



Full length article

Situational awareness through security-based analysis of controllability and observability in power grids

Mohammed Asiri ^a, Neetesh Saxena ^{a,*}, Subhash Lakshminarayana ^b^a School of Computer Science and Informatics, Cardiff University, UK^b School of Engineering, University of Warwick, UK

ARTICLE INFO

Dataset link: <https://www.nyiso.com/load-data>

Keywords:

Power system
Indicators of Compromise
Industrial control systems
Threat detection
Cyber security
ICS
Operational technology

ABSTRACT

The integration of cyber and physical components in modern power grids has introduced vulnerabilities that challenge operators' situational awareness during cyber attacks. Adversarial actions can compromise the grid's controllability, defined as the ability to steer system states through control inputs, and observability, the ability to infer those states from sensor data. This paper addresses the lack of quantitative tools for assessing how Indicators of Compromise (IoCs) affect these key properties. We propose a comprehensive framework that evaluates the operational impact of cyber-physical attacks by quantifying how IoCs degrade system controllability and observability. Using a co-simulation environment, we simulate three coordinated attack scenarios: False Command Injection (FCI), Load Redistribution (LR), and False Data Injection (FDI). The framework introduces novel quantitative indices that capture both structural and operational degradation in controllability and observability capabilities.

Our results reveal that coordinated attacks systematically exploit the interdependencies between cyber and physical layers, causing cascading effects that compromise both controllability and observability simultaneously. In the FCI case, the system experiences a 25.8% drop in controllability and a 32.6% reduction in observability redundancy, with cascading effects on voltage stability and power flow regulation. The LR scenario causes a 13.4% decline in controllability, primarily through operational stress redistribution, while preserving structural connectivity. The false data injection attack conceals physical disruptions, leading to a 33.5% drop in observability margin and masking critical state changes. The proposed indexes successfully quantify system degradation patterns that remain undetected by conventional security mechanisms, revealing compound vulnerabilities where the loss of control coincides with diminished situational awareness.

1. Introduction

The integration of cyber and physical components in modern power grids has fundamentally transformed grid management. This is enabled by technologies such as Wide Area Monitoring Systems (WAMS), automated demand response, and real-time optimization (Ashok et al., 2017; Adhikari et al., 2017). As a result, such advancements allow for sub-second control decisions, dynamic topology reconfiguration, and predictive maintenance capabilities that were not achievable in legacy systems. However, this integration introduces new vulnerabilities in industrial control environments exemplified by the Ukraine attacks in December 2015 and February 2016 (Lee et al., 2016). Cyber-Physical Systems (CPSs) bring together components from the physical and cyber world. However, an attacker can exploit the complex interactions between these different components. This could lead to serious consequences, such as power outages, equipment damage, and substantial financial loss.

A key interest in power grid CPSs is to maintain controllability and observability under cyber-attacks. The controllability refers to the system's ability to reach desired states through appropriate control inputs, while observability involves accurately monitoring and interpreting system states (Kalman, 1968; Kreindler and Sarachik, 1964). Conventional IT controls such as encryption and authentication mitigate unauthorized access and tampering. However, they fail to quantify or reveal how successful intrusions affect the system's ability to maintain stable and reliable operation. In particular, these measures cannot account for how particular cyber attacks impact controllability and observability, which are key properties that define whether operators can accurately assess system states and take corrective control actions. This gap leaves system operators with limited tools for real-time assessment of cyber threats' effects on power systems, crucial for implementing effective countermeasures.

* Corresponding author.

E-mail address: nsaxena@ieee.org (N. Saxena).<https://doi.org/10.1016/j.cose.2026.105156>

Received 13 October 2025; Received in revised form 3 September 2026; Accepted 5 September 2026

Available online 22 September 2026

0167-4048/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

The security landscape of the modern power grid is becoming increasingly complex due to the proliferation of Internet-of-Things (IoT) devices and Distributed Energy Resources (DERs) (Bedi et al., 2018). Key components such as smart meters and smart inverters are now integrated into the grid to enable automated grid-support functions and advanced control schemes. While these advancements boost grid operations, they simultaneously introduce new and significant cyber attack vectors. Cyber attacks on these devices can obscure the true state of the system or mislead operators, severely impacting both controllability and observability.

Prior research has focused on isolated aspects of cyber-physical security, such as Intrusion Detection Systems (IDSs) that process network traffic or physical data for situational awareness. While these aspects are essential, they lack a quantitative framework for understanding the interplay between cyber threats and CPS dynamic properties. For instance, Li et al. (2017) highlight conditions under which a system can remain controllable and observable under adversarial attacks but do not provide metrics quantifying these impacts on the electric grid. Similarly, Jacobs et al. (2020) propose a graph-theoretic approach to achieving cyber-physical observability but employ binary observability assessment that cannot capture the progressive degradation we observe in real attacks. Their work does not quantify how Indicators of Compromise (IoCs) translate into measurable reductions in controllability and observability.

To bridge these gaps, this research aims to develop metrics that quantitatively assess the impact of cyber-attacks compromise on the controllability and observability of cyber-physical power systems. Our novel approach combines cyber security indicators with power system dynamics, providing a holistic framework for threat assessment. In this paper, therefore, we focus on how each indicator in the attack chain affects these system properties with the goal of enhancing situational awareness during ongoing attacks and improving operators' ability to respond effectively by answering the following research questions:

RQ1: How can we develop metrics that accurately quantify the IoC-driven impact on the controllability and observability of cyber-physical power systems, thereby enabling operators to better understand and respond to ongoing cyber threats?

RQ2: How can we quantify the progressive degradation of system controllability and observability under coordinated cyber-physical attacks, and identify the most critical system elements whose compromise leads to maximal operational disruption?

In summary, we make the following contributions.

- We develop a novel set of quantitative indices, including System Control Reach Coverage (SCRC), Structural Controllability Index (SCI), Structural Observability Margin (SOM), and Substation Observability Score (SOSub). Together, these indices capture the degradation of controllability and observability under cyber-physical attacks. They integrate control-theoretic sensitivity analysis with graph-theoretic and measurement-based observability formulations.
- We introduce a systematic methodology for mapping IoCs to operational impacts on power system dynamics. By aligning IoCs with MITRE ATT&CK techniques and associating them with index degradation patterns, we enable real-time threat assessment grounded in system behavior.
- We present an integrated cyber-physical security co-simulator tool to model and evaluate the impact of three attack scenarios: False Command Injection (FCI), Load Redistribution (LR),¹

and False Data Injection (FDI). This setup captures the interdependencies between cyber-layer intrusions and physical-layer impact.

- Through extensive simulation, we demonstrate how coordinated attacks induce cascading degradation in both control authority and measurement integrity. Our results reveal that even localized attacks can lead to system-wide observability fragmentation, voltage violations, and loss of control reach, which remain undetected by conventional security mechanisms.

2. Related work

There are only a few works available in the literature assessing controllability and observability from cyber-attacks points of view. The importance of situational awareness in cyber threats cannot be overstated (Asiri et al., 2025). Kholidy discusses how attackers often use automated tools to exploit vulnerabilities, necessitating a robust situational awareness framework for detecting and responding to evolving threats (Kholidy, 2021). This aligns with the findings of Nafees et al. (2023), who similarly advocate for integrating cyber-physical situational awareness metrics with power system indicators to improve security. However, these works primarily focus on detection frameworks without quantifying how cyber threats degrade controllability and observability metrics.

To accurately assess the security and resilience of power systems, it is crucial to evaluate how cyber threats affect the controllability and observability of the integrated cyber-physical system. This assessment should include a comprehensive evaluation of how cyber-attack compromises impact both the communication networks and the physical power infrastructure, rather than examining them separately. Recent security-focused studies have begun addressing this need. Abbaspour et al. (2019) examine resilient control design under false data injection attacks, using Luenberger observers combined with neural networks for anomaly detection. While their approach demonstrates how False Data Injection (FDI) attacks degrade system observability through compromised measurements, it provides detection accuracy metrics rather than continuous quantification of observability degradation. Similarly, another work (Syrmakesis and Hatziaargyriou, 2024) categorize cyber resilience methods against FDI attacks but primarily offers qualitative frameworks rather than quantitative controllability/observability metrics under varying threat intensities.

Other security-aware approaches include Taha et al. (2016), who discuss risk mitigation strategies that maintain observability through secure measurements, particularly in scenarios where data integrity is compromised. Additionally, other works have shown that false data injection attacks can significantly undermine the observability of power systems, requiring robust countermeasures to identify and mitigate such vulnerabilities (Bae et al., 2017; Xie et al., 2021). However, these works focus on binary assessment (attack/no-attack) rather than quantifying the progressive degradation of system controllability and observability as attacks evolve.

In contrast, traditional controllability and observability studies assume benign operating conditions without considering cyber threats. These studies define observability and controllability metrics that are tailored to power system oscillations. Zhao et al. (2023) introduce Mode Comprehensive Observability Metric (MCOM) and Mode Comprehensive Controllability Metric (MCCM) using Frobenius norm-based quantification, providing sophisticated mathematical frameworks for oscillation analysis. These metrics play a crucial role in measuring the system's ability to effectively respond to disturbances and uphold stability across different scenarios but assume perfect measurements and no adversarial interference. Similarly, Sun et al. (2022) analyze robust structural controllability using zero-forcing set theory and graph-theoretic approaches. They provide only binary structural analysis without considering measurement integrity or communication vulnerabilities.

¹ Note that while Load Redistribution attacks are often classified as a subset of FDI in the literature, we distinguish between attacks that manipulate control setpoints to cause actual physical load changes (LR) and those that only corrupt measurement data without altering the physical system state (FDI), as each produces distinct operational impacts requiring different mitigation strategies.

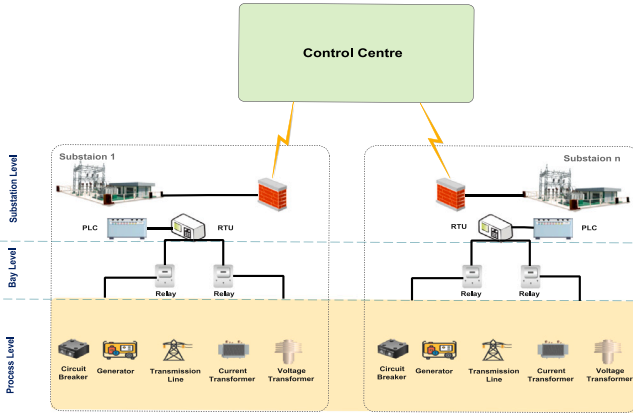


Fig. 1. Power system model with a control center and multiple substations.

Traditional observability analysis extensively focuses on Weighted Least Squares (WLS) state estimation and optimal Phasor Measurement Unit (PMU) placement using Integer Linear Programming (ILP) approaches (Wang et al., 2019; Raju and Kumar, 2016). While these methods optimize system performance, they assume measurement reliability and do not account for adversarial data manipulation. The System Observability Redundancy Index (SORI) provides reliability assessment but lacks quantitative metrics for observability degradation under cyber threats (Awon et al., 2018).

Our work addresses this critical gap by developing IoC-based quantitative metrics that provide continuous assessment of controllability and observability degradation under cyber-attacks. Unlike existing binary detection methods or traditional non-security-aware frameworks, our approach quantifies how specific IoCs progressively degrade system controllability and observability indexes and metrics in order to help operators to assess system degradation in real-time as attacks evolve.

3. System and attack model

In this section, we explain the details of the system and attack model for this work.

3.1. System model

We present a cyber-physical power transmission system designed to monitor and control electrical power flow across a network of transmission substations. Fig. 1 illustrates the system, which consists of three primary layers: a centralized control center, an intermediate communication network, and multiple transmission substations.

The control center serves as the operational hub, housing the SCADA master station that collects telemetry data and issues control commands to field devices. The SCADA system integrates with a Human-Machine Interface (HMI) to provide operators with real-time visualization and control capabilities of the power grid status. The communication infrastructure provides the data exchange layer between the control center and substations, utilizing industrial protocols for command and data transmission. This network layer incorporates security mechanisms, including firewalls and intrusion detection systems, to protect against cyber threats.

Each substation operates as a local monitoring and control node within a hierarchical framework. At the highest level, a Programmable Logic Controller (PLC) serves as the substation automation controller, aggregating data from downstream devices. At the bay level, Remote Terminal Units (RTUs) and protective relays collect system status information and supervise operations. These components interface with the process level, comprising physical devices such as circuit breakers, generators, and transmission lines. Circuit breakers operate as

binary-state components, while generators and line flows are monitored through analogue signals. Protective relays control circuit breakers to isolate faulty circuits, transmitting status data through the RTUs to the substation automation controller for further processing.

Within the control center, the SCADA master continuously polls substation controllers to maintain a comprehensive view of the grid's status, including circuit breaker positions, generator outputs, and transmission line loadings. The system operates on industrial communication protocols, with Modbus being a prevalent protocol for master-slave communication between the control center and field devices. In a real-world scenario, this communication takes place over either an insecure network or a Virtual Private Network (VPN), creating opportunities for adversaries to manipulate data at the substation level. The control center implements protection logic through its control instructions, automatically triggering circuit breakers when faults are detected, thereby enhancing system reliability.

