Bridging Semantic Mesh, CityGML, and Gaussian Splatting for Urban Modelling and Visualization

Deni Suwardhi [1,3], Muhammad Arif Sudibyo [2], Agus Ambarwari [2], Wahyunan Andika [3], Muhammad Ihsan [2], Ratri Widyastuti [1], Widiatmoko Azis Fadilah [4], Arnadi Murtiyoso [4], Pierre Grussenmeyer [4], Fabio Remondino [5]

[1] Spatial System and Cadastral Research Group, Institut Teknologi Bandung (ITB), Indonesia – (deni.suwardhi, ratri)@itb.ac.id
[2] Postgraduate Programmes, Institut Teknologi Bandung (ITB), Indonesia – (25124002, 35121002, 33222026)@mahasiswa.itb.ac.id
[3] PT Inovasi Mandiri Pratama, Spatial Information Company, Indonesia - wahyunan.andika@inovamap.com
[4] Université de Strasbourg, CNRS, INSA Strasbourg, ICube Laboratory UMR 7357, Photogrammetry and Geomatics Group, 67000, Strasbourg, France– (widiatmoko-azis.fadilah, arnadi.murtiyoso, pierre.grussenmeyer)@insa-strasbourg.fr
[5] 3D Optical Metrology (3DOM) Unit, Bruno Kessler Foundation (FBK), Trento, Italy – remondino@fbk.eu

Keywords: Semantic Mesh, CityGML, Gaussian Splatting, UAV, Photogrammetry, Visualization.

Abstract

Urban digital twin systems require 3D city representations that reconcile semantic structure, geometric reliability, simulation capability, and photorealistic real-time rendering. Existing approaches usually prioritize a single modelling paradigm, limiting their ability to support both analytical and visualization needs. CityGML provides standardized semantics and topology but often lacks surface realism. Semantic mesh models preserve geometric detail suitable for environmental simulations but provide limited hierarchical semantics. In contrast, neural radiance-field approaches such as 3D Gaussian Splatting (3DGS) enable photorealistic rendering at interactive frame rates but do not explicitly encode topology or structured semantics. This study establishes a comparative framework linking LiDAR-derived CityGML, semantic mesh, 3D Gaussian Splatting, and Triangle Splatting within a unified urban modelling workflow. UAV data acquired using a DJI ZENMUSE L2 sensor serve as the geometric backbone for reconstructing CityGML LoD1–LoD2 models. The semantic model is transformed into a textured triangular mesh, while radiance-based models are generated from the same imagery using multiple 3DGS implementations and a triangle splatting framework. Comparative evaluation investigates geometric coherence, semantic preservation, and radiance consistency to identify structural correspondences across the representations. The results reveal complementary modelling layers that can be systematically mapped rather than treated as competing alternatives. Based on these findings, the paper proposes a conceptual foundation for a unified 3D urban model capable of transforming consistently into semantic, surface-based, and radiance-based representations for adaptive urban digital twin systems. Data are freely accessible for research purposes at https://github.com/3DOM-FBK/urban-representation-fusion/.

1. Introduction

The rapid development of urban digital twin technologies has significantly increased the demand for three-dimensional (3D) city representations capable of supporting multiple analytical and visualization tasks simultaneously. Urban digital twins are expected to support semantic interpretation of urban objects, geometric analysis, numerical simulations, and realistic visualization within a single integrated environment (Ketzler et al., 2020). However, in practice, no single data representation is currently able to satisfy all these requirements simultaneously. As a result, urban modelling workflows typically rely on multiple heterogeneous models, each optimized for specific tasks but rarely examined within a unified and interoperable framework.

One of the most widely adopted standards for semantic 3D city modelling is CityGML, which provides structured object hierarchies, topological relationships, and thematic attributes for representing urban environments (Kolbe, 2009; Gröger and Plümer, 2012). The level-of-detail (LoD) concept in CityGML enables scalable representations ranging from simplified block models to detailed building components. Due to its structured semantics and explicit topology, CityGML has been widely applied in urban planning, spatial analysis, infrastructure management, and administrative information systems. Nevertheless, CityGML is primarily designed to represent explicit geometric structures and semantic relationships rather than surface radiance characteristics. Consequently, it is not optimized for photorealistic real-time visualization at the scale of entire urban environments.

In contrast, surface-based representations such as dense triangle meshes reconstructed from UAV photogrammetry or airborne LiDAR provide highly detailed geometric descriptions of urban surfaces. These models preserve fine geometric structures of buildings, vegetation, and terrain that are often simplified in semantic city models. When enriched with semantic labels, semantic mesh models can serve as input for various environmental and urban analyses, including airflow simulation, microclimate modelling, and solar radiation analysis. Despite their geometric richness, however, mesh representations typically lack standardized hierarchical semantic structures and formal topological relationships comparable to those available in CityGML. This limitation complicates their integration with structured urban information models used in digital twin systems.

