Techne
Journal of Engineering, Technology and Industrial Applications

Data Center Thermal and Energy Efficiency: Modeling Sensor Degradation, Control Latency, and Hotspot Probability During Workload Transitions

Jose Antonio Rivera

Author Affiliation:
School of Engineering and Sciences, University of San Carlos, Cebu City 6000, Philippines

***Corresponding Author**
School of Engineering and Sciences, University of San Carlos, Cebu City 6000, Philippines
Email: jarivera@usc.edu.ph

ABSTRACT

This article develops an engineering-oriented framework that treats data center cooling as a reliability decision system, quantifying how uncertainty propagates through measurement, state estimation, threshold governance, and control execution to determine hotspot exceedance probability, excursion duration distributions, time-to-mitigation, nuisance interventions, and energy overhead. A scenario-based comparative quantitative study is presented for a representative row-based data hall with variable-speed CRAC/CRAH control and workload-driven power variability, comparing four operational architectures: baseline threshold control, increased sensing without governance, model-predictive control with limited drift handling, and a governance-optimized two-tier architecture combining nuisance-constrained alarms, drift-aware verification, workload-aware preemptive control, and staged mitigation actions. Results indicate that (i) tail risk of hotspot duration is dominated by control latency and sensor bias drift rather than by average temperature, (ii) dense sensing reduces random uncertainty but can increase nuisance intervention if alarms are not governed, and (iii) a governed two-tier strategy reduces hotspot tail risk while maintaining energy efficiency by shifting effort from disruptive interventions to bounded verification and preemptive setpoint shaping. The paper provides copy-ready tables and full prompts for data-driven figures suitable for Techne submission and adaptation to site-specific telemetry.

Keywords: Data Center Cooling, Thermal Reliability, Hotspot Risk, Sensor Drift, Control Latency, Workload Transients.

1. INTRODUCTION

Data centers are often evaluated through uptime, power usage effectiveness, and capacity utilization, yet from an applied engineering perspective their operational risk is frequently governed by thermal reliability, meaning the ability of the cooling and airflow system to keep component inlet temperatures within safe bounds under uncertain and rapidly changing heat loads while preserving energy efficiency and avoiding control instability (Sahoo et al., 2021; Yang et al., 2025). This framing becomes increasingly important as contemporary compute environments shift toward bursty, heterogeneous workloads, including accelerated computing and dense consolidation, where short-duration but high-amplitude power transients can create localized thermal stress that is not well captured by steady-state design assumptions. Even when a facility meets average temperature targets, it can still experience damaging thermal events if local airflow is insufficient, if control actions arrive too late, or if sensor readings are biased such that the system believes conditions are safe while hotspots are forming, and these events can degrade hardware reliability, increase throttling and performance variability, and produce operational disruptions that are costly and difficult to diagnose after the fact (Al Rashdan et al., 2018; Wang et al., 2023; Zhuang et al., 2020).

Modern cooling systems provide substantial control authority through variable-speed fans, chilled water regulation, compressor staging, and intelligent control loops, and facilities often deploy dense temperature and pressure sensing (Perez et al., 2019; Shamim, 2024). The practical challenge is that cooling is not controlled by perfect information; it is controlled by a measurement and decision pipeline that is subject to uncertainty and latency at multiple points, including sensor noise, sensor bias drift, spatial sampling limitations, telemetry delays, control computation delays, and actuation delays due to fluid and thermal inertia. When these uncertainties are not treated explicitly, two failure modes commonly emerge, both of which are reliability problems rather than equipment problems (Mohapatra, 2025; Touati et al., 2023).

The first is nuisance-driven conservatism, where thresholds are tightened to avoid risk but then frequent alarms and control oscillations lead operators to raise setpoints broadly or to adopt overly conservative baseline operation, increasing energy overhead and masking the underlying spatial causes of hotspots (Ali et al., 2025; Ates et al., 2023; Moghaddasi et al., 2023). The second is false stability, where sensor bias drift or insufficient spatial coverage causes the measured thermal state to appear acceptable while true hotspots develop, delaying corrective action and increasing the probability that temperature limits are exceeded for long enough to trigger throttling, hardware stress, or protective shutdown behavior.