Steady-State Control Loop Model: We model supervisory control as a steady-state setpoint-to-state loop. Let N denote the number of buses. Let $x := [\theta^T V^T]^T \in \mathbb{R}^{n_x}$ denote the power-flow state, where $\theta \in \mathbb{R}^N$ and $V \in \mathbb{R}^N$ are the bus voltage angles and magnitudes, respectively, and $n_x = 2N - 1$ after fixing a reference angle. Let $u = [P_{\text{gen},1}, P_{\text{gen},2}, \dots, P_{\text{gen},n_g}]^T$ denote the vector of generator active-power setpoints, where n_g is the number of generators and $j \in \{1, \dots, n_g\}$ indexes generators. Let d denote exogenous demand (active and reactive loads). The steady-state operating point is defined by the AC power-flow equations, written compactly as $g(x, u, d) = 0$, where $g(\cdot)$ stacks the nodal active- and reactive-power balance constraints. Let $P \in \mathbb{R}^N$ and $Q \in \mathbb{R}^N$ denote the vectors of nodal net active- and reactive-power injections, respectively. Around a converged solution (θ^*, V^*) , small perturbations satisfy:

$$\begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix} = J \begin{bmatrix} \Delta \theta \\ \Delta V \end{bmatrix}, \quad J = \begin{bmatrix} \partial P / \partial \theta & \partial P / \partial V \\ \partial Q / \partial \theta & \partial Q / \partial V \end{bmatrix}_{(\theta^*, V^*)}, \quad \begin{bmatrix} \Delta \theta \\ \Delta V \end{bmatrix} = J^{-1} \begin{bmatrix} \Delta P \\ \Delta Q \end{bmatrix}.$$

Here, ΔP and ΔQ represent nodal net-injection perturbations at the operating point. Under supervisory setpoint actions with fixed demand ($\Delta d = 0$), the perturbation is induced by Δu through active-power injection changes at generator buses, so ΔP is the bus-wise injection change associated with Δu (with zero entries at non-generator buses), and ΔQ is zero for the setpoint action itself. This establishes an explicit analytical mapping from setpoint changes to $(\Delta \theta, \Delta V)$ via the inverse Jacobian.

A setpoint change Δu_j at generator j induces Jacobian-based sensitivities $\partial V_i / \partial P_{\text{gen},j}$ for bus voltages and $\partial F_e / \partial P_{\text{gen},j}$ for branch real-power flows $F_e = f_e(\theta, V)$; these quantify the steady-state control influence of u on the controlled variables (V, F) .

Actuation and communication are represented at the supervisory layer by authenticated dispatch of u to field devices. Cyber events that remove, delay, or alter specific actuation paths are modeled as reductions of the effective control set and modifications to the corresponding sensitivities; no dynamic frequency-control mechanisms are invoked in this formulation.

Measurements follow $z = h(x) + v$, where $z \in \mathbb{R}^m$ is the measurement vector, $x \in \mathbb{R}^{n_x}$ is the system state vector containing bus voltage magnitudes and phase angles, $h(\cdot)$ is the nonlinear mapping that relates system states to expected measurements based on network topology and parameters, and $v \in \mathbb{R}^m$ is the measurement noise vector. The linearized estimator uses $H = \partial h / \partial x$ evaluated at (θ^*, V^*) . Observability at the operating point requires $\text{rank}(H) = n_x$ after fixing a reference angle; if the network partitions into I observable islands, each island requires its own reference to ensure a well-posed estimation problem.

3.2. Threat model

Adversary Profile and Access Vector. The adversary attempts to disrupt the stability and operation of a cyber-physical smart grid system. Their goal is to compromise the system's ability to accurately monitor

and control power distribution by exploiting vulnerabilities in communication channels and control mechanisms. Operating externally, they gain unauthorized access through methods such as exploiting weak authentication, breaching VPNs, or intercepting data exchanges between the control center and substations. Once inside, the attacker can issue false commands, alter operational settings, or manipulate measurement data to mislead operators and destabilize the power grid.

Adversary's Objectives and Capabilities. We assume the adversary can carry out attacks with different levels of sophistication. The attacker can gather information about the system's architecture, identify critical components, and analyze communication patterns by leveraging network reconnaissance techniques such as passive packet sniffing and active scanning. This information can be used later to map control system assets and formulate targeted attack strategies. Moreover, the adversary can take advantage of the Modbus/TCP protocol, which lacks authentication mechanisms and encryption. They can modify control commands as they wish between the control center and substations (López-Morales et al., 2024). We can envision that an adversary has high technical capabilities that enable them to craft control messages that resemble legitimate SCADA commands, making them difficult to detect.

Adversary's Knowledge and Attack Progression. The adversary operates with evolving knowledge of both network and power system topologies, which expands progressively during attack execution. Beginning with limited visibility, they exploit initial access to conduct reconnaissance, map system architecture, and escalate privileges. This phase reveals critical operational data, including SCADA command structures and PLC addressing schemes essential for logic implementation. Accessibility evolves from external network penetration to internal system control, ultimately enabling the injection of malicious commands and modification of data using legitimate system credentials. These actions are blended with normal operations, delaying detection until physical anomalies emerge in the power system.

We selected three attack scenarios based on their documented effectiveness in real-world incidents and their distinct impacts on system controllability and observability. These attacks represent critical vulnerabilities identified in modern power grids that can cascade into widespread physical failures. In the first attack scenario (a), the adversary targets the direct control pathway by sending unauthorized commands to open circuit breakers at critical generators. This attack vector gained prominence following the 2015 Ukraine incident where adversaries manually opened breakers, causing immediate loss of generation capacity that affected 225,000 customers (Liang et al., 2016). If this attack is successful, potential impacts include power imbalances triggering under-frequency load shedding and voltage instability propagating through interconnected buses (Ju et al., 2017). Additionally, there is a risk of potential generator damage from out-of-step conditions if re-synchronization is attempted without proper procedures.

In the second attack scenario (b), the adversary conducts load manipulation by altering setpoints within the SCADA database, causing artificial load variations that stress transmission lines and generators beyond operational limits. Unlike general FDI attacks that only falsify measurement data, such an attack represents a structured form of FDI that maintains total system load while creating localized stress concentrations that can cause line outages or even cascading failures (Xiang et al., 2016). In the third attack scenario (c), the adversary executes false data injection by altering measurement values from substation PLCs. The adversary aims to corrupt the observability foundation required for secure operation by presenting misleading system states to operators while concealing developing instabilities in the physical power system. An attack of this nature directly threatens the control and safety mechanisms of the system. Operators might take incorrect actions, like curtailing generation when it is desperately needed, due to receiving false system data (Abbaspour et al., 2019). Similarly, protection systems, relying on corrupted information, could either trip unnecessarily or fail to engage during genuine threats.

4. Our approach

In this section, we present our approach as depicted in Fig. 2, discuss specific use cases for this work, identify potential indicators, and outline how our metrics cover different stages of an attack.

Controllability in CPS refers to the ability of the system operator to influence the state of the power grid through control actions, under the joint constraints imposed by both physical dynamics and the cyber-layer infrastructure. While classical control theory uses the Kalman rank condition to assess whether a system's state can be driven to any desired configuration through a sequence of valid control inputs (Guo et al., 2017), this assumes flawless execution of commands. In real-world CPS operations, controllability is additionally constrained by the security, reliability, and latency of communication networks and actuators. Cyber threats, including Denial-of-Service (DoS), malicious command injection, and unauthorized access, can degrade or delay control actions. Thus, a system that is theoretically controllable may become practically uncontrollable in adversarial environments, necessitating a definition that integrates both physical and cyber considerations (Li et al., 2024; Ge et al., 2023).

Using the state vector $\mathbf{x}$ and the generator control vector $\mathbf{u}$ as defined in Section 3.1, we formalize steady-state controllability through sensitivity analysis. We define the set of controlled quantities $Q = \{q_1, q_2, \dots, q_{n+m}\}$, where the first n elements correspond to bus voltage magnitudes V_i for $i \in N = \{1, 2, \dots, n\}$ (the set of buses), and the remaining m elements correspond to branch power flows F_e for $e \in E = \{1, 2, \dots, m\}$ (the set of branches). For notational simplicity, we use a unified index $i \in \{1, 2, \dots, n+m\}$ to refer to any controlled quantity q_i . The generator set is denoted $G = \{1, 2, \dots, n_g\}$, with $j \in G$ indexing individual generators.

The relationship between generator setpoints and these controlled quantities is characterized through sensitivities evaluated at the steady-state operating point (θ^*, V^*) . The sensitivities of the controlled quantities with respect to the generator setpoints are expressed as $\partial q_i / \partial u_j|_{(\theta^*, V^*)}$, where u_j denotes the j th element of the control vector $\mathbf{u}$ (i.e., $P_{\text{gen},j}$). Given fixed thresholds $\epsilon_V = 0.01$ pu/MW for bus voltages and $\epsilon_F = 0.05$ MW/MW for branch flows, we define the indicator variable:

$$\delta_i = \begin{cases} 1, & \text{if } \max_j \left| \frac{\partial q_i}{\partial u_j} \right| > \epsilon, \\ 0, & \text{otherwise.} \end{cases}$$

where ϵ equals ϵ_V for voltage quantities and ϵ_F for flow quantities. In words, this indicator determines whether controlled quantity i can be sufficiently influenced by at least one generator. If the maximum sensitivity across all generators exceeds the threshold, the element is deemed controllable ($\delta_i = 1$); otherwise, it lacks adequate control authority ($\delta_i = 0$). Finally, the steady-state control coverage is defined as

$$\kappa = \frac{\sum_i w_i \delta_i}{\sum_i w_i},$$

which quantifies the fraction of electrically weighted elements that remain responsive to admissible setpoint actions under cyber and communication constraints, where w_i represents the electrical importance of element i .

It is important to distinguish our approach from traditional controllability metrics such as the Kalman rank condition, which are designed for dynamical systems with state-space representations $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$. In contrast, supervisory control of power systems is primarily concerned with maintaining steady-state equilibrium points. Our framework evaluates whether generator setpoints can influence bus voltages and line flows to keep the system within operational limits. By using sensitivities derived from algebraic power-flow equations, this approach directly captures steady-state control authority and provides a practical basis for the controllability assessment framework presented in this paper.

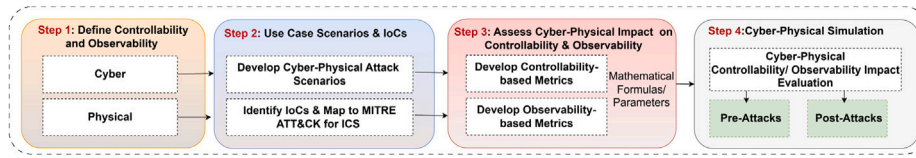


Fig. 2. Overview of our approach.

Observability, on the other hand, refers to the ability to infer the internal state of the power system from available sensor measurements (Cheng et al., 2024), accounting not only for the spatial distribution of sensors but also for the integrity and security of the data they provide. In classical control theory, observability is determined by the ability to reconstruct state variables from output measurements, typically assessed using the rank condition of the observability matrix (Moore, 1981). In traditional power systems, this concept is operationalized through the strategic placement of Phasor Measurement Units (PMUs) and SCADA sensors. However, in a cyber-physical environment, true observability is further compromised by cyber threats such as false data injection, communication delays, and sensor tampering. These threats can lead to incomplete or corrupted state information, meaning that even a mathematically observable system may become partially or fully unobservable in practice. Therefore, the definition must integrate both the sufficiency of measurement data and the trustworthiness of sensor outputs under adversarial conditions, with considerations for dynamic degradation and recovery of measurement integrity. The measurement model in Section 3.1 is augmented with an attack vector as follows:

$$z = h(x) + v + a$$

where $a \in \mathbb{R}^m$ represents the attack vector containing false data injected by adversaries. The linearized estimator uses the Jacobian matrix $H = \partial h / \partial x$ evaluated at the operating point (θ^*, V^*) .

After fixing a reference angle, global observability at the operating point requires $\text{rank}(H) = n_x$, with $n_x = 2N - 1$ for an N -bus system. Topological changes or measurement loss can partition the network into I observable islands, where I is the number of such islands, and each island requires an independent angular reference. Let m_{eff} denote the effective number of usable measurements after bad-data detection and validation, that is, the count of measurements deemed trustworthy once corrupted or suspicious data are removed. The minimum usable-measurement count required for full observability across I islands is $n_x + (I - 1)$. For substation-level assessment, consider a substation j with local state variables (θ_j, V_j) and corresponding sub-Jacobian matrix H_j . H_j is formed by selecting from H the columns associated with these states. Since each substation has $n_j = 2$ state variables $(\theta_j \text{ and } V_j)$, local observability requires $\text{rank}(H_j) = n_j$. The normalized ratio $\frac{\text{rank}(H_j)}{n_j} \in [0, 1]$ serves as a measure of the local state estimation capability, where a value of 1 indicates full observability of the substation states.

4.1. Identify IoCs based on use case scenarios

An adversary can perform various cyber attacks targeting critical components of the power system through various attack vectors. These scenarios demonstrate different approaches where attackers, with varying levels of system knowledge, can compromise system operations through the control center system and PLC devices. The attacks range from direct malicious command injection to subtle data manipulation, each with distinct impacts on system controllability and observability. We discuss three specific scenarios in the remainder of this section. The scenarios studied here are implemented as malicious modifications to actuation and measurement pathways and are evaluated through their operational consequences on controllability and observability. The proposed indices quantify the resulting degradation in control authority and measurement integrity. They do not, by themselves, attribute

whether the initiating cause is a remote cyber intrusion or an insider issuing equivalent actions through authorized interfaces. In operational deployments, provenance is established by correlating the observed degradation with cyber evidence available from the OT infrastructure. Our work focuses on impact quantification once the affected commands or measurements have been identified.

