International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
SJIF (2025): 8.036

# An AI-Assisted Framework for Secure Frequency Regulation and Disturbance Mitigation in Renewable-Rich Interconnected Power Networks

**M. Ramakrishna Raju[1], N. Visali[2]**

[1]Research Scholar, Department of Electrical & Electronics Engg., JNTUA, Anantapur-515002, Andhrapradesh, India
Corresponding Author Email: *ramakimathephd[at]gmail.com*

[2]Senior Professor, Department of Electrical & Electronics Engg., JNTUA, Anantapur-515002, Andhrapradesh, India

**Abstract:** ***The increasing integration of renewable energy sources and the growing dependence on communication networks have introduced new challenges to reliable frequency regulation in interconnected power systems. Although conventional automatic generation control (AGC) strategies perform satisfactorily under normal operating conditions, their effectiveness can be affected by corrupted measurements, communication delays, and cyberattacks. This paper investigates an AI-assisted secure frequency regulation framework that combines a confidence-gated multi-stage fuzzy PID (AI-MSFPID) controller with measurement-trust assessment and signal reconstruction. The framework is evaluated using a deregulated three-area power system incorporating wind and solar generation, battery energy storage systems, aggregated electric vehicles, and flexible AC transmission system (FACTS) devices. A database of 2,430 nominal simulations establishes the reference dynamic performance, while cybersecurity is examined through 126 initial cases and 105 targeted revalidation cases involving different measurement and communication attacks. The initial assessment identifies denial-of-service (DoS) attacks as a major concern, with a mean cyber-resilience index (CRI) of 0.6361. In the subsequent targeted revalidation, all 105 cases exhibit stable responses, achieving an overall mean CRI of 0.8486 and a minimum value of 0.7500. The revalidated DoS cases record a mean CRI of 0.8357. However, replay attacks continue to produce an average residual dynamic degradation of 28.85%, indicating that maintaining stability does not necessarily eliminate the effects of cyber disturbances. As the initial and revalidated assessments involve different test conditions, their results are considered separately. The findings highlight the importance of combining intelligent control with reliable measurement processing and attack-specific assessment to support secure frequency regulation in renewable-rich interconnected power networks. Further investigation using independent validation and real-time implementation is needed to establish the effectiveness of the individual security mechanisms.***

**Keywords:** AI-assisted automatic generation control, cyber resilience, measurement reconstruction, renewable-integrated power networks

## 1. Introduction

The increasing integration of renewable energy sources and the growing complexity of interconnected power systems have made reliable frequency regulation an important challenge in modern power-system operation. Automatic generation control (AGC) plays a crucial role in maintaining the balance between generation and demand by restoring system frequency and regulating scheduled power exchanges between interconnected control areas. In deregulated electricity markets, this task becomes more demanding because generation companies (GENCOs) and distribution companies (DISCOs) participate in contractual power transactions that must be honoured while maintaining system stability. Variations in wind and solar generation further influence the power balance, increasing the need for responsive and coordinated control strategies [1, 2, 3, 4, 5].

Considerable research has been devoted to improving AGC performance through advanced control techniques and supplementary power-system technologies. Conventional PI and PID controllers have gradually been complemented by fuzzy logic, adaptive control, and intelligent optimisation methods to improve dynamic response under changing operating conditions. The integration of battery energy storage systems, electric vehicles, and renewable energy resources has also created new opportunities for supporting frequency regulation [6, 7, 8, 9]. In addition, investigations involving deregulated AGC and flexible AC transmission system (FACTS) devices have demonstrated the importance of contract-based scheduling and supplementary network compensation in regulating frequency deviations and tie-line power exchanges [1, 2, 3, 4, 10, 11, 12].

Despite these developments, reliable AGC operation increasingly depends on the availability and accuracy of measurements exchanged through communication networks. Frequency deviations and tie-line power measurements are essential inputs for calculating the area control error (ACE), which determines the corrective action of the secondary frequency controller. Any manipulation or interruption of these measurements can therefore affect the controller's decisions, even when the physical operating condition of the power system has not changed. Cybersecurity studies have identified false-data injection (FDI), denial-of-service (DoS), replay attacks, and communication delays as important threats to interconnected frequency regulation [5, 13, 14, 15].

These challenges highlight the need for AGC strategies that address not only dynamic performance under normal operating conditions but also the reliability of control decisions when measurement information becomes uncertain or compromised. Although intelligent controllers can adapt their response to varying system conditions, their effectiveness depends on the quality of the information supplied to them. A controller may provide satisfactory frequency regulation under normal operation yet respond


Licensed Under Creative Commons Attribution CC BY

incorrectly when corrupted measurements are interpreted as genuine disturbances.

Against this background, the present study investigates an AI-assisted secure frequency regulation framework that combines confidence-gated intelligent control with measurement-integrity assessment and signal reconstruction. The approach retains a multi-stage fuzzy PID controller as the baseline and permits AI-assisted gain adjustments only when the supervisory decision satisfies a prescribed confidence requirement. A separate measurement-security mechanism assesses the reliability of received feedback and supports the reconstruction of potentially compromised measurements.

The investigation focuses on whether these mechanisms can support stable frequency regulation under different cyberattack conditions and how their performance varies with the nature and severity of the disturbance. Particular attention is given to the distinction between detecting unreliable information and reducing its effects on the physical power system. The study therefore examines attack-specific cyber-resilience and residual dynamic degradation rather than relying exclusively on conventional settling-time or error-index comparisons.

**1) Research Questions and Contributions**

The growing dependence of automatic generation control (AGC) on communication networks and intelligent decision-making has introduced challenges that extend beyond conventional frequency regulation. While advanced controllers can improve system performance under normal operating conditions, their effectiveness may be compromised when measurements are corrupted or control decisions are based on unreliable information. This concern is particularly important in interconnected power systems, where an incorrect control action in one area can influence frequency and tie-line power exchanges in neighbouring areas.

A major challenge in developing secure AGC strategies is distinguishing genuine operating disturbances from abnormalities introduced through measurement and communication channels. Changes in load demand or renewable generation produce actual variations in the power balance, whereas cyberattacks may alter the information received by the controller without initially affecting the physical system. Both situations can influence the calculated area control error (ACE), making it necessary to assess measurement reliability before determining the appropriate corrective action.

The increasing use of AI-assisted controllers introduces another important consideration. Although intelligent algorithms can select control parameters according to prevailing operating conditions, their decisions may become uncertain when the system encounters unfamiliar disturbances or corrupted measurements. A practical intelligent control framework should therefore provide a means of limiting uncertain decisions while maintaining a dependable baseline control strategy.

In addition, detecting a cyberattack does not necessarily eliminate its effect on frequency regulation. Even when suspicious measurements are identified, the controller may continue to experience degraded performance unless suitable feedback information is available. Measurement reconstruction and the subsequent dynamic response must therefore be considered alongside attack detection.

Against this background, the present investigation addresses four principal research questions:

a) How can a confidence-gated AI supervisory mechanism regulate the application of adaptive controller gains while retaining the baseline MSFPID response when the learned decision is uncertain?
b) How can measurement-trust assessment and signal reconstruction be incorporated into the ACE feedback pathway to limit the influence of corrupted measurements?
c) How do different measurement and communication attacks, including false-data injection, replay, delay, denial of service, and tie-line manipulation, affect the regulation performance of selected controller–FACTS architectures?
d) What can be established from the initial cybersecurity assessment and targeted revalidation regarding system stability, residual dynamic degradation, and sensitivity to attack severity?

