference distribution
P'	Probability set associated with the sliding distribution
p_t'	Probability estimate of x_t with respect to the sliding distribution
w	Window size for CKL drift detection
$D_{KL}(P \parallel P')$	Kullback–Leibler divergence between P and P'
θ	Tolerance value used in CKL drift accumulation
N_w	Size of the reference/sliding window in CKL analysis
B	Number of histogram bins used in KL divergence estimation
$O(\cdot)$	Asymptotic complexity notation

5.2 Problem Formulation

Consider a multivariate LMP data stream with seasonal non-stationarity. At each discrete time $t \in \{0, 1, \ldots\}$, a new data point $x_t \in \mathbb{R}^p$ is acquired, where $p \geq 1$ represents the dimensionality of the original data space. In our context, p corresponds to the number of different LMP nodes. Each dimension in the data captures the LMP values at a particular node, making the dataset multidimensional. Unlike previous studies, which assume the data stream to be independent and identically distributed (iid), we drop these assumptions allowing for more flexibility to handle *short term* abrupt anomalies and *long term* changes in underlying data distribution. To address these challenges, Change Point Detection (CPD) a.k.a. quickest or sequential change detection, is suited to the problem designed to detect small deviations from baseline data with minimal false positive rates. CPD is widely used for unsupervised anomaly detection in various fields, e.g., industrial monitoring [101–103]. CPD assesses whether incoming data points deviate significantly from expected patterns by comparing them against predefined estimated distributions.

Let a short-term anomaly occur in the observed data, such as a stealthy LMP manipulation, sudden system malfunction, or unexpected event. This anomaly manifests at unknown time t, referred to as the change-point, and causes an immediate and distinct drift in the data. Before time t_1, referred to as the anomaly/drift change-point, the system operates under normal conditions, characterized by stable statistical properties. However, at change-point t_1, these statistical properties are disrupted, and the system's behavior deviates from the expected pattern. Let the density functions (pdfs) of x_t under nominal (pre-change) and anomalous (post-change) conditions be f_{x_0} and f_{x_1}, respectively modeling the data stream as:

$$x_t \sim \begin{cases} f_{x_0}, & t < t_1, \\ f_{x_1}, & t \geq t_1, \end{cases} \tag{5.1}$$

$f_{x_1} \neq f_{x_0}$ indicates a distributional shift due to anomaly.

Objective 1: Our goal is to detect the change-point t_1 as quickly as possible controlling the rate of false alarms. The framework of CPD is well-suited for this task, focusing on minimizing the worst-case detection delay subject to a constraint on the false alarm rate. A well-known approach is the minimax problem [97]. The objective is to minimize the worst-case detection delay while maintaining a bound on the average false alarm period. Let T denote the stopping time at which a change is declared, and let $\mathbb{E}_t$ denote the expectation if the change occurs at time t. The worst-case average detection delay is defined as:

$$J(T) = \sup_t \operatorname{ess\,sup}_{\mathcal{F}_t} \mathbb{E}_t[(T - t)^+ \mid \mathcal{F}_t], \tag{5.2}$$

where $(\cdot)^+ = \max(0, \cdot)$, $\mathcal{F}_t$ is the history of observations up to time t, and ess sup denotes the essential supremum. The minimax problem is formulated as:

$$\inf_T J(T) \quad \text{subject to} \quad \mathbb{E}_\infty[T] \geq \beta, \tag{5.3}$$

where $\mathbb{E}_\infty[T]$ is the average false alarm period (the expected stopping time when no change occurs), and β is a predefined lower bound on the average false alarm period. If f_{x_0} and f_{x_1} are known, the optimal solution to the minimax problem is the CuSum algorithm [100]. CuSum uses the Log-Likelihood Ratio (LLR) as the test statistic to accumulate evidence of a change. The LLR at time t is:

$$\ell_t = \log \frac{f_{x_1}(x_t)}{f_{x_0}(x_t)}. \tag{5.4}$$

The CuSum decision statistic g_t at t is updated recursively as:

$$g_t = \max(0, g_{t-1} + \ell_t), \quad g_0 = 0. \tag{5.5}$$

A change is declared when the decision statistic exceeds a predefined threshold h, i.e.:

$$T = \inf\{t : g_t \geq h\}. \tag{5.6}$$

In practice, it is often difficult to model or estimate the high-dimensional f_{x_0} and f_{x_1}. When f_{x_1} is unknown, generalized CuSum can be applied. If f_{x_0} is known and has a parametric form, slight deviations from the parameters of f_{x_0} can be detected using a Generalized Likelihood Ratio (GLR). In high-dimensional settings, modeling the full multivariate pdf f_{x_0} is challenging. To address this, we propose extracting informative univariate summary statistics from the high-dimensional data stream and performing the anomaly detection in a lower-dimensional space. This reduces computational complexity and enhances robustness while still capturing key changes in the underlying process.

Beyond abrupt anomalies, real-world data streams exhibit gradual shifts in their statistical properties, known as *concept drift*. Drifts arise from evolving external conditions, such as seasonal demand fluctuations or long-term market shifts. Unlike abrupt anomalies, concept drifts do not occur instantaneously but manifest as a smooth transition over time. Unlike conventional methods, which assume a fixed baseline distribution, concept drifts require adaptive detection mechanisms to ensure robustness in non-stationary environments. The major challenge is distinguishing abrupt anomalies from gradual drift, as conventional change-point detection techniques may misclassify drift as a sequence of anomalies, leading to increased false alarms.

Objective 2: The goal is to extend the anomaly detection model with adaptive capabilities to account for changes in the baseline distribution due to concept drifts. This adaptation aims to preserve model's robustness, ensuring reliable anomaly detection even as the data distribution evolves over time.

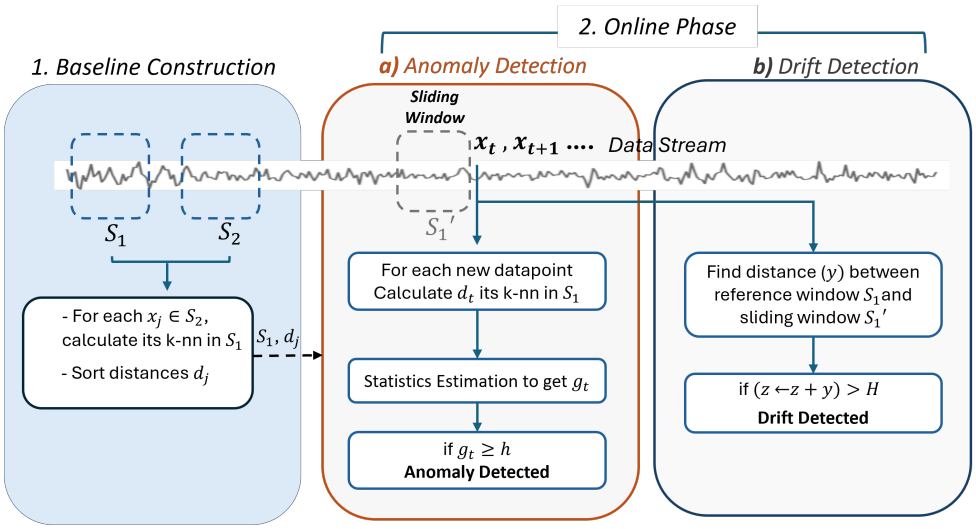


Figure 5.1: Overview of the proposed drift-aware unsupervised anomaly detection framework.

5.3 Methodology

5.3.1 Overview

To achieve objectives 1 and 2, we introduce a robust framework based on CPD that efficiently: (i) identify LMP anomalies and (ii) adaptively adjust to seasonal variations in normal LMP behavior. Our framework is based on the following (see Fig. 5.1): (1) The distributions before and after a change point, as well as the change point, are unknown. (2) Normal LMP values are non-iid, exhibiting temporal dependencies and seasonal patterns. The framework integrates anomaly detection and drift detection modules, each designed to address distinct challenges in streaming LMP data. The anomaly detection module targets sudden and abrupt deviations in short-term LMP sequences, while the drift detection module identifies gradual distributional shifts over time. The process begins by constructing a baseline distribution representing normal LMP behavior. For each incoming data point, we assess its deviation from this baseline using the Geometric Entropy Minimization-based Anomaly Detection (GEMAD) model. If the data point is significantly distant, it is classified as an LMP anomaly, triggering an alert. To ensure adaptability, the drift detection module operates over sliding windows of fixed length. Two statistical measures are employed to quantify divergence between the current window and the baseline window: (i) Cumulative Absolute Difference (CAD) and (ii) Cumulative KL Divergence (CKL). If the cumulative deviation surpasses a predefined threshold, a distributional drift is detected. This *triggers* a re-training of the GEMAD model to reconstruct an updated baseline, which is subsequently used for both anomaly detection and drift detection.

5.3.2 Constructing the Baseline Distribution

The objective of this phase is to establish a robust baseline distribution that represents the normal behavior of LMPs. The process begins by taking as input a *nominal dataset* X, which consists of historical LMP values known to be free from anomalies. This is feasible in practice, as the system generates data points at each sampling interval, enabling the collection of nominal data during normal operation [92]. The dataset X is randomly partitioned into two subsets: S_1 and S_2, with sizes N_1 and $N_2 = N - N_1$, respectively. For each data point $x_j \in S_2$, we compute the sum of the Manhattan distances to its k-nearest neighbors (k-NNs) within the subset S_1. The Manhattan distance, defined as the sum of the absolute differences across each dimension, is chosen over the Euclidean distance due to its robustness against high-dimensional noise and ability to emphasize deviations in specific LMP features. The sum of distances for each point $x_j \in S_2$ is:

$$d_j = \sum_{i=1}^{k} e_j(i), \tag{5.7}$$

where $e_j(i)$ denotes the Manhattan distance between x_j and its i-th nearest neighbor within S_1. These aggregated distances $\{d_j\}$ are then sorted in ascending order. The acceptance region is determined by selecting the smallest $(1 - \alpha)$ fraction of these distances, where α is the chosen significance level. For a new data point x_{test}, the sum of its distances to its k-NNs in S_1, denoted as d_t, is compared against this threshold, i.e., if:

$$\frac{\sum_{x_j \in S_2} \mathbb{I}\{d_t \geq d_j\}}{N_2} > 1 - \alpha, \tag{5.8}$$

then x_{test} is considered as an 'outlier' at significance level α. $\mathbb{I}\{\cdot\}$ is the indicator function that evaluates to 1 if the condition is true and 0 otherwise.

5.3.3 GEMAD: LMP Anomaly Detection Model

Our GEMAD model departs from [92] by introducing the Manhattan distance instead of Euclidean one, enhancing sensitivity to temporal LMP patterns [104]. A decay memory mechanism [105] further refines detection by controlling false discovery rates, ensuring responsiveness to persistent anomalies. It also helps prevent prolonged anomaly indications after an attack has ended, enabling faster recovery to normal system behavior. GEMAD continuously monitors the data stream, estimates statistics, and detects anomalies when monitored values exceed predefined thresholds. The detection relies on **empirical probability estimation** of a new data point x_l, based on its Manhattan distance to k-NNs in the reference set S_1:

$$\hat{p}_t = \frac{1}{N_2} \sum_{x_j \in S_2} \mathbb{I}\{d_j > d_t\}, \tag{5.9}$$

Algorithm 3 GEMAD: Anomaly Detection Model

Input: S_1, d_j and parameters: k, α, λ, h.
Initialization: the reference set $S'_1 \leftarrow S_1$.
for each new point x_t: **do**
    Get the k-NNs of x_t from S_1 and compute d_t.
    $\hat{p}_t = \frac{1}{N_2} \sum_{x_j \in S_2} \mathbb{I}\{d_j > d_t\}$, $\hat{s}_t = \log\left(\frac{\alpha}{\hat{p}_t}\right)$
    $g_t \leftarrow \max\{0, \lambda.g_{t-1} + \hat{s}_t\}$
    if $g_t \geq h$ **then**
        Alarm
    else
        Label x_t as $normal$, $S'_1(t) \leftarrow S'_1(t-1) \cup \{x_t\}$
    end if
end for

where S_2 represents baseline data (of N_2 size). A low $\hat{p}_t$ suggests an anomaly. Therefore, a deviation is quantified using the log-likelihood ratio score:

$$s_t = \log\left(\frac{\alpha}{\hat{p}_t}\right). \tag{5.10}$$

The logarithmic transformation amplifies significant anomalies while suppressing minor fluctuations. If $\hat{p}_t$ is much lower than α, s_t increases, prompting further analysis. To track persistent anomalies, the detection score g_t incorporates historical data using a decay memory mechanism:

$$g_t = \max\{0, \lambda \cdot g_{t-1} + s_t\}, \tag{5.11}$$

where λ is the decay factor (with $0 \leq \lambda \leq 1$) that controls the weight of past deviations. This adaptive approach enhances anomaly detection by filtering temporary fluctuations while capturing significant statistical shifts. If g_t exceeds a threshold h, an *anomaly alert* is triggered. The step-by-step anomaly detection mechanism implemented in GEMAD is detailed in Algorithm 3,

5.3.4 CAD Drift Detection Model

In addition to detecting point-wise anomalies, our mechanism addresses concept drift by leveraging statistical patterns over extended time periods. It involves monitoring changes in data distribution by evaluating distances between two data windows: S_1 and S'_1. S_1 represents a fixed reference set defined above, while S'_1 is a sliding window over the incoming data stream. For each incoming data point, we compute two distances: d_t from S_1 and d'_t from S'_1. The difference between the two distributions is quantified as: $y = |d_t - d'_t|$. To capture the cumulative effect of distributional changes, we maintain a *cumulative drift metric*, z, which is incrementally updated

Algorithm 4 CAD Drift Detection

Input: Threshold H.
Initialization: $z \leftarrow 0$, S'_1 from GEMAD (Algorithm 3).
for each new point x_t: **do**
    Call GEMAD to label x_t and obtain d_t.
    if x_t is $normal$: **then**
        Get the k-NNs of x_t from S'_1 and compute d'_t.
        $y = |d_t - d'_t|$, $z = z + y$
        if $z \geq H$ **then**
            Drift detected; reset z and retrain model
        end if
    end if
end for

Algorithm 5 CKL Drift Detection

Input: Threshold H, tolerance θ, window size w
Initialization: $z \leftarrow 0$, S'_1 from GEMAD (Algorithm 3).
for each new point x_t: **do**
    Call GEMAD to label x_t and obtain d_t.
    if x_t is $normal$: **then**
        Append p_t to P and get the k-NNs of x_t from S'_1.
        Compute d'_t and p'_t in (5.7), (5.9), append p'_t to P'
        if len$(P') = w$ **then**
            $D_{KL}(P\|P')$ in (5.12), $z \leftarrow z + |D_{KL} - \theta|$
            if $z \geq H$ **then**
                Drift detected; reset y and retrain model
            end if
        end if
    end if
end for

$z \leftarrow z + y$. Once z exceeds a predefined threshold H, a concept drift is detected indicating a significant shift in underlying data. Upon detection, the baseline set S_1 is updated with the most recent window S'_1, thus, the mechanism adapts to evolving data. Algorithm 4 describes the CAD drift detection process.