Use Case 1: An adversary aims to disrupt the power system by performing a false command injection attack targeting circuit breaker operations at critical generators. The attacker has gained unauthorized access to the network but lacks complete system knowledge. If an adversary targets a command from a malicious source, such as a fabricated command to detach a generator, the IDS will be able to detect the command and prevent its execution on the power system. By exploiting vulnerabilities in the communication protocols between the master PLC and the outstation device, the adversary sends malicious commands to open the circuit breakers of two critical generators in the system. Prior to issuing these commands, the attacker identifies critical communication paths and asset dependencies by mapping the SCADA environment. The attacker has two options to achieve this. First, they can sniff the communication and access detailed system configuration files from the targeted controller, which provide insights into address schemes and control logic. This enables the attacker to map observed network traffic to specific devices and functions. Second, the adversary can compromise the SCADA system to gain full control and complete knowledge.

Use Case Steps. ❶ The attack begins when the adversary gains access to the target network through a VPN, exploiting weak authentication or using stolen credentials. Once inside, ❷ they harvest credentials from the engineering workstation and map the network to identify critical assets, including SCADA systems and PLCs. With a clear understanding of system dependencies established, ❸ the attacker sends a legitimate-looking Modbus/TCP packet to open the circuit breaker. ❹ The master PLC executes the command, disconnecting a targeted generator from the power grid. ❺ Then, the control center logs the status change without immediate anomaly detection. ❻ The power system subsequently exhibits power flow imbalances and system instability. Moreover, voltage instabilities rapidly spread across the network. This forces other generators into sudden, compensatory adjustments and exacerbates operational imbalances within the power system.

Use Case 2: An adversary seeks to destabilize the power system by altering load distribution across selected substations to cause operational stress without changing the total system load. The attacker has gained unauthorized access to the HMI and modified load set points without operator awareness.

Use Case Steps. Initially, the attacker follows the same preliminary steps ❶ and ❷ as described in use case 1. ❸ Next, they identify and monitor load set points within the SCADA database, collecting critical operational data to prepare for the manipulation. Their goal here is to select specific load buses to increase or decrease demand in a way that shifts power flows toward less redundant parts of the network. ❹ Then, using the acquired insights, the attacker modifies the load set points and propagates the altered values to the outstation PLCs. ❺ As a result, generators and transmission lines experience stress due to unexpected load variations. Such targeted adjustments incrementally redirect power flows toward constrained network segments. The system operates beyond its stability limits and accumulates stress, which can lead to localized overload conditions and eventual equipment failure.

Use Case 3: An adversary executes a false data injection attack by compromising sensor measurements to deceive operators and disrupt power

Table 1
Summary of IoCs and mapped MITRE ATT&CK techniques used in attack scenarios.

Attack scenario	Layer	Indicators	MITRE ATT&CK technique	Technique ID
Scenario 1	Cyber	•Unauthorized access attempts detected on Engineer workstation (<i>IoC-A₁</i>)	Exploitation of Remote Services	T0866
		•Unauthorized login with valid credentials on SCADA (<i>IoC-A₂</i>)	Valid Accounts	T0859
		•Unusual port scanning or network reconnaissance activities (<i>IoC-A₃</i>)	Network Connection Enumeration, Network Sniffing	T0840, T0842
		•SCADA logs indicate a command to open the circuit breaker that was not initiated by an authorized operator (<i>IoC-A₄</i>)	Unauthorized Command Message, Modify Parameter	T0855, T0836
		•Increased frequency of Modbus/TCP commands in a short time frame (<i>IoC-A₅</i>)	Denial of Control	T0813
		•Modbus function codes that are not typically used for regular operations or unexpected (<i>IoC-A₆</i>)	Unauthorized Command Message	T0855
	Physical	•Sudden loss of generation capacity from the disconnected generator (<i>IoC-A₇</i>)	Loss of Availability , Manipulation of Control	T0826, T0831
		•Voltage and frequency deviations exceeding normal operational limits (<i>IoC-A₈</i>)	Manipulation of Control, Modify Parameter	T0831, T0836
		•Generators unexpectedly ramp up to compensate for the lost capacity (<i>IoC-A₉</i>)	Manipulation of Control	T0831
Scenario 2	Cyber	•Unauthorized modifications to setpoint values in the control center (<i>IoC-B₁</i>)	Manipulation of Control, Modify Parameter	T0831, T0836
		•Increased data traffic between the control center and field devices (<i>IoC-B₂</i>)	Denial of Service	T0814
		•Abnormal frequency of control commands issued from the control center to outstation devices (<i>IoC-B₃</i>)	Denial of Control	T0813
	Physical	•Imbalanced load distribution across grid areas (<i>IoC-B₄</i>)	Manipulation of Control	T0831
		•Transmission line overloads, risking physical damage (<i>IoC-B₅</i>)	Manipulation of Control	T0831
		•Anomalous power flow patterns inconsistent with operational norms (<i>IoC-B₆</i>)	Manipulation of Control	T0831
Scenario 3	Cyber	•Anomalous Modbus/TCP traffic with unexpected function codes (<i>IoC-C₁</i>)	Unauthorized Command Message	T0855
		•Discrepancies between redundant sensor measurements and expected values (<i>IoC-C₂</i>)	Manipulation of View	T0832
		•Inconsistent data timestamps across devices (<i>IoC-C₃</i>)	Spoof Reporting Message	T0856
		•Unusual delays or losses in data transmission between devices (<i>IoC-C₄</i>)	Denial of Service, Loss of View, Impair Process Control	T0814, T0829
	Physical	•Anomalies in power flow measurements inconsistent with operational conditions (<i>IoC-C₅</i>)	Manipulation of Control	T0831
		•Frequency deviations or oscillatory behavior in the grid (<i>IoC-C₆</i>)	Manipulation of Control	T0831
		•Overheating or mechanical stress in transformers or lines (<i>IoC-C₇</i>)	Manipulation of Control	T0831

system stability. Targeting outstation PLCs, the attacker injects falsified data into the control stream, undermining system observability and control accuracy.

Use Case Steps. ❶ and ❷ same as in use case 1. ❸ The attacker then conducts network enumeration to identify outstation PLCs responsible for transmitting generator output and substation voltage data to the SCADA system. ❹ Leveraging this position, the attacker injects falsified sensor readings and control signals into the SCADA stream, which is designed to reflect normal operating conditions even when critical components are offline. ❺ The operator relies on the manipulated SCADA display and remains unaware of the generator's actual offline state. As a consequence, no corrective actions occur. The system then operates under false assumptions. This allows underlying instabilities to persist and potentially escalate. Table 1 provides a comprehensive mapping of the cyber and physical indicators identified in each attack scenario to their corresponding MITRE ATT&CK technique identifiers.

4.2. Metric definitions and computation

In this work, several indices have been developed to quantify the degradation of controllability and observability in cyber-physical

power systems under attack. Such indices are composed of metrics derived from power system principles as sub-components. By using these indices, the continuous quantification of system degradation has been meaningful toward providing operators with actionable thresholds for intervention, transforming previously undetectable progressive attacks into measurable degradation patterns with specific response triggers before cascading failures occur. All indices are derived from steady-state linearizations of the AC power-flow equations and the measurement model. Specifically, we use the control-to-state sensitivity matrix $\Sigma = \left[\frac{\partial q_i}{\partial u_j} \right]$, which captures the influence of generator setpoints on system states, and the measurement Jacobian $H = \frac{\partial h}{\partial x}$, which characterizes the relationship between system states and sensor outputs. These matrices form the analytical foundation for the metrics defined below. To ensure consistent interpretation, the indices here are steady-state, operational surrogates of controllability and observability rather than dynamic state-space tests.

At the solved operating point, ground truth denotes the ability to realize a minimum, bounded change in a designated quantity $q_i \in \{V_i, F_e\}$ without violating steady-state network constraints. Let u denote the admissible set of generator set-point changes defined by operating

limits, and use the sensitivity matrix Σ . For element i , we implement a local operational criterion by testing $\max_{j \in \mathcal{U}} |\partial q_i / \partial u_j|$ against materially relevant thresholds and then aggregating element-level decisions with the electrical-importance weights used in this work.

4.2.1. System Control Reach Coverage (SCRC)

In power systems, the fundamental question during cyber attacks is: can operators still control critical infrastructure? SCRC index addresses this by measuring how many system elements remain responsive to available control inputs, providing a quantitative answer to whether sufficient control authority exists to maintain stability. SCRC quantifies the extent to which control inputs in a power system can effectively influence critical system elements. We evaluate steady-state control reach from the sensitivities $\frac{\partial q_i}{\partial u_j}$ in Σ , aggregating element-wise influence tests into a normalized, electrically weighted coverage. Formally, SCRC is defined as:

$$\text{SCRC} = \frac{\sum_{i \in N} w_i \delta_i}{\sum_{i \in N} w_i}$$

where:

- N represents the set of system elements (buses and branches)
- w_i is the electrical significance weight of element i
- δ_i is the controllability indicator for element i

For each element $i \in N$, the controllability indicator δ_i is determined by:

$$\delta_i = \begin{cases} 1, & \text{if } \max_{j \in G} \left| \frac{\partial x_i}{\partial u_j} \right| > \epsilon \\ 0, & \text{otherwise} \end{cases}$$

where:

- G denotes the set of generator buses
- u_j represents the control input from generator j
- x_i represents the state variable of element i (voltage magnitude V_i for buses or real power flow F_e for branches)
- ϵ denotes the sensitivity threshold, specifically $\epsilon_V = 0.01$ pu/MW for voltage sensitivities and $\epsilon_F = 0.05$ MW/MW for flow sensitivities. Where MW/MW represent a dimensionless ratio that quantifies the change in branch real-power flow (MW) per unit change in generator real-power injection (MW). These thresholds are informed by industry practices and technical literature (PowerWorld Corporation, 2025; Arefifar et al., 2016; Xu et al., 2023; Davis and Overbye, 2010).

The electrical significance weight w_i captures the operational importance of element i and can be formulated as:

$$w_i = f(\text{Load}_i, \text{Centrality}_i, \text{VoltageLevel}_i)$$

Common weighting schemes include:

1. **Uniform weighting:** $w_i = 1$
2. **Load-based weighting:** $w_i = \max(\text{Load}_i, 0.1)$ for bus $i \in \mathcal{N}$
3. **Flow-based weighting:** $w_i = \max(|F_e|, 0.1)$ for branch $e \in \mathcal{E}$

The sensitivity $\frac{\partial x_i}{\partial u_j}$ quantifies the response of state variable x_i to a marginal change in control input u . For each system element, we compute:

- **For buses:** $\frac{\partial V_i}{\partial P_{gen}}$ – voltage sensitivity to generator output change
- **For branches:** $\frac{\partial F_e}{\partial P_{gen}}$ – flow sensitivity to generator output change

The maximum absolute sensitivity across all available control inputs determines controllability:

$$\max_{j \in G} \left| \frac{\partial x_i}{\partial u_j} \right|$$

The SCRC yields values ranging from 0 to 1. When the value is close to 1, it reflects a system with high controllability. In such cases, most of the elements that carry significant weight respond well to the available control actions. On the other hand, a value near 0 indicates poor controllability. This suggests that even the most critical elements in the system exhibit little or no response to control inputs.

4.2.2. Structural Controllability Index (SCI)

While SCRC reveals whether control inputs can influence system elements, it does not capture whether the control paths themselves are viable under stressed conditions. SCI fills this gap by evaluating both the existence of control paths and their operational quality, recognizing that cyber attacks often leave systems connected but operating beyond safe limits. The SCI is designed to quantify the extent to which critical control points in a power system remain interconnected and operate within desired power flow parameters. It reflects the system's ability to maintain controllability by measuring the impact of such indicators on system connectivity and power flow integrity. To achieve this, we designed the index by integrating two critical dimensions: structural connectivity and operational stability. In structural connectivity, we examine whether critical control points remain connected through viable and redundant pathways. On the other hand, we assess whether these pathways are functioning effectively by maintaining stable flows and voltages through the operational stability component. In a steady-state loop, SCI reflects the conditions under which setpoint perturbations propagate effectively through the network, with congestion and voltage violations reducing the practical effect of sensitivities in Σ . The following subsections describe each component in detail before presenting the combined SCI formulation.

Structural Connectivity Component: For the structural connectivity component S , we need to evaluate the robustness of connections between critical control points C . The set C represents the critical control points whose interconnection is essential for maintaining system controllability. These points are identified based on their operational importance and vulnerability. Thus, We can express C as:

$$C = C_g \cup C_l \cup C_n \cup C_v$$

where C_g represents generator buses, C_l denotes high-load buses (top 20% by demand), C_n comprises high-centrality buses (80th percentile of betweenness centrality), and C_v includes voltage-critical buses (voltage magnitude below 0.95 p.u.). These critical thresholds are defined empirically based on referential scenarios as in Yan et al. (2014), Eppstein and Hines (2012). We model the power system as an undirected graph:

$$\mathcal{G} = (\mathcal{V}, \mathcal{E})$$

where $\mathcal{V} = \{1, 2, \dots, n\}$ denotes the set of buses and $\mathcal{E} = \{(i, j) | i, j \in \mathcal{V}, i \neq j\}$ comprises the set of transmission lines and transformers. Before evaluating connectivity, we assess the quality of each branch using the loading ratio:

$$\rho_e = \frac{\sqrt{(P_e^{\text{act}})^2 + (Q_e^{\text{act}})^2}}{\text{LineLimMVA}_e + \epsilon}$$

where P_e^{act} and Q_e^{act} are the branch's active (MW) and reactive (MVar) power flows, LineLimMVA_e is its MVA rating, and ϵ is a small positive constant included to prevent division by zero. We then assign a discrete quality score q_e based on ρ_e .

$$q_e = \begin{cases} 1.0, & \rho_e \leq 1.0 \quad \text{Normal operation} \\ 0.8, & 1.0 < \rho_e \leq 1.2 \quad \text{Moderate overloading} \\ 0.6, & 1.2 < \rho_e \leq 1.5 \quad \text{Heavily overloading} \\ 0.3, & \rho_e > 1.5 \quad \text{Critical overloading} \\ 0.0, & \text{branch out of service or disconnected} \end{cases}$$

For transformer branches with $\rho_e > 1.0$, an additional 5% penalty is applied: $q_e \leftarrow 0.95 \cdot q_e$, to reflect their more stringent operational constraints (Choi and Moon, 2009).

