CCC 2025 **Proceedings of the Creative Construction Conference** (2025)
Edited by: Miroslaw J. Skibniewski & Miklós Hajdu & Žiga Turk
https://doi.org/10.22260/CCC2025/0021

SIMPLIFICATION AND ENRICHMENT OF IFC MODELS FOR COST ESTIMATION

Daiva Marcinkeviciute[1], Lukas Schildknecht[1], Manfred Huber[1], Marc Pancera[1], Joel Gschwind[1], Mischa Badertscher[2], Juan Crevillen[2]

1 FHNW, Muttenz, Switzerland
2 keeValue AG, Brugg, Switzerland

Abstract
Building Information Modelling (BIM) has become a cornerstone for data-driven project management. However, the information requirements for BIM models used in cost estimation are often too high, making them difficult to meet, especially during the early stages of a project.

This study explores the possibility to reduce these requirements by using semantic enrichment (SE) for extracting additional insights from low-detail models. To address this, key information necessary for cost estimation was identified, distinguishing between the parameters best extracted directly from the model and those that can be computed algorithmically through geometric and topologic analysis. A custom algorithm was then developed to calculate these properties automatically by either generating new construction parts or recalculating existing ones and evaluating their properties (such as complexity). The approach was tested on 30 sample models and evaluated by a panel of 16 architects and BIM/digitalization experts. While the solution works with complex models, its core approach is the intelligent analysis and enrichment of volumetric models. To derive necessary information from higher level models, simplification and abstraction was adopted.

The outcomes confirmed that the code successfully enriched models with varying levels of detail and the enriched BIM models could provide the required data for the cost estimation effectively. Detailed cost estimation uses 180 parameters: currently 45% are extracted from volumetric models, 45% from low-detail models, and 10% added manually. The algorithm will be integrated into a productive web-based cost estimation tool.

This method significantly accelerates BIM based cost estimation during early design stages, reducing manual effort and reliance on detailed models. Additionally, the procedure is adaptable to other use cases, such as life cycle cost calculations, further enhancing the practicality and adoption of BIM workflows in early-phase decision-making.

Peer-review under responsibility of the scientific committee of the Creative Construction Conference 2025.

Keywords: BIM, Cost estimation, IFC, Information requirements, Semantic Enrichment.

1. Introduction

The Architecture, Engineering, and Construction (AEC) industry is undergoing a significant digital transformation, driven by advances in technology and the growing adoption of Building Information Modelling (BIM). Model-based cost estimation is among the main research areas in BIM [1]. It enhances cost evaluation by automatically updating material quantities in response to design changes [1].

Traditionally, cost estimation from BIM models has required highly detailed and comprehensive input, often demanding intricate modelling and extensive data that are either unavailable or unreliable during early project stages. Furthermore, inaccuracies in detailed input data can lead to flawed cost estimates.

Corresponding author email address: daiva.marcinkeviciute@fhnw.ch

This dependency on complex, error-prone models has created a gap in leveraging BIM for cost estimation, limiting its effectiveness and discouraging its adoption [2], [3], [4].

This study aims to reduce the information requirements for BIM based cost estimation by utilizing enriched BIM data to compensate for missing or imprecise information. It is a part of the "keeValue to BIM" initiative. KeeValue is a cost estimation system that leverages a construction cost database and artificial intelligence (AI) to generate reliable estimates from minimal building data [5]. The initiative integrates BIM data into cost calculation algorithms in the keeValue application. It uses semantic enrichment (SE) to enhance the BIM model and automatically fill the information gaps. It extracts the necessary information from BIM models and processes it to generate missing data, such as absent building elements, or assesses aspects of the development, such as its complexity, e.g., based on the spatial arrangement.

By combining the enriched BIM data with the KeeValue application, accurate cost estimation can be directly derived from simple, low-detail BIM models.

2. Background

BIM is widely used for cost estimation in construction, primarily through a bottom-up approach, where costs are calculated by summing quantities and unit prices [6]. These quantities can be derived from IFC models, as demonstrated in various studies [2], [4], [7]. A key limitation of this approach is its reliance on detailed BIM models, which are often underdeveloped in the early phases of a project, leading to reduced accuracy in cost estimation [3]. Missing information could be supplemented through a process called semantic enrichment (SE) [8]. This term has emerged for the enrichment of information in the context of BIM, though it remains a relatively young field with limited literature [9], [10].

Methods of SE can be classified into 2 categories: SE from internal information, where new information is generated solely from the existing data within the BIM model, such as automatic determination of room usage from geometric analysis [10], and SE from external information, which may include image data, point clouds for Scan2BIM, integration with GIS data, utilizing ontologies or semantic web techniques like RDF and SPARQL [11]. Xue et al. described a conceptual model for SE, which identifies 3 methods: semantic reasoning, involving rule-based algorithms and simulations; semantic registration, which uses iterative trial-and-error procedures; and semantic segmentation, applying supervised machine learning methods [12]. Bloch and Sacks emphasize that the choice between rule-based or data-based methods depends on the specific issue at hand [10].

