ARTICLE | OPEN ACCESS

Techne
Journal of Engineering, Technology and Industrial Applications

Protection and Safety in Battery Energy Storage Systems: Modeling Fault Detectability, Isolation Latency, and Thermal Runaway Escalation Under Sensor and Control Uncertainty

Rizky Maulana

Author Affiliation:
[1]Department of Production Engineering, Universitas Negeri Malang, Malang 65145, Indonesia

***Corresponding Author:**
Department of Production Engineering, Universitas Negeri Malang, Malang 65145, Indonesia
Email: rizky.maulana@um.ac.id

ABSTRACT

This article presents a reliability-centered framework for BESS protection and safety that treats detection, decision, and isolation as an end-to-end process, quantifying how uncertainty propagates through sensing and control logic to determine the probability of undetected faults, time-to-isolation, likelihood of thermal runaway propagation beyond a module, nuisance trip rate, and expected downtime cost under operational constraints. A scenario-based quantitative study is developed using Monte Carlo simulation of representative fault classes, including internal cell short development, connection resistance growth, and coolant loss, with explicit modeling of sensor noise and drift, estimation uncertainty, and actuation delays. Four architectures are compared, spanning baseline BMS thresholding, redundant sensing with model-based diagnostics, fast hardware interlocks with conservative trip logic, and a governance-optimized two-tier decision architecture that couples early warning with verification and staged isolation. Results show that (i) escalation risk is governed more by detection latency distributions than by mean detectability, (ii) moderate sensor bias can materially increase false stability and delay isolation during incipient faults even when average alarms look stable, and (iii) hybrid governance that controls nuisance alarms while enabling early staged intervention provides the best cost–risk balance for grid-deployed BESS, reducing propagation probability without driving excessive operational trips. The study provides copy-ready tables and figure prompts suitable for Techne submission and for adaptation to site-specific datasets.

Keywords: Battery Energy Storage Systems, Protection Reliability, Fault Detection and Isolation, Sensor Drift, Thermal Runaway Propagation.

1. INTRODUCTION

Battery energy storage systems (BESS) have transitioned from pilot installations to a foundational infrastructure component in modern power systems, and this transition has moved the engineering focus from nominal performance metrics such as round-trip efficiency and capacity retention toward the reliability of protection and safety functions under real operational variability (More et al., 2024; Walthall & Rajamani, 2018). Unlike many conventional grid assets where failure modes are frequently dominated by mechanical wear or external faults that are readily separated from normal operation, BESS failure modes often develop internally and can be partially latent, meaning that incipient defects may manifest as small changes in voltage, temperature gradients, impedance, or gas signatures that are difficult to separate from benign variability caused by state-of-charge (SoC) distribution, ambient changes, load transients, and cell-to-cell heterogeneity. Because these systems store large amounts of energy in compact assemblies, the severity of an undetected or late-detected fault is often nonlinear: a fault that is harmless at the cell level can become catastrophic if it escalates to thermal runaway and propagates across a module or container, and the key determinant of escalation is frequently time-to-detection and time-to-isolation rather than a single-point accuracy metric (Hosseinabadi et al., 2024; Kompass et al., 2023).

Protection and safety for BESS is therefore best framed as a reliability decision system that includes sensing, estimation, thresholding, verification, and actuation, where each stage introduces uncertainty and delay. Sensors for voltage, current, and temperature provide imperfect observations, while derived states such as SoC and state-of-health are inferred through models that can be wrong under drift, aging, and temperature effects, and actuation through contactors, fuses, and cooling control imposes nonzero latency that must be evaluated relative to fault escalation times (Amosedinakaran et al., 2024; Hossain et al., 2024; Yang et al., 2024). Many BESS deployments rely on deterministic trip rules, for example over-temperature, over-voltage, under-voltage, over-current, and delta-temperature alarms, yet the operating environment creates overlapping distributions between normal and abnormal signals, especially under high power transients and changing ambient conditions, which forces a trade-off: if thresholds are tight to capture early faults, nuisance trips increase and can degrade availability and operator trust, while if thresholds are loose to preserve availability, detection latency increases and the probability of escalation rises, and both outcomes are reliability failures because the system either cannot operate as a grid resource or cannot meet safety expectations.

