# Multi-perspective assessment of greenery visibility from streets, windows, and drones using deep learning and CIM

Siyuan Meng[a], Maosu Li[b*], Jinfeng Xie[c], Fan Xue[d]

[a]Department of Real Estate and Construction, The University of Hong Kong, Pokfulam, Hong Kong, China;
[b]Thrust of Urban Governance and Design, The Hong Kong University of Science and Technology (Guangzhou), Nansha, Guangzhou, China;
[c]Department of Urban Planning and Design, The University of Hong Kong, Pokfulam, Hong Kong, China;
[d]Department of Real Estate and Construction, The University of Hong Kong, Pokfulam, Hong Kong, China.
National Center of Technology Innovation for Digital Construction Hong Kong Branch, The University of Hong Kong, Pokfulam, Hong Kong, China

**ARTICLE HISTORY**
Compiled May 14, 2026

**ABSTRACT**

Urban greenery includes 3D volumetric greenery and provides visual benefits to billions of urban dwellers. Although high-rise buildings and drones are expanding urban activities vertically, most greenery studies are limited to street-level observations, relying on 2D/2.5D view imagery. This paper presents a greenery volume visibility (GVV) assessment method by integrating 3D deep learning and city information model (CIM). A 3D GVV index ($GVVI$) is defined for three groups of observing perspectives: street, window, and drone. Semantic segmentation using deep learning detects greenery volumes in photorealistic 3D models, then $GVVI$s are calculated across all perspective groups. Hot spot maps and disparity analysis comparing the three sub-indices reveal the urban greenery volumes overlooked in traditional ground-level studies. Experimental results on a 4.92 $km^2$ area of Kowloon, Hong Kong, showcased the proposed method's efficiency (23,490 $GVVI$s in 256.45 hours); Comparative results between street and drone/window perspectives revealed 37.6% and 31.7% greenery entirely overlooked in traditional ground-level spatial analysis. This paper contributes in three aspects: a 3D GVV index definition, an efficient deep-learning-based geospatial analysis method, and novel quantitative 3D evidence for informed urban greenery planning and maintenance, optimal building designs, and low-altitude drone route planning.

**KEYWORDS**
Urban green volumes; Visibility assessment; Deep learning; City information model; Greenery planning and maintenance

Word Count: 6210

* Corresponding author. Email: maosuli@hkust-gz.edu.cn

*This is the peer-reviewed author's version of the paper:*
Meng, S., Li, M., Xie, J., & Xue, F. (2026). Multi-perspective assessment of greenery visibility from streets, windows, and drones using deep learning and CIM. *International Journal of Geographical Information Science, (in press).*
DOI: https://doi.org/10.1080/13658816.2026.2667807

## 1. Introduction

Urban greenery encompasses vegetation in green spaces within urban agglomerations, such as woodlands, parks, gardens, streets, and squares (Konijnendijk *et al.* 2006). Urban greenery provides a range of amenity values, e.g., leisure opportunities and aesthetic enjoyment for residents (Kong *et al.* 2007). Viewing urban greenery improves human health, well-being, life contentment, and productivity (Ulrich 1984, Jiang *et al.* 2021), whether from walking on the street or from homes (Wang *et al.* 2022, Elsadek *et al.* 2020, Li *et al.* 2026). The value of different greenery types is associated with economic factors, such as housing prices (Panduro and Veie 2013). Spatial planning of green spaces integrated with greenery impact assessments (e.g., noise, air pollution, visual amenity value) on residents can enhance their life satisfaction (Neema and Ohgai 2010). Therefore, measuring greenery visibility supports evidence-based decision-making in maintaining urban green spaces, improving city image, and revitalizing city blocks.

Urban greenery visibility is usually assessed at the street level, whereas new perspectives have recently emerged for its assessment. The expanding urban activities include drone-view-based activities and vertical living. Residents in high-rise, high-density cities often experience less nature at street level but benefit more from green views observable from high-rise windows (Elsadek *et al.* 2020, Ko *et al.* 2020). Drone-level views offer a valuable means to capture and assess views from above-ground greenery visibility. They are increasingly used for collecting greenery information, exploring virtual environments, and experiencing urban sightseeing (Ecke *et al.* 2022, Perperidou and Kirgiafinis 2022, Zhang *et al.* 2024). However, a greenery visibility assessment method, including street, window, and drone perspectives, is limited in the current literature.

Urban greenery visibility is often imbalanced in terms of spatial distribution (Cimburova *et al.* 2023), especially in high-rise, high-density cities. Given the importance of viewing greenery for comfort and general well-being, a quantitative assessment of this uneven spatial distribution from different perspectives is necessary. This assessment can further support applications such as greenery planning and maintenance, optimal building design, and low-altitude drone route planning.

Traditional methods operate in 2D or 2.5D, neglecting the 3D nature of GVV. Manual assessment of GVV remains labor-intensive and cost-prohibitive. 2D and 2.5D data, such as street view imagery (SVI) and Digital Surface Models (DSM), were applied to estimate the perception of greenery from a specific observation point, i.e., green view index, or to assess the GVV (Yang *et al.* 2020, Cimburova *et al.* 2023). However, current methods focus on a single perspective, usually at street level, neglecting the multi-perspective GVV assessment. Besides, the 2D-based method failed to locate specific greenery, and viewshed-based methods using 2.5D data may be insufficient to accurately represent the real physical environment and often require substantial computational costs due to the need for stacked computations for each GVV-observation pair.

The emerging deep learning and CIM provide new opportunities for assessing 3D GVV. CIM, which originates from digitalized systems of urban elements and environment (Xue *et al.* 2021), integrates with photorealistic urban mesh models, virtual cameras from any perspective, and 3D rendering techniques for consistent color representation (Li *et al.* 2022, Liang *et al.* 2017). Deep learning has been widely employed to address many geographical issues, from semantic segmentation of SVI to 3D photorealistic mesh models, such as Stratified Transformer (Biljecki and Ito 2021, Lai *et al.*

2022, Li *et al.* 2025). The integration of CIM and deep learning enables the efficient computation of GVV from any perspective.

This paper, therefore, introduces a novel GVV assessment method using deep learning and CIM (Figure 1). The main focus includes assessing visibility from building windows for urban living, low-altitude drones, and traditional pedestrian perspectives. Firstly, the greenery volume visibility index ($GVVI$) is defined. Secondly, a 3D semantic segmentation model detects different greenery volumes from photorealistic mesh models, and the Greenery Volume Layer (GVL) with a grid-level color index database is displayed in 3D CIM. Afterward, the $GVVI$ of each volume is calculated using street, window, and drone-view images. Finally, the 'hot spots' map and disparities of $GVVI$ and three sub-indices are analyzed.

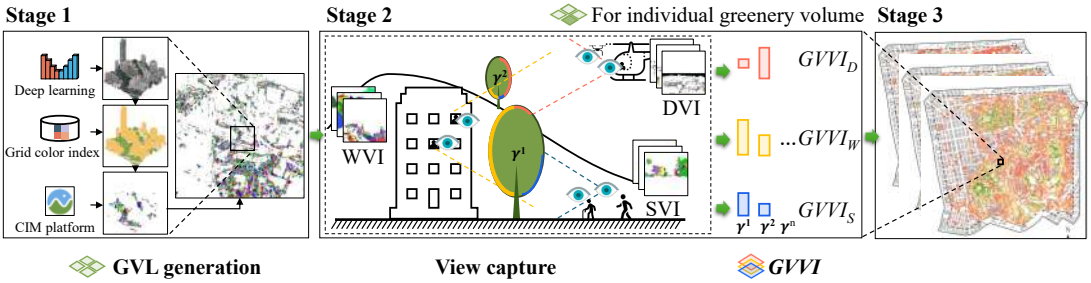


Figure 1.: Conceptual illustration of GVV assessment

The contribution of this paper is threefold. First, a metric $GVVI$ is defined to measure GVV types, including street, park, and garden, from different perspectives (street, window, and drone levels). Second, an automatic three-step assessment method that is effective and efficient, utilizing deep learning and CIM. Third, $GVVI$ hot spots and disparities highlight new quantitative 3D evidence for informing urban greenery planning and maintenance, optimizing building designs, and planning sightseeing 3D routes for low-altitude drones.

## 2. Literature Review

### *2.1. Traditional methods to measure greenery visibility*

Greenery valuation methods established the monetary value of trees, including individual tree conditions and amenity functions such as visual contribution (Helliwell 2008, Roberto Barbosa *et al.* 2010, Doick *et al.* 2018, Cimburova *et al.* 2023). The amenity value of greenery refers to the characteristics that contribute to the aesthetic appreciation of people, i.e., the visibility of greenery. Instead of the green view index, which assesses the residents' greenery exposure degree at street/building level (Li *et al.* 2023b), greenery visibility highlights the visual contribution of greenery to surrounding residents (Cimburova *et al.* 2023). Thus, assessing the visibility of greenery is key to informed decision-making related to greenery maintenance, design, and renovation.

Greenery assessment method is transitioning from traditional overhead views to relying more on human-centric visibility data sources. The Normalized Difference Vegetation Index (NDVI) from satellite imagery effectively quantifies large-scale greenery (Huang *et al.* 2021), but its urban application is limited by the lack of vertical greenery data and low resolution (Yu *et al.* 2022). Researchers, therefore, have turned to eye-level data, such as SVI, to assess the visual quality and amenity value of greenery

(Li *et al.* 2015, Yu *et al.* 2016, Cimburova *et al.* 2023).

Methods for assessing eye-level green view include field surveys, image-based analysis, and geospatial modeling (Yu *et al.* 2016, Wang *et al.* 2023). Since field surveys are labor-intensive, requiring excessive time to collect data (Doick *et al.* 2018), image-based methods, such as SVI and CIM-based window view image (WVI), offer eye-level visual evidence (Li *et al.* 2015, Liu *et al.* 2023, Xie *et al.* 2025). However, the method is proven to have both building and greenery coverage bias (Fan *et al.* 2025, Huang *et al.* 2025). For example, SVI only provided information at photographic observation points and failed to consider the greenery within large green spaces (e.g., gardens and parks) (Huang *et al.* 2025). The issues of availability and data quality (e.g., camera pose variation, weather conditions, blurriness) have been reported in previous literature (Rui and Cheng 2023, Tang *et al.* 2023). Furthermore, the current CIM-based methods, which extract green view through 2D deep learning, are unable to locate greenery in images (Li *et al.* 2024). Thus, the image-based methods suffer from quality, reliability, and coverage limitations, apart from locating greenery, which together make them unsuitable for GVV assessment.

Several studies have employed the viewshed-based method to simulate greenery visibility using 3D data sources, such as DSM and point clouds. For instance, greenery assessments of street-level exposure and tree-level visibility were developed (Cimburova and Blumentrath 2022, Cimburova *et al.* 2023, Tang *et al.* 2023, Pyka *et al.* 2022). While viewshed simulations offer a human perspective in 3D, the accuracy of occlusion surfaces derived from 2.5D data sources is limited in reflecting the real physical environment and require large amount of computation cost (Cimburova and Blumentrath 2022). Besides, the acquisition of point clouds for a large-scale GVV assessment is costly (Tang *et al.* 2023).

### *2.2. Recent advances in CIM and deep learning*

A 3D CIM is a digital representation of the physical and functional attributes of a city, serving as a collaborative knowledge repository (Song *et al.* 2017, Xue *et al.* 2021). CIM-based research has demonstrated its ability to simulate human perspectives from the building level by leveraging its photorealistic texture and high-performance capability. Therefore, photorealistic mesh models demonstrate certain advantages in representing the 3D greenery volumes compared with traditional data sources.

The high-performance capability of CIM makes it feasible for quick positioning and capturing virtual views from any location. Specifically, WVI was generated to assess different urban view types using deep learning, quantify eye-level green view from upper-floor perspectives, and measure window view distance (Li *et al.* 2022, 2023b, 2025). Moreover, the level of detail 2 CityGML-based model was employed to generate WVI and identify visible greenery (Bolte *et al.* 2024). Therefore, real-time greenery observation from higher elevation is promising, thanks to increasingly accurate and realistic urban environment datasets worldwide (Yu and Gong 2012, HKPlanD 2019, HKLandsD 2022).