To investigate these questions, the study employs a renewable-integrated deregulated three-area power-system model incorporating conventional generation, wind and solar resources, battery energy storage systems, aggregated electric vehicles, and supplementary FACTS support. A balanced database of 2,430 nominal simulations provides the reference operating conditions for the cybersecurity investigation.

The security-focused assessment consists of 126 initial attack and reference cases involving selected controller–FACTS architectures and operating conditions. The observations obtained from this assessment are followed by 105 targeted revalidation cases examining replay attacks, communication delays, and denial-of-service conditions. This approach enables the investigation to identify weaknesses observed during the initial assessment and examine the behaviour of the security mechanisms under the selected revalidation scenarios.

The principal contribution of the present work is the integrated assessment of confidence-aware intelligent control, measurement-integrity processing, and cyberattack response within a common AGC framework. Rather than concentrating solely on conventional controller performance indices, the study considers how unreliable feedback information influences control decisions and the resulting physical response of the interconnected system.

A further contribution is the use of attack-specific revalidation to distinguish stable operation from effective mitigation. The reported results show that some cyber disturbances continue to produce measurable dynamic degradation even when the system remains stable. This highlights the importance of examining resilience indicators together with physical regulation performance.


Licensed Under Creative Commons Attribution CC BY

The investigation is based on a deterministic simulation model and the specified cyberattack profiles. Consequently, the findings describe the behaviour of the evaluated configurations under the investigated conditions. Although the framework incorporates confidence-gated control and measurement reconstruction, their individual contributions require further matched experimental comparisons before independent improvements can be established conclusively.
The nominal simulation database is used as a reference for the security analysis, while the principal emphasis of this paper remains on intelligent control under uncertain measurements, cyberattack assessment, and targeted revalidation. This provides a distinct research focus from related investigations concerned primarily with response-surface optimisation and comparative controller–FACTS performance.

## 2. System and Security Architecture

The proposed framework integrates intelligent automatic generation control (AGC), renewable energy resources, energy storage, flexible AC transmission system (FACTS) devices, and measurement-integrity assessment within a deregulated three-area interconnected power system. The architecture is designed to examine frequency regulation under both conventional operating disturbances and cyber-induced measurement abnormalities. It consists of three interconnected layers: the physical power-system and market model, the AI-assisted AGC and FACTS control layer, and the measurement-security layer. Figure 1 illustrates the secure-regulation information flow, while Figure 2 presents the interconnected three-area system and its contractual power-exchange arrangement.

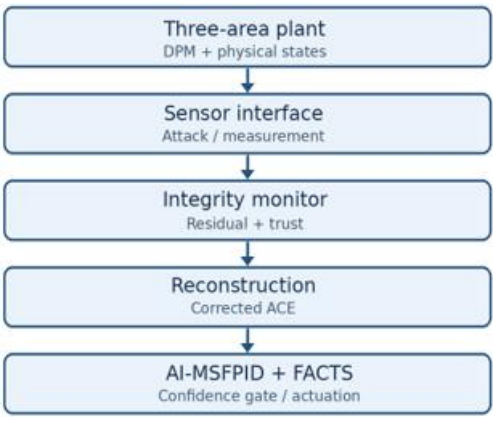


**Figure 1:** Conceptual secure-regulation information flow (schematic).

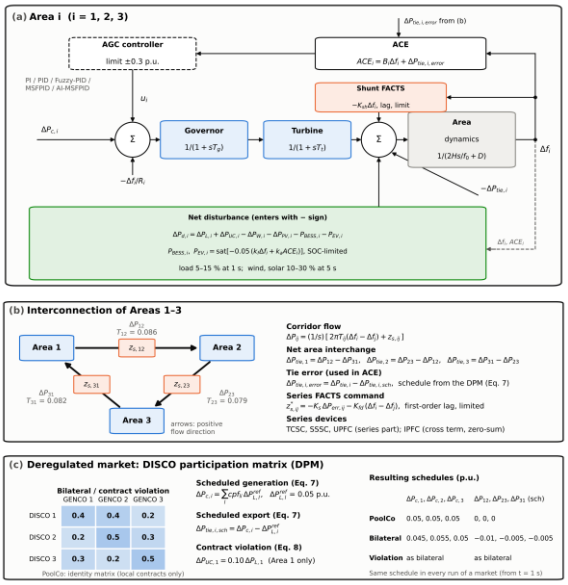


**Figure 2:** Three-area AGC system, interconnected tie-lines and deregulated participation matrix (source-project system model).

### 1) Contract-Aware Three-Area Regulation

The investigated system comprises three interconnected control areas, each represented by an equivalent conventional generating unit with a first-order speed governor, non-reheat turbine, and aggregate generator–load dynamics. Wind and solar generation introduce renewable power fluctuations, while battery energy storage systems (BESS) and aggregated electric vehicles (EVs) provide supplementary active-power support. The three areas exchange power through interconnected tie-lines, with secondary frequency regulation maintaining system frequency and scheduled interchange.
Deregulated electricity transactions are represented through a DISCO participation matrix (DPM), which defines the contractual relationships between generation companies (GENCOs) and distribution companies (DISCOs). Three market scenarios are considered: PoolCo, bilateral trading, and contract violation. In the reference PoolCo arrangement, generation is contracted locally without scheduled cross-area interchange. Bilateral trading permits contractual imports and exports between areas, whereas contract violation introduces additional uncontracted demand without modifying the original bilateral schedule [1–3].

A distinction is maintained between scheduled and actual tie-line power. The area control error (ACE) is determined using the deviation of actual interchange from its scheduled value together with the local frequency deviation. This ensures that legitimate contractual transfers are not interpreted as regulation errors. It is particularly important for cybersecurity assessment because manipulation of measured tie-line power can alter the calculated ACE even when the contractual schedule and physical power balance remain unchanged.

The dynamic response of each area depends on mechanical generation, local demand, renewable power variations, storage support, and net tie-line exchange. The governor–turbine system provides conventional generation control, while BESS and aggregated EVs contribute supplementary power during disturbances. In the nominal operating


Licensed Under Creative Commons Attribution CC BY

configuration, BESS and EV participation are each maintained at 5% in every area.

The model therefore provides a common physical and contractual basis for examining both genuine operating disturbances and cyber-induced abnormalities. Load and renewable fluctuations alter the actual power balance, whereas measurement attacks may initially affect only the information supplied to the controller. Their different origins must be considered when evaluating the resulting frequency and tie-line responses.

**2) AI-Assisted AGC and FACTS Coordination**

Five AGC strategies are considered in the overall simulation framework: proportional–integral (PI), proportional–integral–derivative (PID), fuzzy-PID, multi-stage fuzzy PID (MSFPID), and AI-assisted MSFPID (AI-MSFPID). The conventional controllers provide reference strategies, while the MSFPID structure serves as the baseline for AI-assisted supervisory control.