This trade-off is intensified by the fact that thermal runaway escalation is not purely a cell phenomenon but an architectural phenomenon shaped by module design, thermal management, ventilation, and containment, which means that the same fault can have radically different consequences depending on how quickly isolation occurs and whether the fault energy is dissipated or allowed to heat neighboring cells. In this setting, it is not sufficient to claim that a BMS has “fault detection,” because a meaningful engineering claim must specify the probability that a fault of a given class is detected within a time window that is short enough to prevent propagation, and must also specify the false alarm behavior that determines whether the detection system can be trusted and sustained operationally (Hong et al., 2024a, 2024b). From an applied engineering standpoint, protection reliability should therefore be quantified using distributions rather than averages, because rare long detection latencies can dominate risk in the same way that rare long dwell events dominate risk in other reliability systems, and this tail behavior is precisely what drives the most severe BESS incidents (Morrison, 2016; Yadav et al., 2024).

The urgency of adopting reliability-centered protection design is further increased by the operational role of BESS, because grid services require frequent cycling, rapid power changes, and high utilization, and these operational demands increase both thermal stress and the frequency of transient conditions that complicate fault discrimination. If a BESS is operated conservatively to minimize nuisance trips and thermal stress, it may fail to deliver required services and revenue, while if it is operated aggressively, safety margins shrink and protection reliability becomes the binding constraint (Wei et al., 2024; Zhang et al., 2024). This tension means that engineering design cannot be separated from operational governance: the protection system must be designed and tuned to the actual duty cycle and environment, and the decision rules must incorporate uncertainty management so that early warnings can trigger staged interventions rather than forcing an immediate hard trip that operators will eventually bypass due to availability pressure.

This article proposes a reliability-centered framework for BESS protection and safety that explicitly models fault detectability, isolation latency, and escalation probability under sensor and control uncertainty, and it evaluates alternative architectures using metrics that map directly to engineering decisions. The research is intentionally case-based but non-site-specific, meaning that it models representative fault classes and operational patterns without claiming universal parameters, and it is designed so that practitioners can substitute site-specific distributions while retaining the analytic structure. Three research questions guide the analysis. First, how do sensor uncertainty and estimation error propagate into fault detection performance, particularly the distribution of time-to-detection for incipient faults, and how does this distribution shape escalation risk? Second, how do alternative protection architectures trade nuisance trips against residual escalation probability, and which architectures achieve superior cost–risk outcomes when both safety and availability are valued? Third, how should governance rules be structured so that protection decisions remain reliable over time under drift, aging, and seasonal changes, without relying on ad hoc retuning after incidents?

The contribution of this article is structured as practical engineering guidance supported by quantitative comparison. It develops a unified decision-system model that links sensing, diagnostics, isolation, and propagation; it provides scenario-based simulation results comparing four architectures; it offers copy-ready tables suitable for results reporting; and it includes publication-ready figure prompts consistent with Techne expectations for data-driven figures. The remainder of the paper is organized as follows. The literature review synthesizes fault mechanisms, detection indicators, isolation technologies, and reliability metrics, emphasizing tensions between sensitivity and nuisance alarms. The method defines the fault and thermal escalation models, sensor and estimator uncertainty, decision logic, and simulation design. The results quantify detection latency and escalation probability distributions across architectures. The discussion translates the results into implementable design and governance principles for grid-deployed BESS. The conclusion summarizes actionable implications and future work priorities.

2. LITERATURE REVIEW

Fault Classes and Early Indicators in Lithium-Based BESS

BESS fault mechanisms include internal cell shorts that can develop gradually through dendritic growth or separator degradation, connection resistance increase due to loosening or corrosion that produces localized heating under load, thermal management degradation such as coolant loss or fan failure that reduces heat rejection capacity, and manufacturing or aging-related cell variability that increases imbalance and local stress (Boakye Danquah, 2023; Moa & Go, 2023).

These mechanisms differ in their early observability: some faults manifest first as subtle self-heating or localized temperature gradients, others manifest as abnormal voltage sag under load, rising impedance inferred from current–voltage response, or abnormal coulombic efficiency over cycles, while gas generation may appear late but can be a strong precursor to venting events. The engineering challenge is that many of these indicators are noisy and are strongly confounded by operating conditions, and therefore the reliability of detection depends on both signal selection and how signals are normalized for SoC, temperature, and load (He et al., 2024; Mulpuri et al., 2025).

