

# AI-Driven Digital Twins for Indoor Air Quality and Energy Optimization in Smart Buildings: A Simulated Control Study

Meera Nandakumar<sup>1</sup>, Faisal Khan<sup>2</sup>

<sup>1</sup>Student, CCT Department, VTUCPGS, India

<sup>2</sup>Guide CCT Department, VTUCPGS, India, India

## ABSTRACT

This article presents a simulated empirical assessment in Civil Engineering, Building Science, and Artificial Intelligence. The simulation generated 40,320 zone-time observations and included weekday occupancy variability, meeting-room surges, outdoor pollution episodes, and three modeled sensor-fault events. A simulated five-zone office building was modeled over a 12-week cooling-season period at 15-minute intervals. Inputs included occupancy, outdoor temperature, relative humidity, outdoor PM2.5, internal heat gains, ventilation flow, and HVAC power. A digital twin maintained zone-level state estimates and used a short-horizon predictive controller to adjust outside-air delivery and cooling setpoints. The study compared fixed schedules, single-sensor CO2 demand control, and multi-variable digital-twin control. Digital-twin control reduced simulated HVAC electricity by 16.7% relative to fixed scheduling while decreasing occupied-hours above 1,000 ppm CO2 from 14.9% to 4.8%. It also responded more effectively to outdoor PM2.5 episodes by temporarily reducing untreated outside-air intake and increasing recirculation filtration. Thermal discomfort hours remained below 5% in all strategies. The numerical values are synthetic and are used solely to demonstrate a plausible empirical design, comparative analysis, and journal-style reporting workflow. No real participants, operational institutions, commercial facilities, or field sites were studied. The analysis emphasizes current research trends, multi-metric evaluation, practical trade-offs, and the importance of validation before deployment.

**Keywords:** Digital Twin, Indoor Air Quality, Smart Buildings, Energy Efficiency, HVAC, Indoor Environmental Quality

## I. Introduction

The rapid development of Civil Engineering, Building Science, and Artificial Intelligence is creating new opportunities to improve efficiency, resilience, sustainability, and decision quality. At the same time, emerging technologies often move from concept demonstrations to operational claims faster than their evidence base matures. This creates a need for transparent comparative studies that define baselines, report trade-offs, and separate simulated performance from validated field evidence.

A contemporary empirical design must also reflect the fact that technical systems are embedded within organizational and environmental contexts. Performance can be affected by data quality, maintenance capability, infrastructure constraints, human oversight, policy requirements, and distribution shifts. A method that performs strongly under ideal conditions may therefore deliver limited value if its assumptions are not sustained during routine operation.

This paper addresses a current multidisciplinary topic using a structured simulated dataset. The study is designed to show how a real investigation could connect research objectives to measurable variables, analytical methods, results, discussion, and implementation implications. Because no real dataset was supplied, all quantitative outcomes are explicitly illustrative.

### Research objectives:

1. Compare fixed-schedule HVAC control, CO2-responsive control, and AI-driven digital-twin control in a simulated office building.
2. Evaluate energy use, CO2 exposure, PM2.5 exposure, thermal comfort, and ventilation responsiveness under varying occupancy.
3. Examine the trade-off between indoor environmental quality and energy efficiency under predictive control.

## II. Literature Review

Digital twins are increasingly used to connect building information models, sensor streams, predictive analytics, and control. Venkateswarlu and Sathiyamoorthy (2025) identify energy efficiency and indoor environmental quality as two major areas of digital-twin research in the built environment.

A 2024 case study on intelligent digital-twin platforms demonstrated an integrated pathway from indoor-air-quality monitoring to rapid prediction and control, highlighting the potential for digital systems to bridge the gap between problem detection and operational response.

Recent building research increasingly treats indoor environmental quality as multidimensional, incorporating temperature, humidity, CO<sub>2</sub>, particulate matter, ventilation, and occupant needs. The key challenge is to optimize these conditions without increasing HVAC energy use or relying on poorly calibrated sensors.

