Automatic 3D Building Model Generation for Energy Digital Twins

Oscar Roman[1,2], Giorgio Agugiaro[3], Ken Arroyo Ohori[3], Maarten Bassier[4], Elisa Mariarosaria Farella[1], Fabio Remondino[1]

[1] 3DOM 3D Optical Metrology, Bruno Kessler Foundation, Trento, Italy - oroman@fbk.eu, elifarella@fbk.eu, remondino@fbk.eu
[2] University of Trento, EICS and DII Department, Trento, Italy - oscar.roman@unitn.it
[3] 3D Geoinformation group, Department of Urbanism, Faculty of Architecture and the Built Environment, Delft University of Technology, Delft, The Netherlands - g.agugiaro@tudelft.nl, k.ohori@tudelft.nl
[4] Department of Civil Engineering, TC Construction - Geomatics, KU Leuven – Faculty of Engineering Technology, Ghent, Belgium maarten.bassier@kuleuven.be

Keywords: Automation in constructions, BEM, Deep Learning, Energy simulations, Scan-to-BIM.

Abstract

Digital Twins in the Architecture, Engineering, and Construction (AEC) domain support monitoring, simulation, and increasing levels of automation in building management across scales. Energy Digital Twins are particularly demanding, requiring (i) simulation-grade geometry and (ii) persistent topology and semantics across monitoring- and scenario-driven updates. This paper proposes a unified multi-representation EDT in which (i) a watertight, solid, and (ii) a topology-preserving B-Rep are co-maintained through a mapping layer that preserves object identity and links geometry to a typed property graph. Building on this, the presented Scan-to-Energy Digital Twin pipeline converts raw point clouds into multi-level EDT instances by integrating Scan-to-BIM reconstruction, topological modelling, semantic enrichment and parser–transformer–writer interoperability modules. The graph-backed EDT enables reversible export to epJSON and gbXML (optionally IFC), supporting scenario-based EnergyPlus simulations and incremental retrofit updates, such as insulation thickness and window thermal transmittance value changes. Validation on a set of four buildings achieves 0.86–0.89 mAPv and schema-valid exports, demonstrating the effectiveness of our end-to-end approach for interoperable energy analysis, monitoring, and operational decision support.

1. Introduction

Recent literature (Klar et al., 2024) shows that building Digital Twins (DTs) require robust frameworks linking dynamic, physical and digital assets through geometric consistency and structured data association. In the energy domain, an Energy Digital Twin (EDT) must support (i) *detailed, simulation-grade geometry*, (ii) *continuous performance monitoring*, and (iii) *scenario-based simulations* for decision support. This calls for complementary geometric representations and a data structure that preserves object identity, topology and semantics across state changes, ideally grounded in reality-capture point clouds (Hu and Cai, 2024). Workflows typically combine point-cloud segmentation and classification (Lu et al., 2025; Lee et al., 2025), 3D element reconstruction (Schlenger et al., 2025; Chen et al., 2025), monitoring and visualization (Wang et al., 2025) and semantic enrichment (Büyükdemircioğlu and Oude Elberink, 2025).

Despite these advances, automatic pipelines still rarely produce DT-ready, watertight volumetric models for physics-based simulation (Agugiaro et al., 2018). Reconstructions often exhibit topological inconsistencies that undermine adjacency reasoning and zoning (Horna et al., 2009), yield biased volumetric estimates (Biljecki et al., 2016) and remain difficult to exchange via strict formats such as IFC or gbXML due to valid geometry and topology requirements (Yang et al., 2022).

This work addresses these gaps with an end-to-end framework that unifies (i) geometric and (ii) semantic structures into an interoperable, simulation-ready DT.

1.1 Aim of the work

This paper presents the Scan-to-Energy Digital Twin (Scan-to-EDT) framework, an end-to-end pipeline that transforms raw point clouds into EDTs for monitoring and energy simulation.

The contributions from this article are threefold:

- *Present the Scan-to-EDTs framework,* demonstrating the high-fidelity geometric and topological correctness at a high level of detail (including rooms and openings) and assessing end-to-end robustness across diverse datasets.
- *Unify data via a knowledge graph layer,* constructing a building-scale graph from the Scan-to-EDTs framework outputs to standardize semantics and creating gbXML and epJSON results for simulation and data management.
- *Guarantee interoperability,* ensuring gbXML/IFC and Energy Plus (U.S. Department of Energy, 2025) compliance to enable exchange across modelling and simulation platforms.

2. Digital Twins: modelling and data structures

In the built environment, DTs can be characterized along three axes (Wang et al., 2025): (i) geometric level of detail (LOD), (ii) data-integration capacity, and (iii) functional purpose. Complementing this, (Klar et al., 2024) describe DT maturity in terms of integration, interaction and automation, progressing from basic monitoring (DTL1) to simulation-enabled models (DTL2) and, ultimately, interoperable DTs supporting predictive maintenance (DTL3). Achieving such maturity requires a structured workflow that transforms raw point clouds into operational DTs by (1) semantic classification of building elements, (2) reconstruction of watertight, topology-explicit B-Rep geometry, and (3) enrichment with semantics and physical properties to match the target DT level. The following sections review the state of the art across these stages.

2.1 Recent advances in semantic and instance segmentation

A panoptic segmentation approach (Robert et al., 2024), combining semantic labels and instance IDs, underpins robust and topology-aware 3D semantic reconstruction. Modern transformers, such as Point Transformer v3 (PTv3) (Wu et al., 2024) and structured multi-network frameworks (e.g. Mosaic3D,

Lee et al., 2025), embed panoptic reasoning in both architecture and training, achieving high precision on large-scale scenes and producing structured 3D assets. These assets support downstream scene understanding and can be organized as hierarchical polytree graphs with parent–child relationships. For example, the superpoint hierarchy in Superpoint Transformer (Robert et al., 2024) yields analysis-ready 3D structures (Zhou et al., 2025; Steinke et al., 2025). These graph formalisms crucially unify segmentation, towards a strong reconstruction and dynamics across graph cuts, structured reconstruction, physics-based models and dynamic graph prediction.