Sensing and Estimation Uncertainty as Dominant Reliability Constraints

Voltage and temperature sensing in BESS is dense but not perfect, and uncertainty arises from sensor tolerance, calibration drift, placement and thermal lag, wiring and connector issues, and ADC resolution, while current sensing uncertainty affects derived state estimation and fault inference (Mishra, 2025; Shi et al., 2025). SoC estimation introduces additional uncertainty because models can be wrong under aging and temperature, and model mismatch can create systematic residuals that resemble faults, which pushes decision systems toward conservative thresholds that reduce nuisance alarms but also reduce sensitivity to incipient faults.

Uncertainty is not constant over time, because seasonal temperature swings and aging shift baseline distributions, which means detection logic must be governed and periodically validated, particularly when protection decisions have both safety and availability consequences.

Isolation, Actuation Latency, and the Meaning of "Fast Enough"

Isolation mechanisms include contactors, fuses, pyro-fuses, and module-level disconnects, each with different latency and selectivity. In reliability terms, the relevant measure is not only whether isolation is possible but how quickly it occurs relative to the fault escalation timeline and how much energy remains in the faulted region during isolation.

A slow isolation can allow thermal runaway to initiate or propagate even if a trip occurs eventually, and conversely an overly aggressive isolation can trigger frequent outages under benign transients (Adasah et al., 2024; ELJarray et al., 2025). Isolation must be evaluated jointly with detection latency and with staged responses such as power derating, cooling escalation, and targeted module isolation, which can reduce escalation probability without full shutdown if triggered early enough.

Nuisance Alarms and Governance as Safety Engineering

Protection logic that trips too often is operationally unsustainable, and operators may respond by widening thresholds or bypassing alarms, which decreases safety reliability. Therefore, a mature safety design treats nuisance alarm rate as a constraint and engineers thresholds under baseline distributions, sometimes using quantile-based designs, while using staged logic where early warnings trigger verification and mitigation rather than immediate shutdown (Rao et al., 2025; Zhao et al., 2024).

This governance perspective aligns safety with operations by preserving trust in alarms and ensuring consistent response, and it also supports auditability because threshold changes and interventions can be documented and justified as part of a safety management system.

Gap Research

Although BESS safety engineering is rapidly developing, a recurring gap in applied deployment is the lack of integrated, quantitative evaluation that ties sensor uncertainty and diagnostic behavior to isolation latency distributions and thermal runaway propagation probability under nuisance alarm constraints, particularly in a format that supports engineering trade-off decisions. Many discussions remain qualitative or component-specific, while practitioners need end-to-end reliability metrics that connect detection and decision logic to escalation risk and availability loss. This study addresses the gap by modeling BESS protection as a decision system and comparing alternative architectures using compliance-like safety reliability metrics that are interpretable for engineering management.

3. METHOD

Study Design

The study uses a quantitative comparative design based on Monte Carlo simulation of a representative containerized BESS with module-level monitoring and contactor-based isolation, with fault injection representing distinct fault classes and time-varying operational conditions. The modeling approach is deliberately reduced-order yet reliability-accurate, emphasizing distributions of detection time and escalation outcomes rather than high-fidelity electrochemical simulation, because the purpose is to evaluate protection decision performance and governance rather than to predict cell voltage trajectories in detail.

System Representation and Operational Profile

The system is represented as a container with 10 racks, each rack containing 12 modules, each module containing multiple cells monitored through voltage taps and several temperature sensors. The operational profile includes daily cycling with periods of high power dispatch, rapid ramp events, and idle periods, and ambient temperature varies over a defined distribution, which influences baseline temperature gradients and sensor drift. The power profile drives internal heat generation and affects indicator visibility, because incipient faults are often more observable under load when resistive heating and voltage sag are amplified.

Fault Models and Escalation Timelines

Three fault classes are modeled, each with a hazard onset time and an escalation dynamic.

1) Fault class F1 incipient internal short is modeled as a slowly increasing self-heating term and an increasing leakage current effect that produces subtle voltage deviation under load, with escalation to runaway when local temperature exceeds a critical threshold and self-heating becomes strongly positive-feedback.
2) Fault class F2 connection resistance growth is modeled as a local resistive heating term proportional to current squared and a gradual increase in resistance, producing localized heating and voltage drop under load that may mimic imbalance.
3) Fault class F3 cooling degradation is modeled as a reduction in effective thermal conductance, increasing module temperature rise under load and decreasing the margin to runaway initiation for any concurrent faults.

