An Environment-Aware Indoor-Outdoor Integrated Digital Twin for Healthy Mobility

Yan Li, Chenming Ye, Wenxuan Shi, Wenqing Zhang, Yuyang Zhang, Teng Hu, Zhizhong Kang*

School of Land Science and Technology, China University of Geosciences (Beijing), Beijing, China – yanli.yny@cugb.edu.cn; yechenming@email.cugb.edu.cn; shiwenxuan@email.cugb.edu.cn; zhangwenqing@email.cugb.edu.cn; yuyond@ncut.edu.cn; huteng@cugb.edu.cn; zzkang@cugb.edu.cn
*Corresponding author

Keywords: Digital Twin; Indoor-Outdoor Integration; LOD4 Semantic Modeling; Geometry-Physics Coupling; Mobile Sensing.

Abstract

Existing building digital twins treat indoor environments as static geometric containers, ignoring the dynamic coupling between ventilation structure states and indoor environmental quality. Furthermore, managing indoor and outdoor spaces as separate data silos prevents the continuous assessment of occupant exposure across building boundaries. This paper proposes an environment-aware, indoor-outdoor integrated digital twin framework coupling geometric entity states with physical environmental fields for healthy mobility assessment. The framework utilizes a three-layer architecture. First, the Geometric-Semantic Layer provides a seamless LOD4 model with topologically stitched spaces, modeling ventilation facilities as first-class entities with mutable state attributes (Full Closed, Half Open, Full Open). Second, the Physical Field Layer maps mobile sensing data (PM2.5, CO2) onto semantic entities using a semantic-constrained method, treating walls and closed windows as aggregation barriers. Finally, the Behavioral Response Layer combines entity-level pollution values with pedestrian counts to compute a cumulative Crowd Exposure Index (CEI). Implemented on a Cesium platform, the framework was validated through a week-long university building experiment. Results show indoor PM2.5 in a fully enclosed study room averaged 61.2 μg/m³—1.6 times the outdoor level and 4.1 times the WHO guideline. This resulted in a CEI 12 times higher than in outdoor transit areas. Semantic correlation confirms the "Full Closed" window state primarily drives pollutant accumulation. This validates the framework's core geometry-physics coupling, demonstrating its potential to guide intelligent ventilation interventions and healthy building management.

1. Introduction

People spend approximately 90% of their time in indoor environments, where concentrations of certain pollutants can be two to five times higher than those found outdoors (EPA, 2023). In high-density public buildings such as university campuses, the combination of poor ventilation, high occupancy, and prolonged dwell times can create substantial yet under-quantified cumulative health risks. While outdoor air quality has been widely monitored through urban sensor networks and spatial interpolation techniques (Zhang et al., 2025), continuous environmental mapping across the critical transition zone — where occupants move from outdoor paths into enclosed indoor rooms—remains a major spatial modeling challenge.

Digital Twin (DT) technology offers a promising paradigm for bridging this indoor-outdoor gap. However, existing spatial modeling frameworks for building DTs suffer from critical topological and semantic limitations. First, building DTs are predominantly treated as static geometric containers. Although they faithfully represent the 3D coordinates of walls and roofs, they fail to model the mutable semantic attributes of building entities — particularly the operational states of ventilation structures (e.g., windows fully closed vs. half open) (Ding et al., 2023). Second, indoor and outdoor environments are managed as disjointed data silos: Building Information Modeling (BIM) handles the interior, while Geographic Information Systems (GIS) manage the exterior, lacking topological continuity across the building envelope. This prevents continuous tracking of occupant mobility and environmental exposure from outdoor transit paths into specific indoor zones (Zhao et al., 2025). Furthermore, mapping sparse mobile sensing data into continuous indoor fields using unconstrained spatial interpolation (e.g., Inverse Distance Weighting) inherently produces physically incorrect "through-wall" artifacts, as these traditional algorithms ignore the complex 3D geometric boundaries of indoor topographies.

To address these spatial modeling challenges, this paper proposes an environment-aware, indoor-outdoor integrated digital twin framework. Moving beyond the "static geometric mirror" paradigm, our approach establishes a topologically stitched LOD4 model enriched with dynamic ventilation states extracted from SLAM point clouds. We introduce a novel semantic-constrained mapping method that actively utilizes 3D geometric boundaries as interpolation barriers, thereby preventing through-wall artifacts. By overlaying pedestrian flows onto this semantically constrained spatial field, we compute a cumulative Crowd Exposure Index (CEI) to quantify continuous human mobility risks.

The framework is implemented on a Cesium-based geospatial platform and validated through a multi-sensor mapping experiment. The core contributions of this study to the ISPRS community are summarized as follows:
(1) An indoor-outdoor integrated digital twin modeling framework: We propose a hierarchical DT architecture that topologically bridges indoor spaces and outdoor networks, shifting from static geometry visualization to dynamic, state-aware spatial modeling.
(2) Semantic state extraction and topological stitching from SLAM point clouds: We construct a seamless LOD4 model where transition nodes (doors) stitch the indoor-outdoor graph. Crucially, ventilation facilities are modeled as first-class entities with mutable operational states directly extracted from SLAM point clouds, enhancing the semantic richness of the 3D model.
(3) Semantic-constrained spatial interpolation for indoor environments: To overcome the "through-wall" artifacts of

traditional spatial interpolation in complex indoor topographies, we develop a mapping method that uses LOD4 geometric boundaries and dynamic window states as physical constraints, significantly improving the spatial accuracy of indoor environmental fields.
(4) Spatio-temporal analysis of healthy mobility: Validated through a campus experiment, the integration of spatial mapping and real-time pedestrian flows enables the quantification of a continuous Crowd Exposure Index (CEI), demonstrating how dynamic 3D semantics can actively support spatio-temporal environmental analysis.

