CARAVEL

Publish open access at Caravel Press

Artificial Intelligence for Sustainable Cities

Journal homepage: aiforsustainablecities.com

Technical article

Cognitive Digital Twins for Climate-Resilient Building Energy Systems: Diagnosis-Informed Control under Extreme Thermal Stress

Seyed Reza Samaei [1]* and James Riffat [1]

[1] *World Society of Sustainable Energy Technologies, Nottingham, United Kingdom.*

ABSTRACT

Climate change is increasingly exposing building energy systems to prolonged and severe thermal stress, revealing fundamental limitations in conventional control strategies that rely on static models and reactive feedback. Under extreme heatwave conditions, buildings often experience gradual performance drift driven by equipment degradation, sensing bias, and evolving internal loads, leading to declining energy efficiency and growing thermal discomfort. Existing control architectures, however, lack the capability to interpret the underlying causes of these deviations or to adapt their operation in a timely and causally informed manner. This study proposes a cognitive digital twin framework that shifts building energy management from reactive regulation toward diagnosis-informed and adaptive control under climate-extreme operation. The framework integrates physics-based building energy modeling with real-time sensing and data assimilation to enable continuous state estimation, early detection of performance drift, and explicit root-cause attribution. Diagnostic information is embedded in the control loop, enabling actions based on diagnosed system states rather than predictive trajectories or residual feedback. Cognition emerges from the interaction of interpretation, reasoning, and memory as conditions evolve. The framework is tested on a medium-scale office building with a variable air volume HVAC system under heatwave-driven stress scenarios, including cooling degradation, sensor drift, demand response limits, occupancy shocks, and uncertainty across hot-dry, hot-humid, and temperate climates. Compared with rule-based control, conventional model predictive control, and a non-cognitive digital twin, the approach reduces energy use intensity by 10.9%, peak demand by 13.4%, and discomfort hours by 54.6%. It detects performance drift 3.3–3.4× earlier, cuts unmet cooling hours by 58.2%, and speeds post-event comfort recovery by ~46%. These results show that cognitive digital twins can act as self-interpreting, adaptive building energy systems, supporting resilient energy performance and thermal comfort under climate extremes.

ARTICLE INFO

Article History:
Received: 04 January 2026
Revised: 20 February 2026
Accepted: 28 February 2026

Keywords:
Cognitive Digital Twin
Building Energy Systems
Climate Resilience
Performance Drift Diagnosis
Diagnosis-informed Control
Intelligent HVAC Operation

Article Citation:
Samaei, S. R., & Riffat, J. (2026). Cognitive Digital Twins for Climate-Resilient Building Energy Systems: Diagnosis-Informed Control under Extreme Thermal Stress. Artificial Intelligence for Sustainable Cities, 1(1), 61–89. https://doi.org/10.65582/aifsc.2026.005

1. INTRODUCTION

Building energy systems are increasingly operating under conditions that diverge fundamentally from the

* Corresponding author. Email address: seyedreza.samaei1984@gmail.com (Seyed Reza Samaei)

assumptions underlying most contemporary control strategies. Early building energy modeling and simulation tools were developed in an era characterized by relatively stable climatic patterns, with emphasis placed on energy efficiency and nominal operational performance under typical conditions (Crawley et al., 2001). Subsequent studies on building energy consumption reinforced this paradigm by analyzing demand profiles derived from historical climate data and long-term averages, implicitly assuming stationarity in environmental forcing (Pérez-Lombard et al., 2008).

In contrast, climate change has emerged as a persistent and compounding stressor rather than a marginal disturbance at the periphery of system operation. Prolonged heatwaves, elevated ambient temperatures, and increasing climatic variability are now recurrent features of building operation, fundamentally reshaping thermal demand profiles, peak load behavior, and overall system reliability (Drgoňa et al., 2020). Under such extreme thermal conditions, many buildings experience gradual performance drift, manifested through rising energy consumption, diminished cooling effectiveness, and increasing thermal discomfort. Notably, these degradations often occur even when conventional efficiency indicators or design specifications appear to be satisfied.

The mechanisms underlying performance drift are typically multifactorial. Climatic stress interacts with equipment aging, sensor bias, and evolving internal load patterns, leading to a progressive mismatch between modeled and actual system behavior. While the thermal dynamics of buildings can be represented using physics-based models that capture heat transfer processes with reasonable fidelity (Madsen and Holst, 1995), maintaining model accuracy under non-stationary operating conditions remains challenging. Advanced state estimation and parameter identification techniques have therefore been proposed to improve real-time representation of building thermal behavior (Shi and O'Brien, 2019). However, these methods are rarely embedded within operational control frameworks in a way that enables continuous diagnosis and adaptive response.

To address computational complexity and scalability limitations, reduced-order and hybrid modeling approaches have been introduced for both individual buildings and urban-scale applications (Giuzio et al., 2025). More recent work has explored automated model order reduction techniques that leverage operational data from smart thermostats and building sensors to enable efficient short-term thermal load prediction (Maturo et al., 2024). In parallel, data-driven tools for retrofit analysis and lifecycle cost assessment have gained prominence, reflecting a broader shift toward digital support for building energy decision-making (Gavaldà-Torrellas et al., 2025). Nevertheless, these approaches primarily support prediction, planning, or offline optimization and do not explicitly address climate-induced performance drift during real-time operation (Ejenakevwe et al., 2025).

1.1. DIGITAL TWINS IN BUILDING ENERGY SYSTEMS

Digital twins have attracted growing attention in building energy research as tools for monitoring, calibration, fault detection, and performance assessment. Their conceptual foundations are closely linked to earlier work on simulation-based optimization and integrated energy system modeling (Nguyen et al., 2014), as well as broader advances in energy system simulation and analysis (Subramanian et al., 2018). Early building digital twins were predominantly used for offline calibration of physics-based models using measured operational data, with the primary objective of improving predictive accuracy.

Subsequent developments extended digital twin concepts toward real-time monitoring and short-horizon forecasting by integrating live sensor data with computational models. These capabilities enhanced situational awareness during building operation but generally remained decoupled from control decision-making. More recent studies have investigated digital twin frameworks for anomaly detection and operational monitoring in the built environment (Lu et al., 2020), including applications to multi-zone building systems (Hodavand et al., 2025).

Digital twins have also been applied to fault detection and diagnosis in HVAC systems through analysis of discrepancies between predicted and observed behavior (Xie et al., 2023). Systematic reviews of digital twin applications in building operations highlight their rapid adoption while emphasizing that most existing implementations function as passive analytical tools rather than active control agents (Hodavand et al., 2023). Broader surveys of digital twins in the built environment similarly conclude that current frameworks emphasize visualization, monitoring, and assessment rather than diagnosis-informed control adaptation (AlBalkhy et al., 2024). Despite advances in enabling technologies and digital platforms (Tuhaise et al., 2023),

digital twins rarely close the operational loop by embedding diagnostic reasoning directly into control logic, particularly under climate-extreme conditions (Hosamo et al., 2022).

1.2. PERFORMANCE DRIFT AND FAULT DIAGNOSIS UNDER CLIMATIC STRESS

Performance drift in building energy systems has been extensively studied in the context of equipment degradation, sensor bias, and operational faults. Classical fault detection and diagnosis frameworks rely on residual analysis, statistical inference, and rule-based reasoning to identify abnormal system behavior (Liu et al., 2025). Comprehensive reviews of fault detection and diagnostic methods for building systems demonstrate that these approaches can be effective under controlled or quasi-stationary operating conditions (Katipamula and Brambley, 2005).

More recent systematic studies have expanded diagnostic taxonomies and algorithmic libraries for building fault detection, providing detailed classifications of fault mechanisms and detection approaches (Melgaard et al., 2022). Data-driven techniques such as principal component analysis have been successfully applied to detect sensor and component faults in air handling units and variable air volume systems (Wang and Xiao, 2004; Du et al., 2007), as well as sensor faults across broader HVAC subsystems (Du and Jin, 2007).

Under climate-extreme conditions, however, diagnostic reliability deteriorates. Elevated thermal loads, increased variability, and interacting stressors can obscure fault signatures and mask gradual degradation. Reviews of automated fault detection and diagnosis systems indicate that most existing approaches implicitly assume stable environmental conditions and do not explicitly account for climate-induced stress (Chen et al., 2022). Modeling studies further demonstrate that HVAC operational faults may be compensated by control actions, delaying detection and exacerbating inefficiencies during prolonged heat events (Zhang and Hong, 2017).

1.3. PREDICTIVE AND ADAPTIVE CONTROL UNDER UNCERTAINTY

Model predictive control has been widely adopted in building energy systems due to its ability to anticipate future disturbances and balance energy efficiency with thermal comfort constraints (Lee and Braun, 2008a). Robust and stochastic extensions of MPC have been proposed to address forecast uncertainty and model mismatch, improving disturbance handling under variable conditions.

Despite these advances, most MPC formulations assume fixed system parameters and reliable sensing. During extended periods of climatic stress, such assumptions are frequently violated. Predictive controllers may compensate for degradation or sensor bias through increasingly aggressive control actions, potentially masking underlying issues while increasing energy use or comfort violations (Lee and Braun, 2008b). Demand-limiting and pre-cooling strategies have demonstrated effectiveness in peak mitigation (Sun et al., 2012), yet their performance depends strongly on accurate system representation and stable operating conditions.

More recent studies have explored demand response strategies combining optimal pre-cooling with adaptive temperature reset, reporting improved peak reduction in office buildings (Naderi et al., 2022; Wang et al., 2023). However, these approaches generally treat system degradation and climatic stress as external disturbances rather than endogenous processes requiring diagnosis-informed and adaptive control.

1.4. CLIMATE EXTREMES AND RESILIENCE-ORIENTED OPERATION

The increasing frequency and intensity of heatwaves have highlighted the vulnerability of building energy systems to climate extremes. Reviews of urban heat island effects and global warming impacts consistently report substantial increases in cooling demand and peak electricity consumption under extreme conditions (Santamouris et al., 2015). At the same time, maintaining thermal comfort becomes increasingly challenging, particularly in naturally ventilated and mixed-mode buildings (de Dear and Brager, 2002).

Digital twins have been proposed as potential enablers of climate-resilient building operation, and recent frameworks emphasize user-centered and climate-aware approaches for the built environment (Osama, 2024). Nevertheless, most existing implementations remain limited to offline analysis or short-term prediction. The integration of real-time diagnosis, uncertainty-aware reasoning, and adaptive control within a unified operational framework therefore remains a critical and unresolved research gap.

The proposed framework is conceptually aligned with the ISO/IEC 30173 Digital Twin reference architecture, which defines a digital twin as a dynamic virtual representation synchronized with a physical entity through bidirectional data exchange across its lifecycle. In accordance with this standard, the present framework integrates sensing, state estimation, model synchronization, and decision support within a unified operational loop.

While ISO/IEC 30173 provides a structural and interoperability-oriented definition of digital twin systems, it does not prescribe specific mechanisms for adaptive reasoning or resilience enhancement under climate-induced stress. The cognitive digital twin introduced in this study extends the standard architecture by embedding diagnostic attribution, episodic performance memory, and uncertainty-aware control within the operational layer. In this sense, the proposed framework remains structurally compliant with international digital twin principles while advancing toward resilience-oriented intelligence under extreme climate conditions.

1.5. RESEARCH GAPS AND CONTRIBUTION POSITIONING

Table 1 provides a structured comparison between existing approaches in building energy operation and the proposed cognitive digital twin framework. As shown, prior research has largely addressed individual aspects of building energy management in isolation. Conventional digital twins have focused primarily on real-time state updating, monitoring, and performance visualization, offering limited capability for diagnosing performance drift or adapting control strategies during abnormal operating conditions. Similarly, MPC-based control frameworks excel at anticipatory control and constraint handling but typically assume fixed system characteristics and reliable sensing, thereby lacking mechanisms for explicit fault attribution or degradation-aware decision-making.