The AI-MSFPID controller introduces bounded adjustments to the proportional, integral, and derivative components of the baseline MSFPID law. Its supervisory policy was developed using 729 teacher simulations. Candidate classification approaches included k-nearest neighbours (KNN), support vector machine error-correcting output codes (SVM-ECOC), and bagged decision trees. KNN was retained in the reported implementation.

A confidence-gating mechanism determines whether the AI-recommended gain adjustment is accepted. When the classifier confidence satisfies the prescribed threshold of 0.6, the bounded adjustment is applied. Otherwise, the controller reverts to the unmodified MSFPID configuration. The reported fallback fraction is approximately 23.46%, indicating that the baseline controller was selected whenever the supervisory decision did not meet the confidence requirement.

This arrangement provides a defined control response under uncertain AI decisions. However, classifier confidence reflects the supervisory decision mechanism and should not be interpreted as a direct probability of closed-loop safety. Furthermore, returning to the baseline controller does not independently protect the system against corrupted feedback measurements. Measurement reliability is therefore addressed through a separate security layer.

Supplementary network support is provided through six FACTS configurations: No FACTS, thyristor-controlled series capacitor (TCSC), static synchronous series compensator (SSSC), static synchronous compensator (STATCOM), unified power flow controller (UPFC), and interline power flow controller (IPFC). These configurations influence the dynamic response through their respective series, shunt, or coordinated compensation channels.

FACTS devices affect the physical response of the interconnected network but are not treated as independent cyberattack-detection mechanisms. Selected controller–FACTS combinations are instead considered distinct closed-loop architectures for evaluating frequency regulation under compromised measurement and communication conditions.

**3) Measurement-Integrity Assessment and ACE Reconstruction**

The measurement-security layer is positioned between the received feedback signals and the AGC controller. Its purpose is to assess the consistency of measurement information and limit the influence of unreliable feedback on secondary frequency regulation.

The investigated attack mechanisms include false-data injection (FDI), replay attacks, communication delays, denial of service (DoS), and tie-line measurement manipulation. These disturbances affect the feedback pathway in different ways. FDI introduces incorrect values, replay attacks substitute previously recorded information, communication delays reduce measurement timeliness, and DoS interrupts measurement availability. Tie-line manipulation changes the perceived interchange used in ACE calculation.

The security mechanism evaluates discrepancies between the received ACE and a reference trajectory generated from the corresponding attack-free simulation. Residual-based trust indicators determine the extent to which received information is retained or replaced by the reference signal. The reconstructed ACE is subsequently supplied to the controller. This approach separates two forms of uncertainty. Physical uncertainty arises from variations in load demand and renewable generation, while information uncertainty results from measurement manipulation or communication disruption. A genuine operating disturbance requires corrective generation and storage support; an unreliable measurement requires assessment of the feedback information before it is used for control.

The AI supervisory confidence and measurement-trust indicators serve different purposes. The former determines whether an AI-recommended gain adjustment is accepted, whereas the latter assesses the reliability of information entering the AGC feedback pathway. These mechanisms are complementary but are not interchangeable.

The cybersecurity assessment comprises 126 initial attack/reference cases and 105 targeted revalidation cases across selected controller–FACTS architectures and operating conditions. The initial investigation examines several attack mechanisms, while the targeted revalidation focuses on replay, delay, and DoS profiles. Attack-free reference responses provide the basis for interpreting cyber-induced deviations.

The reconstruction mechanism uses a stored attack-free reference trajectory. Consequently, the demonstrated approach is reference-assisted simulation reconstruction rather than an independently validated online state estimator. This distinction is maintained when interpreting the cybersecurity results.

**4) Mathematical Formulation of Measurement Trust and Reconstruction**

The following discrete-time expressions reproduce the measurement and command processing in


Licensed Under Creative Commons Attribution CC BY

agc_simulate_cyber_research_case_v87d.m. They use the stored attack-free reference trajectory; hence the demonstrated reconstruction is reference-assisted simulation reconstruction, not an independently validated live-state estimator.

$$r_m(k) = \|ACE_a(k) - ACE_r(k)\|_2 \quad (1)$$

$$r\mathcal{d}(k) = \|[ACE_a(k) - ACE_a(k-1)] - [ACE_r(k) - ACE_r(k-1)]\|_2 \quad (2)$$

$$\sigma_m = \max(TrustScale, 10^{-6}) \quad (3)$$

$$T_a(k) = clip_{(0,1)}\{\exp[-r_m(k)/\sigma_m]\} \quad (4)$$

$$T\mathcal{d}(k) = clip_{(0,1)}\{\exp[-r\mathcal{d}(k)/\max(0.002, 0.5\sigma_m)]\} \quad (5)$$

$$T_m(k) = 0 \text{ (active DoS); } \min(T_a, T\mathcal{d}) \text{ (active replay/delay); } T_a \text{ (otherwise)} \quad (6)$$

$$\alpha(k) = \min\{1, T_m(k)/\max(T_{th}, eps)\} \quad (7)$$

$$ACE_u(k) = \alpha(k)ACE_a(k) + [1-\alpha(k)]ACE_r(k) \quad (8)$$

$$ACE_u(k) = ACE_r(k) \text{ (mitigated active DoS); } ACE_a(k) \text{ (no mitigation)} \quad (9)$$

$$r_u(k) = \|u_a(k) - u_r(k)\|_2 \quad (10)$$

$$\sigma_u = \max(0.02, 10\sigma_m);\ T_u(k) = clip_{(0,1)}\{\exp[-r_u(k)/\sigma_u]\} \quad (11)$$

$$\beta(k) = \min\{1, T_u(k)/\max(T_{th}, eps)\} \quad (12)$$

$$u_u(k) = \beta(k)u_a(k) + [1-\beta(k)]u_{sa}\not{f}_e(k) \quad (13)$$

$$Tc(k) = \min[T_m(k), T_u(k)] \quad (14)$$

$$k\mathcal{d}_{et} = \text{first k in attack interval with } Tc(k) < T_{th} \quad (15)$$

$$t\mathcal{d}_{et} = \max\{0, t(k\mathcal{d}_{et}) - t_{start}\} \quad (16)$$

$$DR = (1/N_a)\ \Sigma_{k\in a}\ 1[Tc(k) < T_{th}]\backslash t \quad (17)$$

Here $ACE_a$ is the attacked/received three-area ACE vector, $ACE_r$ is the archived attack-free reference ACE, and $ACE_u$ is the ACE passed to the controller. The dynamic residual is zero at the first sample. $clip_{(0,1)}$ denotes saturation to [0,1]. $T_m$ and $T_u$ denote measurement and command trust; $Tc$ is their minimum. $u_a$ is the attacked command, $u_r$ is the reference command and $u_{sa}\not{f}_e$ is the internally calculated controller output using $ACE_u$.

Equations (7)–(9) apply only while an attack is active and mitigation is enabled. For active DoS, the implementation substitutes $ACE_r$ directly. The command equations apply when the active attack type includes COMMAND or HYBRID; otherwise command trust is set to unity. If no threshold crossing occurs, the code records infinite detection delay; for active DoS it explicitly assigns zero detection delay. The detection rate is the fraction of sampled attack-interval points below the trust threshold, not a case-level classification accuracy.