2.2 Reconstruction techniques for DT models

Although BIM models are largely static, within DTs they often serve as the geometric and semantic base. Long-standing bridging ontologies (El-Mekawy and Östman, 2010) aim to link BIM/IFC to semantic models, such as CityGML/CityJSON, which can support more dynamic representations. Recent topology-aware BIM reconstruction methods using supervised learning (Gan et al., 2025) enrich such models with real-time and historical data. This linkage is further enabled by topology-explicit models (Kolbe et al., 2005), which expose incidence and containment relations, thereby supporting robust integration and reasoning (Agugiaro et al., 2018). These models can target specific LODs from building to city scale, supporting urban energy simulation (Gao et al., 2025) and block-scale benchmarks for downstream applications, such as TUM2TWIN (Wysocki et al., 2025). Standardized data formats are thus essential, particularly those encoding geometric topology (Yang et al., 2022), such as B-Rep in gbXML or solid-based BIM/IFC schemas, to guarantee interoperability and consistent, reliable data exchange. Reconstruction methods range from image-based LoD4 generation (Pantoja Rosero et al., 2024) to data-driven single-building LoD4 semantic modelling (SYNBUILD 3D; Mayer et al., 2025). Given this diversity, standards, particularly for B-Rep topology in formats such as gbXML, are essential to ensure interoperability and reliable exchange (Yang et al., 2022).

2.3 Layers and Types of Digital Twins for Buildings

Geometric DTs form the twins’ base, providing accurate 3D representations for visualization, documentation and asset mapping; furthermore, (i) semantic or informational twins enrich the geometry with materials, systems and performance attributes to support facility management, maintenance planning and building operation, often developed in Virtual Reality (VR) platforms; (ii) analytical or simulation twins, instead, couple the model with measured or synthetic data to enable energy analysis, structural assessment and occupancy studies (Wang et al., 2025); and (iii) operational twins, finally, integrate real-time IoT and sensor data for continuous monitoring (DT Level 1, DTL1), simulation-based forecasting and optimization (DTL2), predictive maintenance (DTL3) and adaptive building control, forming a continuum from static models to dynamic, data-driven systems spanning design and operation (Hu & Cai, 2024).

3. Methodology

The proposed framework (Figure 1) presents an end-to-end solution for constructing hierarchical and multi-level DTs, combining volumetric and surface-based models for energy monitoring and scenario analysis.

The pipeline begins with data acquisition and preprocessing, then advances through (i) Scan-to-BIM (Section 3.1) solid-geometry reconstruction and energy-element extraction via supervised operating-state classification and per-class mesh reconstruction; (ii) topological modelling (Section 3.2); (iii) graph-based network reconstruction (Section 3.3); (iv) semantic enrichment (Section 3.4); (v) internal format conversion to IFC, gbXML and epJSON (Section 3.5); and (vi) scenario-driven simulation (Section 3.6). The results are visualized and served via a customized API. The following sections detail the methodology adopted in this work.

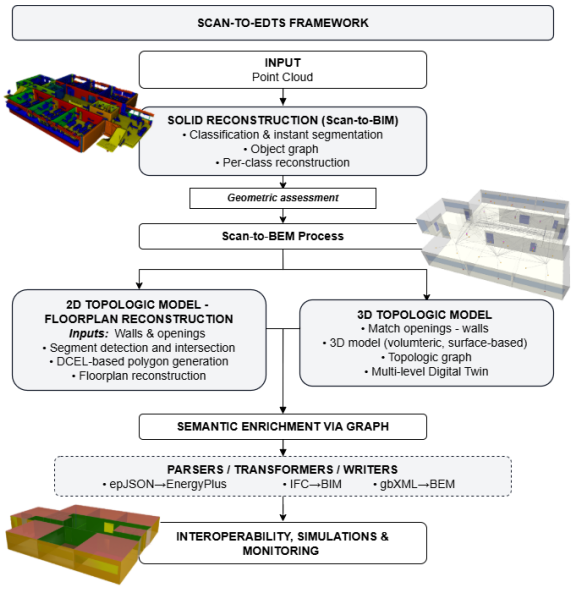

Figure 1. Workflow of the proposed Scan-to-EDTs framework.

3.1 Scan-to-BIM process

In general workflows, automatic BIM creation involves (i) point-cloud classification and instance segmentation, (ii) structural-graph construction linking instances to classified points, and (iii) per-class solid reconstruction of building elements (Figure 2).

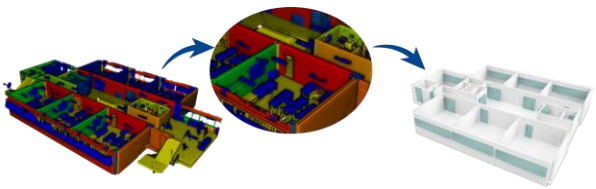

Figure 2. Simplified solid model generated via Scan-to-BIM.

The proposed method builds on Roman et al. (2024b) for Scan-to-BIM mesh generation, while advancing DT multi-level topology reconstruction and strengthening interoperability modules (IFC/gbXML/epJSON) for scenario-driven simulation; classification is adopted off-the-shelf and treated as an upstream dependency. Let $\mathcal{D}$ denote the input point cloud, partitioned into semantic classes as follows (Eq. 1):

$$\mathcal{D} = \mathcal{U} + \mathcal{F} + \mathcal{C} + \mathcal{W} + \mathcal{L} + \mathcal{O} \quad (1)$$

where $\mathcal{F}$, $\mathcal{C}$, $\mathcal{W}$, $\mathcal{L}$, $\mathcal{O}$ and $\mathcal{U}$ denote floors, ceilings, walls, columns, openings (windows and doors) and unclassified elements. Following Roman et al. (2024b), PTv3 (Wu et al., 2024) is used for instance segmentation. In parallel, the Scan-to-BEM step (Roman et al., 2024b) relies on manual visual annotation of energy-relevant elements in $\mathcal{D}$, assigning each a class and persistent *object_id*. From the segmented point cloud, a typed RDF property graph is constructed: nodes (e.g., *Wall, Opening, Mesh, PointSet*) store semantic and geometric

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-1-2026-447-2026 | © Author(s) 2026. CC BY 4.0 License.

attributes, while edges encode relations (e.g., *boundedBy, derivedFrom, inStorey*).
The graph then drives a modular, per-class computational-geometry reconstruction (Roman et al., 2024b), preserving node-level traceability from segments to meshes. The resulting watertight, simplified solid mesh constitutes the Scan-to-BIM output; this mesh-based model (Figure 2) is subsequently processed by the EDTs framework, leveraging $\mathcal{W}$ and $\mathcal{O}$ to generate the multi-level DT.