Several methods have been explored to identify and correct the use of semantic IFC entities, including the use of AI and image recognition to analyse geometric representations [13], [14], decision tree algorithms and machine learning techniques [15], [16], expert knowledge-based approaches [17], [18] and more recently, the use of large language models to identify objects and augment them with missing properties from external databases [19]. Generating key relationships, such as hosting/hosted/adjacent, has also been explored [16], [20]. Research on generating new elements within BIM models remains limited, although there have been efforts to generate missing spatial objects such as buildings, storeys, or spaces [16], as well as HVAC systems [21]. The tool Topologic introduces the concept of rooms (cells) as the basis for generating building parts, where walls, columns, and slabs can be derived from cell boundaries rather than being explicitly modelled [22]. Meanwhile the tool Hypar generates building storeys, facades, roofs, slabs and other elements based on volumetric building models [23].

Simplification is a crucial topic for this paper, as it plays an essential role in facilitating the SE processes. BIM models, due to their complexity, often require simplification for various applications such as urban visualization, GIS integration and rule-checking, thus simplification is commonly employed to extract lightweight but sufficiently detailed data from these models [24]. While basic simplification methods, such as finding the bounding box or using the octree of the elements, are available [25], they often fail to maintain the more complex yet necessary geometric aspects. Other studies have focused on simplifying specific parts of the building, particularly the external envelope or internal spaces, by turning them into simplified prisms. For example, He et al. uses footprint-based generalization to group buildings into simplified volumes suitable for urban visualization [26]. Kim and Li focuses on simplifying indoor spaces, ensuring that key geometric features are preserved while reducing overall complexity [24]. Zhu et al. proposed a framework for generating simplified solid building models for GIS integration using a semantics-based method that included identifying external objects, distinguishing between different slabs, and generating valid external walls [27]. Another methods for finding and simplifying building envelope include first applying the correct semantics to filter the model and then by performing a series of geometric operations in 3D to extract the external envelope [28] or involve first extracting a point cloud from BIM models and then extracting this point cloud's exterior boundary using the 3D alpha shape algorithm [29].

The characteristics of information that can be enhanced through SE have not been fully explored yet [9]. Existing methods remain in the research stage and are not industry-ready, requiring manual operation [8] and programming, querying, or expert knowledge. They often focus on specific problems, such as object identification [8], rather than addressing a particular purpose like cost estimation and its specific needs. However, the selection of appropriate tools and methods for SE should depend on the nature of the problem [9]. Bloch suggests that a potential tool could offer a narrow scope for SE, focusing on a specific domain by integrating the SE engine into a receiving application [9].

Furthermore, most research has been based on a limited set of "idealized" models, which tend to be simpler than real-world industry models. Industry models are likely to be harder to interpret, emphasizing the need for solutions capable of handling such complexities.

3. Methodology and Findings

The study was conducted by software engineers with expertise in AEC and geometric analysis, in collaboration with professionals in cost estimation and BIM modelling, and was supported by a panel of 16 architects and BIM/digitalization experts from keeValue's network of clients and partners. It explored how SE can reduce information requirements for BIM models in cost estimation by extracting insights from low-detail models.

3.1. Preliminary Steps

A total of 180 parameters were defined for detailed cost estimation, encompassing quantity take-offs, complexity assessments, materials and logical attributes (e.g., presence of an atrium: True/False). Please see the table below for a small sample of these parameters (Table 1).

Table 1. A sample of parameters

Parameter	Description
ord_form_baseplate	complexity of baseplate [score from 1 to 8]
ord_form_roof	complexity of roof [score from 1 to 8]
area_insulation_wall_inside	area of insulation for internal walls [m2]
area_insulation_wall_outside	area of insulation for external walls [m2]
num_buildings	count of separate buildings in the model [#]
num_appartments	count of apartments [#]
table_elevators	list of elevators with their vertical heights [m] and number of storeys [#]
area_window_inside	area of internal windows [m2]
area_window_outside_aboveground_east	area of external windows above ground facing east [m2]
area_window_outside_aboveground_northeast	area of external windows above ground facing northeast [m2]
area_door_outside_aboveground_north	area of external doors above ground facing north [m2]

To evaluate the possibilities of SE, the study examined which parameters can be automatically derived from an IFC model's information structure, geometry, or topology. To assess the current state of modelling in the construction sector, approximately 40 industry models were collected, that varied significantly in terms of content, level of detail, and adherence to modelling guidelines, reflecting the diverse practices within the industry. Then recurring data patterns were assessed to determine data presence, accuracy, and the relative cost of manual versus automated completion of missing information.

Next, the study explored rule-based and data-driven approaches for parameter derivation, ensuring applicability across different building types (avoiding a focus on specific categories like single-family homes). Parameters were categorized as directly extractable (counting objects and aggregating base quantities), requiring rule-based calculations, or neither practical to input nor reliably calculable. They were also ranked by their importance for cost estimation and the effort needed for data input.

Finally, the study determined the minimum information requirements for BIM models by consolidating insights from the previous steps, identifying critical data for parameter calculations, and prioritizing the most essential information.