The advancing deep learning facilitates novel approaches for extracting data on urban greenery. Deep learning enables computational models with multiple processing layers to learn multiple-level abstracted representations of data (LeCun *et al.* 2015). Previous research achieved semantic segmentation in 2D (e.g., SVI) and 3D (e.g., point cloud) urban-level data (Biljecki and Ito 2021, Li *et al.* 2025). Multi-view-based, volumetric-based, and point-based methods have been developed to segment 3D point

clouds (Guo *et al.* 2021), such as Stratified Transformer (Li *et al.* 2023a). Thus, the Stratified Transformer can attach the photorealistic models with semantic information.

In summary, GVV is warranted in its own right, but also critical for evaluating its impact on landscape and urban studies. Current literature lacks a quantitative assessment method of GVV from non-traditional perspectives, such as drones and windows. 3D CIM represents a promising data source providing realistic views from drone, window, and street level observations. Deep learning, meanwhile, can be incorporated for the semantic segmentation of a CIM and in the analysis of greenery volumes.

## 3. GVV assessment method

Figure 2 shows an overview of the three-stage GVV assessment method in this paper. In the first stage, the method applies 3D deep learning to the photorealistic mesh model to create a GVL of multi-class greenery. Then, street, window, and drone views are captured from the GVL-enriched CIM using road networks, windows on building footprints, and drone positions above DTM. Finally, the output $GVVI$ values are calculated and analyzed for urban greenery maintenance and other purposes.

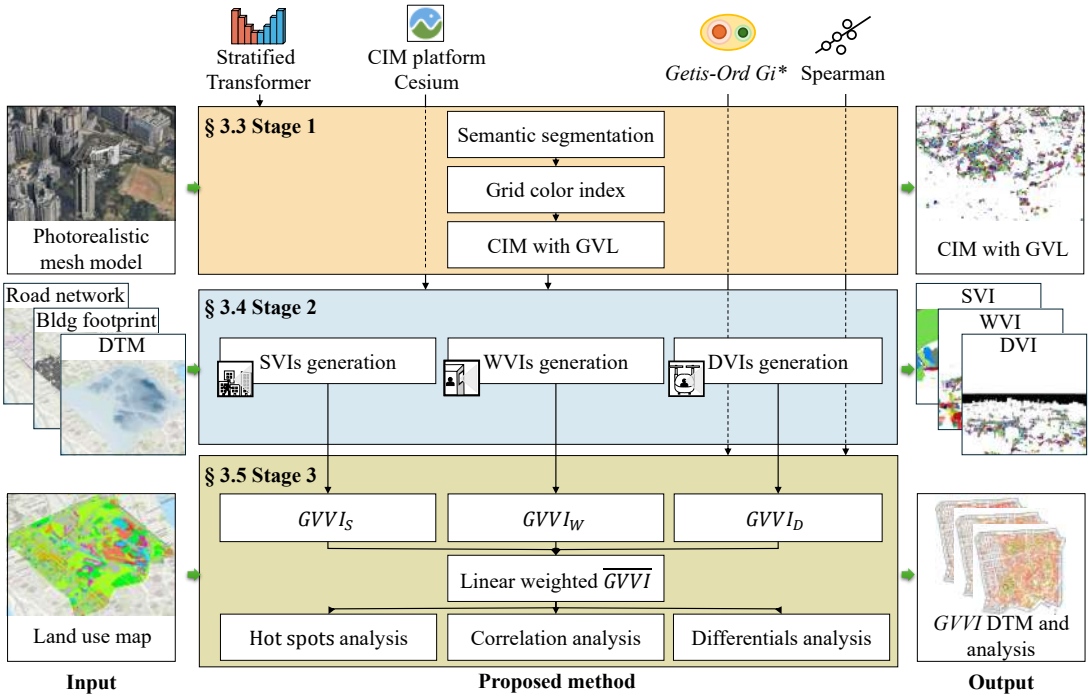


Figure 2.: Flow chart of the proposed green volume visibility assessment method

### *3.1. GVVI definition*

For a given greenery volume $\gamma \in \Gamma$, the $GVVI^{\gamma} = [GVVI^{\gamma}_{\mathrm{S}}, GVVI^{\gamma}_{\mathrm{W}}, GVVI^{\gamma}_{\mathrm{D}}]$ is a vector of three sub-indices. These sub-indices represent the aggregated visibilities of $\gamma$-th volume from three distinct observing perspectives: streets (S), windows (W), and drone positions (D), respectively. For each sub-index $i \in \{\mathrm{S}, \mathrm{W}, \mathrm{D}\}$, Equation 1 shows

the calculation for the total number of visible pixels ($N_i^\gamma$) in all views ($V_i$) in group $i$:

$$N_i^\gamma = \Sigma_{v \in V_i} ||pixels(v, \gamma)||, \tag{1}$$

where *pixels* represents the function that returns all the pixel set of $\gamma$-th volume in the view $v$, and $||\cdot||$ returns the cardinality of the set. Which means that, the *pixels* of the $\gamma$-th volume will be counted $v$ times in each sub-index $i$, to evaluate GVV from all observation points. The $GVVI_i^\gamma$ is the normalized number of pixels in Equation 2.

$$GVVI_i^\gamma = \frac{N_i^\gamma - N_{\min}}{N_{\max} - N_{\min}} = \frac{N_i^\gamma - min_{j \in \Gamma}(N_i^j)}{max_{j \in \Gamma}(N_i^j) - min_{j \in \Gamma}(N_i^j)}. \tag{2}$$

Each $i$-th sub-index $GVVI_i^\gamma \in [0, 1]$, and the vector $GVVI^\gamma = [GVVI_{\mathrm{S}}^\gamma, GVVI_{\mathrm{W}}^\gamma, GVVI_{\mathrm{D}}^\gamma] \in [0, 1]^3$ is well bounded value in the unit cube. Therefore, $GVVI$ values of all green volumes $\Gamma$ are a list of vectors $[GVVI^1, GVVI^2, \ldots, GVVI^J]$, where $J = ||\Gamma||$ is the cardinality of the greenery volumes.

A weighted index $\overline{GVVI}$ aggregates the three sub-indices using observing-perspective-specific weights, as depicted in Equation 3.

$$\overline{GVVI}^\gamma = \frac{w_{\mathrm{S}} \cdot GVVI_{\mathrm{S}}^\gamma + w_{\mathrm{W}} \cdot GVVI_{\mathrm{W}}^\gamma + w_{\mathrm{D}} \cdot GVVI_{\mathrm{D}}^\gamma}{w_{\mathrm{S}} + w_{\mathrm{W}} + w_{\mathrm{D}}}, \tag{3}$$

where $w_{\mathrm{S}}$, $w_{\mathrm{W}}$, and $w_{\mathrm{D}}$ are non-negative weights denoting the relative weights of each group of observing perspectives.

### *3.2. Stage 1: Greenery volume layer (GVL)*

In this stage, all grid-level greenery volumes are detected and highlighted within the input 3D CIM. This stage includes semantic segmentation, colorized indexing, and integration with the 3D CIM platform.

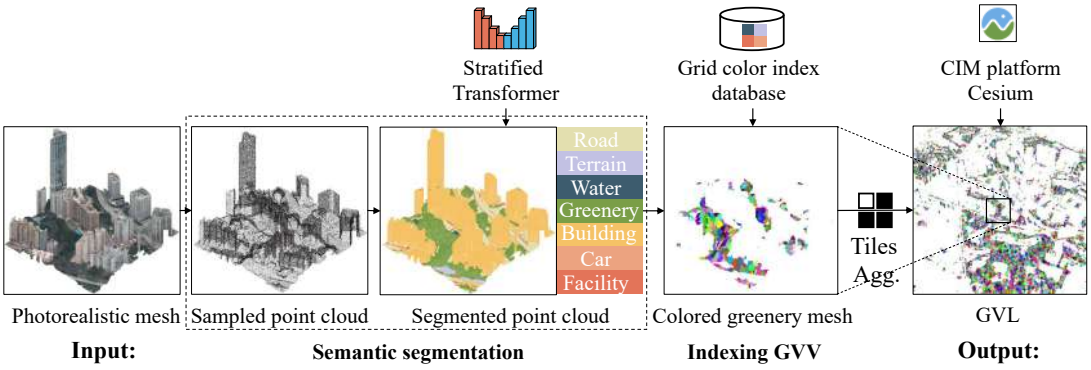


Figure 3.: Creation of greenery volume layer (GVL) using 3D deep learning

Firstly, an urban point cloud is sampled on the photorealistic mesh model at a density of 10 points/m$^2$. Then, the Stratified Transformer is utilized for the 3D semantic segmentation. The Stratified Transformer is a deep learning model capable of capturing long-range contexts, showcasing robust generalization capabilities, and achieving high performance (Lai *et al.* 2022). This model has proven to be effective in segmenting

large-scale urban point clouds (Li *et al.* 2023a). The HRHD-HK of high-rise, high-density cities serves as the training dataset, encompassing seven classes (i.e., road, terrain, water, greenery, building, car, and facility), where the greenery class refers to vegetation such as trees, bushes, and grass. etc, as listed by Li *et al.* (2023a). The training is performed for 500 epochs with a batch size of 1 to balance computational cost and accuracy, while all other parameters follow the default settings of the Stratified Transformer. Semantic segmentation is validated using mean intersection over union (mIoU). One annotator manually assigned ground-truth labels for greenery to a random sample comprising 10% of the point cloud tiles in the study area. The selected tiles cover greenery within hillsides, parks, gardens, and streets that typically occur in high-density cities.

We assign a grid index to the greenery area based on the segmented point cloud. The size of the grid is set to 10m, corresponding to the average tree diameter in Hong Kong (Kong *et al.* 2017). The photorealistic mesh model is segmented based on the greenery point cloud and the grid index; simultaneously, a unique and opaque color, serving as an indexing ID, is randomly attached to each greenery grid mesh. The non-greenery mesh is colored white. The segmented mesh is converted into 3D tiles via CesiumLab and then displayed on the CIM Cesium platform. The colored mesh models are differentiated by their tiles. The final output of this stage is the city-scale 3D CIM with GVL, as shown in Figure 3.

### *3.3. Stage 2: Automatic views capture and indices generation*

All three groups of views, i.e., SVIs, WVIs, and DVIs, are captured based on the city-scale GVL enriched CIM and the sample points generated in this stage (Figure 4).

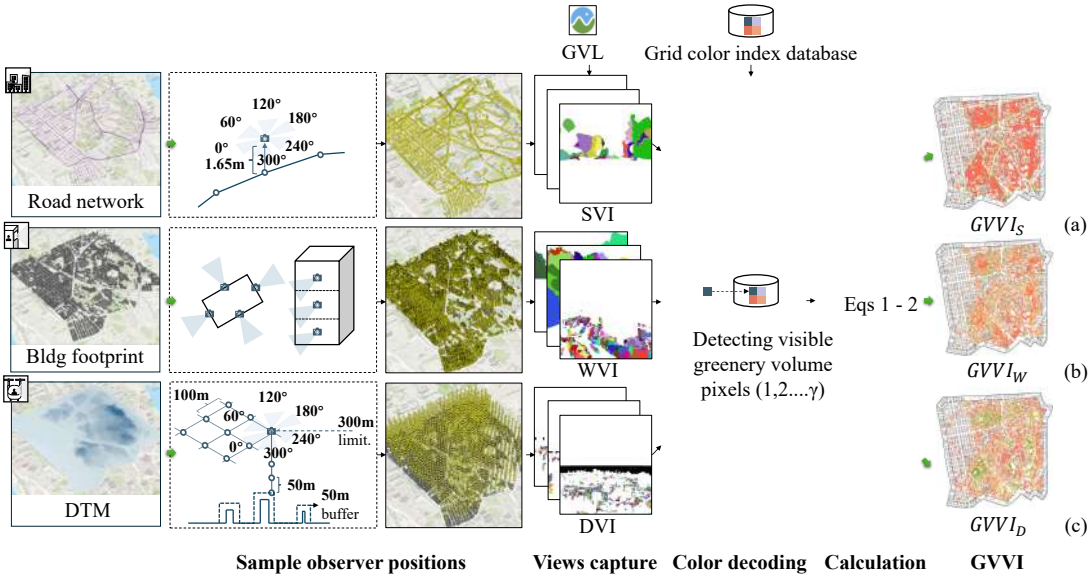


Figure 4.: The process of capturing views. (a) SVI; (b) WVI; and (c) DVI

The sample points of SVIs are generated based on the road network to simulate both pedestrian and car eye-level perception simultaneously. We sample points at an interval of 30m, which is the average street length in the research area (Gong *et al.* 2018). For each point, the $z$ value is determined by adding 1.65m to the DTM (average

height of adults in Hong Kong) (NCD Risk Factor Collaboration (NCD-RisC) 2016). The heading of the virtual camera is classified into six directions (0°, 60°, 120°, 180°, 240°, 300°), with a 60-degree field of view (FOV) (Meng and Zheng 2023).