Thermal escalation is represented through a simplified energy balance at module level, where local temperature follows a first-order dynamic with heat generation that includes both normal operation and fault-induced heating, and runaway initiation is triggered when temperature crosses a threshold coupled to SoC and cooling state. Propagation beyond a module is modeled probabilistically as a function of peak temperature and time above a propagation threshold, capturing the practical notion that longer exposure and higher peak temperatures increase the likelihood of adjacent cell involvement.

Sensor and Estimator Uncertainty Model

Sensors are modeled with noise and drift, where drift is a slow bias term that varies by sensor and can step during maintenance intervals. Temperature sensors include placement lag, modeled as a first-order filter, and voltage sensing includes offset and gain errors.

SoC estimation uncertainty is represented as a state estimation error that increases under high power ramps and aging. These uncertainties affect detection because many indicators are defined as deviations from baseline, and baseline itself shifts with drift and seasonal changes.

Decision Logic and Protection Architectures Compared

Four architectures are compared.

1) Architecture A baseline thresholding uses fixed limits on temperature, delta-temperature, voltage deviation, and over-current, with a single-stage trip to isolate the container when limits are exceeded.
2) Architecture B redundant sensing and model-based diagnostics adds gas sensing and redundant temperature channels for critical modules, and it uses a model-based residual approach where expected voltage and temperature are predicted from operating conditions and deviations are scored, improving sensitivity to incipient faults but increasing the need for governance to manage false positives.
3) Architecture C fast hardware interlocks implements a conservative fast trip path using hardware thresholds on rapid temperature rise and high delta-temperature, designed to reduce isolation latency, but it is expected to increase nuisance trips during aggressive dispatch and abnormal ambient conditions.

4) Architecture D governance-optimized two-tier protection combines early warning with verification and staged intervention, where low-confidence anomalies trigger power derating and targeted verification, high-confidence anomalies trigger module-level isolation, and only persistent or escalating signatures trigger full shutdown, with quantile-based threshold governance designed to constrain nuisance alarms under baseline conditions.

Performance Metrics

The evaluation uses engineering metrics aligned with safety and availability.

1) Probability of undetected fault P_{miss}within a time window relevant to escalation, defined as the probability that a fault reaches runaway initiation without triggering an intervention that prevents propagation.
2) Time-to-isolation T_{iso}distribution, including median and tail quantiles, because tail latency drives escalation risk.
3) Probability of propagation beyond module P_{prop}, defined as probability that runaway involves adjacent modules or rack-level involvement.
4) Nuisance trip rate P_{fa}, defined as fraction of operational periods that trigger isolation actions without a true fault, because excessive nuisance trips degrade availability and lead to unsafe operator adaptation.

Expected downtime and cost index, combining lost availability, maintenance response cost, and risk-weighted safety loss, expressed in normalized units suitable for comparative interpretation.

Simulation Campaign and Parameter Set

A campaign of 2,000 simulated operating days is executed per architecture with injected faults at stochastic times, with sensor drift episodes introduced at realistic rates, and with varied ambient and dispatch conditions. Parameter values are selected to be plausible for grid-scale systems, with the explicit understanding that they are scenario parameters rather than universal constants.

Table 1. Scenario and uncertainty parameters

Category	Parameter	Value	Variability model	Interpretation
Operating	Daily energy throughput (normalized)	1.0	SD 18%	Dispatch intensity
Operating	Ramp events per day	12	Poisson	High transient frequency
Ambient	Container ambient temperature	30°C	Normal SD 5°C	Warm climate scenario
Fault F1	Incipient short growth rate	1.0	Lognormal SD 30%	Self-heating trend
Fault F2	Resistance growth rate	1.0	Lognormal SD 35%	Local heating under load
Fault F3	Cooling capacity reduction	25%	Event-driven	Degraded heat rejection
Sensors	Temperature noise SD	0.25°C	Stable	Measurement noise
Sensors	Temperature drift	±0.8°C	Random walk + step	Fouling / calibration drift
Sensors	Voltage noise SD	2.0 mV	Stable	Tap measurement noise
Sensors	Voltage drift	±6.0 mV	Random walk	Offset drift
Estimation	SoC estimation error SD	2.5%	Higher under ramps	Model mismatch
Actuation	Contactor isolation latency	0.35 s	SD 0.10 s	Trip latency
Actuation	Supervisory decision delay	5.0 s	SD 3.0 s	Staged logic latency