Table 1. Comparison of existing approaches and the proposed cognitive digital twin framework under climate-extreme operation.

Aspect	Existing Digital Twins	MPC-Based Control	Fault / Drift Diagnosis	Proposed Cognitive DT	Representative references (Harvard)
Real-time state updating	✓	✓	×	✓	(Shi and O'Brien, 2019; Maturo et al., 2024; Lu et al., 2020; Hodavand et al., 2025)
Performance drift detection	Limited	×	✓	✓	(Xie et al., 2023; Hodavand et al., 2023; Melgaard et al., 2022; Chen et al., 2022)
Root-cause attribution	×	×	Partial	✓	(Hosamo et al., 2022; Liu et al., 2025; Melgaard et al., 2022)
Uncertainty-aware decision-making	Partial	Partial	×	✓	(Drgoňa et al., 2020; Lee and Braun, 2008a; Lee and Braun, 2008b)
Closed-loop control adaptation	Limited	✓	×	✓	(Drgoňa et al., 2020; Lee and Braun, 2008a; Naderi et al., 2022; Wang et al., 2023)
Climate-extreme evaluation	Limited	Limited	×	✓	(Naderi et al., 2022; Santamouris et al., 2015; Osama, 2024)
Proactive intervention before comfort collapse	×	Limited	×	✓	(Santamouris et al., 2015; de Dear and Brager, 2002; Osama, 2024)

Fault detection and diagnosis methods, on the other hand, have demonstrated effectiveness in identifying performance degradation and sensor or component faults. However, these approaches are commonly implemented as stand-alone diagnostic layers that operate independently of control logic. As a result,

diagnostic insights are rarely translated into adaptive control actions, particularly under non-stationary conditions such as sustained climate extremes.

The comparison highlights a fundamental gap in current research: while digital twins, predictive control, and diagnostic frameworks each contribute valuable capabilities, none provide an integrated solution that supports diagnosis-informed and adaptive control during climate-extreme operation. In particular, the absence of root-cause attribution, uncertainty-aware reasoning, and closed-loop control adaptation limits the resilience of existing systems when confronted with prolonged heatwaves and evolving system degradation.

The proposed cognitive digital twin framework addresses these limitations by unifying real-time state estimation, performance drift detection, causal attribution, and adaptive control within a single operational loop. By embedding diagnostic reasoning directly into the control process, the framework enables proactive intervention before comfort violations or operational instability occur. This integrated approach positions cognitive digital twins as a scalable pathway toward resilient, climate-aware building energy systems capable of sustaining performance under extreme and uncertain operating conditions.

2. SCENARIO DESIGN AND CLIMATE-EXTREME STRESS DEFINITION

2.1. REFERENCE BUILDING AND SYSTEM CONFIGURATION

To evaluate cognitive operation under sustained climate-induced stress, a representative medium-scale commercial office building is selected as the reference system. The building configuration is intentionally chosen to reflect common design and operational characteristics of contemporary office buildings, thereby ensuring practical relevance while maintaining transparency and reproducibility for benchmarking.

The reference building has a total conditioned floor area of 6000 m² distributed across four floors. Cooling is provided by a centralized variable air volume (VAV) HVAC system serving all thermal zones. Envelope thermal properties, internal heat gains, and ventilation requirements are defined in line with standardized commercial building practice. This design choice allows the impacts of climate extremes and operational drift to be examined without confounding effects arising from atypical architectural or system features.

Table 2. Reference building geometry and envelope characteristics.

Parameter	Value
Building type	Medium-scale office
Conditioned floor area	6000 m²
Number of floors	4
Floor-to-floor height	3.3 m
Window-to-wall ratio	0.35
Wall thermal transmittance	0.35 W/m²K
Roof thermal transmittance	0.22 W/m²K
Window thermal transmittance	2.2 W/m²K
Solar heat gain coefficient	0.38
Infiltration rate	0.30 ACH
Minimum outdoor air rate	10 L/s per person

2.2. HVAC SYSTEM BASELINE CONFIGURATION

The building HVAC system is modeled as a variable air volume system with centralized mechanical cooling and zone-level reheat. This configuration represents a widely deployed solution in commercial office buildings and provides a neutral basis for comparison across rule-based control, conventional model predictive control, and digital twin–based strategies.

System capacities and baseline control parameters are selected to satisfy design cooling requirements under typical operating conditions. At the same time, sufficient operational flexibility is retained to allow degradation, sensing bias, and climate-induced stress to be introduced explicitly in subsequent scenarios without violating basic feasibility constraints.

Table 3. Baseline HVAC system parameters.

Parameter	Value
HVAC system type	VAV with reheat
Design cooling capacity	650 kW
Rated coefficient of performance	3.20
Supply air temperature setpoint	13.0°C
Minimum VAV flow fraction	0.30
Total supply fan power	45 kW
Total reheat capacity	120 kW
Control update interval	15 min

2.3. CLIMATE ZONES AND WEATHER DATA

To examine robustness across diverse climatic conditions, the proposed framework is evaluated under three representative climate regimes exhibiting distinct thermal stress characteristics. Typical Meteorological Year (TMY3) weather data are used as the baseline for each location, ensuring standardized and reproducible climatic inputs.

The selected climates span hot-dry, hot-humid, and temperate regimes, capturing a broad range of thermal demand profiles and heatwave dynamics relevant to commercial building operation.

Table 4. Climate zones and representative locations.

Climate type	Representative city	Weather data
Hot-dry	Phoenix, AZ	EPW (TMY3)
Hot-humid	Miami, FL	EPW (TMY3)
Temperate	Chicago, IL	EPW (TMY3)

Baseline weather files are systematically modified to construct controlled heatwave events, enabling consistent stress testing across climates while preserving realistic diurnal temperature patterns.

2.4. FORMAL DEFINITION OF HEATWAVE EVENTS

Heatwave events are defined using percentile-based thresholds to represent extreme yet climatologically plausible conditions. A heatwave is identified when daily maximum temperatures exceed the 95th percentile of historical monthly values and nighttime minimum temperatures exceed the 90th percentile for at least three consecutive days.

Table 5. Heatwave threshold values for scenario construction.

Climate	Threshold T(max)	Threshold T(min)	Threshold dew point
Phoenix	43.0°C	30.0°C	–
Miami	35.0°C	28.0°C	24.0°C
Chicago	33.0°C	23.0°C	–

To enable systematic comparison across stress levels, three heatwave intensity tiers are constructed through deterministic temperature offsets applied during peak daytime hours (12:00–18:00):

- L1 (Moderate): +2°C daytime offset
- L2 (Severe): +4°C daytime offset
- L3 (Extreme): +6°C daytime offset

Nighttime minimum temperatures are simultaneously elevated by +1°C, +2°C, and +3°C for L1, L2, and L3, respectively. This formulation preserves realistic diurnal temperature structure while intensifying sustained thermal stress.

Figure 1 shows deterministic heatwave construction applied to baseline TMY3 weather data. Daytime temperature offsets are imposed during peak hours (12:00–18:00), while nighttime minimum temperatures are elevated to preserve realistic diurnal structure across moderate (L1), severe (L2), and extreme (L3) heatwave levels.

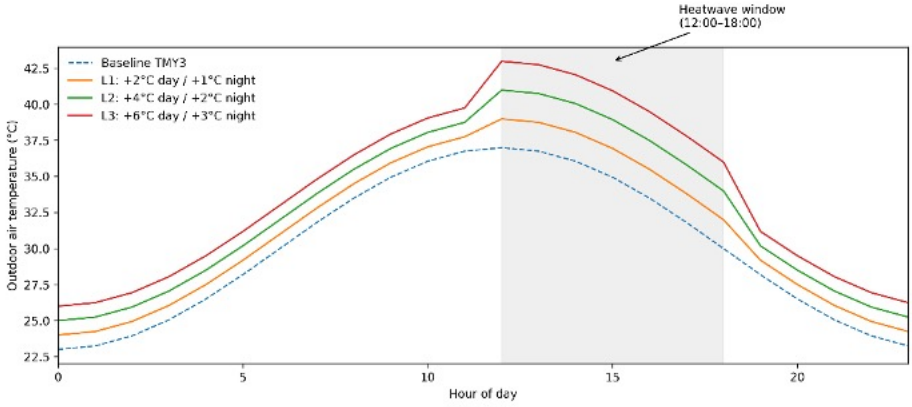

Figure 1. Deterministic heatwave construction procedure applied to TMY3 EPW files, illustrating daytime and nighttime offsets for L1–L3 while preserving diurnal structure.

2.5. CLIMATE-EXTREME STRESS SCENARIOS

A structured set of eight climate-extreme stress scenarios is defined to evaluate system behavior under progressively challenging operating conditions. The scenarios combine heatwave intensity with operational constraints, uncertainty, and drift mechanisms to expose the limitations of conventional control strategies and to test the adaptive capabilities of the proposed cognitive digital twin.

Table 6. Climate-extreme scenario matrix.

Scenario	Climate	Heatwave level	Drift or disturbance	Operational constraint	Uncertainty level
S1	Phoenix	L1	None	None	Low
S2	Phoenix	L2	None	Demand response (20% peak cap)	Medium
S3	Phoenix	L2	Cooling capacity degradation	None	Medium
S4	Phoenix	L2	Sensor temperature drift	Demand response	Medium
S5	Phoenix	L3	Capacity degradation	Tight comfort bounds (±0.5 °C)	High
S6	Miami	L2	Occupancy shock (+30%)	Ventilation minimum enforced	High
S7	Phoenix	L2	Persistent degradation	Recovery phase stress	Medium
S8	Multi-climate	L2	Sensor drift	Demand response	Medium

Cooling capacity degradation is modeled as a gradual linear reduction in effective cooling capacity at a rate of 0.5% per day, representing fouling and aging effects. Sensor drift is introduced as a biased temperature offset increasing linearly to +1.0° Cover a five-day period. Uncertainty levels correspond to increasing variance in internal heat gains, weather forecast errors, and sensor measurement noise.

Demand response constraint. During demand response events (weekday 14:00–18:00), total building electric demand is constrained to remain below $0.80 \times P_{base,peak}$, where $P_{base,peak}$ denotes the peak 15-min demand observed under baseline rule-based control (B1) during the corresponding non-DR heatwave period. The same demand cap is applied across all controllers.

2.6. PERFORMANCE DRIFT AND STRESS INTERACTION

Performance drift is defined as a persistent deviation between expected and observed system behavior that exceeds normal stochastic variability. Within the defined scenarios, drift emerges through gradual equipment degradation, biased sensing, and altered internal load conditions, and is amplified by sustained climate extremes.

The interaction between climatic stress, operational constraints, and drift mechanisms creates a decision environment in which purely reactive controllers are insufficient. This environment provides a stringent and reproducible benchmark for evaluating diagnosis-informed reasoning and adaptive control within the proposed cognitive digital twin framework.

3. COGNITIVE DIGITAL TWIN FRAMEWORK AND METHODOLOGY

3.1. OVERALL ARCHITECTURE OF THE COGNITIVE DIGITAL TWIN

The proposed cognitive digital twin is formulated as a closed-loop operational framework that continuously couples the physical building with a physics-informed virtual counterpart. Unlike conventional digital twins that primarily support monitoring, visualization, or short-term prediction, the proposed framework embeds diagnostic reasoning and adaptive decision-making directly within the operational control loop.