We generate the sample points of WVIs based on the building footprints. The midpoint of each simplified footprint edge is used as the sampling point. Three height levels-low, medium, and high-are sampled to simulate the window view from different building floors. Subsequently, a virtual camera with a 60-degree FOV is positioned at specified locations on the CIMs to capture the exterior window views, according to Li *et al.* (2022, 2023b).

The sample points of DVIs are generated based on the DTM and the 3D building footprints. Firstly, the sample points' altitudes are set to mark a 50m buffer away from buildings and terrain, because of the noise and privacy concerns of the residents. The upper limit of the altitude of DVI is set according to the local airspace classification.

DVI was captured at 50 m vertical intervals, recording visibility changes caused by altitude while balancing computational cost and precision (Seifert *et al.* 2019). In addition, sampling is performed at intervals of 100m in the horizontal direction because of the high-flying speed of the drone, contrary to vehicles on the ground (Xiang *et al.* 2024). The virtual camera is oriented in six directions (0°, 60°, 120°, 180°, 240°, and 300°) for capturing images. A 60-degree FOV is used to represent the central visual field of a human (Meng and Zheng 2023, Tara *et al.* 2021).

Mesh penetration artifacts occur in some images, particularly at sampling points close to the mesh surface and in regions where mesh fusion occurs (e.g., greenery and buildings) at the bottom of the photorealistic model. This leads to inaccuracies in the display of greenery volumes and, consequently, miscalculations of GVV. Hence, the images with model penetration are filtered using a Support Vector Machine based on the image deep feature (totaling 2,048 features) extracted by Inception V3 in Orange (version 3.36) (Szegedy *et al.* 2016). We used 400 views (200 true, 200 false), with 320 for training and 80 for testing, as the dataset.

Finally, we calculate the $GVVI_{\mathrm{S}}$, $GVVI_{\mathrm{W}}$, $GVVI_{\mathrm{D}}$, and $GVVI$ using an automated Python script, based on captured and filtered views for analysis. Each color captured in the views represents the visible pixels of a unique GVV from the observing perspectives. The results are calculated using Equations 1, 2, and 3 with equal weight combination.

### *3.4. Stage 3: GVVI analysis*

The *Getis-Ord Gi** analysis within ArcGIS Pro is utilized to reveal the spatial cluster locations of high and low values, namely the greenery hot spot map. Each of the three sub-indices is visualized and analyzed using the greenery hot spot maps. In the statistical analysis, volumes within a 100m buffer of the research area boundary are excluded due to insufficient observations. A sensitivity analysis of weight adjustment is conducted to evaluate the impact of weight combinations on the distribution of greenery hot spots. Spearman's coefficient was used to analyze the correlations and the independence of indices between the three sub-indices of $GVVI$ and NDVI data from 2018.

The differentials between street and drones, street and windows, are calculated, as shown in Equation 4, where $i \in \{\mathrm{W}, \mathrm{D}\}$. Following this, hot spot maps of differentials are computed. The identified hot spots are extracted to demonstrate the total number of identified differential hot spots.

$$GVVI_{i-\mathrm{S}} = GVVI_i - GVVI_{\mathrm{S}} = [GVVI_i^1 - GVVI_{\mathrm{S}}^1, GVVI_i^2 - GVVI_{\mathrm{S}}^2, \ldots, GVVI_i^J - GVVI_{\mathrm{S}}^J] \quad (4)$$

Greenery distribution across different land-use types has been shown to impact human well-being (Bahr 2024). Therefore, assessing the association between $GVVI$ differentials and land use type can further support such applications. The two differentials $GVVI_{\mathrm{D-S}}$ and $GVVI_{\mathrm{W-S}}$ indicate the extent to which the proposed $GVVI$ values could differentiate themselves from the traditional studies based on land use type at street and ground levels.

## 4. Experimental setting

### *4.1. Study area*

The study area, located in Kowloon, Hong Kong, spans 4.92 km$^2$ with an elevation range of approximately 100m (Figure 5). The study area is a well-developed urban area with small variations in buildings over the past 10 years, totaling 150 new buildings HKBD (2023). It comprises 6,294 buildings, with an average height of 25.6m (the lowest is 1.5m, and the highest is 229m). Natural landscapes (such as parks and gardens) are distributed in different locations in the study area, such as hillsides, residential areas, and along streets. Because of the line-of-sight blockage caused by terrain elevation changes and buildings, the amenity value of greenery to individuals varies considerably, making it an ideal location for this research.

### *4.2. Data*

The photorealistic mesh model and land use map (Figures 5c and f) were provided by the Planning Department of Hong Kong (HKPlanD 2019, 2023). The road network used for SVI generation was produced by the Transport Department of Hong Kong (HKTD 2019). Building footprints in the study area were extracted from the iB1000 digital map of Hong Kong (HKLandsD 2025), which has been continuously updated since 2014. A 300m upper limit of DVI sampling point is set according to the uncontrolled airspace classification of China (CAAC 2023), given that Hong Kong has no official documentation specifying vertical airspace divisions.

### *4.3. Validation metrics*

The validation of the proposed method is conducted based on SVI and DVI. Real-world SVI captured over a span of three years (2014, 2017, 2023) is used to validate the eye-level green view from the street perspective. Real-world SVI is segmented using Mask2Former model with Swin Transformer backbone (Cheng *et al.* 2022) trained on the Cityscapes dataset. The analysis involves calculating the average greenery coverage at sample points along each street. Real-world DVI (totaling 48 DVIs covered eight sampling points and $60° \times 6$ headings) is collected using a drone (DJI Mini3 Pro) in a public garden within the study area. One annotator manually assigned ground-truth labels for greenery in the study area. Spearman's coefficient between the normalized greenery ratio in DVI pairs is used to validate the drone-level accuracy.

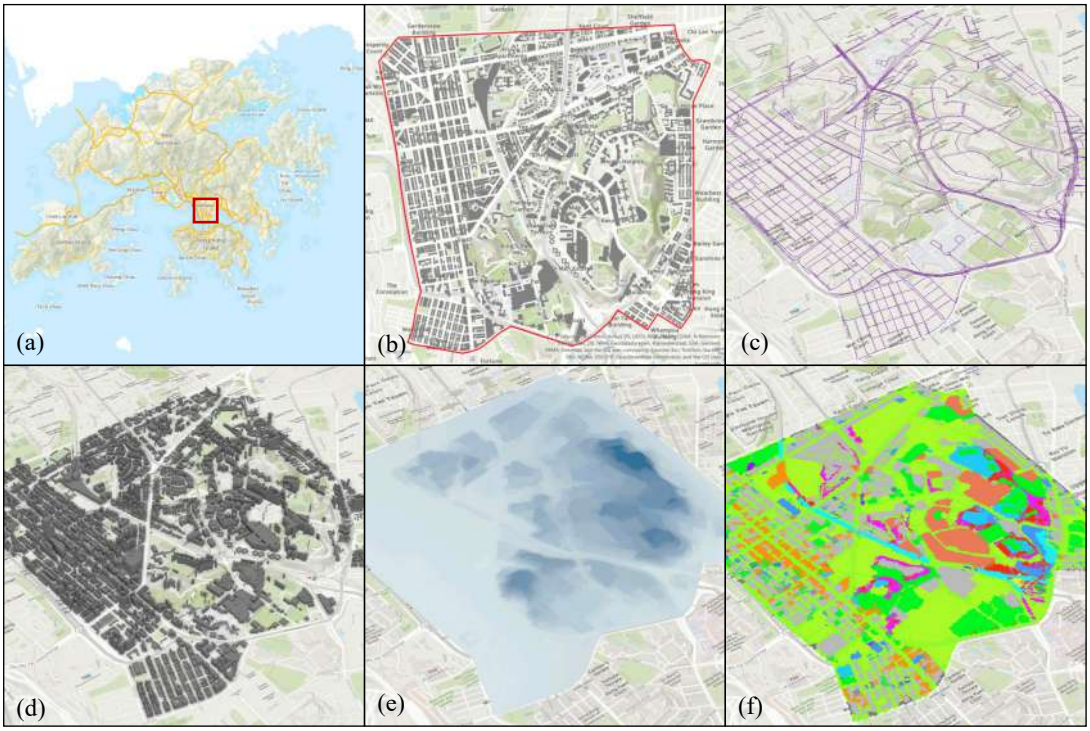


Figure 5.: The study area is Kowloon, Hong Kong. (a) The location of the study area; (b) the 4.92 $km^2$ study area; (c) 2D road network (111.6km in total); (d) 6,294 building footprints; (e) DTM indicating a mixture of hilly and flat areas; (f) land uses map

A method-level validation with baseline methods was conducted to demonstrate effectiveness and efficiency. Spearman's coefficient was calculated to assess the effectiveness of the GVV results, comparing our method with that of Cimburova *et al.* (2023); and the street-level green view, comparing with Cimburova and Blumentrath (2022). The validation ground truth included 18 volumes manually annotated by one annotator, and the street-level green view, which was calculated based on real-world SVI (2017). The 18 greenery volumes were selected based on combinations of three distance levels to the road (close, middle, far) and three levels of surrounding building density (high, middle, low), with two samples for each combination.

### *4.4. Computational environment*

The deep learning training environment was set up as follows: a high-performance computing (HPC) server with Dual Intel Xeon 6226R (16-core), 384GB RAM, and NVIDIA V100 (32GB) SZM2 GPU. The training process was implemented using Pytorch (version 1.8) and Python (version 3.7). The view capture phase for GVV was completed by Cesium (version 1.99), the script for $GVVI$ calculation was completed by Numpy (version 1.24) and Pandas (version 2.0.3), the $GVVI$ analysis was completed in ArcGIS Pro (version 3.3) based on a desktop computer with 13th Gen Intel(R) Core(TM) i7-13700K 3.40 GHz, 128 GB RAM, and NVIDIA RTX A4000 GPU.

## 5. Results

A total of 109,081 views were captured and filtered from the GVL-enriched 3D CIM (17,718 from streets, 73,939 from windows, and 17,424 from drones). The SVM used to filter penetrated views achieved 95% precision, recall, and F1-scores, and removed 14% of views at the window and street levels. Figure 6 illustrates examples of captured SVIs, WVIs, and DVIs, respectively.

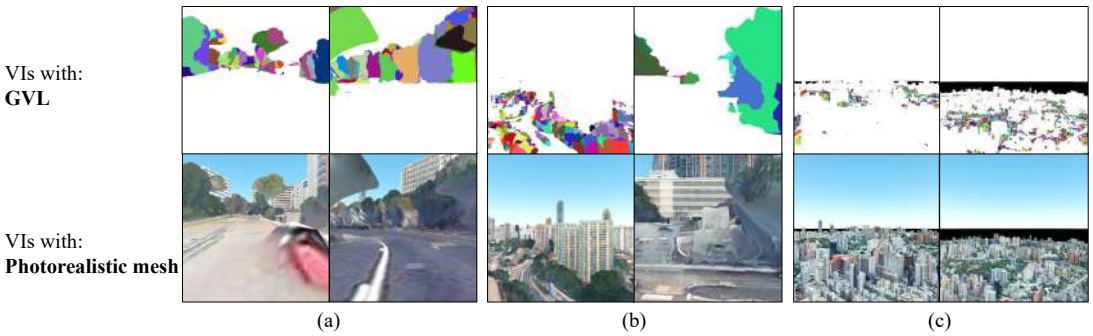


Figure 6.: Comparison of SVI, WVI, and DVI captured based on GVL CIM and the photorealistic model. (a) SVI; (b) WVI; (c) DVI

The segmentation accuracy of greenery achieved IoU = 86.27%, while other classes achieved IoU = 95.94% (mIoU = 91.11%). The greenery types, such as street-side shrub and roof-top greenery, were often misclassified as classes such as facility, building, and road, after manual inspection. These may be caused by low precision at the bottom of the photorealistic models. A total of 23,490 greenery volumes were finally obtained. The greenery volumes within the buffer zone were 20,655 (up to 2.06km$^2$).