5.3.5 CKL Drift Detection Model

Our second model, CKL, leverages Kullback-Leibler (KL) divergence to quantify distributional shifts in probability estimates over time. Unlike CAD, which monitors the distance difference w.r.t. a single point, CKL considers the entire distribution of probability estimates thus being more robust against noisy fluctuations and gradual changes in data. CKL compares the distributions of two data windows: a reference distribution S_2 and a sliding distribution S'_1, as described earlier. As detailed in Algorithm 5, for each incoming point x_t, GEMAD computes a probability

estimate p_t reflecting the likelihood of x_t belonging to the same distribution as the reference set S_2. If x_t is considered 'normal', its associated probability is stored in probability set P. Simultaneously, a secondary probability estimate p'_t is computed to evaluate the distance of x_t, determining its relative positioning within the sliding distribution. The computed p'_t is then stored in probability set P', representing the probability estimated from the most recent points in S'_1. Once the number of probability estimates reaches the window size w, CKL computes the KL divergence between distributions P and P' quantifying the probability distribution deviation from the reference set:

$$D_{KL}(P \parallel P') = \sum_i P(i) \log\left(\frac{P(i)}{P'(i)}\right). \tag{5.12}$$

To reduce sensitivity to variations and false alarms, we introduce an accumulated drift z aggregating absolute KL deviations from a tolerance value θ over time. If z exceeds a critical threshold H, a concept drift event is detected, thus, model retraining is triggered.

5.4 Complexity Analysis

5.4.1 Time Complexity

For the GEMAD procedure, the dominant cost per incoming sample x_t arises from the k-nearest neighbor (k-NN) query over the reference set S_1 of size N_1 and the aggregation over the set $\mathcal{S}_2$ of size N_2. The k-NN query over S_1 requires $O(N_1)$ time per sample. The computation of the k-NN distance is $O(k)$ with small constant k, and the estimation step over $\mathcal{S}_2$ requires $O(N_2)$ time. All other updates and threshold comparisons are constant time. Hence, the overall per-sample complexity is $O(N_1 + N_2)$.

For the CAD drift detector, each sample is processed sequentially. The primary cost stems from the k-NN retrieval over the maintained set S'_1 of size N_1, which requires $O(N_1)$ time. All additional operations, including cumulative updates and drift decision checks, are $O(1)$. Therefore, the per-sample complexity is $O(N_1)$.

For the CKL drift detector, the computational burden similarly arises from the k-NN query over the reference window of size N_w, resulting in $O(N_w)$ time. Additionally, the KL divergence computation over histograms with B bins incurs $O(B)$ time. Thus, the overall per-sample complexity is $O(N_w + B)$.

5.4.2 Space Complexity

The overall memory requirement is dominated by stored reference samples and distribution summaries. GEMAD maintains two reference sets, S_1 and $\mathcal{S}_2$, resulting in $O(N_1 + N_2)$ space complexity. The CAD detector stores the historical comparison set S'_1 of size N'_1, leading to $O(N'_1)$ space complexity. The CKL detector maintains a sliding reference window of size N_w and histogram representations with B bins, yielding $O(N_w + B)$ memory usage. In all cases, additional scalar variables and thresholds require only constant $O(1)$ space and do not affect the overall asymptotic complexity.

5.5 Experimental Evaluation

We evaluate our melanisms using synthetic and real datasets assessing our models in terms of (i) effectiveness in detecting stealthy manipulations of LMP and adapting to changes in data distribution, (ii) robustness against varying attack intensities, and (iii) efficiency in balancing the trade-off between update costs and reliability.

5.5.1 Experiments setup

Datasets

Synthetic Dataset: The open-source Matpower case NPCC 140-bus system is used to generate both normal and manipulated LMPs using the Optimal Power Flow (OPF) function. The dataset provides hourly normal LMPs for year 2020 [76]. To introduce LMP anomalies, we implement a stealthy FDIA that manipulates Transmission Line Rate (TLR) readings reported to control center following the methodology in [38]. The attack spans one week as an adopted assumption in FDIA-based parameter attacks, executed in three randomly selected weeks within an 18-week period. The resulting time series comprises 3,024 points.

CAISO Dataset: The public CAISO dataset provides LMPs at 15-minute intervals across 12 locations [106]. The dataset, free from adversarial attacks, spans a period of six months, resulting in 17,280 time steps. Its extended duration and seasonal variations make it well-suited for evaluating drift detection over long-term dynamics.

Table 5.2: Summary of dataset characteristics, including number of features (# buses/nodes), length, presence of anomalies, and stationarity test results, indicating non-stationarity with low KPSS test p-values against 0.05 null-hypothesis.

Dataset	Features	Length	Anomalies?	Stationarity
Synthetic	140	3,024	Yes	0.0182
CAISO	13	17,280	No	0.0212

Baselines

We compare the stealthy FDIAs detection performance of our approach against state-of-the-art nonparametric models. For a fair comparison, we optimize the parameters of each baseline model: **CuSum** [100]: A well-known model detecting changes by accumulating differences between a test statistic and its pre-change mean. It is implemented as a univariate test monitoring a targeted bus using the first week of data to estimate baseline statistics with decision threshold $h = 5$. **ODIT** [107]: A sequential, nonparametric anomaly detection method based on GEM with parameters $k = 2$, $h = 5$, window size 168. **GEM** [92]: The most relevant baseline estimating statistics offline used for real-time monitoring with parameters $k = 2$, $h = 5$, window size 336.

Parameter Settings

Algorithm 3 is configured with a one-week window (168 time steps) for offline training using k-NN with $k = 2$ for distance computation. For online anomaly detection, the significance level is $\alpha = 0.05$, a decay memory factor of 0.98 ensures gradual forgetting of past observations and anomaly detection threshold $h = 5$. Algorithms 3 and 4 use a window size of 168, with anomaly detection thresholds $y \geq 3.2$ and $y \geq 5$, respectively. The number of model updates and retraining is tracked throughout the evaluation to assess adaptation efficiency.

Evaluation Metrics

The performance of all the methods is evaluated using standard metrics, including Detection Rate (DR), precision, F1-score, Area Under the Curve (AUC), and False Alarm Rate (FAR). The DR (a.k.a. recall) is defined as Detection Rate $= \frac{TP}{TP+FN}$, where TP and FN represent true positives and false negatives, respectively. The precision, which measures the proportion of detected anomalies that are actual anomalies, is given by Precision $= \frac{TP}{TP+FP}$, where FP denotes false positives. The F1-score is computed as $F1 = 2 \times \frac{\text{Precision} \times \text{DR}}{\text{Precision+DR}}$. The AUC evaluates the model's ability to distinguish between normal and anomalous instances. Finally, the FAR is given by FAR $= \frac{FP}{FP+TN}$, where TN represents true negatives.

5.5.2 Detection Performance Comparison

We compare the detection performance of our mechanisms against CuSum, ODIT, and GEM in terms of DR, Precision, F1-score, AUC, and FAR before and after adaptation as shown in Table 5.3 and Fig. 5.2. Our models (CAD, CKL) consistently achieve superior performance across all metrics, highlighting their robustness in detecting stealthy LMP manipulation attacks. Compared to CuSum and ODIT, our models exhibit significantly higher detection accuracy while maintaining a lower false alarm rate. CAD achieves F1-score 0.9482, and CKL reaches 0.9629, significantly outperforming CuSum (0.0976) and GEM (0.3480). Similarly, the AUC

for CAD (0.9796) and CKL (0.9516) indicate a strong ability to distinguish subtle and stealthy anomalies in LMP streams. Although GEM provides a high detection rate (0.91), it suffers from an elevated false alarm rate. This shortcoming arises from the long-term dependencies of cumulative monitoring statistics, which cause the model to take a prolonged time to adapt and return to a normal state after an attack. Our models overcome this limitation by employing a decay memory mechanism that dynamically adjusts to evolving data distributions. By gradually reducing the influence of older observations, our models prevent prolonged anomaly indications once an attack has ended, ensuring faster recovery to normal system behavior. Moreover, CAD and CKL consistently demonstrate superior detection capabilities by incorporating drift awareness into the anomaly detection process. After adaptation, CAD achieves detection rate 0.9821 with F1-score 0.9482, while CKL reaches 0.9519 with F1-score 0.9629; both significantly outperforming the baselines. Despite the presence of concept drift, the FAR for CAD and CKL remain considerably lower (0.2668 and 0.1208) compared to all baselines, ensuring a more reliable and stable detection framework. Table 5.3 compares model performance before and after a distributional change in the LMP data. Here, "before" refers to the period evaluated using the original baseline learned from the initial nominal data, while "after" refers to the period following the change, where non-adaptive models continue to rely on the previous baseline and adaptive models update their reference distribution after drift detection.

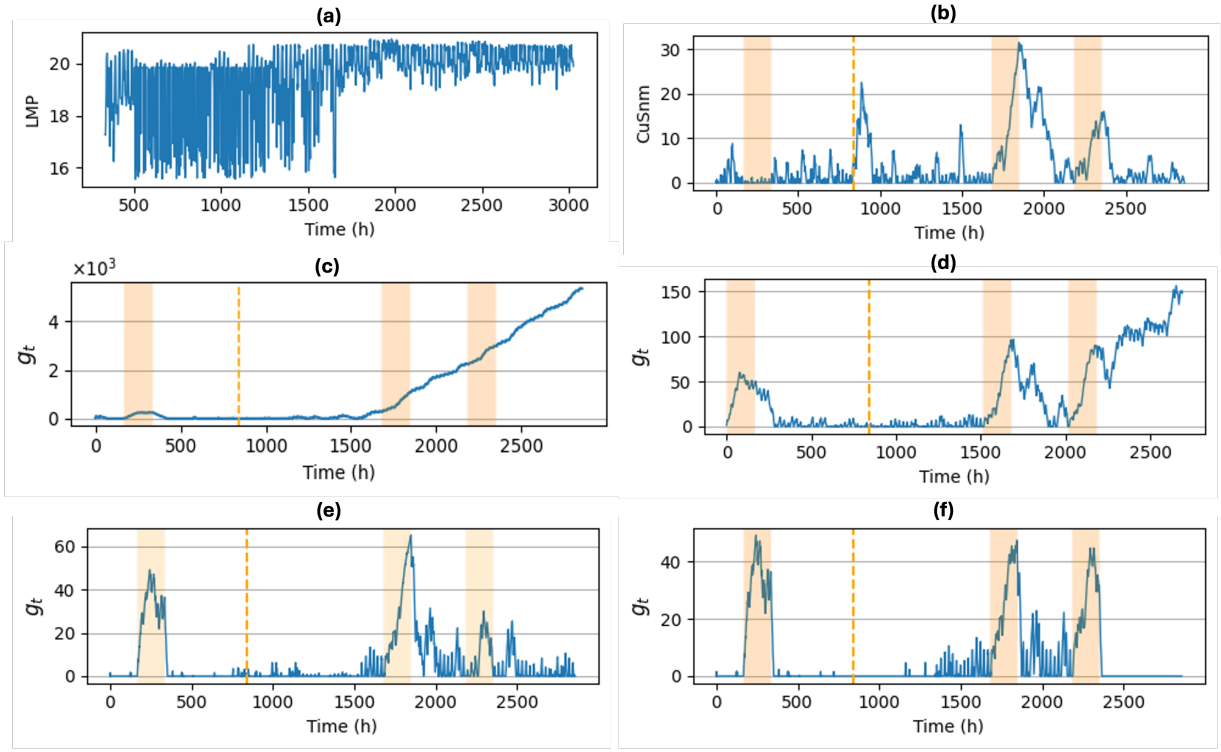


Figure 5.2: Visual representation of all methods detecting manipulated LMPs. (a) Original time-series; (b) CuSum; (c) ODIT; (d) GEM; (e) CAD; (f) CKL. Highlighted orange regions indicate attack duration.

Table 5.3: Performance Metrics Before and After Change for Different Models

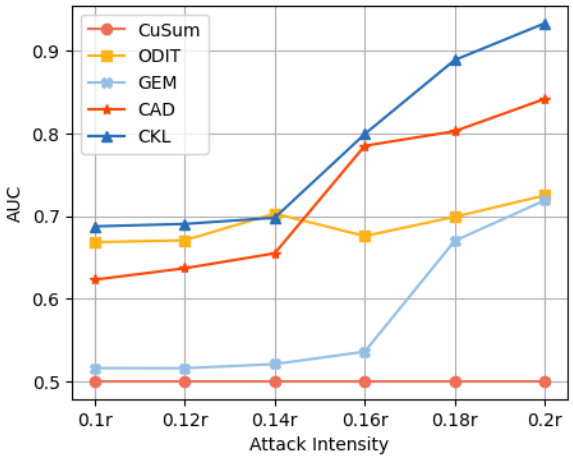
Model	DR	Precision	F1-score	AUC	FAR
CuSum (Before)	0.1190	0.0826	0.0976	0.3388	0.4413
CuSum (After)	0.9642	0.2250	0.3648	0.6503	0.6634
ODIT (Before)	0.9880	0.3023	0.4630	0.6133	0.7614
ODIT (After)	1.0000	0.2250	0.3096	0.5546	0.8906
GEM (Before)	0.9107	0.5751	0.7050	0.7872	0.3363
GEM (After)	0.8244	0.2591	0.3943	0.6504	0.5234
CAD (Before)	0.9821	0.9166	0.9482	0.9776	0.0269
CAD (After)	0.9136	0.4342	0.5886	0.8401	0.2334
CKL (Before)	0.9821	0.9166	0.9482	0.9776	0.0269
CKL (After)	0.9519	0.5955	0.7422	0.9156	0.1208

Table 5.4: LMP values under different attack intensities

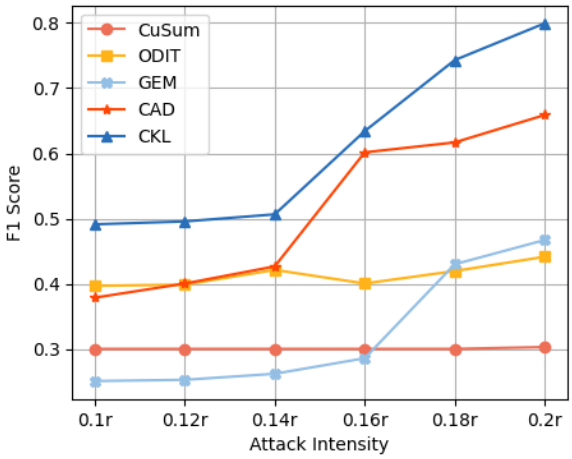
Attack intensity	No FDIA	$0.1r$	$0.12r$	$0.14r$	$0.16r$	$0.18r$	$0.2r$
LMP ($/MWh)	15.6	15.61	15.64	15.7	15.77	16.03	16.2
Total illegal profit in a week ($/MWh)	0	59.83	81.08	121.22	160.8	227.7	269.47

5.5.3 Impact of Diverse Attack Intensity

We compare the robustness of all models against varying attack intensities. Attack intensity refers to the magnitude of manipulation applied to the system. It represents the proportion of the original measurement that has been altered by the attacker. The notation “r" denotes the standard deviation of the measurement noise, and attack intensities like $0.1r, 0.12r, \ldots, 0.2r$ indicate that the injected false data deviates from the true value by 10%, 12%, ..., 20% of this noise level. As attack intensity increases, the injected false data becomes more significant, making the attack more impact on system operations, such as LMP changes and total illegal profits (as shown in Table 5.4). However, higher attack intensities do not necessarily imply easier detection. For instance, while $0.2r$ represents the strongest attack in terms of deviation, it remains stealthy because its effect on LMP series is subtle and visually indistinguishable from normal fluctuations; see Fig.5.2(a). This makes it more challenging to detect through simple observation, despite its significant impact on financial gains for the attacker. Conversely,



(a) AUC of baseline models under different attack intensities.



(b) F1 score of baseline models under different attack intensities.

Figure 5.3: Performance comparison under varying attack intensities using AUC and F1 score metrics.

lower attack intensities, e.g.,$0.1r, 0.12r$, introduce minimal changes making them even harder to detect while still being capable of generating illegal profits. Fig. 5.3 shows the superiority of both models in distinguishing attack intensities of $0.16r$ and above indicating that GEMAD is responsive to stronger attacks.