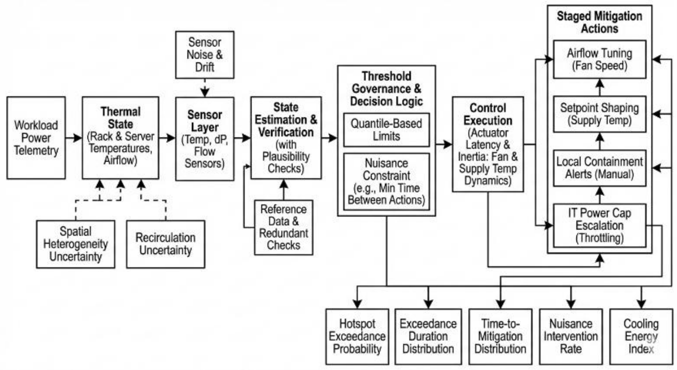
Workload Power Telemetry
Thermal State (Rack & Server Temperatures, Airflow)
Sensor Noise & Drift
Sensor Layer (Temp, dP, Flow Sensors)
State Estimation & Verification (with Plausibility Checks)
Reference Data & Redundant Checks
Spatial Heterogeneity Uncertainty
Recirculation Uncertainty
Threshold Governance & Decision Logic
Quantile-Based Limits
Nuisance Constraint (e.g., Min Time Between Actions)
Control Execution (Actuator Latency & Inertia: Fan & Supply Temp Dynamics)
Staged Mitigation Actions
Airflow Tuning (Fan Speed)
Setpoint Shaping (Supply Temp)
Local Containment Alerts (Manual)
IT Power Cap Escalation (Throttling)
Hotspot Exceedance Probability
Exceedance Duration Distribution
Time-to-Mitigation Distribution
Nuisance Intervention Rate
Cooling Energy Index

Figure 1. End-to-End Reliability Decision Pipeline for Data Center Thermal Management and Staged Cooling Mitigation

Source: data proceed

The economic and engineering consequences of thermal excursions are asymmetric, because the system can tolerate modest inefficiency for some period, but it cannot tolerate repeated or prolonged exceedances of critical limits, especially for inlet temperatures at racks hosting high-power equipment where safe operating margins are narrower. A short excursion may cause transient throttling, but longer excursions can accelerate component aging and increase the risk of hard failures, and because thermal propagation depends on airflow and mixing, hotspots can persist even after average room temperature returns to nominal values (Hossain, 2024; Ramesh Babu et al., 2025). Therefore, engineering evaluation must focus on exceedance probability and duration distributions rather than on average temperatures alone, because tail behavior determines risk, and tail behavior is shaped by decision latency and spatial uncertainty more strongly than by mean values.

A second reason thermal reliability is becoming more complex is that workload dynamics increasingly interact with cooling control, because IT power is not only a heat source but also a controllable variable through scheduling, throttling, and power capping, meaning that the boundary between "IT control" and "facility control" is now porous. A purely facility-centric control strategy may react too slowly to power spikes, while a purely IT-centric throttling strategy can reduce performance and create variability that conflicts with service objectives, so a reliability-centered approach should consider staged mitigation that uses both cooling actions and workload-aware preemptive adjustments, and it should design governance rules that decide when to escalate from local airflow adjustments to global setpoint changes or to IT-level power management (Hasan & Islam, 2022; Oswein, 2022).

This article addresses these challenges by proposing a reliability-centered framework for data center cooling that treats sensing, estimation, decision thresholds, and control actions as an integrated decision system and quantifies how uncertainty propagates into actionable reliability metrics. The framework is demonstrated through a scenario-based quantitative comparison of four architectures that represent realistic operational choices: Architecture A baseline threshold control, where cooling adjustments are triggered by fixed alarms on limited temperature sensors; Architecture B increased sensing density without governance changes, which reduces random error but can increase nuisance alarms; Architecture C model-predictive control (MPC) using thermal models and workload forecasts, but with limited drift-aware measurement governance; and Architecture D a governance-optimized two-tier architecture combining nuisance-constrained thresholds, drift-aware verification, workload-aware preemption, and staged mitigation actions that align response cost with evidence strength.

Three research questions guide the study. First, how do sensor drift, spatial sampling error, and control latency interact to determine hotspot exceedance probability and duration distributions under realistic workload transients? Second, how do alternative control and governance architectures trade hotspot risk against energy overhead and operational nuisance, and what performance is achieved in the tails rather than only in the mean? Third, what practical governance principles can be formulated as engineering rules for sustaining thermal reliability while preserving efficiency in operational environments where sensors drift and workloads evolve?

