Energy & Buildings 325 (2024) 115035

Contents lists available at ScienceDirect

Energy & Buildings

journal homepage: www.elsevier.com/locate/enbuild

A knowledge graph-based framework to automate the generation of building energy models using geometric relation checking and HVAC topology establishment

Meng Wang *, Georgios N. Lilis, Dimitris Mavrokapnidis, Kyriakos Katsigarakis, Ivan Korolija, Dimitrios Rovas

The Bartlett Institute for Environmental Design and Engineering, University College London, London, UK

ARTICLE INFO

Keywords:
Building digital twins
BIM2BEM
Ontology
HVAC topology
Geometric relation checking

ABSTRACT

Building Energy Models (BEM) are widely utilized throughout all stages of a building's lifecycle to understand and enhance energy usage. However, creating these models demands significant effort, particularly for larger buildings or those with complex HVAC systems. While a substantial amount of information can be extracted from Building Information Models (BIM) — which are increasingly accessible and provide necessary data for geometric and HVAC contexts — this information is not readily usable in setting up BEM and typically requires manual translation. To address this challenge, this paper introduces a BIM-to-BEM (BIM2BEM) framework that focuses on automating the generation of HVAC parts of BEM models from BIM data. Core to the methodology is the extraction of HVAC system topologies from the BIM model and the creation of a knowledge graph with the HVAC topology. The topology transformation unfolds in three key stages: first, a geometry-induced knowledge graph is established by examining the geometric relationships among HVAC elements; second, this graph is converted into an informative HVAC topology with enhanced properties from additional data sources; and finally, the informative topology is simplified into a BEM-oriented HVAC topology compliant with BEM platforms such as EnergyPlus. A case study of a large university building with a complex HVAC system showcases that the proposed framework achieves automatic and precise generation of building performance simulation models. The model's predictions are then validated against actual measurements from the building.

1. Introduction

Accurate representation of Heating, Ventilation, and Air Conditioning (HVAC) systems is essential for energy modelling throughout the building lifecycle and for digital twin applications that support building operational management [1]. While the importance of HVAC topology representation is widely acknowledged, manual instantiating remains a significant challenge. Automating the extraction of HVAC topology information can greatly enhance its applicability across design and operational domains [2]. Building Information Modelling (BIM), which provides detailed and comprehensive building and HVAC data, can support the generation of HVAC topology that aligns with the input requirements of Building Energy Models (BEM). To fully exploit this potential within the BIM-to-BEM (BIM2BEM) process, a seamless transformation framework is required, with an emphasis on integrating HVAC data from BIM into the BEM context. This framework is underpinned by two critical sub-processes: (a) the automated extraction of HVAC topology from BIM data and (b) the subsequent mapping of this topology into a specific BEM context. Hence, the present study proposes a knowledge graph-based approach, focusing on HVAC systems, which employs geometric relation-checking techniques to automate the establishment of HVAC topology. The relevant information within this topology is then effectively transferred into the geometric BEM model, resulting in a simulation-ready model with a fully detailed HVAC configuration.

1.1. Literature review

Early work by Bazjanac et al. [3], showcased the potential but also challenges of BIM2BEM methods. This potential is realised through BIM's parametric modelling capabilities and interoperability standards, which make it an effective data source for building energy models [4]. BIM provides a detailed and accurate representation of architectural

* Corresponding author.
E-mail address: wang.meng@ucl.ac.uk (M. Wang).

https://doi.org/10.1016/j.enbuild.2024.115035
Received 10 September 2024; Received in revised form 30 October 2024; Accepted 7 November 2024
Available online 14 November 2024

and HVAC information, including geometric data, system connections, and specific properties, making it a reliable foundation for performance simulations [5]. Unlike other data sources, a BEM derived from BIM preserves the intricate details of the building design, thereby minimising human assumptions and biases. Moreover, simulation uncertainties often stem from imperfections and insufficient information in data sources—issues that BIM effectively resolves by providing a reliable and comprehensive representation of the building [6].

The process of generating a BEM from BIM has been the subject of extensive research, covering aspects of the transformation process and focused on approaches for multi-domain data integration covering a wide range of domains such as building geometry, building envelopes, energy systems, occupant behaviour, and control strategies. The transfer of building geometry information is a core area of interest within this domain, attracting significant research into efficient methodologies and effective data sources. Many geometry-related algorithms have been developed to tackle the challenge of extracting the second-level space boundary (2LSB) topology from BIM models, which is essential for accurately creating building geometry models for performance simulations [7–10]. In exchanging information from BIM to energy-modelling tools, two data formats are frequently used: the Green Building XML schema (gbXML) and Industry Foundation Classes (IFC). The stated purpose of gbXML is to be an interoperable data format for energy-related information exchanges and is a widely adopted exchange scheme for building geometry descriptions, as reported by Yang et al. [11]. One of the benefits of gbXML is its ability to minimise the manual effort involved in restructuring building models for energy modelling, thereby streamlining the handling of input data. However, a persistent challenge remains in extracting the necessary information for data exchange in either format (gbXML or IFC). Tools that can process geometric data from BIM models to extract second-level space boundaries (2LSB) are indispensable, regardless of the container format used. Delgado et al. [12] utilized Revit to export gbXML from BIM data captured by drones, which was then imported into a streamlined BIM2BEM workflow for modelling a single-family wooden house in Portugal. Another Revit-based workflow considering gbXML as the BIM input data format was proposed by Elnabawi [13], who emphasised that the lack of standardisation results in variations between modellers and may cause delays in its application. Furthermore, the IFC schema is a comprehensive BIM data standard encompassing a broad range of building data and managing complex, multidisciplinary information, making it highly versatile for collaboration throughout the building lifecycle. Its broad support by BIM tools and regular updates to align with industry needs encourage scholars to explore its integration into BIM2BEM workflows. Ramaji et al. [14] investigated the study on IFC-based BIM2BEM workflow to translate the open BIM standards into OpenStudio data format, concentrating on architectural information conversion. Chen et al. [15] proposed a BIM2BEM toolkit, AutoBPS-BIM, combined with a chiller optimisation module to evaluate energy loads, demonstrating the potential for extending the BIM2BEM process. Giannakis et al. [16] proposed an automatic BIM2BEM workflow integrating a series of tools, including the Revit IFC Exporter, Common Boundary Intersection Projection (CBIP) tool, and SimModel XML tool, to generate an EnergyPlus Input Data File (IDF) with precise building geometric configurations. Moreover, the interoperability of building geometry and envelope information between gbXML and IFC was also investigated by Kamel et al. [17], demonstrating its potential to enhance BIM2BEM transformation. Among the various types of building information, research on geometric data within the BIM2BEM domain is relatively advanced, with several effective workflows already established that provide proof of concept. However, their scalability remains in question, as many fail to perform effectively in practice. Hence, studies recently have increasingly focused on generating reliable BEM geometric models from imperfect BIM sources, as well as extending these workflows to other information types, such as building envelopes and mechanical systems.

Another important domain in the transformation is the handling of building systems, such as HVAC, where accurate transformation leads to more realistic building simulation models with detailed settings for active components. Pinheiro et al. [18] emphasized that the IFC Model View Definition (MVD) standardizes information exchange within specific domains, including HVAC systems, thereby streamlining BEM development and reducing modelling time, which can further support a semi-automatic BIM2BEM workflow. Autodesk Insight, a tool integrated with Revit, facilitates building performance simulation within the BIM2BEM process by incorporating HVAC settings, as reported by Gonzalez et al. [19]. EnergyPlus, the “gold standard” for whole building simulation, excels in modelling energy use, HVAC systems, and thermal performance, making it ideal as the simulation engine for the BIM2BEM process. Barone et al. [20] generated a preliminary building energy model for EnergyPlus using an automated BIM-to-gbXML-to-OpenStudio import routine, in which the HVAC system was modelled manually using the OpenStudio Application GUI. Li et al. [21] developed an automatic transformation approach to convert HVAC information from IFC data to OpenStudio by leveraging a data-driven reverse engineering algorithm. Their approach’s effectiveness and precision were showcased through a case study of an office building. Besides using EnergyPlus as the simulation engine, Modelica provides flexible, component-based modelling of complex building systems, making it another viable option to support dynamic simulation in the BIM2BEM process. Kim et al. [22] developed a Modelica library for BIM-based building simulation via an Object-oriented physical modelling (OOPM) approach. They investigated the system interface between BIM and BEM, concentrating on component-level simulation. Sayegh et al. [23] presented a semi-automated BIM2BEM workflow using Modelica and BuildingSysPro, providing templates for HVAC settings. They introduced a two-storey building as the case study to highlight the manual intervention required to detect and correct errors in BIM. IDA Indoor Climate and Energy (IDA ICE) was also used by Hosamo et al. [24]. Based on the above studies, whole-building simulation models require specific inputs and syntax. Despite these differences, certain key information areas must be addressed in all models. Tools like EnergyPlus include the geometric context of zone boundaries, as well as HVAC system topology. Meanwhile, the need for higher-quality BIM data has increased as the accuracy of transforming geometric and HVAC data improves. Low-quality BIM data can significantly hinder the automation of the BIM2BEM process, often requiring manual adjustments that reduce both efficiency and scalability.

While stating and meeting the information and exchange requirements is essential, the data structure adopted to represent relevant information is also important. If taking as an example the HVAC topology, which is of interest, using a data format like gbXML is an option. Yet, this information might also be useful in other analysis contexts, such as operation and maintenance or Digital Twins. To support information reuse across domains and specific applications, it is useful to select an ontology as a structured framework for representing and organizing relevant knowledge representation. Such an approach is interesting, focusing on the core concepts, relationships, and entities rather than requirements linked to a specific information exchange. By separating knowledge extraction — such as Building or HVAC system topologies — from how this information will be used, a clearer separation of concerns is achieved. Such structures have improved data interoperability and ensured seamless interaction across multiple domains, including energy performance and spatial analysis [25].

There are several ontologies used in the building domain [26], with extensive discussions on their potential applications. Of special interest for HVAC systems are the Brick Schema [27], Flow Systems Ontology (FSO) [28], and TUBES System Ontology (TSO) [29]. FSO and TSO are compatible with BIM data, focusing on energy and mass exchange in building system models at varying levels of granularity. At the same time, Brick provides a more general framework for standardizing building assets. Wu et al. [30] extended Brick to facilitate automatic build-

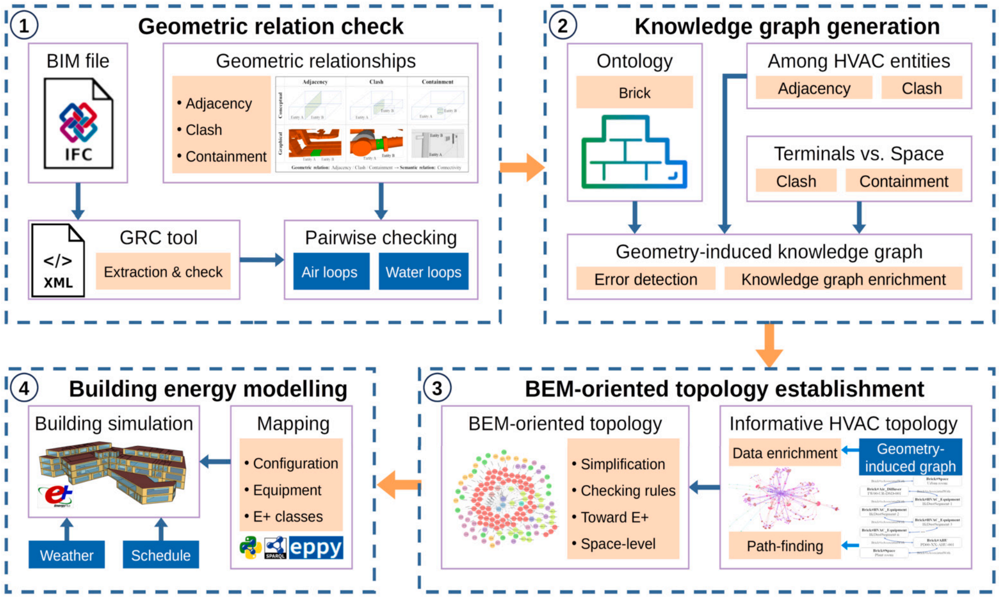


Fig. 1. The proposed framework of this study.

ing energy modelling using a rule-based reasoning method for thermal zoning. Their work concentrated on the automatic generation of building geometry yet with a simple hypothetical HVAC system. Fjerbæk et al. [31] automated Modelica-based simulations of heating systems using FSO to provide detailed representations of individual components and structural information of the heating system. They noted that applying this approach to larger, more complex systems would reveal outcomes that regular practitioners find difficult to quantify. These studies demonstrated the capacity of ontological frameworks to structure and integrate building data for performance simulations, facilitating more detailed and accurate model configurations [32]. Ontologies streamline data integration by addressing the challenge of managing heterogeneous data, offering a solution to improve the reliability and precision of building energy simulations, especially for complex buildings with intricate HVAC systems.

