Citywide assessment of perceived window view openness using photorealistic 3D model imagery and automatic machine learning

Maosu Li[1], Fan Xue[2], and Anthony G.O. Yeh[3*]

Abstract

High window view openness is valued by urban dwellers, especially in high-density cities, for stress relief and living satisfaction. Quantifying human-perceived window view openness is therefore significant for building and environmental applications, e.g., building envelope design, housing, and urban planning. However, existing measures of view openness are either objective, thereby overlooking human perception, or subjective, with limited scalability. Leveraging geovisualization of photorealistic 3D City Information Models (CIMs) and automatic Machine Learning (ML), we introduce a Human-Perceived Window View Openness Index (HPWVOI) together with a citywide assessment method. First, HPWVOI is defined on three calibrated objective view descriptors, i.e., proportions of sky and building views, and landscape distance. Then, we assess HPWVOIs via automatic linear ML regression based on openness ratings and objective view descriptors of representative window views generated from 3D CIMs. Experimental results in Hong Kong confirmed that the method was automatic to generate the representatives (≤ 3.13 s per image), the view descriptors (≤ 0.02 s per image), and the HPWVOIs (≤ 0.001 s per image). The linear ML models were accurate (R^2 = 0.9191/0.8040 in cross/independent validation) and interpretable for HPWVOIs of citywide windows. The 3D CIM-based approach with citywide quantitative perception evidence enables new solutions for decision-making in built environment improvement, generative architectural and urban design, and housing valuation.

Keywords: Window view; Visual openness; Human perception; City Information Modeling; Machine learning; 3D GIS

[1] Maosu LI, Assistant Professor.
Thrust of Urban Governance and Design, The Hong Kong University of Science and Technology (Guangzhou), Nansha, Guangzhou, China; Email: maosuli@hkust-gz.edu.cn, ORCID: https://orcid.org/0000-0002-9970-4053

[2] Fan XUE, Associate Professor.
Department of Real Estate and Construction, The University of Hong Kong, Pok Fu Lam, Hong Kong SAR, China; Hong Kong Branch, National Technology Innovation Center for Digital Construction, The University of Hong Kong, Pokfulam, Hong Kong SAR, China;
Email: xuef@hku.hk, ORCID: http://orcid.org/0000-0003-2217-3693

[3] Anthony G.O. YEH, Chair Professor.
Department of Urban Planning and Design, The University of Hong Kong, Pok Fu Lam, Hong Kong SAR, China; Hong Kong Branch, National Technology Innovation Center for Digital Construction, The University of Hong Kong, Pokfulam, Hong Kong SAR, China;
Email: hdxugoy@hku.hk, ORCID: http://orcid.org/0000-0002-0587-0588

[*]: Corresponding author, Tel: +852 3917 2721 Email: hdxugoy@hku.hk

1. Introduction

Windows are significant media for urban dwellers to access the outside world, especially regarding the large proportions (87%) of time people spend indoors (Li et al. 2022). Particularly, in densely populated cities, urban dwellers enjoy high window view openness for its physical and mental benefits, e.g., stress relief and increased life satisfaction (Kim et al. 2022; Wang & Munakata 2024; Li et al. 2026a). Generally, a high window view openness is characterized by high sky proportion and landscape distance, and low visual obstruction (Kaplan et al. 1989; Stamps III 2005). With 70% of the global population projected to live in urban areas by 2050 (UN 2018), high-density cities have been solutions worldwide for high living efficiency and convenience. However, urban planners face difficulties in providing adequate visual openness for all urban dwellers in densely populated cities, e.g., Hong Kong. Consequently, assessment of window view openness is vital for urban planners, designers, and other decision-makers to inform prioritized planning interventions, e.g., renewal of aging buildings with low view openness, for healthy high-density urban development.

Assessment of human-perceived window view openness has potential impacts for multiple building and environmental applications, e.g., human-centered built environment improvement, building mass design, and housing. The perceived view openness reflects urban dwellers' sense of how open the urban landscape feels through windows, implicitly integrating objective descriptors, e.g., sky proportion and landscape distance. Based on human-perception evidence, urban managers can identify aging buildings and city blocks with limited openness for prioritized improvements (Masoudinejad & Hartig 2020; Li et al. 2023; Li et al. 2026b). Building experts and architects can compare the perceived openness of design alternatives to support decisions that enhance urban dwellers' living satisfaction (Laovisutthichai et al. 2021). Property agents can add perceived openness into housing valuation and transaction aligned with buyer and render preferences (Ling 2022).