More recently, neural rendering approaches have emerged as a new paradigm for representing 3D environments. Among these approaches, 3D Gaussian Splatting (3DGS) has attracted considerable attention due to its ability to generate real-time photorealistic visualization through anisotropic Gaussian-based radiance representations (Kerbl et al., 2023). This method allows immersive exploration of complex urban scenes with high visual realism and efficient rendering performance. However, radiance-based representations do not explicitly encode geometric structures or semantic information, making them less suitable for spatial analysis, simulation, or structured querying. Recent developments such as Triangle Splatting by Held et al. (2025) attempt to incorporate triangle primitives into neural radiance modelling, suggesting a possible intermediate representation between structured meshes and unstructured Gaussian splats.

These differing characteristics indicate that urban digital twin systems cannot rely on a single representation paradigm. Instead, they require approaches that enable multiple representations, including semantic models, surface-based meshes, and radiance-based visualizations, to coexist and complement each other when derived from the same dataset. This need is consistent with the concept of urban digital twins, which emphasizes the integration of heterogeneous data sources and modelling paradigms to support both analytical and visualization tasks within a unified digital environment (Lehtola et al., 2022).

Although various paradigms for 3D urban modelling have rapidly evolved, many approaches remain developed within relatively separate research communities. These communities often emphasize different modelling priorities, leading to fragmentation in digital urban modelling. Consequently, each paradigm optimizes different aspects, including semantics, geometry, or visual appearance, making their integration within urban digital twin systems challenging. Despite these advances, systematic studies that explicitly examine the relationships between these modelling paradigms within a unified framework remain limited.

This study explores the concept of a unified urban data representation, defined as a core representation capable of transforming into multiple modelling paradigms. In this context, the representation should support transformation into structured urban models (CityGML), semantically enriched meshes, and visual radiance-field representations. To investigate this idea, we develop a comparative modelling framework that generates several urban representations from the same UAV dataset consisting of LiDAR point clouds and aerial imagery. By systematically comparing these representations in terms of geometric consistency, semantic preservation, and visual appearance, the study aims to identify structural relationships between modelling paradigms as a foundation for developing transformable urban data models for digital twin applications.

2. Related Works

2.1 Semantic 3D City Models and Urban Digital Twins

Semantic 3D city models have become an essential component in the development of urban information infrastructures and digital twin systems. One of the most widely adopted standards for representing such models is CityGML, which provides a semantic data model for describing urban objects including buildings, transportation networks, vegetation, and terrain (Kolbe, 2009; Gröger and Plümer, 2012). The level-of-detail (LoD) concept supports scalable representations from simple volumetric models to detailed building structures, while CityGML 3.0 introduces improved support for dynamic objects and interoperability (Kutzner et al., 2020).

To address complexity and implementation challenges associated with the XML-based encoding of CityGML, alternative formats such as CityJSON have been developed. CityJSON provides a compact and easy-to-use JSON-based encoding of the CityGML data model, facilitating data exchange, web integration, and software development (Ledoux et al., 2019). These formats play complementary roles in supporting scalable and web-based urban digital twin systems.

Alongside these developments, the generation of 3D city models has also been widely studied, particularly from airborne LiDAR and photogrammetric data. Existing approaches can be broadly categorized into model-driven and data-driven methods, as well as hybrid and learning-based approaches, each with different trade-offs in terms of geometric accuracy, robustness, and flexibility. Data-driven methods enable detailed reconstruction from point clouds, while model-driven approaches provide more structured and efficient representations but rely on predefined geometric templates (Jarząbek-Rychard et al., 2025).

Recent research highlights the importance of semantic 3D city models in supporting integrated simulations, decision-making, and urban management within digital twin environments (Lehtola et al., 2022; Kasprzyk et al., 2024; Kang, 2024). However, such models typically rely on simplified geometric representations and are not optimized for detailed surface reconstruction or photorealistic visualization. This limitation motivates the integration of semantic models with alternative representations, such as surface meshes and radiance-based models, within multi-representation urban modelling frameworks.

2.2 Semantic Mesh Representations

While semantic city models focus on structured representation of urban objects, surface-based mesh models reconstructed from photogrammetry or LiDAR provide more detailed geometric descriptions of urban environments. Dense triangle meshes can capture the complex surface geometry of buildings, vegetation, and terrain with high geometric fidelity.

To enhance their analytical value, several studies have explored semantic enrichment of mesh models. For example, Gao et al. (2021) introduced the SUM dataset, a benchmark dataset for semantic segmentation of urban meshes that supports machine learning applications and urban scene understanding. By assigning semantic labels directly to mesh surfaces, semantic mesh models enable object-level analysis while preserving geometric detail. In addition, recent benchmark datasets such as the Hessigheim 3D (H3D) dataset further demonstrate the potential of combining high-resolution UAV LiDAR point clouds and textured meshes for semantic segmentation and multi-modal analysis (Kölle et al., 2021).

Hybrid approaches combining mesh geometry and semantic city models have also been investigated. Willenborg et al. (2018) demonstrated that integrating mesh-based surface representations with semantic city models can improve the accuracy of solar potential analyses. These studies suggest that mesh models can complement semantic city models by providing higher geometric resolution for simulation tasks.

However, mesh representations typically lack standardized hierarchical schemas and formal topological structures comparable to those defined in CityGML. As a result, interoperability between mesh models and semantic city models remains an open challenge in urban modelling workflows.