For each critical-bus pair $\{i, j\} \in C$, we need to assess the resilience of control pathways under adversarial conditions in which attackers strategically target high-value links. Traditional graph-theoretic methods typically use shortest path algorithms to minimize communication latency or Euclidean distance. However, in a CPS context under attack, these shortest paths often pass through critical bottlenecks. When such bottlenecks are compromised, all control between the nodes can be lost. In addition, power flow physics imposes branch thermal limits. This means that a viable path must not only exist but must also support sufficient capacity.

To address these constraints, we use disjoint widest paths instead of shortest paths. Power system control requires both redundancy and capacity. Disjoint paths provide alternative routes if primary ones are compromised, and path width reflects the actual power transfer capability needed during emergencies. To assess the robustness of control pathways, we consider not just a single route but up to K independent routes between critical points. These routes are required to be disjoint. This requirement ensures that if an outage or cyber-attack disrupts one pathway, alternative pathways remain available. Each route is evaluated based on its weakest link (*bottleneck quality*), which reflects the limiting capacity along that path.

The process first identifies the most reliable available pathway and evaluates its bottleneck quality. Then, it searches for additional non-overlapping pathways to capture redundancy. By repeating this process up to K times, the method quantifies both whether a connection exists and the diversity and strength of alternative routes. The resulting pairwise connectivity score, Q_{ij} , represents the maximum bottleneck quality across the K identified routes. This score provides a measure of both pathway redundancy and overall control delivery capacity. Algorithm 1 summarizes the formal procedure used to compute Q_{ij} .

Algorithm 1: Edge-Disjoint Widest Paths for Connectivity Score

Q_{ij}

Input: Graph $G = (\mathcal{V}, \mathcal{E})$, critical-bus pair $\{i, j\} \in C$, maximum paths K
Output: Pairwise connectivity score Q_{ij}

- 1 Annotate every branch $\ell \in \mathcal{E}$ with its quality q_ℓ ;
- 2 Initialize $G_0 = G$;
- 3 **for** $k = 1$ **to** K **do**
- 4 Compute the maximum-spanning tree MST_k of G_{k-1} using edge-weights q_ℓ ;
- 5 Let $\pi_k(i, j)$ be the unique path between i and j in MST_k ;
- 6 **if no path exists then**
- 7 **break**;
- 8 **end**
- 9 Define bottleneck quality $\beta_k = \min_{\ell \in \pi_k(i, j)} q_\ell$;
- 10 Remove one branch achieving β_k from G_{k-1} to form G_k ;
- 11 **end**
- 12 Set the pairwise connectivity score $Q_{ij} = \max_{1 \leq k \leq K} \beta_k$;
- 13 **return** Q_{ij} ;

Summing Q_{ij} across all critical pairs yields the **connectedness metric** N_{cc} to capture system-wide control capability rather than isolated path quality:

$$N_{cc} = \sum_{i, j \in C, i < j} Q_{ij}$$

The connectedness metric N_{cc} is then normalized by the total number of critical pairs, $N_{tc} = \binom{n_c}{2}$, where $n_c = |C|$, to compute the structural connectivity score S :

$$S = \frac{N_{cc}}{N_{tc}} \in [0, 1].$$

This score is bounded within $[0, 1]$, which ensures consistent interpretation across systems of varying sizes and enables percentage-based degradation assessment. Here, $S \rightarrow 1$ indicates robust structural redundancy, while $S \rightarrow 0$ signals fragmented critical-point connectivity.

Operational Stability Component: While the structural dimension assesses topology, operational stability is measured by deviations from

expected power flows. Modern power systems experience fluctuations due to load variations, renewable intermittency, or contingencies. However, excessive deviations compromise controllability (Bessa et al., 2019).

To quantify branch-level deviations from baseline power flows, each branch e is assigned a weight $w_e = |P_e^{\text{base}}| + \epsilon$, where $\epsilon > 0$ ensures numerical stability for lightly loaded branches.

The weighted mean deviation ΔP_{mean} aggregates discrepancies between actual (P_e^{act}) and baseline (P_e^{base}) flows:

$$\Delta P_{\text{mean}} = \frac{\sum_{e \in \mathcal{E}} w_e |P_e^{\text{act}} - P_e^{\text{base}}|}{\sum_{e \in \mathcal{E}} w_e}.$$

This metric prioritizes deviations in heavily loaded branches, as they pose greater risks to grid stability. The deviation factor D is then normalized against the total actual power flow, $P_{\text{total}} = \sum_e |P_e^{\text{act}}|$:

$$D = 1 - \frac{\Delta P_{\text{mean}}}{P_{\text{total}}}.$$

$D \rightarrow 1$ implies minimal deviation (operational consistency), while $D \rightarrow 0$ reflects significant divergence from baseline expectations. While S captures the presence of structural connectivity, it does not quantify the operational robustness of the resulting paths. To incorporate connection quality, we compute the average branch quality factor, denoted by:

$$Q_{\text{avg}} = \frac{1}{|\mathcal{E}|} \sum_{e \in \mathcal{E}} q_e$$

We then scale the structural score to obtain an adjusted connectivity measure.

$$S_{\text{adj}} = S \times Q_{\text{avg}}$$

This adjustment penalizes networks in which connectivity exists but is supported by branches with low quality factors. Under stressed operating conditions (e.g., Q_{avg} , nominally connected routes may be ineffective due to congestion, thermal constraints, or voltage-related limitations (Greenwood and Taylor, 2014). Consequently, S_{adj} better reflects the true controllability potential of the network under stressed conditions.

In parallel, voltage security is incorporated by penalizing widespread voltage violations. Let:

$$n_{\text{viol}} = \left| \{i \in \mathcal{V} \mid V_i < 0.90 \text{ p.u.} \vee V_i > 1.10 \text{ p.u.}\} \right|$$

be the number of buses outside the acceptable voltage band. We then set:

$$V_{\text{score}} = 1 - \frac{n_{\text{viol}}}{|\mathcal{V}|},$$

so that the operational component is diminished in proportion to the fraction of buses experiencing voltage excursions. Voltage violations directly impair controllability through protection system activations. Loads are disconnected when voltage drops below 0.90 p.u. due to undervoltage load shedding, while generators are tripped when voltage exceeds 1.10 p.u. due to overvoltage protection (Otomega and Van Cutsem, 2007). These actions progressively remove control nodes from the network, reducing the system's ability to respond to control inputs. By incorporating V_{score} , the SCI captures this loss of control resources caused by voltage-triggered disconnections, ensuring the index reflects actual control capability rather than theoretical connectivity.

The final index combines structural and operational metrics:

$$\text{SCI} = \alpha S_{\text{adj}} + (1 - \alpha) D V_{\text{score}}$$

where $\alpha \in [0, 1]$ is a tuning parameter that balances topological connectivity and operational stability. Higher values of α focus on the network structure, whereas lower values devote more attention to the power flow alignment with the desired power flows. This index captures not only the integrity of the most critical control paths but also the ability of the system to meet operational setpoints. The voltage score V_{score} serves

as a penalty factor that reduces the operational component when voltage violations are widespread to reflect the diminished controllability under such conditions.

The resulting index spans $[0, 1]$. Higher values imply that the power system retains more of its control structure, allowing operators to retain greater controllability over power flows. On the other hand, a lower SCI value implies that significant portions of the control structure have been compromised, and power flows deviate considerably from nominal conditions. Such a state could severely impair the controllability of the system thus destabilizing the grid under current load conditions.

4.2.3. Structural Observability Margin (SOM)

To quantify measurement redundancy and resilience against data loss or topological reconfigurations, we propose the SOM index. This index is particularly relevant in cyber-physical attack scenarios, where measurement loss and network fragmentation can progressively erode state estimation capabilities. By quantifying how close the system is to becoming unobservable, SOM supports timely operator intervention and enhances situational awareness. Unlike classical observability assessments, which yield a dichotomous classification (observable vs. unobservable) based on the full-rank condition of the measurement Jacobian matrix, SOM provides a continuous, dynamic index indicating the proximity of the system to the unobservability threshold. This margin jointly incorporates (i) the retained measurement count after bad-data elimination and (ii) the existence of observable islands, each necessitating an independent reference angle (slack bus). To our knowledge, no prior work has introduced a real-time, normalized observability metric that integrates both measurement validity screening and network decomposition into islands (Fotopoulou et al., 2022; Gou and Abur, 2000).

Formally, let:

- N denotes the number of buses, yielding a state vector of dimension $n_x = 2N - 1$ (assuming a single slack bus for angle reference),
- $\mathbf{H} \in \mathbb{R}^{m_{\text{eff}} \times n_x}$ represent the post-processing Weighted Least-Squares (WLS) measurement Jacobian with m_{eff} valid measurements retained after bad-data removal,
- I indicates the number of observable islands identified via topology-processing observability analysis.

We define the unscaled margin:

$$M = m_{\text{eff}} - [n_x + (I - 1)]$$

Here, M represents the surplus number of trustworthy measurements available beyond the minimum required for observability. A positive value of M indicates redundancy that can tolerate measurement losses or attacks, whereas a negative value implies a shortage of measurements, leaving parts of the system unobservable. The normalized SOM is then given by:

$$\text{SOM} = \frac{M}{m_{\text{eff}}} \in [-\infty, 1].$$

The just-observable boundary occurs at $m_{\text{eff}} = n_x - (I - 1)$, equivalently $\text{rank}(\mathbf{H}) = 2N - I$.

- $\text{SOM} = 0 \rightarrow$ System is just-observable (minimal measurements for full observability).
- $\text{SOM} > 0 \rightarrow$ System has measurement redundancy where SOM represents the fraction of surplus measurements.
- $\text{SOM} < 0 \rightarrow$ System is unobservable where $|\text{SOM}| \times m_{\text{eff}}$ equals the number of additional independent measurements needed.
- $\text{SOM} \rightarrow 1 \rightarrow$ Approaches maximum redundancy as $m_{\text{eff}} \gg n_x - (I - 1)$.

Beyond a simple binary pass/fail, SOM offers a granular, real-time assessment of observability robustness which empowers operators to monitor the progressive degradation or restoration of state estimation feasibility (Fotopoulou et al., 2022).

4.2.4. Substation Observability Score (SOSub)

While SOM evaluates system-wide observability resilience, operators need substation-level metrics to diagnose localized visibility loss. SOSub addresses this need by enabling fine-grained visibility into which substations are at risk of observability loss during cyber attacks, allowing for targeted mitigation and improved operational awareness. We thus propose the SOSub index, a novel metric assigning each substation a continuous score $\in [0, 1]$ reflecting its ability to determine voltage magnitude and phase angle from available measurements. Traditional approaches rely on binary measurement-presence indicators or system-wide meter placement studies. In contrast, SOSub directly assesses local state solvability via sub-Jacobian rank analysis (Gol et al., 2012). Concretely, H_j is formed by selecting from $\mathbf{H}$ the columns associated with (θ_j, V_j) , and $\text{SOSub}_j = \text{rank}(H_j)/n_j$ with $n_j = 2$, yielding a localized rank-based assessment consistent with the global Jacobian test. For substation j (with $n_j = 2$ states: voltage and angle):

- Let $\mathbf{H}_j \in \mathbb{R}^{m_{\text{eff}} \times n_j}$ denote the sub-Jacobian formed by selecting columns of $\mathbf{H}$ corresponding to substation j 's states,
- $r_j = \text{rank}(\mathbf{H}_j)$ represent the number of locally observable states.

The SOSub is defined as:

$$\text{SOSub}_j = \frac{r_j}{n_j} \in [0, 1].$$

Here, $\text{SOSub}_j = 1$ implies full local observability (both states resolvable), $\text{SOSub}_j = 0.5$ indicates partial observability (one state unobservable), and $\text{SOSub}_j = 0$ corresponds to complete local observability loss (e.g., an electrically isolated or unmonitored substation). By evaluating sub-Jacobian rank—rather than merely tallying measurements, SOSub provides a more precise and granular identification of substations vulnerable to observability loss during contingency scenarios. This enables fine-grained diagnostics and enhances situational awareness in modern power systems (Abur and Exposito, 2004).

4.3. Cyber-physical controllability and observability impact evaluation

IoC-to-Metric Procedure. For each IoC listed in Table 1, we apply a deterministic update to steady-state parameters and then recompute the indices defined in Section 4.2. The updated parameters are the effective control-input set $U_{\text{eff}} \subseteq U$ (available generator setpoints), the branch-quality scores $q_e \in [0, 1]$, the weighted mean flow deviation ΔP_{mean} , the count of buses with voltage-band violations n_{viol} , the usable-measurement count m_{eff} after screening, and the number of observable islands I . This procedure specifies the computation pipeline used in all case studies and provides a consistent basis for interpreting index variations. To assess the operational impacts of coordinated cyber-physical attacks on power system controllability and observability, we implement a comparative evaluation framework that quantifies system degradation through pre-attack and post-attack analysis.