2. Related Work

This section reviews the three technical threads that converge in our work — indoor-outdoor integrated digital twin modeling, environmental field integration in building DTs, and mobile sensing with spatial interpolation for indoor environments—and identifies the gaps at their intersection.

2.1 Indoor-Outdoor Integrated Digital Twin Modeling

The construction of building digital twins has advanced rapidly over the past decade, driven by the convergence of Building Information Modeling (BIM) and Geographic Information Systems (GIS). BIM provides rich semantic and geometric representations of building interiors, while GIS excels at managing large-scale outdoor spatial data including terrain, transportation networks, and urban infrastructure. The integration of these two domains — commonly referred to as GeoBIM—has been recognized as essential for applications that span the indoor-outdoor boundary, such as emergency evacuation, pedestrian navigation, and facility management (Ma and Ren, 2017; Teo and Cho, 2016).

The CityGML standard, particularly at Level of Detail 4 (LOD4), provides a framework for representing both exterior building shells and interior structures including rooms, corridors, and building installations (Gröger et al., 2012). CityGML 3.0 further extends this capability by allowing indoor modeling at multiple levels of detail (Kolbe et al., 2021). Meanwhile, IndoorGML provides a complementary topological representation through its Navigation Route Graph (NRG) for indoor wayfinding. Several studies have addressed the conversion between IFC (the dominant BIM standard) and CityGML to enable cross-domain interoperability (Donkers et al., 2016; Geiger et al., 2015), and Cesium-based platforms have been used to host and visualize these integrated models as 3D Tiles for web-based access (Xu et al., 2020).

However, most existing GeoBIM integration efforts focus on geometric alignment and data format conversion, while the topological continuity across the indoor-outdoor boundary remains insufficiently addressed. Specifically, few studies explicitly model the door or gate entity as a topological transition node that simultaneously participates in both the indoor spatial graph and the outdoor road network. Zhao et al. (2025) proposed a topological modeling approach for integrated indoor-outdoor representation, which our work builds upon by extending it with dynamic ventilation entity states and environmental field coupling.

2.2 Environmental Field Integration in Building Digital Twins

Building digital twins have increasingly incorporated environmental data to support operations beyond pure geometric visualization. Two primary approaches exist in the literature: simulation-driven and sensor-driven.

The simulation-driven approach uses Computational Fluid Dynamics (CFD) to model airflow, temperature distribution, and pollutant transport within the DT environment. While physically rigorous, CFD simulations are computationally expensive, sensitive to boundary condition assumptions, and difficult to update in real time, making them more suitable for design-stage analysis than for operational monitoring (Jiang et al., 2023).

The sensor-driven approach overlays real-time IoT sensor readings onto the BIM or DT model as point-based data layers. Fixed sensor networks have been deployed in buildings to continuously monitor temperature, humidity, CO2, and PM2.5 at discrete locations (Desogus et al., 2021; Opoku et al., 2024). These readings are then associated with the nearest BIM space or visualized as color-coded markers in the DT platform. While practical and real-time, this approach suffers from low spatial resolution—a small number of fixed sensors cannot capture the fine-grained spatial variation of environmental fields across dozens of rooms—and more critically, the sensor data is treated as an independent overlay rather than being causally linked to the geometric configuration of the building.

A key gap in both approaches is the absence of an explicit mechanism that connects the geometric state of building entities — particularly ventilation structures such as windows, doors, and HVAC outlets—to the observed environmental field distribution. In reality, whether a window is open or closed fundamentally alters the air exchange rate and thus the indoor pollutant concentration in the adjacent room. Yet in existing DT frameworks, ventilation facilities are modeled as static geometric objects without mutable operational state attributes. Our work addresses this gap by modeling ventilation entities as first-class semantic objects with detectable states (Full Closed, Half Open, Full Open) and using these states as the explanatory link between building geometry and environmental physics.

2.3 Mobile Sensing and Spatial Interpolation for Indoor Environments

Characterizing indoor environmental fields requires spatially distributed measurements. Fixed sensor networks provide temporal continuity but limited spatial coverage, as deploying a sensor in every room of a large building is often impractical. Mobile sensing — where a human operator or robot carries sensors along a traversal path through the building — offers an alternative that trades temporal continuity for spatial coverage (Hu and Assaad, 2024; Zhang et al., 2025).

Mobile sensing platforms typically integrate environmental sensors with Simultaneous Localization and Mapping (SLAM) devices, generating geo-tagged time series of environmental readings along the movement trajectory. The key challenge is then to extend these sparse, trajectory-bound observations into continuous spatial fields that characterize the entire building. Classical spatial interpolation methods such as Inverse Distance Weighting (IDW) and Kriging have been widely used for this purpose in outdoor environmental mapping. However, their direct application to indoor environments is problematic: these methods assume spatial continuity, meaning that two points that are close in Euclidean distance are assumed to have similar environmental values. Indoors, this assumption fails at every wall boundary—two points on opposite sides of a wall may be only one meter apart in Euclidean distance but exist in entirely different air volumes with potentially very different pollutant concentrations. Applying unconstrained IDW or Kriging to indoor data produces "through-wall" interpolation artifacts where values from one room inappropriately influence the estimate for an adjacent room.

Some studies have explored barrier-constrained interpolation for indoor temperature or air quality mapping, but these typically require manually defining the barrier geometry as an input parameter separate from the building model. In our framework, the barrier information is inherent in the LOD4 semantic model itself: walls are geometric entities with known positions, and the open/closed state of windows and doors — detected in the Geometric-Semantic Layer — determines whether a boundary acts as a barrier or a permeable interface. This tight coupling between the geometric model and the interpolation method is a distinguishing feature of our approach: the DT is not merely a visualization platform onto which environmental data is overlaid, but an active participant in the environmental field computation.