3.2 Topologic Model

The multi-level DT proposed here, built on a topological model for energy applications, is required to be (i) watertight, (ii) 2-manifold and (iii) hierarchically structured to support monitoring (*DTL1*) and simulation (*DTL2*). It is built in two stages:

- A *2D Topologic model* (implementing the *Topologic Map* in Roman et al., 2024a) that defines floorplan as a set of closed polygons (Figure 3, a1-a2);
- A *3D Topologic model* that reconstructs relevant faces, with internal/external walls (Figure 3, b1-b2) and Boolean-subtracts openings from them for the 3D model (Figure 3, c1-c2).

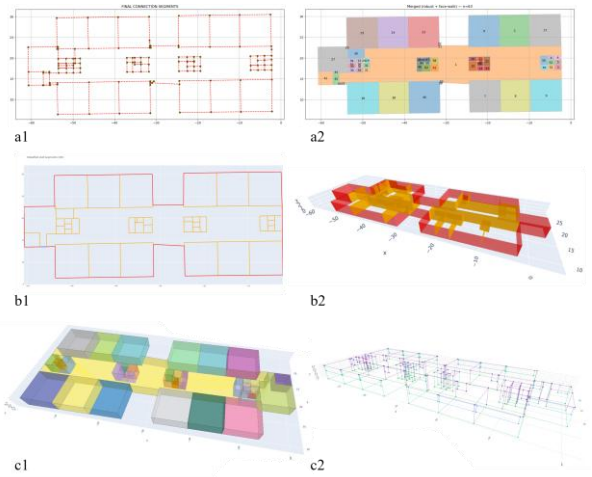

Figure 3. Workflow on the first floor of the School Dataset (Figure 7): (a1) intersections and fine-grained segments; (a2) DCEL floorplan polygons; (b1–b2) internal/external walls; (c1–c2) volumetric and surface-based topologic models.

Topologic models are derived from (i) surfaces aligned with wall centrelines and (ii) axis-aligned bounding boxes representing doors and windows.
After orthogonal projection to the *xy*-plane, wall centrelines are generated by iteratively computing pairwise intersections, splitting segments at intersection points and retaining only the portions coincident with the original wall geometry.
Openings $\mathcal{O}$ are incorporated by clipping the corresponding axis-aligned rectangles and subtracting them from the retained segments, first in the xy-plane and then in 3D, thereby creating gaps for doors and windows. Unique identifiers are assigned to vertices and edges, and a Doubly Connected Edge List (DCEL) is constructed using `half-edges` with origin, `twin`, `next`, and `previous` references. Consistently oriented `half-edge` cycles define planar polygons corresponding to room faces, which are then lifted to 3D for shell and cell construction. Edges bounding polygons on both sides are classified as internal walls, whereas edges with a face on only one side are classified as external walls (Figure 3, b1–b2).
The final 3D structure is a multi-level topology coupling volumes and surfaces:

- ***Volume level:*** Room polygons are extruded between floor and ceiling to create closed cells that partition space. Shared faces encode room–room and room–exterior adjacencies, enabling connectivity and circulation queries.
- ***Surface level:*** Each wall, floor and ceiling face is explicit and grouped under its parent room volume. Faces are consistently oriented and stitched so that the assembly is 2-manifold.

3.3 Graph-based DT network

Volumetric and surface models are unified in the multi-level DT as a typed property multigraph (NetworkX `MultiDiGraph`) persisted in PostgreSQL. The graph encodes a strict hierarchy (*Building, Storey, RoomShell, Face, Edge, Vertex*) with complementary surface- and volume-level topology.
Containment and incidence relations define a directed acyclic subgraph, whereas adjacency is modelled as reciprocal directed-edge pairs in the *MultiGraph*. Nodes and edges carry stable identifiers and descriptive attributes (e.g., material, *U*-value, units). Each *Face* references its parent *RoomShell* and Rooms index their Faces, mirroring the DCEL bidirectionality for reliable adjacency and internal/external wall detection (Figure 3, b1–b2). The same structure is also exposed as RDF via a lightweight schema enabling SPARQL queries: node labels map to `rdf:type`, node attributes to datatype properties and edge labels to object properties (via a custom `nx_to_rdf`). This maintains semantic and topological consistency, supports RDF-based reasoning and enables dynamic updates for simulation workflows (e.g., energy zoning, material, construction variants; Section 3.6). Concretely, the `MultiDiGraph` supports computation (topology analysis, simulation setup, dynamic material updates), while the RDF view serves semantic querying and interoperability.

3.3.1 Typed Property Graph for Model Enrichment

The typed-property multigraph supports incremental enrichment, interoperability and efficient querying. Beyond topology, nodes and edges store extensible attributes, external references and library links. The model is enriched with:

- *layered materials assemblies*, assigned on element types and thickness, with energy-related parameters as conductivity k, density ρ with derived *U-values*;
- *device properties* (asset IDs, sensor types, locations);
- links to national *technical standards* (e.g., setpoints, indoor environmental parameters).

A robust material-assignment workflow was validated as follow. A *construction_layered* assembly is an ordered stack of *Layers*, each referencing a *Material*. From this point, thermal resistance and transmittance are computed as (Eqs. 2–3):

$$R_{\text{tot}} = R_{\text{si}} + \sum_{i=1}^{n} R_i + R_{\text{se}} \tag{2}$$

$$U = \frac{1}{R_{\text{tot}}} \tag{3}$$

For opaque elements (e.g. walls), assemblies are selected by minimizing $\sum_i | t_i - \tau |$ within a 0.025 m tolerance, where τ is the minor mOBB thickness derived from the point-cloud graph, while enforcing internal/external orientation from `thematic_surface`. For windows, U-values are instead assigned directly. Assemblies are exported as epJSON (*Material/Construction*), gbXML (*Construction/Layer*) and IFC (*IfcMaterialLayerSet*). Finally, national standards provide typologies and seasonal setpoints, and the nearest *.epw* file supplies climate data for baseline and scenario simulations.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-1-2026-447-2026 | © Author(s) 2026. CC BY 4.0 License.