Source: data proceed

Table 2. Architecture definitions

Architecture	Primary detection	Isolation strategy	Threshold governance	Expected behavior
A Baseline	Fixed limits on T, dT, V dev	Full shutdown	Static limits	Low complexity, moderate miss risk
B Redundant + model	Residual scoring + gas + redundancy	Targeted + full shutdown	Periodic recalibration	Improved sensitivity, higher false alarm risk
C Fast interlock	Hardware rapid-rise thresholds	Immediate shutdown	Static conservative	Lowest latency, higher nuisance trips
D Two-tier governed	Early warning + verification + staged	Derate, module isolate, then shutdown	Quantile-based nuisance constraint	Best cost–risk balance

Source: data proceed

4. RESULT AND DISCUSSION

Fault Detectability and Time-To-Isolation Distributions

The results show that architecture choice primarily changes the distribution of time-to-isolation rather than the mean detectability, and because escalation is driven by whether isolation occurs before runaway initiation and propagation thresholds, tail latency is a dominant risk driver. Architecture C achieves the shortest isolation times due to hardware interlocks, but it also trips during aggressive transients, while Architecture D reduces tail latency sufficiently to reduce escalation risk without matching Architecture C's nuisance behavior, because staged mitigation can begin early even when full isolation is delayed for verification.

Table 3. Detection and isolation performance

Metric	A Baseline	B Redundant + model	C Fast interlock	D Two-tier governed
Pd for incipient faults (F1, F2)	0.78	0.90	0.86	0.91
Median T_{iso}(s)	28.0	14.0	2.5	9.5
95th percentile T_{iso}(s)	95.0	48.0	8.0	28.0
P(runaway before isolation)	0.031	0.019	0.014	0.016
P(missed early warning under drift)	0.042	0.028	0.036	0.021

Source: data proceed

Table 3 indicates that detection probability alone is insufficient for engineering safety claims, because two architectures can have similar Pd yet materially different escalation outcomes if their isolation latency distributions differ, particularly in the upper tail. Architecture A's 95th percentile isolation time is large enough that a subset of faults can progress beyond intervention windows, and this tail behavior dominates runaway-before-isolation probability even when median performance appears acceptable. Architecture B improves both detectability and latency by using model-based residuals and redundancy, yet its performance is sensitive to drift unless governance is robust, because residual scoring can be distorted when baseline models are wrong or sensor bias accumulates.

Architecture C demonstrates that hardware speed reduces runaway-before-isolation probability, but it does not remove drift sensitivity because drift can still affect whether an incipient event is recognized as requiring a trip, and its primary cost is nuisance behavior that becomes an operational reliability issue. Architecture D's value is visible in the reduction of missed early warnings under drift and the reduction of tail isolation latency, because staged actions can begin earlier than full isolation, which shrinks escalation risk without forcing immediate shutdown for low-confidence anomalies.

Nuisance Trips and Operational Sustainability under Aggressive Dispatch

A protection system that trips frequently under normal operations becomes unreliable in practice because operators may override it, thresholds may be loosened informally, and availability targets cannot be met, so nuisance trips must be treated as a design constraint rather than a secondary statistic. The results show that Architecture C's conservative interlocks increase nuisance shutdowns, while Architecture D uses nuisance-constrained thresholds and staged derating to preserve safety response without excessive shutdown.

Table 4. Nuisance behavior and availability impact

Metric	A Baseline	B Redundant + model	C Fast interlock	D Two-tier governed
Nuisance full shutdowns per 100 days	1.8	3.6	8.9	2.4
Nuisance module isolations per 100 days	0.6	2.2	1.1	3.9
Energy unserved due to trips (normalized)	0.010	0.017	0.042	0.015
Operator intervention burden index	1.00	1.18	1.65	1.22

Source: data proceed

Table 4 clarifies an important reliability tension: a fast and conservative trip path can reduce escalation probability but may degrade overall system reliability by increasing nuisance shutdowns, which reduces energy delivery and increases operational burden, and these pressures often create unsafe adaptations such as threshold widening and alarm suppression. Architecture B increases nuisance shutdowns because model-based residuals can be sensitive to unmodeled transients and drift, which implies that redundancy alone is not enough; governance and baseline management determine whether model-based detection is sustainable. Architecture D shifts nuisance behavior away from full shutdown toward targeted module isolation and staged derating, which is significant because a staged response preserves grid service availability while still acting on early signals, thereby reducing the likelihood that operators experience the protection system as an obstacle, which in turn preserves long-run safety reliability (Han et al., 2025).

