Towards CSG-Based Urban Modelling through Structured Geospatial Data Integration: Preliminary Results

Elisavet Tsiranidou[1], Muataz S. A. Albadri[1], Aiara Brea-Portals[1], Antonio Fernández[1], Lucía Díaz-Vilariño[1]

[1] CINTECX, Universidade de Vigo, GeoTECH group, 36310 Vigo, Spain -
(elisavet.tsiranidou, muatazsafaaabed.albadri, aiara.brea, antfdez, lucia)@uvigo.gal

Keywords: Urban space reconstruction, 3D city models, Airborne LiDAR, OpenStreetMap, Urban space representation, Urban form indicators.

Abstract

Urban environments consist of multiple interconnected spatial entities whose integrated representation is essential for urban modelling and analysis. However, existing three-dimensional (3D) modelling approaches often focus on the reconstruction of individual object classes, limiting the generation of integrated urban-space models. This paper investigates the use of Constructive Solid Geometry (CSG) for the automatic generation of urban-space models through the integration of heterogeneous geospatial datasets, including airborne Light Detection and Ranging (LiDAR), OpenStreetMap (OSM), and cadastral information. The proposed approach leverages semantically structured urban entities derived from these datasets, including buildings, vegetation, transportation spaces, and water bodies, and organises them into two urban-space representations. These are subsequently combined using CSG Boolean operations to construct volumetric urban models, enabling the generation of topologically consistent representations while preserving semantic classes and functional attributes. The method is evaluated through a case study in Luxembourg. Furthermore, the generated CSG models support the derivation of urban-form indicators, demonstrating their potential for quantitative urban characterisation.

1. Introduction

Urban environments are composed of diverse and interconnected elements, including buildings, transportation infrastructure, vegetation, water bodies, and open spaces. The increasing availability of geospatial datasets, such as aerial laser scanning (ALS), cadastral information, and volunteered geographic information, has significantly expanded the possibilities for representing urban environments in digital form (Jeddoub et al., 2025). These digital representations support a wide range of applications, including urban planning, environmental assessment, and urban analytics (Caprari et al., 2022; Omrany et al., 2025).

Numerous approaches have been proposed for 3D urban modelling. Existing methods range from geometric reconstruction techniques based on point clouds and imagery to semantic city models that represent urban entities and their relationships (Biljecki et al., 2017; Ying et al., 2020). Automatic reconstruction workflows have further enabled the generation of large-scale 3D city models directly from geospatial datasets, facilitating their application in urban planning and analysis (Ledoux et al., 2021). Peters et al. (2022) demonstrated the automated reconstruction of LoD1 and LoD2 building models for the entire building stock of the Netherlands, illustrating the scalability of large-scale urban modelling workflows. Similarly, Pađen et al. (2022) proposed an automated workflow that integrates cadastral, topographic, and airborne LiDAR data to reconstruct simulation-ready 3D city models containing terrain, buildings, roads, water bodies, and vegetation. However, these models were specifically designed for computational fluid dynamics simulations and did not focus on the representation of semantically structured urban spaces or their organisation for urban analysis.

The increasing demand for integrated urban representations has stimulated the development of modelling approaches capable of describing cities beyond collections of individual objects. Kolbe (2009) introduced City Geography Markup Language (CityGML) as a semantic information model for representing urban entities together with their geometry, attributes, and relationships, establishing a foundation for interoperable 3D city models. More recently, Kutzner et al. (2020) presented the conceptual model of CityGML 3.0, extending the representation of urban environments through improved support for spaces, transportation infrastructure, and semantic relationships between urban objects. In parallel, Dembski et al. (2020) demonstrated the implementation of an urban digital twin for the city of Herrenberg, illustrating how heterogeneous urban datasets can be integrated to support planning and decision-making processes. Similarly, White et al. (2021) explored the use of digital twins as data integration platforms for smart-city applications, highlighting the importance of combining geometric and semantic information within coherent urban representations.

Despite the significant advances in semantic city modelling, challenges remain in representing urban environments as coherent spatial systems rather than collections of individual thematic layers. Labetski et al. (2023) highlighted that many contemporary urban models continue to organise information according to separate thematic domains, limiting the explicit representation of relationships between urban entities and spaces. Tan et al. (2023) investigated the integration of heterogeneous datasets within urban digital environments and identified difficulties in maintaining semantic consistency and spatial relationships across multiple data sources. Beyond data integration, the increasing demand for urban analytics has further emphasised the need for representations that explicitly capture the organisation of urban space. Ding et al. (2025) demonstrated that structured urban representations facilitate the derivation of urban indicators and improve the interpretation of spatial interactions between different urban components. These findings suggest that future urban models should move beyond the representation of isolated objects and provide coherent descriptions of urban space capable of supporting both modelling and analytical applications.