Table 1.: Computational time of the proposed automated assessment method

| Stage | Processing | Software | Time (h) |
|---|---|---|---|
| 1 | Semantic segmentation | Stratified Transformer | 83.18 |
| | Grid color index | Python | 11.03 |
| | GVL integrating | CesiumLab | 0.80 |
| 2 | View capture | Cesium | 147.03 |
| | Visibility statistics | Python, Orange 3 | 14.41 |
| 3 | $GVVI$ computation and analysis | ArcGIS Pro | 0.00 |
| | Total | | 256.45 |

Table 1 lists the computational cost of each process in the proposed method. The first stage, generating GVL, took 12.06 hours to process 4.92 km$^2$ of CIM mesh with 90 million triangles and 267 million sampled points. The most time-consuming stage was the view capturing process, taking up to 57% of the total processing time.

### 5.1. $\overline{GVVI}$ and three sub-indices

Figure 7 demonstrates the spatial distribution and hot spot map of $\overline{GVVI}$ and sub-indices. The highest $\overline{GVVI}$ value of 0.44 implied that a high amenity value level from all three perspectives (i.e., street, window, and drone) was uncommon to appear simultaneously. Greenery hot spots of $\overline{GVVI}$ value (7,056 in total) mainly comprised greenery

located close to the major roads, areas with few surrounding high-rise buildings, and greenery with greater canopy height and elevation. The near-edge interiors of public green spaces exhibited moderate levels of $\overline{GVVI}$. In contrast, cold spots of $\overline{GVVI}$ were concentrated in the grass, shrublands, and high-rise, high-density residential areas.

Figure 7a, c, and d demonstrate the spatial distribution of GVVI sub-indices. The street-level greenery received higher visual attention along the major roads. Visibility was also obtained for street greenery scattered in residential areas though at lower values. This difference can primarily be attributed to the challenge of distinguishing central plants when they are surrounded by edge greenery at lower eye levels, as illustrated in Figure 6a. The most visible greenery at the window level was typically located at the edges of natural green spaces or urban park areas, whereas the interior greenery in those areas was less visible. Along-street greenery exhibited lower visibility because of visual obstructions by surrounding buildings. At the drone level, the visible greenery was concentrated in the interior of large green spaces at high altitudes and in areas with fewer surrounding buildings that caused visual obstructions.

Figure 7f, g, and h illustrate three hot spot maps of sub-indices. The hot spots (confidence $>$ 90%) of $GVVI_{\mathrm{S}}$ and $GVVI_{\mathrm{W}}$ were 5,427 (mean = 0.06) and 5,237 (mean = 0.07), respectively. $GVVI_{\mathrm{D}}$ obtained the largest number of hot spots, with $N_D$ = 7,922, mean = 0.17. This indicates that $GVVI_{\mathrm{D}}$ had a higher continuity spatial clustering result because the aerial perspective was less visually obstructed.

## *5.2. Validation*

### *5.2.1. Validation of eye-level green view from street perspective*

Table 2 presents a street-level comparison between CIM-based and real-world SVI captured in different years. All real-world SVIs exhibited strong and positive associations with CIM-based SVI (p-value $<$ 0.001), indicating that the proposed method effectively simulated eye-level green view. The 2017 real-world SVI exhibited the highest coefficient (0.93) and $R^2$ (0.84), and the lowest root mean square error (0.06). The differences in coefficients between 2017 and the other two data sources were less than 0.03, indicating that the greenery visibility simulated using CIM was relatively insensitive to temporal variation. Nonetheless, it is recommended to acquire up-to-date data to reduce the greenery change because of urban renovation and expansion.

Table 2.: Comparison of eye-level green view between CIM-based SVI and real-world SVI captured in different years.

| **Data source** | **Spearm. $\rho$** | **p-value** | **MAE** | **RMSE** | **R2** |
|---|---|---|---|---|---|
| 2014 | 0.90 | $<$0.001 | 0.04 | 0.07 | 0.81 |
| 2017 | 0.93 | $<$0.001 | 0.04 | 0.06 | 0.84 |
| 2023 | 0.92 | $<$0.001 | 0.04 | 0.07 | 0.82 |

The difference between real-world SVI and CIM-based SVI was caused by the following reasons: (1) the roughness of the photorealistic mesh at the ground level; (2) the lack of thin objects in the photorealistic mesh, which are labeled in the Cityscapes dataset (e.g., car, person, pole, and traffic sign), and (3) the assumption of GVL that all greenery is impenetrable, as shown in Figure 8, which is different from the real-world SVI. The impenetrable greenery will result in a higher average GVVI differential of 0.002 according to density parameters and height classification criteria of different

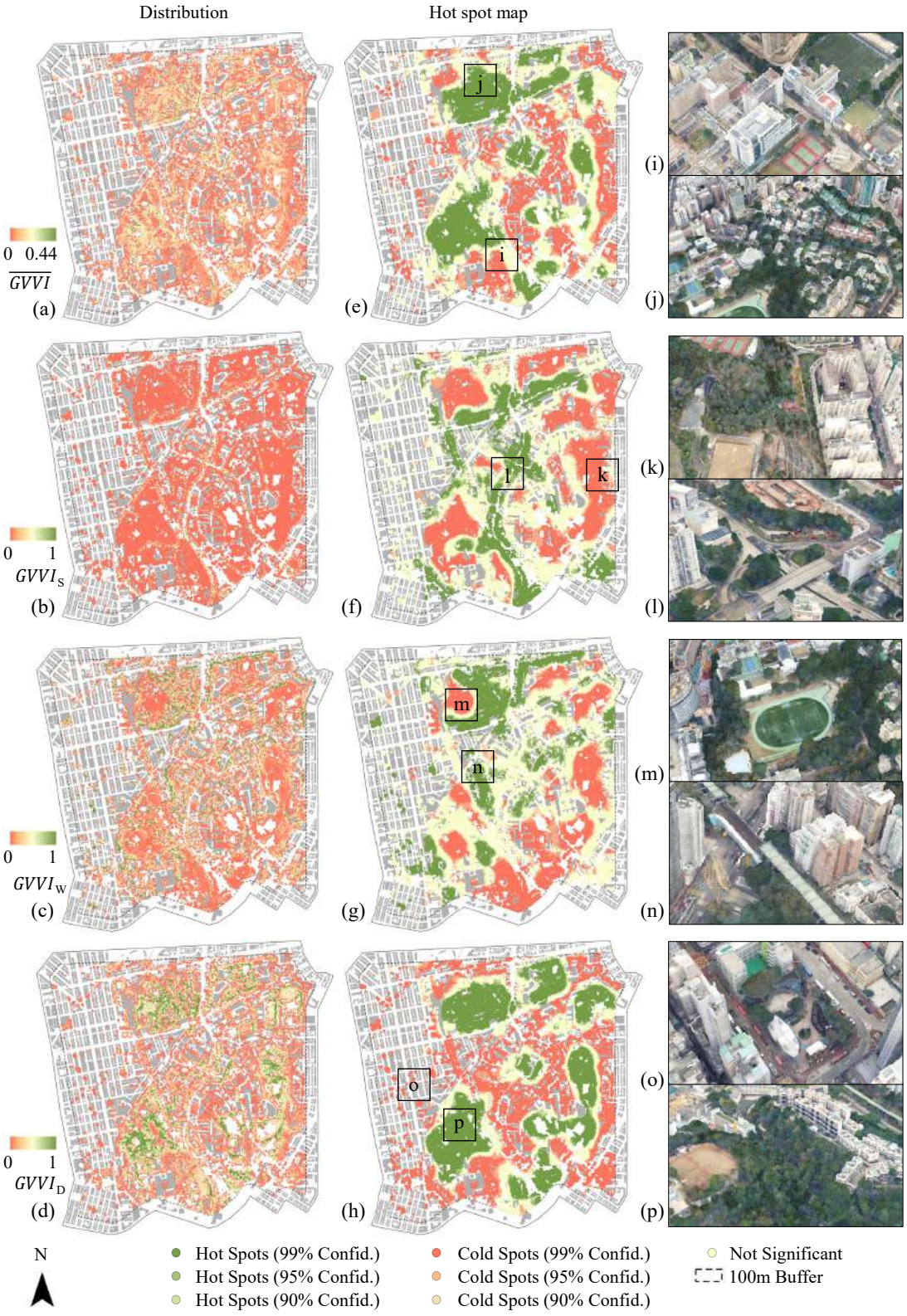


Figure 7.: Scaled distribution and hot spot map results with visible and non-visible areas. (a),(e) $\overline{GVVI}$; (b),(f) $GVVI_S$; (c),(g) $GVVI_W$; (d),(h) $GVVI_D$; (i), (k), (m), (o) non-visible areas; (j), (l), (n), (p) visible areas.

greenery types provided by Li *et al.* (2025). Despite the above limitations, CIM can effectively simulate the perception of greenery as experienced by humans and can further be utilized for GVV computation.

### *5.2.2. Validation of eye-level green view above ground*

Figure 9 shows the comparison between the eye-level green view of CIM and real-world DVI. The overall Spearman's coefficient and $R^2$ of totaling 48 DVI pairs reached 0.95 (p-value $< 0.001$) and 0.92, respectively, demonstrating a satisfactory simulation result of ours. Outlier pairs are mainly caused by vegetation renewal (e.g., changes in height and pattern). Furthermore, due to differences in camera FOV, some distant vegetation with fewer pixels may be obscured by buildings, leading to subtle variations between pairs.

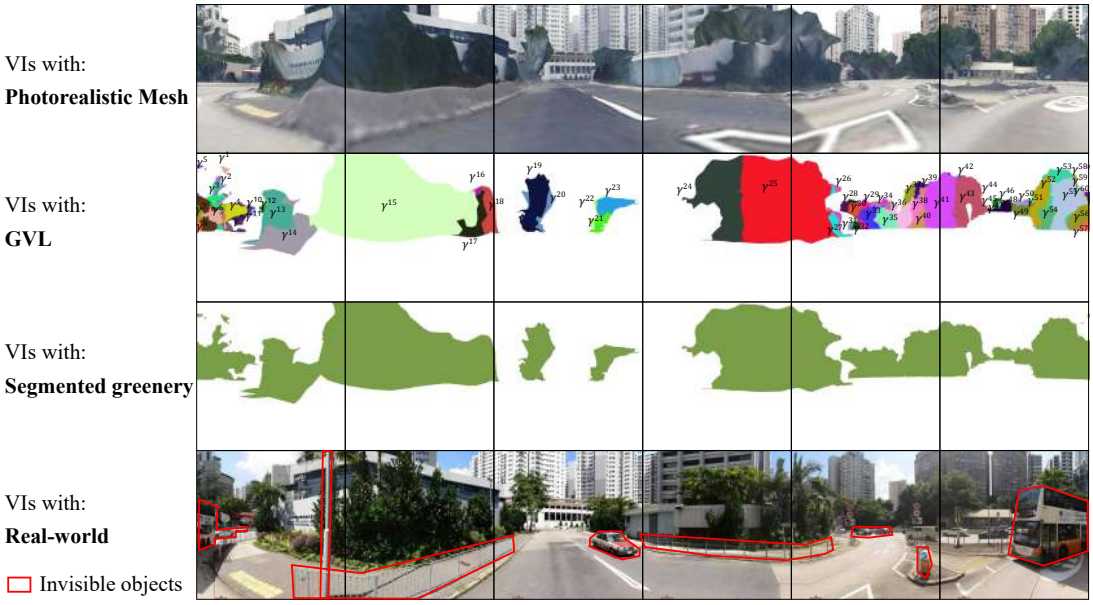


Figure 8.: Illustration of Photorealistic-mesh-based, GVL-based, segmented-greenery-based, and real-world SVIs.