5.5.4 Detection Efficiency Comparison

To compare the efficiency of anomaly detection with drift awareness, we introduce two scenarios: (1) without drift detection and (2) with fixed-interval updates lacking adaptive mechanisms. The experiment uses synthetic and real datasets assuming no anomalies in the real dataset (normal data labeled as '0'), making any detected anomaly a false alarm. Efficiency is measured by FAR and update frequency. The results in Table 5.5 show that our adaptive drift detection CKL and CAD effectively balance the trade-off between update costs and reliability. While the 'No drift detection' approach results in the highest FAR due to outdated models, the 'Fixed update' method significantly reduces FAR but incurs high update costs. In contrast, CKL and CAD achieve an optimal balance maintaining low FAR while minimizing unnecessary updates. CKL demonstrates the best efficiency with the lowest FAR (0.0059 on CAISO, 0.0947 on the synthetic dataset) and the fewest updates, suggesting that it provides a cost-effective and reliable solution for handling concept drift.

Table 5.5: Detection efficiency on CAISO & synthetic data

Method	**CAISO dataset**		**Synthetic dataset**	
	FAR	**Updates**	**FAR**	**Updates**
No drift detection	0.3306	0	0.3123	0
Fixed update (weekly)	0.0404	24	0.0841	18
CAD Model	0.0188	12	0.1483	6
CKL Model	0.0059	11	0.0947	3

5.6 Conclusions

This chapter introduced a drift-aware, market-level framework for detecting stealthy FDIAs in non-stationary Locational Marginal Price (LMP) time series. By integrating online anomaly detection with window-based concept drift monitoring, the proposed approach distinguishes gradual regime shifts from abrupt adversarial manipulation. Experimental results demonstrate improved robustness under evolving market conditions, maintaining stable detection performance while reducing false alarms compared to non-adaptive baselines.

Despite these advances, several limitations remain. First, although the framework operates in a multivariate setting, it does not explicitly model the physical or spatial dependencies among buses in the transmission network, nor does it directly provide localization of manipulated regions within the grid. Moreover, monitoring high-dimensional LMP streams across many buses increases computational cost and may limit scalability in large-scale markets. These limitations motivate the next chapter, which leverages graph-based representations of LMP signals to incorporate spatial dependencies, enable dimensionality reduction, and support anomaly localization under scalable monitoring constraints.

Chapter 6

Physics-Informed Node Selection and Monitoring of Graph-Smooth Signals

6.1 Introduction

Motivated by the limitations identified in the previous chapter, this chapter addresses a remaining practical challenge in electricity market monitoring: scalability and spatial interpretability. While multivariate drift-aware models improve robustness under evolving market conditions, monitoring high-dimensional Locational Marginal Price (LMP) streams across hundreds or thousands of network nodes at short settlement intervals imposes significant computational and statistical burdens. Moreover, anomaly detection in purely feature-space representations provides limited support for localization, making it difficult to attribute detected anomalies to specific physical regions of the transmission network.

Real-time markets operate on short settlement intervals (typically five minutes), with LMPs computed at large numbers of buses. Hereinafter, the terms "nodes" and "buses" are used interchangeably according to context [45, 108]. Monitoring systems must therefore process high-dimensional LMP vectors under strict time constraints. High dimensionality not only increases computational cost but also amplifies noise accumulation and spurious correlations, masking meaningful deviations and reducing detection reliability. These challenges are further compounded by concept drift, which naturally arises from evolving demand patterns, renewable integration, and structural changes in the grid [109, 110]. Continuous full-scale retraining under such conditions is impractical and may introduce instability.

This chapter addresses three tightly coupled research questions: (i) how to monitor a reduced subset of network nodes while preserving detection performance comparable to full-network observability; (ii) how to enable anomaly localization such that detected abnormalities can be traced to their physical origin within the transmission network; and (iii) how to maintain monitoring stability under concept drift without frequent full-scale model retraining.

Related Work. Existing approaches to electricity market monitoring fall into two broad categories. Model-based methods rely on detailed physical system representations and detect deviations from expected system states [58,63,64]. While effective under well-specified assumptions, these approaches require comprehensive system information and are sensitive to modeling inaccuracies [66]. Data-driven methods treat market variables as multivariate time series and focus on detecting abnormal temporal variations [49, 109]. Recent extensions incorporate spatial information through deep learning architectures that learn statistical co-movements among nodes [67]. However, such models capture correlation structure rather than physically grounded dependencies induced by transmission topology, limiting interpretability and stability under non-stationary conditions.

Graph-based methods provide a promising alternative because they explicitly represent dependencies among system components. In particular, Graph Neural Networks (GNNs) have become a dominant learning-based paradigm for graph-structured anomaly detection and have been increasingly applied to FDIA detection and localization in smart grids [111–113]. In parallel, unsupervised and self-supervised graph anomaly detection methods, including graph autoencoders, graph attention models, and contrastive graph-learning frameworks, have been proposed to reduce dependence on labelled attack data [114–117]. These approaches demonstrate the value of graph-aware learning; however, they are often designed for physical measurement or state-estimation monitoring, where the objective is to classify, reconstruct, or localize abnormal grid states. LMP manipulation differs because attacks may remain physically plausible while inducing economically meaningful distortions in market prices. Moreover, high-capacity graph-learning models typically require substantial training data, careful calibration, and repeated adaptation under regime changes, which is challenging in electricity markets where labelled manipulation events are scarce, market conditions are non-stationary, and monitoring must operate within short settlement intervals.

Graph Signal Processing (GSP) offers a more suitable foundation for this setting because it exploits graph structure without requiring deep representation learning. In power systems, spectral and algebraic graph theory have been applied to network partitioning, sensor placement, and fault localization [118–120]. GSP formulations further model physical measurements, such as voltages or state variables, as signals defined over network graphs, enabling detection through spectral smoothness, graph filtering, and localized signal deviations [121–123]. Nevertheless, existing GSP-based approaches are primarily designed for physical-layer monitoring and commonly rely on purely topological or admittance-based graphs. They are therefore not directly tailored to electricity market signals, where LMPs reflect both network constraints and economic congestion patterns. This motivates a market-oriented GSP framework that retains the interpretability and scalability of graph-based monitoring while addressing the limited-data, unsupervised, lightweight, and non-stationary nature of LMP manipulation detection.

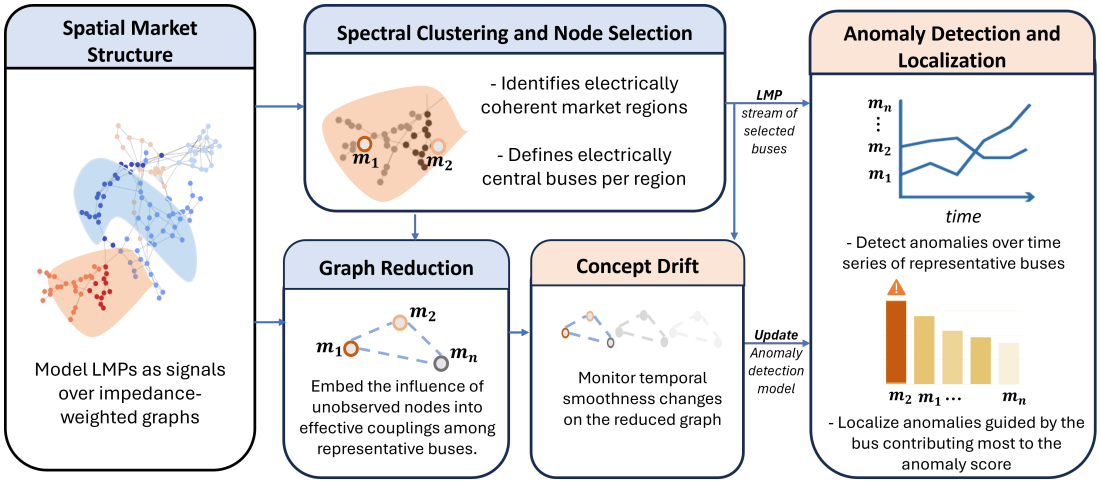


Figure 6.1: Conceptual overview of the proposed GSP-based electricity market monitoring framework. Blue blocks denote offline processing stages, while orange blocks correspond to online monitoring and detection tasks.

Accordingly, the proposed method is positioned as a lightweight alternative to GNN-based graph learning, rather than as an approximation of a GNN. Instead of learning high-capacity node embeddings through data-driven message passing, it exploits the known transmission-network structure directly through graph signal smoothness, spectral clustering, and representative-node monitoring. This design is better suited to LMP manipulation detection, where labelled attack data are limited, market regimes are non-stationary, and online monitoring must remain scalable and interpretable.

Contribution. This chapter introduces a graph signal processing–based framework for electricity market monitoring grounded in a market-oriented graph representation. The proposed graph captures economically meaningful spatial dependencies through impedance-weighted coupling, enabling LMP time series to be modeled as graph signals exhibiting piecewise-smooth spatial structure. The framework leverages spectral properties of this representation to identify coherent market regions and select representative buses that preserve dominant spatial dynamics. Anomaly detection is performed on representative nodes, enabling regional localization while reducing dimensionality. Furthermore, temporal evolution of graph signal smoothness under partial observability is exploited to detect concept drift with reduced sensitivity to localized anomalies. The overall architecture of the proposed graph-based monitoring framework is illustrated in Fig. 6.1. The diagram highlights the separation between offline graph construction and node-selection components and online monitoring components, including anomaly detection and drift monitoring.

The main contributions are summarized as follows:

1. We demonstrate that LMP time series can be modeled as graph signals defined over an impedance-weighted transmission network and exhibit piecewise-smooth spatial structure.

2. We employ spectral clustering to identify electrically coherent market regions and select representative buses that preserve dominant spatial dynamics.

3. We extend anomaly detection to operate on representative nodes, enabling spatial attribution of detected anomalies.

4. We introduce a reduced-order graph signal–based approach for concept drift detection under partial observability.

5. We provide comprehensive experimental evaluation under realistic market stress scenarios, demonstrating improved scalability and stability.

The remainder of this chapter is organized as follows. Section 6.2 introduces the graph-based market representation. Sections 6.3.1, 6.3.2, and 6.3.3 detail node selection, anomaly detection and localization, and drift monitoring mechanisms, respectively. Section 6.5 presents experimental results. Section 6.6 concludes the chapter.

Table 6.1: Notation used in Chapter 6

Symbol	**Description**
$\mathcal{G} = (\mathcal{V}, \mathcal{E})$	Undirected weighted graph representing the power network.
$\mathcal{V}$	Set of nodes (buses).
$\mathcal{E}$	Set of edges (transmission lines).
e_{ij}	Edge connecting nodes i and j.
z_{ij}	Complex impedance of the transmission line between nodes i and j.
w_{ij}	Edge weight between nodes i and j, defined by inverse impedance.
$W = [w_{ij}]$	Weighted adjacency matrix of the graph.
D	Diagonal degree matrix of the graph.
d_{ii}	Degree of node i, equal to $\sum_j w_{ij}$.
L	Graph Laplacian matrix.
L_{norm}	Normalized graph Laplacian.
I	Identity matrix.
$\mathbf{x}$	Graph signal of LMP values over all nodes.
x_i	LMP value at node i.
N	Total number of nodes in the network.
$\mathbf{x}^\top L\mathbf{x}$	Graph-signal smoothness (Dirichlet energy).
$u_1, \ldots, u_r$	Eigenvectors associated with the r smallest eigenvalues of L.
U_r	Spectral embedding matrix formed by the first r eigenvectors.

Continued on next page

Table 6.1 continued from previous page

Symbol	Description
r	Embedding dimension in spectral clustering.
K	Number of clusters in spectral clustering.
$C_1, \ldots, C_K$	Clusters (electrically coherent zones).
C_c	The c-th cluster.
n_c	Number of nodes in cluster C_c.
m_c	Medoid (representative node) of cluster C_c.
$d(i, j)$	Generic distance between nodes i and j.
$d_{\text{elec}}(i, j)$	Impedance-weighted electrical shortest-path distance.
$\pi : i \to j$	Path from node i to node j.
$\text{score}(i)$	Distance-based centrality score for representative selection.
ϵ	Tolerance for selecting near-optimal representative nodes.
M	Set of selected representative nodes.
N_m	Number of representative nodes, i.e., $\lvert M \rvert$.
T	Number of time steps in the monitoring horizon.
$X \in \mathbb{R}^{N_m \times T}$	Reduced multivariate LMP time series over representatives.
$\mathbf{x}_t$	LMP vector over representative nodes at time t.
$x_{i,t}$	LMP at representative node i and time t.
$\mathcal{S}_1, \mathcal{S}_2$	Nominal reference subsets used in GEMAD+ baseline construction.
N_1, N_2	Sizes of $\mathcal{S}_1$ and $\mathcal{S}_2$, respectively.
k	Number of nearest neighbors.
d_t	Average Manhattan-distance deviation of $\mathbf{x}_t$ from nominal neighbors.
$\mathbf{x}_{(j)}$	j-th nearest neighbor of $\mathbf{x}_t$ in the reference set.
$\alpha_{i,t}$	Per-node deviation contribution at node i and time t.
$\hat{p}_t$	Empirical tail probability estimate for deviation d_t.
α	Significance level in anomaly detection.
s_t	Instantaneous anomaly score at time t.
g_t	Cumulative anomaly statistic in GEMAD+.
λ	Forgetting/decay factor in GEMAD+.
h	Anomaly detection threshold.
$\eta_{i,t}$	Normalized contribution of node i to the total anomaly at time t.
R	Set of eliminated nodes, $R = \mathcal{V} \setminus M$.
$L_{MM}, L_{MR}, L_{RM}, L_{RR}$	Block submatrices of L after partitioning into M and R.
L^*	Kron-reduced Laplacian defined on representative nodes.
$\mathbf{x}_M$	Signal restricted to representative nodes M.

Continued on next page

Table 6.1 continued from previous page

Symbol	Description
$\mathbf{x}_R$	Signal restricted to eliminated nodes R.
$\mathbf{x}_R^\star$	Energy-minimizing extension over eliminated nodes.
$Q(\mathbf{x})$	Dirichlet quadratic form on the full graph.
y_t	Graph-signal smoothness on the reduced graph at time t.
$\mathbf{y} = [y_1, \ldots, y_T]^\top$	Time series of reduced-graph smoothness values.
$\bar{y}_t$	Moving-average smoothness statistic over a window.
w	Sliding-window length.
μ_0, σ_0	Baseline mean and standard deviation of $\bar{y}_t$.
r_t	Standardized smoothness deviation statistic.
h_d	CuSum threshold for concept-drift detection.
UF	Update frequency of the drift detector.
N_{updates}	Number of model updates triggered during monitoring.
ε	Small positive constant for numerical stability.
$\text{PM}(b^\star)$	Probability-mass score for cluster-level anomaly localization.
$b^\star$	True anomalous bus.
f_i	Frequency with which representative node i is the main anomaly contributor.
$\hat{b}$	Predicted representative node for localization.
$\text{LE}_{\text{rep}}(b^\star)$	Representative-level localization error.
E	Number of edges in the graph.
I_{km}	Number of iterations in K-means clustering.
$O(\cdot)$	Asymptotic complexity notation.

6.2 Fundamentals

6.2.1 Graph Domain for Energy Market

In this chapter, the power network is modeled as an undirected weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where each bus corresponds to a vertex and each transmission line between buses i and j defines an edge $e_{ij} \in \mathcal{E}$. The edge weight is selected using an *inverse-impedance* model, i.e.,

$$w_{ij} = \frac{1}{|z_{ij}|}, |z_{ij}| > 0, \tag{6.1}$$

which assigns stronger coupling to low-impedance lines. This provides a physically meaningful notion of electrical proximity that reflects how power flows and congestion propagate across the grid. Alternative graph constructions, such as complex-admittance operators [123] or geographical-distance graphs [122], have been used in related applications. The inverse-impedance weighting adopted here is particularly well suited for market variables such as LMPs, whose spatial variation is shaped by power-transfer sensitivities and line impedances. Moreover, defining w_{ij} as $1/|z_{ij}|$, it yields a robust representation under topology changes like line outages, since effective electrical distances remain well defined through alternative feasible paths.