Several recent studies have explored different representations for modelling urban environments according to specific application requirements. Kolbe and Donaubauer (2021) described semantic 3D city models and Building Information Modelling (BIM) as object-based representations that associated urban entities with geometric, semantic, and topological information, providing an interoperable basis for urban information management. Beil and Kolbe (2024) reviewed semantic 3D streetspace models, which explicitly represented transportation spaces together with their semantic characteristics and surrounding urban context to support streetspace-related applications. Other studies focused on network-based representations, where urban environments were modelled through multiscale street-network structures to analyse the spatial organisation of cities. However, these representations primarily captured transportation networks rather than the complete urban environment (Chen et al., 2024). Transportation-oriented representations have also been proposed. For example, Tsiranidou et al. (2025b) introduced a semantic representation of road spaces extracted from point clouds, which was subsequently extended into a multi-level representation enabling progressively refined descriptions of transportation entities (Tsiranidou et al., 2025a). More recently, voxel-based representations have emerged as an alternative modelling approach. Gorte et al. (2024) proposed the integration of heterogeneous 3D datasets within a voxel domain, while Fujiwara et al. (2026) introduced VoxCity, a framework that integrated heterogeneous open geospatial datasets into grid-based semantic 3D city models for urban environmental simulation. Although voxel representations facilitate data integration and large-scale simulations, urban entities are represented as regular volumetric cells rather than explicit geometric objects.

Constructive Solid Geometry offers an alternative modelling paradigm based on the combination of volumetric primitives through Boolean operations. By representing objects as combinations of solids, CSG enables the generation of volumetric models with explicit spatial relationships and consistent topological structures. Wu et al. (2018) employed CSG to reconstruct solid models of individual objects from point clouds by fitting geometric primitives to surface patches and combining them through Boolean operations. Fayolle and Friedrich (2024) reviewed methods for converting unstructured datasets into CSG models and highlighted their advantages in terms of topological consistency and compact model representation. More recently, Albadri et al. (2025) proposed a CSG-based workflow for reconstructing Industry Foundation Classes (IFC)-compatible building models from segmented indoor point clouds, where CSG operations were used to generate solid representations of building components. However, these studies remain focused on reconstructing individual objects or building components and do not address the modelling of complete urban environments composed of multiple semantically distinct urban entities. Consequently, extending CSG to urban-scale modelling requires the prior identification and organisation of coherent spatial entities that can serve as modelling primitives.

In this context, this paper explores the use of CSG for urban-space modelling through the integration of heterogeneous geospatial datasets. The proposed approach combines airborne LiDAR data, OSM data, and cadastral information to decompose urban environments into coherent spatial entities, including buildings, vegetation, transportation spaces, and water bodies. These entities are organised into two urban-space representations, in which the general representation describes the primary urban components, while the refined representation provides a more detailed characterisation of transportation-related areas. The resulting representations are subsequently transformed into volumetric urban models using CSG principles. Furthermore, the generated models are explored through a set of urban-form indicators to illustrate their potential for urban characterisation and analysis. The framework is demonstrated through a real-world case study in Luxembourg, highlighting its ability to support both semantically structured urban modelling and the derivation of analytical information from integrated geospatial datasets.

The rest of the paper is organised as follows: Section 2 describes the proposed method, while Section 3 presents the results obtained from applying the method to a real case study. Finally, Section 4 concludes the paper and outlines the future work.

2. Method

2.1 Input datasets

The proposed framework integrates airborne LiDAR data, OSM data, and cadastral information to represent the different components of the urban environment. Figure 1 provides an overview of the proposed workflow, while the input datasets are summarised in Table 1. ALS data provide 3D information describing terrain and urban objects, including buildings, vegetation, and water bodies. OSM data supply information related to buildings, transportation infrastructure, and land-use features. Cadastral parcels are incorporated to define spatial subdivisions and support the organisation of urban entities within a common framework. The combination of these datasets enables the generation of structured urban-space representations suitable for subsequent modelling and analysis.

Dataset	Geometry	Information
Airborne LiDAR	Point cloud	Terrain, buildings, vegetation, water
OpenStreetMap	Vector	Buildings, roads, land use
Cadastral data	Vector	Parcel boundaries

Table 1. Input datasets used in the proposed framework.

2.2 Urban space decomposition

Following data integration, the urban environment is decomposed into a set of spatial entities representing its principal components. This process combines geometric and semantic information from ALS, OSM and cadastral data to generate a unified representation of urban space. Cadastral parcels provide the initial spatial framework for organising the study area into urban regions, within which buildings, vegetation areas, and water bodies are integrated. All datasets are transformed into a common spatial reference framework and aligned to ensure geometric consistency.