The analysis of parameters and models revealed 2 common levels of detail in the early stages of planning. The most basic level consists of simple volumetric model, described in chapter 3.2.1 “Semantic Enrichment of Volumetric Models”. The next level includes low-detail models, described in chapter 3.2.2 “Semantic Enrichment of Low-Detail Models”.

3.2. Framework

A software module was then developed to extract parameters directly from the model or to derive them through SE. An important requirement was that the evaluation process should operate independently of user intervention, thus missing information was handled using general industry assumptions.

The implementation used Python for its flexibility and extensive library ecosystem. Key libraries included ifcOpenShell for editing IFC files, PythonOCC for 3D geometry processing, and Shapely for 2D geometry operations. 30 standardized sample models, based on real buildings of different types, were created for controlled testing and algorithm validation.

3.2.1. Semantic Enrichment of Volumetric Models

A volumetric model consists of an appropriately defined spatial structure (comprising project, site, building, and building storey objects) that contain space elements representing a whole or part of a building storey. These elements, further referred to as storey volumes, must connect to define the building's shape without intersecting. They should include details on whether a space is external, its usage (residential, commercial, etc.), and whether it is an airspace (excluded from the gross floor area). Optional data includes heating status. The model may also include the terrain shape.

By analysing the topological relationships between the elements of a volumetric model, insights about the building envelope properties and complexity are derived, and new building elements are generated and measured.

Building envelope elements are derived from the faces of storey volumes. The faces are categorized by orientation and relative position to adjacent elements, allowing them to be interpreted as different building components. This process helps to identify important characteristics such as whether an element horizontal or vertical and if it is external or internal. For example, upper horizontal faces of internal storey volumes are categorized as ceilings, roofs, or accessible roofs based on their interactions with other storey volumes (Fig. 1).

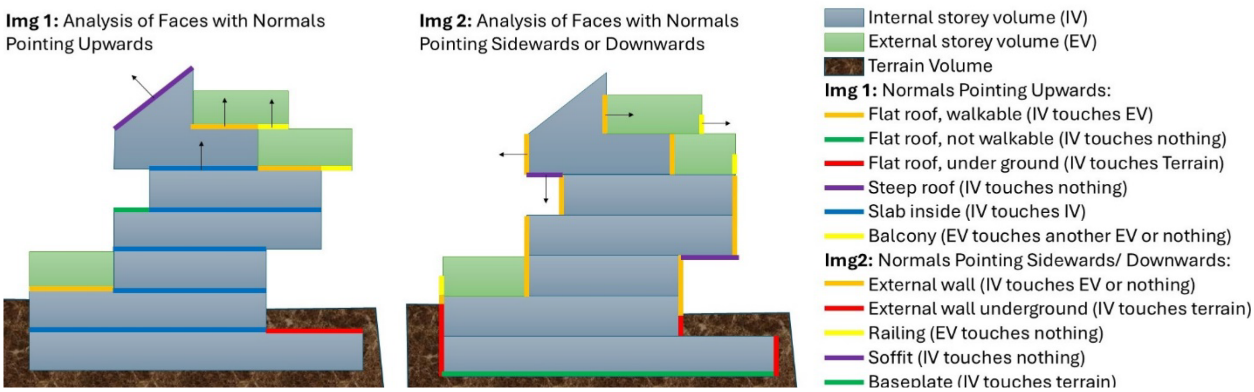

Fig. 1. Rules for generating building envelope elements.

The facades are assigned a cardinal direction (North, North-East, East, …) based on their normal. This analysis is also sensitive to whether terrain geometry is modelled. If terrain is present, elements are classified as underground if they overlap with the terrain volume. If no terrain is modelled, the elements are classified as underground if they lie in building storeys underground.

If heating status information is provided, it helps determine the building's approximate thermal envelope, identifying both internal and external elements that require insulation.

To calculate the complexity of newly generated building envelope components (e.g., facade, roof, balconies, baseplate), all the faces constituting a component are analysed based on their geometric characteristics (e.g., right vs. non-right angles), inclination relative to surrounding elements, and the recurrence of similar forms. These factors are combined to generate a complexity score for each component, helping to quantify the overall building envelope's complexity.

The clustering and merging of storey volumes enable the calculation of building count and dimensions.

3.2.2. Semantic Enrichment of Low-Detail Models

Low-detail models consist of basic building elements like walls, floors, roofs, windows, doors, and stairs, which do not require precise modelling or additional information such as position (inside/outside, above/below ground), material, or thermal properties. They may also include space elements representing rooms with general usage types (main, secondary, circulation, functional) and freely assigned names. These elements were typically provided in the inspected models and contained quantities automatically calculated by the CAD software.

By analysing the topological relationships between building elements, rooms, and elements derived from volumetric models, relationships are identified, missing properties are assigned, and new elements are generated.

For accurate cost estimation, properties such as "inside/outside" and "above/below ground" are calculated for windows, doors, and curtain walls, as this information is often inaccurate in early modelling phases and is not required from the user. This involves comparing the position of opening elements with the facades generated from the volumetric models. If an opening is near a facade and has similar rotation, the component is labelled "external," with directionality based on the facade's cardinal direction. If terrain geometry is available, an opening is labelled "underground" if the majority of it is within the terrain.