Pre-Attack Phase. In this phase, we establish a baseline that characterizes the system's normal operational behavior in the absence of cyber threats. This baseline encompasses not only static measurements but also the system's structural and functional attributes under typical operating conditions. It provides a reference point for detecting deviations in behavior once adversarial activity is introduced. To ensure the simulation remains grounded in realistic conditions, we utilize historical load data from the New York Independent System Operator (NYISO) mapped to the IEEE 39-bus system. The mapping retains both temporal variation and spatial diversity in demand and supports an accurate representation of system dynamics. The resulting baseline enables comparative analysis of how key system-level properties evolve in response to cyber-physical disruptions. Details regarding the simulation environment and data mapping are presented in Section VI-A.

Post-Attack Phase. We evaluate system response under the three coordinated cyber-physical attack scenarios defined previously. For

each attack case, we assess how the system's controllability and observability characteristics change when subjected to false command injection, load manipulation, and false data injection attacks. This multi-scenario approach enables a comprehensive evaluation of different threat vectors and their distinct operational impacts. During each attack simulation, we monitor the behavior of controllability and observability metrics to capture system degradation patterns. The quantitative, metric-based evaluation provides objective measures of how each attack scenario affects critical system functions, from control authority degradation to measurement integrity loss. This systematic monitoring reveals the specific vulnerabilities exposed by different attack types and their cascading effects on system operations. By observing how metrics behave during and after attacks, we can detect and characterize different types of cyber-physical threats, understand how attacks impact situational awareness and control capacity, and use these insights to inform potential defense strategies. The comparative analysis across scenarios supports the identification of attack-specific indicators and enables the development of targeted mitigation approaches for enhanced system resilience.

5. Experimental design and results

This section presents the simulation-based evaluation of the proposed framework. The goal is to assess how the previously defined cyber-physical attack scenarios affect the controllability and observability of the IEEE 39-bus power system. The simulation is conducted using a co-simulation environment that integrates power system dynamics with cyber-layer interactions. We begin by outlining the simulation setup, including the data mapping process used to establish realistic operating conditions. We then evaluate the system's behavior under each attack scenario, using the baseline established in the pre-attack phase as a reference for comparison. Our analysis concludes by determining whether the power system's current state is controllable, observable, and secure.

5.1. Setup

This section establishes the experimental framework used to evaluate our proposed metrics under cyber-physical attacks. We describe the IEEE 39-bus test system with realistic NYISO operational data mapping and present the co-simulation platform that integrates power system dynamics with cyber-layer interactions to enable systematic attack assessment.

5.1.1. IEEE 39-bus model-NYISO data mapping

We use the IEEE 39-bus test system as the foundation for our analysis. The network topology is shown in Fig. 3(a). This system consists of 10 generators, 46 transmission lines, and 19 load buses. This model is widely used for evaluating transmission-level dynamics and control strategies (Pai, 1989). In our work, the power flow and dynamic simulations are executed in PowerWorld Simulator. To ensure realistic operating conditions, we map historical load data from the NYISO onto the load buses of the IEEE 39-bus system. Specifically, we use 5-minute interval load data from NYISO's publicly available archives (New York Independent System Operator, 2025). The NYISO zonal map used in our study is identical to that shown in Fig. 3(b). Each NYISO load zone is proportionally assigned to a subset of the 39-bus system's load buses based on nominal demand levels. We assign these zones to load buses in a two-step process. First, each NYISO zone is mapped to a unique load bus based on nominal load magnitude and spatial proximity. Second, the remaining load buses are assigned reused zones by incorporating data from a different operational year to preserve statistical independence. Thus, we can produce a complete assignment across all 19 load buses while maintaining realism in both spatial and temporal dimensions (Ge and Jiao, 2020).

The load time series are scaled to match the nominal base case profile of the 39-bus system, ensuring consistency in power factor and load magnitude. This preserves the network's steady-state characteristics while introducing realistic temporal dynamics derived from actual grid operations. Control zone definitions used for multi-area monitoring are overlaid on the bus-to-zone mapping, allowing NYISO-derived fluctuations to propagate through the dynamic simulation and influence both local and system-wide observability.

5.1.2. Cyber-physical simulation process

We developed an experimental setup with a co-simulator (Asiri et al., 2023). The co-simulator integrates cyber and physical domain modeling to create an environment and enable systematic assessment of attack scenarios while capturing the interdependencies between communication networks and power system operations. The co-simulation integrates three primary components: CORE network emulator for communication infrastructure modeling, ScadaBR and OpenPLC for control system functionality, and PowerWorld for power system simulation.

Our co-simulation achieves cyber-physical coupling through protocol-based integration using industrial communication standards. OpenPLC functions as the Modbus master, establishing communication with substation devices through Modbus TCP protocol with a 100 ms polling cycle for real-time monitoring and control operations. Power system integration is accomplished through PowerWorld's Simulator Automation Server (SimAuto) interface, which provides programmatic access to simulation data and control functionalities. The bidirectional communication bridge enables the control process loop in OpenPLC to retrieve power system readings and execute control commands within the PowerWorld simulation environment.

The communication network system is modeled using CORE emulator (Ahrenholz et al., 2008). Such an emulator provides a realistic communication network representation with configurable parameters such as data rates, propagation delays, and network topology. The integrated environment enables assessment of cyber-attack impacts on power system controllability and observability through realistic operational scenarios, supporting the validation of our proposed framework for enhanced smart grid security assessment.

5.2. Experimental results

In this section, we present the simulation results for the three attack scenarios: FCI, LR, and FDIA. We evaluate how each attack affects the system's controllability and observability capabilities. The analysis quantifies the operational impacts of each attack vector on grid stability and situational awareness.

5.2.1. Case I: False Command Injection (FCI)

As described in Section 4.1, the attack begins with a covert compromise of the SCADA network, where the attacker gains access and escalates privileges without immediately disrupting system behavior. At this stage, controllability remains theoretically intact, but the adversary's ability to issue unauthorized Modbus commands introduces a latent loss of control authority. Once the attacker transmits a forged command to open circuit breakers at critical generators, the system transitions into a state of partial-to-full controllability loss. Such loss may remain undetected if the operator retains access to SCADA and PLC panels and thus creates a deceptive sense of control. The adversary aims to target critical power components, such as generators and transformers, to damage the power system or cause a service interruption.

After the malicious command executes, the circuit breakers connecting two generators to the grid are simultaneously opened. Consequently, the system experiences significant degradation across multiple dimensions of stability and control. The SCI index quantifies this degradation with a measured score of 0.74, indicating a 25.8% decrease from baseline values, which typically range between 0.97 and 0.99 during

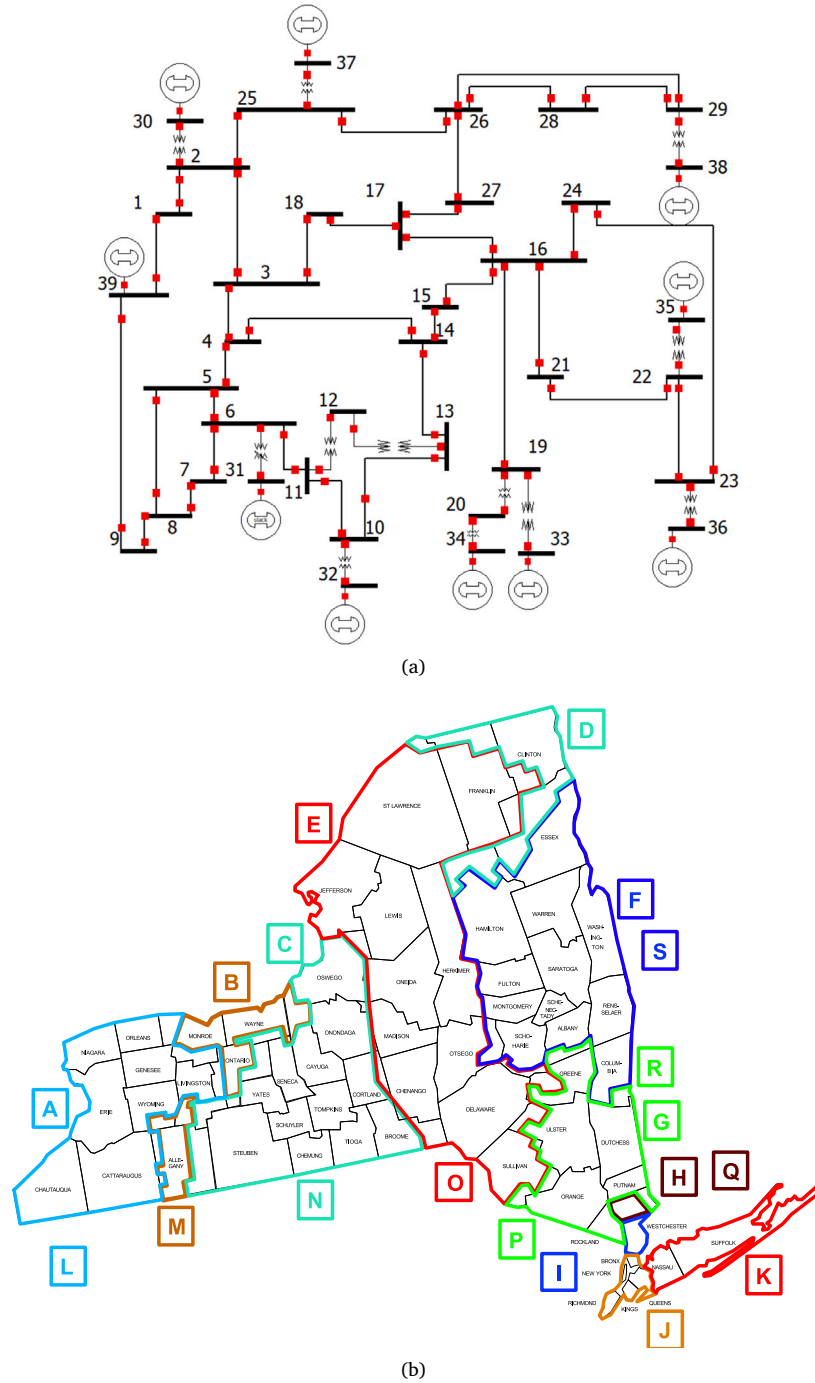


Fig. 3. IEEE 39-bus system mapped into NYISO control area load zone data. (a) IEEE 39-bus system. (b) NYISO map of the 19 control area load zones.

standard operations, as shown in Fig. 4. In contrast, structural connectivity exhibits moderate degradation, decreasing by 10.1% to 0.90. However, the operational stability metric suffers a disproportionately severe decline, falling by 40.2% to 0.59.

Fig. 5 illustrates how this targeted attack triggers cascading effects, observable in branch loading patterns. Prior to the attack, the transmission network maintained uniformly optimal edge quality factors ($q_e = 1.0$). Following generator disruption, specific branches develop critical vulnerabilities, particularly at nodes 2–3 and 2–30 where edge quality deteriorates sharply ($q_e = 0.30$ and 0.28), and corresponding loading ratios reach hazardous levels of 1.85 and 1.60, respectively. Meanwhile, the flow deviation metric increased from 7.54×10^4 MW

before the attack to 2.24×10^6 MW after the attack, which is about 29.7 times higher (a 2,871% increase). This substantial reconfiguration of power flow reflects the system's emergency compensation mechanisms attempting to maintain equilibrium despite the sudden loss of generation capacity.

We observed that the attack caused voltage violations at bus 31, where the generation was compromised. As a result, such an attack indicates the attack's direct impact on local voltage regulation capabilities. The voltage instability further influences the system's critical control infrastructure, where notable reconfiguration occurs following the attack. While generator bus 38 is removed entirely from the critical set, bus 31 remains critical despite its compromised operational status,

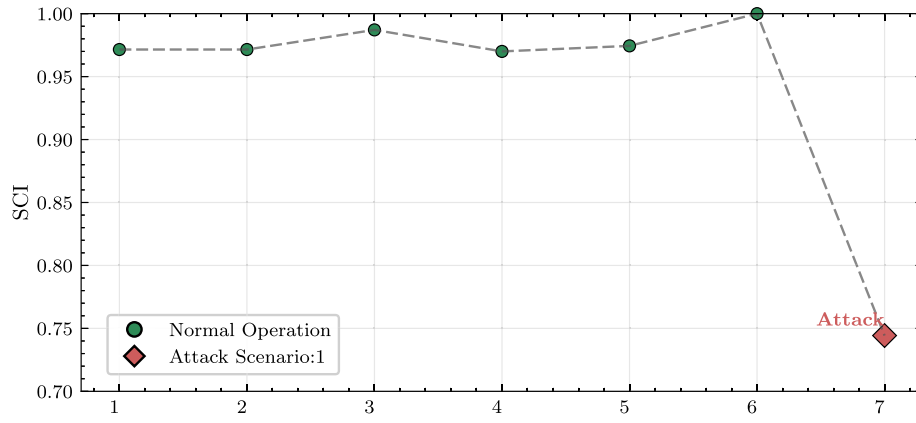


Fig. 4. Temporal evolution of system controllability under attack scenario 1.