Building footprints are extracted from OpenStreetMap and associated with airborne LiDAR information in order to represent the built environment. Vegetation and water entities are

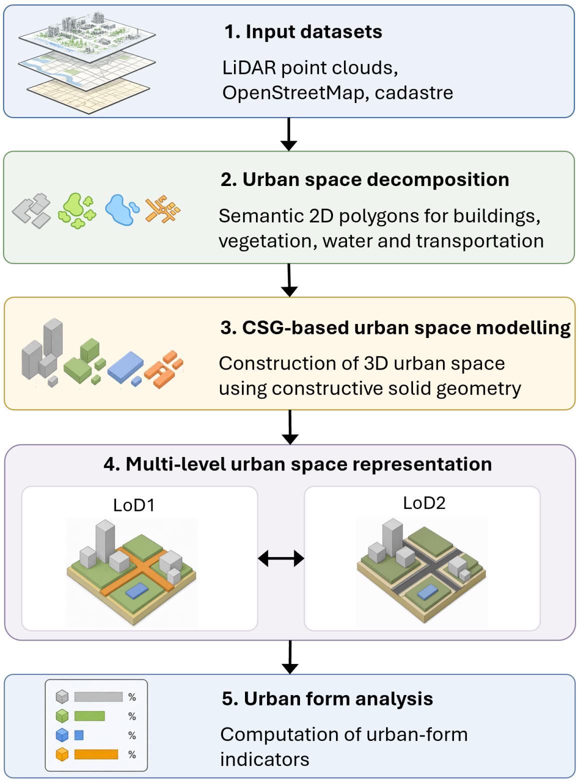

Figure 1. Overall workflow of the proposed framework.

identified from land-cover information and represented as independent spatial categories. Carriageway spaces are generated from OSM road centrelines through a buffering operation controlled by a width parameter (w_1), which transforms the linear transportation network into polygonal corridors. The generated carriageway regions are then, adjusted according to the parcel boundaries to ensure spatial coherence with the remaining urban entities.

The resulting entities are integrated within the cadastral framework to produce a structured representation of urban space. This representation preserves the semantic characteristics of the different urban categories. Figure 2 illustrates extraction of the principal urban entities and their integration into the resulting two-dimensional (2D) urban-space decomposition. The generated urban entities subsequently constitute the basis for the urban-space representations and the volumetric modelling process described in the following sections.

2.3 Multi-level urban space representation

To support urban modelling, the generated urban entities are organised into two urban-space representations:a general urban-space representation and a refined urban-space representation. Rather than focusing exclusively on geometric refinement, the proposed structure differentiates urban spaces according to their functional characteristics and their role within the urban environment.

In the general urban-space representation, the urban environment is represented through its primary components, including buildings, vegetation areas, water bodies, and the remaining open space. The remaining open space corresponds to the portion of the study area not occupied by buildings, vegetation,

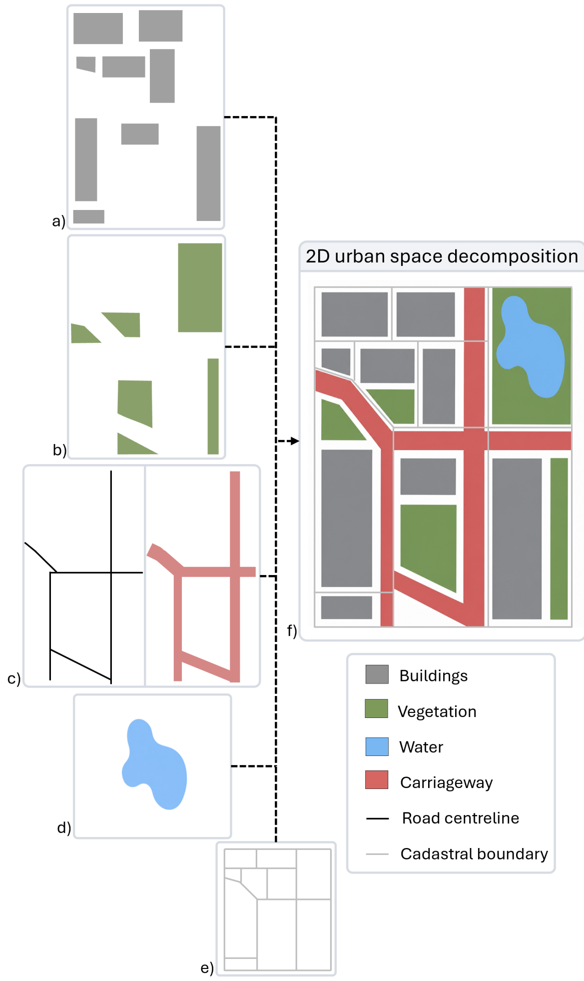

Figure 2. Urban-space decomposition. a) Building layer. b) Vegetation layer. c) Generation of carriageway polygons from the OSM road network. d) Water layer. e) Cadastral framework. f) Resulting 2D urban-space decomposition.

or water bodies. It may include transportation-related areas as well as other open urban spaces. This representation provides a simplified but semantically structured description suitable for characterising the overall organisation of urban space.

The refined urban-space representation introduces a more detailed characterisation of transportation-related areas. Carriageway regions are identified through a buffering operation controlled by a width parameter (w_1). The remaining open space after the extraction of carriageways constitutes the open space associated with the transportation environment, which may include pedestrian areas together with other urban spaces. This refinement provides a more detailed semantic representation of transportation-related areas.