The source implementation uses a profile-specific trust threshold and trust scale; their numerical values should be read from the actual selected attack profile rather than assumed to be universal. The AI classifier confidence threshold of 0.6 is separate from measurement trust. This reference-based formulation must not be described as demonstrated autonomous online reconstruction.

## 3. Simulation Methodology and Validation Design

The proposed secure frequency-regulation framework is evaluated through a structured simulation procedure involving attack-free operating conditions, an initial cybersecurity assessment, and targeted revalidation of selected cyberattack scenarios. The investigation is conducted using the renewable-integrated deregulated three-area AGC model described in Section II. The evaluation considers both physical regulation performance and the behaviour of the measurement-security mechanisms under compromised feedback conditions.

The simulation methodology is organised into three stages. First, the nominal system is examined under different operating patterns, market scenarios, controller structures, and FACTS configurations to establish the reference dynamic behaviour. Second, an initial cybersecurity assessment is performed using selected controller–FACTS architectures and attack conditions. Finally, targeted revalidation is conducted to investigate the response to replay attacks, communication delays, and denial-of-service (DoS) conditions identified as requiring further examination.

### 1) Attack-Free Reference and Numerical Settings

The nominal investigation follows the V8.8 simulation framework, comprising 27 operating patterns, three deregulated market scenarios, five AGC controllers, and six FACTS configurations. The complete experimental arrangement therefore consists of

$$N_{\text{nominal}} = 27 \times 3 \times 5 \times 6 = 2430$$

deterministic simulations.

The three market scenarios are PoolCo, bilateral trading, and contract violation. The five controllers are PI, PID, Fuzzy-PID, MSFPID, and AI-MSFPID, while the six network-support configurations are No FACTS, TCSC, SSSC, STATCOM, UPFC, and IPFC.

The simulations are performed over a 100 s interval using a fixed integration step of 0.02 s and the fourth-order Runge–Kutta numerical method. These settings provide a common basis for comparing the dynamic responses under different operating conditions.

The nominal simulations establish the reference frequency deviations, tie-line power errors, ACE responses, and dynamic performance indices. Although the archived results report stable and settled nominal responses, they do not independently establish the effectiveness of the cybersecurity mechanisms. Furthermore, a separate numerical-convergence investigation using smaller integration steps is not documented in the available results.

For the cybersecurity assessment, seven representative controller–FACTS architectures are selected from the nominal configurations. These architectures are examined under three operating conditions and six cyber states, including an attack-free reference. The initial assessment therefore comprises

$$N_{\text{initial}} = 7 \times 3 \times 6 = 126$$

simulation cases.

This arrangement provides a consistent basis for examining the influence of controller architecture, operating condition, and attack mechanism. However, it represents a selected subset of the overall system configurations rather than the complete nominal factorial design.


Licensed Under Creative Commons Attribution CC BY

**2) Initial Cybersecurity Assessment**

The initial cybersecurity assessment investigates the response of selected controller–FACTS architectures to measurement and communication disturbances. The attack scenarios include false-data injection (FDI), replay attacks, communication delays, denial of service (DoS), and tie-line measurement manipulation, together with the corresponding attack-free reference conditions.

The assessment considers how compromised measurements influence the regulation process and how the implemented security mechanism affects the resulting dynamic response. The archived results are summarised using the cyber-resilience index (CRI), the worst observed CRI, and the mean mitigated composite dynamic performance index.

**Table 1:** Initial Cyber-Resilience Assessment (126 Cases)

| Attack type | Cases | Mean CRI | Worst CRI | Mean mitigated ({}) |
|---|---|---|---|---|
| FDI | 42 | 0.7462 | 0.6588 | 13.6613 |
| Replay | 21 | 0.8097 | 0.6450 | 8.8930 |
| Delay | 21 | 0.8494 | 0.8482 | 7.8906 |
| DoS | 21 | 0.6361 | 0.5851 | 11.6923 |

Note: The remaining 21 cases correspond to attack-free reference conditions. The archived FDI grouping contains 42 cases.

The initial assessment reveals differences in the reported resilience of the investigated attack groups. Communication delay produces the highest mean CRI of 0.8494, while DoS records the lowest mean CRI of 0.6361. The DoS group also exhibits the lowest worst-case CRI of 0.5851, indicating that measurement unavailability is a significant concern within the investigated configurations.

Replay attacks produce a mean CRI of 0.8097 but a lower worst-case value of 0.6450, suggesting that some investigated conditions are more vulnerable than the group average indicates. The FDI group records a mean CRI of 0.7462, with a mean mitigated dynamic index of 13.6613.

These observations provide the basis for selecting attack mechanisms requiring further investigation. However, the reported aggregate indices do not establish the independent contribution of each security component or provide complete information about attack-detection latency and reconstruction accuracy.

The initial assessment is therefore treated as a systematic cybersecurity audit of the simulated system rather than as independent validation against real-world cyberattacks.

**3) Audit-Led Targeted Revalidation**

Based on the initial cybersecurity assessment, targeted revalidation is conducted to examine replay attacks, communication delays, and DoS conditions in greater detail. The revalidation retains the seven selected controller–FACTS architectures and three operating conditions while considering five attack profiles.

The investigated profiles comprise a 5 s replay condition, communication delays of 0.2, 0.5, and 1.0 s, and a DoS condition. The resulting experimental arrangement consists of

$$N_{\text{revalidation}} = 7 \times 3 \times 5 = 105$$

simulation cases.

The use of three delay durations enables the investigation to examine how the reported trust-related and dynamic performance indicators vary with increasing communication delay. The replay and DoS profiles provide additional evidence about the behaviour of the measurement-security mechanism under outdated or unavailable feedback information.

The targeted revalidation results are presented in Table 2.

**Table 2:** Targeted Cybersecurity Revalidation Results (105 Cases)

| Attack profile | N | Mean (J_{}) | Mean CRI | Worst CRI | Mean DCRI | Mean TCI | Degradation (%) |
|---|---|---|---|---|---|---|---|
| Replay, 5 s | 21 | 8.8914 | 0.8550 | 0.7951 | 0.2612 | 0.8669 | 28.85 |
| Delay, 0.2 s | 21 | 7.8906 | 0.8494 | 0.8482 | 1.0000 | 0.9962 | −1.78 |
| Delay, 0.5 s | 21 | 7.7190 | 0.8485 | 0.8454 | 1.0000 | 0.9902 | −4.36 |
| Delay, 1.0 s | 21 | 7.6377 | 0.8541 | 0.8408 | 1.0000 | 0.9793 | −4.68 |
| DoS | 21 | 8.0080 | 0.8357 | 0.7500 | 0.8571 | 1.0000 | 0.00 |

The revalidation results indicate that all 105 investigated cases maintain stable responses under the specified simulation conditions. The overall mean CRI is 0.8486, with a minimum observed value of 0.7500.

The replay profile records a mean CRI of 0.8550 but retains an average residual dynamic degradation of 28.85%. This finding is important because it demonstrates that a stable response and a comparatively high resilience index do not necessarily imply complete mitigation of the attack's dynamic consequences.

For communication delays, the mean trust-related index decreases from 0.9962 at 0.2 s to 0.9793 at 1.0 s. This trend is consistent with increasing temporal mismatch between the received and reference measurements. However, the mean CRI does not vary monotonically with delay duration, indicating that the different performance indicators capture distinct aspects of the system response.