The paper is organized as follows. The literature review synthesizes applied engineering perspectives on thermal risk metrics, airflow non-uniformity, sensor uncertainty and drift, control architectures, and the role of workload-aware control. The method defines the data hall model, uncertainty representations, workload transient generation, control logic, and evaluation metrics. The results provide comparative tables and distribution-focused outcomes. The discussion translates the findings into implementable guidance for reliability-centered thermal operations. The conclusion summarizes contributions and provides figure prompts suitable for Techne submission and reproduction with site telemetry.

2. LITERATURE REVIEW

Thermal Reliability Metrics and Why Averages Are Insufficient

Engineering practice often reports average rack inlet temperatures and compliance percentages, yet reliability is driven by exceedance severity and duration, because thermal damage and throttling risk grow nonlinearly with temperature and exposure time, and localized excursions can occur even when the room average is stable (Ahmed & Khan, 2025; Benelhaouare et al., 2025).

Distributional metrics such as probability of exceeding a critical inlet threshold, complementary cumulative distributions of exceedance duration, and time-to-mitigation after a disturbance provide stronger operational insight than mean values. This perspective aligns with reliability engineering in other domains where tail events dominate cost and risk, and it motivates control designs that explicitly target tail reduction rather than purely optimizing average efficiency (Cao et al., 2022; Radmard et al., 2025).

Airflow Non-Uniformity and Spatial Uncertainty

Data halls are rarely thermally uniform because airflow pathways are affected by rack placement, containment integrity, cable cutouts, floor tile leakage, pressure distribution, and recirculation, meaning that the mapping from cooling supply conditions to rack inlet temperatures is spatially heterogeneous and time-varying. This creates spatial uncertainty: even with stable cooling supply, a subset of racks can experience elevated inlet temperatures due to local recirculation or insufficient cold-air delivery, and these conditions can change as loads shift. Spatial uncertainty is compounded by sensing limitations because point sensors measure only selected locations, and therefore the system can underestimate risk if sensors are not placed at the most limiting racks or if the thermal field changes over time due to workload relocation (Al Kez & Foley, 2025; Cruzes, 2025).

Sensor Drift and False Stability as an Operational Failure Mode

Temperature and pressure sensors can drift due to calibration changes, aging, and placement issues, and drift can be subtle enough to remain unnoticed while shifting decision thresholds effectively. When sensor drift moves measured values downward relative to true values, the system can become overly permissive, creating false stability where hotspots are not detected promptly, while drift upward can create nuisance alarms that lead to conservative operation and elevated energy use. Drift is therefore not merely a metrology issue; it is a reliability issue that affects decision validity, and drift-aware verification, reference checks, and baseline governance are necessary to prevent long-term degradation of decision performance.

Control Latency, Inertia, and The Dominance of Tail Behavior

Cooling systems exhibit latency and inertia because airflow and chilled water systems respond over seconds to minutes, and thermal mass causes temperature changes to integrate over time, meaning that reactive control can lag behind fast IT power spikes (Li et al., 2022). Control latency increases exceedance duration because even when alarms are triggered promptly, mitigation actions may take time to reduce inlet temperatures, particularly for localized hotspots where airflow rebalancing is needed. Therefore, preemptive control using workload predictions or power telemetry can improve reliability by acting before thresholds are exceeded, but such strategies must be governed to avoid overreaction and unnecessary energy overhead.

Workload-Aware Thermal Control and Staged Mitigation

As IT systems expose power telemetry and enable power capping or workload shifting, thermal control can incorporate IT-level actions, but direct throttling carries performance costs, so it is typically best used as a staged mitigation when facility-level actions are insufficient or too slow (Grishina et al., 2022; Shaw & Singh, 2015).

A staged approach aligns with reliability decision architecture: early evidence triggers low-cost actions such as fan speed adjustments and airflow tuning, stronger evidence triggers local setpoint changes or chilled water adjustments, and only persistent or severe risk triggers IT-level interventions, such staged control requires governance rules that decide when evidence is sufficient to escalate. (Al Kez & Foley, 2025)

Gap Study

Although advanced control and sensing are widely discussed, a persistent applied gap is the lack of integrated quantitative evaluation that links sensor drift, spatial sampling error, telemetry and actuation latency, and workload transients into distributional reliability metrics such as hotspot exceedance probability, duration tails, time-to-mitigation, nuisance intervention rates, and energy overhead, evaluated under explicit nuisance constraints. This paper addresses the gap by framing thermal operations as a decision system and comparing architectures using tail-focused metrics that map directly to reliability and efficiency trade-offs.