2.4 CSG-based urban space modelling

The modelling process begins with the generation of a base volumetric representation of the study area using a study-area height parameter (H_{sa}). This base volume defines the spatial extent within which the urban entities are represented. Buildings are subsequently transformed into volumetric primitives using their corresponding footprints and a building-height parameter (H_b), derived from ALS and Digital Terrain Model

https://doi.org/10.5194/isprs-archives-L-4-W2-2026-201-2026 | © Author(s) 2026. CC BY 4.0 License.

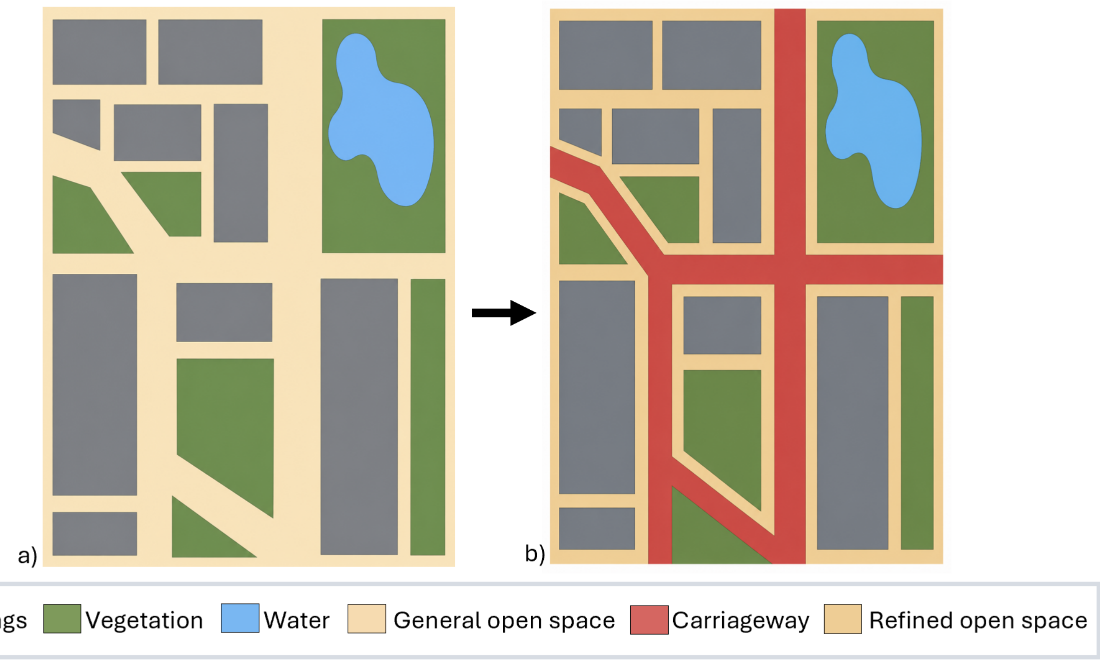

Figure 3. Transition from the general to the refined urban-space representation. a) General urban-space representation. b) Refined urban-space representation.

(DTM) information. Vegetation and water entities are represented using predefined height parameters (H_v) and (H_w), respectively. In the refined urban-space representation, carriageway regions are generated using a carriageway-height parameter (H_c).

The generated primitives are subsequently organised through Boolean *difference* operations to produce a volumetric representation of urban space. Buildings, vegetation, and water bodies are differentiated from the base study-area volume, resulting in a collection of semantically distinct volumetric entities. In the general urban-space representation, the remaining volume corresponds to the remaining open-space primitive after the removal of buildings, vegetation, and water bodies. In the refined urban-space representation, additional Boolean *difference* operations are applied to identify carriageway regions within the remaining open-space primitive. The resulting decomposition preserves the semantic distinction between urban-space categories while maintaining their spatial relationships within the model. Figure 4 illustrates the transformation from the multi-level urban-space representation to the corresponding CSG-based urban model.

Since the urban entities are derived from heterogeneous geospatial datasets, positional inconsistencies or overlapping volumes may occur after their transformation into 3D primitives. To ensure topological consistency, neighbouring volumetric primitives are therefore evaluated for potential geometric conflicts. *Intersection* tests are performed only between spatially adjacent entities, enabling the identification of unintended volumetric overlaps. Whenever overlaps are detected, Boolean *difference* operations are applied to remove the intersecting volumes and maintain the spatial integrity of the model. This validation process ensures that the final CSG representation consists of semantically distinct and geometrically consistent urban entities.

3. Results

3.1 Generated urban space models

The proposed workflow was applied to the study area to generate two urban-space representations. The resulting models integrate the extracted urban entities into semantically structured representations while preserving the spatial relationships between buildings, vegetation, transportation spaces, and water bodies.

Figure 5 presents the 2D urban-space representations. In the general representation, the urban environment is decomposed into buildings, vegetation, water bodies, and the remaining open space. The refined representation further characterises the remaining open space by identifying carriageway regions, while the remaining open space is preserved as a separate urban-space category.