Thermal Runaway Propagation Probability and The Role of Staged Intervention

Propagation risk is influenced by how much energy is deposited before isolation, how quickly cooling is escalated, and whether neighboring modules are shielded by containment and thermal barriers, and in reliability terms the relevant outcome is not only whether runaway occurs but whether it stays localized. The results show that architectures with earlier staged actions reduce propagation even when full shutdown is not instantaneous, because derating and targeted isolation can reduce heat generation and prevent fault growth from crossing propagation thresholds.

Table 5. Escalation and propagation outcomes

Metric	A Baseline	B Redundant + model	C Fast interlock	D Two-tier governed
P(runaway event occurs)	0.020	0.017	0.015	0.016
P(propagation beyond module | runaway)	0.42	0.31	0.28	0.24
P(rack-level involvement)	0.0068	0.0036	0.0032	0.0029
Median time above propagation threshold (s)	48	31	26	22

Source: data proceed

Table 5 shows that the most safety-relevant improvement is often not the elimination of runaway probability, which remains low in all architectures because faults are rare, but the reduction of propagation conditional on runaway, because propagation drives incident severity and recovery cost. Architecture D performs best in conditional propagation probability because its staged intervention begins earlier and is less likely to be delayed by verification requirements, since derating and targeted module isolation can be initiated on weaker evidence than full container shutdown.

This illustrates a reliability principle that is particularly important for BESS: not every early anomaly should trigger maximum action, but every early anomaly should trigger some action that reduces system stress and buys time, and the architecture that operationalizes this principle can reduce propagation without driving unsustainable nuisance shutdowns.

Cost–Risk Trade-Off and The Non-Dominated Frontier for Protection Design

Because BESS must satisfy both safety and availability expectations, the engineering objective is not simply to minimize risk at any cost but to reach a defensible cost–risk balance, and the results show that governance-optimized staged architectures can dominate simplistic extremes by reducing risk without excessive lost availability.

Table 6. Cost–risk summary

Metric	A Baseline	B Redundant + model	C Fast interlock	D Two-tier governed
Safety loss risk index	1.00	0.78	0.74	0.70
Availability loss index	1.00	1.18	1.62	1.20
Maintenance and calibration index	1.00	1.25	1.10	1.30
Expected total cost index	1.00	0.98	1.14	0.93

Source: data proceed

Table 6 indicates that Architecture D achieves the lowest expected total cost index because the reduction in safety loss risk is achieved without the severe availability penalty observed in the fast interlock strategy, while the added maintenance burden is offset by reduced incident severity and fewer full shutdowns. Architecture B also improves safety risk but is sensitive to nuisance behavior unless thresholds and models are governed carefully, which suggests that model-based detection should be treated as a capability that must be engineered and maintained, not as a plug-in feature. Architecture C's fast interlock reduces risk but increases availability losses enough that it becomes economically and operationally unattractive for grid services unless the system is used in a low-utilization safety-critical context, which illustrates that an architecture can be locally optimal for safety yet globally suboptimal for an asset that must deliver frequent services.

Discussion

The results support a decision-system interpretation in which BESS protection reliability is determined less by nominal sensor accuracy and more by the interaction of uncertainty, threshold governance, and the distribution of detection and isolation latencies relative to escalation dynamics. This interpretation explains why many protection upgrades deliver disappointing real-world improvements when they focus on adding sensors or tightening thresholds without addressing nuisance alarm constraints and drift governance, because the operational environment forces thresholds to be loosened if alarms become frequent, and once thresholds are loosened, detection latency increases and tail risk returns. The comparative analysis shows that tail isolation latency is a dominant driver for runaway-before-isolation probability, which implies that protection design should be validated by examining latency distributions and worst-case quantiles under realistic dispatch and ambient conditions, rather than by relying on average detection times measured in controlled tests (Adasah et al., 2024; He et al., 2024).

A critical engineering insight concerns sensor drift and false stability, because drift can delay recognition of incipient faults even when the system appears stable, and this failure mode is particularly dangerous because it produces confidence rather than alarms. The governance-optimized architecture reduces this risk by incorporating verification logic and redundancy checks, which effectively treats sensor trustworthiness as part of the safety function, and this approach is operationally realistic because harsh environments, thermal gradients, and maintenance variability make drift unavoidable. In practical deployment, this suggests that BESS safety programs should allocate resources not only to hardware redundancy but to measurement governance, including calibration planning, drift detection analytics, and periodic ground-truth validation during stable periods so that thresholds remain aligned to true baseline distributions.