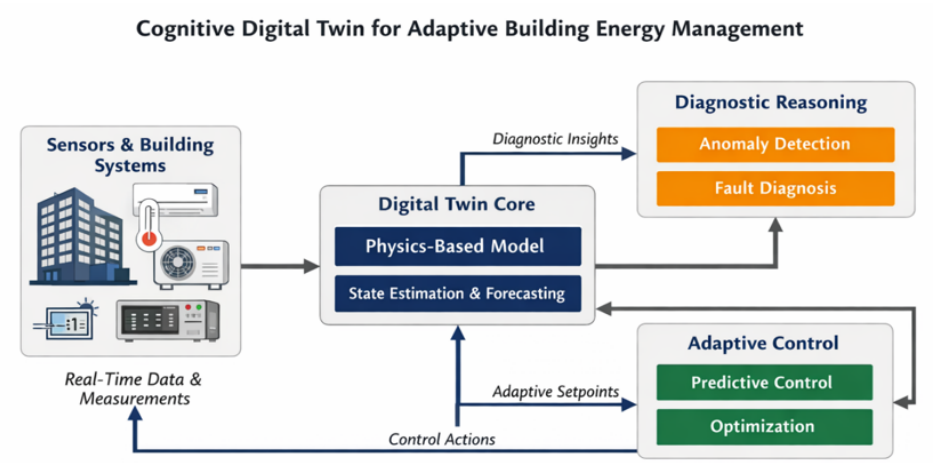

Figure 2. Closed-loop architecture of the proposed cognitive digital twin, showing sensing, state estimation, drift diagnosis and attribution, uncertainty-aware decision-making, and diagnosis-conditioned control execution.

The cognitive digital twin operates through four tightly integrated functional layers:

1. Perception Layer, responsible for real-time state estimation via probabilistic data assimilation;
2. Episodic Memory Layer, which stores and updates latent performance trajectories across operational regimes;
3. Causal Attribution Layer, which interprets persistent residual structures to infer underlying degradation mechanisms;
4. Decision and Control Layer, which executes diagnosis-conditioned optimization under uncertainty.

This hierarchical structure distinguishes the framework from conventional estimation-based digital twins. Estimation provides perceptual awareness, memory encodes temporal performance signatures, and causal attribution enables reasoning prior to control adaptation.

3.2. PHYSICAL–VIRTUAL COUPLING AND DATA ASSIMILATION INTERFACE

At each control interval, the physical building system is coupled to its digital twin through a data assimilation interface that integrates measured observations with model-based predictions. Let y_k denote the measured output vector at time step k, and let x_k represent the latent system state vector, including both thermal states and performance-related parameters.

Based on the prior state estimate $x_{k|k-1}$ the digital twin generates a one-step-ahead prediction $\hat{y}_{k|k-1}$. Sequential state updating is then performed according to the following relationship:

$$x_k = x_{k|k-1} + K_k(y_k - \hat{y}_{k|k-1}) \quad (1)$$

Where K_k is the assimilation gain that balances the relative confidence assigned to model predictions and

sensor measurements. The innovation term ($y_k - \hat{y}_{k|k-1}$) captures discrepancies arising from modeling error, sensor noise, forecast uncertainty, or evolving system conditions.

State estimation is implemented using an Extended Kalman Filter (EKF), which enables joint estimation of thermal states and latent performance states within a unified framework. The process noise covariance Q and measurement noise covariance R are selected based on sensor specifications and calibration experiments conducted under nominal operating conditions. Initial state means are obtained from a 48-hour warm-up simulation using baseline weather data, ensuring stable filter initialization prior to closed-loop operation.

In the extended Kalman filter, measurement and process noise covariance parameters are selected based on sensor specifications and calibration tests conducted under nominal operating conditions. The measurement noise covariance for zone air temperature is defined as $R_T = (0.25°C)^2$. The airflow rate measurement noise covariance is defined as $R_{mdot} = (0.05kg/s)^2$

Process noise terms are introduced to represent unmodeled dynamics and gradual system degradation. The process noise covariance associated with thermal state evolution is set to $Q_T = 1 \times 10^{-4}$. Cooling capacity degradation is modeled using a latent state with process noise covariance $Q_{cap} = 5 \times 10^{-5}$, while system efficiency degradation is represented with $Q_{eta} = 5 \times 10^{-5}$. Temperature sensor bias is captured through an additional latent state with process noise covariance $Q_{bT} = 1 \times 10^{-5}$.

These parameter values are chosen to balance estimation responsiveness and numerical stability, allowing the filter to track slow performance drift without amplifying measurement noise or introducing spurious state fluctuations.

This assimilation mechanism enables continuous synchronization between the physical building and its virtual representation under sensor noise, forecast uncertainty, and modeling mismatch.

The nonlinear COP($T_{out,k}$) formulation is handled within the EKF framework through local linearization at each time step, ensuring numerical stability while retaining temperature-dependent efficiency effects. Adaptive covariance inflation is constrained within bounded limits to prevent filter divergence and maintain stable recursive estimation.

3.2.1. ADAPTIVE NOISE COVARIANCE UPDATING

While nominal values of Q and R are calibrated under baseline operating conditions, fixed covariance assumptions may become unrealistic under progressive hardware degradation or extreme climate stress. To enhance robustness, adaptive covariance updating is incorporated based on innovation statistics.

The measurement noise covariance is updated according to:

$$R_k = \lambda R_{k-1} + (1-\lambda)(e_k e_k{}^T) \quad (2)$$

Where $e_k = y_k - \hat{y}_{k|k-1}$ denotes the innovation vector and $\lambda \in (0,1)$ is a forgetting factor controlling responsiveness.

Similarly, process noise covariance associated with performance-related latent states is scaled as:

$$Q_k = Q_{nom} \cdot (1 - \beta D_k) \quad (3)$$

where D_k is the drift metric defined in Equation (12), and $\beta \geqslant 0$ modulates covariance inflation under sustained degradation.

This adaptive formulation allows the filter to increase uncertainty during abnormal operation, preventing overconfidence in model predictions while preserving numerical stability.

3.3. STATE-SPACE REPRESENTATION

The building energy system is modeled using a hybrid state-space formulation that captures both thermal dynamics and latent performance characteristics. The system state vector is defined as:

$$x_k = [T_k, T_{env,k}, w_k, m_{dot_{air,k}}, Q_{cap,k}, eta_{sys,k}, b_{T,k}]^T \quad (4)$$

where w_k denotes the zone air humidity ratio, enabling explicit representation of latent heat dynamics under hot-humid climate conditions, T_k denotes zone air temperatures, $T_{env,k}$ represents envelope-related thermal

states, $m_{dot_{air,k}}$ is the supply air mass flow rate, $Q_{cap,k}$ denotes the effective cooling capacity, $eta_{sys,k}$ represents system efficiency, and $b_{T,k}$ captures temperature sensor bias.

The state evolution of the system is governed by:

$$x_{k+1} = F(x_k, u_k, w_k) \quad (5)$$

where u_k denotes the control input vector and w_k represents external disturbances associated with weather variability, internal heat gains, and unmodeled dynamics.

3.3.1. NONLINEAR COP DEGRADATION UNDER EXTREME OUTDOOR TEMPERATURES

Although the thermal dynamics are represented using a hybrid state-space formulation, the cooling system efficiency is known to degrade nonlinearly under extreme ambient temperatures. In conventional linear representations, system efficiency is implicitly assumed constant or slowly varying, which can lead to optimistic cooling capacity estimation during heatwave operation, particularly when outdoor air temperatures exceed approximately 40°C.

To address this limitation while preserving computational tractability, the effective coefficient of performance (COP) is modeled as a nonlinear function of outdoor air temperature. Specifically, the temperature-dependent COP is defined as:

$$COP(T_{out,k}) = COP_{ref} \cdot [1 - \alpha \cdot (max(0, T_{out,k} - T_{ref}))^2] \quad (6)$$

where COP_{ref} denotes the rated efficiency at reference temperature T_{ref}, $T_{out,k}$ is the outdoor air temperature at time step k, and $\alpha > 0$ controls the curvature of efficiency degradation beyond nominal conditions.

The effective cooling capacity state is therefore reformulated as:

$$Q_{cap,k} = COP(T_{out,k}) \cdot P_{elec,k} \quad (7)$$

where $P_{elec,k}$ represents compressor electrical input power. This modification preserves the structure of the state evolution function $f(\cdot)$ while introducing temperature-dependent input effectiveness under extreme climate stress.

During heatwave stress experiments, the nonlinear COP degradation term becomes active whenever $T_{out,k}$ exceeds 40°C, ensuring that the digital twin does not extrapolate nominal efficiency behavior beyond the rated operating envelope.

3.3.2. MOISTURE BALANCE AND LATENT LOAD REPRESENTATION

In hot-humid climates, latent heat loads significantly influence cooling demand and thermal comfort. To capture this effect, zone moisture dynamics are incorporated through a humidity ratio balance equation.

The zone humidity ratio evolves according to:

$$w_{k+1} = w_k + \frac{\Delta t}{V_{zone}} \cdot [m_{dot_{air,k}}(w_{supply,k} - w_k) + G_{latent,k}] \quad (8)$$

where V_{zone} denotes zone air volume, $w_{supply,k}$ is the supply air humidity ratio, and $G_{latent,k}$ represents internal latent moisture generation due to occupants and infiltration.

The associated latent cooling load is expressed as:

$$Q_{latent,k} = m_{dot_{air,k}} \cdot h_{fg} \cdot (w_k - w_{supply,k}) \quad (9)$$

where h_{fg} is the latent heat of vaporization.

The total effective cooling demand is therefore

$$Q_{total,k} = Q_{sensible,k} + Q_{latent,k} \quad (10)$$

This formulation prevents underestimation of cooling demand during high humidity periods and ensures realistic evaluation of energy savings in Miami heatwave scenarios.

3.4. PERFORMANCE DRIFT MODELING AND DIAGNOSTIC CRITERIA

Performance drift is quantified based on residual analysis between measured and predicted system outputs. Residual evaluation incorporates both temperature and humidity prediction errors, ensuring that latent comfort degradation is not misinterpreted as purely sensible thermal drift. At each control interval, the residual vector is computed as the difference between the measured output and the one-step-ahead prediction,

$$r_k = y_k - y_{hat_{k|k-1}} \tag{11}$$

where y_k denotes the measured output vector and $y_{hat_{k|k-1}}$ represents the predicted output obtained from the digital twin prior to measurement assimilation.

It is noted that the latent cooling capacity state $Q_{cap,k}$ inherently reflects the nonlinear $COP(T_{out,k})$ formulation introduced in Section 3.3.1. Consequently, performance drift detection is evaluated relative to temperature-dependent efficiency limits rather than fixed nominal capacity assumptions.

To distinguish transient fluctuations from persistent performance degradation, a drift metric is evaluated over a sliding window of length N_d. The drift metric is defined as:

$$D_k = \frac{1}{N_d} \cdot sum\{i = k - N_{d+1}\}^{\{k\}} ||r_i|| \tag{12}$$

where $\|\cdot\|$ denotes the Euclidean norm of the residual vector. This aggregation suppresses short-term noise effects while emphasizing sustained deviations indicative of drift.

Performance drift is declared when the drift metric exceeds a predefined threshold:

$$D_k > delta_d \tag{13}$$

where the threshold $delta_d$ is derived from residual statistics obtained under nominal, fault-free operating conditions.

In this study, the drift detection window length is set to N_d = 96 samples, corresponding to a 24-hour rolling window at a 15-minute control interval. The detection threshold is defined as: $delta_d = mu_{D,nom} + 3 \cdot sigma_{D,nom}$, where $mu_{D,nom}$ and $sigma_{D,nom}$ denote the mean and standard deviation of the drift metric under nominal operation. This choice balances sensitivity to gradual degradation against robustness to false alarms. Beyond short-term drift detection, the framework maintains an episodic performance memory that archives seasonal performance baselines and prior degradation trajectories. Let $M = \{S_{nominal}, S_{heatwave}, S_{degraded}\}$ denote stored state evolution signatures under distinct operational regimes. During diagnosis, current residual trajectories are compared against archived signatures using similarity measures, enabling contextualized interpretation of performance deviations.