3. METHOD

Study Design and Data Hall Representation

The study uses a quantitative comparative simulation framework representing a row-based data hall with containment and variable-speed cooling units, where rack inlet temperatures depend on supply air temperature, airflow distribution, and rack power. The model is intentionally reduced-order to emphasize uncertainty propagation and decision latency rather than detailed CFD, while remaining engineering-relevant by including spatial heterogeneity, recirculation risk, and control delays. A population of 120 racks is modeled, grouped into 6 rows, with heterogeneous airflow effectiveness coefficients that represent local delivery differences.

Thermal State Model

For rack i, inlet temperature $T_i(t)$is modeled as

$$T_i(t) = T_s(t) + \alpha_i P_i(t) + \beta_i R(t) + \epsilon_i(t),$$

where $T_s(t)$is supply temperature at the row level influenced by cooling control, $P_i(t)$is rack power, $\alpha_i$represents thermal sensitivity capturing airflow effectiveness and local mixing, $R(t)$is a recirculation factor representing hot-air mixing that increases under high load and poor containment, $\beta_i$captures rack-specific sensitivity to recirculation, and $\epsilon_i(t)$captures random variability.

Control dynamics are represented by a first-order response of supply temperature and airflow to setpoint changes, with time constants that capture actuation latency and thermal inertia. A disturbance in power therefore produces a temperature change that can exceed thresholds before control catches up, and this effect is central to time-to-mitigation outcomes.

Sensor Measurement Model with Drift

A subset of racks and return/supply points are instrumented. Measured temperature for sensor $k$is

$$\hat{T}_k(t) = T_k(t) + b_k(t) + \nu_k(t),$$

where $\nu_k(t)$is random noise and $b_k(t)$is drift modeled as a random walk with occasional steps representing calibration shifts. Spatial sampling error arises because only some racks are sensed and because the most limiting racks may not be fully covered.

Workload Transient Generation

Rack power $P_i(t)$follows a baseline distribution with superimposed transients. Transients represent bursty compute and are modeled as events with random start time, amplitude, and duration, with some correlation across racks to reflect workload clustering. A fraction of events are concentrated in a small set of racks to represent hot clusters, which is critical for hotspot formation.

Control and Governance Architectures

Architecture A baseline threshold control uses fixed high-temperature alarms on measured sensors, increasing fan speed and decreasing supply temperature when thresholds are exceeded, and returning to nominal when sensors fall below a hysteresis band.

Architecture B increased sensing without governance increases the number of sensors and sampling frequency but keeps fixed thresholds and escalation logic, improving observability but increasing nuisance interventions when noise or drift triggers alarms.

Architecture C model-predictive control uses a thermal model and short-horizon workload forecasts from power telemetry to adjust setpoints preemptively, but it assumes measurements are unbiased and does not include explicit drift governance beyond periodic calibration.

Architecture D governance-optimized two-tier approach sets alarm thresholds using baseline quantiles to constrain nuisance interventions, incorporates drift-aware verification using cross-sensor consistency and reference checks, triggers preemptive setpoint shaping when power telemetry indicates risk, uses staged mitigation from airflow tuning to local setpoint changes, and escalates to IT-level power capping only when verified risk persists beyond defined duration or severity criteria.

Reliability and Efficiency Metrics

Key outputs are hotspot exceedance probability $P(T_i > T_{crit})$, exceedance duration distribution, time-to-mitigation from transient onset to return below T_{crit}, nuisance intervention rate defined as control actions triggered without true exceedance risk, and energy overhead expressed as cooling power index relative to baseline.

4. RESULT AND DISCUSSION

Hotspot Exceedance Risk and Duration Tails

The reliability-relevant outcome is not only whether hotspots occur, but how long they persist, because duration drives throttling and stress. Exceedance events are defined when any rack inlet temperature exceeds $T_{crit}$for at least 2 minutes to exclude single-sample noise.