In general, current studies in the BIM2BEM context focus on investigating the feasibility and applying their approaches to small-scale buildings with simple energy systems or hypothetical configurations rather than large-scale buildings with complex and diverse HVAC systems. The main challenges may lie in handling heterogeneous data and managing imperfect input BIM data, often leading to incomplete HVAC and 2LSB topologies with unexpected errors. While these efforts have achieved some automation, issues like biased information transfer and poor input data quality still necessitate considerable manual modelling work or corrections. Building ontologies and semantic web technologies to structure and harmonise data offer an effective representation of complex information domains and are more flexible compared to standard BIM-based approaches; this flexibility makes such approaches interesting in such problems.

1.2. Motivation and contribution

As previously mentioned, current BIM2BEM transformation methods depend on high-quality BIM models that can be validated through checking tools. However, in practice, BIM models are often developed for design coordination and collaboration rather than specifically tailored for energy modelling with a focus on HVAC systems. Most available BIM data do not meet the stringent requirements for complex Extract, Transform, and Load (ETL) processes needed to generate accurate building simulation models. Therefore, for BIM2BEM workflows to be effective, the transformation processes must be capable of handling BIM models with moderate inaccuracies and design errors. This is particularly important for HVAC systems, where the topology of component interconnections, even with some imperfections, is often sufficient to generate a reliable and accurate BEM for performance simulation.

Hence, this study proposes a framework that integrates advanced techniques to effectively extract HVAC information from BIM models with design errors and leverages semantic web technologies to establish an HVAC topology that accurately reflects the real-world system. Then, the framework explores a method for generating BEM by mapping the HVAC topology to EnergyPlus objects. The framework is applied to a newly constructed building with complex HVAC systems to validate its effectiveness and feasibility. The main contributions of this are: (1) From a scientific perspective, the proposed framework efficiently generates reliable, simulation-ready building models with HVAC configurations that accurately mirror real-world systems, even when the input data contains errors; (2) From a practical perspective, the knowledge graph-based techniques introduced in this study facilitate the robust and efficient integration of heterogeneous data, acting as an intermediary between multi-domain data sources and building performance simulations.

The subsequent sections of the paper offer a more detailed explanation of the proposed approach: Section 2 presents the framework and its key stages, followed by Section 3, which offers an overview of the case study and the available data. Section 4 then delves into the results and discusses effectiveness and limitations. Finally, Section 5 summarize the study's findings and propose avenues for future research.

2. Methodology

The proposed methodology for automating BIM2BEM, focusing on HVAC information, comprises four key stages, as illustrated in Fig. 1.

(1) **Geometric relation check:** In this stage, one or more BIM models containing the geometric information for Architectural (ARC) and Mechanical, Electrical, and Plumbing (MEP) elements are exported in IFC format. A seamless workflow, based on the BIM context and a Geometric Relation Checker (GRC) tool, identifies the geometric relations among MEP and ARC elements. A pairwise checking process is then conducted to obtain the geometric relations of each pair in the MEP-IFC file and between terminals in the MEP-IFC file and spaces in the ARC-IFC file.

(2) **Knowledge graph generation:** In this stage, a geometry-induced knowledge graph for the HVAC domain is generated according to the geometric checking results obtained from step (1). This work

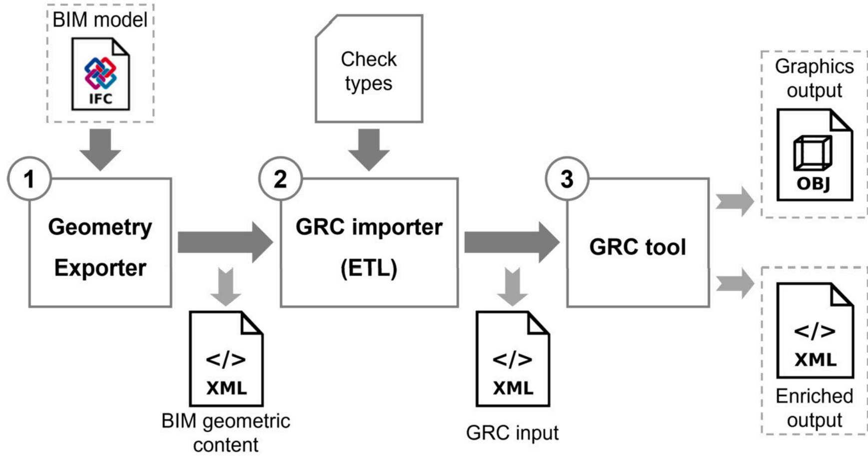


Fig. 2. Diagram of the transfer process from IFC to XML, followed by loading into the GRC tool.

then generates knowledge graphs following the Brick Schema for water loops and air loops, which describe the connections among HVAC elements and the relationships between terminal units and spaces.

(3) **BEM-oriented HVAC topology establishment:** In this stage, a path-finding algorithm is introduced to generate an informative HVAC topology to unlock the links among terminal units, Air Handling Units (AHUs), Mechanical Ventilation with Heat Recovery (MVHR), Fan Coil Units (FCUs), and water pumps. The entities referring to pipe segments, duct segments, fittings, and accessory devices are bypassed. Meanwhile, the device inventory is introduced to enrich the informative HVAC topology with device performance properties. Then, a set of rules is proposed to simplify the informative HVAC topology into a BEM-oriented one conforming to the building performance simulation data requirements.

(4) **Building energy modelling:** In the final stage, a mapping approach is developed to bridge the gap between BEM-oriented HVAC topology and building performance simulation. This mapping approach automatically generates an EnergyPlus model by combining the topology from step (3), the building's BEM geometric content and other data sources. e.g., weather, schedule, and control strategy. Furthermore, a building performance simulation is conducted to evaluate the current design of the HVAC system in terms of indoor environment and energy consumption.

The following four subsections detail each stage of the proposed framework.

2.1. Stage 1: geometric relation check

This initial stage aims to generate geometric relation pairs among the geometric solid representations of the building's MEP elements and space volume geometries. These relations will be used to identify semantic (non-geometric) relations among these elements at a later stage. Three additional subprocesses are involved in this stage, as analysed next.

2.1.1. Geometric information extraction

IFC, an open standard data model used in BIM, is adopted in this approach to describe, share, and exchange construction and facility management information. Developed by buildingSMART International, IFC serves as a neutral format, enabling interoperability between various software platforms in architecture, engineering, construction, and operations [33]. Different IFC standard versions have varied in their support for MEP systems. IFC2x3 provided only basic support for HVAC, while IFC4 significantly expanded and restructured this domain, introducing more detailed entity types and a wider range of property sets essential for BIM2BEM workflows. IFC4.3 offers further refinements, but as it is a recent release, widespread tool support is still pending. Therefore, this study focuses on IFC4 and its updates, like IFC4 ADD2 TC1 (ISO 16739-1:2018), which are widely supported by existing tools.

Given the IFC4 model, information is extracted from relevant entities within the `IfcHvacDomain`, typically including subtypes of `IfcDistributionFlowElement`. These encompass energy conversion devices (`IfcEnergyConversionDevice`), duct or pipe segments, fittings, terminals, flow-moving devices such as fans or pumps, and flow control elements like valves, dampers, and air terminal boxes.

Fig. 2 shows the data manipulation diagram of the GRC tool used in this work to check the geometric relationship between elements in BIM models. The IFC Geometry Exporter [34], a component integrated into the cloud platform, was used to derive geometric representations of the specified elements. It extracts the specific shape representation (`IfcShapeRepresentation`) from the IFC files and organizes the geometric information into a hierarchical tree-structured format in XML files. These generated XML files then serve as input for the GRC tool to detect geometric relationships between pairs of elements. With no prior information available on these relationships, the GRC tool accelerates the process by first applying Axis-Aligned Bounding Boxes (AABB) for an initial check, followed by a more detailed analysis. As AABB requires only a few multiplications and is significantly faster than full clash detection, it is feasible to evaluate all possible pairs, with a complexity of $O(n^2)$. As a result, an OBJ file and an XML file are outputted, encapsulating the geometric relationships needed for further analysis. The GRC tool is currently under development and not yet available for public use. Following sufficient beta testing, it will be released as part of a tool bundle, including a Graphical User Interface (GUI) that allows users to visually verify the detected geometric relationships.

2.1.2. Geometric relationship among entities

Although relationships `IfcRelConnectsPorts` and `IfcRelConnectsPortToElement` describe the connection between the elements in MEP files, some of these connections may not be explicitly modelled, and the relationship would be missing. This might not be important for many uses of the BIM model, but if there is a need to identify connectivity between flow segments, then this information is crucial. One approach is to ask the modeller to model these port connections, but this might not be practical or feasible. To overcome these issues, this work proposes a workflow to address these missing connections. This workflow uses three types of geometric relations that reveal other non-geometric connections between pairs of elements. These relations can be classified as adjacency, clash, and containment. Fig. 3 illustrates these geometric relations using conceptual and graphical representations, with more details provided below.

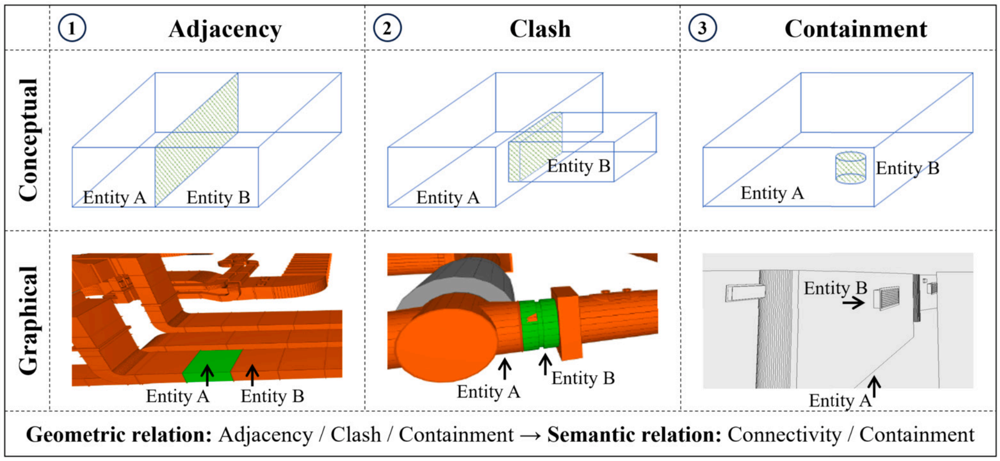


Fig. 3. Conceptual and graphical illustration of geometric relationships, including adjacency, clash, and containment.