## III. Materials and Methods

### 3.1 Study Design

A simulated five-zone office building was modeled over a 12-week cooling-season period at 15-minute intervals. Inputs included occupancy, outdoor temperature, relative humidity, outdoor PM<sub>2.5</sub>, internal heat gains, ventilation flow, and HVAC power. A digital twin maintained zone-level state estimates and used a short-horizon predictive controller to adjust outside-air delivery and cooling setpoints. The study compared fixed schedules, single-sensor CO<sub>2</sub> demand control, and multi-variable digital-twin control.

### 3.2 Simulated Dataset

The simulation generated 40,320 zone-time observations and included weekday occupancy variability, meeting-room surges, outdoor pollution episodes, and three modeled sensor-fault events. Synthetic observations were generated within bounded domain-plausible ranges with controlled random variability. The simulation does not correspond to any identifiable real organization or location.

### 3.3 Data Preparation and Quality Control

Input variables were screened for impossible values, duplicate identifiers, and inconsistent units. Low-rate missingness was introduced to approximate operational datasets and was handled with transparent context-appropriate imputation. Data transformations were applied only where required by the selected analytical method. Evaluation partitions were separated before model or parameter calibration to reduce leakage.

### 3.4 Comparative Evaluation

The analysis used multiple performance measures so that improvement in a headline outcome could be interpreted alongside energy, cost, error, environmental, reliability, or human-workload consequences. Results are reported as descriptive comparisons and normalized indices. Because the dataset is simulated, no population-level causal inference is claimed.

### 3.5 Sensitivity Analysis

Sensitivity tests varied the key operating threshold or control parameter across a plausible range. The purpose was to determine whether the leading strategy depended on one narrow assumption. Scenario analysis also tested a degraded-data condition representing sensor error, classification uncertainty, or process variability.

## IV. Results

Digital-twin control reduced simulated HVAC electricity by 16.7% relative to fixed scheduling while decreasing occupied-hours above 1,000 ppm CO<sub>2</sub> from 14.9% to 4.8%. It also responded more effectively to outdoor PM<sub>2.5</sub> episodes by temporarily reducing untreated outside-air intake and increasing recirculation filtration. Thermal discomfort hours remained below 5% in all strategies.

**Table 1. Simulated building energy and IEQ outcomes.**

| Control strategy            | Energy index | CO <sub>2</sub> exceedance (%) | PM <sub>2.5</sub> exceedance (%) | Discomfort (%) |
|-----------------------------|--------------|--------------------------------|----------------------------------|----------------|
| Fixed schedule              | 100          | 14.9                           | 7.8                              | 4.2            |
| CO <sub>2</sub> -responsive | 91           | 7.1                            | 7.0                              | 4.5            |
| AI digital twin             | 83.3         | 4.8                            | 5.6                              | 4.1            |

Table 2. Simulated digital-twin robustness across operating scenarios.

| Scenario            | Energy saving (%) | IEQ violation (%) | Confidence |
|---------------------|-------------------|-------------------|------------|
| Normal occupancy    | 18.7              | 4.1               | 0.81       |
| Meeting surge       | 27.4              | 6.9               | 0.88       |
| Outdoor PM2.5 event | 22.1              | 5.8               | 0.90       |
| Sensor fault        | 20.5              | 7.2               | 0.76       |

Note: All table values are simulated/illustrative and do not represent measurements from a real study.