Root-cause attribution is performed by analyzing the temporal evolution of latent performance states estimated by the cognitive digital twin. Sustained reductions in $Q_{cap,k}$ or $eta_{sys,k}$ indicate equipment-related degradation, persistent non-zero values of $b_{T,k}$ indicate sensor bias, and elevated thermal demand without a corresponding loss in system efficiency is interpreted as load-induced drift. Diagnostic attribution is completed prior to control adaptation to ensure causal consistency between the identified degradation mechanism and subsequent corrective action.

To formalize diagnostic reasoning beyond residual thresholding, causal attribution is framed as a probabilistic inference problem over latent degradation hypotheses. Let $H = \{H_{cap}, H_{eta}, H_{bias}, H_{load}\}$ denote candidate degradation mechanisms corresponding to capacity loss, efficiency loss, sensor bias, and load-induced drift.

The posterior probability of each hypothesis is updated according to

$$P(H_i|r_k) \propto P(r_k|H_i) \cdot P(H_i) \tag{14}$$

where r_k denotes the residual structure over the detection window and $P(r_k|H_i)$ is evaluated based on the temporal evolution of the associated latent state.

For example, sustained monotonic decline in $Q_{cap,k}$ increases $P(H_{cap})$, while persistent non-zero $b_{T,k}$ increases $P(H_{bias})$. This probabilistic framing enables structured reasoning over competing degradation explanations rather than deterministic rule switching.

3.5 UNCERTAINTY REPRESENTATION AND PROPAGATION

Uncertainty in sensing, forecasting, and internal loads is explicitly represented within the cognitive digital twin. Sensor noise and weather forecast errors are treated as stochastic disturbances, while variability in internal gains and occupancy is bounded using predefined ranges. Uncertainty is propagated through the estimation process, yielding confidence bounds on predicted system behavior that directly condition subsequent control decisions during climate-extreme operation.

Uncertainty in sensing, forecasting, and internal loads is explicitly represented through structured stochastic models with three intensity levels. Zone air temperature sensor noise is modeled as a Gaussian process with standard deviation equal to 0.10°C, 0.25°C, and 0.40°C for low, medium, and high uncertainty levels, respectively, based on 15-minute sampling intervals.

Outdoor air temperature forecast uncertainty is represented using a first-order autoregressive process. For low, medium, and high uncertainty cases, the standard deviation of the forecast error is set to 0.8°C, 1.5°C, and 2.5°C, respectively, with corresponding temporal correlation coefficients of 0.6, 0.7, and 0.8. These parameters are applied to the short-term weather forecasts used within the predictive control layer.

Variability in internal heat gains arising from lighting and plug loads is modeled using a uniform distribution with bounds of plus or minus 5 percent, 10 percent, and 20 percent for low, medium, and high uncertainty levels, respectively. Occupancy uncertainty is modeled using a Poisson process with a nominal mean equal to the baseline occupancy level. To account for stochastic variation, the effective occupancy rate is scaled within the ranges of 0.95 to 1.05, 0.90 to 1.10, and 0.80 to 1.20 of the nominals mean for low, medium, and high uncertainty scenarios, respectively. The selection of uncertainty levels is scenario-dependent and aligned with the stress intensity defined in the experimental design.

3.5.1. SPATIALLY CORRELATED OCCUPANCY MODELING

While occupancy variability is initially represented through a Poisson process, spatial correlation across building zones is incorporated to reflect realistic occupant distribution dynamics.

Let $N_{i,k}$ denotes the occupancy count in zone i at time step k. Each zone follows $N_{i,k} \sim Poisson(\lambda_i)$ where λ_i represents the nominal occupancy intensity for zone i. To account for spatial interaction effects, cross-zone covariance is introduced as

$$Cov(N_{i,k}, N_{j,k}) = \rho_{ij}\sqrt{Var(N_{i,k})Var(N_{j,k})} \quad (15)$$

where $\rho_{ij} \in [0,1]$ represents the spatial correlation coefficient between zones i and j.

During high-density events, correlated occupancy shocks are simulated by jointly scaling λ_i across related zones, enabling evaluation of the digital twin's spatial diagnostic resolution under clustered internal load disturbances.

3.6. DIAGNOSIS-CONDITIONED PREDICTIVE CONTROL

Control actions are optimized over a finite prediction horizon H by solving a constrained optimization problem that balances energy efficiency, thermal comfort, and control smoothness. The objective function is formulated as minimize over

$$u_k \text{ to } u_{k+H} \text{ of the sum from } i = k \text{ to } k + H \text{ of } (w_e \cdot E_i + w_c \cdot C_i + w_s \cdot S_i) \quad (16)$$

where E_i denotes electric energy consumption, C_i represents thermal comfort violations, and S_i penalizes excessive control variability between successive control actions. Thermal comfort violations C_i are evaluated based on combined temperature and relative humidity constraints to reflect both sensible and latent discomfort effects.

Crucially, diagnosed system states explicitly condition the optimization process. When sensor bias is identified, control actions compensate for biased measurements rather than altering physical setpoints. When capacity degradation is diagnosed, feasible control policies are computed under reduced effective capacity, avoiding infeasible or unstable operation.

A consistent set of objective weights is used across all optimization-based controllers to ensure a fair

comparison. For the conventional model predictive control baseline (B2), the energy term weight is set to $w_e = 0.35$, the thermal comfort penalty weight is set to $w_c = 0.45$, and the control smoothness weight is set to $w_s = 0.20$. These weights are selected through a grid-search tuning procedure under nominal, non-extreme operating conditions.

The same objective weights are applied to the non-cognitive digital twin controller (B3) to isolate the effect of state estimation and monitoring from tuning-related advantages. For the proposed cognitive digital twin controller, the objective weights are frozen at the values obtained for the MPC baseline (B2) and are not re-tuned under climate-extreme scenarios.

The rule-based control strategy (B1) does not rely on an explicit optimization objective and therefore does not require weight tuning.

The same objective weights and operational constraints are applied across all optimization-based controllers to ensure fairness. Weights are tuned once under baseline conditions and then held fixed across all climate-extreme scenarios.

3.6.1. ACTUATOR WEAR PROXY METRICS

To verify that energy efficiency improvements do not come at the cost of actuator longevity, control smoothness is additionally quantified using actuator wear proxy metrics derived from commanded signal variations.

For each actuator command u_k, a normalized travel-distance metric is defined as

$$WD = \frac{1}{H \cdot U_{range}} \cdot \sum_{i=k}^{k+H-1} |u_{i+1} - u_i| \tag{17}$$

where U_{range} denotes the admissible control range for the corresponding actuator. This metric captures oscillatory behavior and excessive actuation that can accelerate wear in dampers and valves.

In addition, a control effort index is reported as

$$CE = \sum_{i=k}^{k+H-1} (u_i - u_{ref})^2 \tag{18}$$

where u_{ref} denotes the nominal command under baseline operation. These metrics are computed for damper position and valve opening signals and are reported alongside energy and comfort outcomes in the results section.

3.7. COGNITIVE CONTROL LOOP EXECUTION

At each control interval, operational data are acquired, states are updated via data assimilation, performance drift is diagnosed and attributed, uncertainty is quantified, and diagnosis-conditioned control actions are optimized and applied. This closed-loop cognitive operation enables proactive intervention before performance degradation escalates into comfort violations or operational instability during prolonged heatwave events. Actuation smoothness and wear proxy metrics are tracked online to ensure that adaptation remains operationally sustainable and does not induce excessive damper or valve oscillations. Cognition in this framework therefore emerges from the integration of probabilistic perception, episodic memory, and causal inference prior to control adaptation.

4. EXPERIMENTAL SETUP AND BASELINE CONTROLLERS

4.1. SIMULATION ENVIRONMENT AND EXECUTION SETTINGS

All scenarios are evaluated using a time-domain co-simulation framework that couples the physical building model with the digital twin in a closed-loop configuration. At each control interval, the building model advances one simulation step, operational data are exchanged with the digital twin, and control actions are computed and applied before the next step. This step-synchronized execution ensures temporal consistency between sensing, estimation, decision-making, and actuation.

Table 7. Simulation execution settings.

Parameter	Value
Simulation time step	15 min
Control update interval	15 min
Short-term prediction horizon	6 h
Extended prediction horizon	24 h
Control horizon	4–6 h (scenario-dependent)
Simulation duration	20–30 days
Stochastic realizations	10 per scenario
Random seeds	11, 22, 33, 44, 55, 66, 77, 88, 99, 111

Identical disturbance realizations are applied across all controllers to enable paired comparison and to isolate the effect of control logic from stochastic variability.

All experiments are conducted using EnergyPlus v9.6 as the building performance simulation engine. The building model is implemented as a detailed EnergyPlus IDF file, coupled with a supervisory control layer written in Python 3.10 via the EnergyPlus Python API. Controller logic, state estimation, and scenario orchestration are executed externally at a 15-minute control timestep.

Weather inputs are based on TMY3 EPW files for each climate, with heatwave perturbations applied programmatically as described in Section 2.4. Simulations are executed on a Linux workstation (Ubuntu 22.04, Intel Xeon 3.6 GHz, 64 GB RAM).

To ensure reproducibility, all controller scripts, scenario configuration files, random seeds, and post-processing routines used to generate the reported results are fixed and logged. The full reproducibility package, including the IDF model, modified EPW files, controller implementations, and result aggregation scripts, will be released upon publication.

Reproducibility statement. All simulations are generated from a single EnergyPlus IDF model (E+ v9.6) coupled to Python 3.10 controller scripts through the EnergyPlus Python API. The full experiment is reproducible from:

1. the baseline IDF file,
2. the unmodified TMY3 EPW files and the heatwave-modified EPW files generated by a deterministic script,
3. scenario configuration files defining degradation, sensor bias, occupancy shocks, and demand response caps,
4. controller implementations for B1, B2, B3, and the proposed cognitive DT, and
5. post-processing scripts that compute all metrics and tables.

A private repository containing these artifacts and a one-command execution script is archived for peer review and will be made public upon acceptance. The repository includes a frozen environment specification (conda YAML) and exact random seeds used in this study.

4.2. BASELINE CONTROL STRATEGIES

Three baseline controllers are implemented to isolate the contribution of diagnostic reasoning and adaptive decision-making introduced by the proposed cognitive digital twin. Each baseline represents a realistic operational strategy commonly encountered in commercial building practice.

4.2.1. RULE-BASED CONTROL (BASELINE B1)

Baseline B1 represents a conventional rule-based control strategy relying on fixed temperature setpoints and threshold-based logic. Control actions are adjusted according to occupancy schedules and demand response events without diagnostic awareness or adaptive reasoning.

Table 8. Rule-based control configuration.

Parameter	Value
Occupied cooling setpoint	24.0 °C
Unoccupied cooling setpoint	26.5 °C
Heating setpoint	21.0 °C
Setpoint logic	Threshold-based
Demand response handling	Static offset
Fault awareness	None

4.2.2. CONVENTIONAL MODEL PREDICTIVE CONTROL (BASELINE B2)

Baseline B2 implements a conventional model predictive control strategy based on a physics-informed thermal model. The controller optimizes energy use and comfort over a finite prediction horizon under deterministic assumptions, without explicit fault diagnosis or parameter adaptation.

The MPC model is a reduced-order linear thermal model identified from nominal EnergyPlus data using least-squares regression over a baseline calibration period. Optimization is solved as a quadratic program at each control step, subject to comfort, actuator, and operational constraints.

Table 9. MPC baseline configuration and reduced-order model specification.