The DoS revalidation group records a mean CRI of 0.8357 and a minimum value of 0.7500. Although these values are higher than those obtained for the initial DoS group, the two assessments involve different test definitions and should not be interpreted as a paired before-and-after estimate of mitigation effectiveness.

Similarly, the negative degradation percentages reported for the delay profiles are retained according to the original project convention. Their interpretation depends on the exact reference and normalization used in the archived calculations. The targeted revalidation therefore provides additional evidence about the behaviour of the investigated architectures under selected cyberattack profiles, while leaving broader


Licensed Under Creative Commons Attribution CC BY

generalisation and component-level causal attribution for further investigation.

**4) Performance Metrics and Validation Considerations**

The performance assessment considers conventional AGC dynamic measures together with cybersecurity-related indicators. The principal dynamic measure is the composite performance index (J_{}), which combines time-weighted frequency error, squared frequency error, peak frequency and tie-line deviations, control effort, and storage-related penalties.

The cybersecurity assessment additionally reports the cyber-resilience index (CRI), detection-related index (DCRI), trust-related index (TCI), and residual dynamic degradation. These indicators provide complementary information about the system's response to compromised measurements.

The CRI is used to summarise the resilience of the evaluated configuration under the source implementation's definition. DCRI and TCI provide additional information about detection-related and measurement-trust behaviour. Their numerical values are interpreted according to the archived project definitions rather than treated as universally standardised cybersecurity measures.

An important consideration is the distinction between numerical stability, attack detection, measurement reconstruction, and mitigation effectiveness. A system may remain stable while exhibiting substantial performance degradation. Similarly, an attack may be detected without completely restoring the accuracy of the feedback information or eliminating the resulting frequency deviations. The evaluation therefore considers the reported indices collectively rather than relying on a single measure to establish cybersecurity effectiveness.

For meaningful comparison, the physical response under an attack should be evaluated against an attack-free reference involving the same operating condition, contractual schedule, controller configuration, and supplementary support settings. This distinction is particularly important because changes in load and renewable generation may produce significant frequency deviations even in the absence of a cyberattack.
The archived cybersecurity results provide aggregate performance values for the investigated attack groups. However, complete case-level measurement traces, detection timestamps, and reconstruction-error trajectories are not reproduced in the available summary. Consequently, no additional detection-delay distributions, confidence intervals, or reconstruction-accuracy improvements are inferred.

The initial cybersecurity assessment and targeted revalidation also represent different experimental sets. Although both investigate selected controller–FACTS architectures, their attack profiles and conditions are not fully matched. Differences between their group averages therefore cannot be interpreted as direct causal estimates of improvement.

The simulation evidence is considered at three levels. The nominal investigation establishes the attack-free dynamic reference. The initial cybersecurity assessment identifies the relative weaknesses of the investigated attack mechanisms. The targeted revalidation provides additional evidence about stability, resilience indicators, and residual degradation under selected attack profiles.

Further validation would require matched comparisons with and without individual security mechanisms, numerical-convergence studies, independent operating conditions, and real-time or hardware-in-the-loop testing. These investigations would help establish the separate contributions of AI confidence gating, measurement-trust assessment, and signal reconstruction.

The present methodology provides a structured basis for evaluating secure frequency regulation in the investigated three-area simulation environment while maintaining a clear distinction between demonstrated results and validation requirements that remain open.

## 4. Security-Focused Results and Discussion

**1) What the initial audit revealed**

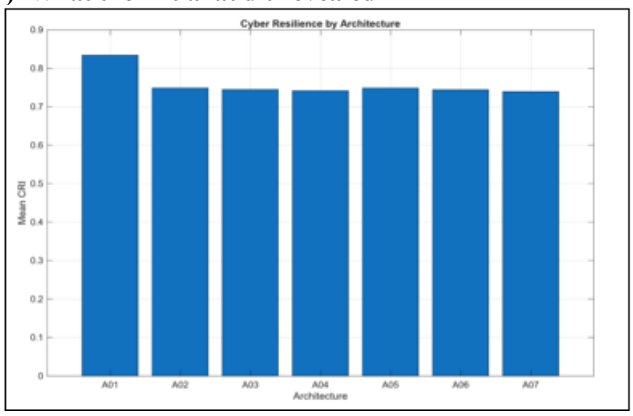


**Figure 3:** Cyber-resilience index for seven selected architectures (archived plot)

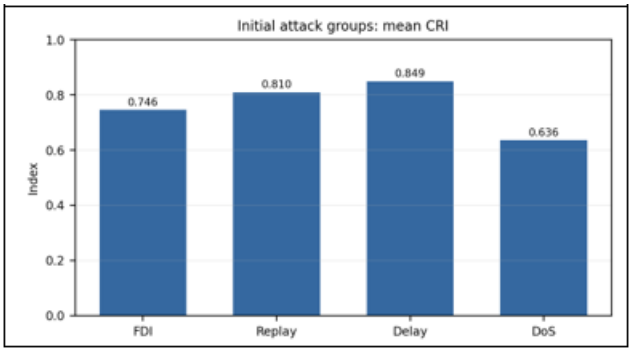


**Figure 4:** Mean cyber-resilience index by attack type in the initial 126-case assessment (source-summary values).

The initial 126-case summary indicates strong attack dependence. DoS has the lowest reported mean CRI (0.6361; minimum 0.5851), followed by FDI (mean 0.7462). Replay has mean 0.8097, and delay has mean 0.8494 in its initial tested conditions. The relative values motivate a more targeted investigation, but they cannot establish that any one mechanism is universally most damaging across different attack severities.

The initial DoS group has the lowest reported mean CRI and worst CRI among the initial attack types. This identifies DoS as a priority for further testing within the initial case


Licensed Under Creative Commons Attribution CC BY

definitions. FDI also produces a lower mean index than the initial replay and delay groups, consistent with the possibility that direct manipulation of feedback can disrupt regulation. However, the attack profiles differ in mechanism and intensity, so their means should not be interpreted as an intrinsic universal severity ranking.

The initial mitigated Jdynamic also differs across attack categories. The FDI group has a reported mean of 13.6613, replay 8.8930, delay 7.8906 and DoS 11.6923. These quantities reflect both the attacked regulation response and the chosen scenario mix. They should be interpreted alongside the resilience index, rather than as a substitute for it.