The generated 2D representations constitute the input for the subsequent CSG-based modelling process. Figure 6 illustrates the corresponding 3D urban-space models derived from the general and refined urban-space representations. The resulting volumetric models preserve the semantic organisation established during the decomposition stage while providing a coherent 3D representation of the urban environment.

3.2 Urban-form analysis

The generated CSG urban-space models were further analysed to derive urban-form indicators describing both the spatial distribution and the volumetric characteristics of the built environment. While conventional 2D indicators provide a general description of the urban structure, the generated CSG model enables the computation of volumetric indicators that explicitly exploit the 3D representation.

As a 2D indicator, building coverage (I_b) is computed as

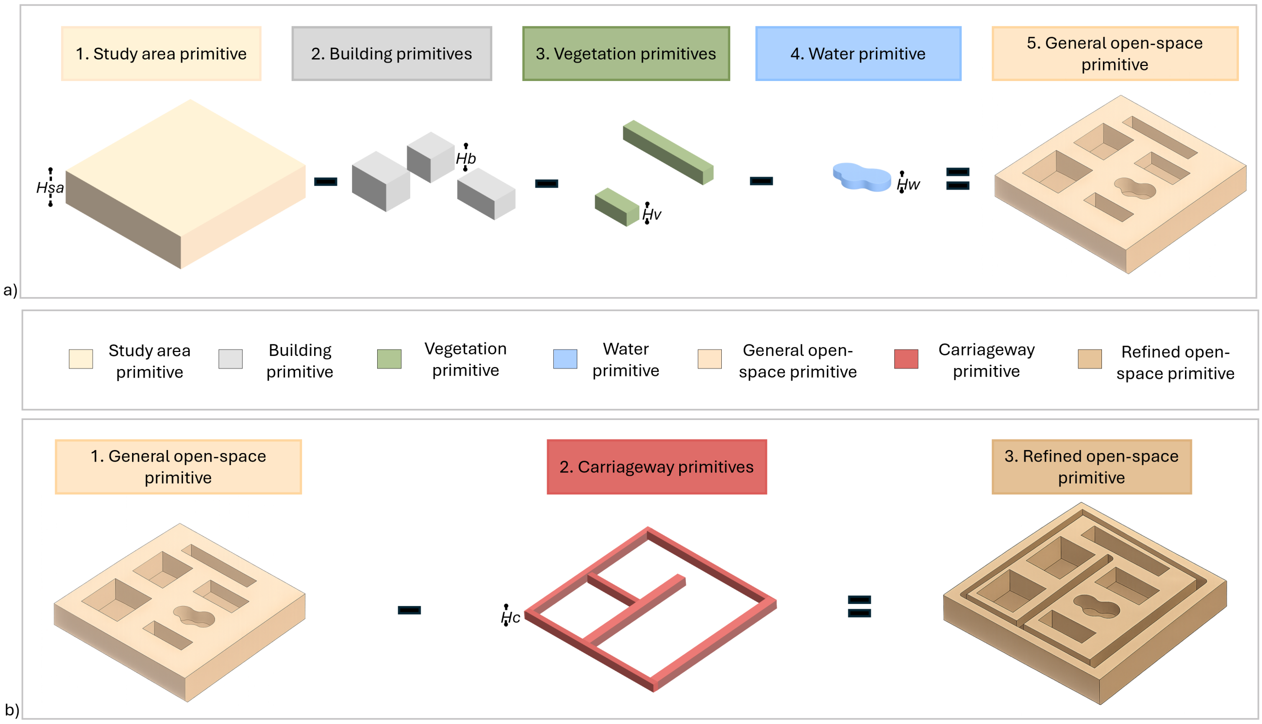

Figure 4. Generation of volumetric primitives through CSG Boolean *difference* operations. a) Derivation of the general open-space primitive. b) Derivation of the refined open-space primitive.

$$I_b = \frac{A_b}{A_{sa}} \times 100, \tag{1}$$

where A_b denotes the total building footprint area and A_{sa} represents the total study-area area.

To exploit the volumetric nature of the generated CSG model, the total built volume (V_b) is computed as

$$V_b = \sum_{j=1}^{n_b} V_{b,j} = \sum_{j=1}^{n_b} A_{b,j} H_{b,j}, \tag{2}$$

where $V_{b,j}$, $A_{b,j}$, and $H_{b,j}$ denote the volume, footprint area, and height of building j, respectively, and n_b is the total number of buildings.

To normalise this quantity with respect to the study-area size, the built-volume density (I_{bvd}) is computed as

$$I_{bvd} = \frac{V_b}{A_{sa}}. \tag{3}$$

The total built volume and built-volume density incorporate the vertical dimension of the urban environment. Consequently, these indicators demonstrate the capability of the proposed CSG representation to support 3D urban-form analyses that cannot be obtained directly from 2D datasets.

Table 2 summarises the urban-form indicators derived from the generated CSG urban-space models.

Indicator	Unit	Value
Building coverage	%	44.8
Total built volume	m^3	2,359,480
Built-volume density	m^3/m^2	7.47

Table 2. Urban-form indicators derived from the generated CSG urban-space models.