Despite growing evidence regarding the benefits of perceived window view openness, citywide assessment remains scarce because it is difficult to obtain large-scale observations of realistic window views and associated human evaluations. To examine correlations between objective view descriptors and human-perceived view openness, researchers in environmental psychology, architecture and built environment, and landscape management often took photos of window views in the real world and achieved findings through control experiments. Three typical correlated descriptors include proportions of sky views (Masoudinejad & Hartig 2020) and building views (Stamps III 2005; Ode et al. 2008) and view distance (Stamps III 2005; Krukar et al. 2021). However, they find it challenging to collect window views citywide to testify and amplify the findings' impacts, given high labor costs and privacy concerns (Li et al. 2022; 2025). Taking a photo with a camera is a high labor cost and even impractical to collect window views in the real world without room access permission. Although traditional 3D GIS visibility analytical tools, e.g., Line of Sight, can turn the real-world photography into a simulation, the current raytracing analyses on simplified cityscape models, e.g., LoD 1 buildings, identical trees, and Digital Terrain Models (DTMs) rendered generated window views unrealistic. The landscape abstraction's significant and identified bias (Daniel & Meitner 2001; Lange 2001) for the human observers' responses has hindered a reliable citywide assessment of perceived window view openness. Thus, a key research gap remains: how to systematically calibrate objective window view descriptors to estimate human-perceived view openness of citywide windows for healthy development in high-density cities.

Linear weightings for calibrations are often employed in multi-criteria analyses for urban applications, e.g., planning and design, given their simplicity and interpretability for public acceptance (Li et al. 2023). Compared with non-linear Machine Learning (ML) models, linear models, e.g., linear regression, calibrate multiple objective window view descriptors with interpretable equations and coefficients as weights. Therefore, this study presents a photorealistic 3D CIM-based linear ML approach to the assessment of human-perceived view openness at the citywide scale. First, a Human-Perceived Window View Openness Index (HPWVOI) is defined on three typical objective view openness descriptors, i.e., proportions of sky and building views and landscape distance, calibrated by human openness ratings. Then, we design a three-step assessment approach: i) automatic generation of citywide photorealistic and type-depth window views from photorealistic 3D CIMs, ii) collection of representative window views with quantified objective descriptors and perceived openness, and iii) automatic linear ML-based calibration of the descriptors to estimate HPWVOI. Last, a set of estimated HPWVOIs is visualized at apartment and building levels for potential urban applications, e.g., housing marketing and built environment improvement.

The contribution of this study is two-fold:

i. The defined HPWVOI represents human perception of window view openness based on three typical objective view openness descriptors. The HPWVOI combining existing descriptors with human ratings extends traditional unidimensional objective representations of window-view openness.

ii. Our 3D CIM-based method extends the scalability of the assessment of human-perceived window view openness. The generation of citywide window views from photorealistic 3D CIMs, together with their representatives, enable a comprehensive representation of human perception of window view openness. Linear ML regression offers models with satisfactory performance, together with interpretable equations and weights for citywide estimation of HPWVOIs.

The remainder of the paper is organized as follows. In Section 2, the literature is reviewed. In Section 3, we introduce the proposed HPWVOI, the citywide assessment method, and 2/3D visualization of HPWVOIs. The experiments are reported in Section 4. Last, we discuss and conclude the study in Sections 5 and 6, respectively.

2. Literature review

2.1 Studies on human-perceived window view openness

Researchers have studied the human perception of window view openness in environmental psychology, architecture, built environment, and landscape management. Assessed objective descriptors for window view openness include proportions of sky, building, ground, and greenery views and view distance (Stamps III 2005; Masoudinejad & Hartig 2020). Among these, the major impactful descriptors comprise two types: i) obstruction, i.e., building view proportion (Stamps III 2005), and ii) prospect, i.e., sky view proportion (Masoudinejad & Hartig 2020), and view distance (Krukar et al. 2021). By contrast, the impacts of other descriptors, e.g., proportions of ground and greenery views, varied with spatial context. For example, in high-rise, high-density urban areas, ground and greenery view proportions showed insignificant correlations with perceived openness, likely because many high-floor open window views contained little visible ground (Wang & Munakata 2023), while street trees and hillside vegetation often acted as visual obstructions (Stamps III 2005).