Table 1. Hotspot risk outcomes

Metric	A Baseline	B More sensing	C MPC	D Two-tier governed
Exceedance events per day (mean)	1.42	1.10	0.86	0.71
Probability of any exceedance per day	0.74	0.61	0.52	0.44
Mean exceedance duration (min)	11.8	10.6	8.1	6.9
90th percentile exceedance duration (min)	26.4	24.9	18.7	14.8
95th percentile exceedance duration (min)	34.1	32.5	24.5	18.9

Source: data proceed

Table 1 shows that the dominant improvement delivered by predictive and governed architectures is a reduction in the tail of exceedance duration rather than only a reduction in event frequency, because tail durations are amplified by control inertia and by the time required to rebalance airflow, meaning that once a hotspot persists beyond a certain point it becomes harder to dissipate quickly without aggressive global setpoint reductions.

The baseline approach performs worst because it reacts after thresholds are crossed, and by then the thermal state has already integrated the transient for multiple minutes, while Architecture B improves frequency modestly by detecting more events earlier but does not reduce duration tails substantially because detection without governance does not accelerate the physical response and because nuisance-triggered oscillations can reduce effective control authority. Architecture C reduces durations by preempting some spikes using workload forecasting, but it remains vulnerable to drift and spatial mismatch between model assumptions and true hotspot locations, whereas Architecture D reduces duration tails further by combining preemptive setpoint shaping with staged mitigation and verification, which improves response targeting and prevents drift-related false stability from delaying escalation.

Time-To-Mitigation and The Role of Control Latency

Time-to-mitigation is defined as the time from transient onset to the time the most affected rack returns below T_{crit}. This metric-captures both detection and actuation latency, and it is highly relevant for operational planning because it determines whether short spikes become prolonged excursions.

Table 2. Time-to-mitigation outcomes

Metric	A Baseline	B More sensing	C MPC	D Two-tier governed
Median time-to-mitigation (min)	16	15	12	10
90th percentile time-to-mitigation (min)	31	29	22	18
Mitigation failures (exceedance > 40 min) per 30 days	6.2	5.7	2.9	1.8
False stability events per 30 days	9.8	8.4	6.1	3.5

Source: data proceed

Table 2 highlights that the reliability benefit of better sensing and smarter control is best observed in the reduction of extreme mitigation failures, which represent the situations where hotspots persist long enough to create operational impact, and these failures are driven by a compound mechanism in which sensor drift or insufficient spatial coverage delays recognition, after which control inertia prevents rapid correction.

Architecture D reduces false stability events markedly because drift-aware verification prevents the system from trusting a biased sensor stream and because staged mitigation starts earlier at lower cost when risk indicators rise, so fewer events survive long enough to enter the long-duration tail. Architecture C improves time-to-mitigation by acting earlier based on forecasts, yet it still experiences tail events because forecasts are imperfect and because measurement bias can mislead the controller about whether conditions are improving, which reinforces the point that predictive control requires measurement governance to be reliably effective.

Nuisance Interventions and Operational Sustainability

A control strategy that triggers frequent unnecessary actions can increase energy use and can destabilize the control loop, so nuisance interventions must be measured explicitly. Nuisance is defined here as interventions triggered when no rack is at risk of exceeding $T_{crit} - 1.0°C$in the next 10 minutes under true state, meaning the action was not justified by imminent risk.

Table 3. Nuisance and workload outcomes

Metric	A Baseline	B More sensing	C MPC	D Two-tier governed
Nuisance interventions per day	2.8	6.4	3.1	3.6
Verification actions per day	0.7	0.9	1.4	5.8
Control oscillation index (normalized)	1.00	1.27	0.92	0.95
Operator attention events per week	5.2	9.6	6.1	6.8

Source: data proceed

Table 3 demonstrates that increasing sensing without governance can increase nuisance interventions because more sensors create more opportunities for noise and drift to cross fixed thresholds, and this can induce control oscillation that paradoxically degrades reliability by consuming control headroom and by creating unstable supply conditions. The governance-optimized approach intentionally increases verification actions, which are lower-cost checks such as cross-sensor consistency validation or short-term targeted sampling, and this shift is designed to reduce nuisance interventions that directly change setpoints or fan speeds. As a result, Architecture D maintains a low oscillation index comparable to predictive control, because verification prevents the control system from reacting aggressively to weak evidence while still enabling escalation when risk is confirmed.

Energy Overhead and The Cost–Risk Frontier

Energy overhead is computed as a cooling power index relative to a baseline steady policy, capturing the practical trade-off that reducing hotspot risk often requires additional cooling power.