**2) Revalidated replay, delay and DoS behaviour**

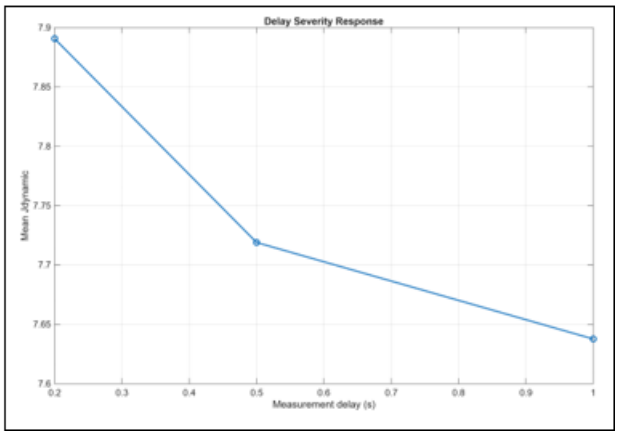


**Figure 5:** Dynamic-index response to measurement delay (archived plot).

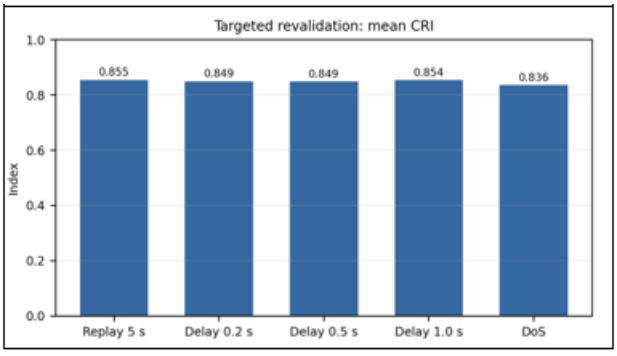


**Figure 6:** Mean cyber-resilience index across the five targeted revalidation profiles (21 cases per profile).

All 105 targeted mitigated cases are reported stable, with overall mean CRI 0.8486 and minimum 0.7500. The DoS revalidation group has mean CRI 0.8357, compared with 0.6361 in the earlier audit; this cross-set difference must not be presented as a controlled treatment effect. Replay at 5 s has mean CRI 0.8550 but retains 28.85% mean residual degradation. Thus detection or stable operation does not imply elimination of attack-induced performance loss.
Across the revalidated delay profiles, the reported TCI decreases from 0.9962 at 0.2 s to 0.9902 at 0.5 s and 0.9793 at 1.0 s. This monotonic trend is consistent with an indicator sensitive to delay severity within these tested cases. It does not establish calibration for arbitrary network delays.

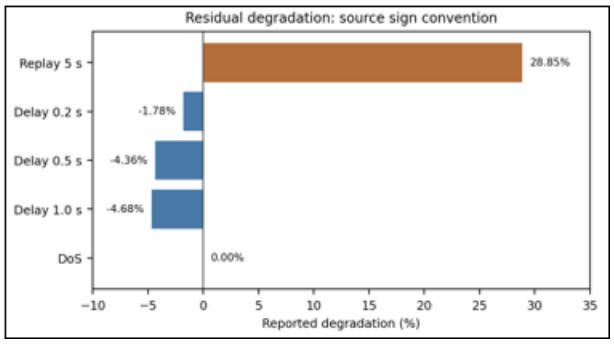


**Figure 7:** Reported residual dynamic degradation in targeted revalidation; negative values retain the source sign convention.

The revalidated replay group reports mean CRI 0.8550, minimum 0.7951, mean DCRI 0.2612 and mean TCI 0.8669. Its mean residual degradation of 28.85% is particularly important: a system may detect a replayed signal yet continue to exhibit a substantial performance penalty. This is evidence against using detection success alone as a sufficient cybersecurity metric.

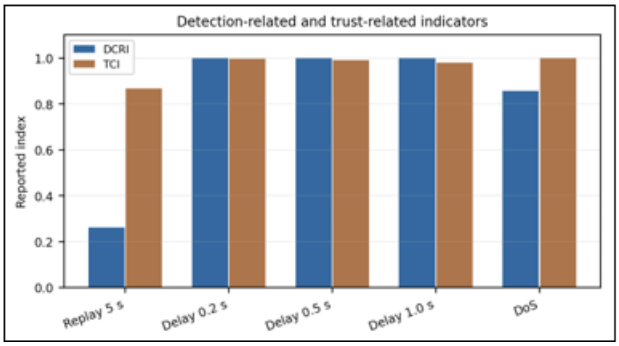


**Figure 8:** Detection-related (DCRI) and trust-related (TCI) indicators across targeted attack profiles.

For the three delay profiles, the mean CRI values remain near 0.85, whereas the reported TCI values decrease with longer delay. This makes the TCI trend more informative about severity within the tested delay range than the aggregate CRI alone. The reported negative degradation percentages for the delay cases must be interpreted according to the project's degradation sign convention; they should not automatically be called physical improvements without reviewing the calculation.

The revalidated DoS group has mean CRI 0.8357, minimum 0.7500, mean DCRI 0.8571 and mean TCI 1.0000. Its reported mean mitigated Jdynamic is 8.0080. These results show the performance of the revised mitigation implementation on the targeted DoS profile, but do not isolate which detector, reconstruction or controller component produced the outcome.


Licensed Under Creative Commons Attribution CC BY

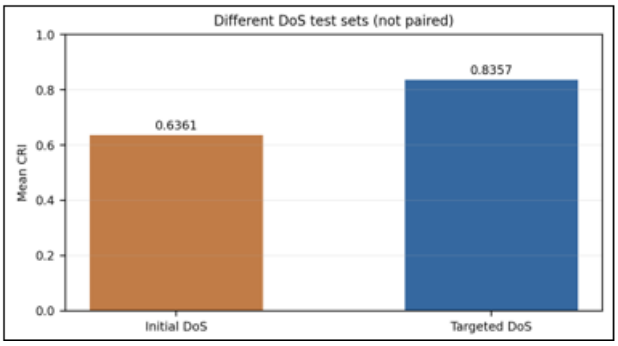


**Figure 9:** Reported mean DoS cyber-resilience indices in the initial and targeted test sets; these are not matched before-and-after observations

The distinction between initial and targeted results is especially important for DoS. The initial DoS group contains 21 cases with mean CRI 0.6361 and minimum 0.5851, whereas the targeted DoS profile contains 21 cases with mean CRI 0.8357 and minimum 0.7500. Both summaries are relevant to the engineering narrative: the first motivates further audit, and the second documents the reported behavior under revalidation. They cannot be subtracted as though they were a paired experiment unless the operating conditions, attack definitions and mitigation settings are shown to match.

Replay presents a different difficulty. A replayed measurement can remain smooth and within an expected range even though it is temporally stale. The reported targeted replay CRI of 0.8550 does not remove the importance of the 28.85% residual dynamic degradation. A suitable figure would therefore overlay the true measurement, replayed signal and reconstructed value over the attack interval, followed by the resulting frequency and ACE. Without that trace, the aggregate table supports only a qualified statement about residual degradation.

For delayed feedback, the reported trust-related indicator decreases as delay rises from 0.2 to 1.0 s. The direction is informative, but the numerical difference is small and the available data do not establish a universal detection threshold. A delay may interact with the derivative filter, controller gains and network support; an evaluation of severity discrimination should therefore include both the indicator trend and the resulting dynamic response. The source summary supports the former and provides aggregate dynamic indices for the latter.

