Scaling Open 3D City Models: Implementation and Validation of an AI-driven Automated Generation Tool "AI City Model Maker beta version"

Takahiro OOHATA [1], Masahiro SHINODA [1], Ryosuke TAKATSUKA [1], Yasuhito NIINA [2], Mayumi MIZOBUCHI [2]

[1] Realglobe Inc., Japan, shinodamasahiro@realglobe.jp
[2] Asia Air Survey Co., Ltd., Japan, mym.mizobuchi@ajiko.co.jp

Keywords: 3D city model, Project PLATEAU, AI, Automated tool

Abstract

This study presents an AI-based tool for the automatic generation and updating of 3D city models to support the expansion of Project PLATEAU in Japan. High-quality models at Level of Detail (LOD) 2 or higher are essential for applications such as urban planning and disaster management, but their production currently relies on labor-intensive manual processes. To address this issue, the proposed system, AI City Model Maker (Beta Version), integrates multiple AI techniques to generate buildings, roads, city furniture, and vegetation models from heterogeneous geospatial data, including imagery, DEM, and point clouds.

The tool is designed to comply with stringent quality requirements, such as positional accuracy and logical consistency. However, fully automated processing alone remains insufficient to meet these standards, particularly when using commonly available datasets. Therefore, a hybrid workflow combining automated processing and human verification is adopted, with a target cost reduction of 30-50%. The beta version has been tested in multiple real-world environments, demonstrating promising performance, especially for road and urban object modeling. The results indicate the potential to significantly improve efficiency while maintaining practical usability for municipal applications.

1. Introduction

Project PLATEAU, launched by the City Bureau of the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) of Japan in fiscal year 2020, aims to promote the digital transformation of urban development, including the social implementation of smart cities. As its foundational data, the project advances the development, utilization, and open dissemination of 3D city models. One of the major challenges in this initiative is the time and cost required to develop 3D city models. At present, high specification of 3D city models with Level of Detail (LOD) 2 or higher are primarily created manually by skilled technicians. Consequently, when attempting to develop such models over large areas, issues such as labour shortages and increased costs become significant obstacles.

In PLATEAU Vision 2023, a target has been set to complete the development of 3D city models for 500 cities across Japan by the end of fiscal year 2027 (with approximately 340 cities completed as of the end of fiscal year 2025). To expand this coverage, substantial reductions in development time and cost are required. Furthermore, once a 3D city model is created, it immediately begins to become outdated due to temporal changes in urban elements such as buildings, roads, and urban facilities, including construction and demolition. Therefore, periodic updates of 3D city models are highly desirable. Although it is expected that 3D city models will be updated in accordance with the update cycle of municipal urban planning base maps (typically every five years), there is concern that high costs may hinder timely updates.

To help address these challenges, this study focuses on the development of a tool for the automatic generation and updating of 3D city models.

2. Related Work

Research on object-based 3D city model generation compliant with CityGML has advanced rapidly in recent years, driven by the growing availability of open data and progress in automated generation techniques. Zhao et al. (2025) proposed a method for automatically generating city-scale 3D models in CityGML format by integrating OSM, DEM, and existing LOD2 data, covering buildings, roads, urban facilities, and vegetation. Their study targets LOD0 to LOD2 and demonstrates applicability to urban simulation and decision-support systems by enabling parametric control of roof shapes, road widths, and the placement of urban facilities.

City4CFD (TU Delft, 2023) is an open-source software tool that automatically reconstructs LOD2-equivalent building models as well as terrain and roads by combining point cloud data with building polygons. It is primarily designed for urban-scale computational fluid dynamics (CFD) analyses. Supporting standard formats such as CityGML and CityJSON, City4CFD enables the integrated generation of urban models that include not only buildings but also other ground surface elements, thereby offering high practical utility.

Furthermore, Zhu et al. (2025) from the Technical University of Munich (TUM) constructed the GlobalBuildingAtlas, which covers approximately 2.75 billion buildings worldwide using satellite imagery and machine learning techniques. This dataset is characterized by its global-scale provision of building polygons, height information, and LOD1-equivalent 3D building models. While it offers high coverage and consistent quality, its scope is limited to buildings and does not support the automatic generation of comprehensive city object models that include roads or urban facilities, nor does it provide detailed geometric representations beyond LOD2.