3. Methodology

We propose a hierarchical indoor-outdoor integrated digital twin framework implemented on a Cesium-based geospatial platform. The methodology consists of a three-layer coupled architecture (Section 3.1), LOD4 semantic modeling with topological stitching and ventilation state detection (Section 3.2), semantic-constrained environmental field mapping (Section 3.3), and crowd exposure index computation (Section 3.4).

3.1 Overall Framework

The framework is formulated as a three-layer coupled architecture (Figure 1). The Geometric-Semantic Layer hosts the LOD4 semantic model with topologically connected indoor-outdoor spaces and ventilation entities carrying mutable state attributes. The Physical Field Layer transforms mobile sensing data into entity-level environmental values using semantic-constrained mapping, where walls and closed windows act as aggregation barriers. The Behavioral Response Layer combines entity-level pollution values with pedestrian counts to compute cumulative exposure indices.

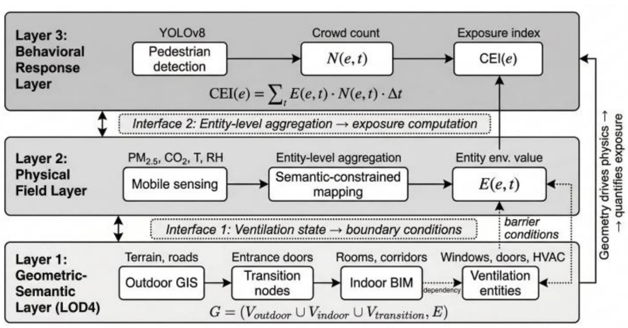

Fig. 1. Three-layer coupled architecture showing the Geometric-Semantic Layer, Physical Field Layer, and Behavioral Response Layer, with two inter-layer coupling interfaces.

The layers are coupled through two interfaces: (1) ventilation entity states from the Geometric-Semantic Layer define the boundary conditions for the Physical Field Layer; (2) entity-level environmental values from the Physical Field Layer are combined with occupancy data in the Behavioral Response Layer to yield exposure assessments.

3.2 LOD4 Semantic Modeling, Topological Stitching, and State Detection

3.2.1 Indoor LOD4 Model Construction. The indoor environment is captured using a handheld SLAM device integrated with the portable sensing kit. The raw point cloud undergoes statistical outlier removal and optional down-sampling to ~2,500 points/m². For multi-story buildings scanned in a single traversal, horizontal cross-sections along the Z-axis separate the point cloud into floor slices.

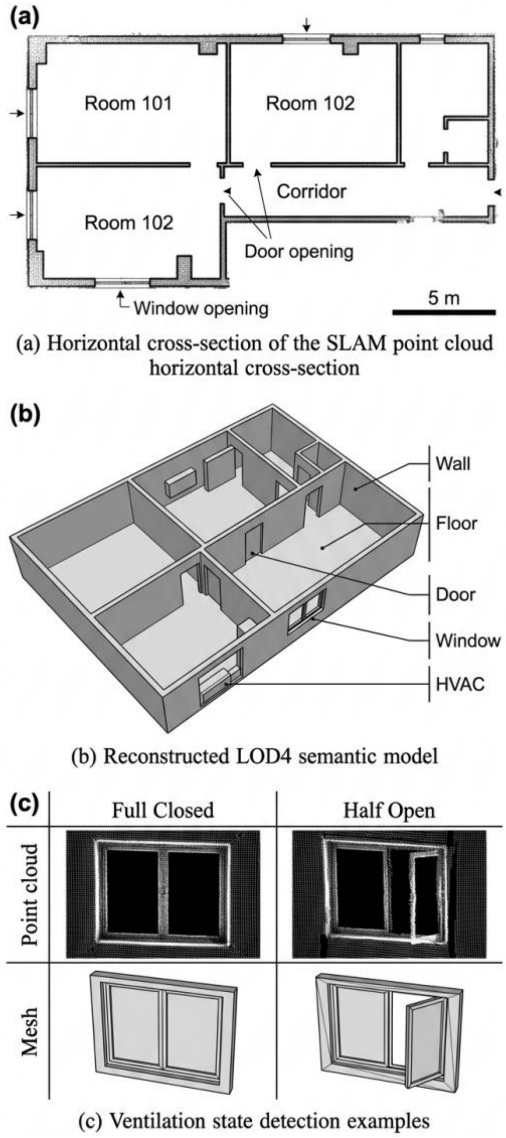

Fig. 2. LOD4 semantic model construction and ventilation state detection. (a) Horizontal cross-section of the SLAM point cloud at 1.2 m height, with room boundaries, door and window openings annotated; (b) Reconstructed LOD4 mesh model with components labeled by type (ceiling removed for visibility); (c) Window state detection: point cloud appearance (top) and mesh result (bottom) for Full Closed and Half Open states.

The LOD4 model is constructed through manual interpretation in CloudCompare. Horizontal cross-sections at multiple heights (0.5 m, 1.2 m, and 2.2 m above the floor, thickness ≈ 0.1 m) produce clear floor-plan views: walls appear as dense point lines, doors as full-height gaps, and windows as gaps visible only at mid-height (Figure 2a). Component dimensions are measured using point-picking and recorded on annotated screenshots. Building components—walls, floors, ceilings, door frames, window frames—are segmented using manual polygon selection.