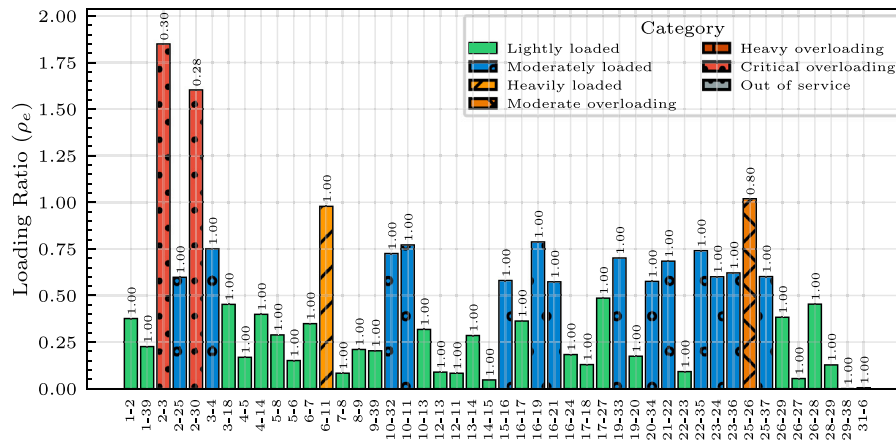


Fig. 5. Comparative analysis of branch loading ratios (ρ_e) and edge quality scores (q_e): Scenario 1.

specifically due to its newly emerged voltage violations. This finding emphasizes an important characteristic of the SCI index: its ability to adaptively identify critical infrastructure based on both structural importance and operational violations. From a controllability perspective, systems must monitor not only the connectivity of critical control points but also their operational states. Generators that become inoperable may still remain essential to system controllability due to their influence on voltage stability and network dynamics.

Further, the SCRC index offers deeper insight into how the attack scenario impairs the system's capacity to regulate its critical elements using available control actions, as detailed in Table 2. The attack at buses 31 and 38 reduced the SCRC from 0.88 to 0.78. This change reflects an 11.36% decline in overall system controllability. The degradation is evident across both voltage and power flow domains. Bus controllability dropped from 97.4% to 87.2%, while branch controllability experienced a more pronounced reduction from 80.4% to 70.5%. In addition, maximum voltage sensitivity rose from 1.00 to 1.15 pu/MW. Maximum flow sensitivity nearly doubled, increasing from 5.66 to 10.04 MW/MW. These shifts reinforce the operational stability deterioration previously observed in the SCI results.

Sensitivity analysis further revealed that the remaining generators must exert significantly greater influence to maintain system control post-attack, consistent with the cascading effects observed in branch loading patterns. Average flow sensitivity rose by 56.25% from 0.096 to 0.150 MW/MW, indicating increased stress on control resources. Notably, a 20% reduction in generation capacity led to a disproportionate 13.3% loss in controllable elements (10 out of 75). Branches connected to buses 31–6 exhibited the most severe controllability degradation, forming localized vulnerability zones with limited control reach. These

zones align with areas showing critical edge quality deterioration in the SCI assessment. Consequently, the SCRC results quantitatively confirm that the attack not only undermines system stability metrics but also fundamentally erodes the power system's ability to regulate power flow across its critical elements.

Beyond controllability degradation, the FCI attack significantly compromises system observability, thereby undermining the operator's situational awareness and the reliability of state estimation. The SOM index declined from 0.442 to 0.297, marking a 32.8% reduction in measurement redundancy. This sharp drop indicates that the system retains only a 29.8% redundancy margin, critically close to the observability threshold where state estimation becomes infeasible. In contrast to the moderate 10.1% decline in structural connectivity observed in controllability metrics, observability suffers a disproportionate deterioration, comparable to the 40.2% drop in operational stability. A summary of the observability and redundancy findings is presented in Table 3.

Observable islands increase from 1 to 9 following the attack, fragmenting the measurement topology by 800%. Each island requires an independent angular reference, which prevents unified system-wide state estimation. As a result, blind spots emerge between regions, compromising inter-area oscillation monitoring and coordinated control actions. This fragmentation in measurement coverage directly corresponds to the power flow redistribution patterns identified in the controllability analysis. In that analysis, flows increased by 2,871% as the system struggled to maintain equilibrium.

Substation-level observability reveals degradation patterns that align with the targeted attack locations. Before the attack, only bus 31 (the reference bus) exhibited partial observability ($SOSub = 0.5$). After the generator disruption, buses 28, 29, 30, and 38 also entered this degraded state. Consequently, average system observability declined from

Table 2

Control reach coverage analysis of system parameters under attack Scenario 1.

System parameter	Normal scenario	Under attack	Change (%)
Overall SCRC	0.88	0.78	-11.36
Voltage Controllability	0.97	0.87	-10.31
Flow Controllability	0.80	0.70	-12.50
Avg. Generator Reach	7.8	6.2	-20.51
Max Voltage Sens. (pu/MW)	1.00	1.15	+15.00
Max Flow Sens. (MW/MW)	5.66	10.04	+77.39
Avg. Flow Sens. (MW/MW)	0.096	0.150	+56.25
Total Controllable Elements	75	65	-13.33

Note: Avg. Generator Reach represents the mean number of system elements (buses and branches) that each active generator can effectively control (i.e., influence beyond established sensitivity thresholds of 0.01 pu/MW for voltage and 0.05 MW/MW for flow). This metric quantifies how the control capability of individual generators degrades even beyond the simple reduction in generation resources.

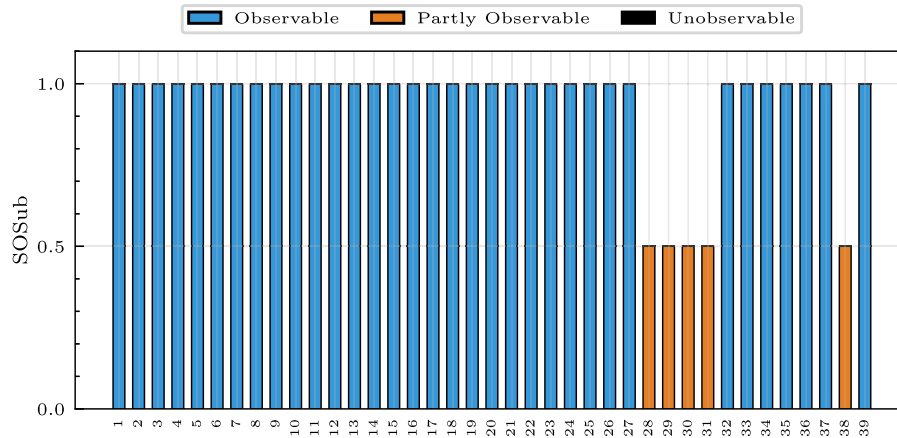


Fig. 6. SOSub scores under attack, showing partial observability (SOSub = 0.5) at buses 28, 29, 30, 31, and 38 and full observability (SOSub = 1.0) for all other substations.

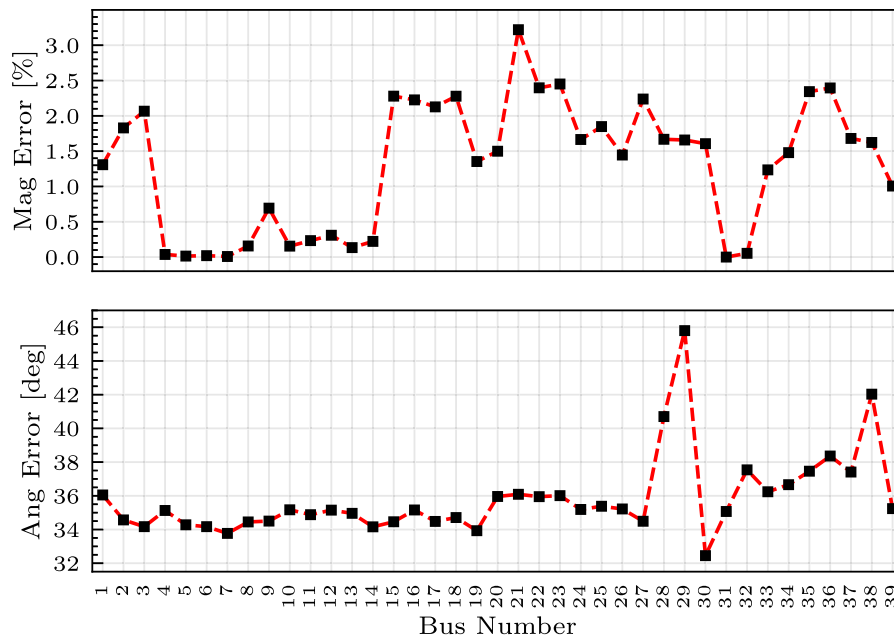


Fig. 7. Voltage angle and magnitude error under Attack Scenario 1.

0.987 to 0.935, as shown in Fig. 6. These newly affected substations correspond precisely to the critical edge quality deterioration zones identified in the SCI assessment. In particular, branches 2–3 and 2–30 experienced the most severe loading stress, with edge quality values of 0.30 and 0.28, respectively.

A closer inspection of state estimation accuracy in Fig. 7 reveals catastrophic deterioration. The deviation of state estimation surged from 0.7904 to 25.393. Angular estimation errors were the primary contributor, increasing from 1.117° to 35.911°, while voltage magnitude errors doubled from 0.0384 to 0.1199 p.u. These errors exceed

Table 3

Observability & redundancy analysis across normal and attack scenario.

Metric	Pre-attack	Post-attack	Change (%)
SOM	0.442	0.297	−32.8%
Observable Islands	1	9	+800%
Measurement Redundancy Index	44.2	29.8	−32.6%
Average SOSub	0.987	0.935	−5.27%
Unreliable Measurements	69/207	86/207	+24.6%

Note: The 800% increase in observable islands corresponds to the formation of nine islands (from one in the normal state to nine under attack, i.e., a $\times 9$ increase).

operational thresholds for secure system operation, rendering state estimates unreliable for real-time decision-making.

Measurement effectiveness analysis shows that 41.5% of measurements (86 out of 207) become unreliable under the extreme operating conditions induced by the attack. This degradation occurs without any physical damage to the measurement infrastructure. Instead, the severe operating point violates fundamental assumptions underlying state estimation algorithms. Unreliable measurements are concentrated around buses 31 and 38, creating zones where operators lose visibility precisely where enhanced monitoring is most critical.

The correlation between observability and controllability metrics reveals compound vulnerabilities. The 11.36% reduction in SCRC translates into a diminished ability to validate control actions through state estimation. Operators can no longer reliably assess system responses to control inputs. The near doubling of maximum flow sensitivity from 5.66 to 10.04 MW/MW in the controllability domain corresponds to increased estimation uncertainty in the observability domain. In this context, small measurement errors propagate into large deviations in state estimation.

The operational implications extend beyond numerical degradation. With only 29.8% measurement redundancy remaining, any additional measurement loss could lead to complete observability failure. The 35.911° average angular error prevents accurate contingency analysis, security assessment, and optimal power flow calculations. Under such conditions, operators must abandon model-based decision support tools and rely solely on direct telemetry. This shift sacrifices comprehensive situational awareness in favor of localized visibility. The observability analysis thus confirms that coordinated attacks trigger cascading degradation across multiple system dimensions. A loss of generation capacity leads to compromised visibility, which in turn amplifies operational vulnerabilities.

5.2.2. Case II: Load Redistribution (LR)

In this scenario, the attacker executes a coordinated load redistribution by altering demand values at specific buses while keeping the total system load unchanged. The attack targets the eastern corridor of the network particularly the area spanning buses 21, 28, and 29—by increasing demand at these locations and reducing it at central buses such as 4, 8, and 20. The objective is to redirect power flows toward a less redundant and more constrained section of the grid, exploiting its limited meshing to introduce hidden operational stress. All changes remain within normal load variation thresholds, making them unlikely to trigger forecasting or anomaly detection systems. This strategy leverages topological insight rather than internal system metrics, aiming to subtly degrade system controllability without causing immediate violations or alarms.

Quantitative analysis reveals significant differences in system controllability between normal operating conditions and the load redistribution attack scenario. Under normal conditions, the SCI index consistently ranged between 0.96 and 0.98, indicating robust controllability. This performance is supported by high structural connectivity (1.00) and operational stability values between 0.92 and 0.96. The detailed SCI scores and component breakdowns for both baseline and attack scenarios are presented in Table 4.

Table 4

Comparative analysis of SCI index across normal and attack scenarios.

Scenario	Metric	Range	Attack	Change (%)
Scenario 1	Overall SCI	0.97–1.00	0.74	−25.80
	S	1.00–1.00	0.90	−10.08
	D	0.95–1.00	0.59	−40.19
	ΔP (MW)	1.10×10^4 – 2.96×10^5	2.24×10^6	$\times 29.7$
Scenario 2	Overall SCI	0.96–0.98	0.84	−13.40
	S	1.00–1.00	1.00	0.00
	D	0.92–0.96	0.68	−29.20
	ΔP (MW)	2.21×10^5 – 4.42×10^5	1.77×10^6	$\times 5.33$

When subjected to a load redistribution attack, the system previously in a carefully balanced state experienced substantial stress. The overall SCI dropped by 13.4% to 0.84, as shown in Fig. 8. This degradation followed a distinct pattern across components. Structural connectivity remained unaffected, maintaining its perfect score of 1.00. However, operational stability absorbed the full impact, declining by 29.2% to 0.68. This behavior suggests that while the topology of critical connections remained intact, the system's ability to sustain stable operational parameters was significantly compromised.

As illustrated in Fig. 9, edge quality assessment revealed clear differences in loading patterns between normal and attack conditions. During normal operations, approximately 90% of branches exhibited quality factors (q_e) of 1.0, indicating nominal performance. Under attack conditions, however, specific branches experienced overloading. Branches 26–28, 28–29, and 29–38 were most affected. Branch 26–28 showed heavy overloading, with a quality factor of 0.60 and a loading ratio of 1.25. Branch 28–29 experienced moderate overloading, with a quality factor of 0.80 and a loading ratio of 1.10. The most severe impact occurred at branch 29–38, where the quality factor dropped to 0.30 and the loading ratio reached 1.65, well beyond normal operational thresholds. These localized vulnerabilities highlight the non-uniform distribution of stress across the network under attack conditions. Consequently, total system power flow increased significantly. While normal conditions saw flows ranging from approximately 2.21×10^5 to 4.42×10^5 MW, the attack scenario pushed this value to 1.77×10^6 MW. This substantial increase demonstrates the attack's ability to redistribute internal power flows while maintaining overall system stability.