This process does not apply to walls or ceilings/roofs because their position relative to inside/outside or above/below ground cannot always be easily determined, given that these components often span multiple conditions (e.g., half inside and half outside). Additionally, facades, slabs, and roofs are generated from the volumetric model, making this approach redundant.

To identify elevators, rooms labelled "elevator", “lift” or similar are found by searching for specific substrings in room names. If the rooms are vertically aligned, they are grouped together to form an elevator. Similarly, rooms labelled "sanitary”, “wc” or related terms are gathered and analysed for potential common shafts. Elevators must overlap fully, while sanitary rooms can overlap partially or be adjacent. If these rooms are vertically aligned or share a wall of sufficient size, they are grouped into a common shaft, and the necessary number of shafts is determined. The proximity of these shafts to each other as well as to the technical rooms also gauge the complexity of the sanitary system.

A geometric analysis of rooms, doors, and stairs calculates the number and composition of apartments. Residential rooms are identified by comparing the low-detail model with the volumetric model and selecting the rooms within residential storey volumes. Residential rooms that are connected or share doors or a staircase are grouped together to form an apartment, and the total number of such clusters is counted as the number of apartments. This is the only analysis that is used exclusively for residential buildings. All other types of simplification and SE can be used for all building types.

3.2.3. Simplification for Volumetric Models

Although the volumetric model is of low-detail, it must still adhere to strict requirements for consistent volumetric analysis. To address common issues with volumetric models, simplifications are applied to correct modelling irregularities, allowing for a broader range of industry models — often created for other purposes — to be accepted and later used for cost calculations.

One example of simplification in this project is terrain geometry, which plays a crucial role in determining which building elements are above or below ground. Terrain geometry, however, can vary significantly in both form and quality. It may be modelled as either a volume or a surface, and buildings might be cut out of the terrain in inconsistent ways. In such cases, a procedure was developed to recalculate the geometry when terrain is imported. First, the upper surface of the terrain volume is identified, either by analysing the volume or using the surface, depending on how the terrain is modelled. Points are then sampled from the surface, and an average plane is calculated to determine the inclination of the terrain. Using this new surface, a simplified terrain volume is generated (Fig. 2), which can be used for further calculations of underground building elements. While the areas extracted using simplified terrain are slightly less precise, the method prevents much bigger discrepancies caused by inconsistent modelling practices.

Another example involves the generation of missing storey volumes. Storey volumes are essential for generating building envelope elements and assessing complexity. However, they are rarely modelled explicitly, while rooms are common. To address this, a procedure was developed to generate storey volumes by simplifying the rooms. This process manipulates the faces on each building storey through offsetting, merging, and extrusion operations to create the common storey volume. The resulting volumes are then adjusted to eliminate any overlaps between storeys (Fig. 2). This method only works when the gaps between room volumes fall within an allowed tolerance.

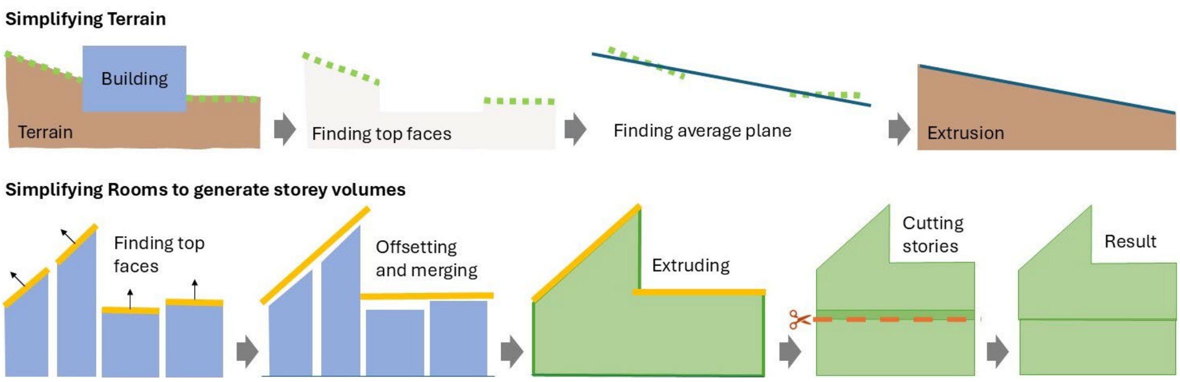

Fig. 2. Rules for simplification of terrain and rooms

3.2.4. Manual Parameters

In early-stage models, certain information, such as building materials, is often absent or inaccurate, or in case of volumetric models even impossible to input. Early on, however, a quick analysis is often needed to compare the cost effects of using different materials or achieving a specific energy standard. Manually entering this information is not practical, as volumetric models lack efficient methods for such input, and low-detail models would require users to repeatedly update materials and properties for each element with every design iteration.

To address this challenge, a user interface was developed, allowing users to manually input critical parameters, such as building materials, eco and energy standards, and HVAC system details. These parameters are some of the most difficult to input into the model or reliably assess from low-detail models. Thus, given that early cost estimation does not require highly detailed data, manual adjustments is a practical and flexible solution. This approach enables users to experiment with various cost scenarios without having to modify the BIM model for every iteration.