Although these studies all automate the generation of CityGML-compliant city models, many remain primarily within the research domain, and examples of social implementation as practical tools or applications in municipal workflows are still limited. Building upon these prior works, the novelty of this study lies in providing AI-based automatic generation modules as a usable tool, enabling users themselves to create 3D city models. In doing so, this research contributes to expanding the coverage and accessibility of 3D city models.

3. Concept

3.1 Target Market

As observed in related works, the automatic generation of 3D city models has been actively studied worldwide. However, the required performance of 3D city models varies depending on their intended applications. For instance, in use cases such as games and entertainment, visual appearance is prioritized even if the models differ somewhat from real-world conditions. In contrast, applications in disaster management and urban planning require high accuracy for decision-making purposes. Meanwhile, in applications such as autonomous driving, the models must be dynamic and interpretable by systems rather than visually appealing. Thus, the required performance of 3D city models differs according to their purposes.

In this study, in order to contribute to the advancement of Project PLATEAU, we focus on surveying accuracy so that the models can be utilized in municipal operations such as disaster management and urban planning. At the same time, we target a wide range of levels of detail (LODs), from low to high (Figure 1).

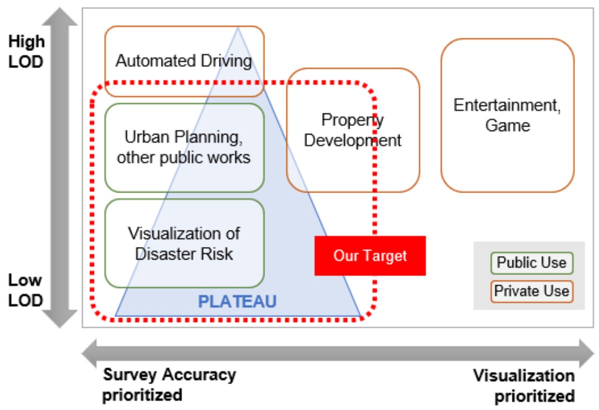

Figure 1. Urban Digital Twin Market and Our Target

3.2 Comparison with Other Tools and Products

In addition, compared with other commercial tools and products, this study is characterized by its wide coverage of urban features (Table 1). This is because the aim is to support a broad range of municipal applications, including disaster management and urban planning.

Feature		AI City Model Maker	HxMap	BuildingReconstruction	Zenrin 3D	AW3D	DYNAMIC MAP PLATFORM	MAP IV	Google/Map Box
Building	LOD0	○			○	○			○
	LOD1	○	○	○	○	○			△
	LOD2	○	○	○	△	△			△
	LOD3	○							
	Attribute	○			△				
Road	LOD1	○				○	○	○	△
	LOD2	○					○	○	
	LOD3	○					○		
City Furniture	LOD1	○					○ (Roadside)	○ (Roadside)	
	LOD2	○					○ (Roadside)	○ (Roadside)	
	LOD3	○					○ (Roadside)	○ (Roadside)	

Table 1. Tools/Products Comparison

4. Methodology

4.1 Requirements

In Japan, Project PLATEAU has developed the "Standard Data Product Specification for 3D City Models" (MLIT, 2025), which is compliant with CityGML, an international standard established by the Open Geospatial Consortium (OGC). This specification serves as a comprehensive guideline for the development, management, and utilization of 3D city models, ensuring their interoperability and consistency across different applications and stakeholders. In particular, it defines detailed acquisition criteria and rigorous quality requirements for 3D city models intended to support municipal operations, including urban planning, infrastructure management, and disaster risk reduction undertaken by local governments throughout Japan.

The specification not only standardizes data structures and formats but also establishes measurable benchmarks for data quality. For instance, when generating 3D city models from point cloud data or aerial imagery at Level 2500, strict positional accuracy requirements are imposed. Specifically, the absolute positional accuracy must achieve a horizontal standard deviation within 1.75 m and a vertical standard deviation within 0.66 m. These accuracy thresholds are designed to ensure that the resulting models are sufficiently precise for practical use in urban-scale analyses and decision-making processes.

Beyond positional accuracy, the specification incorporates a comprehensive quality framework aligned with the ISO 19100 series of geographic information standards. This includes multiple dimensions of data quality, such as completeness, logical consistency, and thematic accuracy. Completeness refers to the degree to which all required features are present without omission or redundancy, while logical consistency ensures that the data structure adheres to predefined rules regarding topology and conceptual relationships. Thematic accuracy, on the other hand, concerns the correctness of attribute information associated with spatial features, which is essential for reliable semantic interpretation and analysis.