For the above research, the high labor cost of the experiments, particularly the difficulty of collecting realistic window views in large areas, limits the applicability of prior approaches to citywide assessment. First, experimental pipelines typically rely on manual steps, e.g., image collection, quantification of objective view descriptors, and statistical analyses (Lin et al. 2022). Second, respondent-rated window view images were often limited to either locally handmade 3D scenes or a few real-world sites, e.g., university campuses and communities (Abd-Alhamid et al. 2023). Consequently, statistical models (e.g., linear or logistic regression), based on local window views, were less generalized to other regions.

2.2 Current objective assessment of urban view openness

Landscape openness from lookouts and windows has been assessed in GIS and architecture via visibility analysis. For instance, the openness of landscapes from adjacent lookouts has been assessed by view distance through visibility analysis tools, e.g., *isovist* (Weitkamp et al. 2014) and viewshed (Pyka et al. 2022). The line-of-sight lengths between the observers and landscape elements have been aggregated into a 3D visible volume to represent landscape openness (Zhu et al. 2024). By contrast, street-level openness is also assessed from street-view imagery using sky-view proportion (Xia et al. 2021).

However, current approaches have failed to capture human-perceived openness of citywide window views for three observed reasons. First, single quantified descriptors, e.g., view distance and sky view proportion, are not equal to perceived openness. For instance, extremely short view distance (e.g., within 5 m) signifies enclosure (Stamps III 2005), whereas differences in perceived openness diminish beyond large distances (e.g., > 1,200 m) (Weitkamp et al. 2011). Second, current assessments fail to generate realistic window views for either human rating collection or citywide assessment: Visibility analyses typically output numeric ray tracing measures without photorealistic views, while street view imagery is car-centric and confined to the ground level (Li et al. 2025). Last, current visibility analyses for window view openness were mainly conducted on simplified models, e.g., cubic buildings, generic trees, and coarse terrains, which limited realism and scalability (Zhu et al. 2024; Li et al. 2025). Consequently, a research gap remains in the large-scale assessment of perceived window view openness, calling for the integration of next-generation 3D GIS techniques and datasets to generate realistic, citywide window views for perceived openness estimation.

2.3 Evolving photorealistic 3D City Information Modeling and machine learning

Photorealistic 3D CIMs are increasingly available to support multi-perspective urban view generation (Meng et al. 2026). These textured mesh models, reconstructed from aerial imagery, include products, e.g., Google Photorealistic 3D tiles (2023) covering 2,500 cities around the globe, and 3D maps of 266 cities from governmental agencies, e.g., Hong Kong Planning Department (HKPlanD) (2019), Hong Kong Lands Department (2019), Singapore Land Authority, and mainland China's 3D Realistic Geospatial Scene Scheme. The high-fidelity appearances bring simulations closer to reality (see Figures 1a and b). Geovisualization of photorealistic 3D CIMs through 3D WebGIS technologies, such as 3D Tiling, has enabled a wide range of urban applications, e.g., arbitrary-perspective generation of street and drone views for greenery assessment (Meng et al. 2026), and integrated visual–physical nature exposure analytics for urban renewal (Li et al. 2023).

Emerging photorealistic 3D CIMs also enable the citywide generation of window views. Certain researchers have tried to generate window views from photorealistic 3D CIMs (Li & Samuelson 2020; Li et al. 2021; Law et al. 2026). The basic idea is to place a virtual camera at a window location within a geovisualization platform (e.g., Cesium, ArcGIS Pro, or Unity), and then the rendered scene as its window view is generated using OpenGL-based graphics. Building on this workflow, researchers have quantified objective characteristics of window views openness, including view types (e.g., natural and urban) (Li et al. 2021), visual elements (e.g., greenery and buildings) (Li et al. 2022; Park & Ergan 2026), and viewing distances (Li et al. 2025). However, existing objective methods overlook human perception. The potential of generated high-quality and diverse window view imagery (see Figures 1, 7, and 10) has not been leveraged to collect respondents' openness ratings for citywide view assessment. Thus, the next generation of assessment methods should integrate photorealistic 3D CIMs, existing view generation pipelines, and human perception surveys to transform objective window view measurement into citywide perception assessment.