### *5.2.3. Method-level validation*

The proposed method exhibited high efficiency, accuracy, and multi-perspective assessment (Table 3). Although the method of Cimburova and Blumentrath (2022) achieved higher efficiency, it failed to compute GVV, while that of Cimburova *et al.* (2023) required significant computational cost for viewshed analysis. Both methods lack multi-perspective assessment. Our proposed method required 24.6h to calculate $GVVI_S$, leading to an approximately 40% saving in time. Our results showed a strong correlation with the real-world SVI for both GVV and eye-level green view, with Spearman's $\rho = 0.89$ and 0.93 (p-values $< 0.01$), and approximately 0.2 higher than previous methods, indicating enhanced assessment accuracy.

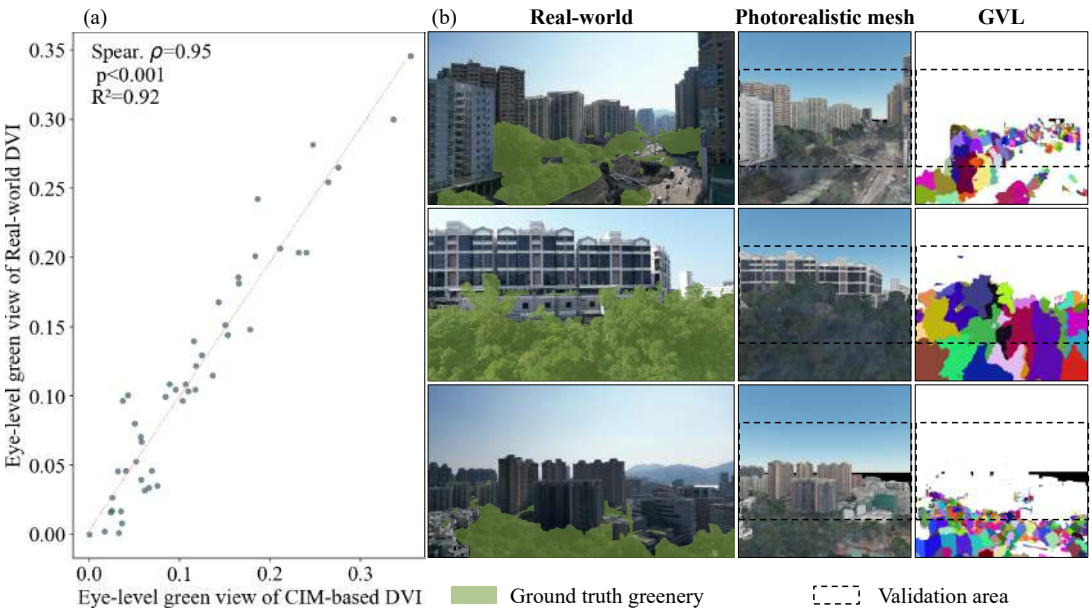


Figure 9.: Scatter plot and comparison between eye-level green view of CIM and real-world DVI. (a) Scatter plot; (b) illustrative examples.

Table 3.: Validation of the proposed method with viewshed-based GVV assessment

| Method | Type | Time cost (h) | Spearm. $\rho$ (GVV) | Spearm. $\rho$ (eye-level) |
|---|---|---|---|---|
| Cimburova and Blumentrath (2022) | Viewshed | 1.3 | \ | 0.78 |
| Cimburova *et al.* (2023) | Viewshed | 41.6 | 0.70 | \ |
| Ours | CIM | 24.6 | 0.89 | 0.93 |

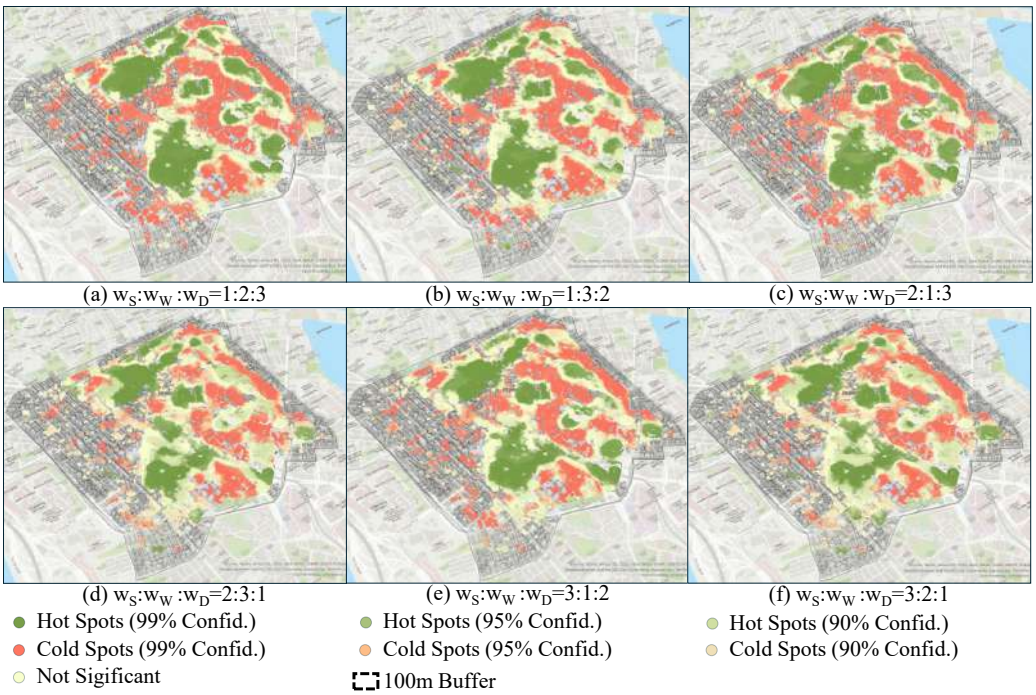


Figure 10.: Hot spot maps of greenery's $\overline{GVVI}$ with different weight combination

## 5.3. *GVVI analysis*

### *5.3.1. Sensitivity of weight adjustment*

The linear weighted $\overline{GVVI}$, with different weight groups, identified 6,840 hot spots and 7,512 cold spots GVV on average, as shown in Figure 10a-f. Those groups that emphasize SVI and WVI exhibited lower $\overline{GVVI}$ on average; for example, about 0.04 in Figure 10d and f. This difference indicates that $GVVI_{\mathrm{D}}$ contributed more significantly to $\overline{GVVI}$. The clusters from different weights share a similar spatial distribution pattern, but vary in certain areas, e.g., greenery in open space with relatively high elevation or dense greenery along streets.

### *5.3.2. Correlations between the sub-indices*

The Spearman's correlation analysis is shown in Figure 11, with all p-values $< 0.0001$. The correlation between $GVVI_{\mathrm{S}}$ and $GVVI_{\mathrm{W}}$ ($\rho = 0.34$) was weak but significant, indicating that a green volume more visible from streets was likely to be more visible from windows. The correlation between streets and drones was weaker ($\rho = 0.19$). This is because pedestrians mainly observe the greenery's bottom part, whereas a drone's perspective offers a clearer view of the tree canopy and less of the bottom part due to high elevation. The coefficient between $GVVI_{\mathrm{W}}$ and $GVVI_{\mathrm{D}}$ was 0.40, indicating a moderate positive association between greenery visibility in WVI and DVI.

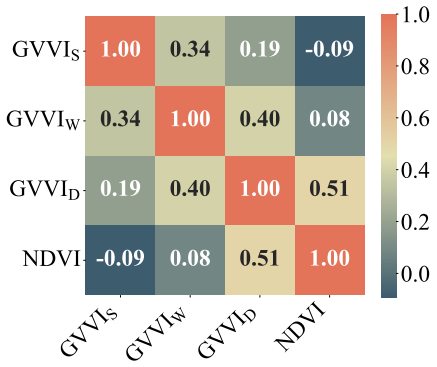


Figure 11.: Correlation metrics of the scaled $GVVI$ sub-indices and NDVI, all p-values $< 0.0001$

The correlation between NDVI and the three sub-indices shows weak associations, especially at the street and window levels, as shown in Figure 11. The negligible correlation between NDVI and $GVVI_{\mathrm{S}}$ ($\rho = -0.09$) and $GVVI_{\mathrm{W}}$ ($\rho = 0.08$) indicates that greenery visibility of pedestrians and residents was significantly different from traditional top-view data sources. Although the correlation between NDVI and $GVVI_{\mathrm{D}}$ showed a moderate positive association ($\rho = 0.51$), greenery in areas such as open spaces, streets, and high-density residential areas displayed a notable differential compared with NDVI (ranging from -0.88 to 0.39) according to manual validation.

### *5.3.3. Differentials and the correlations to land use*

Figure 12a shows the number of clustered hot spots of $GVVI_{\mathrm{D\text{-}S}}$ and $GVVI_{\mathrm{W\text{-}S}}$ in different land use types. The government land areas had the highest number of hot spots of differentials, mainly because of the larger area size and higher altitude (e.g.,

hilltops). $GVVI_{\text{D-S}}$ had 7,757 hot spots of 0.77km$^2$, equivalent to 37.6% of the total greenery area within the buffer zone. Except for government areas, open spaces, private residential areas, utilities, and woodland contributed most significantly to $GVVI_{\text{D-S}}$ (N$\gtrsim$800). Meanwhile, $GVVI_{\text{W-S}}$ had fewer hot spots (N = 6,539, 31.7% of the total greenery area), with government areas being the primary contributor, followed by private residential, open spaces, and transportation. Sensitivity analysis of DVI sampling parameters' impact on $GVVI_{\text{D-S}}$ hot spots is shown in Appendix C1.

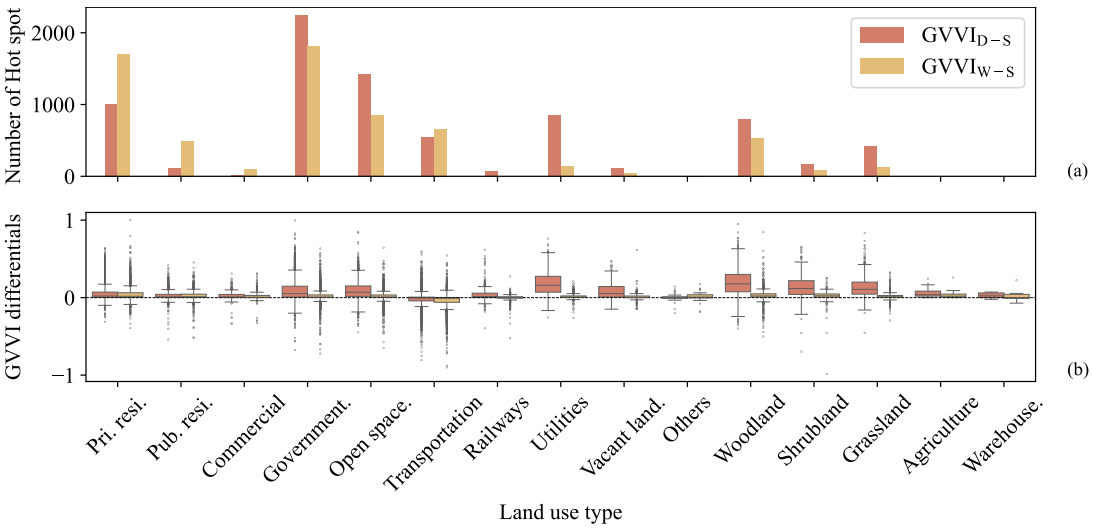


Figure 12.: Hot spots number and differentials distribution of $GVVI_{\text{D-S}}$ and $GVVI_{\text{W-S}}$. (a) Number of hot spots; (b) differential distribution.