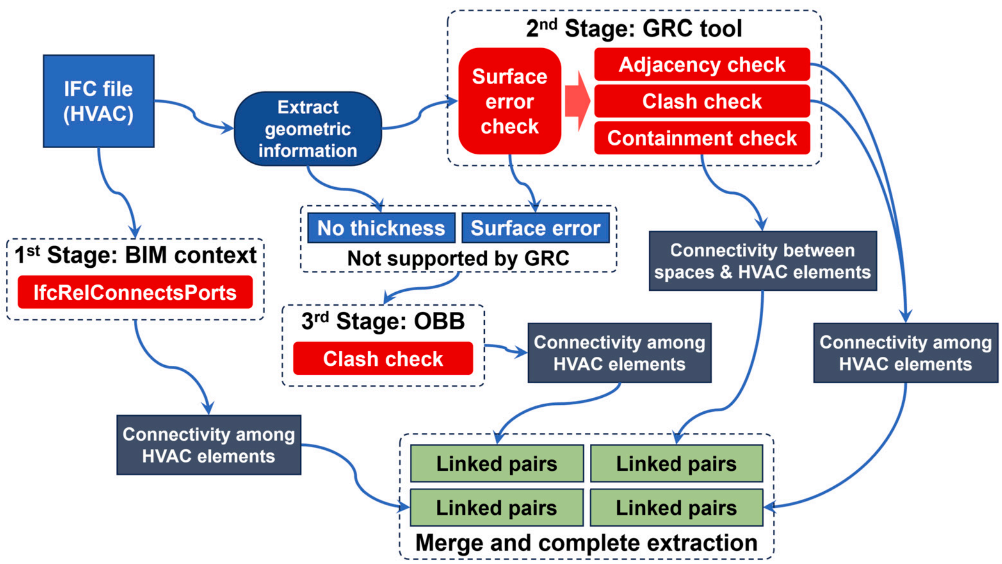


Fig. 4. The proposed geometric relationship check workflow with three stages.

Adjacency: An adjacency relation between two solids exists when the boundary representation of one solid has surfaces lying within the same plane and intersecting with surfaces of the other solid's boundary representation. Part 1 of Fig. 3 illustrates the adjacency between two elements. This adjacency occurs in pairs of ducts, especially rectangular-shaped ones, as illustrated in the bottom image of part 1 of Fig. 3. Unlike other geometric relations, adjacency detection uses a threshold value to determine the allowable distance between the intersecting surface planes. The adjacency will not be detected if the distance between these planes exceeds this threshold value. The adjacency relation is used to identify connectivity between MEP elements.

Clash: A clash relation between two solids exists when surfaces or portions of surfaces of the boundary representation of one solid are fully contained within the solid volume of the other. Part 2 of Fig. 3, illustrates the clash relation between two elements. This relation is detected among MEP element pairs and space volumes, such as between spaces and terminal units, terminal units and ducts/pipes, ducts and ducts, pipes and pipes, and ducts/pipes and systems. Clash occurrences offer valuable insights related to MEP element connectivity, especially when these connectivity relations in the BIM context are missing.

Containment: A containment relation between two solids occurs when the surfaces of one solid's boundary representation are entirely within the volume of the other. This relation establishes links between spatial entities and terminal units, precisely positioning HVAC terminal units within a building's spatial structure. This relationship ensures that air terminal units, like diffusers, are fully contained within the building space, as illustrated in part 3 of Fig. 3.

As analysed next, the above relations induce non-geometric semantic relations among MEP and building space entities.

2.1.3. Geometric relationship check workflow

This workflow is designed to discover all current or potential connections among the MEP entities. It comprises three stages, displayed in Fig. 4. First, this workflow identifies all explicitly modelled connections (designated by `IfcRelConnectsPorts` and `IfcRelConnectsPortToElement`) and compiles a list of connected element pairs. Subsequently, the GRC tool is utilized to uncover additional geometric relationships among these element pairs. The GRC tool can also flag elements with geometric surface errors, which cannot pass through certain geometric relation checks, thus directing them to the Oriented Bounding Boxes (OBB) [35] checking process in the third stage. Additionally, elements with planar geometric representations, such as duct end caps, which lack volume and cannot be handled by the GRC tool, are also routed to the third stage. Ultimately, all linked element pairs from each stage are merged into a consolidated list of linked pairs, maximizing the extraction of connections from MEP-IFC files.

2.2. Stage 2: knowledge graph generation

This stage aims to create an instance of a knowledge graph of HVAC components and related topological information using available information from the BIM model. The proposed approach uses available modelled information and geometric relationship checking to be resilient to potential modelling issues and errors in the BIM model. The resulting knowledge graph captures the HVAC components and their connectivity and can also be useful for the BIM2BEM framework.

2.2.1. Ontology and geometry-induced knowledge graph

Using multiple ontologies concurrently to represent building information is generally more efficient than creating a customised ontology for each project or designing a single comprehensive ontology to encompass all knowledge domains. Integrating multiple ontologies enables the development of a robust digital representation, potentially incorporating one or more knowledge graphs to accurately reflect the building's information [36,37]. Specifically, in this study, Brick Ontology, developed by the Brick Consortium, is introduced to represent HVAC components, building spaces, and representation of HVAC systems' topology. Brick provides a consistent vocabulary to describe all MEP-related and architectural elements and their interconnections, facilitating the integration of diverse data sources and enabling interoperability across different systems. Additionally, an ETL tool, namely Knowledge Graph Generator [38], was employed to facilitate the transformation of HVAC elements from IFC files into corresponding entities within the knowledge graph.

Within the Brick framework, the `brick:HVAC` abstract class provides semantic descriptions for various components of HVAC systems, covering a wide range of elements from subsystems (e.g., `brick:AHU`, `brick:FCU`, and `brick:Chiller`) to terminal units (`brick:Air_Diffuser`), pumps (`brick:Water_Pump`), VAV boxes (`brick:VAV`), radiators (`brick:Radiator`) among others. In addition to the HVAC class, `brick:Space` and `brick:Room` classes are essential for describing the spaces and rooms within buildings. Since Brick lacks a fully corresponding relationship to precisely represent the direction-agnostic geometric connections obtained from the previous subsection, `brick:isAssociatedWith` (other custom relationships are also applicable) is utilised to enrich the geometry-induced knowledge graph that will subsequently be transformed into an HVAC topology. To avoid name duplication, the IFC global unique identifier (GUID) can be transferred into the knowledge graph, either to name the entity or to be stored as a property of the entity.

2.2.2. Error detection and knowledge graph generation

While the geometric relation-checking workflow can extract most connections, some may still be absent due to the data quality of the inputted IFC file. Addressing these issues in this stage is crucial to ensure the accuracy of the generated HVAC topology in subsequent stages. This study proposes a set of rules to identify potential issues in establishing connections between entities, ensuring the completeness of knowledge graphs. These rules are designed to cover most entities within the HVAC domain. For instance, an error will be flagged if a pipe or duct segment entity has fewer than two linked entities. This ruleset is detailed in Table 1.

Based on these rules, entities with issues are identified and displayed in a BIM-viewing platform for verification purposes. Users can then manually address the problematic entities by supplementing missing connections or removing non-existent ones in the graph, thereby enabling the creation of a complete graph.

Table 1
Geometry-induced knowledge graph issue ruleset.

Item A	Relation/Property	Item B
Pipe segment	< 2 linked entities	-
Duct segment	< 2 linked entities	-
Duct fitting	< 2 linked entities	-
Pipe fitting	< 2 linked entities	-
Tee pipe	< 3 linked entities	-
Pipe segment	linked to	Duct segment
Duct segment	linked to	Pipe segment
Air diffuser	not linked to	Duct segment
Air Handling Unit (AHU)	not linked to	Non-space entity
Fan Coil Unit (FCU)	not linked to	Pipe segment
Air diffuser	not linked to	Space entity
Air diffuser	linked to	> 1 space entities
Terminal unit	not linked to	Space entity
Terminal unit	linked to	> 1 space entities

2.3. Stage 3: BEM-oriented HVAC topology establishment

This stage establishes a BEM-oriented HVAC topology from the HVAC system's geometry-induced knowledge graph. Obtaining the informative HVAC topology is essential initially, as detailed in the subsequent subsection.

2.3.1. Informative HVAC topology generation

Path-finding algorithms, e.g., depth-first search, can achieve the logical link between any two entities via specific relationship (predicate) checks, offering greater flexibility and efficiency than simple queries when handling customized constraints in complex graphs [39]. Using this capability, these methods can be introduced to establish logical links between air terminals and systems (including FCUs, AHUs, and MVHRs), radiators/FCUs/AHUs and pumps/Heat Interface Units (HIUs), and other critical, logical links between different HVAC equipment. Additionally, the Variable Air Volume (VAV) box can be identified in the path-finding process, and some entities can be designated as inaccessible to avoid achieving unreasonable links. It is better to apply the path-finding process within the water- and air-loops by splitting the generated HVAC graph into separate water- and air-loop sub-graphs. In this way, a significant reduction of computing resource requirements can be achieved.

Initially, the `brick:isAssociatedWith` relation is used to represent the detected geometric relations of the second stage that refer to semantic connections of HVAC elements. Regarding space entities, the `brick:isLocationOf` is introduced to indicate where the assets (e.g., terminal units and systems) are situated. To add the notion of direction, the `brick:feeds` is used to indicate the directional links between AHUs/MVHRs/FCUs and air terminals in air loops. In water loops, the same relation can be used to describe the direction of the water supply. In the case study of this work, the airflow direction is determined according to the air terminals' asset labels, in which there is a special string to indicate whether the label refers to an air supply terminal or an air return terminal. The schematic diagram in Fig. 5, illustrates how the informative HVAC topology is generated from the geometry-induced knowledge graph.

The outlined approach proves effective in generating an informative HVAC topology. However, essential data for building energy modelling, including capacity, efficiency, and control strategy, may be lacking in MEP-BIM models below Level Of Development 400 (LOD400). To address this, the study establishes an equipment inventory based on manuals provided by the estate team, encompassing all assets in the HVAC topology. This inventory is then integrated to enrich the topology with equipment performance data. Additionally, room temperature control strategies can be obtained from building management system operation manuals supplied by the estate team.

Finally, ensuring the correctness and completeness of the generated informative HVAC topology is essential. Considering that a few subtle logical errors might originate from the BIM inference, this work introduces an additional data quality checking rule set of the generated informative HVAC topology. Potential issues are displayed in the following Table 2. Given that MVHR is not explicitly included in the Brick Ontology, this work utilises the Dedicated Outdoor Air System (DOAS) class as a representation. EnergyPlus also offers a module for modelling DOAS, which can be customised to closely resemble MVHR.

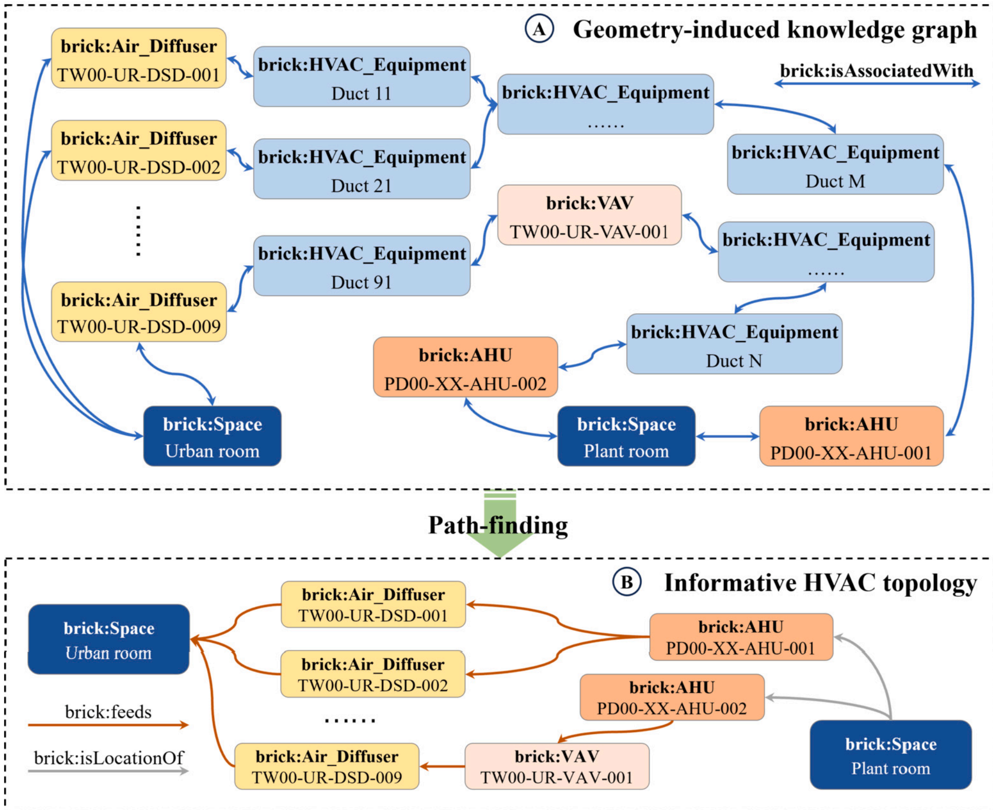


Fig. 5. Schematic diagram of generating informative HVAC topology based on the path-finding process.