Some parameters, such as building site conditions, excavation, and protection of existing structures are handled as manual inputs, even though they could be evaluated automatically using GIS integration. This approach was not implemented due to time constraints.

4. Validation

To illustrate and validate the results of simplification and SE, 30 IFC models were analysed using the described algorithm and the enriched models were manually inspected and validated. 2 apartment building models (A by MBA Immobilien AG and B by Konzeptwerk GmbH) containing both volumetric and low-detail representations were selected to demonstrate the results of SE.

First, the terrain was simplified into a prism (Fig. 3). The volumetric models were then analysed alongside the terrain volume to generate a building envelope containing all relevant information. The newly generated elements were successfully classified as internal or external, above or underground, and insulated or not. Roofs were further categorized as steep or flat and walkable or not. Slabs were identified as baseplates, soffits, or regular slabs, while facades were assigned cardinal directions (Fig. 4). Complexity was calculated for each element group (Fig. 5).

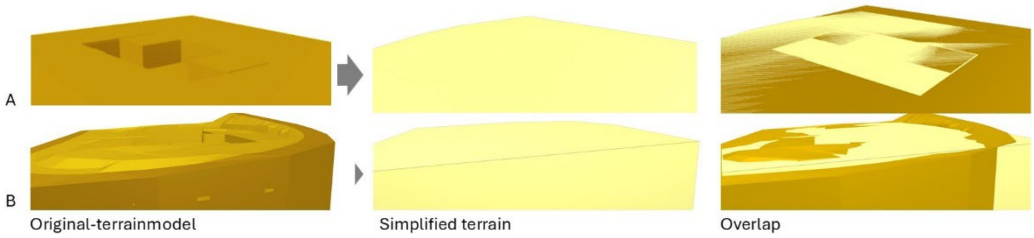

Fig. 3. Terrain simplification for buildings A and B

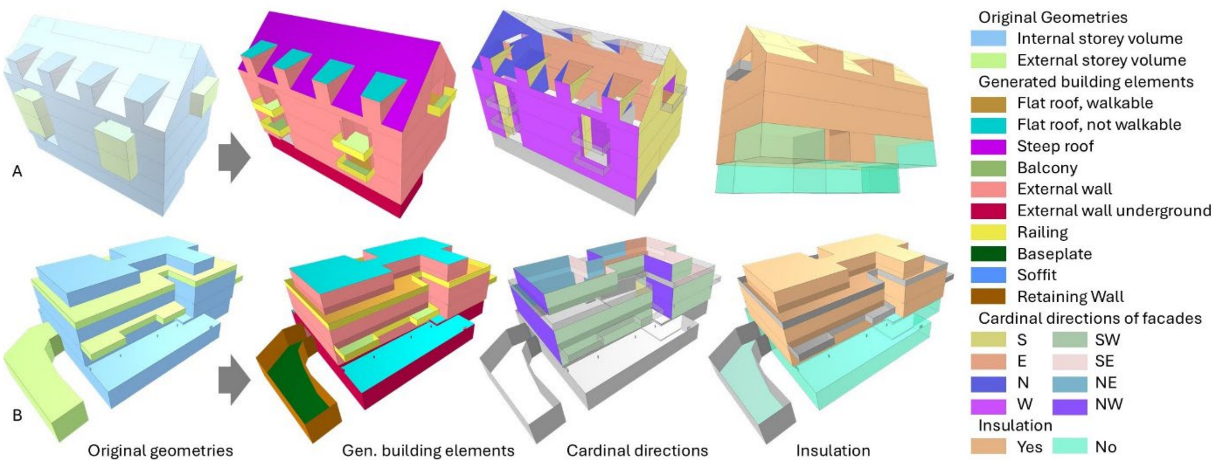

Fig. 4. Generated envelope elements for buildings A and B

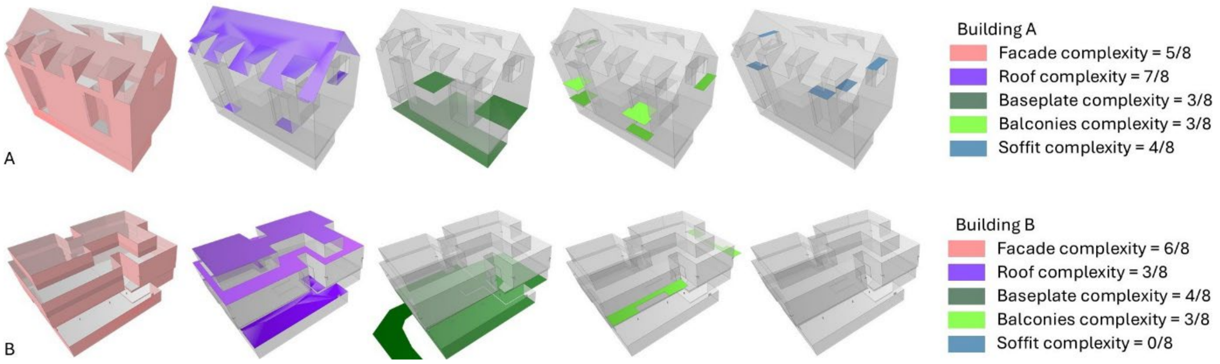

Fig. 5. Complexity calculation for buildings A and B

The buildings were analysed to determine their count and correct dimensions. As windows and doors were provided, their position relative to the ground level (above or underground), placement (inside or outside), and cardinal direction were identified (Fig. 6). The analysis of rooms extracted various details, including the count of apartments and the complexity of the sanitary system (Fig. 7).