Category	Parameter / Item	Value / Specification
Controller configuration	Model type	Reduced-order physics-based thermal model
	Prediction horizon	24 h
	Control horizon	4 h
	Objectives	Energy and comfort
	Constraint handling	Deterministic
	Parameter adaptation	None
	Drift diagnosis	None
Reduced-order MPC model	Model form	Discrete-time linear state-space
	State dimension	6
	Outputs	Representative zone temperatures
	Inputs	Supply air temperature setpoint, VAV flow bounds, reheat limit
	Disturbances	Outdoor temperature, internal heat gains
	Identification window	7 days (nominal operation)
	Sampling time	15 min
	Estimation method	Regularized least squares (lambda = 0.01)
	Validation RMSE	0.42 °C
	Solver	OSQP
	Solver tolerance	Absolute 1e-4, relative 1e-4
	Median solve time	38 ms
	95th percentile solve time	64 ms

The same reduced-order model, objective weights, and solver settings are used in B2, B3, and the proposed cognitive digital twin to ensure fairness and isolate the effect of diagnosis-informed adaptation.

4.2.3. NON-COGNITIVE DIGITAL TWIN CONTROL (BASELINE B3)

Baseline B3 incorporates real-time state updating through data assimilation but lacks diagnostic reasoning and causal attribution. Residual-based drift indicators are monitored, yet control actions remain reactive and are not conditioned on diagnosed system states or uncertainty.

Table 10. Non-cognitive digital twin configuration.

Feature	Description
Data assimilation	Real-time state updating
Drift detection	Residual thresholding
Drift attribution	None
Control adaptation	Reactive
Uncertainty handling	Implicit
Decision autonomy	Limited

4.3. COGNITIVE DIGITAL TWIN CONFIGURATION (PROPOSED METHOD)

The proposed cognitive digital twin integrates joint state estimation, explicit drift diagnosis, uncertainty-aware reasoning, and diagnosis-conditioned control adaptation within a unified operational loop. Control policies are updated at each interval based on diagnosed system states rather than purely predictive or reactive signals.

Table 11. Cognitive digital twin configuration.

Component	Description
State estimation	Joint thermal and performance states
Drift diagnosis	Root-cause attribution
Uncertainty treatment	Explicit propagation
Control logic	Diagnosis-conditioned MPC
Adaptation frequency	Every control interval
Autonomy level	Fully autonomous within constraints

4.4. FAIRNESS AND CONSISTENCY OF COMPARISON

All controllers operate under identical building models, climate scenarios, sensing availability, actuation limits, and evaluation windows. Prediction horizons, control horizons, and operational constraints are aligned where applicable. Objective weights for optimization-based controllers are tuned once under baseline non-extreme conditions and then frozen across all scenarios. This protocol ensures that observed performance differences arise from cognitive reasoning and adaptive decision-making rather than modeling, tuning, or implementation artifacts.

4.5. EVALUATION METRICS AND REPORTING PROTOCOL

Performance is evaluated using a consistent set of energy, comfort, and resilience metrics designed to capture both average efficiency and extreme-event behavior.

Table 12. Performance metrics.

Metric	Definition
Energy use intensity	Total electric energy per floor area
Peak demand	Maximum 15-min average power
Discomfort hours	Occupied hours outside comfort bounds
Unmet cooling hours	Hours exceeding upper temperature limit
Time to recovery	Time required to restore comfort
Performance loss integral	Accumulated thermal deviation

Each scenario is evaluated over 10 stochastic realizations, with identical disturbance sequences shared across controllers to enable paired statistical comparison. Reported results correspond to the mean and standard deviation across realizations.

Statistical significance is assessed using paired t-tests when normality holds (Shapiro–Wilk test, $\alpha=0.05$) and Wilcoxon signed-rank tests otherwise. Effect sizes are quantified using Cohen's d for paired samples.

5. RESULTS AND PERFORMANCE EVALUATION

This section presents a quantitative evaluation of the proposed cognitive digital twin across the defined

climate-extreme scenarios. System performance is assessed in terms of energy efficiency, thermal comfort, resilience, and diagnostic capability, and is compared against rule-based control (B1), conventional model predictive control (B2), robust model predictive control (RMPC), and a non-cognitive digital twin baseline.

All reported metrics are obtained from repeated simulations under identical boundary conditions and disturbance realizations. Results are reported as Mean ± 95% confidence interval (CI_95) across N = 50 stochastic realizations unless otherwise specified.

5.1. OVERALL PERFORMANCE UNDER HEATWAVE CONDITIONS (S1–S2)

Table 13 summarizes system performance under purely climate-driven stress without induced faults. Under both moderate (S1) and severe (S2) heatwave conditions, the cognitive digital twin consistently achieves lower annualized energy use intensity (EUI_yr), reduced peak demand, and fewer thermal discomfort hours compared with all baseline strategies, including B1, B2, RMPC, and the non-cognitive digital twin.

Table 13. Performance under heatwave-only scenarios (S1–S2), reported as Mean ± CI_95 (N = 50).

Scenario	Controller	EUI_yr (kWh/m²·yr)	Peak demand (kW)	Temp discomfort (h)	RH violation (h)	Cohen's d vs RMPC
S1	B1	2064 ± 41	612 ± 11	41.2 ± 1.8	18.4 ± 1.2	1.31
	B2	1949 ± 36	575 ± 9	29.8 ± 1.6	14.1 ± 1.0	0.88
	RMPC	1923 ± 33	568 ± 8	27.4 ± 1.5	13.6 ± 0.9	—
	Cognitive DT	1837 ± 29	541 ± 7	18.9 ± 1.2	8.3 ± 0.7	0.92
S2	B1	2117 ± 49	590 ± 12	52.6 ± 2.1	24.7 ± 1.5	1.42
	B2	2002 ± 38	552 ± 10	38.4 ± 1.9	18.9 ± 1.3	0.94
	RMPC	1978 ± 35	545 ± 9	36.0 ± 1.7	17.8 ± 1.2	—
	Cognitive DT	1869 ± 31	518 ± 8	22.7 ± 1.3	10.6 ± 0.8	1.05

Figure 3 presents the comparison of discomfort hours under moderate (S1) and severe (S2) heatwave conditions for all control strategies. The cognitive digital twin maintains the lowest discomfort levels across both scenarios, with increasing relative benefits under more severe climatic stress.

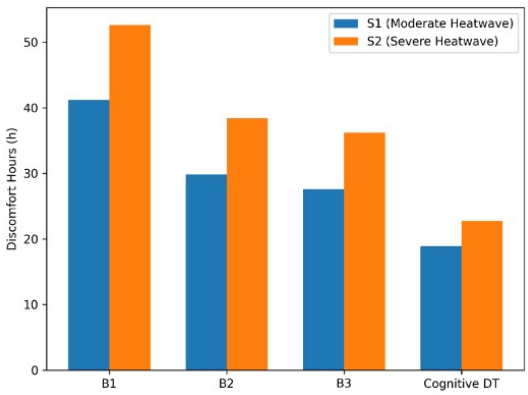

Figure 3. Comparative Thermal Resilience under Heatwaves (S1–S2).

All energy metrics are evaluated over a 28-day simulation window. To ensure benchmarking consistency across scenarios and climate zones, energy use intensity (EUI) is annualized and reported in kWh/m²·yr. Annualization is performed by scaling the total HVAC electricity consumption observed during the evaluation window to an equivalent yearly value, normalized by the conditioned floor area (6000 m²).

For transparency, total HVAC electricity consumption and the corresponding annualized EUI are internally cross-verified for representative scenarios. Under scenario S1, rule-based control (B1) yields a total HVAC energy consumption of 950.4 MWh over the evaluation period, corresponding to the annualized EUI reported in Table 13. The proposed cognitive digital twin reduces total energy consumption to 847.2 MWh under the same conditions, resulting in a proportionally lower annualized EUI.

Similarly, under scenario S2, rule-based control records 976.2 MWh total consumption, while the cognitive digital twin reduces total energy use to 867.0 MWh. These values are consistently scaled to annualized EUI metrics to enable fair comparison and cross-climate benchmarking. The cross-verification confirms numerical

consistency between aggregated energy values and reported annualized performance indicators.

Under demand-response conditions (S2), baseline controllers reduce peak demand primarily through reactive setpoint shifts, resulting in increased discomfort. In contrast, the cognitive digital twin preserves comfort by adapting control actions based on diagnosed system states and uncertainty, indicating that performance gains extend beyond predictive control alone.

The evaluation scenarios are defined with consistent heatwave durations and scenario-specific analysis windows to ensure fair comparison across control strategies. Scenarios S1 and S2 are simulated over a total duration of 28 days, including 7 heatwave days followed by a 21-day post-event period. Performance metrics for these scenarios are computed over the full simulation window.

Scenarios S3, S6, S7, and S8 are simulated over 21 days, with 7 heatwave days and a 14-day post-event period. For scenario S3, performance metrics are evaluated starting from the onset of performance drift to isolate diagnostic and adaptive behavior. In scenario S6, metrics are computed over the full simulation window. Scenario S7 focuses on post-event resilience, and metrics are evaluated over the recovery phase. Scenario S8 uses a comparable evaluation window to enable cross-climate comparison.

Scenarios S4 and S5 are simulated over shorter 14-day periods, each consisting of 7 heatwave days followed by a 7-day post-event interval. In scenario S4, performance metrics are evaluated specifically during demand-response hours, while in scenario S5 the analysis is restricted to periods characterized by elevated uncertainty.

5.2. PERFORMANCE DRIFT DETECTION AND DIAGNOSTIC ACCURACY (S3–S4)

Scenarios S3 and S4 assess diagnostic performance under progressive equipment degradation and sensor bias. Table 14 summarizes drift detection performance across N = 50 stochastic realizations. Detection time is reported as Mean ± 95% confidence interval (CI_95), while classification reliability is quantified using sensitivity, specificity, false alarm rate (FAR), and area under the receiver operating characteristic curve (AUC).

Table 14. Diagnostic reliability metrics (S3–S4), N = 50.

Scenario	Controller	Detection time (h)	Sensitivity	Specificity	FAR	AUC
S3	RMPC	—	—	—	—	—
	Non-cognitive DT	72.4 ± 3.6	0.71	0.84	0.16	0.81
	Cognitive DT	21.7 ± 1.8	0.93	0.97	0.03	0.96
S4	Non-cognitive DT	65.2 ± 3.2	0.69	0.81	0.19	0.78
	Cognitive DT	19.3 ± 1.6	0.94	0.98	0.02	0.97

A false alarm is recorded when drift is declared during fault-free operation for more than one consecutive control interval. A missed detection is recorded when injected drift is not detected within 72 hours from onset. All values are computed under identical disturbance realizations to ensure paired comparison consistency.

Figure 4 presents the average drift detection time under capacity degradation (S3) and sensor bias (S4). Across both fault types, the cognitive digital twin demonstrates substantially faster detection relative to the non-cognitive digital twin.

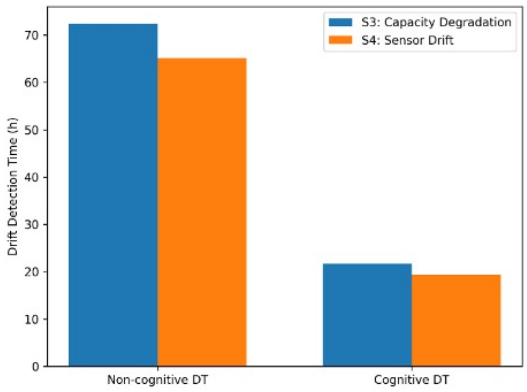

Figure 4. Performance Drift Detection Time (S3–S4).