Table 2
Informative HVAC topology issue ruleset.

Item A	Relation/Property	Item B
Air terminal	> 1 linked entities	VAV/AHU/DOAS/FCU
Air terminal	not linked entities	-
FCU	not linked to	Water pump entity
AHU	not linked to	Water pump entity
Radiator	not linked to	Hot water pump entity
Water pump	not linked to	Supply-side entity

The above rules are applied to detect errors in the informative HVAC topology only. Although the path-finding process can also infer some missing links between the above entities, it would be preferable to fix the errors manually through validation against the design drawings or other data sources.

2.3.2. BEM-oriented HVAC topology establishment

Given the introduction of a newly constructed building equipped with complex HVAC systems featuring diverse energy technologies, EnergyPlus has been chosen as the platform for building performance simulation. To align with EnergyPlus's programming syntax, a BEM-oriented HVAC topology is developed by simplifying the detailed HVAC topology, placing emphasis on zone- and space-level granularity. Thus, this study presents a straightforward and effective method for transforming the informative HVAC topology into its BEM-oriented counterpart, as illustrated in Fig. 6. The method can be summarized into four parts, outlined as follows:

Space and FCU/VAV bridging: In the informative HVAC topology, VAV boxes and FCUs are not directly connected to individual spaces; instead, the connection is made through intermediary components, such as air terminals (e.g., diffusers and grilles). To adapt these connections into usable links within a zone-level BEM-oriented HVAC topology, certain scenarios require the creation of direct links between air terminals and VAV/FCU, as depicted in Part A of Fig. 6. However, in some cases, air terminals cannot be bypassed, especially when they are directly linked to AHUs or MVHRs, or when they are interconnected with FCUs and MVHRs in series (this will be addressed by a decoupling method described below).

Terminal units merging at space level: Generally, terminal units within a single space are typically supplied by the same system according to design conventions. Additionally, for economic reasons and air distribution in ductwork, the terminal units in the same space usually have similar performance [40]. As a result, it is practical to consolidate similar units into a single hypothetical terminal unit with a capacity equal to the combined capacities of the individual units. It is important to note that this approach can only be applied to air terminals serving the same space and supplied by the same source (AHU/MVHR). This method not only complies with the modelling rules of widely used pre-processing tools such as Eppy (via HVACTemplate) and OpenStudio but is also well-suited for modelling a large building equipped with hundreds of diffusers. The schematic diagram is depicted in Part B of Fig. 6.

Hot/chilled water-loop construction: Given the simulation platform's dictionary, it's essential to introduce a distinct entity for representing water loops. Within the introduced overall framework, all equipment associated with a specific water loop, such as pumps, buffer vessels, valves, and degassing units, are replaced by a generated `brick:Water_Loop` class entity. Subsequently, connections are established between these water-loop entities and terminal units or other system entities, e.g., AHUs. This process is illustrated in Part C of Fig. 6.

FCU and MVHR decoupling: As the connections between FCU and MVHR are not directly predefined in the simulation platform's dictionary, this study introduces a decoupling technique to transform the FCU-MVHR connections from series to parallel feeding the same space. The approach involves relocating the air mixing point from the front of the FCU to the back. Part D of Fig. 6 illustrates an example of decoupling the merged FCU-MVHR terminal units.

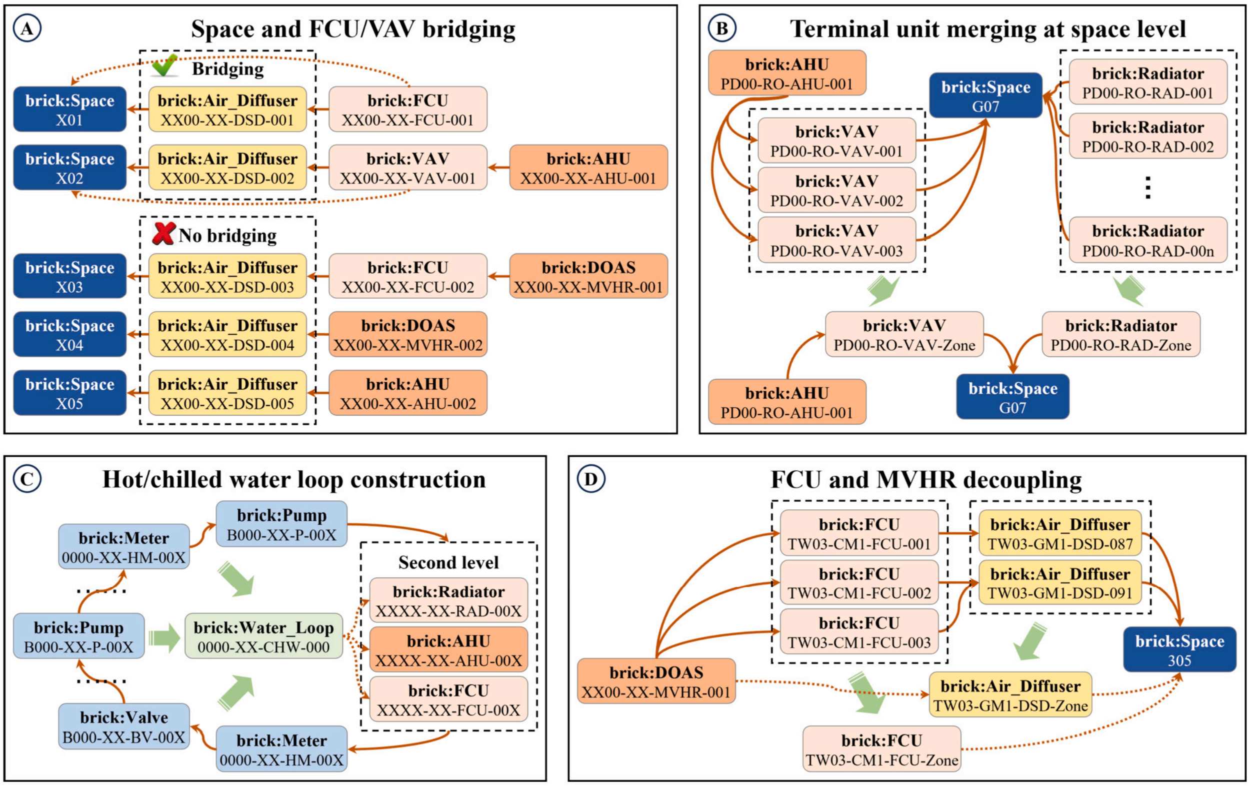


Fig. 6. Schematic diagram of the methods to transform informative HVAC topology into BEM-oriented HVAC topology.

2.4. Stage 4: building energy modelling

This study presents an automated framework for creating a BEM model by integrating data from BIM and other sources into an HVAC topology. The simulations are conducted using EnergyPlus, which requires an IDF file (BEM model) containing comprehensive descriptions of building geometry, material properties, equipment details, system definitions, climatic conditions, and other simulation parameters. Through the simulation process, a thorough evaluation of building performance regarding energy usage, indoor comfort levels, and environmental effects can be attained.

Regarding the IDF file structure, the BEM model can be segmented into two components: the geometrical module and the non-geometrical module. The geometrical module within IDF files encompasses the physical characteristics and layout of the building. It includes detailed specifications regarding building geometry, such as building envelopes and their thermal properties. Additionally, internal partitions, zones, and their connections are defined to accurately represent the spatial configuration and thermal zones of the building. The geometric BEM model can only be generated from the architectural BIM model if the BIM model provides a high level of detail regarding building geometry, openings, and construction, as demonstrated in our previous work [16,34]. However, due to data limitations, only a simplified BIM model with space entities is accessible in this work, which cannot support the generation of a complete geometric model for EnergyPlus simulation. Therefore, given this constraint and the research focus, a geometric BEM model in IDF format, without active system settings, was manually created based on architectural drawings, as detailed in Section 3. The HVAC topology was then used to enrich the geometric BEM model with detailed HVAC information, resulting in a simulation-ready BEM model for EnergyPlus simulation.

The HVAC topology and IDF editing processes are seamlessly integrated into a single coding platform by leveraging a Python environment with RDFLib and Eppy. The procedure can be outlined as follows: Each space entity is correlated with a corresponding thermal zone within the Zone class of EnergyPlus. The attributes of space entities related to internal heat gains are transferred to the appropriate fields within the People, Lights, and ElectricEquipment classes. Similarly, properties associated with control strategies are linked to the field within the ZoneControl:Thermostat class. Additional data, such as hourly schedules, are extracted from several CSV files, with the properties providing only their indexing information. Moreover, there are three approaches to translating HVAC-related information from the topology into fields in the IDF file. The first approach involves using the classes HVACTemplate:Zone, HVACTemplate:System, and HVACTemplate:Plant to assign fields by parsing the entities and in the generated HVAC topology, offering convenience and error-proofing. The second approach entails introducing the class ZoneHVAC and other device-level HVAC-related classes, demanding comprehensive and highly detailed information for each HVAC component. The third approach, which can be considered as the combination of them, utilizes the HVACTemplate classes to establish the structure for HVAC-related classes, generating an initial IDF file and then refining equipment details based on properties of the obtained HVAC topology. In this study, the third approach was adopted due to the complexity of the HVAC system and the completeness of metadata provided by the estate team. The schematic diagram illustrating the mapping from the HVAC topology to the BEM model can be found in Fig. 7. Finally, the time cost and simulation results, in terms of space's indoor temperatures and energy use, are both used to validate the proposed framework's performance.

3. Case study

This section delves into the case study across four dimensions: providing a building overview, detailing data preprocessing of BIM files, defining additional parametric values for building performance simulation, and describing the modelling environment.

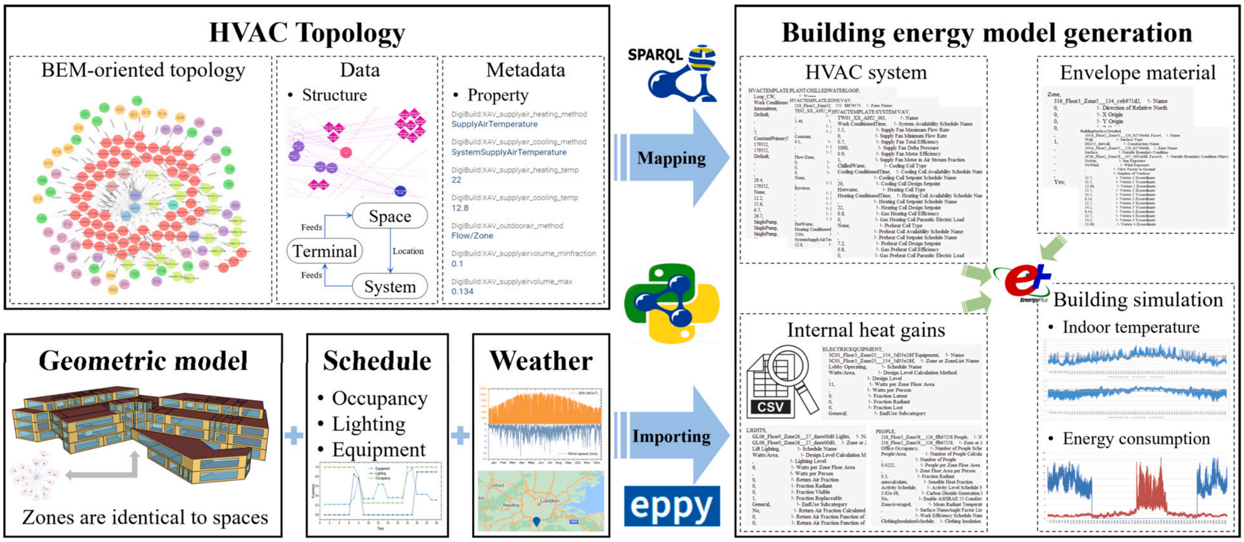


Fig. 7. Schematic diagram of the mapping from HVAC topology to BEM model.