### 3) Architecture-dependent cyber response

**Table 3:** Architecture-Wise Cyber Revalidation Summary.

| ID | Controller | FACTS | Mean CRI | Worst CRI | Mean DCRI | Mean TCI | Degradation (%) |
|---|---|---|---|---|---|---|---|
| A01 | PI | No FACTS | 0.8309 | 0.7500 | 0.6000 | 0.9658 | -6.58 |
| A02 | MSFPID | No FACTS | 0.8512 | 0.8086 | 0.8617 | 0.9664 | 3.92 |
| A03 | AI-MSFPID | No FACTS | 0.8496 | 0.7951 | 0.8632 | 0.9653 | 5.94 |
| A04 | AI-MSFPID | SSSC | 0.8493 | 0.7951 | 0.8609 | 0.9672 | 6.07 |
| A05 | AI-MSFPID | STATCOM | 0.8603 | 0.8027 | 0.8624 | 0.9649 | 4.72 |
| A06 | AI-MSFPID | UPFC | 0.8498 | 0.8027 | 0.8589 | 0.9678 | 4.91 |
| A07 | AI-MSFPID | IPFC | 0.8488 | 0.7951 | 0.8586 | 0.9682 | 6.27 |

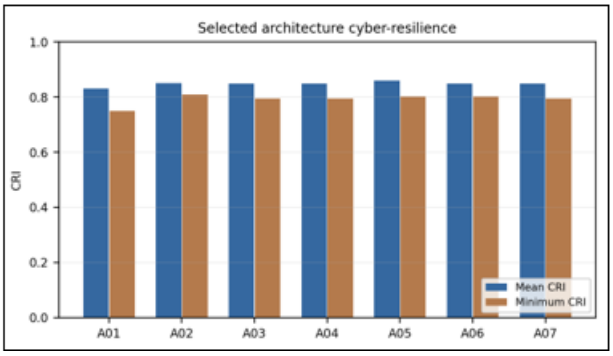


**Figure 10:** Mean and minimum CRI for the seven selected controller–FACTS architectures after targeted revalidation

The seven architectures have reported mean CRI values between 0.8309 and 0.8603. AI-MSFPID with STATCOM has the largest mean CRI among these selected architectures, while AI-MSFPID with IPFC has the largest reported mean TCI. Differences are modest and conditional on the chosen cases. The best attack-free dynamic configuration need not be the best configuration under measurement attacks.

The architecture comparison is based on seven representative pairings rather than all 30 controller–FACTS combinations under cyberattack. Its reported CRI spread is relatively narrow, so small differences among architectures should not be overinterpreted as robust rankings. The network-support configuration may influence the physical trajectory used by the detector and reconstruction layer, but the existing summary does not prove that FACTS devices independently improve cyberattack detection.

An informative next experiment would fix each disturbance pattern, market, attack onset, attack amplitude and controller–FACTS pair while switching security modules on and off. The baseline, confidence-only, detection-only and full-reconstruction cases would then yield matched response differences. This would distinguish the effect of safe AI fallback from the effect of reconstructed measurements and reveal whether their combination provides more than either component separately.

A first ablation would hold the controller and plant fixed while enabling and disabling the confidence gate. The response variables would include fallback frequency, Jdynamic, peak frequency deviation, peak tie-line error and ACE. This would show whether accepting a learned correction under attack changes the regulation outcome compared with remaining on MSFPID. It would also identify cases where fallback is triggered but the underlying measurements are still corrupted.

A second ablation would hold the controller policy fixed while enabling and disabling measurement reconstruction. Under the same attack trace, the experiment would compare


Licensed Under Creative Commons Attribution CC BY

the received signal, reconstructed signal and attack-free reference. Reconstruction RMSE, detection delay and the resulting ACE trajectory would jointly indicate whether the security layer improves information quality and whether that improvement reaches the plant output. Neither of these ablations can be reconstructed reliably from the current aggregate CRI table alone.

A third ablation would compare detection-only and detection-plus-reconstruction configurations. A detector may correctly identify a corrupted measurement yet leave the controller operating on a held or stale value. If reconstruction improves the ACE trajectory, this should be demonstrated with matched time-domain plots and the same attack realization. This is especially important for replay and DoS, where a signal may remain numerically plausible despite being temporally invalid.

Finally, the architecture comparison should examine sensitivity to the controller and FACTS configuration after the security module is held fixed. This would separate physical disturbance rejection from the measurement-security contribution. The seven selected architectures provide an initial sample, but the complete 30-cell controller–FACTS design under attack would require additional simulations before a universal architecture ranking could be justified.

**4) Limits of the evidence and ablation requirements**

The current archive supports attack-response comparisons and targeted revalidation, but the aggregated tables do not provide a complete four-way matched ablation separating (1) baseline MSFPID, (2) confidence-gated AI without reconstruction, (3) detection without reconstruction and (4) the complete reconstructed-feedback framework. Such a test is required to attribute a measured resilience gain specifically to the confidence gate or reconstruction layer. It should use identical attack realizations, operating patterns, market schedules and controller–FACTS choices. Figures 5–10 are descriptive visualizations of the archived summary values; they do not represent new MATLAB simulations or matched causal estimates.

Independent validation should also inspect the correspondence between the reported attack labels and the actual modified signals in the MATLAB implementation. For example, replay should use a previously recorded measurement segment, delay should preserve the intended lag history, and DoS should specify whether measurements are held, dropped or reconstructed. These implementation choices materially change the meaning of a cyber-resilience result and should be documented in a reproducibility appendix.

For numerical verification, representative attack-free and attacked runs should be repeated at smaller integration steps while holding attack timing fixed in physical seconds. Agreement of peak frequency, tie-line error, settling time and resilience metrics would support numerical convergence. For external validation, the measurement-security layer should be tested against independent attack traces or a hardware-in-the-loop platform. These are recommended validation steps, not results already established by the current archive.

The published literature establishes that AGC is sensitive to feedback quality, but its numerical results are often based on different models, operating assumptions and cyberattack definitions. A result from a two-area model with one attack type cannot be compared directly with a three-area deregulated model containing scheduled interchange and multiple support devices. The proper literature comparison therefore identifies the threat model, security mechanism, controller assumptions and validation evidence, rather than ranking studies solely by their reported settling times.

The present paper contributes an explicit audit-and-revalidation narrative. This is valuable because an initial mitigation claim may be undermined by poor detectability or insufficient severity discrimination. A revalidation stage can reveal whether a modified implementation behaves consistently under selected profiles. However, the resulting performance should still be described as conditional on those profiles, not as a guarantee against adaptive adversaries or unseen attacks.

In the context of related work by the same authors, the nominal 2,430-run data and plant model are shared infrastructure. Their reuse is appropriate for defining reference conditions if disclosed, but the paper should not claim that the nominal controller ranking is a second independent discovery. The cyberattack-specific results and analysis should remain the main contribution and be identified explicitly in the cover letter.

**5) Relationship to previous research**

The deregulated-market modelling follows the established DPM literature [1, 2, 3, 4]; the role of FACTS and storage is motivated by earlier LFC studies [10, 11, 12]. The security emphasis differs from attack-specific FDI and DoS investigations [5, 13, 14, 15] by jointly documenting confidence gating, measurement reconstruction and an audit/revalidation sequence. However, differences in power-system models and attack profiles preclude direct numerical claims of superiority over those publications.

The findings also clarify the relationship between the cyber paper and related nominal AGC investigations. The common plant model and baseline controller data are methodological context. The independent focus here is the security-layer logic, attack-specific audit, targeted revalidation and interpretation of residual regulation degradation. This boundary should be disclosed to the editor, especially where nominal numerical summaries overlap with other submitted manuscripts.

An engineering consequence of the results is that the best nominal controller–FACTS pairing need not maximize the selected cyber-resilience index. The nominal results emphasize frequency recovery and error accumulation; the cyber results emphasize continued regulation with unreliable information. These objectives can favor different architectures, and the reported seven-architecture comparison should not be extended to all 30 possible pairings without new attack simulations.