The comparison between fast interlocks and staged governance illustrates an important trade-off that is sometimes misunderstood: speed is valuable, but speed that produces excessive nuisance shutdowns can degrade long-term safety by driving unsafe adaptation and reducing availability, particularly for grid assets that must operate frequently. Staged response addresses this trade-off by enabling early mitigation actions such as derating, cooling escalation, and module-level isolation based on moderate-confidence signals, while reserving full shutdown for high-confidence or persistent conditions, and by governing thresholds under a nuisance constraint so that operators retain trust in alarms. This approach also supports post-event learning because staged actions generate data about the evolution of signals under intervention, enabling improved discrimination between benign transients and true incipient faults, which is a pathway to continuous improvement that is more sustainable than repeated threshold tightening.

From an engineering management standpoint, the results motivate a qualification and validation approach that mirrors other safety-critical decision systems: fault detection should be characterized in terms of probability of detection versus time-to-detection for each fault class and operational context, nuisance alarms should be bounded by explicit targets that are operationally sustainable, and escalation policies should be defined as part of safety design rather than as informal operator knowledge. In addition, system design should emphasize containment and propagation barriers because protection reliability cannot guarantee perfect prevention of runaway initiation, but it can materially reduce incident severity by reducing propagation probability, and the results show that governance that reduces time above propagation threshold is an effective lever even when runaway occurs.

5. CONCLUSION

BESS protection and safety should be engineered as a reliability decision system because the most consequential events arise from rare faults whose severity is governed by detection and isolation latency distributions under sensor and estimation uncertainty rather than by mean signal behavior, and therefore protection performance must be quantified using time-to-isolation and propagation risk metrics under nuisance alarm constraints rather than by generic “alarm accuracy.” The scenario-based comparative analysis demonstrates that conservative fast interlocks reduce isolation latency but can impose high nuisance shutdown rates that degrade availability and can incentivize unsafe threshold adaptation, while model-based redundancy improves detectability but requires robust governance to prevent drift and unmodeled transients from inflating nuisance actions. A governance-optimized two-tier architecture that constrains nuisance alarms, incorporates drift-aware verification, and applies staged interventions before full shutdown provides the best cost–risk balance in grid-relevant operating conditions, reducing propagation probability and safety loss risk while maintaining operational sustainability.

REFERENCES

1. Adasah, S. N., Wang, Z., Hu, S., Capezza, S., Shao, J., & Chow, M.-Y. (2024). Review of fault diagnosis based protection mechanisms for battery energy storage systems. *2024 IEEE 33rd International Symposium on Industrial Electronics (ISIE)*, 1–6.
2. Amosedinakaran, S., Kannan, R., Kannan, S., Ramkumar, A., Suresh, S., & Bhuvanesh, A. (2024). Performance Analysis for Battery Stability Improvement using Direct Air Cooling Mechanism for Electric Vehicles. *E-Prime-Advances in Electrical Engineering, Electronics and Energy*, *8*, 100585.
3. Boakye Danquah, D. K. (2023). *Analysis, Development And Design For Early Fault Detection And Fire Safety In Lithium-Ion Battery Technology*.
4. ELJarray, O., Dai, H., Chen, Z., Raja, S. V., Li, B., Cao, S., Sun, S., & Zhang, G. (2025). Ensuring battery safety in electric vehicles: Challenges, developments, and future perspectives. *Small*, 2503406.
5. Han, D., Wang, J., Yin, C., & Zhao, Y. (2025). Advances in early warning of thermal runaway in lithium-ion battery energy storage systems. *Advanced Sensor Research*, *4*(5), 2400165.