The computed indicators provide a quantitative description of the generated urban-space model. Buildings occupy 44.8% of the study area, indicating a compact built environment. More importantly, the generated CSG model contains a total built volume of 2,359,480 m^3, corresponding to a built-volume density of 7.47 m^3/m^2. These volumetric indicators exploit the 3D nature of the generated CSG model by combining building footprints with their corresponding heights, demonstrating the additional analytical capabilities provided by the proposed modelling framework.

4. Conclusion

The proposed framework demonstrates the potential of integrating airborne LiDAR, OpenStreetMap, and cadastral data to generate semantically structured urban-space models through a constructive solid geometry approach. By decomposing the urban environment into coherent spatial entities and organising them into two urban-space representations, the methodology enables the representation of buildings, vegetation, water bodies, and transportation-related areas within a unified volumetric model while preserving their spatial relationships. The generated models provide a structured representation of urban space suitable for visualisation, analysis, and interoperability with standard 3D modelling environments.

The results obtained for the Luxembourg case study illustrate

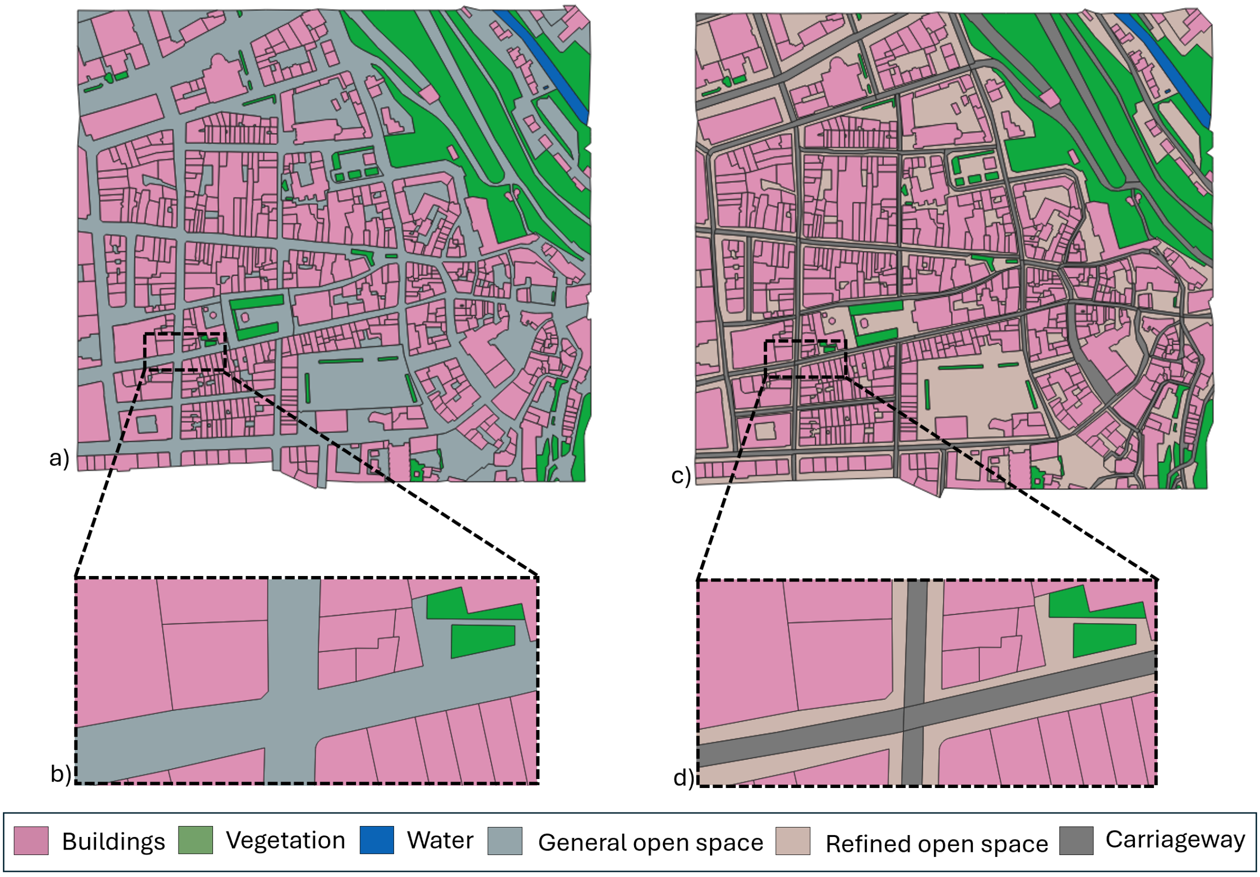

Figure 5. 2D urban-space representations. a) General urban-space representation. b) Zoom-in of the a). c) Refined urban-space representation. d) Zoom-in of the c).

the capability of the proposed workflow to produce coherent CSG urban-space models from heterogeneous geospatial datasets. The two urban-space representations provide progressively richer semantic descriptions of transportation-related areas while preserving the overall semantic organisation of the urban environment. Furthermore, the generated models support the derivation of urban-form indicators directly from the semantic representation, demonstrating their potential for the quantitative characterisation of urban environments.