An overview of the quality requirements defined in the specification is presented below (Table 2). These criteria collectively ensure that 3D city models maintain a high level of reliability and usability for operational purposes. For completeness, errors due to excess or omission in newly developed datasets must remain within 10% based on sampling inspection. Logical consistency is subject to the strictest requirement, with zero allowable errors in conceptual integrity, topological relationships, and related validation checks. Positional accuracy, as described above, is quantitatively defined for each level of detail, with Level 2500 requiring a horizontal standard deviation of within 1.75 m and a vertical standard deviation of within 0.66 m. Finally, thematic accuracy requires that attribute data be completely error-free, ensuring consistency between geometric representations and their associated descriptive information.

Completeness	For newly developed datasets, errors due to excess or omission shall be within 10% based on sampling inspection.
Logical consistency	Zero errors are permitted for conceptual consistency, topological consistency, and related checks.
Positional accuracy	For Level 2500, the horizontal positional standard deviation shall be within 1.75 m, and the vertical standard deviation shall be within 0.66 m.
Thematic accuracy	Attribute accuracy shall have zero errors.

Table 2. Overview of Quality Requirements for 3D city models

Through this multi-dimensional quality framework, Project PLATEAU provides a robust foundation for the standardized production and utilization of 3D city models. This approach

https://doi.org/10.5194/isprs-archives-L-4-W1-2026-213-2026

facilitates their reliable application in a wide range of domains, including smart city development, digital twin implementation, and evidence-based policymaking.

4.2 Architecture

The proposed tool, the AI-based automatic generation system for 3D city models named “AI City Model Maker (Beta Version)”, is designed to streamline and automate the production of standardized urban digital twin data. The tool ingests quality-assured input datasets, including aerial or terrestrial images, Digital Surface Models (DSM), building footprint polygons, and Digital Elevation Models (DEM). By integrating these heterogeneous datasets within a unified processing pipeline, the system enables efficient and consistent generation of multi-scale 3D urban representations.

Specifically, the tool automatically generates a wide range of urban objects at different Levels of Detail (LOD). For buildings, both LOD1 and LOD2 models are produced, with LOD2 models incorporating roof structures and enhanced geometric detail. In addition, the system performs AI-based texture image upscaling, which improves the visual quality of building façades and enhances realism in visualization applications. Road infrastructure is also generated at LOD1 and LOD2, capturing essential geometric and structural characteristics of the urban road network. Furthermore, the tool supports the generation of higher-detail objects, such as city furniture at LOD3 (e.g., streetlights, traffic signs, and other roadside elements) and vegetation models at LOD3, enabling more precise representation of the urban environment.

The generated outputs are exported in multiple widely used formats, including CityGML for standardized data exchange and semantic modeling, OBJ for visualization and rendering purposes, and SHP (Shapefile) for GIS-based analysis. In addition, texture image datasets are also output to support photorealistic rendering workflows. This multi-format export capability ensures interoperability with various downstream applications, such as GIS platforms, simulation environments, and visualization engines.

The overall system architecture is illustrated in Figure 2. The architecture consists of a modular pipeline in which diverse input datasets are processed through a set of AI-driven 3D modeling modules. These modules are responsible for generating individual city components, such as buildings, roads, city furniture, and vegetation, while also managing associated texture generation processes. The system is supported by backend components, including job management, data storage, and logging subsystems, which enable scalable and traceable processing of large datasets.

Another key feature of the tool is its flexible deployment configuration. The system is implemented in both cloud-based and desktop-based environments, allowing users to select the most suitable operational mode depending on their computational resources, network connectivity, and data volume. The cloud-based version supports large-scale processing and collaborative workflows, while the desktop-based version enables local execution in environments where data transfer or network access may be constrained.

Through this architecture and functionality, the AI City Model Maker (Beta Version) provides a comprehensive and scalable solution for the automated generation of high-quality 3D city models. It contributes to reducing manual modeling efforts while ensuring consistency with standardized data specifications, thereby facilitating the practical implementation of digital twin applications in urban planning, infrastructure management, and related domains.