3.4 Parser-Transformer-Writer (PTW) for interoperability

The graph model provides a robust, multi-layer organization comprising: (i) *a strict hierarchy* for containment and inheritance; (ii) *a topological layer* capturing areas/volumes, adjacencies and spatial reasoning; (iii) *a semantic layer* encoding material, assemblies, devices and their properties; and (iv) *a compliance layer* parameterized by national energy codes and standards (Figure 4). Building on this, three parser–transformer–writers (PTW) are derived to generate: (i) epJSON for direct ingestion by EnergyPlus (U.S. Department of Energy, 2025), (ii) IFC (optional in this workflow), and (iii) gbXML. The formats (Table 1) are complementary: IFC provides a comprehensive object-oriented BIM with detailed geometry and relations (in STEP format), gbXML targets energy-exchange with a zone-oriented XML structure and simplified geometry, and epJSON is a flexible JSON format optimized for detailed EnergyPlus simulations.

3.4.1 Interoperability: epJSON for EnergyPlus

In recent years, the EnergyPlus community has supplemented the traditional *.idf* format with a JSON-based input schema (*epJSON*) to streamline data ingestion and interoperability.

In this work, the *epJSON* representation is generated from the graph-based DT through a deterministic mapper that traverses stable identifiers across the building topology, including building, storey, room/shell, face, edge and vertex and produces EnergyPlus object graphs consistent with the EnergyPlus Input Data Dictionary (IDD). Topological relations, such as adjacency and boundary classification type, guide the creation of objects, including `Zone`, `Surface:HeatTransfer` and `ZoneMixing`, while layered materials are mapped to `Material` and `Construction` definitions with derived *U*-values. Semantic references to national technical standards are resolved into `Schedule`, `DesignSpecification`, `SetpointManager` and related Heating, Ventilation and Air Conditioning (HVAC) objects to ensure schema compliance and simulation readiness.

3.4.2 Interoperability: IFC for BIM

The second Parser–Transformer–Writer (PTW) module derives the IFC model from the graph. The parser ingests the Scan-to-EDTs CSV output from the DT, including thematic surface types, parent identifiers, relationships, materials, constructions and reconstructs the hierarchical and topological graph in memory. The hierarchy is rebuilt from *parent_id*, which encodes the underlying topology, while *thematic_surface* types map custom outputs to IFC classes. The transformer maps DT classes and relations to (IFC3x4) schema, entities and associations, adding geometry, placements and materials. Hierarchical relations become the IFC spatial structure (*Project, Site, Building, Storey, Space*); adjacencies drive space boundaries and openings; material and layer data assemble constructions. The writer then serializes a valid IFC, assigning names and *GlobalIds* from DT identifiers for traceability. This modular design supports incremental export, schema-level validation and openBIM alignment while preserving DT semantics (Figure 5).

	IFC	gbXML	epJSON
Format file	STEP (EXPRESS)	XML (XSD schema based)	JSON
Encoding	Plain text, line-based	XML tags	JSON with nested keys
Structure	Object graph network	Hierarchical Tree	Key-value
Application	BIM coordination	Energy modelling	Energy simulation
Orientation	Geometry-oriented	Space and zone centric	Simulation oriented
Geometry detail	High/medium detail (LOD)	Planar simplified surfaces	Planar geometry, vertex array

Table 1. Features of IFC, gbXML, epJSON format files.

3.4.3 Interoperability: gbXML for energy applications

The third Parser–Transformer–Writer (PTW) module exports analysis-ready gbXML from the graph-based DT, preserving geometric fidelity and thermal semantics. The format is an *Extensible Markup Language* (XML) schema with a *Geography Markup Language* (GML)-like hierarchical structure (elements, attributes, coordinates and explicit spatial relationships). The parser extracts the spatial hierarchy, surface topology, materials, openings and georeferencing data and reconstructs the logical building structure from the *parent_id* links and classified surface types. The transformer maps graph entities to gbXML constructs

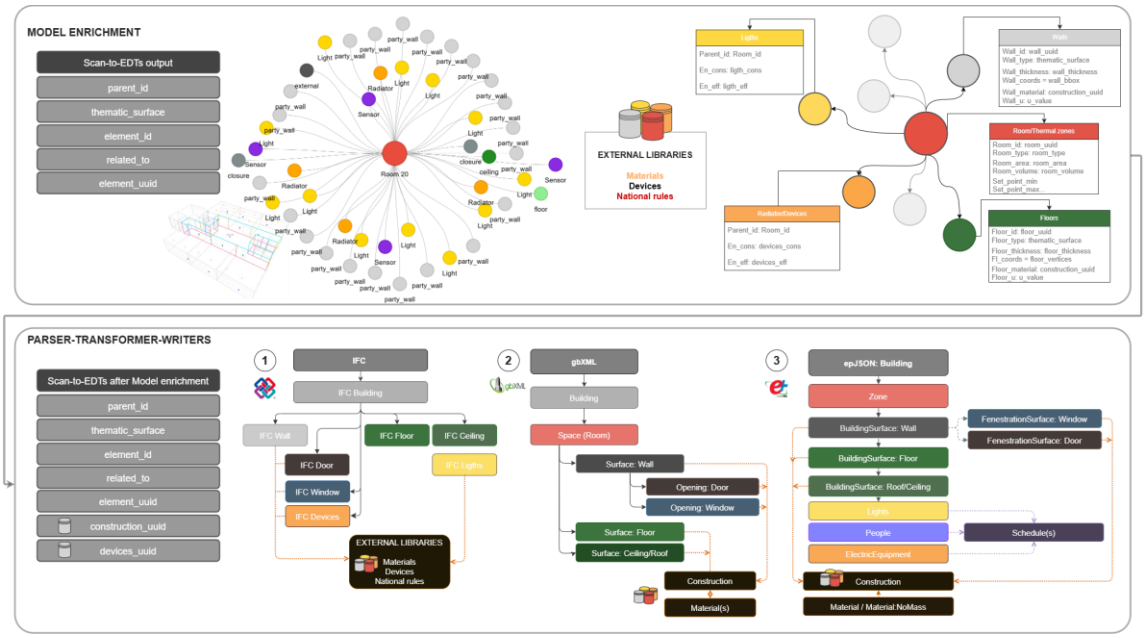

Figure 4. Top: Scan-to-EDT output fields and related graph, showing how semantics are embedded and enriched from external libraries. Bottom: From the enriched graph, the Parser–Transformer–Writer pipeline generates IFC, gbXML and epJSON models.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-1-2026-447-2026 | © Author(s) 2026. CC BY 4.0 License.