Moreover, the SCRC index provides a dynamic perspective on how the same load redistribution attack affects system controllability margins. Under attack conditions, the SCRC value declined from 0.8824 to 0.8588. This 2.67% reduction indicates a subtle yet operationally significant degradation in system controllability. Although the change appears modest, it reflects a measurable weakening in the system's capacity to execute coordinated control actions in response to disturbances.

At the bus level, controllability remained largely unaffected. Of the 39 buses, 38 retained effective voltage control, corresponding to 97.4% coverage. This result suggests that local voltage observability was preserved under the constrained attack scenario. In contrast, branch-level controllability declined from 80.4% to 76.1%, with only 35 of 46 branches remaining controllable. The 5.3% reduction in flow control reach indicates a diminished capacity to regulate power transfers across the network.

Sensitivity analysis offers further insight into the system's dynamic response. Maximum voltage sensitivity remained constant at 1.00 pu/MW, indicating that the most responsive voltage control point was unaffected. However, maximum flow sensitivity decreased from 5.66 MW/MW to 4.40 MW/MW, reflecting a redistribution of control stress away from the most heavily loaded branches. Average voltage sensitivity increased slightly from 0.1106 to 0.1121 pu/MW. At the same time, average flow sensitivity rose from 0.0959 to 0.1336 MW/MW, a 39% increase. These shifts suggest a broader dispersion of control effort and a corresponding increase in the effort required to achieve targeted flow adjustments.

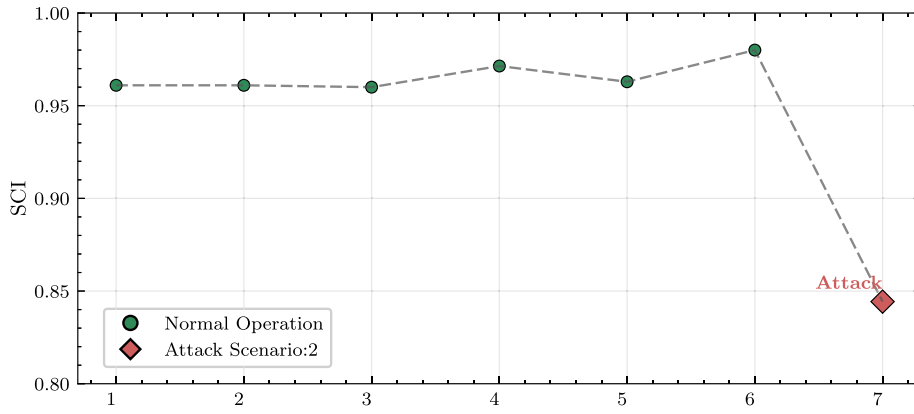


Fig. 8. Temporal evolution of system controllability under attack scenario 2.

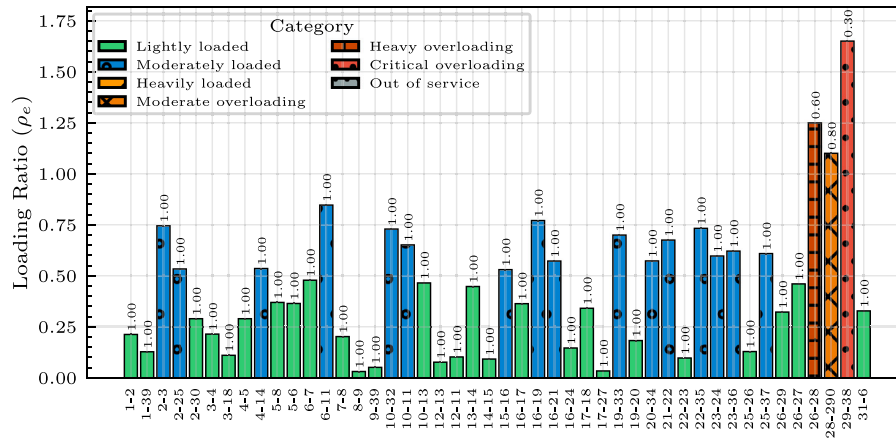


Fig. 9. Comparative analysis of branch loading ratios (ρ_e) and edge quality scores (q_e): Scenario 2.

The attack was deliberately constrained to $\pm 20\%$ load variations and applied selectively to a subset of buses. These parameters enabled a gradual erosion of controllability margins while avoiding detection by conventional monitoring systems. Unlike abrupt or randomized load perturbations, which typically trigger alarms, this low-intensity strategy remained below standard detection thresholds.

5.2.3. Case III: False data injection (FDIA)

The controllability impact observed here closely mirrors the effects described in use case 1. In this scenario, the adversary escalates the attack vector by executing an FDIA strategy augmented with breaker status spoofing. Unlike the direct command injection used in use case 1, this attack introduces a deceptive layer that conceals physical topology changes by falsifying circuit breaker status indicators. As a result, although generators are physically disconnected, the system continues to report them as operational—creating a critical cyber–physical incongruity between the cyber representation and the actual physical state of the grid.

The impact of this stealthy attack is profound. Network observability deteriorates significantly, as reflected by the SOM index dropping to 0.1979. This represents a 33.37% reduction relative to use case 1, leaving only a 19.8% redundancy margin. Such a collapse in observability places the system on the brink of operational blindness, with minimal resilience to further measurement or component losses.

A particularly insidious feature of this FDIA is its ability to preserve the appearance of a single, coherent observable island. While use case 1 exposed system fragmentation and the formation of multiple observable islands, the status-spoofed scenario maintains the illusion of network integrity. This false continuity delays operator detection and response,

critically impairing situational awareness during a period of heightened vulnerability.

Measurement integrity is also severely compromised. Of the 207 measurements analyzed, 111 (53.6%) are flagged as bad data—an increase of 29% over compared to case 1. This surge in data rejection results from the irreconcilable mismatch between actual system conditions and the falsified topology model. The state estimator, confronted with this inconsistency, is unable to reconcile expected and observed system behavior.

Paradoxically, the spoofed scenario exhibits fewer partially observable substations, limited to bus 31, compared to use case 1, which affected buses 28, 29, 30, 31, and 38. This apparent improvement in observability metrics is misleading and dangerous. It reinforces a false sense of system health precisely when conditions are deteriorating. The absence of observable island fragmentation further compounds this illusion.

State estimation accuracy reveals a nuanced degradation pattern. While the total estimation deviation (23.9632) is slightly lower than in use case 1 (25.3930), the angular estimation error remains severe (33.89° versus 35.91°). Interestingly, voltage magnitude errors improve significantly (0.0380 versus 0.1199 p.u.). This asymmetric impact on estimation quality—where some metrics appear to improve while overall system integrity declines—creates a hazardous monitoring environment. In such a context, partially valid results may inadvertently reinforce operator trust in a compromised system.

5.3. IoC-to-metric quantitative mapping

The cyber-layer IoCs in Table 1 are intended to be instantiated from standard OT telemetry, such as SCADA/EMS audit trails and

Table 5
IoC-to-Parameter-to-Metric mapping.

IoC	Parameter perturbation	Metric impact	Quantitative effect
Scenario 1: FCI			
A1	†	–	Precursor only
A2	†	–	Enables attack
A3	†	–	Target identification
A4	$U_{\text{eff}} \downarrow, q_e \downarrow, \Delta P_{\text{mean}} \uparrow, n_{\text{viol}} \uparrow, I \uparrow$ (if boundary measurements are lost)	SCRC↓, SCI↓, SOM↓ (when $m_{\text{eff}} \downarrow$ and/or $I \uparrow$), SOSub↓	Observed SCRC –11.36% (case study)
A5	†	–	Detection indicator
A6	†	–	Detection indicator
A7	$U_{\text{eff}} \downarrow, \Delta P_{\text{mean}} \uparrow, q_e \downarrow$	SCRC↓, SCI↓	Physical confirmation
A8	$n_{\text{viol}} \uparrow$	SCI↓ (via V_{score})	Stability degradation
A9	$\Delta P_{\text{mean}} \uparrow, q_e \downarrow$ (congestion)	SCI↓, local SCRC↓	System stress
Scenario 2: LR			
B1	$\Delta P_{\text{mean}} \uparrow, q_e \downarrow$ (corridor), U_{eff} unchanged	SCI↓, minor SCRC↓	Observed SCI –13.4%, SCRC –2.67%
B2	$m_{\text{eff}} \downarrow$	SOM↓, SOSub↓	Communication degradation
B3	$U_{\text{eff}} \downarrow$ (if commands fail)	SCRC↓	Control congestion
B4	$\Delta P_{\text{mean}} \uparrow, q_e \downarrow$	SCI↓	Operational stress
B5	$q_e \downarrow$ (specific lines)	SCI↓ (via S_{adj})	$q_e \in \{0.60, 0.80, 0.30\}$
B6	$\Delta P_{\text{mean}} \uparrow, q_e \downarrow$	SCI↓, local SCRC↓	Flow redistribution
Scenario 3: FDIA			
C1	†	–	Detection indicator
C2	$m_{\text{eff}} \downarrow, I \uparrow$ (if boundary measurements are lost)	SOM↓, SOSub↓	Observed SOM 0.297 → 0.1979
C3	$m_{\text{eff}} \downarrow$	SOM↓	Data staleness
C4	$m_{\text{eff}} \downarrow, I \uparrow$ (possible)	SOM↓, SOSub↓	Measurement loss
C5	$m_{\text{eff}} \downarrow$ (screening), $\Delta P_{\text{mean}} \uparrow$ (if trusted)	SOM↓; SCI↓ if $\Delta P_{\text{mean}} \uparrow$	Compound effect
C6	No direct index parameter; $n_{\text{viol}} \uparrow$ (if voltage-band limits are violated)	SCI↓ (via V_{score})	Stability indicator
C7	$q_e \downarrow$	SCI↓ (via $S_{\text{adj}}, Q_{\text{avg}}$)	Physical degradation

Note: ↓ indicates decrease, ↑ indicates increase. † denotes diagnostic IoCs that enable attack detection without causing parameter perturbations. — indicates no direct impact. Quantitative effects shown where measured in experiments.

protocol/network observations (e.g., Modbus/TCP transaction traces). These cyber indicators provide corroborating evidence of cyber compromise and help distinguish cyber-originated manipulation from local manual action. In this section, we use the identified IoCs to update the corresponding steady-state parameters and then recompute the proposed indices to quantify the resulting operational impact. For each detected IoC listed in Table 1, the steady-state parameters are updated and the indices defined in Section 4.2 are recomputed. The updated parameters include the effective control-input set $U_{\text{eff}} \subseteq U$ (available generator setpoints), branch-quality scores $(q_e) \in [0, 1]$, the weighted mean flow deviation ΔP_{mean} , the number of buses with voltage-band violations n_{viol} , the usable-measurement count m_{eff} after screening, and the number of observable islands I . This computation procedure follows the methodology in Section 5 and is applied uniformly across all scenarios. The indices are evaluated on these parameters as follows:

- $\text{SCRC} = (\sum_i w_i \delta_i) / (\sum_i w_i)$ with $\delta_i = 1$ if $\max_{u \in U_{\text{eff}}} |\partial q_i / \partial u| > \varepsilon(\cdot)$ and 0 otherwise.
- $\text{SCI} = \alpha S_{\text{adj}}(\{q_e\}) + (1 - \alpha) D(\Delta P_{\text{mean}}) V_{\text{score}}(n_{\text{viol}})$.
- $\text{SOM} = [m_{\text{eff}} - (n_x - (I - 1))] / m_{\text{eff}}$ with $n_x = 2N - I$.
- $\text{SOSub}_j = \text{rank}(H_j) / 2$.

The metrics respond to these parameter changes according to their respective formulations. SCRC is sensitive to reductions in U_{eff} and the suppression of control sensitivities below threshold levels. SCI is jointly influenced by structural connectivity degradation (via q_e) and operational instability (via ΔP_{mean} and n_{viol}). SOM decreases monotonically with measurement loss and with fragmentation into multiple observable islands, while SOSub captures localized observability degradation at substations affected by compromised measurements. Table 5 consolidates the complete mapping between IoCs, affected parameters, and metric impacts.

Under FCI scenario, initial reconnaissance IoCs (A1–A3) provide no direct parameter changes. Malicious control actions (A4) are executed to remove generation resources from service, directly reducing U_{eff} . A cascade of stress is triggered across the network. Branch stress is elevated with degraded q_e . Generation-demand imbalance raises ΔP_{mean} . Voltage limits can be breached, increasing n_{viol} . The grid may splinter into islands when boundary or bridging measurements are lost, increasing I . These disruptions produce measurable reductions in SCRC and SCI, reflecting a weakened ability to stabilize the system. SOM deteriorates as m_{eff} is reduced and I increases simultaneously. Abnormal communication patterns (A5–A6) remain diagnostic unless

successful actuation is achieved. Physical-layer indicators (A7–A9) provide tangible confirmation through observable generation loss, voltage excursions, and compensatory ramping.

In contrast to the direct physical disruptions of FCI, the LR scenario operates differently by manipulating control setpoints for certain loads (B1) to redistribute demand within the grid. Total system consumption remains unaltered. The system responds by adjusting generation dispatch and tapping reserves, increasing ΔP_{mean} . Transmission corridors experience degraded q_e with reduced stability margins. U_{eff} remains unaffected. The operational component of SCI experiences significant degradation due to flow imbalances. Structural connectivity may remain intact. Supporting IoCs (B2–B3) can impair communication reliability, reducing m_{eff} and potentially affecting SOM and SOSub. Physical indicators (B4–B6) manifest as flow imbalances and anomalous power patterns. These further degrade branch quality and flow stability. SCRC experiences modest degradation in heavily loaded regions where control actions lose efficacy.