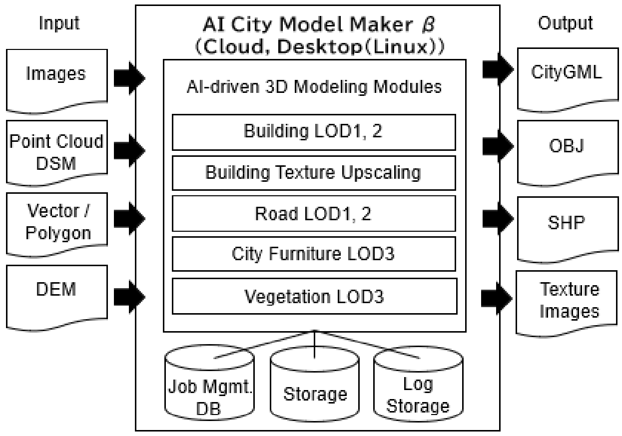

Figure 2. Overview of architechture

4.3 Adopted Technologies

For building model generation, the workflow is primarily based on the open-source software Auto-Create-bldg-lod2-tool (MLIT, 2023), which provides a standardized pipeline for producing LOD2 building geometries from geospatial input data. In the proposed system, this workflow is extended through the integration of well-established deep learning models, including ResNet and TransUNet, to improve both geometric reconstruction and semantic segmentation performance. ResNet is utilized for robust feature extraction from image data, enabling the identification of structural elements such as roofs and façades, while TransUNet contributes to precise segmentation by combining convolutional neural networks with transformer-based architectures. This hybrid approach allows the system to more accurately capture complex roof shapes and building boundaries, thereby enhancing the quality and consistency of the generated LOD1 and LOD2 building models.

For road model generation, the system employs Mask2Former, a recent high-performing architecture for universal image segmentation. This model enables the extraction of road surfaces and related features from imagery with high accuracy, even in complex urban environments where occlusions and heterogeneous surface materials are present. The segmentation outputs are subsequently converted into structured geometric representations, allowing for the generation of road models at multiple Levels of Detail.

In the case of city furniture and vegetation modeling, the tool adopts a multi-stage processing pipeline that combines several AI techniques and geometric processing methods. Point Transformer V3 is applied to point cloud data to capture fine-grained spatial relationships among points, enabling accurate segmentation of irregular objects such as trees and street furniture. YOLOX is used for object detection tasks in image data, enabling the localization of discrete urban elements such as traffic signs and streetlights. In addition, ResNet is further employed for feature extraction and classification, supporting semantic differentiation among detected objects. The outputs of these models are then refined using DBSCAN clustering, which groups spatially coherent point sets and removes noise, followed by parametric modeling techniques that generate structured geometric representations at LOD3. This combination of learning-based and rule-based approaches ensures both flexibility and robustness in modeling diverse urban objects.

The system architecture supports both cloud-based and desktop-based deployment environments, providing flexibility for different operational requirements. The cloud-based version leverages Amazon Web Services (AWS), including Elastic Container Service (ECS) for scalable container orchestration, SageMaker for training and deploying machine learning models, CloudWatch for system monitoring and logging, and S3 for efficient storage and management of large-scale geospatial datasets. This configuration enables high-performance processing of extensive datasets and supports collaborative workflows across distributed environments.

In contrast, the desktop-based version is designed for local execution and adopts a Linux-based environment with Docker for containerized application management and SQLite for lightweight database functionality. This setup allows users to run the system in environments with limited network connectivity or strict data governance requirements, while maintaining consistency with the cloud-based processing pipeline.

4.4 Expected Data Generation Workflow

The quality of the input data has a significant influence on the generation of 3D city models, and this dependency is particularly evident in the proposed AI-based tool. Since the system relies on automated processing of geospatial datasets, the accuracy, resolution, and completeness of the input data directly affect the geometric precision and semantic correctness of the generated models. Consequently, ensuring the suitability of input datasets is a critical prerequisite for achieving outputs that comply with standardized quality requirements.

In Japan, many local governments routinely acquire aerial imagery with a spatial resolution of approximately 25-40 cm as part of their administrative activities, including urban planning and property taxation. In addition to such imagery, nationwide coverage of topographic maps at Level 2500 is available, providing a consistent and accessible source of geospatial information. From a practical perspective, the reuse of these existing datasets as input to automated modeling systems is both cost-effective and scalable, and therefore represents a realistic approach for large-scale deployment of 3D city model generation.

However, despite these advantages, there are inherent limitations associated with the use of such datasets for fully automated processing. At this level of spatial resolution and positional accuracy, it remains difficult to reliably extract fine geometric features, such as detailed roof edges and complex roof structures, solely through automated methods. Similarly, accurately identifying vertical relationships in complex urban environments - such as grade-separated road intersections, overpasses, and underpasses - poses significant challenges. These issues arise due to insufficient resolution, occlusions, and ambiguities in the input data, which can lead to errors or inconsistencies in the generated models.