(*Space, Zone, Surface, Opening, Construction*), embedding geometry, orientation and material properties, and uses topology to infer adjacency, boundary type and exposure conditions. The writer outputs a validated gbXML document with DT-consistent identifiers for downstream energy simulation tools.

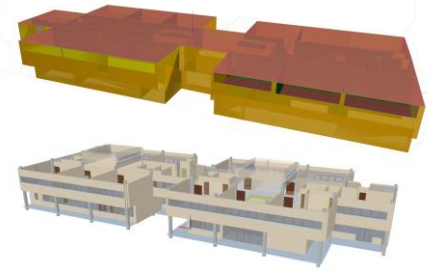

Figure 5. The gbXML and IFC file for the school building.

3.5 Data-driven scenarios

The modular algorithm generates epJSON files directly from retrofit inputs and instantiates simulation scenarios (Figure 6). From the main graph, the module derives the current building state (*baseline scenario*, Figure 6, step 1). In the retrofit stage (step 2), it targets specific elements (e.g., external walls $\mathcal{W}_{ext}$ or openings $\mathcal{O}$) via typed queries on *thematic_surface*, leveraging DCEL-derived internal/external typology and stable UUIDs, optionally refined with PostgreSQL/PostGIS spatial filters.

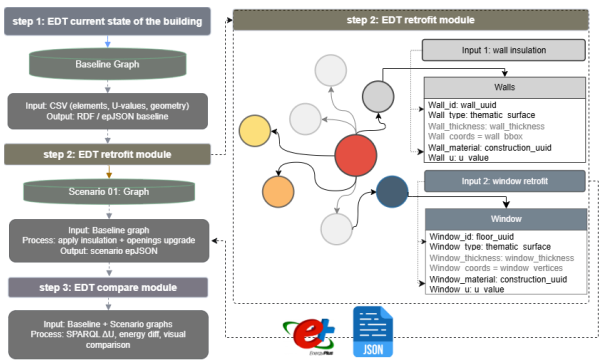

Figure 6. Modular pipeline for scenarios creation.

For each element, the algorithm retrieves its *uuid*, element type and *construction_id*, then follows the chain to *layer_id* and *material_id* to calculate the $U-value$ (Wm^2 K^{-1}). At the moment, two retrofit classes are supported:

- insulation of opaque envelope elements such as walls and roofs by reassigning or adding layers;
- upgrade of exterior openings toward a specified U_{target}.

Updates are transactional, such as append/replace insulation or set U_{target}, and the modified *layers* and resulting U are written to the scenario epJSON for simulation. Scenarios are executed into EnergyPlus with recorded inputs and timestamps for direct baseline comparison (step 3).

4. Results

As described above, the multi-level DT framework comprises: (i) Scan-to-BIM solid mesh modelling (classification and instance segmentation) (Section 4.1); (ii) volumetric and surface-based topological modelling (Section 4.2); and (iii) multi-level DT generation. Quality is evaluated on the datasets shown in Figure 7 using task-specific metrics. For Scan-to-BIM versus point-cloud classification, mIoU and F1 (Eqs. 4–5) are computed on matched minimum oriented bounding boxes (mOBBs):

$$IoU = \frac{|A \cap B|}{|A \cup B|} \tag{4}$$

$$F1 = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \tag{5}$$

Where A and B are the predicted and ground-truth (GT) OBBs, $|A \cap B|$ and $|A \cup B|$ denote intersection and union volumes, and TP, FP, FN are matched predictions, unmatched predictions, and unmatched GT OBBs, respectively. For the Topologic model, we additionally report surface/volume Precision and Recall (Eqs. 6–7), and vertex average precision (mAP_v, Eq. 8).

$$\text{Precision} = \frac{TP}{TP + FP} \tag{6}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{7}$$

$$\text{mAP}_\text{v} = \frac{1}{|\mathcal{T}|} \sum_{\eta \in \mathcal{T}} \left(\int_0^1 P_\eta(R)\, dR \right) \tag{8}$$

Where $|\mathcal{T}|$ is the set of distance thresholds, η is the threshold, $P_\eta(R)$ is the precision at recall (R) for that threshold.
The mAP_v measures then how accurately vertex positions match GT vertices within a distance threshold η_v of any GT vertex.

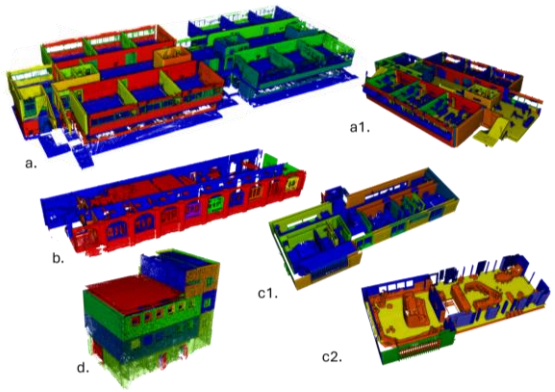

Figure 7. Classify datasets used in this work: (a) School; (a1) School subset; (b) Santa Chiara building; (c) House 1, (c1) floor plan, (c2) first floor; (d) Campus Gent, Block N.

4.1 Scan-to-BIM results

The proposed framework was evaluated on four distinct datasets (Table 2, Figure 7), including a private residence (a), two public school buildings (b, c) and a heritage complex building (d).

	House 1	***Campus - Block N***	***School***	***Heritage***
Acquisition	MLS	TLS	TLS	TLS
Point density	6 mm	5 mm	5 mm	5 mm
# Floorplans	2	4	2	1

Table 2. Datasets for Scan-EDTs framework validation.

Table 3 reports the average results for the classification and instance segmentation task. The model shows high and consistent performance across the different building datasets, achieving strong accuracy for large structural classes, such as floors and ceilings, with mIoU around 92.73–94.35% and lower accuracy for smaller or detailed elements such as columns, doors and windows with mIoU around 65.33–72.61 %. Metrics have been evaluated with thresholds 0.05, 0.065 and 0.10 m.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-1-2026-447-2026 | © Author(s) 2026. CC BY 4.0 License.

Mean results remain stable across datasets with low deviation (about 1–3%, above 5% for unclassified components), indicating robust generalization despite challenges in fine details.
The Scan-to-BIM pipeline (Section 3.1) produces simplified mesh-based solids as parametric objects (Figure 2); for graph-level analysis, each object is represented by its mOBB. Comparing these mOBB with those from point-cloud classification provides precision and recall metrics (Table 3).