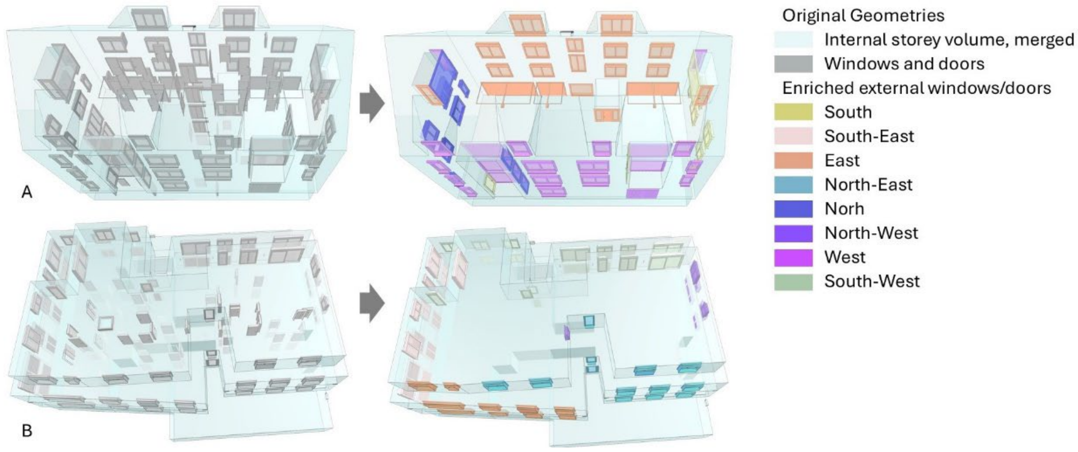

Fig. 6. Windows and doors classification for buildings A and B

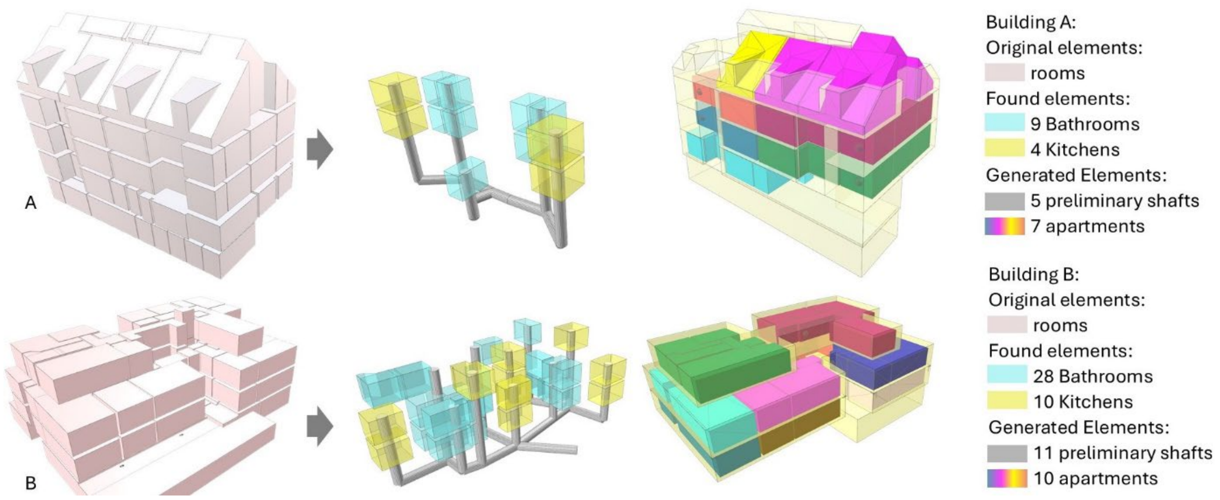

Fig. 7. Sanitary-shaft and apartment generation for buildings A and B

Additional smaller analyses were performed, and certain properties, such as address and building part count, were directly read from the model. The results confirmed that the code successfully enriched the models, producing plausible and sufficient outcomes. An isolated issue was observed in Building A, where certain kitchens were not detected due to non-standard naming. This discrepancy is considered insignificant in the context of the overall analysis.

When storey volumes are absent, rooms are simplified and used to generate them. For the illustration, 3 IFC models available online (buildings C, D and E) were used to generate storey volumes (Fig. 8). Since storey volumes are an abstraction rather than an exact representation of the building shape, minor discrepancies of 20–30 cm occurred. However, these differences are considered negligible, as they do not significantly affect the total area calculations.