2.3 Neural Radiance Representations for Urban Scenes

Recent advances in neural rendering have introduced new methods for representing 3D environments using radiance fields rather than explicit geometric surfaces. Among these methods,

https://doi.org/10.5194/isprs-archives-XLIX-B4-2026-81-2026

3D Gaussian Splatting (3DGS) has gained significant attention due to its ability to produce photorealistic rendering with real-time performance (Kerbl et al., 2023). By representing scenes using anisotropic Gaussian primitives, 3DGS enables efficient rendering of complex environments while preserving high visual fidelity. Recent studies have also highlighted the potential of 3DGS for urban-scale reconstruction, while emphasizing challenges related to scalability, geometric accuracy, and evaluation in complex urban environments (Yan et al., 2025).

Subsequent studies have explored extensions and improvements of the Gaussian splatting paradigm. For example, Jiang et al. (2025) proposed integrating LiDAR observations into Gaussian splatting frameworks to improve reconstruction accuracy for large-scale environments. Open-source libraries such as gsplat have further facilitated the adoption of Gaussian splatting methods for research and practical applications (Ye et al., 2024). Other research efforts have addressed scalability issues by optimizing memory management and GPU utilization for large radiance field models (Zhao et al., 2025).

Several software implementations have also emerged to support practical workflows for Gaussian splatting reconstruction and visualization. Implementations such as LichtFeld Studio and Postshot enable neural rendering pipelines from photogrammetric datasets (LichtFeld Studio, 2025; Jawset, 2024).

Despite their impressive visualization capabilities, radiance-based representations typically do not explicitly encode geometric surfaces or semantic information. This limitation restricts their use for analytical modelling tasks.

2.4 Hybrid Radiance–Geometry Representations

To bridge the gap between geometric modelling and radiance-based rendering, recent research has explored hybrid representations and integration strategies that combine surface primitives with radiance field concepts. One such approach is Triangle Splatting, which introduces triangle primitives into radiance-based rendering frameworks (Held et al., 2025). By incorporating explicit geometric elements into neural rendering pipelines, triangle splatting provides a potential intermediate representation between traditional mesh models and purely radiance-based representations such as Gaussian splats.

More recent developments have further extended this direction by introducing alternative hybrid formulations. For instance, Convex Splatting represents scenes using smooth convex primitives for radiance rendering, while MeshSplatting enables differentiable rendering directly on mesh-based surfaces, strengthening the integration between geometric models and radiance-based representations (Held et al., 2025; Held et al., 2026). In addition, approaches such as MiLo demonstrate how radiance-based models can be coupled with mesh generation processes, enabling bidirectional interaction between neural rendering and explicit geometry reconstruction.

These emerging approaches suggest new opportunities not only for representation, but also for tighter integration between geometric modelling and neural rendering techniques within urban visualization pipelines. However, systematic investigations that compare semantic models, mesh representations, and radiance-based methods within a unified modelling and integration framework remain limited.

In this context, the present study investigates how CityGML, Semantic Mesh, and radiance-based representations can be generated from a common dataset and comparatively analysed in terms of geometric structure, semantic preservation, and visualization performance.

3. Data and Methodology

This study establishes a comparative workflow to generate and evaluate multiple 3D urban representations derived from a common UAV dataset. The workflow is designed as an initial step toward exploring the concept of a unified urban data representation. The methodology integrates LiDAR-based semantic modelling, surface-based mesh generation, and neural radiance-field reconstruction. By deriving CityGML, Semantic Mesh, 3D Gaussian Splatting, and Triangle Splatting models from the same dataset, the framework enables a systematic comparison of geometric consistency, semantic structure, and visualization performance. The overall workflow is illustrated in Figure 1 and consists of six main stages: data acquisition and preprocessing, CityGML reconstruction, semantic mesh generation, neural radiance-field modelling, cross-representation comparison, and structural analysis toward a unified urban representation.

3.1 Data Acquisition and Pre-processing

The dataset was acquired using a DJI M350 UAV equipped with a DJI Zenmuse L2 sensor, which integrate a LiDAR and a camera, enabling simultaneous capture of LiDAR point clouds and aerial imagery. The survey was conducted over the Institut Teknologi Bandung (ITB) campus area in Bandung, West Java, Indonesia, covering approximately 25 hectares. The LiDAR data provide a geometrically reliable representation of urban structures, while the images support photogrammetric reconstruction and neural rendering.

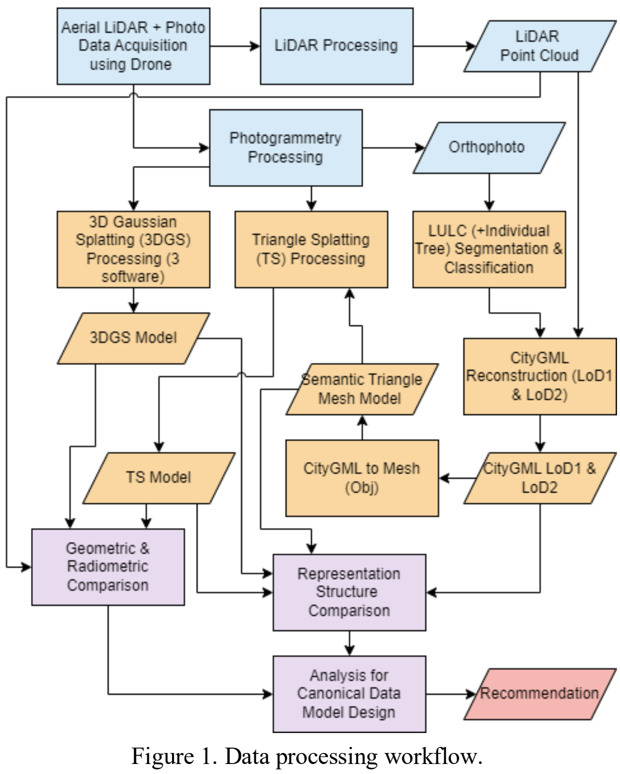

Figure 1. Data processing workflow.