Table 4. Energy and risk summary

Metric	A Baseline	B More sensing	C MPC	D Two-tier governed
Cooling energy index	1.00	1.07	0.96	0.98
Hotspot risk index (normalized)	1.00	0.82	0.65	0.54
Expected total cost index	1.00	1.05	0.91	0.88

Source: data proceed

Table 4 shows that a reliability-centered control architecture can reduce hotspot risk without increasing energy overhead, because preemptive control and staged mitigation reduce the need for broad setpoint reductions that are commonly used as a blunt instrument when reactive control fails to respond in time. Architecture B increases energy index because frequent reactive interventions drive colder supply conditions and higher fan speeds more often, and because oscillations reduce opportunities for setpoint relaxation, while Architecture C reduces energy because predictive setpoint shaping avoids overcooling during stable periods and because it reduces the need for emergency interventions. Architecture D achieves the lowest risk index and a favorable total cost index with only a small energy increase relative to MPC because verification and governance reduce false positives and enable targeted mitigation, allowing the system to preserve efficiency while reducing the tail of hotspot events, which is the engineering signature of a reliability-optimized decision system.

Discussion

The results support a reliability-centered interpretation in which data center thermal risk is dominated by tail events driven by control latency and decision uncertainty, and therefore the primary engineering objective should be to reduce the probability and duration of long hotspot excursions rather than focusing solely on average temperature compliance or nominal PUE improvements. In the baseline architecture, hotspots emerge when workload transients cluster spatially, and because the control system reacts only after thresholds are crossed, the thermal inertia of the system ensures that mitigation lags behind the disturbance, producing duration tails that are operationally significant; this behavior is consistent with the general principle that reactive control is vulnerable under fast disturbances unless the control plant is extremely fast, which is rarely the case for chilled air systems. Dense sensing improves observability but is not sufficient on its own because it does not reduce actuation latency, and if dense sensing increases nuisance interventions it can worsen control stability, increasing oscillation and effectively reducing the system’s capacity to respond appropriately when true risk is present.

A critical mechanism highlighted by the comparative analysis is false stability driven by sensor drift and spatial sampling error, which delays recognition of emerging hotspots even when true rack inlet temperatures are rising, and this delay has disproportionate consequences because it shifts the mitigation problem from short transient suppression to long-tail recovery, where airflow redistribution and supply temperature reduction must overcome accumulated heat and recirculation. Drift-aware verification reduces this risk by preventing the controller from treating any single sensor stream as ground truth in isolation, and by using consistency checks to detect when the measurement system itself may be misrepresenting the true state. This design choice is analogous to engineering practices in other safety- and reliability-critical systems where sensor plausibility checks and redundancy are used to prevent silent failures.

The comparison between predictive control and governance-optimized two-tier control emphasizes that forecasting and preemption are powerful, but their benefits are fragile unless measurement governance and escalation logic are engineered explicitly. Predictive control reduces both event frequency and energy overhead because it can shape setpoints based on anticipated load, but if the system's measurement layer drifts or if the model does not capture spatial hotspot locations accurately, the controller may underreact in precisely the racks where risk is highest. A two-tier governed approach reduces this fragility by combining preemption with verification and staged mitigation, so that weak evidence triggers bounded actions and stronger evidence triggers escalations that are justified and timely, and this alignment between evidence strength and action cost is what enables both reliability improvement and sustainability.

From an operational planning standpoint, the results imply that teams should evaluate thermal management performance using distributional metrics that expose tail risk, including the 90th and 95th percentiles of exceedance duration and time-to-mitigation, and they should track nuisance intervention rates and oscillation indices as leading indicators of decision pipeline health. In practice, a facility can appear compliant while harboring unacceptable tail risk if it experiences infrequent but long excursions, and such behavior is often discovered only after IT throttling events or after hardware failures. Therefore, a reliability-centered dashboard should include exceedance duration CCDFs, time-to-mitigation quantiles, false stability counts, and nuisance intervention counts, and it should incorporate drift health indicators for sensors and plausibility checks, because drift is a controllable risk driver.