The study also illustrates why a security claim requires more than one number. A high CRI may coexist with measurable


Licensed Under Creative Commons Attribution CC BY

replay degradation; a monotonic delay indicator may coexist with only modest changes in the aggregate resilience score. A transparent paper should report the security indicator together with the physical response metric, the attack definition and the tested mitigation mechanism. Such joint reporting is a stronger contribution than simply asserting that the controller is cyber-resilient.

## 5. Conclusions

This study investigated an AI-assisted secure frequency regulation framework for a renewable-integrated, deregulated three-area interconnected power system. The proposed approach combines a confidence-gated AI-MSFPID controller with measurement-integrity assessment and reference-assisted signal reconstruction to examine the behaviour of automatic generation control under different cyberattack conditions. The investigation was supported by 2,430 nominal simulations, 126 initial cybersecurity assessment cases, and 105 targeted revalidation cases involving selected controller–FACTS architectures.

The initial cybersecurity assessment identified denial-of-service (DoS) attacks as a significant concern, with a mean cyber-resilience index (CRI) of 0.6361 and a minimum value of 0.5851. In the subsequent targeted revalidation, all 105 investigated cases maintained stable responses, achieving an overall mean CRI of 0.8486 and a minimum value of 0.7500. The revalidated DoS group recorded a mean CRI of 0.8357. However, replay attacks continued to produce an average residual dynamic degradation of 28.85%, demonstrating that stable operation does not necessarily eliminate the effects of compromised measurements. The communication-delay results also revealed a consistent reduction in the trust-related indicator as delay increased, highlighting the importance of considering measurement reliability alongside conventional dynamic performance.

The findings demonstrate the value of assessing intelligent controller decisions and measurement integrity within a common frequency-regulation framework. The confidence-gated AI supervisor provides a defined fallback to the baseline MSFPID controller when its decision confidence is insufficient, while the measurement-security layer enables reference-assisted processing of unreliable feedback. The results provide evidence of the behaviour of the complete framework under the investigated simulation conditions, although the independent contribution of each security mechanism has not yet been established.

Further research should focus on matched ablation studies, independent verification of cybersecurity performance indices, and validation of reconstruction using online estimates rather than stored attack-free reference trajectories. Numerical-convergence analysis and real-time or hardware-in-the-loop testing would also help establish the practical applicability of the proposed approach. These developments would provide a stronger foundation for implementing reliable and cyber-resilient frequency regulation in renewable-rich interconnected power networks.

## References

[1] A. Pappachen, A. Peer Fathima, Critical research areas on load frequency control issues in a deregulated power system: A state-of-the-art-of-review, Renewable and Sustainable Energy Reviews 72 (2017) 163–177. DOI: 10.1016/j.rser.2017.01.053.

[2] A. Demiroren, H.L. Zeynelgil, GA application to optimization of AGC in three-area power system after deregulation, International Journal of Electrical Power & Energy Systems 29(3) (2007) 230–240. DOI: 10.1016/j.ijepes.2006.07.005.

[3] Y. Prajapati, V.N. Kamat, Secondary frequency regulation/automatic generation control under deregulated power system along with renewable energy sources using electric vehicle/distributed energy storage systems, 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) (2016). DOI: 10.1109/ICEEOT.2016.7755293.

[4] K. Chatterjee, R. Shankar, R. Bhushan, Impact of energy storage system on load frequency control for diverse sources of interconnected power system in deregulated power environment, International Journal of Electrical Power & Energy Systems 79 (2016) 11–26. DOI: 10.1016/j.ijepes.2015.12.029.

[5] A.M. Mohan, N. Meskin, H. Mehrjerdi, A comprehensive review of the cyber-attacks and cyber-security on load frequency control of power systems, Energies 13(15) (2020) 3860. DOI: 10.3390/en13153860.

[6] S.K. Pandey, S.R. Mohanty, N. Kishor, A literature survey on load–frequency control for conventional and distribution generation power systems, Renewable and Sustainable Energy Reviews 25 (2013) 318–334. DOI: 10.1016/j.rser.2013.04.029.

[7] M. Ranjan, R. Shankar, A literature survey on load frequency control considering renewable energy integration in power system: Recent trends and future prospects, Journal of Energy Storage 45 (2022) 103717. DOI: 10.1016/j.est.2021.103717.

[8] X. Liu, S. Qiao, Z. Liu, A survey on load frequency control of multi-area power systems: Recent challenges and strategies, Energies 16(5) (2023) 2323. DOI: 10.3390/en16052323.

[9] M. Wadi, A. Shobole, W. Elmasry, I. Kucuk, Load frequency control in smart grids: A review of recent developments, Renewable and Sustainable Energy Reviews 189 (2024) 114013. DOI: 10.1016/j.rser.2023.114013.

[10] I.A. Chidambaram, B. Paramasivam, Control performance standards based load-frequency controller considering redox flow batteries coordinate with interline power flow controller, Journal of Power Sources 219 (2012) 292–304. DOI: 10.1016/j.jpowsour.2012.06.048.

[11] K. Peddakapu, M.R. Mohamed, M.H. Sulaiman, P. Srinivasarao, A.S. Veerendra, P.K. Leung, Performance analysis of distributed power flow controller with ultra-capacitor for regulating the frequency deviations in restructured power system, Journal of Energy Storage 31 (2020) 101676. DOI: 10.1016/j.est.2020.101676.

[12] I.A. Khan et al., Load frequency control in power systems with high renewable energy penetration: A


Licensed Under Creative Commons Attribution CC BY

strategy employing PIλ(1+PDF) controller, hybrid energy storage, and IPFC-FACTS, Alexandria Engineering Journal 106 (2024) 337–366. DOI: 10.1016/j.aej.2024.06.087.

[13] X.-C. ShangGuan, Y. He, C.-K. Zhang, L. Jin, L. Jiang, M. Wu, J.W. Spencer, Switching system-based load frequency control for multi-area power system resilient to denial-of-service attacks, Control Engineering Practice 107 (2021) 104678. DOI: 10.1016/j.conengprac.2020.104678.

[14] T. Weng, Y. Xie, G. Chen, Q. Han, Y. Tian, L. Feng, Y. Pei, Load frequency control under false data inject attacks based on multi-agent system method in multi-area power systems, International Journal of Distributed Sensor Networks 18(4) (2022). DOI: 10.1177/15501329221090469.

[15] A. Sargolzaei et al., Resilient control design for load frequency control system under false data injection attacks, IEEE Transactions on Industrial Electronics 67(9) (2020) 7951–7962. DOI: 10.1109/TIE.2019.2944091.

## Author Profile

**M. Ramakrishna Raju,** Research Scholar, Department of Electrical & Electronics Engg., JNTUA, Anantapur-515002, Andhrapradesh, India at present working as Senior Lecturer in EEE, Governemnt Polytechnic, Gannavaram, Krishna District-521101, Andhrapradesh, India. Corresponding Author Email: *ramakimathephd[at]gmail.com*

**N. Visali**, Senior Professor, Department of Electrical & Electronics Engg., JNTUA, Anantapur-515002, Andhrapradesh, India


Licensed Under Creative Commons Attribution CC BY