Figure 12b demonstrates the $GVVI$ differentials distribution of $GVVI_{\text{D-S}}$ and $GVVI_{\text{W-S}}$ against land-use types. More than 81% and 76% of greenery volumes exhibited positive values for $GVVI_{\text{D-S}}$ and $GVVI_{\text{W-S}}$, respectively. Not surprisingly, the transportation area was the primary land-use type with negative value differentials. In contrast, other land-use types predominantly showed positive differentials for both $GVVI_{\text{D-S}}$ and $GVVI_{\text{W-S}}$, suggesting that greenery in these areas was more visible from residential buildings and drones. Furthermore, the differentials of $GVVI_{\text{W-S}}$ were smaller than those for $GVVI_{\text{D-S}}$, indicating that the contribution of GVV increased with elevation of the observing points.

## 6. Discussion

### *6.1. GVV assessment method and results*

The proposed $GVVI$ metric includes multi-perspective greenery visibility, street, window, and drone, compared with the previous GVV assessment, which was limited to street-level only. The proposed GVL based on deep learning and CIM can efficiently capture views from any perspectives and assess GVV. GVL enables localizing greenery volume from remote observer points (e.g., windows, drone views). Besides, the proposed method can assess eye-level green view and GVVI simultaneously without complex coding using a reusable and customizable GVL. The proposed method reduces the time cost by 40% and increases the correlation by 0.19 compared to the viewshed-based method. Therefore, it enables efficient, effective, and multi-perspective

GVV assessment.

The analysis results confirmed the imbalanced spatial distribution of 3D GVV. A total of 37.6% differentials from drone, and 31.6% from window, compared to traditional street-level studies, were identified. Correlation metrics highlight the necessity of assessing 3D greenery visibility, according to the correlation between $GVVI_{\mathrm{S}}$ and $GVVI_{\mathrm{W}}$ (0.34), as well as $GVVI_{\mathrm{S}}$ and $GVVI_{\mathrm{D}}$ (0.19). A multi-level strategy that balances ground-level experience with overhead landscape aesthetics is therefore required to ensure residents can access adequate and consistent visible greenery from different perspectives. The hot spots, distribution, and weight adjustment map can serve as a coarse-to-fine guide for greenery maintenance. The three sub-indices and NDVI can further be combined as a comprehensive, complementary, and down-to-top assessment strategy for assessing different levels of greenery visibility.

### *6.2. Potential applications*

The $\overline{GVVI}$ and hot spot map identified the amenity value level of urban greenery and provided new insights into priority areas for maintenance. The proposed $GVVI$ can serve as the foundational data for the government authorities to prioritize green areas with high $GVVI$ value and to make informed decisions regarding urban green space management, including planting, removal, or maintenance activities. The multi-perspectives $GVVI$ provides a significant reference framework across multiple fields, including urban planning, landscape architecture, forestry management, and drone-based environmental assessment.

The application of GVL can be expanded. For instance, it can be used to determine how the height and canopy size of greenery can best serve surrounding areas by simulating greenery volumes in CIM. Reorganization of the three-dimensional pattern of the urban landscape may influence the take-off, route trajectory, and landing of drones (Perperidou and Kirgiafinis 2022). $GVVI_{\mathrm{D}}$ could help recommend optimal drone take-off and landing sites for cities, balancing aesthetic and cultural factors with other drone indicators, such as noise and wind speed (Wild 2024). The drone route planning can be optimized based on the $GVVI_{\mathrm{D}}$ to balance the observed greenery resolution and cost (Ecke *et al.* 2022). The proposed method also inspires CIM-based simulations for multi-perspective visions for agents, i.e., physical assets (e.g., windows, lookouts, and surveillance), individuals (e.g., wheelchair users, children, and cyclists), and robotics (e.g., drones, unmanned cars, and vessels). This capacity thus enables applications in safety, urban accessibility, and potential conflicts analysis in order to support a future human-machine society (Meng *et al.* 2025).

The weights of $\overline{GVVI}$ can be determined through experts' experience, surveys, and local policies. Greenery managers can assign a personalized weighting to alter the emphasis of $\overline{GVVI}$ according to the focus of management, for example, by prioritizing greenery with higher $GVVI_{\mathrm{W}}$ (Meng and Zheng 2023). Additionally, other indicators, such as surrounding air pollution, the noise impact of drones, greenery health condition, greenery maintenance cost, urban heat island, and global warming effect (Chauhan *et al.* 2025, Zheng *et al.* 2025), can be combined with $GVVI$ to comprehensively evaluate the ecosystem value and maintenance priority of greenery in future research. Optimization-based multiple-criteria decision analysis can also be used for $GVVI$ assessment to recognize the order of priority (Li *et al.* 2023b, Zhou and Xue 2023).

### *6.3. Limitations and future directions*

This research has certain limitations. From the CIM perspective, the limitations include low precision at the bottom of the photorealistic models, the absence of realistic occlusion caused by the window frame, opaque GVL assumption against the permeability of greenery, and low efficiency in view capture. From the view sampling perspective, the sampling parameter based on experience may lead to assessment bias, and the static views-capture approach failed to reflect human-environment interactions. From the drone views perspective, limitations include over-reliance on simulated data and under-representation of actual trajectories of different types of drone views. Besides, model segmentation accuracy may be affected when applied to CIM across regions outside Hong Kong due to differences in feature distribution. Finally, this paper excludes the impact factors (e.g., urban morphology, seasons, and temporal data availability) on $GVVI$ across different perspectives.

Therefore, future work can focus on several key aspects. These include the development of high-precision CIM models and the design of high-performance algorithms for OpenGL-based rendering. Additionally, evaluating sampling parameter bias, simulating viewpoints driven by diverse user types and video data, and modeling realistic drone trajectories are necessary to advance future research. Further, the applicability validation of datasets (e.g., HRHD-HK) across regions using transfer learning (Weiss *et al.* 2016) should be explored. Finally, systematic analyses of seasonality, transferability, and temporal data availability impact on $GVVI$ assessment are required (Han *et al.* 2023).

## 7. Conclusion

This paper presents an effective, efficient, and automatic method to evaluate GVV. In theory, the proposed method defines a new $GVVI$ metric to unify street, window, and drone-level GVV. In methodology, the proposed method introduces GVL for enabling the 3D CIM-based GVV assessment using deep learning and color index. In application, the proposed method is implemented through a case study in Hong Kong. Linear-weighted $\overline{GVVI}$ values and hot spot clustering, covering 2.06 $km^2$ of greenery volumes, provided an analysis of multi-perspective greenery visibility.

The results confirmed the uneven contribution of greenery visibility. The identification of an additional 37.6% of greenery at the drone level, and 31.7% at the window level–overlooked in traditional street-level studies–demonstrates the value of the proposed method for greenery management. The Spearman's correlation between $GVVI_{\mathrm{S}}$ and $GVVI_{\mathrm{W}}$ (0.34), and that with $GVVI_{\mathrm{D}}$ (0.19) highlighted the necessity of 3D greenery visibility. Greenery located closest to main roads, surrounded by a few high-rise buildings, and with high canopy height, should be given priority for maintenance. The quantitative differentials of hot spot distributions across land-use types provided a reliable reason for stakeholders to develop targeted maintenance strategies.

The results of the proposed method demonstrate the utility of assessing 3D $GVVI$ and can serve as foundational geospatial data for urban planners and managers to prioritize greenery design and maintenance. The GVL attached to the photorealistic mesh can represent a close-to-reality urban built environment, enhance the semantic representation of greenery in 3D CIM, simulate greenery visibility from any perspective, and provide efficient GVV assessment compared with previous methods, thereby further supporting applications such as drone route planning and greenery maintenance.

Limitations of this study include model penetration, the time-consuming process for capturing views, a lack of simulation of human-environment interactions and drone trajectories, as well as the exclusion of factors such as urban morphology, seasons, and temporal data availability. Future research directions are needed, such as incorporating a higher precision photorealistic mesh/CIM, developing high-performance algorithms for OpenGL rendering, simulating real drone trajectories and human-environment interactions, evaluating sampling parameter bias, validating dataset applicability across regions, analyzing seasonality, transferability, temporal data availability of CIM impact on $GVVI$ assessment, integrating additional assessment indicators, and exploring potential applications.

## Acknowledgments

The work presented in this paper was supported by the Hong Kong Research Grants Council (RGC) (No. T22-504/21-R), and in part by the Department of Science and Technology of Guangdong Province (GDST) (2023A1515010757).

## Declaration of interests

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

## Data and codes availability statement

The code and data are shared privately at https://figshare.com/s/46caea113650a8bfc03c for review purposes. The code and data will be made publicly available upon acceptance.

## Declaration of generative AI in writing

A university self-hosted LLM GenAI was used to assist with proofreading the work to correct grammatical and connection errors in the main text, and to improve readability of Abstract, Introduction, and Conclusion sections. We declare no use of GenAI to produce any new content in the work.

## References

Bahr, S., 2024. The relationship between urban greenery, mixed land use and life satisfaction: An examination using remote sensing data and deep learning. *Landscape and Urban Planning*, 251, 105174. doi:10.1016/j.landurbplan.2024.105174.

Biljecki, F. and Ito, K., 2021. Street view imagery in urban analytics and GIS: A review. *Landscape and Urban Planning*, 215, 104217. doi:10.1016/j.landurbplan.2021.104217.

Bolte, A.M., *et al.*, 2024. The green window view index: automated multi-source visibility analysis for a multi-scale assessment of green window views. *Landscape Ecology*, 39 (3), 71. doi:10.1007/s10980-024-01871-7.

CAAC, 2023. *National airspace basic classification method*. Civil Aviation Administration of China, Government of PR China. Available from: https://www.caac.gov.cn/XXGK/XXGK/TZTG/202312/P020231222621680839714.pdf.

Chauhan, S., *et al.*, 2025. Urban heat stress, air quality and climate change adaptation strategies in UK cities. *Frontiers of Engineering Management*, 1–17. doi:10.1007/s42524-025-4029-y.

Cheng, B., *et al.*, 2022. Masked-attention mask transformer for universal image segmentation. *In*: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 1290–1299. doi:10.1109/CVPR52688.2022.00135.

Cimburova, Z. and Blumentrath, S., 2022. Viewshed-based modelling of visual exposure to urban greenery–An efficient GIS tool for practical planning applications. *Landscape and Urban Planning*, 222, 104395. doi:10.1016/j.landurbplan.2022.104395.

Cimburova, Z., Blumentrath, S., and Barton, D.N., 2023. Making trees visible: A GIS method and tool for modelling visibility in the valuation of urban trees. *Urban Forestry & Urban Greening*, 81, 127839. doi:10.1016/j.ufug.2023.127839.

Doick, K.J., *et al.*, 2018. CAVAT (Capital Asset Value for Amenity Trees): valuing amenity trees as public assets. *Arboricultural Journal*, 40 (2), 67–91. doi:10.1080/03071375.2018.1454077.

Ecke, S., *et al.*, 2022. UAV-based forest health monitoring: A systematic review. *Remote Sensing*, 14 (13), 3205. doi:10.3390/rs14133205.

Elsadek, M., Liu, B., and Xie, J., 2020. Window view and relaxation: Viewing green space from a high-rise estate improves urban dwellers' wellbeing. *Urban Forestry & Urban Greening*, 55, 126846. doi:10.1016/j.ufug.2020.126846.

Fan, Z., Feng, C.C., and Biljecki, F., 2025. Coverage and bias of street view imagery in mapping the urban environment. *Computers, Environment and Urban Systems*, 117, 102253. doi:10.1016/j.compenvurbsys.2025.102253.

Gong, F.Y., *et al.*, 2018. Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. *Building and Environment*, 134, 155–167. doi:10.1016/j.buildenv.2018.02.042.

Guo, Y., *et al.*, 2021. Deep learning for 3D point clouds: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43 (12), 4338–4364. doi:10.1109/TPAMI.2020.3005434.

Han, Y., *et al.*, 2023. Mapping seasonal changes of street greenery using multi-temporal street-view images. *Sustainable Cities and Society*, 92, 104498. doi:10.1016/j.scs.2023.104498.