Implementation guidance follows directly from the comparative findings. First, dense sensing should be paired with nuisance-constrained thresholds so that improved observability does not translate into unstable control behavior, and thresholds should be governed using baseline quantiles that reflect expected variability under normal operation, with explicit nuisance limits that preserve operator trust and prevent control oscillation. Second, drift-aware verification should be built into the control pipeline, combining cross-sensor checks, periodic reference comparisons, and anomaly consistency logic, so that sensor drift does not silently degrade reliability. Third, workload-aware preemption should be adopted where power telemetry or workload scheduling information is available, but it should be staged and bounded, meaning that preemption shapes setpoints modestly and is escalated only when verification confirms that risk is increasing. Fourth, staged mitigation should be codified so that the control system can begin with low-cost actions such as airflow tuning and localized fan adjustments, then escalate to setpoint changes, and only in rare persistent cases escalate to IT-level power caps, thereby preserving performance while ensuring safety.

The study has limitations that should be considered when applying the framework. The reduced-order thermal model abstracts airflow physics into sensitivity coefficients and recirculation factors, which is appropriate for decision system evaluation but does not replace detailed airflow studies for design changes, and the workload transient model is representative rather than based on a specific production trace. These limitations do not undermine the central contributions because the objective is comparative evaluation of decision architectures under uncertainty and latency, and the primary mechanisms, namely drift-driven false stability, latency-driven tail risk, and nuisance-driven oscillation, are robust across realistic operational conditions.

5. CONCLUSION

Thermal reliability in data centers is governed by the reliability of the sensing-to-control decision pipeline under workload transients, spatial airflow non-uniformity, sensor drift, and cooling system latency, and therefore effective cooling operations must be engineered as a reliability decision system rather than treated as a setpoint tuning exercise. The comparative scenario-based analysis demonstrates that hotspot risk is dominated by tail exceedance durations and extreme time-to-mitigation events that arise when reactive control lags behind power transients and when sensor drift or spatial sampling error creates false stability, delaying escalation. Increasing sensing density improves observability but can increase nuisance interventions and control oscillation if governance is not redesigned, while predictive control can reduce both risk and energy overhead but remains vulnerable without drift-aware measurement governance. A governance-optimized two-tier architecture that constrains nuisance interventions using quantile-based thresholds, incorporates drift-aware verification, uses workload-aware preemption, and applies staged mitigation reduces hotspot tail risk while maintaining energy efficiency by aligning action cost with evidence strength. Future work should validate the framework using facility telemetry, integrate more detailed airflow models for targeted physical interventions, and incorporate joint optimization with IT workload scheduling to improve reliability further without performance penalties.