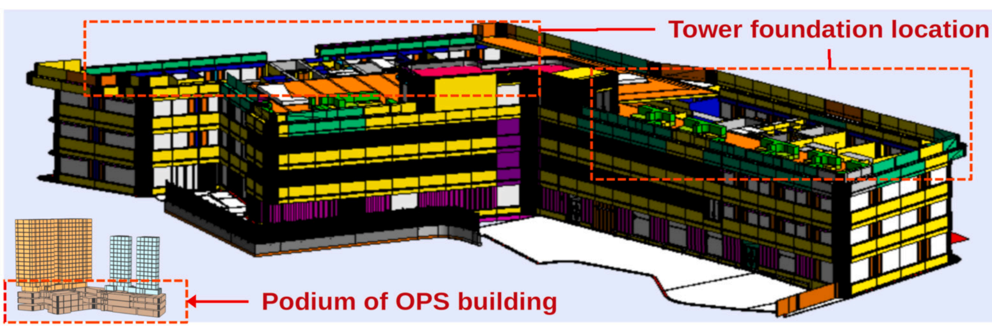


Fig. 8. 3D visualization of the OPS podium.

3.1. Overview of OPS building

The proposed framework was implemented in a newly constructed high-rise building, namely One Pool Street (OPS), situated within the UCL East Campus. OPS serves as a multi-functional building equipped with a state-of-the-art HVAC system with diverse energy technologies, which is monitored and controlled by a Building Management System (BMS). For cooling purposes, an air-cooled chiller is installed to provide chilled water for AHUs and FCUs to respond to space cooling requirements. For heating purposes, the building is connected to a district heating system to supply the heating coils in AHUs, FCUs, and Radiators. Many MVHRs exist to enhance ventilation and improve energy performance via recycling waste heat in several areas. All the aforementioned energy technologies are implemented in the podium section of this building, which is the basis where two towers (east and west) are founded. The podium has common areas from the ground floor to the second floor, as shown in Fig. 8, where an Autodesk Navishworks model is displayed. The areas where the two towers are based are shown as “holes” in the top slab in this figure.

The west and east towers, designated for residential purposes, do not require cooling and only have radiators for heating purposes. Therefore, compared to the towers, manually constructing the building energy model for such a complex energy system in the podium area would be arduous, extremely time-consuming, and even susceptible to mistakes, which makes it an ideal scenario for assessing the effectiveness of the proposed framework with a focus on HVAC information. Meanwhile, the informative HVAC topology would be highly beneficial for building management services as it effectively handles heterogeneous data.

3.2. Data preprocessing

A simple and optimal scenario would involve inputting a single MEP-IFC file into the geometric relationship check workflow to establish the HVAC topology. However, due to the large size and complexity of the geometric data at LOD 300-400 granularity, there’s a memory issue when running GRC as the number of pair checks for N elements grows polynomially $O(n^2)$. For N elements, the number of pair checks is $\frac{N(N-1)}{2}$ and for $N \approx 20.000$ the number of pair checks approaches 200 million pairs. Thus, it becomes necessary to partition the entire MEP-IFC file into two smaller ones: one containing air-loop domain data and one containing water-loop domain data. This approach optimally reduces the size of IFC files, minimizing the need for manual intervention. Fig. 9 illustrates the raw data, comprising the ARC-BIM model (Part 1) and MEP-BIM model (Part 2) in Revit, as well as the exported IFC files from the Revit files for the proposed framework (parts A, B and C). Additionally, alongside the BIM models, schedules for lighting, occupancy, electrical equipment, and control strategies are stored in a few CSV files, facilitating indexing during the parsing of HVAC topology.

As shown in Fig. 9, the available ARC-BIM model is relatively basic, containing only space entities and lacking sufficient details on openings and construction to generate a complete geometric BEM model. Therefore, in this work, the geometric module of the IDF file was manually created using Grasshopper-Honeybee, based on the detailed design drawings. Subsequently, the initial IDF file was automatically enriched with comprehensive and detailed HVAC settings, using the proposed framework and the corresponding generated HVAC topology to make it ready for simulation.

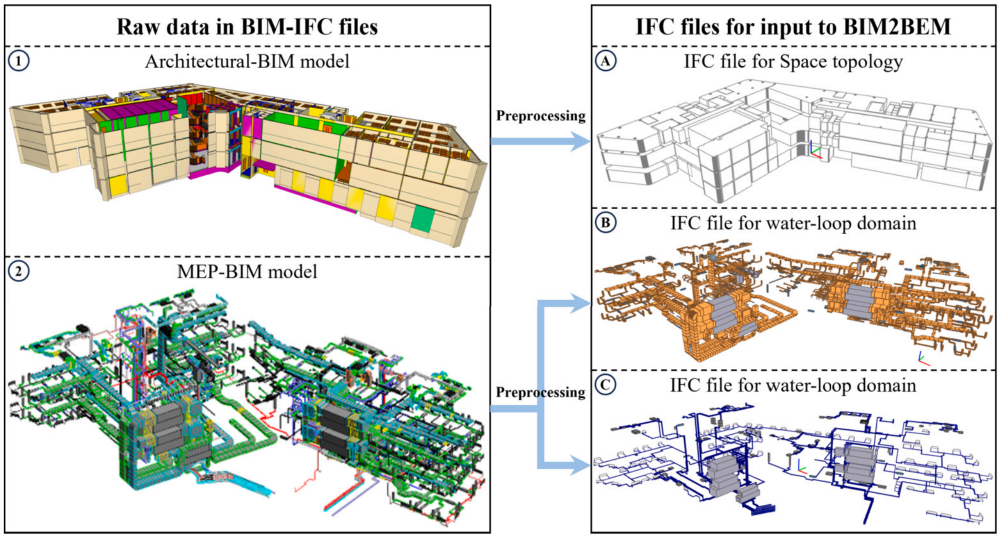


Fig. 9. 3D visualization of BIM input data and the exported IFC files for use in the BIM2BEM framework.

Table 3
Parametric setting for building envelope and operation.

Domain	Item	Value	Data source
Envelope	Roof External wall Floor Window	0.18 W/(m^2 K) 0.26 W/(m^2 K) 0.18 W/(m^2 K) 1.6 W/(m^2 K)	Building regulation [41]
Operation	Primary function rooms Secondary function rooms	21 + -1 °C for room temperature at daytime 18 °C for supply air temperature at daytime	BMS manual

Table 4
Parametric setting for internal heat gains of several main spaces [42,43].

Item	Unit	Office	Auditorium	Lobby	Lab	Toilet
Lighting	W/m^2	7.5	5	5	12.5	5
People	Person/m^2	0.11	0.34	0.12	0.10	0.11
Equipment	W/m^2	11.99	1.78	5.27	8.73	4.57

3.3. Additional parametric setting

Due to data limitations, several parameters critical for building performance simulation remain unavailable in addition to BIM data and device manuals. Specifically, detailed information on construction and electrical systems is not accessible or integrated into the BIM models, resulting in missing data on the thermal properties of the building envelope and internal heat gains, which are required for performance simulation. To address this issue, building regulations most closely aligned with the construction characteristics and age of the new building were selected. This approach ensures that the building envelope parameters presented are as realistic as possible, given the current data availability. Tables 3 and 4 outline the key parameters required for building performance simulation and their respective sources. Further collection of actual construction data and detailed interior design parameters could facilitate updates to the existing parametric settings, thereby improving the reliability of building performance simulations. However, given the current data availability for this case study, only floor plans and a simplified ARC-BIM model are accessible, while the HVAC data is comparatively more detailed and comprehensive. As a result, in this study, HVAC information is directly sourced from the actual BIM files and design manuals, while certain building-related parameters have been reasonably simplified and, where necessary, assumed.

3.4. Modelling environment

The entire framework was executed on a standard laptop equipped with a 13th Gen Intel(R) Core(TM) i7 processor and 32 GB of RAM. Data preprocessing of BIM models was carried out in Revit 2018 to export the IFC files, while the BIM2BEM framework was developed within the Python environment, utilizing RDFLib for parsing topology and Eppy for BEM generation. Additionally, building performance simulation was performed by invoking the EnergyPlus 9.5 simulation engine.

4. Results and discussions

The proposed framework was conducted as mentioned in the above sections, and this section elaborates upon the results and discussions in three aspects, including the generated HVAC topology, framework performance, as well as limitations and future work.

4.1. Knowledge graph and HVAC topology

Although the whole MEP-IFC file and ARC-IFC file (space topology) can be imported into the geometric relationship checking workflow directly, the out-of-memory issue can not be avoided if a regular laptop is used. Thus, after the data preprocessing, three IFC files regarding space, air-loop subsystem, and water-loop subsystem were generated for geometric relation check. The conducted pair checks between terminal units/AHUs/Radiators/FCUs and spaces were approximately 115,000, which took several minutes. The conducted pair checks among the elements in the air-loop subsystem and the water-loop were approximately 38,000,000 pairs and 101,000,000 pairs, respectively. Moreover, the path-finding process based on depth-first search algorithms was implemented to establish the structure of the informative HVAC topology.

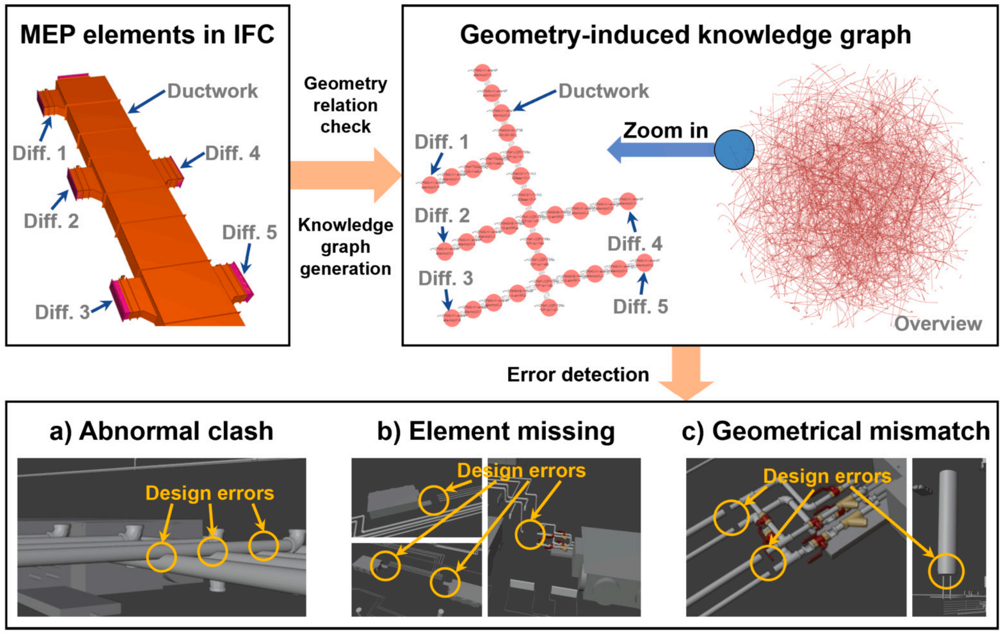


Fig. 10. Geometry-induced HVAC knowledge graph and error detection associated with a) abnormal clash, b) element missing, and c) geometric mismatch.