Each segmented component is converted to a mesh surface via Delaunay 2.5D triangulation on the best-fitting plane, with a maximum edge length constraint (0.3 m) to preserve door and window openings as natural holes. All component meshes are merged into a unified model with semantic attributes (entity type, room ID, floor level, ventilation state) stored as scalar fields (Figure 2b).

3.2.2 Outdoor Scene and Topological Stitching. The outdoor environment is represented on Cesium using global terrain and satellite imagery. The pedestrian transit network is digitized as a node-edge graph from OpenStreetMap or manual GIS digitization.

We define the unified topological graph as:

$$G = (V, E)$$

where:

$$V = V_{\text{outdoor}} \cup V_{\text{indoor}} \cup V_{\text{transition}}$$
$$E = E_{\text{outdoor}} \cup E_{\text{indoor}} \cup E_{\text{transition}}$$

V_{outdoor} consists of pedestrian path nodes; V_{indoor} consists of space nodes (Rooms, Corridors, Staircases); and $V_{\text{transition}}$ consists of building entrance Door entities. The stitching procedure identifies, for each entrance door d, the nearest outdoor node p^* within a distance threshold δ:

$$p^* = \arg\min_{p \in V_{\text{outdoor}}} \text{dist}(d, p), \quad \text{subject to} \quad \text{dist}(d, p^*) \leq \delta$$

where $\text{dist}(\cdot, \cdot)$ is the Haversine distance and $\delta = 30$ m. A transition edge (p^*, d) is then added to $E_{\text{transition}}$. Connectivity is verified by testing reachability from all outdoor nodes to all indoor rooms.

3.2.3 Ventilation Facility State Detection. Ventilation facilities carry mutable operational state attributes that serve as the bridge between building geometry and environmental physics. Window states are classified into three categories through visual interpretation of the SLAM point cloud (Figure 2c):

- Full Closed: dense reflected points across the window opening; no outdoor points visible.
- Half Open: partial coverage, with glass reflections on one side and outdoor points visible on the other; or sash displacement of 5 – 25 cm from the wall plane.
- Full Open: no sash point cloud; distant outdoor points clearly visible through the opening.

Classification is performed by a human operator and cross-validated against photographs taken during scanning. HVAC states (active/inactive) are recorded from field observations. All states are stored as scalar field attributes in the mesh model and serve as (1) explanatory variables for indoor air quality analysis (Section 4) and (2) barrier conditions for environmental mapping (Section 3.3).

3.3 Semantic-Constrained Environmental Field Mapping

3.3.1 Sensing Platform and Sampling Protocol. A portable all-in-one sensing kit (135 mm × 262 mm × 130 mm; 5 kg) integrates 10 sensor types including chemical sensors (CO_2, PM2.5), temperature/humidity sensors, and cameras for pedestrian detection. All readings are time-synchronized and geo-tagged via the SLAM trajectory.

Data collection was conducted in a university teaching building during autumn. Sampling was performed for one week, from 06:00 to 22:00 at 2-hour intervals (9 rounds/day, 63 rounds total). Each round, the operator traversed outdoor paths, building entrances, corridors, and target rooms on the 1st and 2nd floors.

3.3.2 Semantic-Constrained Mapping Method. Rather than applying spatial interpolation across the building, we assign each sensor reading to the semantic entity in which it was collected and aggregate within each entity. The LOD4 model boundaries act as natural barriers preventing cross-room data mixing.

Step 1: Point-to-Entity Assignment. For each reading s_i with coordinates (x_i, y_i, z_i) at time t, the entity membership is determined by testing whether (x_i, y_i) falls within the 2D boundary polygon of each room or corridor defined in the LOD4 model. For outdoor readings, the point is assigned to the nearest road segment. Each reading receives a label $\text{entity}(s_i) = e$.

Step 2: Entity-Level Aggregation. For entity e and sampling round t:

$$E(e,t) = \frac{1}{n} \sum_{\substack{i:\,\text{entity}(s_i)=e \\ \text{time}(s_i)=t}} c_i$$

where c_i is the pollutant concentration of reading i and n is the count of readings assigned to entity e during round t.

This approach treats each enclosed room as a well-mixed air volume — a reasonable assumption at room scale (<100 m^2) where convective mixing homogenizes the air within minutes. Unlike conventional IDW or Kriging, it produces no "through-wall" artifacts because readings from one room never influence the estimate for an adjacent room. The open/closed state of windows is recorded as a covariate for the correlation analysis in Section 4.

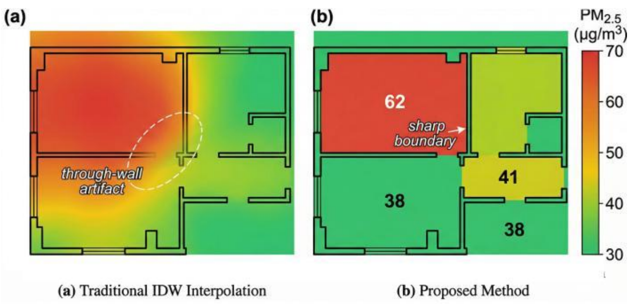

Fig. 3. Comparison of semantic-constrained mapping versus conventional IDW interpolation. (a) Traditional IDW interpolation of PM2.5 showing through-wall artifacts at room boundaries; (b) Proposed semantic-constrained mapping showing clear room-level isolation with sharp concentration boundaries at walls.

3.4 Crowd Exposure Index

3.4.1 Pedestrian Flow Detection. Onboard cameras capture images along the traversal trajectory. A pre-trained YOLOv8 model detects and counts visible persons in each frame. Counts are assigned to entities via the same point-to-entity procedure. For each entity e and round t, the occupancy $N(e,t)$ is taken as the maximum count across all frames within that entity during that round, to compensate for the limited camera field of view.