Evolving automatic ML techniques can calibrate objective view descriptors for perceived openness while balancing interpretability and accuracy. Compared with nonlinear models, e.g., an ensemble of trees and deep neural networks, linear models, e.g., ordinary least squares, ridge, and lasso, learn additive relationships between multiple indicators and the target, generating explicit interpretations, e.g., equations and coefficients (Jordan & Mitchell 2015; Li et al. 2023). The interpretations support practitioners' decision-making and improve acceptance of practices in multiple fields, e.g., aging-friendly built environment improvement (Liu et al. 2023). Additionally, the automatic hyperparameter tuning (e.g., grid search) can finetune parameters, e.g., regularization strength in ridge and lasso regressions, to maximize the predictive performance (Hastie et al. 2009). These automatic ML techniques have been applied in environment management and digital construction (Liang & Xue 2023).

In summary, citywide assessment of human-perceived window view openness remains limited, given the lack of citywide photorealistic window views and reliance on simplified visibility analysis. Emerging geovisualization of photorealistic 3D CIMs, together with automatic ML, opens the door to assessment of the human-perceived openness of window views citywide.

3. Research methods

3.1 Definition of HPWVOI

Figure 1c shows HPWVOI defined on three typical objective descriptors of window view openness, calibrated by human perception via Equation (Eq.) 1.

$$HPWVOI = F\,(f_{\text{sky}}, f_{\text{bldg}}, f_{\text{dist}}) \in [0, 1]. \qquad (1)$$

F is a linear calibration function bridging three objective descriptors and the *HPWVOI* of a photorealistic window view (see Figure 1b), a replica of the real-world photo shown in Figure 1a. The normalized values of *HPWVOI* range from 0 to 1, with higher values indicating greater perceived openness.

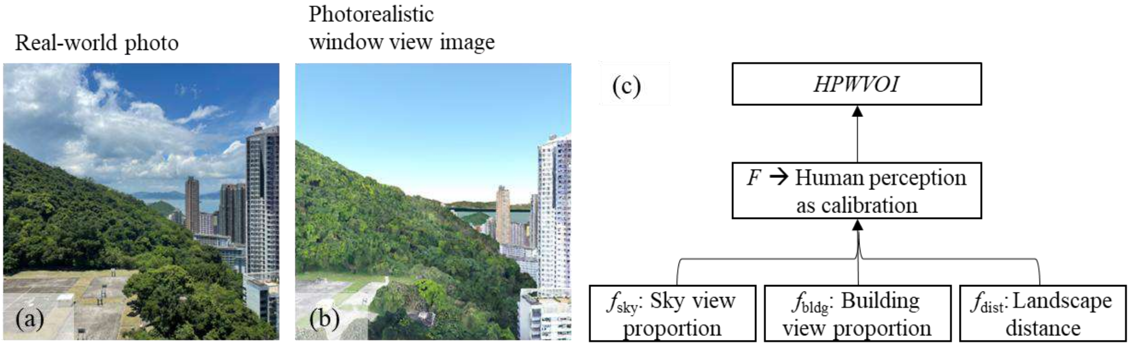

Figure 1. Definition of HPWVOI. (a) Real-world photo, (b) photorealistic view image as a replica, and (c) composition of HPWVOI.

We extract the three descriptors, f_{sky}, f_{bldg}, and f_{dist}, from the photorealistic window view image. The three descriptors quantify the proportion of the view covered by sky, the proportion covered by buildings, and the distance to non-sky landscape, as defined below,

$$f_{\mathrm{sky}} = f\,(sky) = f\,(\text{number of pixels belonging to sky} / n), f_{\mathrm{sky}} \in [0, 1], \tag{2}$$

$$f_{\mathrm{bldg}} = f\,(bldg) = f\,(\text{number of pixels belonging to building} / n), f_{\mathrm{bldg}} \in [0, 1], \tag{3}$$

$$f_{\mathrm{dist}} = f\,(dist) = f\,(\text{depths of } m \text{ pixels belonging to landscape}), m \leq n, f_{\mathrm{dist}} > 0, \tag{4}$$