As a result, there are cases in which it becomes difficult for the automatically generated outputs to fully satisfy the quality requirements described in Section 4.1, particularly with respect to positional accuracy, logical consistency, and completeness. To address these challenges, the proposed tool is not designed as a fully autonomous system, but rather as part of a hybrid workflow that incorporates both automated processing and human intervention.

The overall workflow is illustrated in Figure 3.

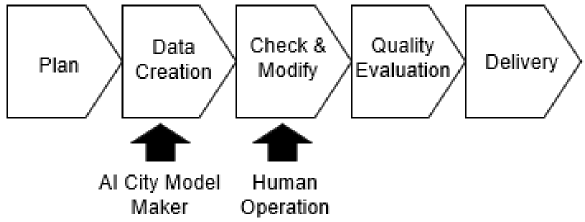

Figure 3. Data Generation Workflow

5. Implementation and Validation

5.1 AI City Model Maker beta version

The user interface of the developed beta version of the AI City Model Maker is shown in Figure 4. It is available on the website: digitaltwin.realglobe.jp/index.html (1 Jun 2026).

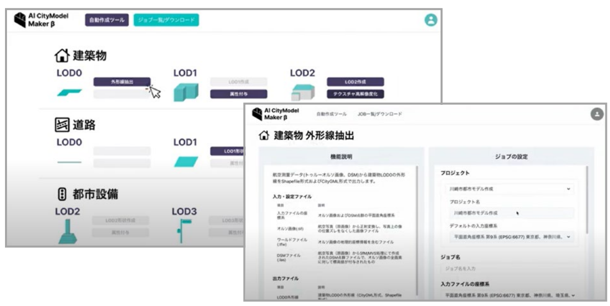

Figure 4. AI City Model Maker beta version

The AI City Model Maker (beta version) is a system under development designed to automate the generation and continuous updating of 3D city models as part of a digital twin infrastructure. The tool is being developed within the framework of Project PLATEAU and aims to scale 3D urban model coverage across Japan through automated workflows.

Technically, the system integrates multiple data sources, including 2D-footprint, aerial imagery, Mobile Mapping System (MMS) point clouds, to construct urban geometries with a high level of detail. It supports the automatic extraction of positional and attribute information from urban facility datasets and converts them into structured 3D representations. Furthermore, the system is being extended to generate fine-grained geometries from high-density point clouds, enabling detailed modeling of infrastructure elements and urban objects.

The system emphasizes efficient processing of large-scale geospatial data, incorporating optimizations for computational performance and scalability. It also includes a quality management framework designed to ensure the reliability of outputs through validation and verification processes. Moreover, ongoing research focuses on dynamic updating mechanisms that allow the city models to reflect temporal changes in urban environments, supporting near real-time digital twin applications.

Overall, the AI City Model Maker represents a comprehensive approach to automating 3D city model production by combining

AI-driven data processing, multi-source integration, and scalable system design, thereby facilitating practical implementation of digital twins in urban planning, infrastructure management, and disaster response contexts.

5.2 Performance Metrics and Targets

In implementing the tool, target values were systematically defined to evaluate both its technical performance and its economic impact. These targets were established for three key aspects: (i) the overall cost reduction achieved across the entire 3D city model production workflow, (ii) cost reduction at the level of individual feature types, and (iii) the degree of automation in 3D model generation required to realize these cost reductions (Table 3). By structuring the evaluation framework in this multi-layered manner, it becomes possible to assess not only the total efficiency gains but also the contribution of each feature type and processing step to the overall performance.

The overall cost reduction targets were set with a phased approach, reflecting the expected evolution of the tool as it matures from prototype to operational deployment. Specifically, a 30% reduction was targeted for fiscal year (FY) 2026, followed by a more ambitious target of 50% by FY2027. These progressive targets assume continuous improvement in automation accuracy, processing efficiency, and robustness under diverse urban conditions. In this context, automation plays a critical role, as higher levels of automation directly translate into reduced labor requirements and shorter production times.

At the individual feature level, distinct cost reduction targets were defined based on the characteristics and complexity of each feature type. For building models at LOD2, a cost reduction of 30% was set, accompanied by an automation target requiring that 90% of models achieve Grade A quality under the assumption of simple roof geometry. This reflects the relatively high complexity of building reconstruction, where geometric variability and roof structure diversity present significant challenges for automated processing.