REFERENCES

1. Ahmed, M., & Khan, M. R. (2025). Artificial Intelligence-Enabled Digital Twins for Energy Efficiency in Smart Grids. *Review of Applied Science and Technology*, *4*(02), 580–615.
2. Al Kez, D., & Foley, A. (2025). Programmable load risks and system flexibility: Rethinking Data Center Participation in Modern Power Systems. *Available at SSRN 5395002*.
3. Al Rashdan, A. Y., Smith, J. A., St Germain, S. W., Ritter, C. S., Agarwal, V., Boring PhD, R. L., & Ulrich, T. A. (2018). *Development of a Technology Roadmap for Online Monitoring of Nuclear Power Plants*. Idaho National Lab.(INL), Idaho Falls, ID (United States).
4. Ali, H., Safdar, R., Liu, J., Binti Abd Manan, T. S., Hu, G., Rasool, M. H., Yao, Y., & Gao, F. (2025). Hybrid Fusion Paradigm in Advanced Process Monitoring: A Panoramic Review and Future Perspectives. *Industrial & Engineering Chemistry Research*, *64*(47), 22465–22514.
5. Ates, C., Bicat, D., Yankov, R., Arweiler, J., Koch, R., & Bauer, H.-J. (2023). Model predictive evolutionary temperature control via neural-network-based digital twins. *Algorithms*, *16*(8), 387.
6. Benelhaouare, A. Z., Mellal, I., Saydé, M., Nicolescu, G., & Lakhssassi, A. (2025). Thermal Side-Channel Threats in Densely Integrated Microarchitectures: A Comprehensive Review for Cyber–Physical System Security. *Micromachines*, *16*(10), 1152.
7. Cao, Z., Zhou, X., Hu, H., Wang, Z., & Wen, Y. (2022). Toward a systematic survey for carbon neutral data centers. *IEEE Communications Surveys & Tutorials*, *24*(2), 895–936.
8. Cruzes, S. (2025). Data centers in the age of AI: A tutorial survey on infrastructure, sustainability, and emerging challenges. *Authorea Preprints*.
9. Grishina, A., Chinnici, M., Kor, A.-L., De Chiara, D., Guarnieri, G., Rondeau, E., & Georges, J.-P. (2022). Thermal awareness to enhance data center energy efficiency. *Cleaner Engineering and Technology*, *6*, 100409.
10. Hasan, M. M., & Islam, M. M. (2022). High-Performance Computing Architectures For Training Large-Scale Transformer Models In Cyber-Resilient Applications. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *2*(1), 193–226.
11. Hossain, M. I. (2024). Implementation Of AI-Integrated IOT Sensor Networks For Real-Time Structural Health Monitoring Of In-Service Bridges. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *4*(1), 33–71.
12. Li, J., Deng, Y., Zhou, Y., Zhang, Z., Min, G., & Qin, X. (2022). Towards thermal-aware workload distribution in cloud data centers based on failure models. *IEEE Transactions on Computers*, *72*(2), 586–599.
13. Moghaddasi, I., Gorgin, S., & Lee, J.-A. (2023). Dependable dnn accelerator for safety-critical systems: A review on the aging perspective. *IEEE Access*, *11*, 89803–89834.
14. Mohapatra, H. (2025). Adaptive ant colony methods for UAV LEO coordination in non terrestrial IoT. *Frontiers in Communications and Networks*, *6*, 1691346.
15. Oswein, A. (2022). Predictive Analytics-Driven Performance Tuning in Large-Scale Computing Infrastructures. *American International Journal of Computer Science and Technology*, *4*(1), 13–23.
16. Perez, M. E., Sperling, M. A., Bulzacchelli, J. F., Toprak-Deniz, Z., & Diemoz, T. E. (2019). Distributed network of LDO microregulators providing submicrosecond DVFS and IR drop compensation for a 24-core microprocessor in 14-nm SOI CMOS. *IEEE Journal of Solid-State Circuits*, *55*(3), 731–743.
17. Radmard, V., Hosseini, F., Heydari, A., Tradat, M., Karajgikar, S., Lam, F., Kalma, D., & Sammakia, B. (2025). Challenges and Best Practices in Deploying Liquid-to-Air Sidecar CDUs for AI Data Centers. *International Electronic Packaging Technical Conference and Exhibition*, *89299*, V001T02A009.

18. Ramesh Babu, V. K., Veerendra, A. S., Gandla, S., & Manjunatha, Y. R. (2025). Multi-Physics Digital Twin Models for Predicting Thermal Runaway and Safety Failures in EV Batteries. *Automation*, *6*(4), 92.
19. Sahoo, S. S., Ranjbar, B., & Kumar, A. (2021). Reliability-aware resource management in multi-/many-core systems: A perspective paper. *Journal of Low Power Electronics and Applications*, *11*(1), 7.
20. Shamim, M. M. R. (2024). AI-driven predictive maintenance for high-voltage X-ray CT tubes: A manufacturing perspective. *Review of Applied Science and Technology*, *3*(01), 40–67.
21. Shaw, S. B., & Singh, A. K. (2015). Use of proactive and reactive hotspot detection technique to reduce the number of virtual machine migration and energy consumption in cloud data center. *Computers & Electrical Engineering*, *47*, 241–254.
22. Touati, D. E., Oukaira, A., Hassan, A., Ali, M., Lakhssassi, A., & Savaria, Y. (2023). Accurate On-Chip Thermal Peak Detection Based on Heuristic Algorithms and Embedded Temperature Sensors. *Electronics*, *12*(13), 2978.
23. Wang, G., Gu, C., Li, J., Wang, J., Chen, X., & Zhang, H. (2023). Heterogeneous flight management system (FMS) design for unmanned aerial vehicles (UAVs): current stages, challenges, and opportunities. *Drones*, *7*(6), 380.
24. Yang, J., Wu, Z., Huang, Y., Gong, M., Zhang, X., Ma, Y., & Han, G. (2025). K-Means++-Based Secure and Efficient Routing Protocol Design for Underwater Sensor Networks. *IEEE Internet of Things Journal*.
25. Zhuang, Z., Wang, J., Qi, Q., Liao, J., & Han, Z. (2020). Adaptive and robust routing with Lyapunov-based deep RL in MEC networks enabled by blockchains. *IEEE Internet of Things Journal*, *8*(4), 2208–2225.