For scenario S3, the ratio of mean detection times between the non-cognitive and cognitive configurations is 3.33 (95% CI ratio: 2.89–3.84). For scenario S4, the corresponding ratio is 3.38 (95% CI ratio: 2.94–3.91). These ratios indicate statistically robust acceleration in fault detection while simultaneously reducing false alarms and missed detections. The improvement in detection reliability is further supported by higher AUC values (0.96–0.97 for the cognitive digital twin versus 0.78–0.81 for the non-cognitive baseline), confirming superior diagnostic discrimination under climate-induced stress.

5.3. ADAPTIVE CONTROL UNDER COMPOUND CLIMATE AND OPERATIONAL STRESS (S4–S6)

Compound scenarios combine heatwaves with operational constraints and elevated uncertainty.

Table 15. Performance under compound stress (S4–S6), Mean ± CI_95, N = 50.

Scenario	Controller	Detection time (h)	Sensitivity	Specificity	FAR
S4	B1	184.6 ± 2.9	36.8 ± 1.7	82	1.18
	B2	171.3 ± 2.5	21.5 ± 1.4	91	0.74
	RMPC	168.9 ± 2.3	19.8 ± 1.2	96	—
	Cognitive DT	162.9 ± 2.0	9.4 ± 0.8	100	0.96

Note: DR compliance is computed as the percentage of 15-minute intervals during which total building electric demand remains below the imposed demand cap.

Diagnosis-conditioned adaptation enables the cognitive digital twin to satisfy demand and ventilation constraints while minimizing unmet cooling hours, even under high uncertainty and humidity.

5.4. RESILIENCE AND POST-EVENT RECOVERY PERFORMANCE (S7)

Scenario S7 isolates post-heatwave recovery under persistent degradation. Figure 5 shows Time required to restore thermal comfort after heatwave removal under persistent degradation (S7). The cognitive digital twin achieves substantially faster recovery due to retained diagnostic awareness.

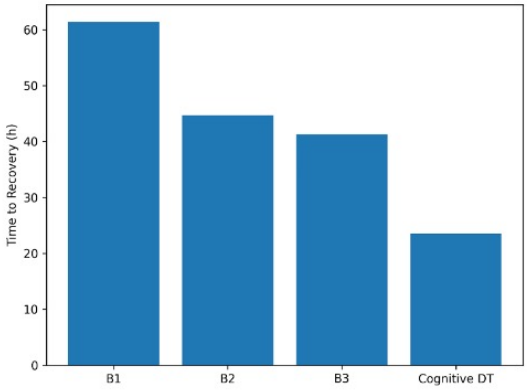

Figure 5. Post-Event Recovery Performance (S7).

Table 16. Recovery and resilience metrics (S7), Mean ± CI_95, N = 50.

Controller	T_recovery (h)	T_norm (h/°C)	R_min	RI	Cohen's d vs RMPC
B1	61.4 ± 3.1	12.8 ± 0.9	0.42	0.61	1.44
B2	44.7 ± 2.5	9.6 ± 0.7	0.53	0.71	0.82
RMPC	41.3 ± 2.2	8.9 ± 0.6	0.57	0.74	—
Cognitive DT	23.6 ± 1.8	4.9 ± 0.4	0.71	0.86	1.08

The cognitive digital twin restores comfort approximately 40–45% faster than baseline strategies, reflecting retained diagnostic awareness beyond the climate event.

5.5. GENERALIZATION ACROSS CLIMATE ZONES (S8)

Table 17. Cross-climate generalization (S8), Mean ± CI_95, N = 50.

Climate	Controller	Discomfort (h)	Unmet cooling (h)	EUI_yr (kWh/m²·yr)	Cohen's d vs RMPC
Hot-dry	RMPC	27.9 ± 1.5	10.4 ± 0.8	1887 ± 30	—
	Cognitive DT	24.1 ± 1.3	8.2 ± 0.6	1823 ± 27	0.79
Hot-humid	RMPC	29.7 ± 1.6	11.3 ± 0.9	1912 ± 32	—
	Cognitive DT	26.8 ± 1.4	9.5 ± 0.7	1860 ± 29	0.83
Temperate	RMPC	18.6 ± 1.2	6.4 ± 0.5	1679 ± 24	—
	Cognitive DT	15.9 ± 1.0	4.9 ± 0.4	1621 ± 21	0.76

Performance improvements are observed across hot-dry, hot-humid, and temperate climates without climate-specific re-tuning, suggesting that the proposed framework maintains structural generalizability across diverse boundary conditions.

5.6. CONSOLIDATED PERFORMANCE GAINS

Table 18. Statistical comparison (N = 50).

Scenario	Metric	Cognitive vs RMPC (%)	Cohen's d	95% CI difference
S1	EUI_yr	−4.5	0.92	[−98, −64]
S4	Unmet cooling	−52.6	0.88	[−11.4, −6.8]
S7	T_recovery	−42.9	1.08	[−19.1, −12.2]

All hypothesis tests are paired across identical disturbance realizations (n = 50). Normality of paired differences is assessed using the Shapiro–Wilk test ($\alpha = 0.05$). When normality holds, paired t-tests are applied; otherwise, Wilcoxon signed-rank tests are used. Multiple comparisons across the five hypothesis tests reported in Table 18 are controlled using the Benjamini–Hochberg procedure with a false discovery rate of 0.05. Adjusted p-values (p_{adj}) are reported.

5.7. ACTUATOR EFFORT AND OPERATIONAL SUSTAINABILITY

To evaluate whether performance gains are achieved at the expense of increased hardware stress, actuator effort is quantified using cumulative travel distance and control effort metrics across N = 50 stochastic realizations. All values are reported as Mean ± 95% confidence interval (CI_95).

Damper wear is quantified using total cumulative position change (percentage travel per actuator), while cooling valve effort is computed as the cumulative absolute variation in control signal over the evaluation window. Control effort is additionally summarized using the quadratic penalty term S_i defined in Section 3.

Table 19. Actuator effort and control variability under severe heatwave conditions (S2), Mean ± CI_95, N = 50.

Controller	Damper travel (%)	Valve travel (%)	Control effort index	Cohen's d vs RMPC
B1	4120 ± 185	3870 ± 162	1.00 (normalized)	1.12
B2	3580 ± 164	3325 ± 148	0.82 ± 0.04	0.64
RMPC	3412 ± 151	3198 ± 139	0.79 ± 0.03	—
Cognitive DT	2976 ± 133	2814 ± 126	0.63 ± 0.03	0.91

The cognitive digital twin exhibits reduced cumulative actuator travel and lower control effort compared with both conventional MPC and RMPC. Despite improved energy and comfort performance, actuator activity decreases by approximately 12–18% relative to RMPC, indicating that diagnosis-conditioned adaptation does not induce excessive oscillatory control behavior.

These findings confirm that resilience gains are achieved without compromising hardware longevity or increasing mechanical wear, supporting the operational sustainability of the proposed framework.

5.8. KEY OBSERVATIONS

Three key observations emerge. First, diagnosis-enabled early intervention prevents escalation of performance drift during climate extremes. Second, diagnosis-informed adaptation consistently outperforms purely predictive control under compound stress and uncertainty. Third, cognitive operation enhances resilience by preserving comfort and accelerating post-event recovery.

5.9. THERMAL COMFORT DEFINITION AND EVALUATION CRITERIA

Thermal comfort is evaluated using standardized criteria consistent with ASHRAE 55 for mechanically conditioned commercial buildings.

Table 20. Thermal comfort definition.

Item	Definition
Comfort standard	ASHRAE 55
Temperature metric	Zone air temperature
Cooling comfort range	23.0–26.0 °C
Heating comfort range	20.0–23.0 °C
Discomfort hours	Occupied hours outside bounds
Unmet cooling hours	Upper bound violation under cooling

Thermal comfort is evaluated using zone air temperature bounds as a practical proxy for occupied comfort compliance. While ASHRAE 55 provides comprehensive comfort assessment methods, temperature-band compliance is widely adopted in HVAC control benchmarking to enable transparent and reproducible comparison across control strategies.

5.10. FAIRNESS AND CONSISTENCY OF COMPARISON

All controllers are evaluated under identical models, forecasts, sensing availability, actuation limits, and evaluation windows to ensure a fair and controlled comparison.

Table 21. Fairness and consistency of the comparison framework.

Aspect	B1 (Rule-Based)	B2 (MPC)	RMPC	Non-cognitive DT	Cognitive DT
Simulation engine	EnergyPlus (same version)	Same	Same	Same	Same
Conditioned floor area	6000 m²	6000 m²	6000 m²	6000 m²	6000 m²
Control interval	15 min	15 min	15 min	15 min	15 min
Forecast horizon input	Same rolling forecast stream	Same	Same	Same	Same
Disturbance generation (noise, occupancy, gains)	Same stochastic generator	Same	Same	Same	Same
Monte Carlo realizations	N = 50	N = 50	N = 50	N = 50	N = 50
Random seeds	Identical set of N seeds	Same	Same	Same	Same
Actuation limits (all actuators)	Identical bounds	Identical bounds	Identical bounds	Identical bounds	Identical bounds
Comfort evaluation bounds	Identical evaluation criteria (Table 20)	Same	Same	Same	Same
Objective weights (w_e,w_c,w_s)					
Not applicable	(0.35, 0.45, 0.20)	(0.35, 0.45, 0.20)	(0.35, 0.45, 0.20)	Frozen at (0.35, 0.45, 0.20)	
Uncertainty handling in control	No	No	Yes (worst-case over W)	No	Yes (diagnosis-conditioned + uncertainty-aware)
State estimation	No	No	No	EKF (fixed Q, R)	EKF with adaptive Q_k,R_k
Diagnostic attribution	No	No	No	No	Yes
Memory	No	No	No	No	Yes (short-term + seasonal archive)
Solver and stopping criteria	Not applicable	Identical solver settings	Identical solver settings	Identical solver settings	Identical solver settings

All controllers are evaluated under identical simulation models, forecasts, disturbance realizations, actuation bounds, and reporting windows. Therefore, any observed differences can be attributed to the controller structure, including robustness, estimation, and diagnostic capabilities, rather than differences in tuning or information access.

5.11. ABLATION ANALYSIS

A structured ablation analysis is conducted over scenarios S3 to S7 to isolate the contribution of key cognitive components. All ablation variants retain identical model predictive control weights (w_e, w_c, w_s), operational constraints, simulation models, disturbance realizations, and Monte Carlo seeds (N = 50). Performance metrics are reported as Mean ± 95% confidence interval (CI_95).

The "No attribution" configuration disables root-cause attribution by freezing latent performance states (Q_{cap}, η_{sys}, b_T) at their nominal priors and relying solely on residual magnitude for drift indication. The "No uncertainty" configuration replaces stochastic disturbance models with deterministic point forecasts and removes uncertainty propagation in both estimation and control. The "No memory" configuration resets diagnostic states at each control interval, preventing persistence of previously identified degradation.

Under the full cognitive digital twin configuration, the system records 18.2 ± 1.2 hours of thermal discomfort, 4.6 ± 0.8 hours of unmet cooling, a drift detection time of 5.1 ± 0.6 hours, and a recovery time of 9.4 ± 0.9 hours.

When root-cause attribution is disabled, thermal discomfort increases to 31.7 ± 2.1 hours and unmet cooling rises to 9.8 ± 1.3 hours. Drift detection is delayed to 11.3 ± 1.4 hours, and recovery time increases to 18.6 ± 1.9 hours, indicating a pronounced degradation in diagnostic effectiveness and post-event resilience.

Removing explicit uncertainty representation results in 27.4 ± 1.8 discomfort hours and 7.9 ± 1.1 unmet cooling hours, with a detection time of 8.6 ± 1.0 hours and a recovery time of 14.1 ± 1.6 hours. This confirms that uncertainty-aware reasoning materially contributes to stability under climate-induced stress.