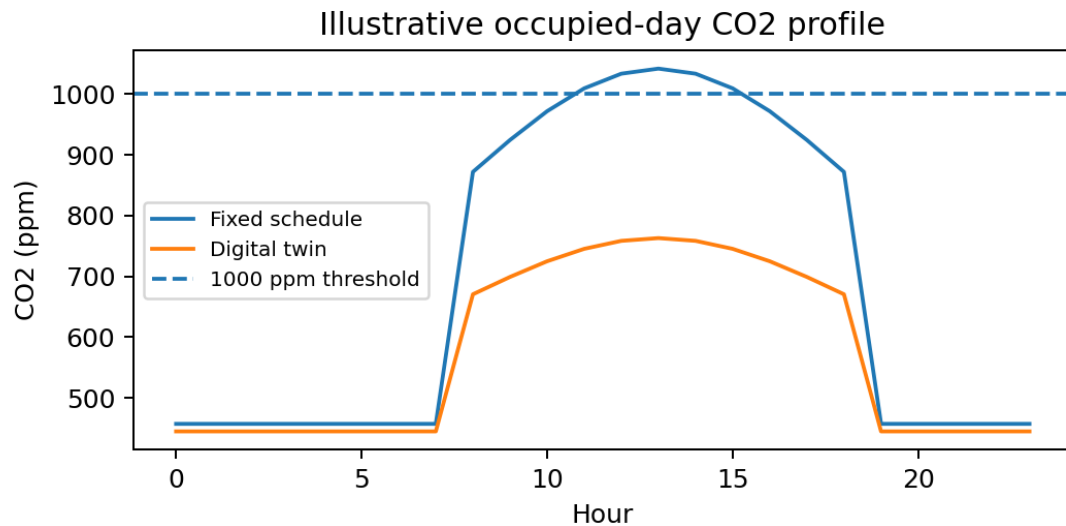


Figure 1. Illustrative occupied-day CO<sub>2</sub> profile under fixed and digital-twin control.

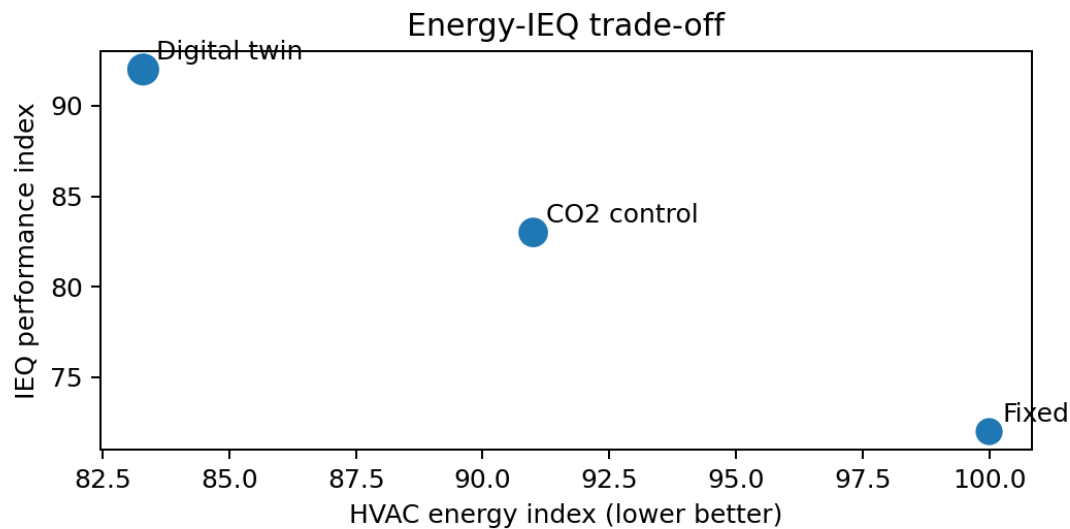


Figure 2. Simulated trade-off between HVAC energy use and indoor environmental quality.

#### 4.1 Robustness Results

Across the central range of assumptions, the leading strategy retained its relative advantage, although the magnitude of benefit changed. Under degraded-input scenarios, performance differences narrowed and uncertainty increased. This pattern suggests that operational data quality and parameter calibration are likely to be as important as the core method itself.