Given a weight matrix $W = [w_{ij}]$, the diagonal degree matrix D has entries $d_{ii} = \sum_j w_{ij}$, representing the total connection strength of each node. The combinatorial graph Laplacian is then defined as

$$L = D - W, \tag{6.2}$$

and normalized as

$$L_{\text{norm}} = I - D^{-1/2} W D^{-1/2}. \tag{6.3}$$

The Laplacian encapsulates both the magnitude and structure of the graph weights, serving as the core operator that governs diffusion, smoothness, and spectral variation of graph signals. Given that L_{norm} is real and symmetric, its eigenvalues are real and non-negative while eigenvectors form an orthogonal graph Fourier basis. Lower eigenvalues correspond to system-wide, slowly varying components of a graph signal, whereas higher eigenvalues capture localized fluctuations around strongly coupled buses. Therefore, the impedance-weighted Laplacian provides a physically grounded spectral domain for analyzing market-driven signals whose spatial patterns are shaped by network impedance and congestion-induced coupling.

6.2.2 LMPs as Graph Signals

Locational Marginal Prices (LMPs) reflect the marginal cost of serving load under the physical and economic constraints of the transmission network. Since such constraints couple LMPs through the impedance-weighted topology, LMPs inherit a spatial organization in which rapid variations are naturally moderated across strongly connected regions of the grid. In effect, the electrical network filters out highly localized discrepancies while allowing broader zonal patterns to persist, making an LMP vector well represented as a graph signal supported on the impedance-weighted domain defined in the previous section.

Let $\mathbf{x} = [x_1, x_2, \ldots, x_N]^\top$ denote a vector of scalar LMPs. Across portions of the network with strong electrical connectivity, LMPs often vary gradually, reflecting the tight interaction among buses within these regions. In contrast, *localized congestion* produces sharp transitions at a small set of boundary edges. These observations motivate a piecewise-smooth description: the signal is smooth across most of the graph, while localized non-smooth behaviour is concentrated near congested interfaces [124]. The spatial variation of $\mathbf{x}$ is captured by the Laplacian quadratic

form, i.e.,

$$\mathbf{x}^\top L\mathbf{x} = \frac{1}{2}\sum_{(i,j)\in\mathcal{E}} w_{ij}(x_i - x_j)^2, \tag{6.4}$$

which assigns larger values when significant differences appear across edges of high electrical weight. This measure highlights coherent regions where the contribution is small, and boundary regions where congestion induces strong local gradients.

This graph-signal interpretation provides the basis for the node-selection and monitoring framework developed in the following sections, where smoothness, spectral structure, and localized deviations are used to support scalable anomaly detection.

6.3 Methodology

6.3.1 Spectral Clustering and Node Selection

Our objective is to partition the network into electrically coherent zones and to select a compact set of representative buses for reduced-order market monitoring. Since LMPs exhibit piecewise-smooth spatial structure, their variation tends to be relatively small within coherent regions and be concentrated along congestion boundaries. This motivates us to adopt spectral clustering on the impedance-weighted graph to extract zones consistent with the underlying electrical coupling.

Spectral Clustering

The spectral clustering begins by embedding the nodes into a low-dimensional space spanned by the smoothest eigenvectors of the normalized Laplacian L. These eigenvectors minimize the quadratic form $x^\top Lx$, thus, capturing the dominant, slowly varying spatial patterns of the network. Once nodes are projected into this subspace, electrically coherent buses map to nearby points, while buses separated by congested or weak interfaces appear more widely separated.

Let $u_1, \ldots, u_r$ be the eigenvectors associated with the r smallest eigenvalues of L forming the spectral embedding $U_r = [u_1\, u_2 \cdots u_r]$. Following standard normalized spectral clustering in [125], we set $r = K$, so that the embedding dimension matches the number of desired clusters. The zonal partitions $\{C_1, \ldots, C_K\}$ are then obtained by applying K-means clustering to the rows of U_r. This groups buses together according to their positions in the low-frequency subspace.

The number of clusters K is selected by examining how the LMP variation is distributed across edges within each candidate partition. Partitions in which only a small share of the variation lies inside clusters, indicate that most changes occur at their boundaries. These are consistent with the piecewise smooth structure of Section 6.2.

Representative Selection

Once the spectral clustering has produced electrically coherent zones, the next step in the dimensional-reduction pipeline is to identify a compact set of representative buses. Such buses summarize the behavior of each cluster. We adopt a medoid-based approach using impedance-weighted electrical distance as the primary metric. This is due to the fact that electrically closer buses have shown to exhibit more similar market-driven responses. This yields the electrically central bus (within each cluster) as a physically interpretable and structurally meaningful representative.

Specifically, for a cluster $C_c = \{i_1, \ldots, i_{n_c}\}$ and a distance function $d(\cdot, \cdot)$, the medoid is defined as

$$m_c = \arg\min_{i \in C_c} \sum_{j \in C_c} d(i, j). \tag{6.5}$$

This denotes that node whose aggregate distance to the remainder of the cluster is minimal. In our work, the distance used for representative selection is the impedance-weighted shortest-path length,

$$d_{\text{elec}}(i, j) = \min_{\pi: i \to j} \sum_{(u,v) \in \pi} |z_{uv}|, \tag{6.6}$$

which reflects the strength of electrical coupling of buses. We then rank buses within each cluster according to their distance-based centrality,

$$\text{score}(i) = -\sum_{j \in C_c} d_{\text{elec}}(i, j), \tag{6.7}$$

and select the top-scoring bus along with any nodes whose scores are within tolerance $\epsilon > 0$ from the maximum.

6.3.2 Adaptive Energy Market Anomaly Detection

Following the representative selection, the monitoring process operates on a *reduced set* of nodes $M \subseteq \mathcal{V}$ that preserve the spatial structure of LMP dynamics. Each representative bus is associated with a time series of LMPs $x_{i,t}$. The collection of these values forms a multivariate time series $X = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T] \in \mathbb{R}^{N_m \times T}$, where $\mathbf{x}_t = [x_{1,t}, x_{2,t}, \ldots, x_{N_m,t}]^\top$ denotes the vector of LMPs across *all* representatives at time t. This reduced representation serves as the input to the subsequent monitoring stage, where temporal and spatial dependencies among representative nodes are jointly analyzed to detect and interpret abnormal market behavior. The objective in this stage is twofold: (i) **detect** the occurrence of abnormal events in the energy market, and (ii) **localize** the buses or regions that predominantly contribute to these deviations.

Anomaly Detection

Our detection framework, coined GEMAD+, substantially extends the GEMAD [109] to the extend that it operates on graph-sampled LMP time series and anomaly localization. In our problem, each representative node corresponds to a selected bus in the constructed graph. The temporal evolution of LMP signals is jointly analyzed across these representative nodes in order to capture both spatial and temporal dependencies in the system behavior.

Baseline: To establish a reference for nominal market behavior, a baseline distribution is first constructed from historical LMP data known to be anomaly-free [92, 109]. The nominal dataset is randomly partitioned into two subsets, $\mathcal{S}_1$ and $\mathcal{S}_2$, containing N_1 and N_2 samples, respectively. The subset $\mathcal{S}_1$ is used to compute k-nearest neighbor statistics, while $\mathcal{S}_2$ serves to estimate the empirical tail probability of the deviation metric introduced later in this section. This two-sample design enables a nonparametric characterization of the normal operating regime without assuming a specific distributional form.

Real-time Monitoring: At each time step t, the current observation $\mathbf{x}_t$ is compared with a reference set of nominal samples $\mathcal{S}_1$ obtained from the historical data. The k nearest neighbors of $\mathbf{x}_t$ in $\mathcal{S}_1$ are identified using the Manhattan distance (L_1 norm), which is adopted instead of the Euclidean distance (L_2 norm) for high-dimensional settings. L_2 norm 'loses' discriminative power when the distances between similar and dissimilar points become nearly equal [104]. The average deviation in L_1 is then computed as:

$$d_t = \frac{1}{k} \sum_{j=1}^{k} \|\mathbf{x}_t - \mathbf{x}_{(j)}\|_1, \tag{6.8}$$

which quantifies the deviation of the current state from its nominal neighborhood. Due to the additive property of the L_1 norm, the global deviation d_t can be decomposed into per-node components as:

$$d_t = \sum_{i=1}^{N_m} \alpha_{i,t}, \qquad \alpha_{i,t} = \frac{1}{k} \sum_{j=1}^{k} \left| x_{i,t} - x_{i,(j)} \right|. \tag{6.9}$$

The term $\alpha_{i,t}$ represents the average absolute deviation of bus i from its nominal behavior at time t. This allows the same computational framework to jointly support anomaly detection and spatial localization.

An empirical tail probability is then estimated as

$$\hat{p}_t = \frac{1}{N_2} \sum_{\mathbf{x}_j \in \mathcal{S}_2} \mathbb{I}\{d_j > d_t\}, \tag{6.10}$$

where $\mathcal{S}_2$ denotes an auxiliary nominal reference set and $\mathbb{I}\{\cdot\}$ is the indicator function. The instantaneous deviation score is given by:

$$s_t = \log\left(\frac{\alpha}{\hat{p}_t}\right). \tag{6.11}$$

We can then define a cumulative statistic

$$g_t = \max\{0,\ \lambda g_{t-1} + s_t\}, \tag{6.12}$$

that tracks persistent departures from the nominal regime. An anomaly alarm is triggered when g_t exceeds a predefined threshold h. The parameters k, λ, α, and h are selected empirically from nominal data.

Anomaly Localization

To localize and identify the sources of detected anomalies, the per-node deviations $\alpha_{i,t}$ are normalized to obtain their relative contributions:

$$\eta_{i,t} = \frac{\alpha_{i,t}}{\sum_{l=1}^{N_m} \alpha_{l,t}}, \qquad \sum_{i=1}^{N_m} \eta_{i,t} = 1. \tag{6.13}$$

The term $\eta_{i,t}$ represents the proportion of the total deviation at time t attributed to node i, providing a spatially interpretable measure of anomaly contribution. Nodes with larger $\eta_{i,t}$ values correspond to stronger relative departures from nominal conditions at that time instant.

While $\eta_{i,t}$ quantifies the relative strength of node-level deviations instantaneously, anomaly localization in our chapter is formulated as an event-centric task. Specifically, the target is to identify nodes or regions that repeatedly dominate anomaly attribution over the duration of an abnormal event, rather than to estimate the cumulative magnitude of deviation. The absolute scale of $\eta_{i,t}$ is influenced by the underlying k-nearest neighbors-based detection mechanism and the temporal accumulation dynamics of the GEMAD+ statistic.

6.3.3 Temporal Evolution of Smoothness in Market Drift Detection

Graph-based representations provide a principled framework for tracking *concept drift* by explicitly modeling the evolving relational structure of the system. A range of graph-based indicators, including graph distance [126], graph entropy [127], and graph-signal smoothness [128], have been proposed to detect such structural changes. Among these, graph-signal smoothness is particularly effective in capturing global variations in spatial coherence while remaining comparatively insensitive to localized disturbances. This distinction is particularly important in energy-market settings, where regime-level shifts manifest as persistent changes in system-wide

coupling patterns rather than as isolated anomalies. In this work, graph-signal smoothness is used exclusively as a trigger for detecting concept drift and initiating adaptive updates of the anomaly-detection baseline, rather than as a direct indicator of localized attacks. To enable scalable monitoring under partial observability, we adopt a reduced-order graph representation that preserves the structural properties of the original network with respect to the previously selected node set M. The following subsections introduce the construction of the reduced graph, define smoothness over this representation, and describe the statistical procedure for concept-drift detection.

Reduced-Order Graph Construction from L to L^*

To obtain a reduced-order representation that is equivalent from the perspective of the selected nodes, the full-network Laplacian L is condensed with respect to M. Let $M \subseteq \mathcal{V}$ denote the previously selected set of representative buses and define $R := \mathcal{V} \setminus M$. After reordering the nodes such that those in M appear first, the Laplacian can be written as

$$L = \begin{bmatrix} L_{MM} & L_{MR} \\ L_{RM} & L_{RR} \end{bmatrix}. \tag{6.14}$$

Eliminating the interior nodes R yields an equivalent reduced graph on M, whose Laplacian is obtained via the Schur complement of L_{RR},

$$L^* = L_{MM} - L_{MR} L_{RR}^{-1} L_{RM}. \tag{6.15}$$

The reduced Laplacian $L^* \in \mathbb{R}^{|M|\times|M|}$ embeds the influence of the eliminated nodes into effective couplings among the retained nodes, thereby preserving the network interactions as observed from M [118]. To make the equivalence explicit, consider a graph signal $\mathbf{x} \in \mathbb{R}^{|\mathcal{V}|}$ partitioned conformably as $\mathbf{x} = [\mathbf{x}_M^\top, \mathbf{x}_R^\top]^\top$. The Dirichlet quadratic form on the full graph is defined as

$$Q(\mathbf{x}) \triangleq \mathbf{x}^\top L \mathbf{x}.$$

For fixed $\mathbf{x}_M$, the energy-minimizing extension over the eliminated nodes satisfies

$$L_{RR}\mathbf{x}_R^\star = -L_{RM}\mathbf{x}_M, \qquad \mathbf{x}_R^\star = -L_{RR}^{-1} L_{RM}\mathbf{x}_M, \tag{6.16}$$

and substituting $\mathbf{x}_R^\star$ into $Q(\mathbf{x})$ yields

$$\min_{\mathbf{x}_R} \mathbf{x}^\top L \mathbf{x} = \mathbf{x}_M^\top L^* \mathbf{x}_M. \tag{6.17}$$

Hence, the reduced Laplacian L^* preserves the Dirichlet energy associated with signals defined on the selected node set M and provides a structurally consistent reduced-order representation of the original network for subsequent monitoring and analysis.

Statistical Detection of Concept Drift Using L^*

Let $\mathbf{y} = [y_1, y_2, \ldots, y_T]^\top$ denote the temporal sequence of graph-signal smoothness values computed on the reduced graph. At each time t, the smoothness is defined as

$$y_t = \mathbf{x}_t^\top L^* \mathbf{x}_t, \tag{6.18}$$

where $\mathbf{x}_t \in \mathbb{R}^{|M|}$ is the vector of LMPs observed at the selected node set M, and L^* is the corresponding Kron-reduced Laplacian. The sequence $\{y_t\}$ captures the temporal evolution of spatial coherence in LMPs as encoded by the reduced network structure. Under nominal operating conditions, y_t fluctuates around a stable mean, reflecting a consistent market regime governed by approximately stationary coupling patterns.

To attenuate short-term fluctuations and emphasize persistent structural changes, a moving-average statistic is computed over a sliding window of length w,

$$\bar{y}_t = \frac{1}{w} \sum_{i=t-w+1}^{t} y_i. \tag{6.19}$$

The nominal mean μ_0 and standard deviation σ_0 of $\bar{y}_t$ are estimated from a baseline period representative of stable market conditions. Deviations from the nominal regime are quantified using the standardized statistic

$$r_t = \frac{\bar{y}_t - \mu_0}{\sigma_0}. \tag{6.20}$$

Sustained departures of r_t from zero indicate the presence of *concept drift*, corresponding to changes in the global spatial organization of LMPs with respect to the reduced graph.