Class	*Std dev [%]*	*mIoU*	*mAcc [%]*
Unclassified	5.12	78.28	85.38
Foors	2.77	94.35	96.84
Ceilings	2.86	92.73	95.77
Walls	2.53	87.51	89.87
Columns	0.85	46.14	72.61
Doors	1.17	51.62	65.33
Windows	1.55	58.83	70.24

Table 3. Classification results using PTv3, average per class.

Walls achieve the highest scores due to their large, planar geometry and good sampling, while doors and windows show a larger F1 mIoU gap, reflecting accurate detection but modest localization errors on thin, small elements.
Table 4 reports the results. Walls show the best performance, with mIoU between 90–96% and F1 between 86–94%. Windows and doors perform moderately, with mIoU and F1 mostly in the 60–70 % range, while columns are lower, with F1 around 81–89 % and mIoU 75–85%. Overall, F1 follows mIoU closely, with walls showing the highest consistency and openings/columns greater variability.

Classes	*House 1*	*Campus - Block N*	*School*	*Heritage*
Columns mIoU	85.2	-	78.3	75.1
Columns F1	86.9	-	81.2	89.1
Doors mIoU	69.55	73.4	76.4	75.1
Doors F1	71.6	75.1	77.9	89.1
Walls mIoU	89.25	91.7	96.3	94.8
Walls F1	86.10	88.6	94.0	92.0
Windows mIoU	67.95	71.3	75.8	76.2
Windows F1	69.85	72.8	76.7	77.5

Table 4. Metrics results of the solid-based model reconstruction.

4.2 Topologic model

For the evaluation of the topological models, two types of assessments are conducted: (i) a geometric assessment, based on the quantitative metrics defined in (Eqs. 4–8) and (ii) a consistent and qualitative assessment in FME® software, comparing manually developed GT models with the automatically derived geometries to analyse reconstructed areas, volumes and the number of identified elements.
The geometric assessment used precision, recall, F1 score, mIoU, and *mAPV* (thresholds of 0.05, 0.075, 0.10 m) for both volumetric and surface-based models.

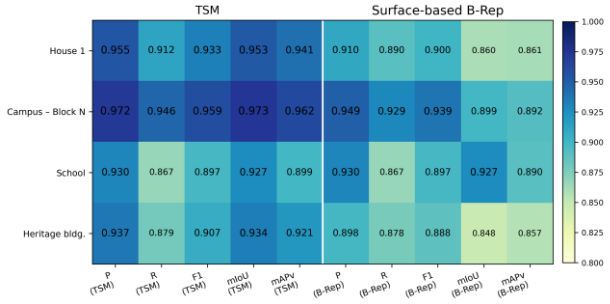

Figure 8. Metrics for Topologic Solid and B-Rep model.

Figure 8 and Table 5 show that the *Topologic Solid Model (TSM)* consistently outperformed the *Surface*-based model, achieving higher mIoU (0.95 vs. 0.88) and mAP_v (0.93 vs. 0.88), along with slightly superior precision, recall and F1 across all datasets, with the largest gain on *Campus Block N* and the smallest on *School* dataset (Figure 7). Finally, geometric/topologic consistency was checked in FME®, confirming correctness of reconstructed elements.

4.3 Graph enrichment and dynamic scenarios

The Scan-to-EDT framework enables on-demand creation of simulation-ready DTL2 instances by substituting materials and constructions for retrofit assessment.
In this proof of concept, each scenario reports annual energy use and results are served via an API for side-by-side comparison. Figure 9 (Section 3.5) contrasts the school-subset baseline with user-defined retrofits (added wall insulation or reduced window U-value). User inputs (insulation thickness in meters or a target window U-value) are translated through dynamic graph operations into epJSON models, executed in EnergyPlus, and returned with timestamps and analytical plots. (Figure 9).

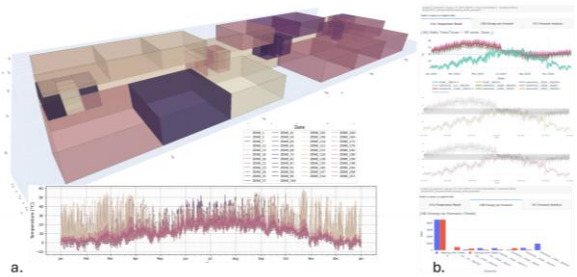

Figure 9. (a) Volumetric model with preliminary EnergyPlus outputs; (b) interface outputs for cross-scenario analysis.

Given an *Insulation_layer* input, exterior walls are targeted, *area* and *construction_id* are retrieved, the layer stack is resolved, and an insulation layer from the material library is inserted. Then, per-layer resistances are recomputed, U is updated (Eqs. 2–3), and the upgraded construction is written to epJSON.
For windows, the *U-value* is set directly to the target and recorded. A proof-of-concept simulation is run on dataset a1 (Figure 7) using *Zone Ideal Loads Air System*. The EnergyPlus model is generated automatically as epJSON (*v25.1*) from the graph: geometry is exported as `BuildingSurface:Detailed` polygons (one zone per room shell), constructions are assembled from the layered-material library, and each zone is controlled by a constant dual-setpoint thermostat (20 °C heating / 26 °C cooling) via `ZoneHVAC:IdealLoadsAirSystem`.
Temperature fidelity is evaluated against GT measurements using the mean absolute error (Eq. 9):

$$E_\theta = \frac{1}{n}\sum_{i=1}^{n} \mid t_{GT,i} - t_{\text{simul},i} \mid, \tag{9}$$

Where $t_{\text{GT},i}$ is the IoT-recorded temperature at time i (ground truth) and $t_{\text{simul},i}$ is the simulated temperature at the same instant. The resulting errors are $E_\theta = 0.81°C$ for daily and $E_\theta = 1.97$ °C for monthly simulations. These results indicate good optimization and stable performance across time scales, while also suggesting room for improvement through finer calibration of boundary conditions, internal gains and material properties.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-1-2026-447-2026 | © Author(s) 2026. CC BY 4.0 License.