Helliwell, R., 2008. Amenity valuation of trees and woodlands. *Arboricultural Journal*, 31 (3), 161–168. doi:10.1080/03071375.2008.9747532.

HKBD, 2023. *Building information and age records*. Building Department, Government of Hong Kong SAR. Available from: https://data.gov.hk/en-data/dataset/hk-bd-opendata-building-information.

HKLandsD, 2022. *3D Photo-realistic Model*. Lands Department, Government of Hong Kong SAR. Available from: https://3d.map.gov.hk/.

HKLandsD, 2025. *Digital topographic map iB1000*. Lands Department, Government of Hong Kong SAR. Available from: https://www.hkmapservice.gov.hk/OneStopSystem/map-search?product=OSSCatB&series=iB1000.

HKPlanD, 2019. *3D Photo-realistic Model*. Planning Department, Government of Hong Kong SAR. Available from: https://www-pland-gov-hk.eproxy.lib.hku.hk/pland_en/info_serv/3D_models/download.htm.

HKPlanD, 2023. *Land utilization in Hong Kong*. Planning Department, Government of Hong Kong SAR. Available from: https://www.pland.gov.hk/pland_en/info_serv/open_data/landu/.

HKTD, 2019. *Road Network*. Transport Department, Government of Hong Kong SAR. Available from: https://static.data.gov.hk/td/road-network-v2/RdNet_IRNP.gdb.zip.

Huang, S., *et al.*, 2021. A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *Journal of Forestry Research*, 32 (1),

1–6. doi:10.1007/s11676-020-01155-1.
Huang, Y., *et al.*, 2025. No "true" greenery: Deciphering the bias of satellite and street view imagery in urban greenery measurement. *Building and Environment*, 269, 112395. doi:10.1016/j.buildenv.2024.112395.
Jiang, B., *et al.*, 2021. Impacts of nature and built acoustic-visual environments on human's multidimensional mood states: A cross-continent experiment. *Journal of Environmental Psychology*, 77, 101659. doi:10.1016/j.jenvp.2021.101659.
Ko, W.H., *et al.*, 2020. The impact of a view from a window on thermal comfort, emotion, and cognitive performance. *Building and Environment*, 175, 106779. doi:10.1016/j.buildenv.2020.106779.
Kong, F., Yin, H., and Nakagoshi, N., 2007. Using GIS and landscape metrics in the hedonic price modeling of the amenity value of urban green space: A case study in Jinan City, China. *Landscape and Urban Planning*, 79 (3), 240–252. doi:10.1016/j.landurbplan.2006.02.013.
Kong, L., *et al.*, 2017. Regulation of outdoor thermal comfort by trees in Hong Kong. *Sustainable Cities and Society*, 31, 12–25. doi:10.1016/j.scs.2017.01.018.
Konijnendijk, C.C., *et al.*, 2006. Defining urban forestry - A comparative perspective of North America and Europe. *Urban Forestry & Urban Greening*, 4 (3), 93–103. doi:10.1016/j.ufug.2005.11.003.
Lai, X., *et al.*, 2022. Stratified transformer for 3D point cloud segmentation. March. doi:10.1109/CVPR52688.2022.00831.
LeCun, Y., Bengio, Y., and Hinton, G., 2015. Deep learning. *Nature*, 521 (7553), 436–444. doi:10.1038/nature14539.
Li, M., *et al.*, 2026. Influence of objective and perceived exposures to urban nature on people's happiness. *npj Urban Sustainability*, 6 (1), 6. doi:10.1038/s42949-025-00306-9.
Li, M., *et al.*, 2023a. HRHD-HK: A benchmark dataset of high-rise and high-density urban scenes for 3D semantic segmentation of photogrammetric point clouds. *In: 2023 IEEE International Conference on Image Processing Challenges and Workshops (ICIPCW)*, October. 3714–3718. doi:10.1109/ICIPC59416.2023.10328383.
Li, M., *et al.*, 2022. A room with a view: Automatic assessment of window views for high-rise high-density areas using City Information Models and deep transfer learning. *Landscape and Urban Planning*, 226. doi:10.1016/j.landurbplan.2022.104505.
Li, M., Xue, F., and Yeh, A.G.O., 2023b. Bi-objective analytics of 3D visual-physical nature exposures in high-rise high-density cities for landscape and urban planning. *Landscape and Urban Planning*, 233, 104714. doi:10.1016/j.landurbplan.2023.104714.
Li, M., Xue, F., and Yeh, A.G., 2025. Efficient and accurate assessment of window view distance using City Information Models and 3D Computer Vision. *Landscape and Urban Planning*, 260, 105389. doi:10.1016/j.landurbplan.2025.105389.
Li, M., Yeh, A.G., and Xue, F., 2024. CIM-WV: A 2D semantic segmentation dataset of rich window view contents in high-rise, high-density Hong Kong based on photorealistic city information models. *Urban Informatics*, 3 (1), 12. doi:10.1007/s44212-024-00039-7.
Li, X., *et al.*, 2015. Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry & Urban Greening*, 14 (3), 675–685. doi:10.1016/j.ufug.2015.06.006.
Liang, J., *et al.*, 2017. Embedding user-generated content into oblique airborne photogrammetry-based 3D city model. *International Journal of Geographical Information Science*, 31 (1), 1–16. doi:10.1080/13658816.2016.1180389.
Liu, D., *et al.*, 2023. Establishing a citywide street tree inventory with street view images and computer vision techniques. *Computers, Environment and Urban Systems*, 100, 101924. doi:10.1016/j.compenvurbsys.2022.101924.
Meng, S., *et al.*, 2025. From 3D pedestrian networks to wheelable networks: An automatic wheelability assessment method for high-density urban areas using contrastive deep learning of smartphone point clouds. *Computers, Environment and Urban Systems*, 117, 102255. doi:10.1016/j.compenvurbsys.2025.102255.
Meng, S. and Zheng, H., 2023. A personalized bikeability-based cycling route recommendation

method with machine learning. *International Journal of Applied Earth Observation and Geoinformation*, 121, 103373. doi:10.1016/j.jag.2023.103373.

NCD Risk Factor Collaboration (NCD-RisC), 2016. A century of trends in adult human height. *eLife*, 5, e13410. doi:10.7554/eLife.13410.

Neema, M.N. and Ohgai, A., 2010. Multi-objective location modeling of urban parks and open spaces: Continuous optimization. *Computers, Environment and Urban Systems*, 34 (5), 359–376. doi:10.1016/j.compenvurbsys.2010.03.001.

Panduro, T.E. and Veie, K.L., 2013. Classification and valuation of urban green spaces-A hedonic house price valuation. *Landscape and Urban Planning*, 120, 119–128. doi:10.1016/j.landurbplan.2013.08.009.

Perperidou, D.G. and Kirgiafinis, D., 2022. Urban air mobility (UAM) integration to urban planning. *In*: *Conference on Sustainable Urban Mobility*. Springer, 1676–1686. doi:10.1007/978-3-031-23721-8_130.

Pyka, K., Piskorski, R., and Jasińska, A., 2022. LiDAR-based method for analysing landmark visibility to pedestrians in cities: case study in Kraków, Poland. *International Journal of Geographical Information Science*, 36 (3), 476–495. doi:10.1080/13658816.2021.2015600.

Roberto Barbosa, M., *et al.*, 2010. Forest fire alert system: a Geo Web GIS prioritization model considering land susceptibility and hotspots–a case study in the carajás national forest, brazilian amazon. *International Journal of Geographical Information Science*, 24 (6), 873–901. doi:10.1080/13658810903194264.

Rui, Q. and Cheng, H., 2023. Quantifying the spatial quality of urban streets with open street view images: A case study of the main urban area of Fuzhou. *Ecological Indicators*, 156, 111204. doi:10.1016/j.ecolind.2023.111204.

Seifert, E., *et al.*, 2019. Influence of Drone Altitude, Image Overlap, and Optical Sensor Resolution on Multi-View Reconstruction of Forest Images. *Remote Sensing*, 11 (10). doi:10.3390/rs11101252.

Song, Y., *et al.*, 2017. Trends and opportunities of BIM-GIS integration in the architecture, engineering and construction industry: A review from a spatio-temporal statistical perspective. *ISPRS International Journal of Geo-Information*, 6 (12), 397. doi:10.3390/ijgi6120397.

Szegedy, C., *et al.*, 2016. Rethinking the inception architecture for computer vision. *In*: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2818–2826. doi:doi.org/10.48550/arXiv.1512.00567.

Tang, L., *et al.*, 2023. Assessing the visibility of urban greenery using MLS LiDAR data. *Landscape and Urban Planning*, 232, 104662. doi:10.1016/j.landurbplan.2022.104662.

Tara, A., Lawson, G., and Renata, A., 2021. Measuring magnitude of change by high-rise buildings in visual amenity conflicts in Brisbane. *Landscape and Urban Planning*, 205, 103930. doi:10.1016/j.landurbplan.2020.103930.

Ulrich, R.S., 1984. View through a window may influence recovery from surgery. *Science*, 224 (4647), 420–421. doi:10.1126/science.6143402.

Wang, L., *et al.*, 2022. Measuring residents' perceptions of city streets to inform better street planning through deep learning and space syntax. *ISPRS Journal of Photogrammetry and Remote Sensing*, 190, 215–230. doi:10.1016/j.isprsjprs.2022.06.011.

Wang, Z., *et al.*, 2023. A view-tree method to compute viewsheds from digital elevation models. *International Journal of Geographical Information Science*, 37 (1), 68–87. doi:10.1080/13658816.2022.2094385.

Weiss, K., Khoshgoftaar, T.M., and Wang, D., 2016. A survey of transfer learning. *Journal of Big Data*, 3 (1), 9. doi:10.1186/s40537-016-0043-6.

Wild, G., 2024. Urban aviation: The future aerospace transportation system for intercity and intracity mobility. *Urban Science*, 8 (4), 218. doi:10.3390/urbansci8040218.

Xiang, S., *et al.*, 2024. Autonomous eVTOL: A summary of researches and challenges. *Green Energy and Intelligent Transportation*, 3 (1), 100140. doi:10.1016/j.geits.2023.100140.

Xie, J., *et al.*, 2025. Semantic segmentation of building facade materials and colors for urban conservation. *npj Heritage Science*, 13 (1), 378. doi:10.1038/s40494-025-01888-4.

Xue, F., Wu, L., and Lu, W., 2021. Semantic enrichment of building and city in-

formation models: A ten-year review. *Advanced Engineering Informatics*, 47, 101245. doi:10.1016/j.aei.2020.101245.

Yang, Y., *et al.*, 2020. Urban greenery, active school transport, and body weight among Hong Kong children. *Travel Behaviour and Society*, 20, 104–113. doi:10.1016/j.tbs.2020.03.001.

Yu, L. and Gong, P., 2012. Google Earth as a virtual globe tool for Earth science applications at the global scale: progress and perspectives. *International Journal of Remote Sensing*, 33 (12), 3966–3986. doi:10.1080/01431161.2011.636081.

Yu, S., *et al.*, 2016. View-based greenery: A three-dimensional assessment of city buildings' green visibility using Floor Green View Index. *Landscape and Urban Planning*, 152, 13–26. doi:10.1016/j.landurbplan.2016.04.004.

Yu, X., *et al.*, 2022. Spatio-temporal monitoring of urban street-side vegetation greenery using Baidu Street View images. *Urban Forestry & Urban Greening*, 73, 127617. doi:10.1016/j.ufug.2022.127617.

Zhang, J., *et al.*, 2024. Exploring geospatial digital twins: a novel panorama-based method with enhanced representation of virtual geographic scenes in Virtual Reality (VR). *International Journal of Geographical Information Science*, 38 (11), 2301–2324. doi:10.1080/13658816.2024.2386064.

Zheng, L., *et al.*, 2025. Harnessing geographic information system and street view imagery for thermal gradient distribution auditing. *Urban Climate*, 59, 102248. doi:10.1016/j.uclim.2024.102248.