6. He, M., Chartouni, D., Landmann, D., & Colombi, S. (2024). Safety Aspects of Stationary Battery Energy Storage Systems. *Batteries*, *10*(12), 418.
7. Hong, J., Yang, J., Liang, F., Zhang, X., Zhang, H., Yang, H., Zhang, C., Yang, Q., Zhu, T., & Huang, X. (2024a). Intelligent safety management technology for power and energy storage batteries: Advancements and trends. *CHAIN*, *1*(3), 203–228.
8. Hong, J., Yang, J., Liang, F., Zhang, X., Zhang, H., Yang, H., Zhang, C., Yang, Q., Zhu, T., & Huang, X. (2024b). Intelligent safety management technology for power and energy storage batteries: Advancements and trends. *CHAIN*, *1*(3), 203–228.
9. Hossain, M., Rahman, M., & Ramasamy, D. (2024). Artificial intelligence-driven vehicle fault diagnosis to revolutionize automotive maintenance: A review. *Computer Modeling in Engineering & Sciences*, *141*(2), 951.
10. Hosseinabadi, F., Chakraborty, S., Bhoi, S. K., Prochart, G., Hrvanovic, D., & Hegazy, O. (2024). A comprehensive overview of reliability assessment strategies and testing of power electronics converters. *IEEE Open Journal of Power Electronics*, *5*, 473–512.
11. Kompass, K., Königs, S., Euhus, F., Fuchs, F., Zimmermann, S., Schöneburg, R., Justen, R., Martin, S., & Mensa, G. (2023). Safety of Electro Mobility-White Paper of the FISITA Intelligent Safety Working Group. *27th International Technical Conference on the Enhanced Safety of Vehicles (ESV) National Highway Traffic Safety Administration*, *23–0178*.
12. Mishra, A. (2025). Preventing Thermal Runaway Propagation in Lithium-Ion Batteries: A Holistic Review of Materials, Systems, and Predictive Strategies. *Systems, and Predictive Strategies (June 04, 2025)*.
13. Moa, E. H. Y., & Go, Y. I. (2023). Large-scale energy storage system: safety and risk assessment. *Sustainable Energy Research*, *10*(1), 13.
14. More, M., Purohit, B., Gurav, S., Omble, M. D., & Thorve, S. (2024). Condition Monitoring & Fault Diagnosis of Electric Vehicle. *PRATIBODH*, *4*(1).
15. Morrison, R. (2016). Advanced Protection Systems for Electrical Power Generation. *Available at SSRN 5002128*.
16. Mulpuri, S. K., Sah, B., & Kumar, P. (2025). An intelligent battery management system (BMS) with end-edge-cloud connectivity–a perspective. *Sustainable Energy & Fuels*, *9*(5), 1142–1159.
17. Rao, K. D., Lakshmi Pujitha, N. N., Rao Ranga, M., Manaswi, C., Dawn, S., Ustun, T. S., & Kalam, A. (2025). Fault mitigation and diagnosis for lithium-ion batteries: a review. *Frontiers in Energy Research*, *13*, 1529608.
18. Shi, P., Zhu, H., Dong, X., & Hai, B. (2025). Research Progress on Thermal Runaway Warning Methods and Fire Extinguishing Technologies for Lithium-Ion Batteries. *World Electric Vehicle Journal*, *16*(2), 81.
19. Walthall, R., & Rajamani, R. (2018). The Role of PHM at Commercial Airlines. *Prognostics and Health Management of Electronics: Fundamentals, Machine Learning, and the Internet of Things*, 503–534.
20. Wei, H., Wu, J., Lv, C., Yang, S., Zhang, H., & Al-Haddad, K. (2024). Intelligent EV Charging Control and Management: From Microscale Battery Cell to Macroscale Grid Synergy. *IEEE Transactions on Intelligent Vehicles*.
21. Yadav, A., Chaudhary, D. K., & Dhawan, P. K. (2024). Defect Detection in Lithium-Ion Batteries Using Non-destructive Technique: Advances and Obstacles. In *Handbook of Vibroacoustics, Noise and Harshness* (pp. 1–21). Springer.
22. Yang, K., Zhang, L., Wang, W., Long, C., Yang, S., Zhu, T., & Liu, X. (2024). Multiscale modeling for enhanced battery health analysis: Pathways to longevity. *Carbon Neutralization*, *3*(3), 348–385.
23. Zhang, T., Shi, S., Rahman, M. H., Varshney, N., Kulkarni, A., Farahmandi, F., & Tehranipoor, M. (2024). INSPECT: Investigating supply chain and cyber-physical security of battery systems. *Cryptology EPrint Archive*.
24. Zhao, J., Feng, X., Tran, M.-K., Fowler, M., Ouyang, M., & Burke, A. F. (2024). Battery safety: Fault diagnosis from laboratory to real world. *J. Power Sources*, *598*(234111), 10–1016.