To formally identify change points in the smoothness trajectory, a cumulative sum (CuSum) test is applied to the standardized sequence $\{r_t\}$ [97]. When the CuSum statistic exceeds a predefined threshold, a concept-drift alarm is raised. Following detection, the reference statistics are updated by re-estimating (μ_0, σ_0) from the most recent stable segment, and the CuSum statistic is reset to enable detection of subsequent regime changes.

6.4 Complexity Analysis

6.4.1 Offline phase

The offline phase consists of graph construction, spectral clustering, representative selection, reduced-order graph formation, and baseline preparation. Building the impedance-weighted graph and Laplacian in (6.2.1) requires traversing all edges, yielding $O(E)$ complexity. Spectral clustering then computes the $r = K \ll N$ smallest eigenvectors of the symmetric normalized Laplacian; while a full eigendecomposition costs $O(N^3)$, sparse iterative methods compute only the required r eigenvectors with complexity approximately $O(rE)$, which scales as $O(rN)$ for sparse transmission networks ($E = O(N)$). Applying K-means to the resulting $N \times r$ embedding requires $O(NKrI)$ operations, where I denotes the number of iterations. Representative (medoid) selection relies on the impedance-weighted shortest-path distance in (6.3.1); computing distances from each candidate node via Dijkstra's algorithm requires $O(E \log N)$ per source, leading to $O(NE \log N)$ overall, which scales as $O(N^2 \log N)$ under sparsity. Forming the Kron-reduced Laplacian $L^\star$ via the Schur complement requires solving linear systems involving L_{RR}; explicit inversion of $L_{RR} \in \mathbb{R}^{(N-N_m)\times(N-N_m)}$ incurs $O((N - N_m)^3)$ time in the worst case, although in practice linear systems are solved directly for improved numerical stability. Finally, baseline preparation for GEMAD+ computes and stores $\{d_j : x_j \in \mathcal{S}_2\}$ according to (6.3.2); for each x_j, distance computation to $\mathcal{S}_1$ costs $O(N_1 N_m)$, selecting the k nearest neighbors requires $O(N_1 + k \log N_1)$, and sorting $\{d_j\}_{j=1}^{N_2}$ costs $O(N_2 \log N_2)$, yielding total complexity

$$O(N_2 \, (N_1 N_m + k \log N_1) + N_2 \log N_2) \, .$$

Overall, the offline complexity is dominated by the spectral embedding ($O(rE)$ for sparse networks) and the Kron reduction ($O((N - N_m)^3)$), in addition to the baseline preparation cost. Importantly, the offline phase is executed once during system initialization (or only upon structural changes in the network), and therefore does not impact the real-time per-step monitoring latency.

6.4.2 Online phase

At each time t, the algorithm processes the reduced observation $x_t \in \mathbb{R}^{N_m}$. Online anomaly detection (GEMAD+) computes distances to the N_1 reference samples in $\mathcal{S}_1$, yielding $O(N_1 N_m)$ complexity, with kNN selection and probability evaluation adding $O(k \log N_1 + \log N_2)$. The localization mechanism introduces no additional asymptotic cost, since the node-level contributions $\alpha_{i,t}$ are obtained directly from the deviation computation and require only normalization. Drift monitoring evaluates the reduced-graph smoothness $y_t = x_t^\top L^\star x_t$, costing $O(N_m^2)$ (for dense $L^\star$), while the moving-average and CuSum updates are $O(1)$. Hence, the overall per-time-step online complexity is

$$O\Big(N_1 N_m + k \log N_1 + \log N_2 + N_m^2\Big) \, .$$

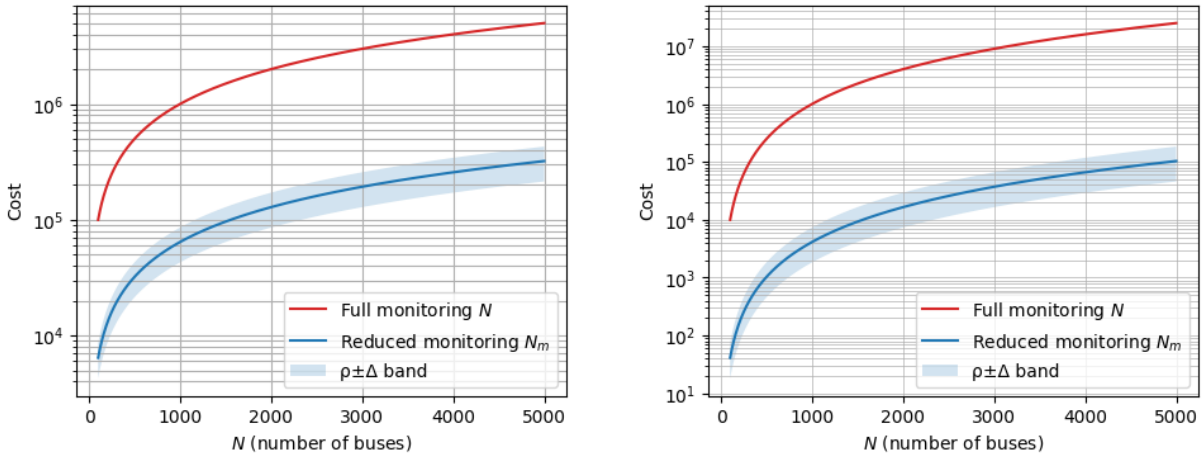


Figure 6.2: Online computational cost as a function of system size N. Left: GEMAD+ anomaly detection. Right: GS-Drift monitoring. Full-network monitoring ($N_m = N$) is compared with reduced monitoring ($N_m = \rho N$); the shaded region reflects the observed variability in ρ.

6.4.3 Full-System vs. Reduced Monitoring Complexity

If the same online monitoring pipeline were applied to the full network (dimension N instead of N_m), the per-time-step complexity would scale as

$$O\left(N_1 N + k \log N_1 + \log N_2 + N^2\right),$$

where the linear term arises from kNN distance computation and the quadratic term from computing the GS-Drift smoothness value $y_t = x_t^\top L x_t$.

Under reduced monitoring, operating in the dimension $N_m \ll N$, the complexity becomes

$$O\left(N_1 N_m + k \log N_1 + \log N_2 + N_m^2\right),$$

with the smoothness term evaluated using the reduced Laplacian $L^\star$. Hence, dimensionality reduction yields an approximate linear reduction factor of N/N_m for the GEMAD+ anomaly-detection term and a quadratic reduction factor of $(N/N_m)^2$ for the GS-Drift component. For the NPCC 140-bus system with $N_m = 6$, this corresponds to a reduction of approximately 23× in the kNN term and more than 500× in the smoothness computation.

Figure 6.2 illustrate the resulting scaling behavior as a function of system size N, assuming an approximately stable monitoring ratio $\rho = N_m/N$ obtained via the proposed node-selection procedure. The reduced framework preserves the asymptotic growth rates of the full system while substantially lowering computational magnitude.

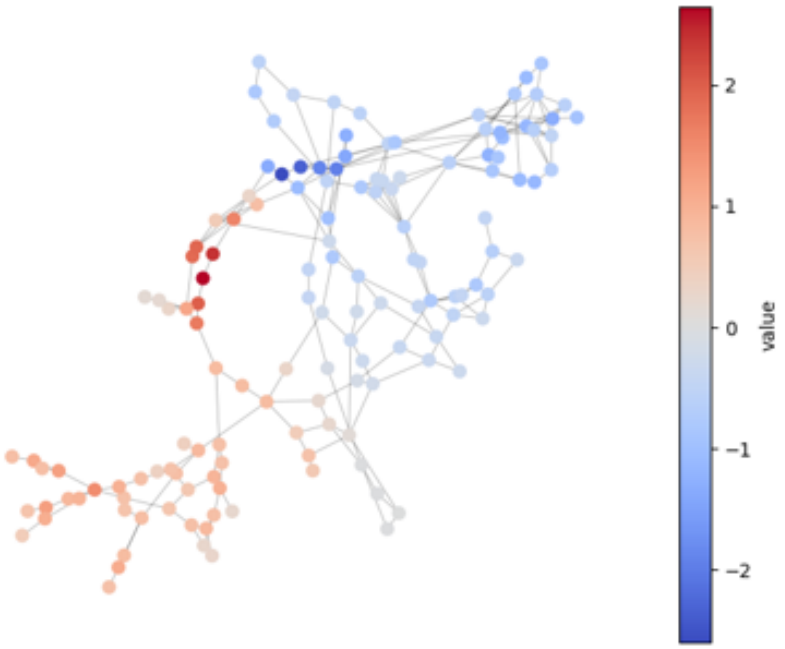


Figure 6.3: Single-snapshot LMP signals exhibiting piecewise-smooth structure on the NPCC 140-bus system.

6.5 Experimental Evaluation

The experimental scenarios used in this work are derived from the Stealthy Manipulated LMP Timeseries (SMLT) dataset, which was explicitly designed to model realistic electricity market conditions under stealthy cyber–physical manipulation. The dataset is grounded in the NPCC benchmark system, a widely adopted test system for market and cyber–physical studies [76]. Operational feasibility constraints are enforced on all attack parameters such that injected perturbations remain within realistic operating ranges, for example through bounded relative deviations from nominal system parameters, ensuring practical relevance and avoiding unrealistic conditions [49].

Before applying the proposed monitoring framework, we verify that the LMP signals in the NPCC system exhibit the piecewise-smooth structure assumed in Section 6.2.2. As shown in Fig. 6.3, the LMP signal varies smoothly within electrically coherent regions, while sharper variations appear around localized boundary interfaces. Quantitatively, the single-snapshot LMP signal has a total graph smoothness of 34.39, with 81.97% of its spectral energy concentrated in the lowest 20% of graph frequencies and a spectral centroid of 0.21. These results indicate that most of the signal energy is concentrated in low graph frequencies, supporting the interpretation of LMPs as piecewise-smooth graph signals on the transmission network.

The experimental evaluation considers four scenarios: Transmission Line Rating (TLR) attacks (Cases 1 and 2), a physical line fault (Case 3), and a strategic market manipulation scenario involving falsified generator ramping constraints in the look-ahead dispatch (Case 4). TLR attacks manipulate transmission line thermal limits to induce artificial congestion without

Table 6.2: Dataset Description

Attribute	Value
System	NPCC 140-bus
Time resolution	Hourly
Duration	15 weeks
Number of features / buses	140
Number of scenarios	4
Types of disturbances	TLR, physical fault, ramp-induced manipulation
Anomalous period	168 hours

Table 6.3: GEMAD+ and GS-Drift Hyperparameters

GEMAD+ Parameters	
w (window length)	366
k (nearest neighbors)	2
λ (forgetting factor)	0.95
α (significance level)	0.05
h (anomaly threshold)	Empirical
GS-Drift Parameters	
w (window length)	168
μ_0, σ_0	Estimated from nominal data
h_d (CuSum threshold)	Empirical

altering load or network topology data, reflecting a cyber–physical market manipulation strategy that has been extensively studied in the literature [38]. Case 3 represents a physical disturbance in the form of a transmission line fault, while Case 4 captures an economically motivated market manipulation strategy based on ramp-induced data falsification, which constrains generator ramping capabilities in the look-ahead dispatch and leads to distorted price formation [87]. Table 6.2 summarizes the key characteristics of the dataset used in this study. Our experimental evaluation addresses the following questions:

Q1: To what extent does the GSP-based node selection preserve the global behavioral characteristics of the energy market while maintaining comparable anomaly-detection performance under reduced monitoring coverage?

Q2: How does graph-signal smoothness contribute to enhancing drift detection and adaptive anomaly-detection stability in terms of update frequency, robustness, and false-alarm rate (FAR)?

Q3: How effective is the proposed framework in accurately localizing anomalies across the four scenarios considered in this chapter?

6.5.1 Impact of Dimensionality Reduction on Detection Performance

The objective here is not only dimensionality reduction for computational efficiency, but also to evaluate whether a compact, information-preserving subset of nodes selected based on physically meaningful similarity maintain or enhance anomaly-detection accuracy. Therefore, we compare our dimensionality-reduction strategy in Section 6.3.1 with several alternative approaches: K-means baseline, conventional PCA, and gLPCA method [129]. Moreover, we consider two spectral-clustering–based variants that rely on purely signal-driven similarity and hybrid (signal–electrical) similarity, respectively. This comparative design allows us to assess the benefit of physically grounded node selection relative to purely data-driven and mixed similarity measures. All approaches retain only a small subset of nodes, as reported in Table 6.6, and GEMAD+ together with the GS-Drift detector are executed using identical hyperparameters to those of the full 140-node configuration. Each method is evaluated under the same market scenarios using the performance metrics: DR, precision, F1 score, AUC, and FAR.

The results in Table 6.6 show that a sole dimensionality reduction severely degrades detection performance, with K-means using raw features yielding the lowest F1 score (0.724) and the highest FAR (0.103). Incorporating PCA improves performance, however, it remains below the full 140-node baseline. In contrast, all four graph-based spectral clustering methods, each leveraging physical or signal-informed similarities, achieve accuracy comparable to or exceeding the full model while using *only* 6–13 nodes (out of 140 in total). Electrical similarity attains the highest AUC (0.953), hybrid similarity achieves the lowest FAR (0.0075), and gLPCA similarity provides the best F1 score (0.924). These results indicate that detection performance is preserved not only by dimensionality reduction, but also, by selecting the 'most appropriate' nodes grounded in the physics and spatial structure of the grid.

6.5.2 Impact of GS-Drift on Concept Drift Detection

In this experiment, we evaluate a graph-signal–based drift detector, hereafter referred to as *GS-Drift*, and compare it with two established methods, CAD and CKL [109], using the aforementioned cases to assess how each approach responds under varying scenarios. Three evaluation metrics are employed in this study. Firstly, the *update frequency (UF)*, defined as $\mathrm{UF} = N_{\mathrm{updates}}/T$, which reflects how often the drift detector triggers model updates and, thus, indicates the responsiveness and efficiency of adaptation. In addition, the *stability* is computed as the normalized mean absolute first-order difference, i.e.,

$$\text{stability} = \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{|y_{t+1} - y_t|}{|y_t| + \varepsilon}, \tag{6.21}$$

Table 6.4: Update Frequency (UF), Stability, and FAR comparison across models.

Case	GS-Drift			CKL			CAD		
	UF	Stability	FAR	UF	Stability	FAR	UF	Stability	FAR
case1	0.0004	0.3064	0.0022	0.0024	0.9703	0.0764	0.0011	0.9907	0.102
case2	0.0004	0.2979	0.0022	0.0020	0.8886	0.0823	0.0011	0.9993	0.0508
case3	0.0004	0.3599	0.0212	0.0016	1.0097	0.0183	0.0015	1.0148	0.0965
case4	0.0004	0.3112	0.0201	0.0020	1.0514	0.0178	0.0015	0.9085	0.0878
Average	0.0004	0.3189	0.0114	0.0020	0.9800	0.0487	0.0013	0.9783	0.0843

where y_t denotes the drift score produced by the corresponding detector and $\epsilon > 0$ is used for numerical stability. Normalization accounts for differences in numerical scale across drift detectors, ensuring that the stability reflects temporal variability rather than absolute magnitude. A lower value (close to 0) corresponds to more temporally robust and noise-resistant behavior. Finally, the FAR is produced by the GEMAD+ anomaly-detection module, as discussed in the previous experiment.

For a fair comparison, GEMAD+ is configured with parameters identical to those reported in Table 6.3, and only representative sets are used as monitoring nodes. The hyperparameter values were selected to maximize the F1 score on nominal validation data, thereby balancing detection rate and false-alarm rate, and were kept fixed across all experiments.