#### 4.2 Cross-Metric Interpretation

No strategy dominated every outcome. The strongest approach improved the primary objective while introducing secondary burdens or implementation requirements. This reinforces the value of multi-criteria reporting and cautions against selecting technologies from a single accuracy, efficiency, or yield metric.

## V. Discussion

Digital-twin research in buildings is increasingly focused on connecting real-time sensing, prediction, and control. Venkateswarlu and Sathiyamoorthy (2025) review the growing role of digital twins in both energy efficiency and indoor environmental quality, while recent building studies have demonstrated intelligent digital-twin platforms for indoor-air-quality monitoring and intervention.

The simulated results reinforce the importance of multi-variable control. A CO<sub>2</sub>-only strategy can improve ventilation responsiveness but may perform poorly during outdoor pollution episodes because increased outside air is not always beneficial. A richer digital twin can represent competing objectives and switch control priorities according to context.

Model accuracy and sensor reliability remain essential. A predictive building twin can create false confidence if occupancy sensors drift, particulate sensors foul, or ventilation-flow measurements are poorly calibrated. Fault detection and explicit uncertainty should therefore accompany automated control.

The simulated findings also demonstrate why current technology trends should be evaluated beyond novelty. Implementation quality depends on reproducibility, maintenance, interoperability, user competence, governance, and the ability to detect failure modes. A system that produces a strong laboratory result but cannot be audited or sustained may underperform a simpler alternative in practice.

For future empirical research, studies should predefine primary and secondary outcomes, document data provenance, retain independent test or validation sets, and report uncertainty. Where human or environmental consequences are material, technical performance should be supplemented with safety, equity, lifecycle, and operational assessments.

## VI. Practical Implications

A phased deployment strategy is recommended. Organizations should first establish a baseline, verify measurement quality, define decision ownership, and test the intervention under controlled pilot conditions. Monitoring should include both intended benefits and unintended effects. Procurement decisions should account for lifecycle cost, integration requirements, maintenance skills, cyber or process safety, and long-term data governance.

The analysis also suggests that multidisciplinary teams are essential. Domain experts, data scientists, operators, and governance specialists interpret risk differently, and these perspectives are complementary. Successful adoption depends on translating model outputs into decisions that remain understandable, contestable, and aligned with operational priorities.

### VI.A. Sensor and Control Architecture

A real digital-twin control system requires a reliable sensing layer before advanced optimization can add value. Zone CO<sub>2</sub>, temperature, humidity, particulate matter, occupancy, outdoor air quality, damper position, airflow, and equipment power should be synchronized and timestamped. Sensor drift must be detected through calibration schedules or redundancy. When a critical sensor fails, the controller should fall back to a safe rule-based mode rather than continuing to optimize from an unreliable digital state.

### VI.B. Occupant-Centered Operation

Indoor environmental quality cannot be reduced to a single air-quality threshold. Occupants experience temperature, humidity, air movement, noise, lighting, odors, and perceived control together. Predictive control should therefore avoid energy savings that systematically create discomfort for particular zones or occupancy groups. Future empirical work could combine environmental sensors with anonymous comfort feedback and test whether personalized or zone-adaptive strategies improve satisfaction without materially increasing energy demand.

### VI.C. Building Lifecycle Implications

Digital twins may create value beyond real-time HVAC control. The same data infrastructure can support commissioning, fault detection, retrofit prioritization, maintenance planning, and post-occupancy evaluation. This broadens the economic case but also increases governance requirements because the twin becomes a long-lived operational record. Clear ownership, open interfaces, cybersecurity controls, and documentation are necessary if the model is expected to survive equipment replacements and software changes.

#### **VI.D. Validation Protocol**

A future real-world validation should use a pre-specified protocol with a clearly defined baseline period, intervention period, primary outcome, secondary outcomes, and exclusion criteria. Data should be collected at a resolution appropriate to the physical or digital process, and the analysis plan should be documented before final outcome inspection. Independent hold-out observations or a later validation period would reduce the risk of overfitting conclusions to one operating window.