The LiDAR point cloud processing produces classified data representing buildings, terrain, and vegetation. In parallel, photogrammetric processing of aerial images generates camera exterior parameters and orthophotos that are subsequently used for neural radiance modelling and landcover map generation, respectively. The image orientation was supported by onboard GNSS/IMU data processed using a post-processed kinematic (PPK) approach, providing an accurate UAV trajectory that was used as input for bundle adjustment. The georeferencing accuracy was further evaluated and refined using ground control derived from RTK GNSS measurements through an ICP-based alignment procedure. The combined dataset serves as a consistent geometric and radiometric foundation for the generation of multiple urban representations.

3.2 CityGML Reconstruction from LiDAR Data

CityGML models are reconstructed from the processed LiDAR point cloud to provide a structured semantic representation of the study area. Building models are generated at LoD1 and LoD2 levels. LoD1 reconstruction is performed using the 3dfier software, which automatically extrudes building footprints and generates block models from point cloud data (Ledoux et al., 2021).

For detailed roof modelling at LoD2, the City3D framework is employed to reconstruct building geometry from airborne LiDAR data (Huang et al., 2022). In addition to building reconstruction, individual trees are modelled using parametric approaches derived from LiDAR and photogrammetric point clouds following the method proposed by Suwardhi et al. (2022). The resulting CityGML dataset therefore includes semantic objects representing buildings and vegetation elements.

3.3 Semantic Mesh Generation

To enable a consistent comparison between different urban representations, the semantic mesh used in this study is derived from the CityGML dataset. The CityGML geometry is converted into a triangle-based mesh using the citygml2obj conversion process (Biljecki et al., 2015), ensuring that mesh surfaces preserve the object-level structure and semantic correspondence of the original CityGML model.

Although dense surface models could also be generated directly from photogrammetric point clouds, such representations typically lack explicit object-level structure and consistent semantic correspondence with CityGML. For this reason, photogrammetric meshes are not used as the primary representation in the comparative analysis, as they would not allow a controlled evaluation across semantic and radiance-based models.

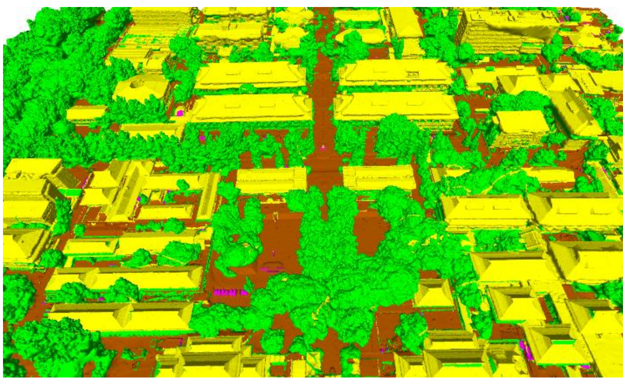

Figure 2. Semantically segmented mesh model of the surveyed urban area.

For illustration purposes, a semantic mesh generated using the SUM framework from photogrammetric data of the ITB campus is presented in Figure 2. This example highlights the capability of segmentation-based approaches to capture detailed surface geometry, but is not included in the core comparative evaluation.

3.4 Neural Radiance Modelling

Neural radiance representations are generated using the available aerial imagery. Two radiance modelling approaches are implemented in this study.

1. First, 3D Gaussian Splatting (3DGS) models are reconstructed using three software implementations: LichtFeld Studio, CLM-GS, and Postshot.
2. Second, Triangle Splatting is applied to generate a radiance representation based on triangle primitives (Held et al., 2025). In this approach, triangle surfaces derived from mesh geometry are optimized with radiance parameters, enabling a surface-aware neural rendering representation.

All radiance models are reconstructed using the same set of aerial imagery and camera parameters obtained from the UAV data processing. This ensures consistent input data for comparing radiance-based representations with semantic city models and surface-based mesh models.

3.5 Representation Comparison Framework

To evaluate the generated urban models, this study applies a representation comparison framework considering three aspects: geometric consistency, radiometric fidelity, and semantic structure. These aspects reflect the fundamental differences between the modelling paradigms investigated in this research, namely semantic city models (CityGML), surface-based representations (semantic mesh), and radiance-based models such as 3D Gaussian Splatting and Triangle Splatting.

The geometric comparison assesses how accurately each representation reconstructs the spatial structure of the scene. The LiDAR point cloud is used as the geometric reference, as it provides a direct and high-resolution measurement of the scene geometry. Other representations are analysed relative to this reference to evaluate their ability to reproduce building geometry, roof structures, and vegetation elements.