In contrast, more aggressive cost reduction targets of 60% were established for road models at both LOD1 and LOD2. For LOD1, the automation target focuses on non-intersection segments, while for LOD2, the target excludes intersections and structural change points, where modeling complexity increases substantially. In both cases, a Grade A achievement rate of 75% was required. These conditions acknowledge that road features are generally more uniform and thus more amenable to automation, provided that complex sections are treated separately.

For city furniture and vegetation at LOD3, the evaluation metrics emphasize detection and classification performance rather than purely geometric accuracy. The automation targets include recall and precision values of at least 70%, along with a requirement that at least 25% of generated models meet Grade A quality. A cost reduction target of 50% was defined for this category, reflecting the expectation that advances in object recognition and segmentation techniques can significantly streamline the modeling of smaller and more numerous urban features.

It should be noted that Grade A denotes models that satisfy the quality requirements defined in Section 4.1, which include criteria related to geometric accuracy, completeness, and semantic correctness. By integrating these quality thresholds into the automation targets, the evaluation framework ensures that cost reductions are not achieved at the expense of model usability or reliability. Overall, the target-setting strategy provides a comprehensive basis for quantitatively assessing the effectiveness of the proposed tool across multiple dimensions of performance.

Feature type	Automation target*	Cost reduction	Overall cost reduction
Building LOD2	Simple roof geometry, Grade A: 90%	-30%	FY2026 : -30% FY2027 : -50%
Road LOD1	Non-intersection segments, Grade A: 75%	-60%	
Road LOD2	Excluding intersections and structural change points, Grade A: 75%	-60%	
City Furniture, Vegetation LOD3	Recall & Precision ≥70%; Grade A models ≥25%	-50%	

* Grade A denotes models that satisfy the quality requirements defined in Section 4.1.

Table 3. Evaluation metrics and target settings (overview)

5.3 Validation Results

The evaluation results obtained at the beta stage of the proposed tool are summarized in Table 4. These results provide an initial indication of the tool's performance across multiple feature types and levels of detail (LOD), under varying urban conditions.

Overall, the findings reveal that the targets set for building models at LOD2 were not achieved by a substantial margin. Specifically, although the target aimed for 90% of models to reach Grade A quality with simple roof geometry, the actual results ranged between 79% and 82% in the test areas of Kaga City and Hiroshima City. This gap suggests that further refinement is required, particularly in accurately capturing roof structures and ensuring consistent geometric reconstruction under diverse architectural conditions.

In contrast, the targets for roads, city furniture, and vegetation were generally achieved or approached, indicating comparatively stable performance in these feature classes. For road features, both LOD1 and LOD2 evaluations focused on cost reduction relative to conventional methods, with a target reduction of 60%. The results demonstrate that while the target was nearly achieved in some cases (e.g., up to -58% for LOD1 and -76% for LOD2), there were also instances of underperformance, particularly in more complex environments, where reductions were limited to -42% and -25%, respectively. These variations suggest that the effectiveness of the proposed tool for road modeling is influenced by local conditions such as road network complexity, data density, and surrounding urban structures.

For city furniture and vegetation at LOD3, the tool showed relatively strong performance in terms of detection and classification accuracy. The recall and precision values exceeded the target threshold of 70%, reaching 72% and 90%, respectively, in the Yokohama City test area. Furthermore, the proportion of Grade A models significantly surpassed the target value of 25%, achieving 47%. These results indicate that the proposed method is particularly effective for extracting and modeling smaller-scale urban objects, likely due to the distinct geometric and spectral characteristics of these features.

However, it is important to note that cases of underperformance were also observed depending on the characteristics of the target area. Factors such as urban density, occlusion, variability in object morphology, and data quality appear to influence the results. Therefore, while the beta version demonstrates promising capabilities, especially for non-building features, further improvements are necessary to enhance robustness and achieve consistent performance across all feature types and geographic contexts.