4.1.1. *Geometry-induced knowledge graph*

Initially, a geometry-induced knowledge graph of the HVAC system was established to directly reflect the geometric relationship among all elements in the inputted IFC files. Fig. 10 provides an overview of the geometry-induced knowledge graph alongside a detailed view of a set of ducts and their digital counterparts within the graph. The graph appears non-structural because there are no discernible logical relationships among these elements. Although this graph may not be directly applicable to generating a BEM model, it offers valuable insights for directly detecting errors in the inputted IFC files using the ruleset outlined in Section 2.2.2. Fig. 10 also shows some of the detected issues by human errors or extraction from the federated BIM model, which can be summarized into three types, i.e., a) abnormal clash, b) missing element, and c) geometric mismatch. Abnormal clashes typically occur in areas with dense element distribution or highly complex pipeline configurations, while missing elements are often found within intricate pipeline networks. Geometric mismatches are irregularly and subtly distributed across various areas. The error detection results underscore the importance of the proposed GRC-based workflow in improving BIM data quality and ensuring the completeness of the HVAC graph. However, despite the accuracy of the error detection in identifying issues, manual intervention for error correction—such as adding or removing links in the geometry-induced knowledge graph—remains necessary.

Furthermore, depth-first search algorithms were employed in the path-finding process to establish logical connections between air-side terminals and air-side systems (e.g., AHUs, MVHRs, and FCUs), as well as between water-side terminals (e.g., AHUs, radiators, and FCUs) and water pumps. These processes can also indirectly detect errors and support the generation of an informative HVAC topology. If a path cannot be found between a particular air terminal and an air-side system, it can be inferred that there may be undetected errors that were not identified by the ruleset. Similarly, if an air-side terminal entity is linked to more than one air-side system, or if a pipe entity is connected to both Chilled Water (CHW) pumps and Low-Temperature Hot Water (LTHW) pumps simultaneously, a double-check of the corresponding BIM elements is required.

4.1.2. *Informative HVAC topology*

Fig. 11 illustrates the overview of the generated informative HVAC topology and several critical air-supply subsystems. In this case, there are four major air supply mechanisms around the building, which include: a) AHU2VAV2Diffuser2Zone, b) MVHR2FCU2Diffuser2Zone, c) AHU2Diffuser2Zone, and d) MVHR2Diffuser2Zone. The device performance information was stored as properties of each entity, as shown in Part e) of Fig. 11. Based on the informative HVAC topology, it can be inferred that, among the nine AHUs, four of them are Variable Air Volume (VAV) systems that supply fresh air to zones via VAV boxes, serving the general areas, such as workspace, meeting rooms, and learning hubs. The remaining five are Constant Air Volume (CAV) systems, which serve an auditorium, toilets, corridors, cupboards, and other ancillary areas. Additionally, six of the AHUs are linked with both the LTHW and CHW pumps for heating and cooling purposes, while three of them are not linked to the CHW pumps as they service the toilets and showers. Moreover, nine MVHR-based ventilation systems for heat recovery purposes are identified. Of these, six are in series with FCUs to serve the common rooms and teaching spaces, while the remaining three isolated systems directly serve laundry rooms and music rooms. Overall, approximately 44% of the zones are served by AHUs and 9% by MVHRs. In the other zones, only FCUs, Radiators, and door heaters are installed to meet the heating or cooling needs of spaces that do not require fresh air.

The generated informative HVAC topology embodies significant benefits in comprehending HVAC systems, particularly for the complex systems in large-volume buildings, through systematically mapping out the logical connections among air-side and water-side components. Each component is documented with detailed properties, ensuring access to critical information about device performance and control methods. A notable feature of the informative HVAC topology is its compatibility with query languages, allowing for efficient retrieval of specific information to troubleshoot and facilitate decision-making. Moreover, this topology enables seamless cross-referencing and validation with schematic drawings of HVAC designs, which can not only mitigate the likelihood of errors or omissions but also facilitate integration with BMS and building performance simulation. Finally, a double-check process against HVAC schematic drawings was implemented to ensure the generated topology was entirely accurate.

4.1.3. *BEM-oriented HVAC topology*

Considering the feasibility and complexity of the proposed framework, a BEM-oriented HVAC topology focusing on thermal zones was established through a transformation process from the informative HVAC topology. With each transformation, the topology becomes less complex

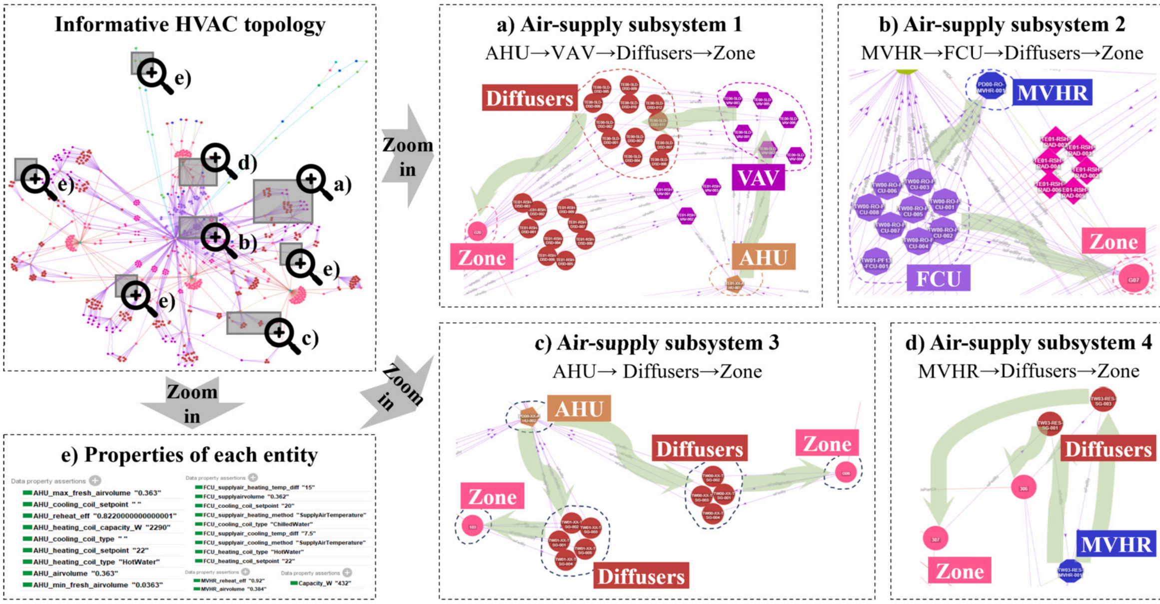
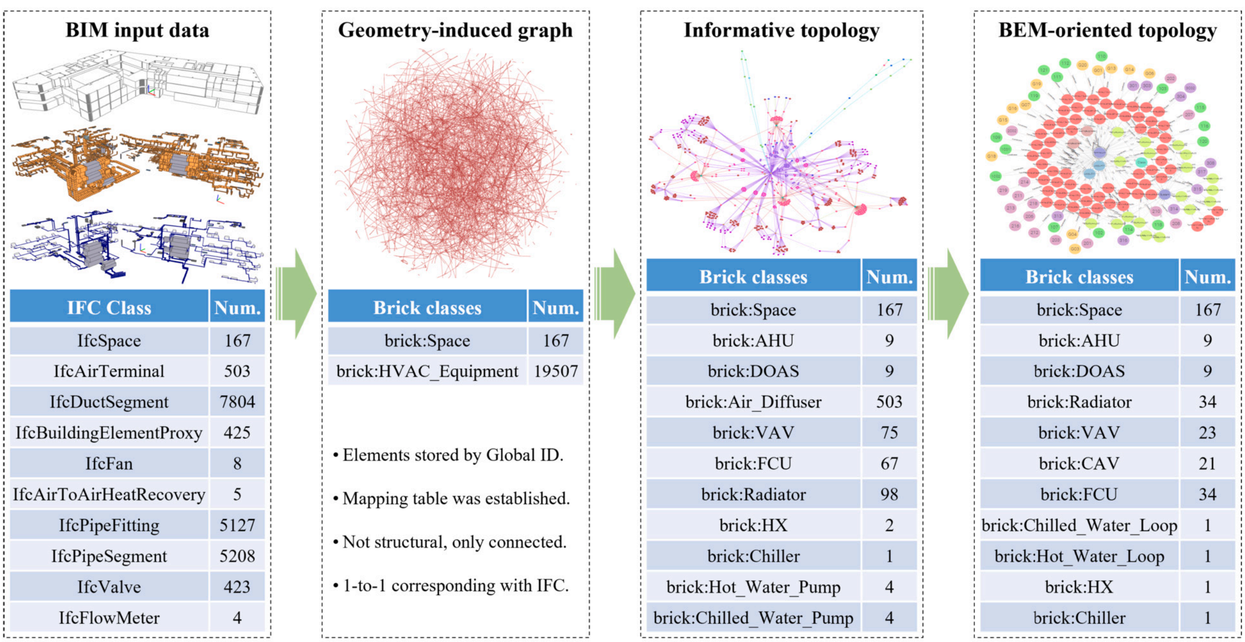


Fig. 11. Informative HVAC topology with four air-supply subsystems (a–d) within the HVAC system, along with an example of entities' properties (e).



IFC Class	Num.
IfcSpace	167
IfcAirTerminal	503
IfcDuctSegment	7804
IfcBuildingElementProxy	425
IfcFan	8
IfcAirToAirHeatRecovery	5
IfcPipeFitting	5127
IfcPipeSegment	5208
IfcValve	423
IfcFlowMeter	4

Brick classes	Num.
brick:Space	167
brick:HVAC_Equipment	19507

Brick classes	Num.
brick:Space	167
brick:AHU	9
brick:DOAS	9
brick:Air_Diffuser	503
brick:VAV	75
brick:FCU	67
brick:Radiator	98
brick:HX	2
brick:Chiller	1
brick:Hot_Water_Pump	4
brick:Chilled_Water_Pump	4

Brick classes	Num.
brick:Space	167
brick:AHU	9
brick:DOAS	9
brick:Radiator	34
brick:VAV	23
brick:CAV	21
brick:FCU	34
brick:Chilled_Water_Loop	1
brick:Hot_Water_Loop	1
brick:HX	1
brick:Chiller	1

Fig. 12. Overview diagrams and entity inventories for BIM data, geometry-induced knowledge graph, informative HVAC topology, and BEM-oriented HVAC topology.

while gaining structural clarity. Ultimately, the BEM-oriented HVAC topology was designed to adhere to the modelling language rules of EnergyPlus. One key point is that, for each space, a single hypothetical terminal entity was created for each type, representing all terminal units of that type, with a capacity equal to the combined capacities of the individual units. Fig. 12 illustrates the entity inventories at each stage, from the input IFC files to the BEM-oriented topology. As shown, the number of entities in the BEM-oriented topology is approximately 300, significantly fewer than those in both the geometry-induced knowledge graph and the informative HVAC topology.

4.2. Framework performance

The proposed BIM2BEM framework aims to generate BEM models automatically, accurately and promptly. The following subsections elaborate on two critical aspects, including time cost and building performance simulation.