Radiometric comparison evaluates visual fidelity by comparing rendered images from each representation with the original UAV imagery. For viewpoints corresponding to UAV camera poses, rendered images are compared with the photographs using PSNR, SSIM, and LPIPS, which measure pixel-wise fidelity, structural similarity, and perceptual similarity.

Finally, semantic comparison examines how urban objects are represented within each modelling paradigm. CityGML provides explicit hierarchical object structures, semantic meshes retain geometry with limited semantic labels, while radiance-based models mainly focus on visual reconstruction without explicit object semantics. Together, these comparisons provide an overall evaluation of the strengths and limitations of each representation.

4. Results and Discussion

4.1 Generation of Multi-Representation Urban Models

Using the workflow described in Section 3, four different representations of the study area were generated from the same UAV dataset: CityGML models, semantic triangle meshes, 3D

Gaussian Splatting (3DGS) models, and Triangle Splatting models.

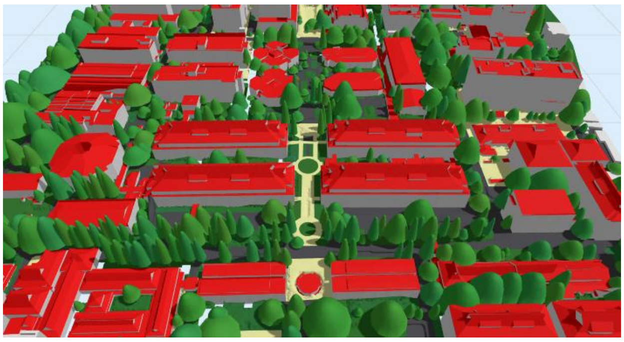

Figure 3. CityGML models generated from LiDAR data.

Figure 4. Instance-level semantic mesh derived from CityGML through citygml2obj conversion.

Figure 5. Illustration of the 3DGS results.

The resulting CityGML model therefore provides a semantically structured representation of the urban environment, including buildings and vegetation components, as illustrated in Figure 3. However, the geometric surfaces are relatively simplified compared to dense surface models derived from photogrammetry or mesh reconstruction.

To enable surface-based analysis, the CityGML geometry was subsequently converted into a semantic triangle mesh using a CityGML-to-OBJ conversion workflow. The resulting instance-level semantic mesh is illustrated in Figure 4. Because the mesh is derived from the simplified CityGML surfaces, the resulting geometry remains less detailed than meshes directly reconstructed from LiDAR or photogrammetric point clouds. However, this approach enables the generation of instance-segmented mesh objects, which are not directly available in segmentation-based mesh frameworks such as SUM. As explained in Section 3.3, this strategy allows semantic mesh instances corresponding to buildings and vegetation objects to be obtained while maintaining compatibility with the CityGML structure.

Figure 5 illustrates the 3D Gaussian Splatting (3DGS) reconstruction results generated from the UAV imagery dataset. The radiance-based representation successfully preserves photorealistic appearance and fine visual details of the urban environment while enabling interactive real-time rendering. Compared to the semantic and mesh-based representations, the 3DGS model emphasizes visual realism and view-dependent appearance rather than explicit topology or semantic structure.

4.2 Visual and Radiometric Characteristics

The results of this evaluation provide insights into the trade-offs between geometric representation and visual realism, which is a key consideration for the development of urban digital twin systems that must support both analytical tasks and immersive visual exploration.

Figure 6. Rendered views from identical input camera poses: (a, top-left) Postshot (Splat3), (b, top-right) Lichtfeld Studio (MRNF), (c, middle-left) CLM-GS, (d, middle-right) Triangle Splatting, (e, bottom-left) Photorealistic Mesh, and (f, bottom-right) References.

Table 1. Radiometric similarity between rendered images and UAV reference images.

Representation	PSNR ↑	SSIM ↑	LPIPS ↓
3DGS – Postshot (Splat3)	20.52	0.685	0.046
3DGS – LichtFeld Studio (MRNF)	19.77	0.641	0.094
3DGS – CLM-GS	18.76	0.622	0.180
Triangle Splatting	18.73	0.598	0.118
PhotoRealistic Mesh	15.42	0.527	0.135

While the visual similarity metrics provide insights into the radiometric quality of the reconstructed scenes, it is equally important to assess the geometric accuracy of the models. Therefore, the following section evaluates the geometric consistency of the 3D Gaussian Splatting reconstructions by comparing them with the LiDAR-derived reference point cloud.

https://doi.org/10.5194/isprs-archives-XLIX-B4-2026-81-2026

The quantitative results reported in Table 1 reveal clear differences in visual reconstruction quality across the evaluated representations. Radiance-based approaches generally achieve higher perceptual similarity with the reference UAV imagery (see Figure 6) compared with surface-based models. In particular, the 3D Gaussian Splatting reconstructions tend to produce higher PSNR and SSIM values, indicating improved radiometric fidelity and structural similarity relative to the original images. This behaviour is expected, as radiance-field representations directly optimise image appearance during training.

Among the evaluated implementations, the differences between LichtFeld Studio, CLM-GS, and Postshot are relatively small in terms of overall visual similarity, although minor variations can be observed depending on the reconstruction strategy and optimisation pipeline used by each software system. In general, these approaches demonstrate the capability of radiance-based representations to reproduce complex visual characteristics of urban environments, including vegetation textures, material reflectance, and lighting variations.