Future work will focus on extending the framework to represent additional urban entities and more detailed levels of representation, as well as evaluating its applicability across different urban contexts and data sources. Further research will also investigate the use of the generated semantic CSG models for more advanced urban-form analyses and urban digital twin applications.

Acknowledgement

This work was partially supported by the project PID2024-157803OB-I00 funded by MICIU/AEI/10.13039/501100011033/ FEDER, UE. The statements made herein are solely the responsibility of the authors.

References

Albadri, M. S. A., González-Cabaleiro, P., Fernández, A., Díaz-Vilariño, L., 2025. Exploring Constructive Solid Geometry for Building Reconstruction from Point Clouds: Preliminary Results. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-4/W15-2025, 1–7.

Beil, C., Kolbe, T. H., 2024. Applications for Semantic 3D Streetspace Models and Their Requirements—A Review and Look at the Road Ahead. *ISPRS International Journal of Geo-Information*, 13(10).

Biljecki, F., Ledoux, H., Stoter, J., 2017. Generating 3D city models without elevation data. *Computers, Environment and Urban Systems*, 64, 1-18.

Caprari, G., Castelli, G., Montuori, M., Camardelli, M., Malvezzi, R., 2022. Digital Twin for Urban Planning in the Green Deal Era: A State of the Art and Future Perspectives. *Sustainability*, 14(10).

Chen, W., Huang, H., Liao, S., Gao, F., Biljecki, F., 2024. Global urban road network patterns: Unveiling multiscale planning paradigms of 144 cities with a novel deep learning approach. *Landscape and Urban Planning*, 241, 104901.

Dembski, F., Wössner, U., Letzgus, M., Ruddat, M., Yamu, C., 2020. Urban Digital Twins for Smart Cities and Citizens: The Case Study of Herrenberg, Germany. *Sustainability*, 12(6).

Ding, L., Xiao, G., Pano, A., Fumagalli, M., Chen, D., Feng, Y., Calvanese, D., Fan, H., Meng, L., 2025. Integrating 3D city data through knowledge graphs. *Geo-spatial Information Science*, 28(2), 780–799.