3.4.2 Cumulative Crowd Exposure Index. The Crowd Exposure Index is defined as:

$$\mathrm{CEI}(e) = \sum_{t} E(e,t) \cdot N(e,t) \cdot \Delta t$$

where $E(e,t)$ is the entity-level pollutant concentration, $N(e,t)$ is the occupancy count, and $\Delta t = 2h$ h is the sampling interval. The summation spans all rounds in the observation period.

CEI captures the joint effect of pollution concentration, population density, and temporal duration. The "double jeopardy" scenario arises when both E and N are simultaneously high—high concentration in an enclosed room with many stationary occupants over extended periods. CEI values are visualized in Cesium through entity-level color coding (green → yellow → red), providing facility managers with an intuitive map of cumulative exposure hotspots.

4. Experiment and Results

4.1 Study Area and Data Overview

The experiment was conducted in a teaching building of a university campus in Beijing, China. The building has two above-ground floors with a total floor area of approximately 2,400 m². The 1st floor contains a high-density study room (Room 101, ~85 m², capacity ~120 seats), classrooms, and a central corridor connecting to two entrances. The 2nd floor contains classrooms (including Room 201, ~80 m²) and a parallel corridor connected to the 1st floor via two staircases. The building is naturally ventilated through operable windows; wall-mounted split air conditioning units are installed but were not actively used during the autumn sampling period.

The SLAM scanning campaign produced a unified point cloud of 38.7 million points covering both floors. After preprocessing (SOR denoising, down-sampling to 2 cm spacing), 31.2 million points were retained. The multi-modal sensing kit collected environmental data over 7 consecutive days in October, from 06:00 to 22:00 at 2-hour intervals, yielding 63 sampling rounds. Each round traversed outdoor transit paths (~200 m), both building entrances, 1st and 2nd floor corridors, the 1st-floor study room, and two 2nd-floor classrooms.

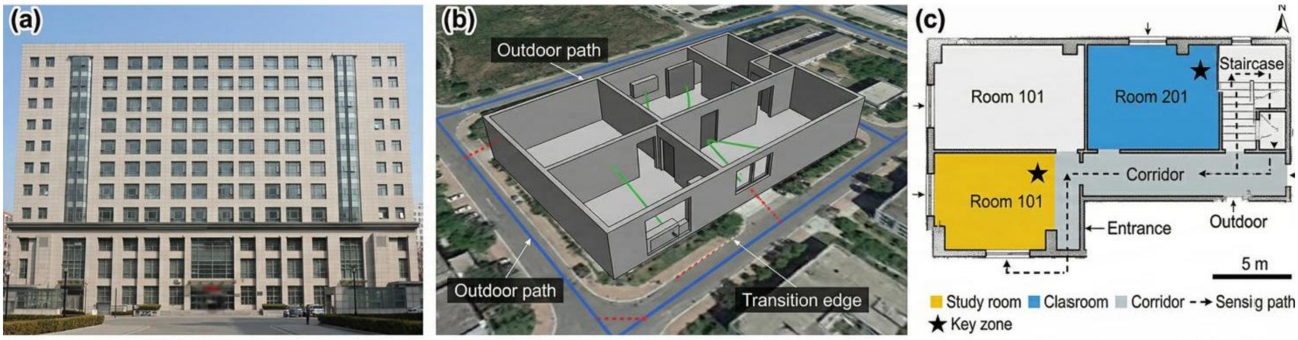

Fig. 4. Experimental site and digital twin model. (a) Exterior photograph of the teaching building; (b) Cesium platform screenshot showing the integrated LOD4 indoor model with outdoor road network and topological edges (blue: outdoor paths, green: indoor connections, red dashed: transition edges at entrances); (c) Annotated floor plan derived from the SLAM point cloud horizontal cross-section, with color-coded functional zones (yellow: 1F study room, blue: 2F classrooms, gray: corridors) and the sensing kit traversal path (arrows).

4.2 LOD4 Model Construction Quality

The LOD4 semantic model comprises 47 mesh components: 28 wall segments, 4 floor surfaces, 4 ceiling surfaces, 6 doors (including 2 building entrances), 4 windows, and 1 HVAC unit. The unified topological graph contains 32 nodes (6 indoor space nodes, 2 transition nodes, 24 outdoor path nodes) and 31 edges (5 indoor, 1 vertical, 2 transition, 23 outdoor). Connectivity verification confirmed that all 4 indoor rooms are reachable from every outdoor path node.

Model quality was evaluated across four metrics, as summarized in Table 1:

Table 1. Quality assessment of the LOD4 model and ventilation state detection.

Metric	Value
Geometric accuracy (Cloud-to-Mesh distance)	Mean: 3.2 cm, SD: 1.8 cm
Semantic annotation accuracy (spot check, 20% sample)	95.0% (19/20 components correct)
Topological connectivity rate	100% (all rooms reachable from outdoor)
Window state detection accuracy	92.3% (12/13 windows correctly classified)
Inter-annotator agreement (Cohen's Kappa)	0.85

The mean cloud-to-mesh distance of 3.2 cm indicates that the manually constructed mesh closely approximates the original point cloud geometry. One window state was misclassified as Full Closed when it was actually Half Open (the narrow opening of ~4 cm fell below the visual detection threshold in the point cloud), yielding a 92.3% detection accuracy. The inter-annotator Kappa of 0.85 indicates strong agreement between two independent operators.

4.3 Geometry-Physics Coupling Verification

This section presents the core experimental finding: the geometric state of ventilation facilities explains the observed indoor-outdoor environmental disparity.