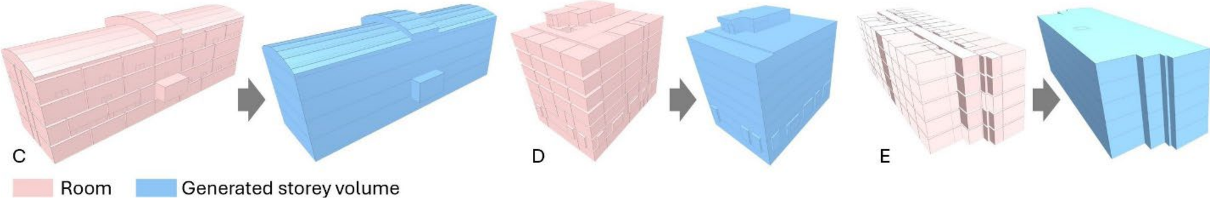

Fig. 8. Room-simplification for buildings C, D and E

5. Discussion and Conclusions

Thorough simplification and SE the information requirements for BIM models are significantly reduced. The panel of experts assessed these requirements as reasonable and easy to implement. Based on the study results, 10% of parameters were found infeasible to extract, 45% can be derived from volumetric models and another 45% from low-detail models. The extracted data is sufficiently accurate for cost calculation, as confirmed by the cost estimation professionals. The most important parameters for cost estimation can be extracted from volumetric models, while additional information from low-detail models can enhance the accuracy of the cost estimation.

SE significantly reduces BIM model information requirements, allowing early-stage cost calculations with simple models. This enhances decision-making and facilitates project comparisons. The approach of SE and automatic correction can extend beyond cost calculations to other building assessments, such as life cycle analysis, increasing the feasibility of BIM models in early project phases.

However, challenges remain, including variations in IFC structures, irregular geometry types, import errors and issues within geometric processing libraries. Excessive geometric details can make some models heavy and error prone. Special cases and exceptions are common in buildings and make it difficult to establish universal rules for all building types. Also, simplification and SE involve a significant amount of abstraction and while they often deliver good results, their accuracy is not guaranteed.

Future research could focus on reducing information requirements through SE for other use cases. Additionally, GIS integration offers potential for evaluating building site conditions, excavation needs, and the protection of existing neighbouring structures.

6. Bibliography

[1] A. Heidari, Y. Peyvastehgar, and M. Amanzadegan, “A systematic review of the BIM in construction: from smart building management to interoperability of BIM & AI,” Architectural Science Review, vol. 67, no. 3, pp. 237–254, Aug. 2023, doi: 10.1080/00038628.2023.2243247.

[2] F. H. Abanda, B. Kamsu-Foguem, and J. H. M. Tah, “BIM – New rules of measurement ontology for construction cost estimation,” Engineering Science and Technology, an International Journal, vol. 20, no. 2, pp. 443–459, Apr. 2017, doi: 10.1016/j.jestch.2017.01.007.

[3] J. Lee, H. Yang, J. Lim, T. Hong, J. Kim, and K. Jeong, “BIM-based preliminary estimation method considering the life cycle cost for decision-making in the early design phase,” Journal of Asian Architecture and Building Engineering, vol. 19, no. 4, pp. 384–399, May 2020, doi: 10.1080/13467581.2020.1748635.

[4] S. Wu, G. Wood, K. Ginige, and S. W. Jong, ‘A technical review of BIM based cost estimating in UK quantity surveying practice, standards and tools’, J. Inf. Technol. Constr. ITcon, vol. 19, pp. 534–563, 2014.

[5] ‘kennwerte’, website. Accessed: Apr. 03, 2025. [Online]. Available: https://www.keevalue.ch/

[6] F. K. T. Cheung, J. Rihan, J. Tah, D. Duce, and E. Kurul, “Early stage multi-level cost estimation for schematic BIM models,” Automation in Construction, vol. 27, pp. 67–77, Nov. 2012, doi: 10.1016/j.autcon.2012.05.008.

[7] A. Fazeli, M. S. Dashti, F. Jalaei, and M. Khanzadi, “An integrated BIM-based approach for cost estimation in construction projects,” Engineering, Construction and Architectural Management, vol. 28, no. 9, pp. 2828–2854, Dec. 2020, doi: 10.1108/ecam-01-2020-0027.

[8] S. Jiang, X. Feng, B. Zhang, and J. Shi, "Semantic enrichment for BIM: Enabling technologies and applications," Advanced Engineering Informatics, vol. 56, p. 101961, Apr. 2023, doi: 10.1016/j.aei.2023.101961.

[9] T. Bloch, "Connecting research on semantic enrichment of BIM - review of approaches, methods and possible applications," Journal of Information Technology in Construction, vol. 27, pp. 416–440, Apr. 2022, doi: 10.36680/j.itcon.2022.020.

[10] T. Bloch and R. Sacks, "Comparing machine learning and rule-based inferencing for semantic enrichment of BIM models," Automation in Construction, vol. 91, pp. 256–272, Jul. 2018, doi: 10.1016/j.autcon.2018.03.018.