The results in Table 6.4 demonstrate the evident superiority of the GS-Drift model. GS-Drift consistently outperforms both CKL and CAD in all evaluation metrics. On average, it reduces the update frequency by 80.4% relative to CKL and by 70.8% relative to CAD, while achieving stability improvements of 66.9% and 67.3%, respectively. In terms of FAR, GS-Drift yields an average reduction of 41.4% compared to CKL and 87.2% compared to CAD.

These improvements stem from two fundamental properties of graph signal smoothness: (1) it captures the spatiotemporal structure of LMP dynamics, allowing the detector to track gradual, system-wide changes in the underlying LMPs distribution; and (2) it is inherently less sensitive to localized anomalies that perturb only small groups of adjacent nodes, since such disturbances typically preserve the broader spatial dependencies in the network. As a result, GSP-based smoothness provides a more reliable indicator of concept drift than the purely temporal CKL and CAD methods, which are more easily disrupted by node-level anomalies and therefore exhibit higher instability. As shown in Fig. 6.4, GS-Drift produces a gradually varying and well-structured drift trajectory, whereas CKL exhibits high-frequency oscillations and CAD generates abrupt jumps and large discontinuities, explaining their higher instability and FAR rates.

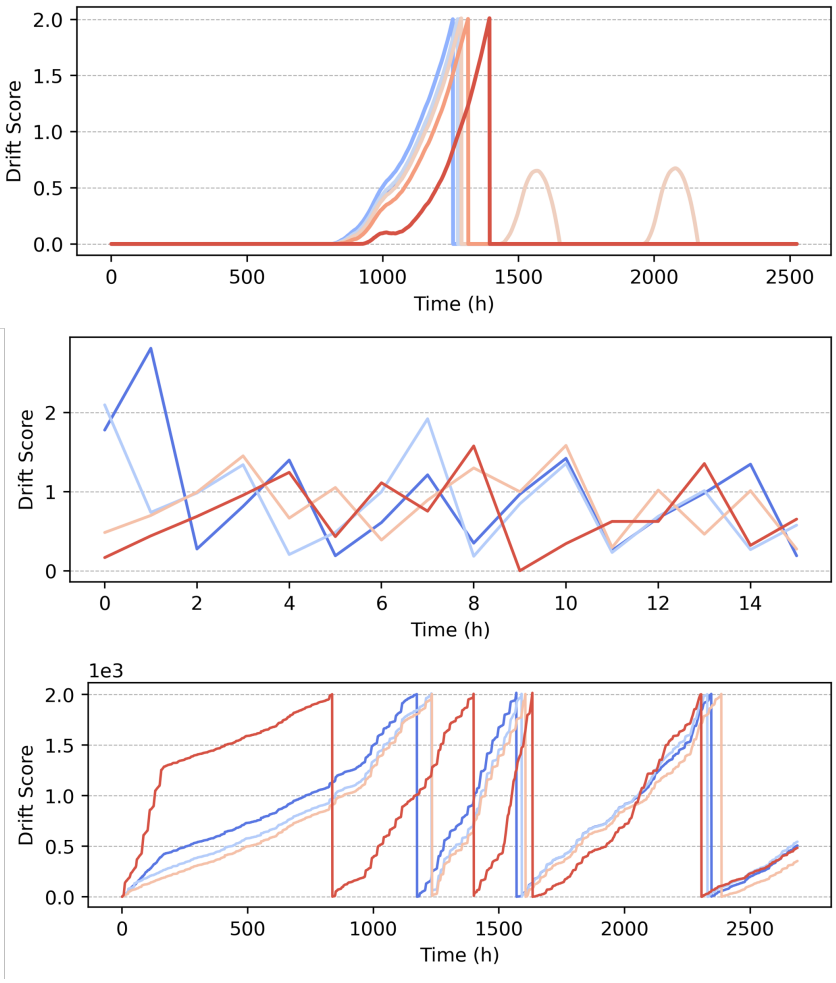


Figure 6.4: Time-series visualization of drift scores produced by GS-Drift (top), CKL (middle), and CAD (bottom).

6.5.3 Anomaly Localization

To evaluate anomaly localization performance, we adopt two complementary metrics inspired by graph-based image processing and signal analysis, where spatial attribution of signal deviations is commonly studied. Unlike classification tasks, anomaly localization does not have a universally established evaluation metric, particularly under partial observability where only representative nodes are monitored. Therefore, we employ two metrics that capture localization performance from both regional and node-level perspectives.

The first metric, probability mass (PM), measures the fraction of anomaly attribution concentrated within the cluster containing the true anomalous bus. This metric reflects whether the detection mechanism correctly identifies the affected electrical region. The second metric, representative-level localization error (LE), quantifies the electrical distance between the true anomalous bus and the representative node most frequently identified as anomalous.

Probability–Mass Score: During an anomalous interval, each representative node reports a per–time-step contribution value $\eta_{i,t}$, and the node that achieves the highest $\eta_{i,t}$ is recorded at each step. Let f_i denote the number of times the representative node i appears as the main contributor. For a given true anomalous bus $b^\star$, let $C(b^\star)$ denote its corresponding cluster. The probability–mass assigned to the true cluster is defined as

$$\mathrm{PM}(b^\star) = \frac{\sum_{i \in C(b^\star)} f_i}{\sum_j f_j}, \tag{6.22}$$

which quantifies the proportion of the alarm frequency driven by anomalies captured within the true cluster. Higher values indicate better cluster-level localization.

Representative–Level Localization Error: To measure node-level precision, we compute the distance between the true anomalous bus and the representative node with the highest overall alarm frequency f_i. Let $\hat{b}$ denote this predicted representative. The node-localization error is given by the shortest-path distance over the impedance-weighted graph:

$$\mathrm{LE}_{\mathrm{rep}}(b^\star) = d(b^\star, \hat{b}), \tag{6.23}$$

where

$$d(i, j) = \min_{\pi: i \to j} \sum_{(u,v) \in \pi} |z_{uv}|, \tag{6.24}$$

is the minimum impedance-path length between the nodes i and j. Smaller values of $\mathrm{LE}_{\mathrm{rep}}$ indicate more accurate node-level localization.

The results in Table 6.5 and in Fig. 6.5 and 6.6 show that electrical similarity yields the best localization performance, achieving both the highest probability mass and the lowest localization error, followed by gLPCA. In contrast, purely data-driven clustering and similarity methods, such as signal similarity or K-means, exhibit substantially poorer results, as reflected in the lower-left and lower-right regions of Fig. 6.6. These findings indicate that physically informed graph structures are essential for preserving anomaly spatial signatures and, consequently, for achieving reliable attribution under market disturbances.

Table 6.5: Localization Performance: Probability-Mass and Representative-Level Error

		$\mathrm{PM}(b^\star)$	$\mathrm{LE}_{\mathrm{rep}}$
K-means	Original	0.2559 ± 0.1305	0.3398 ± 0.1241
	PCA	0.1860 ± 0.1453	0.2508 ± 0.1261
	gLPCA	0.4122 ± 0.1155	0.1154 ± 0.1461
Spectral Clustering	Signal similarity	0.4018 ± 0.2825	0.3589 ± 0.2207
	Electrical similarity	0.8393 ± 0.1416	0.0620 ± 0.0281
	Hybrid similarity	0.2843 ± 0.3322	0.2309 ± 0.1957

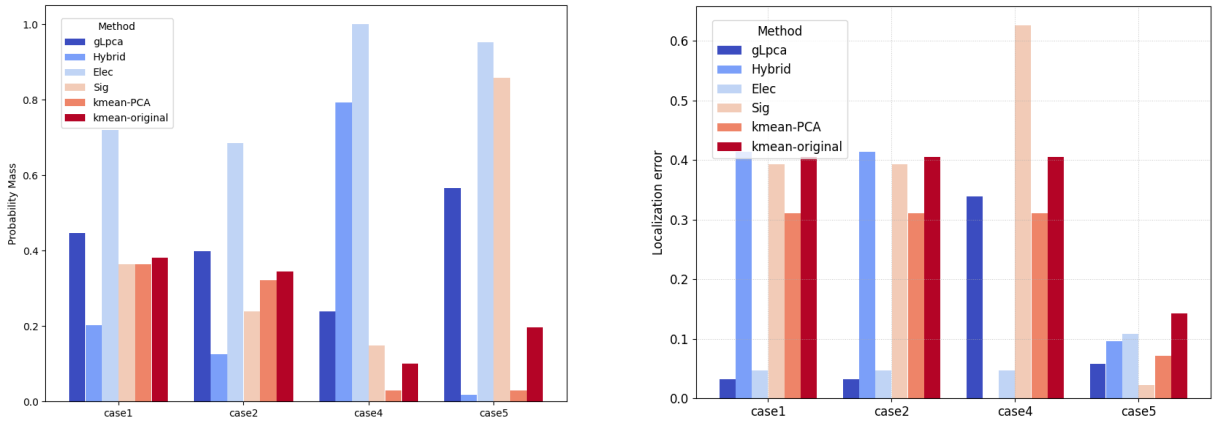


Figure 6.5: Anomaly localization performance: (left) probability–mass score (PM) at the cluster level and (right) representative-level localization error (LE).

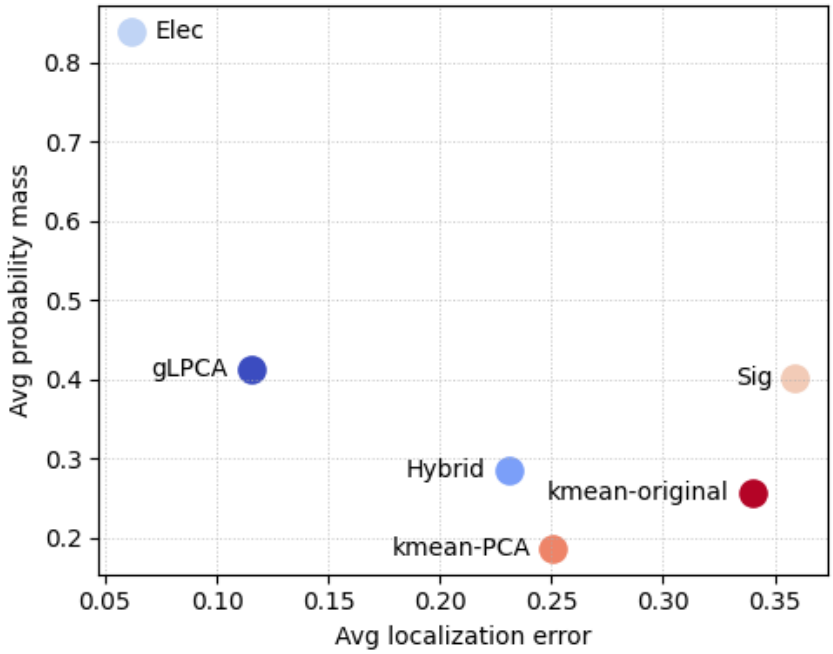


Figure 6.6: Localization error versus probability-mass trade-off across all methods. Higher position and leftward displacement indicate better anomaly localization.

Table 6.6: Detection performance under dimensionality reduction across node-selection methods.

Method	**Variant**	**#Nodes**	**Performance Metrics**				
			DR	Precision	F1	AUC	FAR
Full Dataset	–	140	0.918 ± 0.045	0.912 ± 0.019	0.915 ± 0.013	0.951 ± 0.018	0.017 ± 0.005
K-means	Original	2	0.881 ± 0.066	0.616 ± 0.014	0.724 ± 0.032	0.889 ± 0.031	0.103 ± 0.002
	PCA	2	0.891 ± 0.061	0.919 ± 0.019	0.903 ± 0.021	0.938 ± 0.028	0.015 ± 0.005
	gLPCA	12	0.911 ± 0.036	0.941 ± 0.049	**0.924 ± 0.009**	0.950 ± 0.013	0.011 ± 0.010
Spectral Clustering	Signal similarity	12	0.896 ± 0.064	0.951 ± 0.045	0.920 ± 0.022	0.943 ± 0.027	0.0095 ± 0.0099
	Electrical similarity	6	**0.921 ± 0.043**	0.915 ± 0.018	0.917 ± 0.013	**0.953 ± 0.019**	0.016 ± 0.005
	Hybrid similarity	13	0.875 ± 0.085	**0.960 ± 0.033**	0.912 ± 0.035	0.934 ± 0.039	**0.0075 ± 0.007**

6.6 Conclusions

We introduce a physics-informed graph signal processing framework for reduced-order monitoring of electricity markets under cyber–physical disturbances and concept drift. Using the spatial structure of locational marginal prices through an impedance-weighted graph representation, our approach enables representative node selection, stable anomaly detection, drift adaptation, and effective localization. Experiments over realistic market scenarios demonstrate that the framework preserves detection performance under substantial dimensionality reduction while improving stability and reducing false alarms under non-stationary conditions. Future work includes Laplacian PCA–based dimensionality reduction to enable reconstruction of LMP signals, with the aim of supporting market mitigation under disturbance scenarios. In addition, we plan to evaluate performance under peak-demand periods while manipulation events coinciding with concept drift.

Chapter 7

Discussion and Conclusion

This thesis investigated the problem of detecting financially motivated cyber-attacks in electricity markets, focusing on stealthy manipulation of Locational Marginal Prices (LMPs) through False Data Injection Attacks (FDIAs). The research was motivated by the observation that modern energy markets operate as high-dimensional, non-stationary cyber-physical systems in which traditional residual-based and distribution-dependent detection approaches are fundamentally inadequate. The central research question asked whether adaptive, data-driven monitoring can reliably identify stealthy market manipulation without assuming trusted infrastructure or stationary behaviour.

This thesis demonstrates that adaptive, data-driven monitoring can reliably detect stealthy market manipulation without assuming trusted infrastructure or stationarity, provided that manipulation produces persistent structural deviations in temporal or spatial market behaviour. However, detection becomes fundamentally limited when adversarial actions are either too brief to accumulate statistical evidence or indistinguishable from genuine regime shifts.

7.1 Summary of Findings

The findings show that reliable detection requires moving beyond static anomaly detection toward adaptive monitoring that jointly captures temporal evolution, spatial dependencies, and structural market behaviour. Two fundamental properties of manipulation emerge from the analysis. First, manipulation leaves a measurable footprint through changes in the distribution and temporal evolution of LMPs, even when individual observations remain within normal operating ranges. Unlike natural variability driven by supply–demand dynamics, manipulation introduces structured inconsistencies in price evolution. Second, manipulation is not localized: a price change at one node propagates across the network according to physical connectivity. These coordinated deviations differ from normal spatial correlations and enable localization of anomalous regions under evolving market conditions.

Beyond empirical performance, the results provide broader insight into energy-market cybersecurity. Market manipulation manifests as a structural phenomenon rather than a purely statistical anomaly: adversaries simultaneously exploit physical constraints, temporal settlement mechanisms, and economic incentives. Consequently, effective monitoring must integrate statistical learning with physical system knowledge and market mechanisms instead of relying solely on distributional deviations. Under this perspective, anomaly detection becomes a cyber-physical inference problem rather than a conventional pattern-recognition task.

In doing so, the thesis contributes new knowledge in three ways. First, it establishes that stealthy LMP manipulation leaves observable temporal distributional signatures even when values remain within operational bounds. Second, it shows that manipulation is inherently a network-level phenomenon detectable through spatial propagation rather than isolated anomalies. Third, it demonstrates that reliable monitoring in electricity markets requires integrating statistical learning with physical and economic system structure, reframing anomaly detection as a cyber-physical inference problem. Table 7.1 summarizes the contributions of each chapter and highlights how the proposed approaches progressively address key challenges in electricity market monitoring, including concept drift, real-time detection, and scalability.