In contrast, the semantic mesh representation shows slightly lower radiometric similarity metrics. Although the mesh model preserves detailed geometric surfaces, its visual quality depends on the generated texture maps and does not explicitly model view-dependent radiance effects. As a result, differences between the rendered mesh images and the original UAV photographs may appear in areas with strong illumination changes or complex surface reflectance.

Triangle Splatting provides an intermediate behaviour between purely radiance-based and surface-based representations. By associating radiance parameters with triangular primitives, this representation partially retains geometric structure while producing visually realistic rendering. The resulting metrics therefore tend to lie between those obtained from semantic mesh models and full Gaussian Splatting representations.

Overall, these results highlight the complementary characteristics of the evaluated modelling paradigms. Radiance-based models excel in reproducing visual realism and immersive scene appearance, while mesh-based and semantic models remain essential for representing explicit geometry and structured urban objects. This trade-off between visual fidelity and structural representation further supports the need for integrated modelling frameworks capable of linking multiple representations within urban digital twin environments.

4.3 Geometric Consistency of 3DGS Reconstructions

To quantitatively evaluate the geometric consistency of radiance-based reconstructions, the point sets derived from the 3D Gaussian Splatting (3DGS) models were compared with the LiDAR reference point cloud using the Multiscale Model-to-Model Cloud Comparison (M3C2) distance metric. The comparison was conducted for three different 3DGS implementations: LichtFeld Studio, CLM-GS, and Postshot.
The statistical results of the M3C2 distance measurements are summarized in Table 2.

The mean distance values for all three reconstructions are close to zero, indicating that the radiance-based models do not exhibit significant systematic geometric bias relative to the LiDAR reference dataset. This suggests that the overall geometric structure of the urban scene is preserved during the radiance-field reconstruction process.

Table 2. Geometric comparison between LiDAR point clouds and 3D Gaussian Splatting reconstructions.

3DGS Implementation	Mean Distance (m)	Std Dev (m)	Histogram Classes
3DGS-LichtFeld Studio (MRNF)	0.072	0.903	23,783
3DGS-CLM-GS	-0.134	1.677	21,832
3DGS-Postshot (Splat3)	0.009	0.903	23,533
Triangle Splatting	0.11	1.689	12,659

However, differences in geometric dispersion can be observed between the tested implementations. The Postshot reconstruction shows the smallest standard deviation (0.90 m), indicating the highest geometric consistency with the LiDAR dataset. The LichtFeld Studio reconstruction exhibits moderate dispersion (0.90 m), while CLM-GS produces the largest variance (1.67 m).

These results indicate that although 3D Gaussian Splatting is primarily designed for photorealistic visualization rather than geometric reconstruction, the resulting models can still approximate the underlying LiDAR geometry with reasonable consistency. Cross-sectional comparisons of the reconstructed models in Figure 7 further demonstrate that different radiance-based implementations preserve the overall structural configuration of buildings and terrain despite variations in surface smoothness and geometric sharpness. This finding supports the potential integration of radiance-based representations with geometry-based models within urban digital twin frameworks.

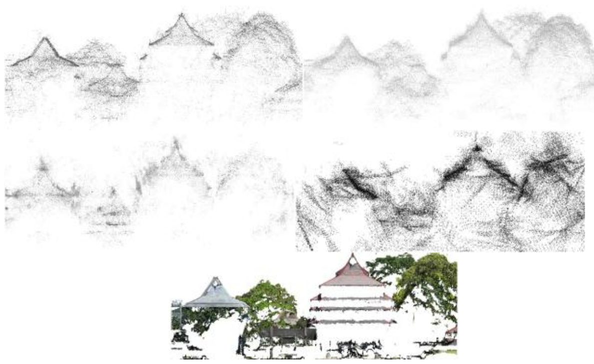

Figure 7. Cross-sectional views of the reconstructed models: (a) Postshot (Splat3), (b) Lichtfeld Studio (MRNF), (c) CLM-GS, (d) Triangle Splatting, and (e, bottom-center) photorealistic mesh (dense point cloud).

4.4 Relationship Between Semantic, Mesh, and Radiance Representations

The comparative experiments conducted in this study highlight fundamental differences between semantic, surface-based, and radiance-based representations, while also revealing their structural relationships.

CityGML represents urban environments through structured semantic hierarchies and explicit topology. Semantic mesh models, on the other hand, preserve detailed surface geometry derived from LiDAR or photogrammetric reconstruction, while maintaining links to semantic entities when object identifiers are retained. In contrast, radiance-based representations such as 3D

Gaussian Splatting model the scene using radiance primitives, enabling photorealistic visualization without explicitly encoding surface geometry or semantic structures.

Despite these differences, the results indicate that the three representations are not mutually exclusive but can be interpreted as complementary layers describing the same urban environment. The LiDAR-derived geometry provides a common spatial reference that ensures geometric consistency across representations.

This relationship suggests that semantic, geometric, and radiance-based representations can be integrated within a unified modelling framework. In such a framework, mesh representations serve as a geometric backbone due to their ability to explicitly encode surface structures while maintaining compatibility with both semantic models and radiance-based approaches.