4.4 Interoperability: Parser-Transformer-Writers outputs

Interoperability remains a core challenge in the AEC domain. The presented graph-based representation enables export to gbXML and optionally IFC for multi-platform use.
Figure 10 shows (i) gbXML output validated in a viewer and (ii) a topological consistency check in FME. Although the reconstructed gbXML is geometrically consistent (Figure 11), the viewer displays both volumetric and surface representations, reflecting the multi-level DT construction, causing full walls to occlude perforated ones. Minor visual mismatches also appear when stacking multiple floors, especially along exterior walls, as each floor is processed independently without enforcing a global exterior envelope.

Dataset	*Scan-to-EDTs/gbXML*	*ΔA [%]*	*ΔV [%]*	T_{FME} *[s]*	T_{gbXML} *[s]*
a	293/293	3.83	5.86	2.1	3.1
a1	130/130	0.92	4.72	1.7	2.5
b	146/146	3.45	8.33	1.6	2.8
c	62/62	2.67	9.10	1.3	2.9
d	58/58	1.25	8.23	1.4	2.7

Table 4. Per-dataset metrics for the gbXML transformer.

Furthermore, a preliminary CityGML 2.0 with Energy ADE 3.0 conversion (Agugiaro and Padsala, 2025) on dataset a1 (Figure 7) confirms the schema compatibility and feasible mapping of hierarchical and topological structures for broader urban or district-scale integration.
From a simulation and energy-efficiency perspective, Table 4 shows that area and volume errors are highest in the most geometrically complex buildings, especially those with many exterior walls and irregular shapes (e.g., the *School* dataset), and in multi-storey buildings (dataset b, *Campus-Block N*), where errors tend to accumulate.

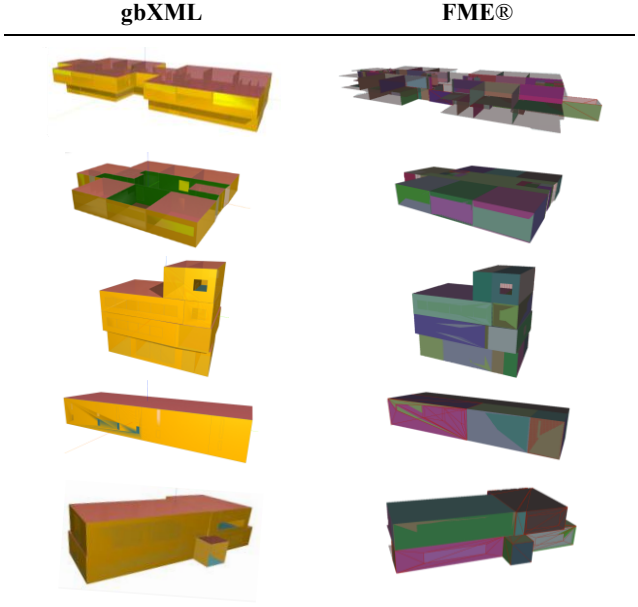

Figure 10. Visual result for gbXML models and visualisation in FME software.

Reconstruction losses mainly arise from upstream classification and detection, with the dominant error due to reconstructing walls from centrelines rather than exterior surfaces, as required by gbXML. Volume errors stem from 2D imperfections propagating into 3D and from using slab mean elevation (z_{mean}) instead of floor elevation (z_{floor}), which inflates volumes and increases ΔV.
Future work will implement centreline-to-exterior-wall and corrected z computation to improve accuracy. Nevertheless, the parse–transform–write (PTW) stage is fast, as all elements parsed in Scan-to-EDTs are converted to gbXML with low conversion runtimes T_{FME} and T_{gbXML}(s) measured on standard hardware Intel® Core™ i7-12700H, 2.30 GHz, 32 GB RAM.

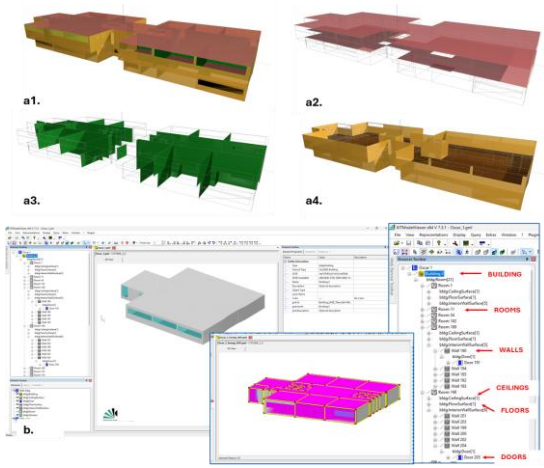

Figure 11. (a) Topologically consistent model: (a1) complete, (a2) ceilings, (a3) internal walls, (a4) floors and walls; (b) ivisualisation of the same model exported to CityGML and Energy ADE 3.0.

5. Conclusions & Future works

The proposed Scan-to-EDT framework is a standards-aware end-to-end pipeline that converts raw point clouds into simulation-ready DTs with validated detailed geometry and topology.
In particular, the framework’s results demonstrated:

- *Validation* (Figure 8): multi-room B-Rep reconstructions achieved 0.86–0.89 mAP_v (vertex-level accuracy) across four buildings, demonstrating robustness to varied layouts.
- *Unified graph layer:* a typed property graph couples geometry, semantics and simulation parameters, enabling consistent scenario updates and reversible gbXML↔epJSON (optionally IFC) transformations.
- *Interoperability:* from the derived graph, gbXML and epJSON instances achieve complete schema-valid exports.

Overall, the framework supports faster and repeatable modelling for energy analysis and indoor monitoring, while leaving room for improvement.
Future work will (i) expand benchmarking; (ii) enhance classification methods for better results; (iii) scale validation to district and city levels; (iv) strengthen alignment with IFC, CityGML 3.0, and Energy ADE; and (v) improve live DT operation through sensor-based calibration, monitoring, and on-demand epJSON generation with comparative analytics.

Acknowledgments

The presented work was carried out within the co-financed program funded by the European Union – Next Generation EU (Mission 4, Component 2, CUP E66E22000050008). Part of the work presented in this paper has been carried out within the European project DigiTwins4PEDs (Utilization of urban digital twins to co-create flexible positive energy systems for districts). The project is funded by the European Commission under the Horizon Europe Partnership scheme (Funding Indicator: 03EN3081A). The DUT Call 2022 also contributes to the Urban Transition Mission of Mission Innovation as part of the MICall 2022 initiative.

References

Agugiaro, G., Benner, J., Cipriano, P., Nouvel, R., 2018. The Energy Application Domain Extension for CityGML: Enhancing interoperability for urban energy simulations. *Open Geosp. Data, Software and Standards*, 3:2. SpringerOpen, United Kingdom.