Unlike FCI and LR, which target physical and control-layer properties, the FDIA attack is characterized by corrupting the measurement data rather than physical manipulation. C4 provides diagnostic indication. C5 introduces stealthy measurement discrepancies that reduce the effective measurement pool. Timestamp anomalies combined with communication delays (C6–C7) render measurements untrustworthy or desynchronized. The grid can splinter into islands when boundary measurements are corrupted, increasing I in the observer's model. Observability metrics are primarily affected. SOM declines according to its formula above. SOSub decreases at substations where measurement redundancy no longer satisfies rank conditions for observability. Physical-layer indicators (C8–C10) either introduce additional measurement corruption or induce equipment stress. This reduces q_e and causes compounding degradation of both observability and controllability.

Our results validate these distinct impacts. FCI scenarios caused an 11.36% reduction in SCRC and a 25.80% decline in SCI. SOM decreased from 0.442 to 0.297. Load redistribution induced a 13.4% reduction in SCI, indicating notably weakened operational connectivity. SCRC decreased by 2.67%. Transmission corridors experienced degraded qualities of $q_e \in \{0.60, 0.80, 0.30\}$, highlighting localized vulnerabilities. The FDIA scenario drove SOM from 0.297 to 0.1979, reflecting a major loss of reliable situational awareness. This resulted from a substantial reduction in m_{eff} . A single-island topology was maintained in the reported case.

6. Conclusions

This paper presents a novel methodology for quantifying how cyber–physical attacks impact power grid operations by linking specific IoCs to fundamental system properties. Unlike existing approaches that rely on binary detection mechanisms, we develop a framework that provides a continuous, quantitative assessment of how cyber threats systematically degrade controllability and observability in real-time.

Our key methodological innovation lies in developing metrics that capture both structural topology changes and operational performance degradation. This enables operators to understand attack progression before critical failures occur, representing a paradigm shift from reactive security monitoring to proactive vulnerability assessment based on measurable system degradation. Our work provides a foundation for developing adaptive defense strategies that respond to progressive system degradation and enable preemptive countermeasures based on quantitative threat assessment. In the future, we will focus on integrating this methodology with machine learning approaches and validating the framework across larger, more complex power systems.

CRedit authorship contribution statement

Mohammed Asiri: Writing – original draft, Visualization, Validation, Software, Methodology, Conceptualization. **Neetesh Saxena:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Subhash Lakshminarayana:** Writing – review & editing, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research work was partially supported by the RITICS (UK) and EPSRC (EP/Y026233/1).

Data availability

This research uses IEEE 39-bus case. Specifically, we use 5-minute interval load data from NYISO's publicly available archives (<https://www.nyiso.com/load-data>).

References

- Abbaspour, A., Sargolzaei, A., Forouzaneshad, P., Yen, K.K., Sarwat, A.I., 2019. Resilient control design for load frequency control system under false data injection attacks. *IEEE Trans. Ind. Electron.* 67 (9), 7951–7962.
- Abur, A., Exposito, A.G., 2004. *Power System State Estimation: Theory and Implementation*. CRC Press.
- Adhikari, U., Morris, T., Pan, S., 2017. WAMS cyber-physical test bed for power system, cybersecurity study, and data mining. *IEEE Trans. Smart Grid* 8 (6), 2744–2753. <http://dx.doi.org/10.1109/TSG.2016.2537210>.
- Ahrenholz, J., Danilov, C., Henderson, T.R., Kim, J.H., 2008. CORE: A real-time network emulator. In: *MILCOM 2008-2008 IEEE Military Communications Conference*. IEEE, pp. 1–7.
- Arefifar, S.A., Ordóñez, M., Mohamed, Y.A.-R.I., 2016. Voltage and current controllability in multi-microgrid smart distribution systems. *IEEE Trans. Smart Grid* 9 (2), 817–826.
- Ashok, A., Govindarasu, M., Wang, J., 2017. Cyber-physical attack-resilient wide-area monitoring, protection, and control for the power grid. *Proc. IEEE* 105 (7), 1389–1407. <http://dx.doi.org/10.1109/JPROC.2017.2686394>.
- Asiri, M., Arunasalam, A., Saxena, N., Celik, Z.B., 2025. Frontline responders: Rethinking indicators of compromise for industrial control system security. *Comput. Secur.* 154, 104421. <http://dx.doi.org/10.1016/j.cose.2025.104421>, URL <https://www.sciencedirect.com/science/article/pii/S0167404825001105>.
- Asiri, M., Saxena, N., Burnap, P., 2023. ARCSG: Advancing resilience of cyber-physical smart grid: An integrated co-simulation approach incorporating indicators of compromise. In: *2023 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. pp. 370–378. <http://dx.doi.org/10.1109/EuroSPW59978.2023.00047>.
- Awon, M., Butt, H.Z., Khalid, H.A., Janjua, A.K., 2018. System redundancy index based optimized phasor measurement unit placement for complete network observability. In: *2018 International Conference on Power Generation Systems and Renewable Energy Technologies. PGSRET, IEEE*, pp. 1–5.
- Bae, J., Lee, S., Kim, Y.-W., Kim, J.-H., 2017. Protection strategies against false data injection attacks with uncertain information on electric power grids. *J. Electr. Eng. Technol.* 12 (1), 19–28.
- Bedi, G., Venayagamoorthy, G.K., Singh, R., Brooks, R.R., Wang, K.-C., 2018. Review of internet of things (IoT) in electric power and energy systems. *IEEE Internet Things J.* 5 (2), 847–870.
- Bessa, R., Moreira, C., Silva, B., Matos, M., 2019. Handling renewable energy variability and uncertainty in power system operation. *Adv. Energy Syst.: Large-Scale Renew. Energy Integr. Chall.* 156–178.
- Cheng, G., Lin, Y., Abur, A., Gómez-Expósito, A., Wu, W., 2024. A survey of power system state estimation using multiple data sources: PMUs, SCADA, AMI, and beyond. *IEEE Trans. Smart Grid* 15 (1), 1129–1151. <http://dx.doi.org/10.1109/TSG.2023.3286401>.
- Choi, J.-H., Moon, S.-I., 2009. The dead band control of LTC transformer at distribution substation. *IEEE Trans. Power Syst.* 24 (1), 319–326.
- Davis, C.M., Overbye, T.J., 2010. Multiple element contingency screening. *IEEE Trans. Power Syst.* 26 (3), 1294–1301.
- Eppstein, M.J., Hines, P.D., 2012. A “random chemistry” algorithm for identifying collections of multiple contingencies that initiate cascading failure. *IEEE Trans. Power Syst.* 27 (3), 1698–1705.
- Fotopoulou, M., Petridis, S., Karachalios, I., Rakopoulos, D., 2022. A review on distribution system state estimation algorithms. *Appl. Sci.* 12 (21), 11073.
- Ge, Q., Jiao, C., 2020. Mitigating the impacts of false data injection attacks in smart grids using deep convolutional neural networks. In: *2020 IEEE 10th International Conference on Electronics Information and Emergency Communication. ICEIEC, IEEE*, pp. 174–177.
- Ge, Y., Li, Y., Xu, T., He, Z., Zhu, Q., 2023. Analysis of controllability in cyber–physical power systems under a novel load-capacity model. *Processes* 11 (10), 3046.

- Gol, M., Abur, A., Galvan, F., 2012. Metrics for success: Performance metrics for power system state estimators and measurement designs. *IEEE Power Energy Mag.* 10 (5), 50–57.
- Gou, B., Abur, A., 2000. A direct numerical method for observability analysis. *IEEE Trans. Power Syst.* 15 (2), 625–630.
- Greenwood, D.M., Taylor, P.C., 2014. Investigating the impact of real-time thermal ratings on power network reliability. *IEEE Trans. Power Syst.* 29 (5), 2460–2468. <http://dx.doi.org/10.1109/TPWRS.2014.2305872>.
- Guo, J., Han, Y., Guo, C., Lou, F., Wang, Y., 2017. Modeling and vulnerability analysis of cyber-physical power systems considering network topology and power flow properties. *Energies* 10 (1), 87.
- Jacobs, N., Hossain-McKenzie, S., Summers, A., Jones, C.B., Wright, B., Chavez, A., 2020. Cyber-physical observability for the electric grid. In: 2020 IEEE Texas Power and Energy Conference. TPEC, pp. 1–6. <http://dx.doi.org/10.1109/TPEC48276.2020.9042494>.
- Ju, W., Sun, K., Qi, J., 2017. Multi-layer interaction graph for analysis and mitigation of cascading outages. *IEEE J. Emerg. Sel. Top. Circuits Syst.* 7 (2), 239–249.
- Kalman, R.E., 1968. Lectures on controllability and observability. centro intern. mat. estivo seminar on controllability and observability. Report identifier: NASA-CR-113869.
- Kholidy, H.A., 2021. Autonomous mitigation of cyber risks in the cyber-physical systems. *Future Gener. Comput. Syst.* 115, 171–187. <http://dx.doi.org/10.1016/j.future.2020.09.002>.
- Kreindler, E., Sarachik, P.E., 1964. On the concepts of controllability and observability of linear systems. *IEEE Trans. Autom. Control* 9, 129–136, URL <https://api.semanticscholar.org/CorpusID:123519029>.
- Lee, R.M., Assante, M.J., Conway, T., 2016. Analysis of the cyber attack on the ukrainian power grid. *SANS Industrial Control Systems, (E-ISAC)* 23, 1–20.
- Li, Y., Ge, Y., Xu, T., Zhu, M., He, Z., 2024. Controllability evaluation of complex networks in cyber-physical power systems via critical nodes and edges. *Int. J. Electr. Power Energy Syst.* 155, 109625.
- Li, Y., Wu, J., Li, S., 2017. Controllability and observability of CPSs under networked adversarial attacks. *IET Control Theory Appl.* 11 (10), 1596–1602. <http://dx.doi.org/10.1049/iet-cta.2016.1226>.
- Liang, G., Weller, S.R., Zhao, J., Luo, F., Dong, Z.Y., 2016. The 2015 Ukraine blackout: Implications for false data injection attacks. *IEEE Trans. Power Syst.* 32 (4), 3317–3318.
- López-Morales, E., Planta, U., Rubio-Medrano, C., Abbasi, A., Cardenas, A.A., 2024. SoK: Security of programmable logic controllers. In: 33rd USENIX Security Symposium (USENIX Security 24). pp. 7103–7122.
- Moore, B., 1981. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Trans. Autom. Control* 26 (1), 17–32.
- Nafees, M.N., Saxena, N., Cardenas, A., Grijalva, S., Burnap, P., 2023. Smart grid cyber-physical situational awareness of complex operational technology attacks: A review. *ACM Comput. Surv.* 55 (10), <http://dx.doi.org/10.1145/3565570>.
- New York Independent System Operator, 2025. Load data. <https://www.nyiso.com/load-data>. (Accessed 15 December 2024).
- Otomega, B., Van Cutsem, T., 2007. Undervoltage load shedding using distributed controllers. *IEEE Trans. Power Syst.* 22 (4), 1898–1907. <http://dx.doi.org/10.1109/TPWRS.2007.907354>.
- Pai, A., 1989. *Energy Function Analysis for Power System Stability*. Springer Science & Business Media.
- PowerWorld Corporation, 2025. Power flow: Voltage setpoint tolerance. URL https://www.powerworld.com/WebHelp/Content/MainDocumentation_HTML/PowerFlow%20Voltage%20Setpoint%20Tolerance.htm. (Accessed: 12 March 2024).
- Raju, V.R., Kumar, S., 2016. An optimal PMU placement method for power system observability. In: 2016 IEEE Power and Energy Conference At Illinois. PECE, pp. 1–5. <http://dx.doi.org/10.1109/PECE.2016.7459248>.
- Sun, Y., Yang, D., Gao, X., Qin, J., 2022. Robust strong structural controllability of complex power systems. *Front. Energy Res.* 10, 913893.
- Syrmakesis, A.D., Hatziaargyriou, N.D., 2024. Cyber resilience methods for smart grids against false data injection attacks: categorization, review and future directions. *Front. Smart Grids* 3, 1397380.
- Taha, A.F., Qi, J., Wang, J., Panchal, J.H., 2016. Risk mitigation for dynamic state estimation against cyber attacks and unknown inputs. *IEEE Trans. Smart Grid* 9 (2), 886–899.
- Wang, G., Giannakis, G.B., Chen, J., Sun, J., 2019. Distribution system state estimation: An overview of recent developments. *Front. Inf. Technol. Electron. Eng.* 20 (1), 4–17.
- Xiang, Y., Ding, Z., Zhang, Y., Wang, L., 2016. Power system reliability evaluation considering load redistribution attacks. *IEEE Trans. Smart Grid* 8 (2), 889–901.
- Xie, B., Peng, C., Yang, M., Kong, X., Zhang, T., 2021. A novel trust-based false data detection method for power systems under false data injection attacks. *J. Franklin Inst.* 358 (1), 56–73.
- Xu, Z., Ma, J., Gao, Y., Li, Y., Yu, H., Wang, L., 2023. Inertia identification and analysis for high-power-electronic-penetrated power system based on measurement data. *Energies* 16 (10), 4101.
- Yan, J., Tang, Y., He, H., Sun, Y., 2014. Cascading failure analysis with DC power flow model and transient stability analysis. *IEEE Trans. Power Syst.* 30 (1), 285–297.
- Zhao, C., Wang, X., Liu, Z., Liu, S., 2023. The observability and controllability metrics of power system oscillations and the applications. *Front. Energy Res.* 10, 1020894.