Rather than replacing one paradigm with another, the findings indicate that future urban digital twin systems may benefit from morphable representations, where semantic, mesh, and radiance models can be transformed or synchronized depending on the requirements of analysis, simulation, or visualization.

4.5 Implications for Unified Urban Data Models

The identified relationships between semantic, mesh, and radiance representations imply that urban digital twin systems require an integrated modelling approach rather than relying on a single representation paradigm.

Among possible integration strategies, surface-based representations, particularly multiresolution semantic meshes, emerge as a promising intermediate layer. Mesh structures explicitly encode surface geometry, can be enriched with semantic information, and provide a suitable geometric basis for radiance-based modelling. This makes them effective in bridging semantic models such as CityGML and radiance-based visualization approaches.

The transformation between representations can be understood as a sequential and interoperable process. Semantic mesh acts as a structured geometric layer that preserves detailed surface information. Through semantic enrichment, it can be transformed into semantic representations such as CityGML or CityJSON, while also serving as a geometric foundation for radiance-field approaches such as 3D Gaussian Splatting.

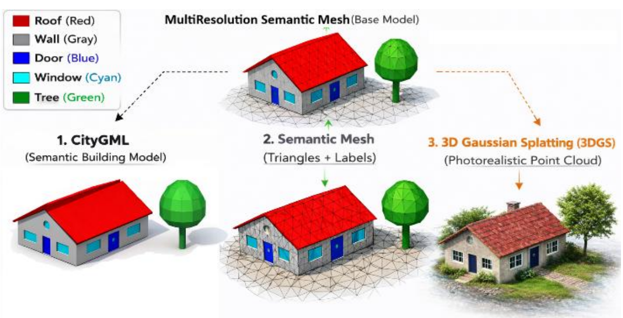

Figure 8. Proposed unified urban data model.

This concept is illustrated in Figure 8, where mesh representation forms the geometric backbone, semantic models provide structured object-level information, and radiance-based models support high-quality visualization within a unified framework.

Although a fully unified data structure has not yet been implemented in this study, the results provide initial evidence that such an integrated approach is feasible. This opens opportunities for future research on data structures, transformation mechanisms, and database architectures that enable seamless interoperability between semantic, geometric, and radiance-based representations in urban digital twin systems.

5. Conclusions

This study investigated the relationships between three major paradigms of 3D urban representation: semantic mesh representations, semantic city models, and radiance-based neural representations. Using a UAV dataset consisting of LiDAR point clouds and aerial imagery, four representations of the same urban environment were generated and analyzed: CityGML models, semantic triangle meshes, 3D Gaussian Splatting reconstructions, and Triangle Splatting representations.

The results demonstrate that each modelling paradigm provides distinct advantages for specific tasks within urban digital twin systems. CityGML offers structured semantic hierarchies and explicit topology that support urban analysis and information management. Semantic mesh representations preserve detailed surface geometry derived from LiDAR or photogrammetric reconstruction, enabling surface-based simulations and geometric analysis. In contrast, radiance-based approaches such as 3D Gaussian Splatting provide highly realistic visualization of urban environments through radiance field representations.

Quantitative comparison between LiDAR point clouds and 3DGS reconstructions using the M3C2 distance metric indicates that radiance-based models can maintain reasonable geometric consistency with the underlying geometric dataset, despite not explicitly storing surface geometry. At the same time, structural analysis highlights fundamental differences between semantic object models, triangle-based surface representations, and radiance-field primitives.

Rather than viewing these modelling paradigms as competing approaches, the results suggest that they should be considered complementary components of urban digital twin environments. The comparative analysis presented in this study indicates that surface-based representations, particularly semantic multiresolution meshes, may provide a promising intermediate structure capable of linking semantic models such as CityGML with radiance-based visualization models.

Based on these findings, the study suggests the potential development of morphable urban data models, where semantic, geometric, and radiance-based representations can be derived from a shared structure depending on application requirements. Such an approach could enable more flexible and interoperable urban digital twin systems capable of simultaneously supporting analysis, simulation, and immersive visualization.

Future work will investigate the design of unified data structures, database architectures, and transformation mechanisms that allow seamless interoperability between CityGML, semantic mesh representations, and radiance-based models within large-scale urban environments.

References

Biljecki, F., Arroyo Ohori, K., 2015. Automatic semantic-preserving conversion between OBJ and CityGML. Eurographics Workshop on Urban Data Modelling and Visualisation, 25–30.

Gao, W., Nan, L., Boom, B., Ledoux, H., 2021. SUM: A benchmark dataset of semantic urban meshes. ISPRS Journal of Photogrammetry and Remote Sensing, 179, 108–120.

Gröger, G., Plümer, L., 2012. CityGML – Interoperable semantic 3D city models. ISPRS Journal of Photogrammetry and Remote Sensing, 71, 12–33.

Held, J., Vandeghen, R., Deliege, A., Hamdi, A., Cioppa, A., Giancola, S., Vedaldi, A., Ghanem, B., Tagliasacchi, A., Van Droogenbroeck, M., 2025. Triangle splatting for real-time radiance field rendering. arXiv preprint, arXiv:2505.19175.