4.3.1 Indoor vs. Outdoor Environmental Profiles. Entity-level environmental values computed via the semantic-constrained mapping method (Section 3.3) reveal a stark indoor-outdoor contrast. During the one-week sampling period, the average indoor PM2.5 concentration across all indoor entities was 58.4 μg/m³, compared to an outdoor average of 38.2 μg/m³ along transit paths—an indoor/outdoor (I/O) ratio of 1.53. The indoor average exceeds the WHO Global Air Quality Guideline for 24-hour PM2.5 exposure (15 μg/m³) (WHO, 2021) by a factor of 3.9.

Indoor CO_2 exhibited an even more pronounced disparity. The 1st-floor study room (Room 101) recorded peak CO_2 levels of 1,800 ppm during afternoon hours (14:00 – 16:00), far exceeding both the typical outdoor baseline (~420 ppm) and the commonly referenced indoor hygienic threshold of 1,000 ppm. The 2nd-floor classrooms showed peak values of ~1,200 ppm during occupied lecture periods, declining to ~500 ppm during unoccupied intervals.

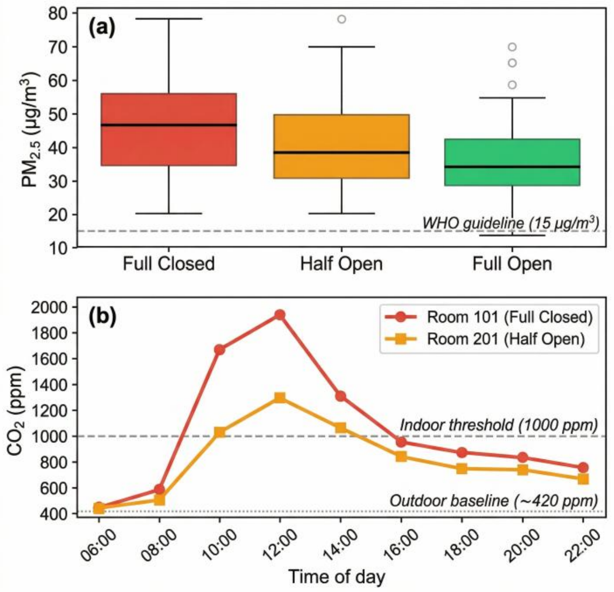

Fig. 5. Correlation between ventilation facility state and indoor air quality. (a) Box plots of entity-level PM2.5 concentration grouped by window state (Full Closed, n=38; Half Open, n=16; Full Open, n=9), with the WHO 24-hour guideline (15 μg/m³) shown as a dashed line; (b) Diurnal CO2 time series for Room 101 (all windows Full Closed) and Room 201 (windows Half Open) on a representative weekday, with indoor threshold (1000 ppm) and outdoor baseline (~420 ppm) indicated.

4.3.2 Ventilation State Attribution. To verify that the window state causally explains the indoor air quality disparity, we compared entity-round pairs grouped by window state. As shown in Figure 5(a), PM2.5 concentrations in Full Closed entities (median: 61.2 μg/m³) were significantly higher than in Half Open entities (median: 42.8 μg/m³) and Full Open entities (median: 35.1 μg/m³). This monotonic decrease confirms that window opening status—a geometric attribute detected from the SLAM point cloud in Section 3.2.3 — is a primary driver of indoor pollutant accumulation.

The CO2 time-series comparison (Figure 5(b)) provides further evidence. Room 101, where all windows remained Full Closed throughout the observation period, exhibited sustained CO2 levels above 1,000 ppm during occupied hours, with minimal diurnal recovery. In contrast, Room 201, where windows were Half Open, showed CO2 peaks during lectures followed by rapid decay toward outdoor baseline levels during breaks — consistent with effective air exchange through the partially open windows.

These results validate the core premise of the Geometric-Semantic Layer: ventilation entity states, as geometric attributes detectable from SLAM point clouds, provide a direct physical explanation for the spatial distribution of indoor environmental fields. The "Full Closed" configuration effectively seals the room from outdoor air exchange, causing both PM2.5 (from infiltrated outdoor particles with no dilution pathway) and CO2 (from occupant respiration with no exhaust pathway) to accumulate.

4.3.3 Semantic-Constrained Mapping Effectiveness. To demonstrate the advantage of the proposed mapping method over conventional interpolation, we applied both approaches to the same set of mobile sensing readings collected during one sampling round. Figure 3 (in Section 3.3) illustrates the spatial distribution of PM2.5 produced by each method.

Using traditional IDW interpolation (ignoring walls), the PM2.5 field exhibits smooth gradients across wall boundaries: Room 101 (measured: 62 μg/m³) and the adjacent corridor (measured: 41 μg/m³) show a gradual transition zone that extends ~3 m into each space, producing a physically meaningless intermediate value of ~52 μg/m³ at the wall location. This "through-wall" artifact arises because IDW treats the wall as transparent and weights readings by Euclidean distance regardless of physical barriers.

Table 2. Summary of environmental parameters and cumulative crowd exposure index across representative entities. CEI is computed over the full 7-day observation period (63 sampling rounds).

Entity	Type	Window State	Mean PM2.5 (μg/m³)	Mean CO2 (ppm)	Mean Occupancy	CEI (μg · person · h /m³)
Room 101 (1F study room)	Room	Full Closed	61.2	1,340	68	58,147
Room 201 (2F classroom)	Room	Half Open	42.8	820	35	12,474
Corridor 1F	Corridor	N/A	44.5	610	12	4,704
Outdoor transit path	Outdoor	N/A	38.2	420	8	4,843

https://doi.org/10.5194/isprs-archives-XLIX-B4-2026-555-2026

Using the proposed semantic-constrained mapping, Room 101 receives a value of 62 μg/m³ and the corridor receives 41 μg/m³, with a sharp discontinuity at the wall boundary. This result faithfully reflects the physical reality: the closed door between Room 101 and the corridor prevents air mixing, and the two spaces maintain distinct pollutant concentrations.