4.2.1. Time cost for BIM2BEM

This research endeavours to achieve automation in BIM2BEM generation, focusing on complex HVAC systems with as little manual intervention as possible. There are three relatively time-consuming parts within this framework, including geometric relation checking, HVAC topology generation, and building performance simulation. Regarding the geometric relation check process within the HVAC system, it took approximately 4 hours and 0.5 hours for the air-loop and water-loop IFC files, respectively, while the checking process between spaces and terminals required only several minutes. Hence, generating a geometry-induced HVAC graph combined with IFC parsing and geometric relation checking took nearly 5 hours, with errors also identified. Then, the manual

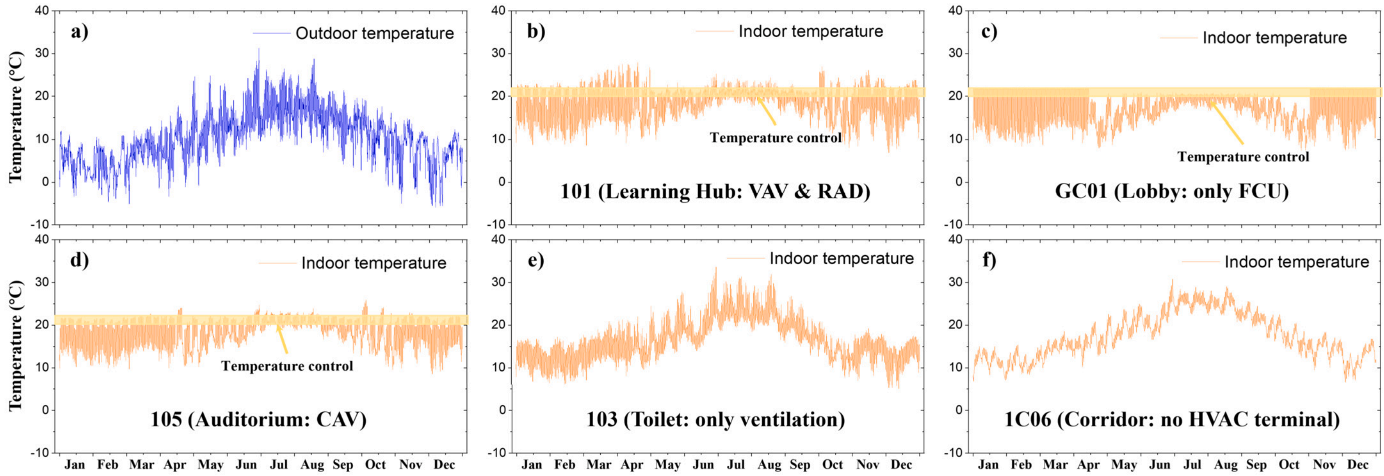


Fig. 13. Whole-year outdoor temperature and indoor temperatures of several typical rooms.

work for fixing errors took 2 hours, which involved adding and deleting the corresponding links between HVAC entities in the graph. Subsequently, the establishment of the informative HVAC topology based on the path-finding process and the simplification to BEM-oriented topology took approximately 2.5 hours and half an hour, respectively. Finally, the BEM generation and execution, which involved refining the IDF file and conducting a whole-year building performance simulation, were completed in under 20 minutes within the Python environment by invoking EnergyPlus as the simulation engine. In summary, automatically achieving the generation of a building energy model (with only a few manual interventions required to fix errors) via the proposed framework required a time investment of fewer than 10 hours.

According to previous references [44,45] on manual modelling times, the process of modelling complex buildings with detailed HVAC configurations typically takes several days or even longer. The time required for manual modelling is strongly influenced by the designer's expertise with the platform and knowledge of the building's systems. The proposed framework, therefore, offers a significant reduction in modelling time, bringing it down to just several hours, even when accounting for errors in the BIM data. This underscores the framework's efficiency and its potential to support and streamline the BIM2BEM process, particularly in relation to HVAC systems, while minimising the need for extensive manual intervention and the likelihood of errors. The effectiveness of the framework was demonstrated in this case study involving a building with a complex HVAC system, which included over 160 spaces, more than 600 terminals, nearly 20 AHUs and MVHRs, and over 60 FCUs.

Although most stages within this framework are automated, data preparation remains a manual task, especially when the data source format, e.g., schedule, weather, and device properties, does not match the input format of the framework. Moreover, a high-quality MEP-BIM model above LOD300 is essential, as lower data quality may lead to errors in generating HVAC topology, such as incorrect or missing links among elements. This significantly amplifies the manual effort needed to rectify errors, thus reducing the efficiency of this framework.

4.2.2. Building performance simulation

This study focuses on two critical indicators, namely indoor temperature and energy use, concerning the building performance simulation of the generated BEM model for a newly constructed building with a total floor area of 8,800 sq.m (comprising 5,300 sq.m of conditioned area and 3,500 sq.m of unconditioned area). The simulated indoor temperatures are to verify the ability of the designed HVAC system to meet the heating/cooling requirements of the building, while the energy use indicates can serve as a reference for evaluating the performance of the designed HVAC system.

Fig. 13 shows the outdoor temperature in London and the indoor temperature of a few typical rooms with their No., function, and HVAC terminal type, including a toilet, a lobby, a corridor, a learning hub, a corridor, and an auditorium. According to the HVAC system manuals, the control strategy was set out to maintain the indoor temperature of 21 +/- 1 °C regarding the spaces supplied by FCUs and AHUs. However, regarding toilets and other ancillary spaces, the control strategy was set out to maintain a supply-air temperature of 18 °C +/- 1 °C. Meanwhile, the whole year was divided into heating, cooling, and transition seasons, while the HVAC system was only in operation for the daytime. The simulation results of indoor temperature can provide a clear picture to designers or managers. In parts b), c) and d) of Fig. 13, it can be seen that the three typical types of terminal units can both maintain the indoor temperature of their respective rooms within the expected range, i.e., 20 °C - 22 °C during the daytime. Thanks to the excellent thermal properties of this newly constructed building, the indoor temperatures do not drop below 10 °C during most of the night in winter. Moreover, regarding the toilet and corridor without heating/cooling terminals, as shown in parts e) and f) of Fig. 13, the indoor temperatures fluctuate significantly even though the ventilation terminals are installed in the toilets. Nonetheless, since the model does not consider door-opening actions, it's possible that the real indoor temperatures in unconditioned rooms might outperform the simulated results. However, maintaining precise temperature control in secondary function rooms/spaces is not as critical as it is in primary function rooms. In this scenario, the existing HVAC design scheme successfully attains the targeted temperature control range, as evidenced by the results derived from the proposed framework.

This work introduced Energy Use Intensity (EUI) to evaluate the energy performance of the building and its HVAC system, whose equations can be found in Equation (1).

$$\mathrm{EUI} = \frac{\sum_{h=1}^{8760} (\mathrm{Load}_{\mathrm{h}}^{\text{electricity, district heating}} \times 1)}{\mathrm{Floorage}} \tag{1}$$

where Load indicates the time-series energy load of the building, and Floorage is the total floorage of the building, including conditioned and unconditioned areas. The number 8760 is the total number of hours in a 365-day year.

Based on the building energy simulation, the case study demonstrated good performance of the newly constructed OPS building, with an EUI (total electricity) of 86.67 $\mathrm{kWh/m^2 \cdot year}$. Since the electricity

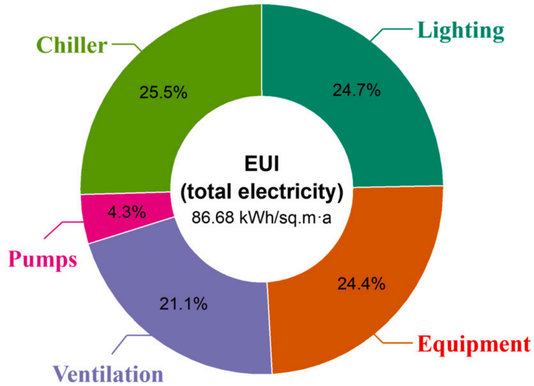


Fig. 14. Breakdown analysis of the simulation results associated with EUI and its components.

meter became active at the beginning of this year, the electricity consumption for the first half of the year, from February 2024 to August 2024, was recorded at 42.06 $\mathrm{kWh/m^2{\cdot}6\,months}$, with the full-year estimate being 84.12 $\mathrm{kWh/m^2{\cdot}year}$. This closely aligns with the simulated results from the proposed framework. Regarding district heating, the simulated EUI is 29.03 $\mathrm{kWh/m^2{\cdot}year}$, lower than the previous academic year's heating energy bill of 46.93 $\mathrm{kWh/m^2{\cdot}year}$. This difference is due to the inclusion of domestic hot water consumption from common areas and accommodation sections outside the podium, which cannot be easily separated, as well as the fact that operational records show the heating water loops were not turned off every night as planned. Consequently, the comparison analysis of the two EUIs highlights the feasibility and credibility of the proposed framework. With the activation of additional meters, more detailed real-time data will become available, facilitating model calibration and further enhancing the framework.

Furthermore, Fig. 14 illustrates a breakdown analysis of the total electricity EUI in terms of lighting, equipment, and HVAC components. It can be seen that the proportion of lighting was similar to the equipment, while the proportion of the HVAC system was slightly higher than the results obtained from previous engineering practical experiences. This result overall aligns well with the ground truth, whereas the little discrepancy arises from the operation mode being based on standard scenarios outlined in building manuals without considering the actual partial operation caused by UCL's soft desk policy. Hence, if more precise schedules of occupant behaviours and system operation can be obtained, it would be possible to refine the model details and achieve more accurate simulation results.

4.3. Limitations and future work

In the proposed framework, defining HVAC topologies is a critical step that relies on high-quality, precise IFC files specific to the HVAC domain. This process is further supported by prior building knowledge to guide the semi-automated generation and validate the resulting BEM. Additionally, supplementary data sources beyond BIM are necessary to enrich the HVAC topology with quantitative information. Without these, only structural information of HVAC systems can be extracted, leaving most device capacity fields set to “Autosize” or based on assumptions. While the GRC-based workflow can switch to bounding box clash detection when surface errors occur, additional checking methods are needed to ensure the accuracy and completeness of HVAC graphs. Meanwhile, this study focuses on generating non-geometric components of BEM, which avoids architectural BIM geometry errors (e.g., slabs, walls, and spaces) directly affecting HVAC graph generation. However, indirect impacts may arise if space volumes are incorrectly defined, hindering the proper placement of terminal devices. Furthermore, human errors often increase the need for manual intervention to correct issues, even though the geometry-induced graph can pinpoint specific problems, such as missing elements or geometric mismatches.

It is important to point out that the proposed framework focusing on HVAC configuration and topology can be integrated into a comprehensive BIM2BEM workflow. The geometric aspect of the BIM2BEM workflow—specifically the automatic generation of an IDF file containing building geometry and constructions, including second-level space boundary topology—has been addressed in our earlier study [8,16,34]. By integrating these two aspects, a complete BIM2BEM workflow that encompasses both passive and active components can be achieved. Future efforts could focus on enhancing the framework's ability to manage varying data quality by incorporating quality-control methods at each stage, ensuring the reliability of the generated building energy models. Simultaneously, the need for supplementary data beyond BIM can be assessed to enrich the HVAC topology, providing a more accurate and complete representation of the real-world system. Additionally, an effective calibration module addressing device performance could be developed to improve model accuracy as more operational data becomes available.

5. Conclusions

This study proposed a framework focused on HVAC information within the BIM2BEM context, aiming to automate the generation of building simulation models primarily using MEP-IFC files. A geometry-induced HVAC graph was established to describe the connectivity among the HVAC elements through the geometric relation check workflow. By combining with the issue ruleset, error detection was performed to enhance data quality. The ontology and path-finding approaches were introduced to produce an informative HVAC topology that provides a comprehensive description of the HVAC system's structure and corresponding devices' properties. A BEM-oriented HVAC topology was established from the informative HVAC topology to represent HVAC-related information in alignment with the coding syntax of the building performance simulation model. Based on the implementation of a case study of a newly constructed building with a complex modern HVAC system, the following conclusions can be drawn:

(1) The proposed framework enables automatic building performance simulation with detailed HVAC settings by integrating IFC files and other data sources, thereby approaching a level of near-complete automation.
(2) The geometric relation check workflow, focusing on the HVAC domain, can achieve the geometry-induced knowledge graph to extract all the physical connectivity among the elements and then identify the precise positions of the errors in the inputted IFC files.
(3) The informative HVAC topology plays a critical role in comprehensively reflecting the system's features by indicating the logical linking between spaces and terminals, as well as terminals and air-handling systems or water-side components, by storing the detailed device information as entity's properties.
(4) When modelling a large-scale building featuring a complex HVAC system, the proposed framework can generate a detailed and reliable building simulation model by significantly reducing modelling time and minimizing the likelihood of errors, in contrast to manual modelling processes.