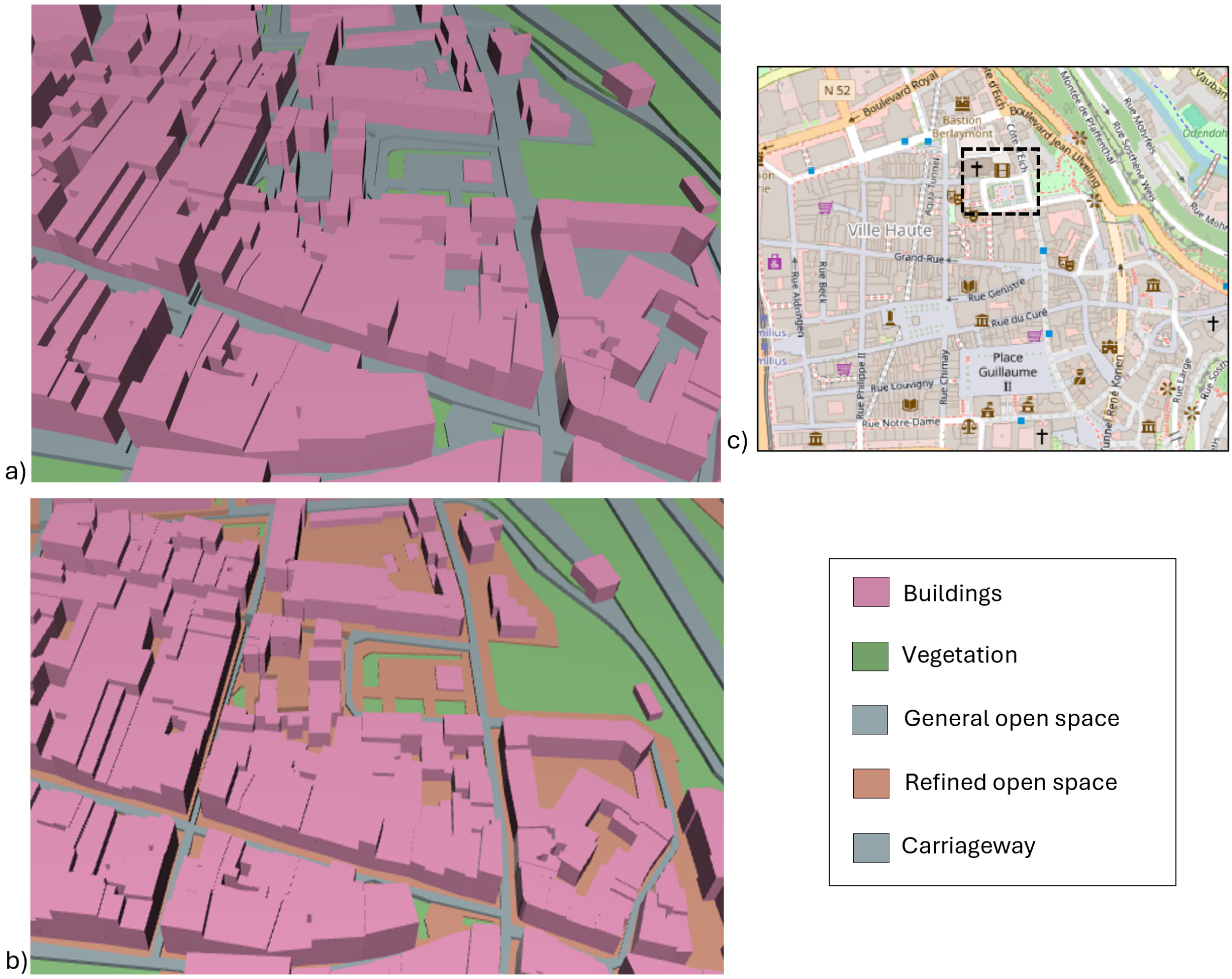

Figure 6. Generated CSG urban-space models. a) General urban-space representation. b) Refined urban-space representation. (c) Location of the study area.

Fayolle, P.-A., Friedrich, M., 2024. A Survey of Methods for Converting Unstructured Data to CSG Models. *Computer-Aided Design*, 168, 103655.

Fujiwara, K., Tsurumi, R., Kiyono, T., Fan, Z., Liang, X., Lei, B., Yap, W., Ito, K., Biljecki, F., 2026. VoxCity: A seamless framework for open geospatial data integration, grid-based semantic 3D city model generation, and urban environment simulation. *Computers, Environment and Urban Systems*, 123, 102366.

Gorte, B., Zlatanova, S., Pilouk, M., Diakite, A., Barton, J., 2024. 3D Data Integration in the Voxel Domain. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4-2024, 133–140.

Jeddoub, I., Nys, G.-A., Hajji, R., Billen, R., 2025. Data integration across urban digital twin lifecycle: a comprehensive review of current initiatives. *Annals of GIS*, 31(3), 367–386.

Kolbe, T. H., 2009. *Representing and Exchanging 3D City Models with CityGML*. Springer Berlin Heidelberg, Berlin, Heidelberg, 15–31.

Kolbe, T. H., Donaubauer, A., 2021. *Semantic 3D City Modeling and BIM*. Springer Singapore, Singapore, 609–636.

Kutzner, T., Chaturvedi, K., Kolbe, T. H., 2020. CityGML 3.0: New Functions Open Up New Applications. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 88, 43–61.

Labetski, A., Vitalis, S., Biljecki, F., Ohori, K. A., Stoter, J., 2023. 3D building metrics for urban morphology. *International Journal of Geographical Information Science*, 37(1), 36–67.

Ledoux, H., Biljecki, F., Dukai, B., Kumar, K., Peters, R., Stoter, J., Commandeur, T., 2021. 3dfier: automatic reconstruction of 3D city models. *Journal of Open Source Software*, 6(57), 2866.

Omrany, H., Mehdipour, A., Oteng, D., Al-Obaidi, K. M., 2025. The uptake of urban digital twins in the built environment: a pathway to resilient and sustainable cities. *Computational Urban Science*, 5(20).

Pađen, I., García-Sánchez, C., Ledoux, H., 2022. Towards automatic reconstruction of 3D city models tailored for urban flow simulations. *Frontiers in Built Environment*, Volume 8 - 2022.

Peters, R., Dukai, B., Vitalis, S., van Liempt, J., Stoter, J., 2022. Automated 3D Reconstruction of LoD2 and LoD1 Models for All 10 Million Buildings of the Netherlands. *Photogrammetric Engineering and Remote Sensing*, 88(3), 165–170.

Tan, Y., Liang, Y., Zhu, J., 2023. CityGML in the Integration of BIM and the GIS: Challenges and Opportunities. *Buildings*, 13(7).

Tsiranidou, E., Agugiaro, G., Fernández, A., Díaz Vilariño, L., 2025a. From point clouds to CityGML 3.0: An approach to multi-granular urban road modelling. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4/W6-2025, 201–208.

Tsiranidou, E., González-Cabaleiro, P., Fernández, A., Díaz-Vilariño, L., 2025b. A framework for road space extraction from point clouds and integration into 3D city models. *International Journal of Applied Earth Observation and Geoinformation*, 143, 104803.

White, G., Zink, A., Codecá, L., Clarke, S., 2021. A digital twin smart city for citizen feedback. *Cities*, 110, 103064.

Wu, Q., Xu, K., Wang, J., 2018. Constructing 3D CSG Models from 3D Raw Point Clouds. *Computer Graphics Forum*, 37(5), 221-232.

Ying, Y., Koeva, M. N., Kuffer, M., Zevenbergen, J. A., 2020. URBAN 3D MODELLING METHODS: A STATE-OF-THE-ART REVIEW. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B4-2020, 699–706.