Held, J., Vandeghen, R., Hamdi, A., Deliege, A., Cioppa, A., Giancola, S., Vedaldi, A., Ghanem, B., Van Droogenbroeck, M., 2025. 3D Convex Splatting: Radiance Field Rendering with 3D Smooth Convexes. Proc. CVPR.

Held, J., Son, S., Vandeghen, R., Rebain, D., Gadelha, M., Zhou, Y., Cioppa, A., Lin, M., Van Droogenbroeck, M., Tagliasacchi, A., 2026. MeshSplatting: Differentiable Rendering with Opaque Meshes. Proc. CVPR.

Huang, J., Stoter, J., Peters, R., Nan, L., 2022. City3D: Large-scale building reconstruction from airborne LiDAR point clouds. Remote Sensing, 14(9), 2254.

Jarząbek-Rychard, M., Rigon, S., Boguslawski, P., and Remondino, F., 2025. Automating 3D Building Modeling: A Comparative Study of Data- and Model-Driven Methods. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-G-2025, 701–707.

Jawset Visual Computing, 2024. Postshot: AI-powered photogrammetry and neural rendering software. Available at: https://www.jawset.com (accessed: 5 March 2026).

Jiang, C., Gao, R., Shao, K., Wang, Y., Xiong, R., Zhang, Y., 2025. LI-GS: Gaussian splatting with LiDAR incorporated for accurate large-scale reconstruction. IEEE Robotics and Automation Letters, 10(2), 1864–1871.

Kang, S., 2024. Use case of national digital twin standard for buildings. FIG Working Week 2024, Accra, Ghana.

Kasprzyk, J.-P., Nys, G.-A., Billen, R., 2024. Towards a multi-database CityGML environment adapted to big geodata issues of urban digital twins. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-4/W10-2024.

Kerbl, B., Kopanas, G., Leimkühler, T., Drettakis, G., 2023. 3D Gaussian splatting for real-time radiance field rendering. ACM Transactions on Graphics, 42(4), 139.

Ketzler, B., Naserentin, V., Latino, F., Zangelidis, C., Thuvander, L., Logg, A., 2020. Digital twins for cities: A state-of-the-art review. Built Environment, 46(4), 547–573.

Kolbe, T. H., 2009. Representing and exchanging 3D city models with CityGML. In: 3D Geo-Information Sciences, Springer, 15–31.

Kölle, M., Laupheimer, D., Schmohl, S., Haala, N., Rottensteiner, F., Wegner, J. D., Ledoux, H., 2021. The Hessigheim 3D (H3D) benchmark on semantic segmentation of high-resolution 3D point clouds and textured meshes from UAV LiDAR and multi-view stereo. ISPRS Open Journal of Photogrammetry and Remote Sensing, 1, 100001.

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(1), 43–61.

Ledoux, H., Arroyo Ohori, K., Kumar, K., Dukai, B., Labetski, A., Vitalis, S., 2019. CityJSON: a compact and easy-to-use encoding of the CityGML data model. Open Geospatial Data, Software and Standards, 4, 4.

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.

Lehtola, V. V., Koeva, M., Oude Elberink, S., Raposo, P., Virtanen, J.-P., Vahdatikhaki, F., Borsci, S., 2022. Digital twin of a city: Review of technology serving city needs. International Journal of Applied Earth Observation and Geoinformation, 114, 102915.

LichtFeld Studio, 2025. LichtFeld Studio: A high-performance C++ and CUDA implementation of 3D Gaussian Splatting. Available at: https://github.com/MrNeRF/LichtFeld-Studio (accessed: 5 March 2026).

Suwardhi, D., Fauzan, K. N., Harto, A. B., Soeksmantono, B., Virtriana, R., Murtiyoso, A., 2022. 3D modeling of individual trees from LiDAR and photogrammetric point clouds by explicit parametric representations for green open space management. ISPRS International Journal of Geo-Information, 11(3), 174.

Willenborg, B., Pültz, M., Kolbe, T. H., 2018. Integration of semantic 3D city models and 3D mesh models for accuracy improvements of solar potential analyses. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-4/W10, 223–230.

Yan, Z., Yin, M., Shao, Y., Remondino, F., 2025. Evaluating 3D Gaussian Splatting for Urban Scene Reconstruction. ISPRS Archives, XLVIII-1/W6-2025, 251–258.

Ye, V., Li, R., Kerr, J., Turkulainen, M., Yi, B., Pan, Z., Seiskari, O., Ye, J., Hu, J., Tancik, M., Kanazawa, A., 2025. gsplat: An open-source library for Gaussian splatting. Journal of Machine Learning Research, 26, 1–17.

Zhao, H., Min, X., Liu, X., Gong, M., Li, Y., Li, A., Xie, S., Li, J., Panda, A., 2025. CLM: Removing the GPU memory barrier for 3D Gaussian splatting. arXiv preprint, arXiv:2511.04951.

Acknowledgements

This research is funded by the Indonesian Endowment Fund for Education (LPDP) on behalf of the Indonesian Ministry of Higher Education, Science and Technology and managed under the EQUITY Program (Contract No. 4298/B3/DT.03.08/2025). The research is also partly supported by the Digital Twins for Cultural Heritage (Twin4CH – https://twin4ch.github.io/) Junior Professorship Chair, financed by the French National Agency for Research (*Agence nationale de la recherche* – ANR)