Overall, this study proposes a standard paradigm for automatic building performance simulation with detailed HVAC settings from BIM data sources, which provides a generalizable technique for building digital twins that help designers assess the energy performance of buildings and HVAC systems conveniently and accurately. Furthermore, the HVAC topologies can not only facilitate automatic BEM generation but also carry substantial implications and benefits for building management systems and operational strategies.

CRediT authorship contribution statement

Meng Wang: Writing – review & editing, Writing – original draft, Validation, Methodology, Conceptualization. **Georgios N. Lilis:** Writing – review & editing, Validation, Conceptualization. **Dimitris Mavrokapnidis:** Methodology. **Kyriakos Katsigarakis:** Software, Resources, Methodology. **Ivan Korolija:** Writing – review & editing, Funding acquisition, Conceptualization. **Dimitrios Rovas:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the Horizon Europe DigiBUILD Project under contract (No. 101069658), and co-funded by the UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee grant number 10040988.

Data availability

Data will be made available on request.

References

[1] Dimitrios Mavrokapnidis, Georgios Lilis, Kyriakos Katsigarakis, Ivan Korolija, Dimitrios Rovas, Semi-Automated Extraction of Hvac System Topology from Imperfect Building Information Models, IBPSA, 2023.
[2] Hongxin Wang, Peng Xu, Huajing Sha, Jiefan Gu, Tong Xiao, Yikun Yang, Dingyi Zhang, Bim-Based Automated Design for Hvac System of Office Buildings—an Experimental Study, Building Simulation, vol. 15, Springer, 2022, pp. 1177–1192.
[3] Vladimir Bazjanac, Ifc Bim-Based Methodology for Semi-Automated Building Energy Performance Simulation, 2008.
[4] Z. Pezeshki, A. Soleimani, A. Darabi, Application of bem and using bim database for bem: a review, J. Build. Eng. 23 (2019) 1–17.
[5] Carla Di Biccari, Filippo Calcerano, Francesca D'Uffizi, Antonio Esposito, Massimo Campari, Elena Gigliarelli, Building information modeling and building performance simulation interoperability: state-of-the-art and trends in current literature, Adv. Eng. Inform. 54 (2022) 101753.
[6] Hervé Pruvost, Raimar J. Scherer, Analysis of risk in building life cycle coupling bim-based energy simulation and semantic modeling, Proc. Eng. 196 (2017) 1106–1113.
[7] Cody M. Rose, Vladimir Bazjanac, An algorithm to generate space boundaries for building energy simulation, Eng. Comput. 31 (2015) 271–280.
[8] Georgios N. Lilis, Georgios I. Giannakis, Dimitrios V. Rovas, Automatic generation of second-level space boundary topology from IFC geometry inputs, Autom. Constr. 76 (2017) 108–124.
[9] Huaquan Ying, Sanghoon Lee, Generating second-level space boundaries from large-scale IFC-compliant building information models using multiple geometry representations, Autom. Constr. 126 (2021) 103659.
[10] Eric Fichter, Veronika Richter, Jérôme Frisch, Christoph van Treeck, Automatic generation of second level space boundary geometry from ifc models, in: Proceedings of Building Simulation 2021: 17th Conference of IBPSA, in: Building Simulation, vol. 17, IBPSA, Bruges, Belgium, September 2021, pp. 1083–1090.
[11] Yikun Yang, Yiqun Pan, Fei Zeng, Ziran Lin, Chenyu Li, A gbxml reconstruction workflow and tool development to improve the geometric interoperability between bim and bem, Buildings 12 (2) (2022) 221.
[12] João M.P.Q. Delgado, Ana S. Guimarães, João Poças Martins, Diogo F.R. Parracho, Sara S. Freitas, António G.B. Lima, Leonardo Rodrigues, Bim and bem interoperability–evaluation of a case study in modular wooden housing, Energies 16 (4) (2023) 1579.
[13] H. Mohamed Elnabawi, Building information modeling-based building energy modeling: investigation of interoperability and simulation results, Front. Built Environ. 6 (2020) 573971.
[14] Issa J. Ramaji, John I. Messner, Ehsan Mostavi, Ifc-based bim-to-bem model transformation, J. Comput. Civ. Eng. 34 (3) (2020) 04020005.
[15] Zhihua Chen, Zhang Deng, Adrian Chong, Yixing Chen, Autobps-bim: a toolkit to transfer bim to bem for load calculation and chiller design optimization, in: Building Simulation, Springer, 2023, pp. 1–12.
[16] Georgios I. Giannakis, Kyriakos I. Katsigarakis, Georgios Nektarios Lilis, Dimitrios V. Rovas, A Workflow for Automated Building Energy Performance Model Generation Using Bim Data, Building Simulation 2019, vol. 16, IBPSA, 2019, pp. 167–174.
[17] Ehsan Kamel, Ali Kazemian, Bim-integrated thermal analysis and building energy modeling in 3d-printed residential buildings, Energy Build. 279 (2023) 112670.
[18] Sergio Pinheiro, Reinhard Wimmer, James O'Donnell, Sergej Muhic, Vladimir Bazjanac, Tobias Maile, Jérôme Frisch, Christoph van Treeck, Mvd based information exchange between bim and building energy performance simulation, Autom. Constr. 90 (2018) 91–103.
[19] Jorge González, Carlos Alberto Pereira Soares, Mohammad Najjar, Assed N. Haddad, Bim and bem methodologies integration in energy-efficient buildings using experimental design, Buildings 11 (10) (2021) 491.
[20] Giovanni Barone, Annamaria Buonomano, Giovanni Francesco Giuzio, Adolfo Palombo, Towards zero energy infrastructure buildings: optimal design of envelope and cooling system, Energy 279 (2023) 128039.
[21] Hang Li, Jiansong Zhang, Improving ifc-based interoperability between bim and bem using invariant signatures of hvac objects, J. Comput. Civ. Eng. 37 (2) (2023) 04022059.
[22] Jong Bum Kim, WoonSeong Jeong, Mark J. Clayton, Jeff S. Haberl, Wei Yan, Developing a physical bim library for building thermal energy simulation, Autom. Constr. 50 (2015) 16–28.
[23] Hasan Sayegh, Georgios N. Lilis, Mathias Bouquerel, Thierry Duforestel, Kyriakos Katsigarakis, Dimitrios Rovas, Automatic modelica bem generation from ifc bim, in: 2024 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv), IEEE, 2024, pp. 437–441.
[24] Haidar Hosamo Hosamo, Merethe Solvang Tingstveit, Henrik Kofoed Nielsen, Paul Ragnar Svennevig, Kjeld Svidt, Multiobjective optimization of building energy consumption and thermal comfort based on integrated bim framework with machine learning-nsga ii, Energy Build. 277 (2022) 112479.
[25] Zhangcheng Qiang, Stuart Hands, Kerry Taylor, Subbu Sethuvenkatraman, Daniel Hugo, Pouya Ghiasnezhad Omran, Madhawa Perera, Armin Haller, A systematic comparison and evaluation of building ontologies for deploying data-driven analytics in smart buildings, Energy Build. 292 (2023) 113054.
[26] Zhelun Chen, Gabe Fierro, Virginia Gori, Henrik Madsen, Anna Marszal-Pomianowska, Ojas Pradhan, Dimitrios Rovas, Francesco Sacco, Sofia Stensson, A. Christian Thilker, Charalampos Vallianos, Jin Wen, Stephen D. White, Data-Driven Smart Buildings: State-of-the-Art Review Energy in Buildings and Communities Technology Collaboration Programme, 2023.
[27] Bharathan Balaji, Arka Bhattacharya, Gabriel Fierro, Jingkun Gao, Joshua Gluck, Dezhi Hong, Aslak Johansen, Jason Koh, Joern Ploennigs, Yuvraj Agarwal, et al., Brick: metadata schema for portable smart building applications, Appl. Energy 226 (2018) 1273–1292.
[28] Ville Kukkonen, Ali Kücükavci, Mikki Seidenschnur, Mads Holten Rasmussen, Kevin Michael Smith, Christian Anker Hviid, An ontology to support flow system descriptions from design to operation of buildings, Autom. Constr. 134 (2022) 104067.
[29] Nicolas Pauen, Dominik Schlütter, Jérôme Frisch, Christoph van Treeck, Tubes system ontology: digitalization of building service systems, in: Proceedings of the 9th Linked Data in Architecture and Construction Workshop, 2021.
[30] Zhaoji Wu, Jack C.P. Cheng, Zhe Wang, Helen H.L. Kwok, An ontology-based framework for automatic building energy modeling with thermal zoning, Energy Build. (2023) 113267.
[31] Esben Visby Fjerbæk, Mikki Seidenschnur, Ali Kücükavci, Kevin Michael Smith, Christian Anker Hviid, From bim databases to modelica-automated simulations of heating systems, in: 14th REHVA HVAC World Congress, TU Delft, 2022, pp. 2437–2443.
[32] Na Luo, Marco Pritoni, Tianzhen Hong, An overview of data tools for representing and managing building information and performance data, Renew. Sustain. Energy Rev. 147 (2021) 111224.
[33] buildingSMART International, Ifc schema specifications, https://technical.buildingsmart.org/standards/ifc/ifc-schema-specifications/, 2024. (Accessed 23 October 2024).
[34] Kyriakos Katsigarakis, Georgios Nektarios Lilis, Dimitrios Rovas, A cloud IFC-based BIM platform for building energy performance simulation, in: EC3 Conference 2021, vol. 2, 2021, pp. 350–357.
[35] Yu-Wei Zhou, Zhen-Zhong Hu, Jia-Rui Lin, Jian-Ping Zhang, A review on 3d spatial data analytics for building information models, Arch. Comput. Methods Eng. 27 (5) (2020) 1449–1463.
[36] Handi Chandra Putra, Tianzhen Hong, Clinton Andrews, An ontology to represent synthetic building occupant characteristics and behavior, Autom. Constr. 125 (2021) 103621.
[37] Nicolas Pauen, Jérôme Frisch, Christoph van Treeck, Integrated representation of technical systems with bim and linked data: tubes system ontology, Autom. Constr. 165 (2024) 105502.
[38] Dimitris Mavrokapnidis, Enabling scalable deployment of data-driven applications across building portfolios, in: EC3 Conference 2023, vol. 4, European Council on Computing in Construction, 2023.
[39] Baozhu Liu, Xin Wang, Pengkai Liu, Sizhuo Li, Xiaofei Wang, Pairpq: an efficient path index for regular path queries on knowledge graphs, in: Web and Big Data: 5th International Joint Conference, APWeb-WAIM 2021, Proceedings, Part II 5, Guangzhou, China, August 23–25, 2021, Springer, 2021, pp. 106–120.
[40] ASHRAE Handbook—HVAC Systems and Equipment, American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE), Atlanta, GA, 2020, Available in I-P and sI editions.

[41] H.M. Government, The building regulations 2010: Part l, https://www.labc.co.uk/professionals/building-regulations-guidance-documents/approved-documents-and-technical-guidance-england/approved-document-l-conservation-fuel-and-power, 2023.
[42] UK NCM, Uk's national calculation method for non domestic buildings, https://www.uk-ncm.org.uk/, 2023.
[43] Ivan Korolija, Ljiljana Marjanovic-Halburd, Yi Zhang, Victor I. Hanby, Uk office buildings archetypal model as methodological approach in development of regression models for predicting building energy consumption from heating and cooling demands, Energy Build. 60 (2013) 152–162.
[44] Roberto Lamberts, Jan L.M. Hensen, Building Performance Simulation for Design and Operation, Spoon Press, 2011.
[45] Brackney Larry, Parker Andrew, Macumber Daniel, Benne Kyle, Building Energy Modeling with Openstudio: A Practical Guide for Students and Professionals, 2018.