[11] F. M. Dinis, J. Poças Martins, A. S. Guimarães, and B. Rangel, "BIM and Semantic Enrichment Methods and Applications: A Review of Recent Developments," Archives of Computational Methods in Engineering, vol. 29, no. 2, pp. 879–895, May 2021, doi: 10.1007/s11831-021-09595-6.

[12] F. Xue, L. Wu, and W. Lu, "Semantic enrichment of building and city information models: A ten-year review," Advanced Engineering Informatics, vol. 47, p. 101245, Jan. 2021, doi: 10.1016/j.aei.2020.101245.

[13] C. Emunds, N. Pauen, V. Richter, J. Frisch, and C. Treeck, IFCNet: A Benchmark Dataset for IFC Entity Classification. 2021.

[14] C. Mirarchi, M. Gholamzadehmir, B. Daniotti, and A. Pavan, "Semantic Enrichment of BIM: The Role of Machine Learning-Based Image Recognition," Buildings, vol. 14, no. 4, p. 1122, Apr. 2024, doi: 10.3390/buildings14041122.

[15] S. Bigdeli, P. Pauwels, S. Verstockt, N. Van de Weghe, and B. Merci, "Semantic Enrichment of a BIM Model Using Revit: Automatic Annotation of Doors in High-Rise Residential Building Models Using Machine Learning," Fire Technology, Oct. 2024, doi: 10.1007/s10694-024-01655-0.

[16] Z. Wang, R. Sacks, B. Ouyang, H. Ying, and A. Borrmann, "A Framework for Generic Semantic Enrichment of BIM Models," Journal of Computing in Civil Engineering, vol. 38, no. 1, Jan. 2024, doi: 10.1061/jccee5.cpeng-5487.

[17] L. Ma, R. Sacks, U. Kattel, and T. Bloch, "3D Object Classification Using Geometric Features and Pairwise Relationships," Computer-Aided Civil and Infrastructure Engineering, vol. 33, no. 2, pp. 152–164, Nov. 2017, doi: 10.1111/mice.12336.

[18] N. Nisbet, Z. Zhang, L. Ma, W. Chen, and M. S. Çıdık, "Semantic correction, enrichment and enhancement of social and transport infrastructure BIM models," Advanced Engineering Informatics, vol. 59, p. 102290, Jan. 2024, doi: 10.1016/j.aei.2023.102290.

[19] K. Forth and A. Borrmann, "Semantic enrichment for BIM-based building energy performance simulations using semantic textual similarity and fine-tuning multilingual LLM," Journal of Building Engineering, vol. 95, p. 110312, Oct. 2024, doi: 10.1016/j.jobe.2024.110312.

[20] M. H. Rasmussen, M. Lefrançois, G. F. Schneider, and P. Pauwels, "BOT: The building topology ontology of the W3C linked building data group," Semantic Web, vol. 12, no. 1, pp. 143–161, Nov. 2020, doi: 10.3233/sw-200385.

[21] H. Wang et al., "BIM-based automated design for HVAC system of office buildings—An experimental study," Building Simulation, vol. 15, no. 7, pp. 1177–1192, Feb. 2022, doi: 10.1007/s12273-021-0883-7.

[22] R. Aish, W. Jabi, S. Lannon, N. Wardhana, and A. Chatzivasileiadi, Topologic: tools to explore architectural topology. 2018.

[23] Yuko Ishizu, Hypar for Generative Design by Andrew Heumann, (Dec. 12, 2020). Accessed: Feb. 10, 2025. [Online Video]. Available: https://www.youtube.com/watch?v=3dDRV-fDbEs

[24] J.-S. Kim and K.-J. Li, "Simplification of geometric objects in an indoor space," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 147, pp. 146–162, Jan. 2019, doi: 10.1016/j.isprsjprs.2018.11.017.

[25] W. Solihin, J. Dimyadi, Y.-C. Lee, C. Eastman, and R. Amor, "Simplified schema queries for supporting BIM-based rule-checking applications," Automation in Construction, vol. 117, p. 103248, Sep. 2020, doi: 10.1016/j.autcon.2020.103248.

[26] S. He, G. Moreau, and J.-Y. Martin, 'Footprint-Based Generalization of 3D Building Groups at Medium Level of Detail for Multi-Scale Urban Visualization', Int. J. Adv. Softw., vol. 5, pp. 378–388, Dec. 2012.

[27] J. Zhu, P. Wu, and C. Anumba, "A Semantics-Based Approach for Simplifying IFC Building Models to Facilitate the Use of BIM Models in GIS," Remote Sensing, vol. 13, no. 22, p. 4727, Nov. 2021, doi: 10.3390/rs13224727.

[28] S. Donkers, H. Ledoux, J. Zhao, and J. Stoter, "Automatic conversion of IFC datasets to geometrically and semantically correct CityGML LOD3 buildings," Transactions in GIS, vol. 20, no. 4, pp. 547–569, Sep. 2015, doi: 10.1111/tgis.12162.

[29] L. Yan, 'Extraction of Exterior Building Envelopes from Building Information Models', 2023, Accessed: Jan. 09, 2024. [Online]. Available: https://repository.tudelft.nl/islandora/object/uuid%3A2ff23b56-dc28-4c71-adf7-a9aecfcce413