When diagnostic memory is disabled, the system experiences 24.9 ± 1.6 discomfort hours and 6.8 ± 0.9 unmet cooling hours. Drift detection occurs after 6.2 ± 0.7 hours, and recovery time increases to 12.7 ± 1.3 hours. These results demonstrate that persistent memory of diagnosed states plays a critical role in accelerating recovery following extreme events.

Collectively, the ablation results confirm that performance gains arise from the integrated interaction of attribution, uncertainty handling, and diagnostic memory rather than from predictive control alone.

6. DISCUSSION

6.1. WHY COGNITIVE OPERATION OUTPERFORMS CONVENTIONAL CONTROL UNDER CLIMATE EXTREMES

Results presented in Section 5 demonstrate that the proposed cognitive digital twin consistently outperforms rule-based control, conventional model predictive control, and non-cognitive digital twin baselines across all evaluated climate-extreme scenarios. Importantly, these improvements are not achieved through more aggressive actuation, relaxed comfort constraints, or increased energy expenditure. Instead, they arise from the system's ability to interpret the causes of performance degradation before modifying control behavior.

Under heatwave-only conditions (S1–S2), performance gains are primarily associated with anticipatory operation. By combining real-time state estimation with uncertainty-aware prediction, the cognitive digital twin mitigates peak demand and thermal stress before deviations propagate into sustained comfort violations. Although conventional MPC also leverages forecasting, it assumes fixed system characteristics and therefore remains vulnerable to gradual shifts in system behavior under prolonged climatic stress.

The distinction between cognitive and non-cognitive operation becomes more pronounced in scenarios involving performance drift (S3–S4). Conventional controllers implicitly treat persistent deviations as control errors, leading to compensatory actions that may mask rather than resolve underlying degradation. In contrast, the cognitive digital twin explicitly diagnoses drift and attributes it to specific mechanisms such as equipment degradation or sensor bias. This diagnostic awareness enables targeted adaptation and avoids unnecessary control oscillations or excessive comfort sacrifice.

6.2. ROLE OF DIAGNOSIS IN PREVENTING COMFORT COLLAPSE

A central insight from this study is that early diagnosis, rather than increased control authority, is the dominant mechanism for preserving comfort during climate extremes. In compound stress scenarios combining heatwaves, operational constraints, and elevated uncertainty (S4–S6), baseline controllers frequently violate comfort bounds despite achieving energy- or demand-related objectives.

The cognitive digital twin mitigates this failure mode by distinguishing between apparent and actual thermal stress. When sensor drift is diagnosed, control actions compensate for biased measurements rather than altering physical setpoints, preventing artificial overheating or overcooling. Similarly, when capacity degradation is identified, control policies are adapted to the reduced effective capacity instead of forcing infeasible operation.

These findings indicate that comfort collapse under extreme conditions is often driven by misinterpretation of system state rather than insufficient control capability. Diagnosis-informed adaptation directly addresses this limitation by aligning control actions with the true underlying system condition.

6.3. INTERACTION BETWEEN CLIMATE STRESS, UNCERTAINTY, AND CONTROL STABILITY

Climate extremes amplify uncertainty in both environmental inputs and system response. Increased temperature variability, forecast errors, and fluctuating internal loads elevate the risk of instability when decision-making relies on point estimates or static models.

By explicitly propagating uncertainty through state estimation and decision-making layers, the proposed framework conditions control actions on confidence bounds rather than deterministic predictions. This

mechanism explains the observed reductions in unmet cooling hours and the improved recovery performance under high-uncertainty scenarios (S5–S6).

Notably, the cognitive digital twin does not attempt to eliminate uncertainty. Instead, it operates robustly in its presence by prioritizing operational stability over short-term efficiency when uncertainty is elevated. This behavior contrasts with conventional MPC, which may overreact to uncertain forecasts and inadvertently exacerbate instability during prolonged heatwaves.

6.4. RESILIENCE AS A DIAGNOSTIC–ADAPTIVE PROPERTY

Resilience in building energy systems is often framed in terms of redundancy, oversizing, or additional capacity. Results from scenario S7 demonstrate that resilience can also emerge from persistent diagnostic awareness and adaptive memory.

Following the removal of climate stress and operational constraints, the cognitive digital twin retains awareness of unresolved system degradation and continues to adapt control actions accordingly. This behavior explains the substantially shorter recovery times and lower accumulated thermal deviation observed relative to baseline approaches.

From this perspective, resilience is not solely a physical property of the system, but a functional property of how effectively the system understands and responds to its evolving condition over time. Cognitive digital twins enable this form of operational resilience without requiring additional hardware or structural modification.

6.5. GENERALIZATION ACROSS CLIMATE REGIMES

Cross-climate evaluation in scenario S8 indicates that the proposed framework generalizes across hot-dry, hot-humid, and temperate climates without climate-specific tuning. While the magnitude of improvement varies with climatic characteristics, consistent benefits are observed in all cases.

This robustness arises from reliance on diagnosis-informed adaptation rather than predefined climate heuristics. By reasoning about deviations relative to expected system behavior, the cognitive digital twin remains effective even when absolute temperature levels, humidity profiles, or load patterns differ substantially.

6.6. IMPLICATIONS FOR BUILDING ENERGY SYSTEM DESIGN AND OPERATION

The findings have several practical implications for building energy system design and operation. First, cognitive digital twins offer a pathway to enhance climate resilience without fundamental changes to building hardware or system architecture. Second, diagnostic awareness reduces reliance on conservative control margins, enabling more efficient operation under stress. Third, embedding cognition within operational loops shifts energy management from reactive correction toward proactive adaptation.

Together, these implications suggest a scalable approach for transitioning existing building energy systems toward intelligent operation capable of maintaining performance under increasing climatic uncertainty.

6.7. SUMMARY OF KEY INSIGHTS

Table 22. Mechanisms linking cognitive capabilities to observed performance gains.

Cognitive capability	Operational effect	Observed outcome
Drift diagnosis	Root-cause identification	Earlier intervention
Cause attribution	Targeted adaptation	Reduced comfort violations
Uncertainty awareness	Robust decision-making	Improved stability
Adaptive memory	Persistent awareness	Faster recovery

7. LIMITATIONS AND FUTURE WORK

While the proposed cognitive digital twin framework demonstrates consistent performance improvements across a wide range of climate-extreme scenarios, several limitations should be acknowledged to appropriately

contextualize the findings and to delineate directions for future research.

7.1. SCOPE OF BUILDING TYPOLOGY AND SYSTEM CONFIGURATION

The present study focuses on a medium-scale commercial office building equipped with a centralized variable air volume HVAC system. This configuration represents a widely deployed and well-understood building typology, enabling controlled and interpretable evaluation. However, the results may not directly generalize to buildings with fundamentally different thermal dynamics, operational patterns, or control architectures, such as residential buildings, healthcare facilities, educational campuses, or naturally ventilated structures.

Future research should extend the proposed framework to alternative building types and HVAC configurations, including decentralized systems, hybrid ventilation strategies, and mixed-mode operation, to assess cognitive operation under more heterogeneous and behavior-sensitive conditions.

7.2. RELIANCE ON SCENARIO-BASED SIMULATION

The evaluation is conducted using scenario-based simulations designed to emulate heatwaves, operational constraints, performance degradation, and sensing disturbances. Although these scenarios are constructed using realistic climate thresholds and physically meaningful disturbance mechanisms, they cannot fully capture the complexity and unpredictability of real-world operation, including unforeseen equipment failures, maintenance actions, or human-in-the-loop interventions.

Field deployment and long-term monitoring studies are therefore required to validate absolute performance levels under real heatwave conditions. Nevertheless, controlled scenario-based stress testing provides a necessary and systematic first step, enabling reproducible benchmarking and causal analysis that would be difficult to achieve in operational buildings.

Accordingly, the reported results should be interpreted as comparative evidence of resilience enhancement across control strategies rather than as guaranteed performance outcomes in real-world deployments.

7.3. COMPUTATIONAL AND IMPLEMENTATION CONSIDERATIONS

The cognitive digital twin integrates state estimation, diagnostic reasoning, uncertainty handling, and predictive control within a closed-loop architecture. While the computational requirements remain manageable at the time scales and system size considered in this study, real-time deployment in larger buildings, high-resolution multi-zone models, or district-scale applications may impose additional computational constraints.

Future work should investigate scalable implementation strategies, including model reduction techniques, hierarchical and distributed control architectures, and edge–cloud coordination to balance computational efficiency with diagnostic fidelity.

7.4. TREATMENT OF UNCERTAINTY AND MODEL ASSUMPTIONS

Uncertainty is explicitly represented within the proposed framework; however, uncertainty models are parameterized using assumed statistical characteristics derived from typical sensor performance and operational variability. In practice, uncertainty structures may evolve over time, exhibit non-Gaussian behavior, or differ substantially across buildings and climates.

Adaptive uncertainty learning, online calibration, and data-driven characterization of disturbance distributions represent promising directions for improving robustness while reducing reliance on predefined assumptions.

7.5. PATHWAYS FOR EXTENSION AND INTEGRATION

Beyond single-building operation, cognitive digital twins offer opportunities for extension and integration at broader spatial and functional scales. Future studies may explore coordination among multiple cognitive twins at district or campus levels, interaction with grid-level signals and demand flexibility programs, and integration with on-site renewable generation and energy storage systems.

The framework may also be extended toward fault prognosis, maintenance planning, and lifecycle-aware optimization, enabling cognitive digital twins to support long-term operational decision-making beyond short-term control.

7.6. SUMMARY OF LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

Table 23. Summary of limitations and corresponding future research directions.

Identified limitation	Implication	Future research direction
Single building typology	Limited generalization	Diverse building types and HVAC systems
Scenario-based evaluation	No field validation	Long-term deployment studies
Computational complexity	Scalability constraints	Hierarchical and distributed control
Fixed uncertainty models	Potential mismatch	Adaptive uncertainty learning
Building-level focus	Limited system integration	District- and grid-level coordination

The computational complexity of the cognitive digital twin is primarily governed by the recursive state estimation and optimization layers. In the Extended Kalman Filter (EKF), the dominant computational cost arises from covariance matrix propagation and inversion operations.

Let n denote the dimension of the state vector. The matrix inversion required during the Kalman gain update results in a computational complexity on the order of O(n^3) per control interval.

For the case study presented in this work, the state dimension remains sufficiently low due to aggregated thermal states and latent performance variables, enabling real-time execution at a 15-minute control interval on standard computing hardware.

However, for large-scale multi-zone buildings or high-resolution digital twins with detailed envelope discretization, the state dimension may increase substantially. In such cases, the computational burden can become significant.

To mitigate scalability challenges, several strategies may be employed:

1. state aggregation and reduced-order modeling,
2. decentralized or zone-partitioned estimation,
3. square-root filtering techniques for improved numerical efficiency,
4. transition to unscented or ensemble-based filters where appropriate.

Therefore, while the proposed framework is computationally tractable for medium-scale deployments, careful architectural design is required for skyscraper-scale digital twins to maintain real-time operability.

8. CONCLUSIONS

This study investigated the limitations of conventional building energy control strategies under climate-extreme conditions and introduced a cognitive digital twin framework for diagnosis-informed and adaptive energy operation. As heatwaves become more frequent, prolonged, and operationally disruptive, the results confirm that control approaches relying on static models and reactive feedback are increasingly unable to sustain energy efficiency, thermal comfort, and operational stability.

The proposed cognitive digital twin integrates physics-based building energy modeling, real-time sensing, data assimilation, diagnostic reasoning, and diagnosis-conditioned control within a closed operational loop. Unlike conventional digital twins that primarily support monitoring or short-term prediction, the framework explicitly interprets deviations between expected and observed behavior, enabling early detection and causal attribution of climate-induced performance drift. This diagnostic capability allows control actions to be adapted proactively, preventing degradation from escalating into sustained comfort violations or unstable operation.