Agugiaro, G. and Padsala, R., 2025. A proposal to update and enhance the CityGML Energy Application Domain Extension, *ISPRS Ann. Phot. Remote Sens. Sp. Inf. Sci*., X-4/W6-2025, 1–8.

Biljecki, F., Ledoux, H. Stoter, J., 2016: An improved LOD specification for 3D building models. *Computers, Environment and Urban Systems*, 59, 25–37.

Buyukdemircioglu, M. and Oude Elberink, S., 2025. Enriching Urban Digital Twins with Energy-related Information from Aerial and Street View Imagery for Precise Urban Climate Modeling. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-M-6-2025, 79–85.

Chen, Z., Wang, Y., Nan, L., Zhu, X. X., 2025. Parametric Point Cloud Completion for Polygonal Surface Reconstruction. *CVPR,* pp. 11749–11758.

El-Mekawy, M., Östman, A., 2010. Semantic mapping: an ontology engineering method for integrating building models in IFC and CityGML. In: *Proceedings of the 3rd ISDE Digital Earth Summit*, Nessebar, Bulgaria, 12–14 June 2010, pp. 1–11.

~ Software. Copyright (c) Safe Software Inc. www.safe.com

Gan, V. J. L., Li, K., Li, M., Halfian, L. B. E., 2025. 3D reconstruction of building information models with weakly-supervised learning for carbon emission modelling in the built environment. *Applied Energy*, 377, 124695.

Gao, W., Agugiaro, G., León-Sánchez, C., 2025. Data-Driven Energy Simulations to Evaluate Positive Energy District Potential in Rotterdam, *ISPRS Ann. Phot. Remote Sens. Spatial Inf. Sci.*, X-4/W7-2025, 25–32.

Horna, S., Meneveaux, D., Damiand, G., Bertrand, Y., 2009. Consistency constraints and 3D building reconstruction. *Computer-Aided Design* 41(1), 13–27.

Hu, W., Cai, Y., 2024. A semi-supervised method for digital twin-enabled predictive maintenance in the building industry. *Neural Computing and Applications*, 36, 15759–15775.

Klar, R., Arvidsson, N., Angelakis, V., 2024. Digital Twins' Maturity: The Need for Interoperability, in *IEEE Systems Journal*, vol. 18, no. 1, pp. 713-724.

Kolbe, T.H., Gröger, G. Plümer, L., 2005. CityGML: Interoperable access to 3D city models. In: P. van Oosterom, S. Zlatanova and E.M. Fendel (Eds.), *Geo-information for Disaster Management*. Springer, Berlin, Heidelberg, pp. 883–899.

Lee, J., Park, C., Choe, J., Wang, Y.-C. F., Kautz, J., Cho, M., Choy, C., 2025. Mosaic3D: Foundation Dataset and Model for Open-Vocabulary 3D Segmentation. *In: Proc. IEEE/CVF (CVPR), 14089–14101.*

Lu, D., Xu, L., Zhou, J., Gao, K., Gong, Z., Zhang, D., 2025. 3D-UMamba: 3D U-Net with state space model for semantic segmentation of multi-source LiDAR point clouds. *Int. Jour. of Applied Earth Observation and Geoinformation*, *136*, p.104401.

Mayer, K., Vesel, A., Zhao, X., Fischer, M., 2025. SYNBUILD-3D: a large, multi-modal, and semantically rich synthetic dataset of 3D building models at Level of Detail 4. *arXiv* preprint, arXiv:2508.21169.

Pantoja-Rosero, B.G., Rusnak, A., Kaplan, F., Beyer, K., 2024. Generation of LOD4 models for buildings towards the automated 3D modeling of BIMs and digital twins. *Automation in Construction*, 168, 105822.

Robert, D., Raguet, H., Landrieu, L., 2024. Scalable 3D Panoptic Segmentation as Superpoint Graph Clustering. *arXiv preprint* arXiv:2401.06704.

Roman, O., Mazzacca, G., Farella, E. M., Remondino, F., Bassier, M., Agugiaro, G., 2024a. Towards Automated BIM and BEM Model Generation using a B-Rep-based Method with Topological Map. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, X-4-2024, 287–294.

Roman, O., Bassier, M., De Geyter, S., De Winter, H., Farella, E. M., Remondino, F., 2024b. BIM Module for Deep Learning-driven parametric IFC reconstruction, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-2/W8-2024, 403–410.

Schlenger, J., Pluta, K., Mathew, A., Yeung, T., Sacks, R., Borrmann, A., 2025. Reference architecture and ontology framework for digital twin construction. *Automation in Construction,* 174, 106111.

Steinke, T., Büchner, M., Vödisch, N., Valada, A., 2025. Collaborative Dynamic 3D Scene Graphs for Open-Vocabulary Urban Scene Understanding. *arXiv preprint*, arXiv:2503.08474.

U.S. Department of Energy’s (DOE) Building Technologies Office (BTO), Energy Plus, 2025.

Wang, Y., Wang, X., Liu, A., Zhang, J., Zhang, J., 2025. Ontology of 3D virtual modeling in digital twin: a review, analysis and thinking. *Jour. of Intell. Manuf.*, 36(1), 95–145.

Wu, X., Jiang, L., Wang, P.-S., Liu, Z., Liu, X., Qiao, Y., Ouyang, W., He, T., Zhao, H., 2024. Point Transformer V3: Simpler, Faster, Stronger. *arXiv preprint*, arXiv:2312.10035.

Wysocki, O., Schwab, B., Biswanath, M. K., Zhang, Q., Zhu, J., Fröch, T., Heeramaglore, M., Hijazi, I., Kanna, K., Pechinger, M., et al., 2025. TUM2TWIN: Introducing the large scale multimodal urban digital twin benchmark dataset. *arXiv* preprint, arXiv:2505.07396.

Yang, Y., Pan, Y., Zeng, F., Lin, Z., Li, C., 2022. A gbXML Reconstruction Workflow and Tool Development to Improve the Geometric Interoperability between BIM and BEM. *Buildings* 12(2), 221.

Zhou, W., Liu, K., Jin, W., Wang, Q., She, Y., Yu, Y., Ma, C., 2025. Advancements in deep learning for point cloud classification and segmentation: A comprehensive review. *Computers & Graphics*, 130, 104238.