Overall, the thesis contributes to the state of the art in two main ways. Empirically, it provides a labelled and reproducible benchmark for studying stealthy LMP manipulation and evaluates detection performance under different attack intensities, drift conditions, and spatial settings. Methodologically, it develops a sequence of data-driven monitoring frameworks that address real-time detection, non-stationary LMP behaviour, and spatial dependence across market nodes. Together, these contributions complement existing residual-based and protection-oriented approaches by moving toward adaptive, scalable, and interpretable market-level detection.

Table 7.1: Summary of thesis chapters and monitoring capabilities.

Chapter	**Problem Addressed**	**Method**	**Handles Concept Drift?**	**Real-Time**	**Scalability**
Ch.2	Background and literature review on LMP manipulation and detection	Systematic literature review	N/A	N/A	N/A
Ch.3	Benchmarking stealthy LMP manipulation	Physics-consistent synthetic benchmarking framework (SMLT)	N/A	N/A	N/A
Ch.4	Real-time detection of LMP anomalies	Incremental change-point detection	Partial	Yes	Limited
Ch.5	Detection under non-stationary markets	Drift-aware anomaly detection	Yes	Yes	Moderate
Ch.6	Scalable monitoring and localization	Graph signal processing + node selection	Yes	Yes	Strong

7.2 Limitations

Despite the effectiveness of the proposed adaptive monitoring framework, certain limitations remain inherent to data-driven detection in non-stationary cyber-physical markets.

1- Observability and Identifiability of Market Manipulation

The proposed framework relies on identifying structural deviations in temporal and spatial market behaviour. Consequently, manipulation must accumulate sufficient statistical evidence to become observable. Short-duration attacks may therefore remain indistinguishable from normal market volatility, particularly when adversarial perturbations remain within typical operational ranges. This reflects a fundamental trade-off between adaptation and sensitivity: mechanisms designed to track legitimate market drift may reduce the contrast of brief adversarial deviations. Economically rational attackers could exploit this property by distributing manipulation across limited time intervals, keeping each occurrence below the statistical detection horizon while gains accumulate over time. A related challenge arises from ambiguity of interpretation. During genuine market regime shifts, such as extreme weather events, sudden demand surges, or transmission congestion, electricity markets naturally experience abrupt price fluctuations and structural changes in Locational Marginal Prices. These legitimate system responses can resemble adversarial manipulation. When such events occur concurrently with an attack, observations may support multiple plausible explanations. In these situations, the limitation becomes one of identifiability rather than sensitivity, as purely data-driven monitoring cannot always uniquely distinguish malicious intervention from genuine system dynamics without incorporating additional contextual information.

2- Synthetic Data Realism

A significant component of the experimental evaluation relies on the SMLT synthetic dataset designed to model stealthy Locational Marginal Price manipulation scenarios. Although the dataset is constructed using physics-consistent constraints and realistic market mechanisms, synthetic environments inevitably simplify real-world market complexity. Real electricity markets involve strategic bidding behavior, unforeseen outages, regulatory interventions, and operational constraints that may not be fully captured in a simulated environment. Therefore, while the synthetic dataset provides a controlled and reproducible benchmarking environment, future validation on additional real market datasets would further strengthen the empirical assessment of the proposed methods.

3- Scalability in Extremely Large ISO Markets

Although the proposed graph-based monitoring framework significantly reduces computational complexity through physics-informed node selection and dimensionality reduction, extremely large electricity markets containing thousands of nodes may still present scalability challenges. Real-time monitoring requirements, combined with high-frequency settlement intervals, could impose strict latency constraints. While the proposed approach improves scalability compared to full-system monitoring, further work may be required to investigate distributed implementations or parallelized architectures for deployment in very large-scale market environments.

4- Potential Adversarial Adaptation

The proposed detection framework assumes adversaries aim to preserve statistical stealth while manipulating market outcomes. However, sophisticated attackers may adapt their strategies once detection mechanisms become known. For instance, adversaries could distribute manipulations across multiple nodes, gradually introduce perturbations over extended time periods, or exploit correlations across market layers to evade detection. Addressing such adaptive adversarial behavior remains an open challenge and motivates future research on adversarially robust anomaly detection and monitoring frameworks tailored to electricity markets.

7.3 Future Research Directions

Short-Term Extensions

In the short term, several methodological extensions could further enhance the proposed framework. Building on **Limitation 1**, concerning observability and identifiability, future work should investigate multi-source contextual monitoring. One promising direction is multi-market integration, where signals from multiple settlement layers, such as day-ahead, real-time, ancillary services, and financial transmission rights markets, are jointly analyzed to identify cross-market inconsistencies indicative of coordinated manipulation. Another important extension involves multi-modal data integration, incorporating contextual information such as generation dispatch levels, congestion indicators, load forecasts, and weather variables to provide a richer representation of market dynamics and improve detection robustness. In relation to **Limitation 2**, future work could also investigate cross-market transferability and broader real-market validation, examining whether models trained in one electricity market environment can partially generalize to others despite differences in network topology, market rules, and settlement mechanisms. This direction is also relevant to **Limitation 3**, as transferable and scalable models could reduce the computational and deployment burden associated with adapting the framework to very large market environments.

Long-Term Directions

From a longer-term perspective, future work should focus on the operational and regulatory integration of data-driven monitoring systems within electricity market infrastructures. Building on **Limitation 3**, concerning scalability and deployment in large ISO markets, this includes investigating how anomaly detection frameworks can support regulatory auditing, automated market surveillance, and real-time alert generation within Independent System Operators (ISOs) or Regional Transmission Organizations (RTOs). Achieving this goal requires addressing practical challenges such as latency constraints, distributed data processing, and human-in-the-loop investigation workflows. Such workflows are also relevant to **Limitation 1**, as they may help contextualize ambiguous alerts during genuine market regime shifts. In response to **Limitation 4**, concerning adaptive adversarial behaviour, future research should also explore adversarially robust detection approaches capable of maintaining reliability under adaptive attack strategies, thereby strengthening the long-term resilience of electricity market monitoring systems.

Bibliography

[1] Q. Zhang, F. Li, Q. Shi, K. Tomsovic, J. Sun, and L. Ren, "Profit-oriented false data injection on electricity market: reviews, analyses, and insights," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 9, pp. 5876–5886, 2020.

[2] G. O. Alsharif, C. Anagnostopoulos, and A. K. Marnerides, "Energy market manipulation via false-data injection attacks: A review," *IEEE Access*, vol. 13, pp. 42 559–42 573, 2025.

[3] P. Cramton, "Electricity market design," *Oxford Review of Economic Policy*, vol. 33, no. 4, pp. 589–612, 2017.

[4] J. L. Sweeney, "The california electricity crisis: Lessons for the future," *Bridge*, vol. 32, no. 2, pp. 23–31, 2002.

[5] Federal Energy Regulatory Commission, "Presentation & report fy2023 report on enforcement," 2023. [Online]. Available: https://www.ferc.gov/news-events/news/presentation-report-fy2023-report-enforcement

[6] Ofgem, "Ofgem requires intergen to pay £37m over energy market abuse," 2023. [Online]. Available: https://www.ofgem.gov.uk/press-release/ofgem-requires-intergen-pay-ps37m-over-energy-market-abuse

[7] R. Deng, G. Xiao, R. Lu, H. Liang, and A. Vasilakos, "False data injection on state estimation in power systems—attacks, impacts, and defense: A survey," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 411–423, 2016.

[8] C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," in *Proceedings of the 18th international conference on evaluation and assessment in software engineering*, 2014, doi: 10.1145/2601248.2601268.

[9] A. Huseinović, S. Mrdović, K. Bicakci, and S. Uludag, "A survey of denial-of-service attacks and solutions in the smart grid," *IEEE Access*, vol. 8, pp. 177 447–177 470, 2020.

[10] Y. Zhang, T. Huang, and E. F. Bompard, "Big data analytics in smart grids: a review," *Energy informatics*, vol. 1, no. 1, pp. 1–24, 2018.

[11] D. e. a. Syed, "Smart grid big data analytics: Survey of technologies, techniques, and applications," *IEEE Access*, vol. 9, pp. 59 564–59 585, 2020.

[12] H. Zhang, B. Liu, and H. Wu, "Smart grid cyber-physical attack and defense: A review," *IEEE Access*, vol. 9, pp. 29 641–29 659, 2021.

[13] S. Tan, D. De, W.-Z. Song, J. Yang, and S. K. Das, "Survey of security advances in smart grid: A data driven approach," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 397–422, 2016.

[14] B. P. e. a. Bhattarai, "Big data analytics in smart grids: state-of-the-art, challenges, opportunities, and future directions," *IET Smart Grid*, vol. 2, no. 2, pp. 141–154, 2019.

[15] T. Reed, *At the abyss: an insider's history of the Cold War*. Presidio Press, 2007.

[16] D. U. Case, "Analysis of the cyber attack on the ukrainian power grid," *Electricity information sharing and analysis center (E-ISAC)*, vol. 388, no. 1-29, p. 3, 2016.

[17] A. S. Musleh, G. Chen, and Z. Y. Dong, "A survey on the detection algorithms for false data injection attacks in smart grids," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2218–2234, 2019.

[18] C. Konstantinou and M. Maniatakos, "Hardware-layer intelligence collection for smart grid embedded systems," *Journal of Hardware and Systems Security*, vol. 3, pp. 132–146, 2019.

[19] C. Konstantinou, M. Sazos, A. S. Musleh, A. Keliris, A. Al-Durra, and M. Maniatakos, "Gps spoofing effect on phase angle monitoring and control in a real-time digital simulator-based hardware-in-the-loop environment," *IET Cyber-Physical Systems: Theory & Applications*, vol. 2, no. 4, pp. 180–187, 2017.

[20] A. Althobaiti, A. Jindal, A. K. Marnerides, and U. Roedig, "Energy theft in smart grids: a survey on data-driven attack strategies and detection methods," *IEEE Access*, vol. 9, pp. 159 291–159 312, 2021.

[21] M. Ismail, M. F. Shaaban, M. Naidu, and E. Serpedin, "Deep learning detection of electricity theft cyber-attacks in renewable distributed generation," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3428–3437, 2020.

[22] X. Liu, Z. Li, X. Liu, and Z. Li, "Masking transmission line outages via false data injection attacks," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 7, pp. 1592–1602, 2016.

[23] G. Liang, S. R. Weller, F. Luo, J. Zhao, and Z. Y. Dong, "Generalized fdia-based cyber topology attack with application to the australian electricity market trading mechanism," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3820–3829, 2017.

[24] Y. Chen, F. Pan, F. Qiu, A. S. Xavier, T. Zheng, M. Marwali, B. Knueven, Y. Guan, P. B. Luh, L. Wu *et al.*, "Security-constrained unit commitment for electricity market: Modeling, solution methods, and future challenges," *IEEE Transactions on Power Systems*, 2022.

[25] Y. Jiang, M. Chen, and S. You, "A unified trading model based on robust optimization for day-ahead and real-time markets with wind power integration," *Energies*, vol. 10, no. 4, p. 554, 2017.

[26] A. Abur and A. G. Exposito, *Power System State Estimation: Theory and Implementation*. CRC press, 2004.

[27] W. W. Hogan, "Virtual bidding and electricity market design," *The Electricity Journal*, vol. 29, no. 5, pp. 33–47, 2016.

[28] S. Baltaoglu, L. Tong, and Q. Zhao, "Algorithmic bidding for virtual trading in electricity markets," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 535–543, 2018.

[29] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 659–666, 2011.

[30] C. Lo Prete and W. W. Hogan, "Manipulation of day-ahead electricity prices through virtual bidding in the us," *IAEE Energy Forum*, pp. 31–32, 2014.

[31] S. P. Karthikeyan, I. J. Raglend, and D. P. Kothari, "A review on market power in deregulated electricity market," *International Journal of Electrical Power & Energy Systems*, vol. 48, pp. 139–147, 2013.

[32] Z. El Mrabet, N. Kaabouch, H. El Ghazi, and H. El Ghazi, "Cyber-security in smart grid: Survey and challenges," *Computers & Electrical Engineering*, vol. 67, pp. 469–482, 2018.

[33] Y. Lin, A. Abur, and H. Xu, "Identifying security vulnerabilities in electricity market operations induced by weakly detectable network parameter errors," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 1, pp. 627–636, 2020.

[34] H. Xu, Y. Lin, X. Zhang, and F. Wang, "Power system parameter attack for financial profits in electricity markets," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3438–3446, 2020.

[35] J. Kim and L. Tong, "On topology attack of a smart grid: Undetectable attacks and countermeasures," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1294–1305, 2013.

[36] D.-H. Choi and L. Xie, "Ramp-induced data attacks on look-ahead dispatch in real-time power markets," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1235–1243, 2013.

[37] M. Esmalifalak, H. Nguyen, R. Zheng, L. Xie, L. Song, and Z. Han, "A stealthy attack against electricity market using independent component analysis," *IEEE Systems Journal*, vol. 12, no. 1, pp. 297–307, 2015.

[38] H. Ye, Y. Ge, X. Liu, and Z. Li, "Transmission line rating attack in two-settlement electricity markets," *IEEE Transactions on Smart Grid*, vol. 7, no. 3, pp. 1346–1355, 2015.

[39] S. Tan, W.-Z. Song, M. Stewart, J. Yang, and L. Tong, "Online data integrity attacks against real-time electrical market in smart grid," *IEEE Transactions on Smart Grid*, vol. 9, no. 1, pp. 313–322, 2016.

[40] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "A framework for cyber-topology attacks: Line-switching and new attack scenarios," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1704–1712, 2017.

[41] M. R. Mengis and A. Tajer, "Data injection attacks on electricity markets by limited adversaries: Worst-case robustness," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 5710–5720, 2017.

[42] N. A. Ruhi, K. Dvijotham, N. Chen, and A. Wierman, "Opportunities for price manipulation by aggregators in electricity markets," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 5687–5698, 2017.

[43] R. Moslemi, A. Mesbahi, and J. Mohammadpour Velni, "Design of robust profitable false data injection attacks in multi-settlement electricity markets," *IET GT&D*, vol. 12, no. 6, pp. 1263–1270, 2018.

[44] C. Liu, H. Liang, and T. Chen, "Network parameter coordinated false data injection attacks against power system ac state estimation," *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1626–1639, 2020.

[45] Q. e. a. Zhang, "Market-Level Defense Against FDIA and a New LMP-Disguising Attack Strategy in Real-Time Market Operations," *IEEE Transactions on Power Systems*, vol. 36, no. 2, pp. 1419–1431, Mar. 2021.

[46] P. K. Jena, S. Ghosh, E. Koley, D. K. Mohanta, and I. Kamwa, "Design of ac state estimation based cyber-physical attack for disrupting electricity market operation under limited sensor information," *Electric Power Systems Research*, vol. 205, p. 107732, 2022.

[47] H. Badrsimaei, R. Hooshmand, and S. Nobakhtian, "Stealthy and profitable data injection attack on real time electricity market with network model uncertainties," *Electric Power Systems Research*, vol. 205, p. 107742, Apr. 2022.

[48] S. D. Roy and S. Debbarma, "Imposter attacks in energy market operation," *IEEE Transactions on Smart Grid*, vol. 13, no. 5, pp. 3836–3839, 2022.

[49] G. Cheng, Y. Lin, J. Yan, J. Zhao, and L. Bai, "Model-measurement data integrity attacks," *IEEE Transactions on Smart Grid*, vol. 14, no. 6, pp. 4741–4757, 2023.

[50] A. Dey and M. V. Salapaka, "Reinforcement learning-based electricity market vulnerability analysis under cyber-topology attack," in *Proc. IEEE Power and Energy Society*, 2023, doi: 10.1109/ISGT51731.2023.10066378.