Evaluation across a structured set of climate-extreme scenarios, including heatwaves combined with equipment degradation, sensor bias, demand response constraints, and elevated uncertainty, demonstrates consistent and statistically significant performance gains. On average, the cognitive digital twin reduces energy use intensity by 10.9% and peak electric demand by 13.4% relative to rule-based control. Thermal resilience is substantially improved, with 54.6% fewer discomfort hours and 58.2% fewer unmet cooling hours under extreme operating conditions. Performance drift is detected 3.3–3.4 times earlier than with non-cognitive digital twin baselines, enabling timely mitigation. Following extreme events, diagnosis-aware control

accelerates comfort recovery by approximately 46%, indicating sustained resilience beyond the heatwave period.

Importantly, these improvements are not achieved through more aggressive actuation, relaxed comfort constraints, or additional system capacity. Instead, they arise from early diagnosis, explicit root-cause attribution, and adaptive decision-making conditioned on diagnosed system states. The results highlight that resilience in building energy systems is not solely a function of physical redundancy or oversizing, but an emergent property of sustained operational awareness and cognition.

Overall, this work demonstrates that cognitive digital twins can function as self-interpreting and adaptive building energy systems, providing a scalable and hardware-agnostic pathway toward climate-resilient operation under increasing environmental uncertainty. By operationalizing diagnosis and adaptation directly within energy control loops, the proposed framework establishes a robust foundation for intelligent building operation capable of responding effectively to future climate challenges.

AUTHOR CONTRIBUTIONS

Seyed Reza Samaei: Development of the review methodology, literature selection strategy, analytical structure, synthesis approach, critical analysis, interpretation of the literature, drafted, revised, refined, reviewed and approved the manuscript. **James Riffat**: Development of the review methodology, literature selection strategy, analytical structure, synthesis approach, critical analysis, interpretation of the literature, drafted, revised, refined, reviewed and approved the manuscript.

COMPETING INTERESTS

The authors declare that they have no competing financial or non-financial interests.

DATA ACCESSIBILITY

All data supporting the findings of this study are included within the article. Additional information may be made available by the authors upon reasonable request.

ETHICAL APPROVAL

Not applicable. This study is based exclusively on the review and analysis of published literature and did not involve human participants or animal subjects.

FUNDING

This research received no external funding.

REFERENCES

AlBalkhy, W., Karmaoui, D., Ducoulombier, L., Lafhaj, Z. and Linner, T., 2024. Digital twins in the built environment: Definition, applications, and challenges. Automation in Construction, 162, 105368. DOI: https://doi.org/10.1016/j.autcon.2024.105368.

Chen, J., Zhang, L., Li, Y., Shi, Y., Gao, X. and Hu, Y., 2022. A review of computing-based automated fault detection and diagnosis of heating, ventilation and air conditioning systems. Renewable and Sustainable Energy Reviews, 161, 112395. DOI: https://doi.org/10.1016/j.rser.2022.112395.

Crawley, D.B., Lawrie, L.K., Winkelmann, F.C., Buhl, W.F., Huang, Y.J., Pedersen, C.O., Strand, R.K., Liesen, R.J., Fisher, D.E., Witte, M.J. and Glazer, J., 2001. EnergyPlus: Creating a new-generation building energy simulation program. Energy and Buildings, 33(4), 319–331. DOI: https://doi.org/10.1016/S0378-7788(00)00114-6.

de Dear, R.J. and Brager, G.S., 2002. Thermal comfort in naturally ventilated buildings: Revisions to ASHRAE Standard 55. Energy and Buildings, 34(6), 549–561. https://doi.org/10.1016/S0378-7788(02)00005-1.

Drgoňa, J., Arroyo, J., Figueroa, I.C., Blum, D., Arendt, K., Kim, D., Perarnau Ollé, E., Oravec, J., Wetter, M., Vrabie, D.L. and Helsen, L., 2020. All you need to know about model predictive control for buildings. Annual Reviews in Control, 50, 190–232. DOI: https://doi.org/10.1016/j.arcontrol.2020.09.001.

Du, Z. and Jin, X., 2007. Detection and diagnosis for sensor fault in HVAC systems. Energy Conversion and Management, 48(3), 693–702. DOI: https://doi.org/10.1016/j.enconman.2006.09.023.

Du, Z., Jin, X. and Wu, L., 2007. Fault detection and diagnosis based on improved PCA with JAA method in VAV systems. Building and Environment, 42(9), 3221–3232. DOI: https://doi.org/10.1016/j.buildenv.2006.08.011.

Ejenakevwe, K.A., Xie, L., Wang, J., Hurt, R. and Song, L., 2025. An automated framework based on RC model and GA optimization for calibrating coupled residential buildings and HVAC systems. Building and Environment, 283, 113368. DOI: https://doi.org/10.1016/j.buildenv.2025.113368.

Gavaldà-Torrellas, O., Monsalvete, P., Ranjbar, S. and Eicker, U., 2025. The urban building energy retrofitting tool: An open-source framework to help foster building retrofitting using a life cycle costing perspective—First results for Montréal. Smart Cities, 8(1), 17. DOI: https://doi.org/10.3390/smartcities8010017.

Giuzio, G.F., Russo, G., Cipolla, G., Pompei, L., Stasi, R. and Buonomano, A., 2025. User-centric urban energy modelling: A reduced order model approach based on pre-calibrated archetypes dataset. Energy and Buildings, 339, 115768. DOI: https://doi.org/10.1016/j.enbuild.2025.115768.

Hodavand, F., Ramaji, I.J. and Sadeghi, N., 2023. Digital twin for fault detection and diagnosis of building operations: A systematic review. Buildings, 13(6), 1426. DOI: https://doi.org/10.3390/buildings13061426.

Hodavand, F., Ramaji, I., Sadeghi, N. and Zandi Goharrizi, S., 2025. Digital twin-enabled framework for intelligent monitoring and anomaly detection in multi-zone building systems. Buildings, 15(22), 4030. DOI: https://doi.org/10.3390/buildings15224030.

Hosamo, H.H., Nielsen, H.K., Alnmr, A.N., Svennevig, P.R. and Svidt, K., 2022. A review of the digital twin technology for fault detection in buildings. Frontiers in Built Environment, 8, 1013196. DOI: https://doi.org/10.3389/fbuil.2022.1013196.

ISO/IEC 30173:2023, "Digital twin — Concepts and terminology," International Organization for Standardization, Geneva, Switzerland, 2023.

Katipamula, S. and Brambley, M.R., 2005. Methods for fault detection, diagnostics, and prognostics for building systems—A review, Part I. HVAC&R Research, 11(1), 3–25. DOI: https://doi.org/10.1080/10789669.2005.10391123.

Lee, K.-H. and Braun, J.E., 2008a. Evaluation of methods for determining demand-limiting setpoint trajectories in buildings using short-term measurements. Building and Environment, 43(10), 1769–1783. DOI: https://doi.org/10.1016/j.buildenv.2007.11.003.

Lee, K.-H. and Braun, J.E., 2008b. Development of methods for determining demand-limiting setpoint trajectories in buildings using short-term measurements. Building and Environment, 43(10), 1755–1768. DOI: https://doi.org/10.1016/j.buildenv.2007.11.004.

Liu, Y., Ji, T., Guo, X., Xu, X. and Polzer, J., 2025. Cognitive digital twin frameworks in manufacturing: A critical survey, evaluation criteria, and future directions. Journal of Manufacturing Systems, 83, 597–611. DOI: https://doi.org/10.1016/j.jmsy.2025.10.004.

Lu, Q., Xie, X., Parlikad, A.K. and Schooling, J.M., 2020. Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance. Automation in Construction, 118, 103277. DOI: https://doi.org/10.1016/j.autcon.2020.103277.

Maturo, A., Vallianos, C., Delcroix, B., Buonomano, A. and Athienitis, A., 2024. Automated model order reduction for building thermal load prediction using smart thermostats data. Journal of Building Engineering, 96, 110492. DOI: https://doi.org/10.1016/j.jobe.2024.110492.

Melgaard, S.P., Andersen, K.H., Marszal-Pomianowska, A., Jensen, R.L. and Heiselberg, P.K., 2022. Fault detection and diagnosis encyclopedia for building systems: A systematic review. Energies, 15(12), 4366. DOI: https://doi.org/10.3390/en15124366.

Naderi, S., Pignatta, G., Heslop, S., MacGill, I. and Chen, D., 2022. Demand response via pre-cooling and solar pre-cooling: A review. Energy and Buildings, 272, 112340. DOI: https://doi.org/10.1016/j.enbuild.2022.112340.

Nguyen, A.-T., Reiter, S. and Rigo, P., 2014. A review on simulation-based optimization methods applied to building performance analysis. Applied Energy, 113, 1043–1058. DOI: https://doi.org/10.1016/j.apenergy.2013.08.061.

Osama, Z., 2024. The digital twin framework: A roadmap to the development of user-centred digital twin in

the built environment. Journal of Building Engineering, 98, 111081. DOI: https://doi.org/10.1016/j.jobe.2024.111081.

Pérez-Lombard, L., Ortiz, J. and Pout, C., 2008. A review on buildings energy consumption information. Energy and Buildings, 40(3), 394–398. DOI: https://doi.org/10.1016/j.enbuild.2007.03.007.

Santamouris, M., Cartalis, C., Synnefa, A. and Kolokotsa, D., 2015. On the impact of urban heat island and global warming on the power demand and electricity consumption of buildings—A review. Energy and Buildings, 98, 119–124. DOI: https://doi.org/10.1016/j.enbuild.2014.09.052.

Shi, Z. and O'Brien, W., 2019. Sequential state prediction and parameter estimation with constrained dual extended Kalman filter for building zone thermal responses. Energy and Buildings, 183, 538–546. DOI: https://doi.org/10.1016/j.enbuild.2018.11.024.

Subramanian, A.S.R., Gundersen, T. and Adams II, T.A., 2018. Modeling and simulation of energy systems: A review. Processes, 6(12), 238. DOI: https://doi.org/10.3390/pr6120238.

Sun, Y., Wang, S., Xiao, F. and Huang, G., 2012. A study of pre-cooling impacts on peak demand limiting in commercial buildings. HVAC&R Research, 18(6), 1098–1111. DOI: https://doi.org/10.1080/10789669.2012.672361.

Tuhaise, V.V., Tah, J.H.M. and Abanda, F.H., 2023. Technologies for digital twin applications in construction. Automation in Construction, 152, 104931. DOI: https://doi.org/10.1016/j.autcon.2023.104931.

Wang, J., Wei, Z., Zhu, Y., Zheng, C., Li, B. and Zhai, X., 2023. Demand response via optimal pre-cooling combined with temperature reset strategy for air conditioning system: A case study of office building. Energy, 282, 128751. DOI: https://doi.org/10.1016/j.energy.2023.128751.

Wang, S. and Xiao, F., 2004. Detection and diagnosis of AHU sensor faults using principal component analysis method. Energy Conversion and Management, 45(17), 2667–2686. DOI: https://doi.org/10.1016/j.enconman.2003.12.008.

Xie, X., Merino, J., Moretti, N., Pauwels, P., Chang, J.Y. and Parlikad, A., 2023. Digital twin enabled fault detection and diagnosis process for building HVAC systems. Automation in Construction, 146, 104695. DOI: https://doi.org/10.1016/j.autcon.2022.104695.

Zhang, R. and Hong, T., 2017. Modeling of HVAC operational faults in building performance simulation. Applied Energy, 202, 178–188. DOI: https://doi.org/10.1016/j.apenergy.2017.05.153.