4.4 Crowd Exposure Assessment

The Crowd Exposure Index (Equation 1) was computed for all monitored entities over the full observation period. The results, summarized in Table 2 and visualized in Figure 6(a), reveal dramatic disparities.

The 1st-floor study room (Room 101) recorded a CEI of 58,147 μg・person・h/m³—approximately 12 times higher than the outdoor transit path (4,843) and 4.7 times higher than the 2nd-floor classroom (12,474). This "double jeopardy" arises from the simultaneous co-occurrence of two adverse factors: Room 101 had both the highest pollutant concentrations (driven by the Full Closed window state) and the highest sustained occupancy (students engaged in all-day self-study, with mean occupancy of 68 persons across sampling rounds).

By contrast, although the outdoor transit path had non-trivial PM2.5 levels (38.2 μg/m³), the low and transient pedestrian density (mean ~8 persons) and short dwell times resulted in a low cumulative CEI. The 2nd-floor classroom occupied an intermediate position: its Half Open windows moderated PM2.5 levels relative to Room 101, and its occupancy was periodic (lecture hours only) rather than continuous. Figure 6(b) further reveals the diurnal pattern of high-exposure episodes, concentrated during daytime occupied hours on weekdays in the enclosed indoor entities.

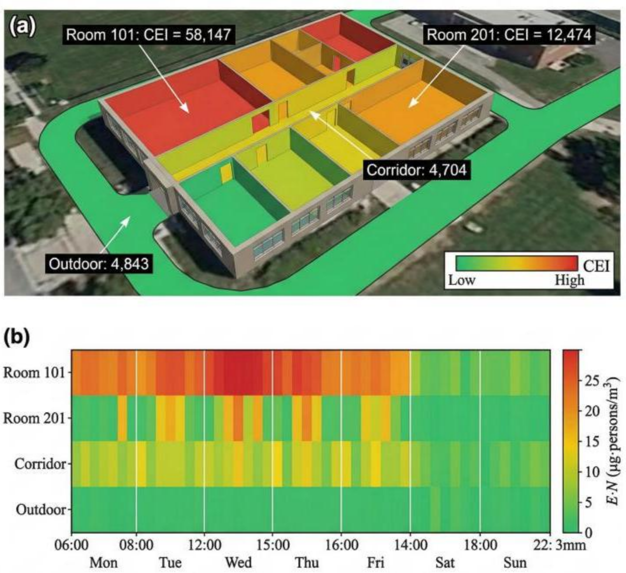

Fig. 6. Crowd Exposure Index visualization. (a) Cesium platform screenshot with entity-level CEI color coding at a representative time step (Wednesday 14:00): Room 101 in red (highest exposure), corridors in yellow, outdoor paths in green (lowest); (b) Spatiotemporal heatmap of instantaneous exposure E(e,t)・N(e,t) across all monitored entities (vertical axis) and 63 sampling rounds over one week (horizontal axis). White vertical lines separate days. Color scale follows the AQI convention: green (low) to red (high).

4.5 Discussion

The experimental results support three key findings.

First, the geometry-physics coupling hypothesis is validated: the geometric state of ventilation facilities (windows fully closed) is the primary explanatory factor for elevated indoor pollutant concentrations. This finding demonstrates the value of modeling ventilation entities as first-class semantic objects with mutable states in the digital twin, rather than treating them as static geometric decorations.

Second, the indoor-outdoor integrated modeling enables continuous exposure tracking across the building boundary. The unified topological graph allows direct comparison of CEI between outdoor transit areas and specific indoor rooms within a single framework, revealing that crossing a building entrance with closed windows increases cumulative exposure risk by an order of magnitude. This cross-boundary perspective would not be possible with separate indoor and outdoor monitoring systems.

Third, the semantic-constrained mapping method produces physically meaningful entity-level environmental values that correctly reflect room-level compartmentalization, avoiding the through-wall artifacts of conventional spatial interpolation. While the method is conceptually simple (entity-level averaging rather than geostatistical interpolation), its effectiveness depends entirely on the availability of the LOD4 semantic model with accurate room boundaries—demonstrating that the Geometric-Semantic Layer is not merely a visualization tool but an active participant in the environmental analysis pipeline.

Several limitations should be noted. The experiment covers a single building during one season (autumn) with a 2-hour sampling interval, limiting the temporal resolution and seasonal generalizability. The window state detection relies on human interpretation of the point cloud, which may miss subtle opening states below ~5 cm. The occupancy estimates from onboard cameras capture only visible persons within the field of view, likely underestimating true room occupancy. Future work will address these limitations through deployment of IoT-based window sensors for real-time state monitoring, extended multi-season campaigns, and integration with building access control systems for more accurate occupancy data.

5. Conclusion

This paper presented an environment-aware indoor-outdoor integrated digital twin framework that couples geometric entity states with physical environmental fields and occupant behavioral responses for healthy mobility assessment. The framework was implemented on a Cesium-based platform and validated through a week-long experiment in a university teaching building. The main conclusions are as follows:

(1) A three-layer coupled architecture was proposed in which the Geometric-Semantic Layer, Physical Field Layer, and Behavioral Response Layer are linked through two explicit coupling interfaces. The ventilation entity state serves as the bridge between building geometry and environmental physics—a mechanism absent from existing building digital twin frameworks that treat environmental sensor data as an independent overlay on static 3D models.