[51] J. Ospina, D. M. Fobes, and R. Bent, "On the feasibility of market manipulation and energy storage arbitrage via load-altering attacks," *Energies*, vol. 16, no. 4, p. 1670, 2023.

[52] P. Verma and C. Chakraborty, "Load redistribution attacks against smart grids–models, impacts, and defense: A review," *IEEE Transactions on Industrial Informatics*, 2024.

[53] A. Farahani, H. Delkhosh, H. Seifi, and M. Azimi, "A new bi-level model for the false data injection attack on real-time electricity market considering uncertainties," *Computers and Electrical Engineering*, vol. 118, p. 109468, 2024.

[54] X. Feng, L. Aniello, and S. Hu, "Multi-transaction and carbon-aware strategies for profit-oriented fdias on electricity trading systems," in *2024 IEEE Power and Energy Society General Meeting (PESGM)*. IEEE, 2024, doi: 10.1109/PESGM51994.2024.10688834.

[55] G. Hug and J. A. Giampapa, "Vulnerability assessment of ac state estimation with respect to false data injection cyber-attacks," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1362–1370, 2012.

[56] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 645–658, 2011.

[57] U. Cali, M. Kuzlu, M. Pipattanasomporn, J. Kempf, and L. Bai, *Digitalization of Power Markets and Systems Using Energy Informatics*. Springer, 2021.

[58] C. Liu, W. He, R. Deng, Y.-C. Tian, and W. Du, "False-Data-Injection-Enabled Network Parameter Modifications in Power Systems: Attack and Detection," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 177–188, Jan. 2023.

[59] W.-H. Liu, F. F. Wu, and S.-M. Lun, "Estimation of parameter errors from measurement residuals in state estimation (power systems)," *IEEE Transactions on power systems*, vol. 7, no. 1, pp. 81–89, 1992.

[60] R. e. a. Tan, "Modeling and mitigating impact of false data injection attacks on automatic generation control," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 7, pp. 1609–1624, 2017.

[61] R. Deng, G. Xiao, and R. Lu, "Defending Against False Data Injection Attacks on Power System State Estimation," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 1, pp. 198–207, Feb. 2017.

[62] C. Liu, M. Zhou, J. Wu, C. Long, and D. Kundur, "Financially motivated fdi on sced in real-time electricity markets: Attacks and mitigation," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1949–1959, 2017.

[63] D. K. Molzahn and J. Wang, "Detection and Characterization of Intrusions to Network Parameter Data in Electric Power Systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3919–3928, Jul. 2019.

[64] X. Li and K. W. Hedman, "Enhancing Power System Cyber-Security With Systematic Two-Stage Detection Strategy," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 1549–1561, Mar. 2020.

[65] P. K. Jena, E. Koley, and S. Ghosh, "An Optimal Scheme for Installation of PMUs and IEDs to Reinforce Electricity Market Immunity Against Data Attacks in Smart Grid," *IEEE Journal of Emerging and Selected Topics in Industrial Electronics*, vol. 4, no. 2, pp. 603–613, Apr. 2023.

[66] H. Badrsimaei, R.-A. Hooshmand, and S. Nobakhtian, "Observable placement of phasor measurement units for defense against data integrity attacks in real time power markets," *Reliability Engineering & System Safety*, vol. 230, p. 108957, Feb. 2023.

[67] Z. Zhang, S. Bu, Y. Zhang, and Z. Han, "Market-level integrated detection against cyber attacks in real-time market operations by self-supervised learning," *IEEE Transactions on Smart Grid*, 2024.

[68] Y. Zhang, J. Wang, and J. Liu, "Attack identification and correction for pmu gps spoofing in unbalanced distribution systems," *IEEE transactions on smart grid*, vol. 11, no. 1, pp. 762–773, 2019.

[69] Y. Chen, Y. Tan, and B. Zhang, "Exploiting vulnerabilities of load forecasting through adversarial attacks," in *Proceedings of the tenth ACM International Conference on Future Energy Systems*, 2019, doi: 10.1145/3307772.3328314.

[70] Y. Chen, M. Sun, Z. Chu, S. Camal, G. Kariniotakis, and F. Teng, "Vulnerability and impact of machine learning-based inertia forecasting under cost-oriented data integrity attack," *IEEE Transactions on Smart Grid*, vol. 14, no. 3, pp. 2275–2287, 2022.

[71] R. e. a. Khan, "Energy sustainability–survey on technology and control of microgrid, smart grid and virtual power plant," *IEEE Access*, vol. 9, pp. 104 663–104 694, 2021.

[72] P. Pal, A. Parvathy, K. Devabalaji, S. J. Antony, S. Ocheme, T. S. Babu, H. H. Alhelou, and T. Yuvaraj, "Iot-based real time energy management of virtual power plant using plc for transactive energy framework," *IEEE Access*, vol. 9, pp. 97 643–97 660, 2021.

[73] J. Ruan, Q. Wang, S. Chen, H. Lyu, G. Liang, J. Zhao, and Z. Y. Dong, "On vulnerability of renewable energy forecasting: Adversarial learning attacks," *IEEE Transactions on Industrial Informatics*, 2023.

[74] B. Biggio and F. Roli, "Wild patterns: Ten years after the rise of adversarial machine learning," in *ACM SIGSAC*, 2018, pp. 2154–2156.

[75] G. O. Alsharif, C. Anagnostopoulos, and A. K. Marnerides, "Energy market manipulation via false-data injection attacks: A review," *IEEE Access*, vol. 13, pp. 42 559–42 573, 2025.

[76] Q. Zhang and F. Li, "A dataset for electricity market studies on western and northeastern power grids in the united states," *Scientific Data*, vol. 10, no. 1, p. 646, 2023.

[77] R. Billinton and D. Huang, "Test systems for reliability and adequacy assessment of electric power systems," *Proc. IEEE Gen. Meet.(PES)*, pp. 8–14, 2015.

[78] F. Li and R. Bo, "Small test systems for power system economic studies," in *IEEE PES General Meeting*, 2010.

[79] A. L. Diniz, "Test cases for unit commitment and hydrothermal scheduling problems," in *IEEE PES General Meeting*, 2010.

[80] J. E. Price and J. Goodin, "Reduced network modeling of wecc as a market design prototype," in *2011 IEEE Power and Energy Society General Meeting*, 2011, pp. 1–6.

[81] R. Palma-Behnke, "The chilean test system for economic and reliability analysis," in *IEEE PES General Meeting*, 2010.

[82] I. Peña, C. B. Martinez-Anido, and B.-M. Hodge, "An extended ieee 118-bus test system with high renewable penetration," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 281–289, 2018.

[83] T. Xu, A. B. Birchfield, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Application of large-scale synthetic power system models for energy economic studies," in *2017 50th Hawaii International Conference on System Sciences (HICSS)*, 2017.

[84] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Grid structural characteristics as validation criteria for synthetic networks," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3258–3265, 2017.

[85] H. Xu, Y. Lin, X. Zhang, and F. Wang, "Power system parameter attack for financial profits in electricity markets," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3438–3446, 2020.

[86] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "A framework for cyber-topology attacks: Line-switching and new attack scenarios," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1704–1712, 2019.

[87] D.-H. Choi and L. Xie, "Ramp-induced data attacks on look-ahead dispatch in real-time power markets," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1235–1243, 2013.

[88] A.-H. Mohsenian-Rad and A. Leon-Garcia, "Distributed internet-based load altering attacks against smart power grids," *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 667–674, 2011.

[89] N. Azizan Ruhi, K. Dvijotham, N. Chen, and A. Wierman, "Opportunities for price manipulation by aggregators in electricity markets," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 5687–5698, 2018.

[90] J. Yang, S. Rahardja, and P. Fränti, "Outlier detection: how to threshold outlier scores?" in *Proceedings of the international conference on artificial intelligence, information processing and cloud computing*, 2019, pp. 1–6.

[91] X. Mao, W. Zhu, L. Wu, and B. Zhou, "Comparative study on methods for computing electrical distance," *International Journal of Electrical Power & Energy Systems*, vol. 130, p. 106923, 2021.

[92] M. N. Kurt, Y. Yılmaz, and X. Wang, "Real-time nonparametric anomaly detection in high-dimensional settings," *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 7, pp. 2463–2479, 2020.

[93] R. Deng, G. Xiao, and R. Lu, "Defending against false data injection attacks on power system state estimation," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 1, pp. 198–207, 2015.

[94] H. Badrsimaei, R.-A. Hooshmand, and S. Nobakhtian, "Observable placement of phasor measurement units for defense against data integrity attacks in real time power markets," *Reliability Engineering & System Safety*, vol. 230, p. 108957, 2023.

[95] D. K. Molzahn and J. Wang, "Detection and characterization of intrusions to network parameter data in electric power systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3919–3928, 2018.

[96] X. Li and K. W. Hedman, "Enhancing power system cyber-security with systematic two-stage detection strategy," *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 1549–1561, 2019.

[97] G. Lorden, "Procedures for reacting to a change in distribution," *The annals of mathematical statistics*, pp. 1897–1908, 1971.

[98] S. Hochreiter and J. Schmidhuber, "Lstm can solve hard long time lag problems," *Advances in neural information processing systems*, vol. 9, 1996.

[99] P. Malhotra, L. Vig, G. Shroff, P. Agarwal *et al.*, "Long short term memory networks for anomaly detection in time series." in *Esann*, vol. 2015, 2015, p. 89.

[100] G. V. Moustakides, "Optimal stopping times for detecting changes in distributions," *the Annals of Statistics*, vol. 14, no. 4, pp. 1379–1387, 1986.

[101] S. Guha, N. Mishra, G. Roy, and O. Schrijvers, "Robust random cut forest based anomaly detection on streams," in *International conference on machine learning*. PMLR, 2016, pp. 2712–2721.

[102] S. Li, Y. Yılmaz, and X. Wang, "Quickest detection of false data injection attack in wide-area smart grids," *IEEE Transactions on Smart Grid*, vol. 6, no. 6, pp. 2725–2735, 2014.

[103] H. Ren, B. Xu, Y. Wang, C. Yi, C. Huang, X. Kou, T. Xing, M. Yang, J. Tong, and Q. Zhang, "Time-series anomaly detection service at microsoft," in *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, 2019, pp. 3009–3017.

[104] S. Sadik and L. Gruenwald, "Research issues in outlier detection for data streams," *Acm Sigkdd Explorations Newsletter*, vol. 15, no. 1, pp. 33–40, 2014.

[105] A. Ramdas, F. Yang, M. J. Wainwright, and M. I. Jordan, "Online control of the false discovery rate with decaying memory," *Advances in neural information processing systems*, vol. 30, 2017.

[106] U.S. Energy Information Administration, "CAISO Wholesale Electricity Market Data," 2025, accessed: 06-Mar-2025. [Online]. Available: https://www.eia.gov/electricity/wholesalemarkets/data.php?rto=caiso

[107] Y. Yilmaz, "Online nonparametric anomaly detection based on geometric entropy minimization," in *2017 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2017, pp. 3010–3014.

[108] A. Rahimi and A. Y. Sheffrin, "Effective market monitoring in deregulated electricity markets," *IEEE Transactions on Power systems*, vol. 18, no. 2, pp. 486–493, 2003.

[109] G. O. Alsharif, C. Anagnostopoulos, and A. K. Marnerides, "Drift-aware unsupervised detection of stealthy fdia towards energy market," in *2025 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2025, pp. 1–7.

[110] M. Mohammadpourfard, Y. Weng, M. Pechenizkiy, M. Tajdinian, and B. Mohammadi-Ivatloo, "Ensuring cybersecurity of smart grid against data integrity attacks under concept drift," *International Journal of Electrical Power & Energy Systems*, vol. 119, p. 105947, 2020.

[111] O. Boyaci, A. Umunnakwe, A. Sahu, M. R. Narimani, M. Ismail, K. R. Davis, and E. Serpedin, "Graph neural networks based detection of stealth false data injection attacks in smart grids," *IEEE Systems Journal*, vol. 16, no. 2, pp. 2946–2957, 2021.

[112] O. Boyaci, M. R. Narimani, K. R. Davis, M. Ismail, T. J. Overbye, and E. Serpedin, "Joint detection and localization of stealth false data injection attacks in smart grids using graph neural networks," *IEEE Transactions on Smart Grid*, vol. 13, no. 1, pp. 807–819, 2021.

[113] A. Takiddin, R. Atat, M. Ismail, O. Boyaci, K. R. Davis, and E. Serpedin, "Generalized graph neural network-based detection of false data injection attacks in smart grids," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 7, no. 3, pp. 618–630, 2023.

[114] K. Ding, J. Li, R. Bhanushali, and H. Liu, "Deep anomaly detection on attributed networks," in *Proceedings of the 2019 SIAM international conference on data mining*. SIAM, 2019, pp. 594–602.

[115] H. Zhao, Y. Wang, J. Duan, C. Huang, D. Cao, Y. Tong, B. Xu, J. Bai, J. Tong, and Q. Zhang, "Multivariate time-series anomaly detection via graph attention network," in *2020 IEEE international conference on data mining (ICDM)*. IEEE, 2020, pp. 841–850.

[116] A. Deng and B. Hooi, "Graph neural network-based anomaly detection in multivariate time series," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 35, no. 5, 2021, pp. 4027–4035.

[117] Y. Liu, Z. Li, S. Pan, C. Gong, C. Zhou, and G. Karypis, "Anomaly detection on attributed networks via contrastive self-supervised learning," *IEEE transactions on neural networks and learning systems*, vol. 33, no. 6, pp. 2378–2392, 2021.

[118] F. Dorfler and F. Bullo, "Kron reduction of graphs with applications to electrical networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, 2012.

[119] A. Pal, G. A. Sanchez-Ayala, V. A. Centeno, and J. S. Thorp, "A pmu placement scheme ensuring real-time monitoring of critical buses of the network," *IEEE transactions on power delivery*, vol. 29, no. 2, pp. 510–517, 2013.

[120] M. Jamei, R. Ramakrishna, T. Tesfay, R. Gentz, C. Roberts, A. Scaglione, and S. Peisert, "Phasor measurement units optimal placement and performance limits for fault localization," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 1, pp. 180–192, 2019.

[121] E. Drayer and T. Routtenberg, "Detection of false data injection attacks in smart grids based on graph signal processing," *IEEE Systems Journal*, vol. 14, no. 2, pp. 1886–1896, 2019.

[122] M. A. Hasnat and M. Rahnamay-Naeini, "A graph signal processing framework for detecting and locating cyber and physical stresses in smart grids," *IEEE Transactions on Smart Grid*, vol. 13, no. 5, pp. 3688–3699, 2022.

[123] R. Ramakrishna and A. Scaglione, "Grid-graph signal processing (grid-gsp): A graph signal processing framework for the power grid," *IEEE Transactions on Signal Processing*, vol. 69, pp. 2725–2739, 2021.

[124] S. Chen, A. Singh, and J. Kovačević, "Multiresolution representations for piecewise-smooth signals on graphs," *arXiv preprint arXiv:1803.02944*, 2018.

[125] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, no. 4, pp. 395–416, 2007.

[126] D. Zambon, C. Alippi, and L. Livi, "Concept drift and anomaly detection in graph streams," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 11, pp. 5592–5605, 2018.

[127] Y. Yao and L. B. Holder, "Detecting concept drift in classification over streaming graphs," in *Proceedings of the IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, 2016, pp. 843–850.

[128] E. Isufi, A. S. U. Mahabir, and G. Leus, "Blind graph topology change detection," *IEEE Signal Processing Letters*, vol. 25, no. 5, pp. 655–659, 2018.

[129] B. Jiang, C. Ding, B. Luo, and J. Tang, "Graph-laplacian pca: Closed-form solution and robustness," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2013, pp. 3492–3498.