Zhou, Q. and Xue, F., 2023. Pushing the boundaries of modular-integrated construction: A symmetric skeleton grammar-based multi-objective optimization of passive design for energy savings and daylight autonomy. *Energy and Buildings*, 296, 113417. doi:10.1016/j.enbuild.2023.113417.

## Appendix A. GVVI sub-indices distribution and correlation

Figure A1a illustrates the value distribution of $GVVI_{\mathrm{S}}$, $GVVI_{\mathrm{W}}$ and $GVVI_{\mathrm{D}}$. The majority of greenery had a $GVVI_{\mathrm{S}}$ of 0 (mean = 0.02, stdev. = 0.07), showing that the visual contribution of greenery volumes at the street level was highly imbalanced. The value of $GVVI_{\mathrm{W}}$ predominantly fell within the range of 0 to 0.05 (mean = 0.04, stdev. = 0.06), suggesting that the majority of greenery had low visibility at the window level. The distribution of $GVVI_{\mathrm{D}}$ was concentrated between 0 and 0.14 (mean = 0.10, stdev. = 0.12), which was reasonable because of the altitude change.

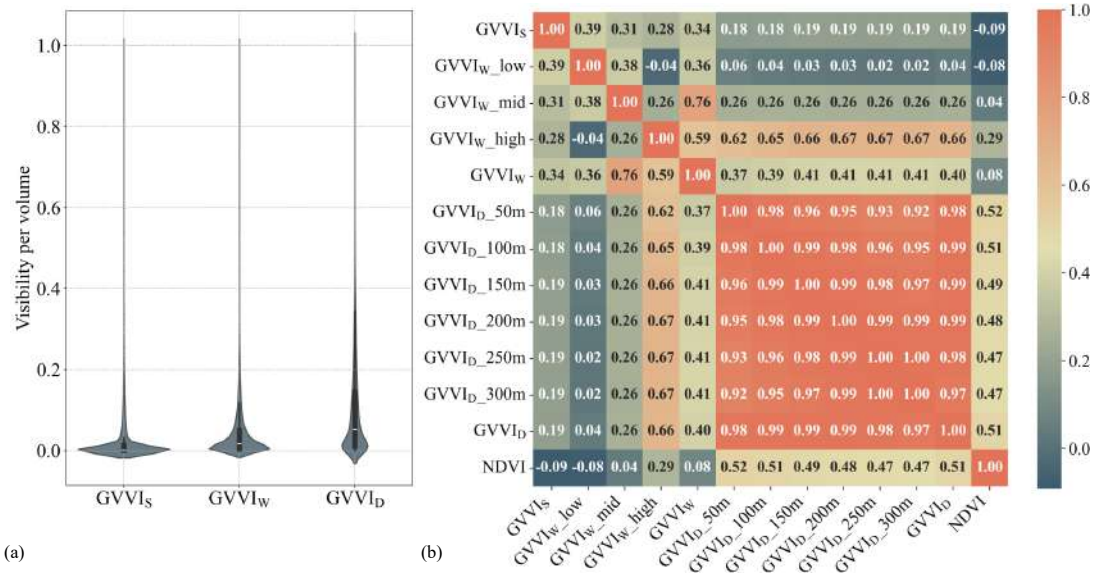


Figure A1.: Distribution of $GVVI$ sub-indices and correlation matrix among sub-indices with different sampling heights and NDVI. (a) Distribution; (b) correlation matrix, all p-values $< 0.0001$

The correlation of $GVVI$ computed based on different height levels of WVI (low: 0-24m, middle: 25-52m, and high: 53-137m, based on Natural Break) and DVI (50 to 300m) is conducted to validate the spatial dependency of sub-indices. The correlation between $GVVI_{\mathrm{W}}$ with increasing height levels and $GVVI_{\mathrm{S}}$ showed a diminishing trend, which logically considers the proximity of the viewpoints at lower height for $GVVI_{\mathrm{W}}$ and $GVVI_{\mathrm{S}}$. All coefficients of $GVVI_{\mathrm{W}}$ at different height levels with other indicators show positive but less moderate associations, indicating that GVV is less spatially dependent at near-ground level. The diverse height levels of $GVVI_{\mathrm{D}}$ exhibit strong associations among themselves, meaning that the vertical disparity of greenery within the research area was under 100m. It is therefore recommended to reduce the number of DVI sampling intervals, especially in areas with small terrain height differences, to reduce the computation cost.

## Appendix B. Comparison of GVV with different DVI FOV

Figure B1 shows the drone views from FOVs equal to 60 and 120 degrees at different sample points' height levels. A total of 17,424 DVIs (FOV = 60) and 8,712 DVIs (FOV = 120) were captured. In general, the average visible pixel of FOV 120 was similar to that of FOV 60, which was lower by 923 only (maximum = 111,929).

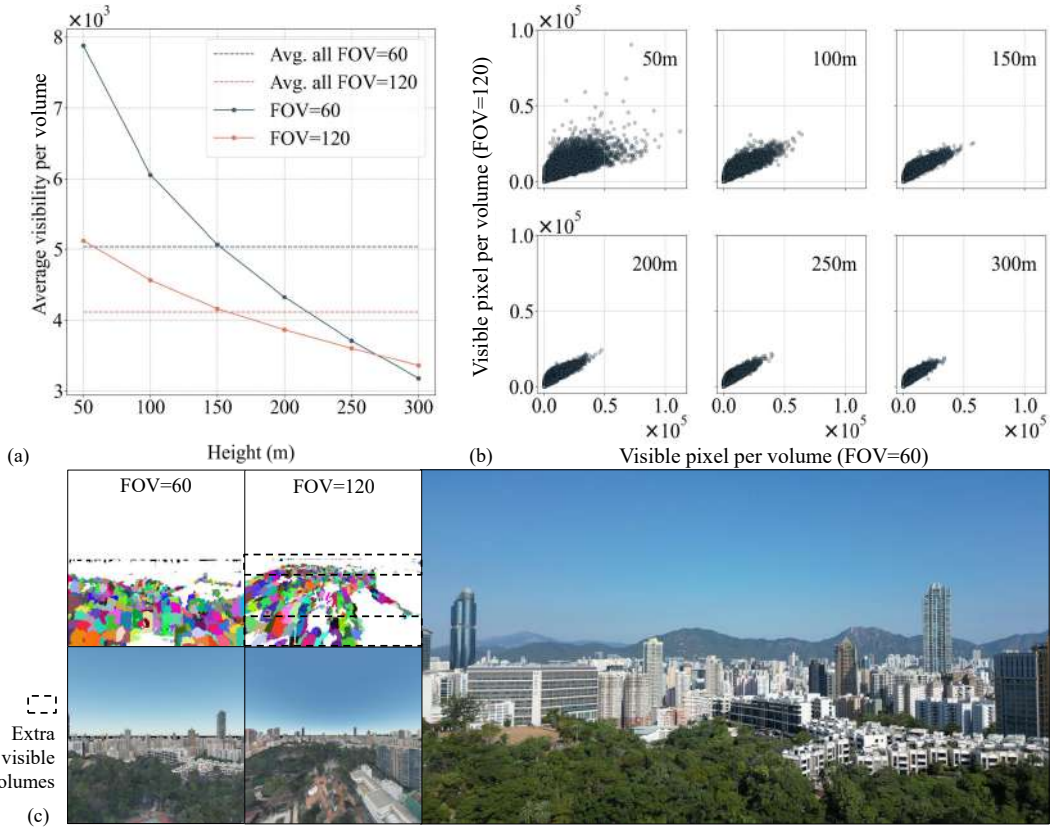


Figure B1.: Visibility difference between FOV equal to 60 and 120 at different heights. (a) average visible pixel per volume; (b) scatter plot of visible pixel per volume's difference; and (c) the DVIs at the same sample point with different FOV, which are displayed with GVL and photorealistic meshes, and DVI captured by real drone.

The result of the FOV of 60 degrees had a higher average visible pixel $N_D$ per volume at the height of 50m, with an average of 7,878 pixels, compared to the FOV of 120. However, the visibility of FOV 120 degrees exceeded that of FOV 60 with the increasing height, as shown in Figure B1a. This was because FOV of 120 degrees could capture greenery close to the bottom of the virtual camera as well as greenery further away (Figure B1c). The Spearman's coefficient was 0.92 for visible pixels per volume at 50m height. The Spearman's coefficient gradually stabilized at around 0.95 as the height increased, as depicted in Figure B1b. Hence, an FOV of 120 is suggested as the data source for calculating the drone-level visibility with high altitudes (e.g., altitude $>$ 100m), as it can capture visible greenery with reduced computational expenses (only half the time required compared to an FOV of 60).

## Appendix C. Sensitivity analysis of DVI sampling position parameters

Figure C1 shows the sensitivity analysis of DVI sampling position parameters impact on $GVVI_{\text{D-S}}$ hot spot differential. The horizontal and height intervals, as well as FOV changing caused less impact on $GVVI_{\text{D-S}}$ hot spot differential, with an average value of 36.3%. Different FOV mainly shared the similar trend on $GVVI_{\text{D-S}}$, but FOV equal to 120 showed a dramatic change curve. For FOV 120, the hot spot differential increased with higher sampling heights and smaller horizontal intervals, while for FOV 60, it decreased as the height increased; both FOVs showed similar trends with respect to the horizontal interval, as shown in Figure C1c and d. This may be caused by the building obstruction demonstrated in Figure B1c. Therefore, the proposed method can robustly evaluate the differential between different observer perspectives, and it is also suggested to evaluate $GVVI_{\text{D}}$ using multiple sampling height levels.

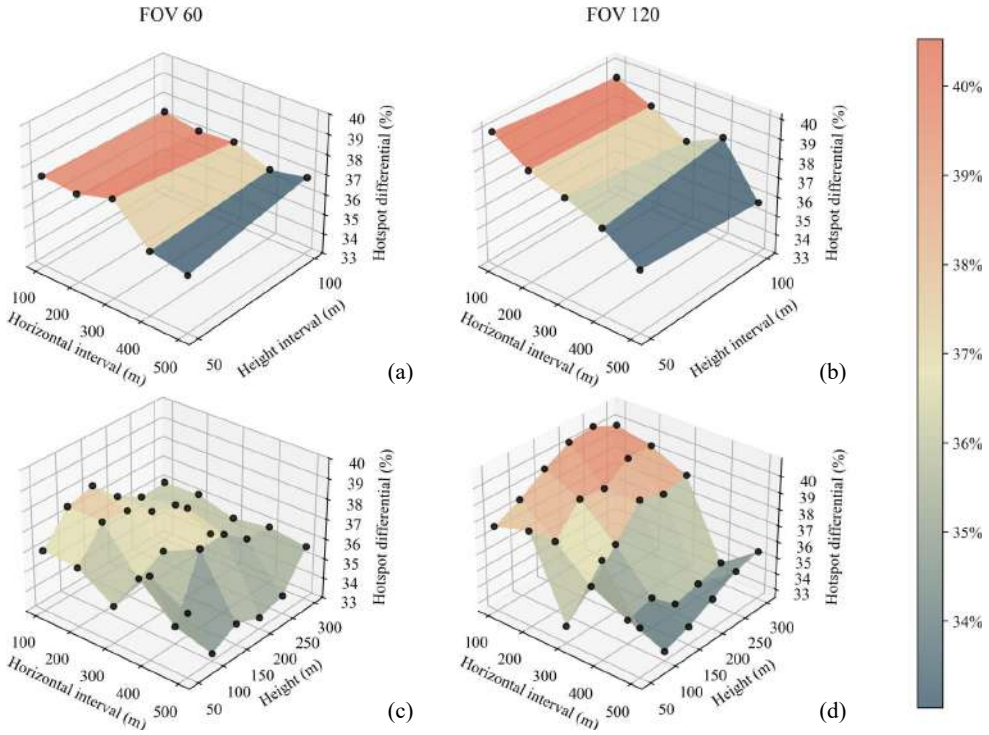


Figure C1.: Sensitivity analysis of DVI sampling position parameters impact on $GVVI_{\text{D-S}}$ hot spot differential, where (a) and (b) shows the horizontal, height interval comparison of FOV 60 and 120, (c) and (d) shows the horizontal, sampling height comparison of FOV 60 and 120.