(2) A seamless LOD4 semantic model was constructed from SLAM point clouds, in which indoor spaces and outdoor transit

networks are topologically stitched through transition nodes at building entrances, forming a unified graph $G = (V_{\text{outdoor}} \cup V_{\text{indoor}} \cup V_{\text{transition}},; E)$. Ventilation facilities (windows, doors, HVAC) are modeled as first-class entities with mutable state attributes (Full Closed / Half Open / Full Open), achieving a state detection accuracy of 92.3% through point cloud interpretation validated against field photographs.

(3) A semantic-constrained environmental mapping method was developed that assigns mobile sensing readings to LOD4 entities and aggregates within each entity, using walls and closed windows as natural barriers. This approach eliminates the through-wall interpolation artifacts inherent in conventional IDW or Kriging methods and produces physically meaningful room-level environmental values without requiring complex geostatistical modeling.

(4) The experimental results revealed a "double jeopardy" in the indoor environment: PM2.5 concentrations in the fully enclosed 1st-floor study room averaged 61.2 μg/m³—1.6 times the outdoor level and 4.1 times the WHO 24-hour guideline—while the cumulative Crowd Exposure Index in this room was 12 times higher than in outdoor transit areas. Semantic correlation confirmed that the Full Closed window state was the primary driver of pollutant accumulation, validating the geometry-physics coupling at the core of the framework.

These findings challenge the conventional assumption of indoor spaces as safe havens and demonstrate that the geometric configuration of ventilation structures—an attribute readily detectable from SLAM point clouds—directly governs indoor environmental quality. The proposed framework enables facility managers to identify high-exposure spatiotemporal hotspots and trigger targeted interventions such as scheduled ventilation or occupancy redistribution.

Future work will extend the framework in three directions: (1) deploying IoT-based window and door sensors for real-time ventilation state monitoring, removing the dependence on periodic SLAM scanning; (2) conducting multi-building, multi-season validation campaigns to assess the generalizability of the geometry-physics coupling relationship under varying climatic conditions; and (3) integrating CFD simulation into the Physical Field Layer to enable predictive environmental modeling—transitioning the digital twin from a monitoring tool to a predictive decision-support platform.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (Grant No. 42371453).

References

Desogus, G., Quaquero, E., Rubiu, G., Gatto, G., & Perra, C. (2021). BIM and IoT sensors integration: A framework for consumption and indoor conditions data monitoring of existing buildings. Sustainability, 13(8), 4496. https://doi.org/10.3390/su13084496

Ding, Y., Zhang, Y., & Huang, X. (2023). Intelligent emergency digital twin system for monitoring building fire evacuation. Journal of Building Engineering, 77, 107416.

Donkers, S., Ledoux, H., Zhao, J., & Stoter, J. (2016). Automatic conversion of IFC datasets to geometrically and semantically correct CityGML LOD3 buildings. Transactions in GIS, 20(4), 547 – 569.

EPA (2023). Indoor air quality. Report on the Environment. U.S. Environmental Protection Agency. https://www.epa.gov/report-environment/indoor-air-quality

Geiger, A., Benner, J., & Häfele, K.-H. (2015). Generalization of 3D IFC building models. In 3D Geoinformation Science (pp. 19 – 35). Springer.

Gröger, G., Kolbe, T. H., Nagel, C., & Häfele, K.-H. (2012). OGC City Geography Markup Language (CityGML) Encoding Standard, Version 2.0. OGC Doc. No. 12-019, Open Geospatial Consortium.

Hu, W., & Assaad, R. H. (2024). A BIM-enabled digital twin framework for real-time indoor environment monitoring and visualization by integrating autonomous robotics, LiDAR-based 3D mobile mapping, IoT sensing, and indoor positioning technologies. Journal of Building Engineering, 86, 108901. https://doi.org/10.1016/j.jobe.2024.108901

Jiang, L., Shi, J., Wang, C., & Pan, Z. (2023). Intelligent control of building fire protection system using digital twins and semantic web technologies. Automation in Construction, 147, 104728.

Kolbe, T. H., Kutzner, T., Smyth, C. S., Nagel, C., Roensdorf, C., & Heazel, C. (2021). OGC City Geography Markup Language (CityGML) Part 1: Conceptual Model Standard, Version 3.0. OGC Doc. No. 20-010, Open Geospatial Consortium.

Ma, Z., & Ren, Y. (2017). Integrated application of BIM and GIS: An overview. Procedia Engineering, 196, 1072 – 1079.

Opoku, D.-G. J., Perera, S., Osei-Kyei, R., Rashidi, M., Bamdad, K., & Famakinwa, T. (2024). Digital twin for indoor condition monitoring in living laboratories: University library case study. Automation in Construction, 157, 105188.

Teo, T.-A., & Cho, K.-H. (2016). BIM-oriented indoor network model for indoor and outdoor combined route planning. Advanced Engineering Informatics, 30(3), 268 – 282.

WHO (2021). WHO Global Air Quality Guidelines: Particulate Matter (PM2.5 and PM10), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide. World Health Organization.

Xu, Z., Zhang, L., Li, H., Lin, Y.-H., & Yin, S. (2020). Combining IFC and 3D tiles to create 3D visualization for building information modeling. Automation in Construction, 109, 102995.

Zhang, Y., Yu, D., Li, D., Li, Y., & Zhang, J. (2025). Three in Motion: A Mobile Study on the Interlinked Dynamics of CO_2, Air Temperature, and PM2.5. Journal of Cleaner Production, 506, 145449.

Zhao, W., Kang, Z., & Ye, C. (2025). Integrated Indoor and Outdoor Topological Modeling. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 48, 1769 – 1775.