Feature type	Target	Result	Test area
Building LOD2	Simple roof geometry, Grade A: 90%	79 to 82%	Kaga City, Hiroshima City
Road LOD1	Cost reduction: -60%	-42 to -58%	Nishiyodogawa Ward, Kikugawa City
Road LOD2	Cost reduction: -60%	-25 to -76%	Nishiyodogawa Ward, Kikugawa City
City Furniture, Vegetation LOD3	Recall & Precision ≥70%	72%, 90%	Yokohama City
	Grade A models ≥25%	47%	Yokohama City

Table 4. Evaluation results of the beta version

5.4 Feedback from Test Users

The beta version of the tool was introduced to nine companies in Japan, including aerial surveying companies, construction consulting firms, and software developers. Among them, six companies actively used the tool and provided feedback.

As a result, the road LOD1 function received very high evaluations, with users indicating that it could be directly integrated into their production workflows. For vegetation, although it is essential for applications such as landscape simulation, the creation of CityGML-compliant LOD3 models typically requires significant manual effort. Therefore, strong expectations were expressed for the capability to automatically generate such models.

In addition to functional aspects, valuable feedback was also obtained regarding improvements for practical use, including the types of input data, visualization of processing workflows, and the clarity of terminology used in the tool interface.

6. Conclusion

This paper focuses on 3D city models at LOD2 and higher and presents the development of a beta version of an AI-driven modeling tool designed to support efficient and scalable urban digital twin creation. As discussed in Section 4.1, the quality requirements for 3D city models defined by Project PLATEAU are highly stringent, particularly in terms of positional accuracy, logical consistency, and completeness. Under the current state of technology, it remains difficult to fully satisfy these requirements through completely automated processing alone, especially when relying on commonly available input datasets.

For this reason, the development of the proposed tool has been carried out under the practical assumption that manual editing and verification by human operators will remain an integral part of the workflow. Within this hybrid framework, the primary objective has been to significantly reduce the total production cost and effort associated with 3D city model generation, with a target reduction of approximately 30-50%, as described in Section 5.2. This target reflects a balanced approach that combines the efficiency of AI-based automation with the reliability of human quality control.

Although these performance and cost reduction targets have not yet been fully achieved at the current beta stage, the tool has already reached a level of maturity sufficient for real-world evaluation. In fact, it has been deployed for pilot testing at six different companies, where it is being assessed under practical operational conditions. Feedback obtained from these pilot implementations is being used to further refine both the modeling algorithms and the overall system design. In parallel, continuous efforts are being made not only to improve the accuracy and robustness of the generated models, but also to enhance the usability and operability of the tool as a practical application, including interface design and workflow efficiency.

Future work will focus on further expanding the capabilities of the system. This includes not only improving the quality and coverage of the currently supported feature types, but also extending the automation to more detailed representations, such as LOD3 building models and LOD3 road models. Achieving these advancements will require the integration of more sophisticated AI techniques and higher-resolution data processing, but they are essential steps toward realizing fully functional and high-fidelity digital twin environments.

Acknowledgement

This work has been carried out as part of the project "Development and Demonstration of an Automated System for the Generation and Updating of 3D City Models" under the Small and Medium Enterprise Innovation Promotion Program (SBIR Phase 3) of the Ministry of Land, Infrastructure, Transport and Tourism (MLIT), Japan.

References

MLIT, 2023: Technical Report on the Development and Open-Sourcing of an AI-Based LOD2 Automatic Generation Tool. github.com/Project-PLATEAU/Auto-Create-bldg-lod2-tool (1 Jun 2026).

MLIT, 2025: Standard Data Product Specification for 3D City Model. mlit.go.jp/plateaudocument/ (1 Jun 2026).

Niina, Y., Nonaka, H., Mizobuchi, M., Jimba, W., Yu, Z., Komori, A., Tsutsui, S., Ohata, T., and Shinoda, M., 2025. Development of an automated 3D city model generation tool using AI technologies, *Journal of the Japanese Society for Artificial Intelligence*, 40(4), pp. 474–481.

TU Delft 3D Geoinformation Group, 2023: City4CFD: Reconstruction of LoD2 3D City Models. github.com/tudelft3d/City4CFD (1 Jun 2026).

Zhao, C., Ogawa, Y., Liao, L., Sekimoto, Y., 2025: Virtual 3D City Model Generation in CityGML, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4/W6-2025, 257–264.

Zhu, X. X., Chen, S., Zhang, F., Shi, Y., Wang, Y., 2025: GlobalBuildingAtlas: an open global and complete dataset of building polygons, heights and LoD1 3D models, *Earth System Science Data*, 17, 6647–6668. doi.org/10.5194/essd-17-6647-2025.

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-L-4-W1-2026-213-2026 | © Author(s) 2026. CC BY 4.0 License.