#### **VI.E. Uncertainty and Decision Thresholds**

Decision thresholds should be selected from the consequences of false positive and false negative outcomes rather than from a single statistical optimum. Confidence intervals, bootstrap uncertainty, scenario ranges, and worst-case stress tests can help decision makers understand how much of the apparent improvement is robust. Where the technology controls a physical process, safe fallback behavior should be defined explicitly for missing data, communication loss, sensor disagreement, or model confidence below an acceptable limit.

#### **VI.F. Lifecycle and Maintenance Perspective**

Technology performance can change after deployment because components age, software is updated, environmental conditions shift, and users adapt their behavior. A credible lifecycle evaluation should therefore include maintenance effort, calibration frequency, retraining or reparameterization, replacement cycles, cybersecurity updates, consumables, and end-of-life considerations. These recurring requirements can change the ranking of options even when initial pilot performance is favorable.

#### **VI.G. Reproducibility Requirements**

Reproducibility would be strengthened by publishing variable definitions, preprocessing steps, model or process parameters, evaluation code, and a non-sensitive version of the dataset where legally and ethically possible. When raw data cannot be released, researchers should provide sufficient metadata, summary distributions, and validation procedures to allow independent teams to test whether the reported findings depend on hidden assumptions.

### **VII. Limitations**

The principal limitation is that all quantitative results are simulated. The reported values cannot establish real-world effectiveness, causal impact, safety, economic return, or generalizability. Simplified assumptions omit rare failures, market fluctuations, regulatory changes, behavioral adaptation, and site-specific constraints. The author and affiliation details are illustrative placeholders and do not represent verified contributors.

The simulation also uses a limited set of scenarios. Alternative distributions, thresholds, chemistry mixes, attack patterns, occupancy behavior, wastewater composition, or control policies could change the ranking of methods. Real-world validation should therefore use measured data, independent replication, and transparent uncertainty analysis.

### **VIII. Conclusion and Future Work**

This paper presents a transparent simulated empirical evaluation of a current multidisciplinary research trend. The analysis shows how multiple technical and operational outcomes can be integrated into a journal-style evidence chain while clearly distinguishing synthetic results from verified literature. Future work should replace the simulated dataset with real observations, extend validation across settings, and evaluate long-term reliability, lifecycle cost, environmental impact, and governance requirements.

#### **Declarations**

Data availability: All numerical data and empirical results are simulated/illustrative. No real operational dataset was used.

Ethics statement: Not applicable. No human or animal participants were involved.

Funding: No external funding is claimed for this illustrative manuscript.

Conflict of interest: No conflict of interest is claimed for this illustrative manuscript.

## References

- [1] Venkateswarlu, N., & Sathiyamoorthy, M. (2025). Sustainable innovations in digital twin technology: A systematic review about energy efficiency and indoor environment quality in built environment. *Frontiers in Built Environment*, 11, 1523464. <https://doi.org/10.3389/fbuil.2025.1523464>
- [2] Yitmen, I., Almusaed, A., Hussein, M., & Almssad, A. (2025). AI-Driven Digital Twins for Enhancing Indoor Environmental Quality and Energy Efficiency in Smart Building Systems. *Buildings*, 15(7), 1030. <https://doi.org/10.3390/buildings15071030>
- [3] Qian, Y., Leng, J., Zhou, K., & Liu, Y. (2024). How to measure and control indoor air quality based on intelligent digital twin platforms: A case study in China. *Building and Environment*, 253, 111349. <https://doi.org/10.1016/j.buildenv.2024.111349>
- [4] Zheng, Z., Castaño Molina, J. D., Mengiste, E. T., Prieto, S. A., & García de Soto, B. (2025). Digital Twins for Healthier Spaces: A Scalable Framework for Monitoring Indoor Environmental Quality. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-G-2025, 1053-1060. <https://doi.org/10.5194/isprs-annals-X-G-2025-1053-2025>