where *sky*, *bldg*, and *dist* indicate sky view proportion, building view proportion, and an array of non-sky landscape distances, respectively, n is the total number of pixels in the window view image, and m is the number of pixels belonging to non-sky landscape elements (e.g., buildings, greenery, water bodies, and roads). f is a set of functions to transform the quantified sky, building, and depth attributes of the photorealistic window view image into multiple variants. We classify f into two categories based on their objectives: i) functions that alter the relationships between the descriptors and HPWVOI, e.g., the natural logarithmic function, *ln*, for establishing a linear relationship, and ii) functions that generate variants to identify the most sensitive to HPWVOI, i.e., the average (*avg*), maximum (*max*), minimum (*min*), sum (*sum*), and standard deviation (*std*) of the landscape-distance array. For example, the summed distance, *sum*(*dist*), refers to the sum of the three-dimensional distances from the window center to all non-sky landscape elements rendered in the image. It reflects the cumulative spatial extent of visible landscapes.

3.2 Proposed assessment method

Figure 2 presents the research framework of the proposed 3D CIM-based method. The inputs include photorealistic 3D CIMs and building footprints with heights and floor numbers (see Figure 2a). Specifically, floor numbers are used to identify sample window locations across different building levels. Figure 2b shows the proposed assessment method with three steps: i) automatic generation of citywide photorealistic and type-depth window views through 3D geospatial scene visualization, ii) collection of representative window views with quantified objective descriptors and perceived openness, and iii) automatic ML-based calibration of objective window view descriptors for HPWVOI. Figure 2c shows the output as a set of visualized HPWVOIs for 3D and 2D urban applications in housing marketing and built environment improvement.

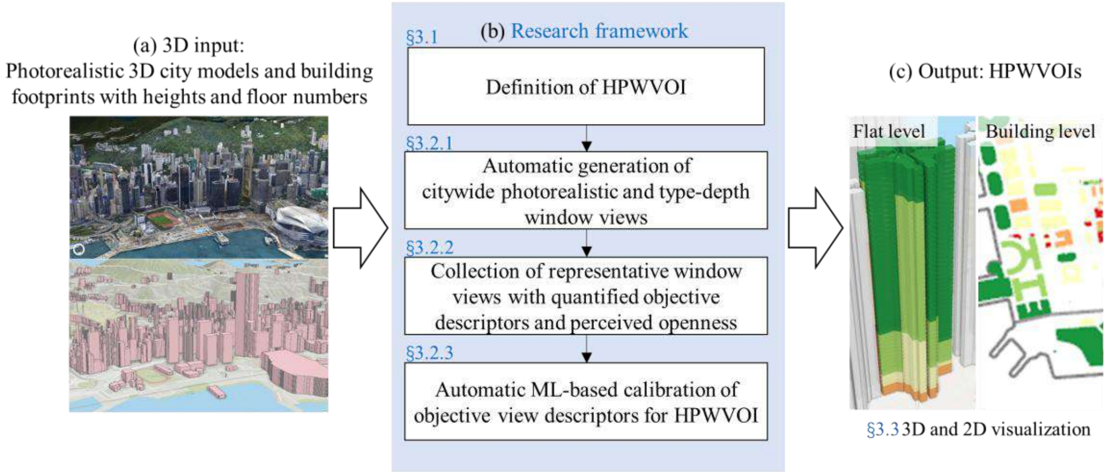

Figure 2. Research framework. (a) Input, (b) definition of HPWVOI and the three-step citywide assessment method, and (c) outputs: 3D and 2D visualized HPWVOIs.

3.2.1 Automatic generation of citywide photorealistic and type-depth window views

To support the collection of window view representatives as well as the objective descriptors and perceived openness in Section 3.2.2, we first generate citywide photorealistic and type-depth window views from photorealistic 3D CIMs. 3D CIM scene is rendered in two styles, i.e., the photorealistic textures and type-depth values, by customizing OpenGL shading (see Figures 3a and b). Specifically, photorealistic 3D CIMs are loaded and rendered on an empty scene of a 3D geospatial platform, e.g., Cesium. We use default shading to render the photorealistic scene (see Figure 3a). For the type-depth scene, we customize OpenGL shaders to encode landscape types and depth in the RGB channels (see Figure 3b), referring to Li's method (2025). Specifically, the R channel stores categorical codes for landscape types obtained via 3D semantic segmentation of the textured meshes (Li et al. 2025). The G and B channels together as a two-digit 256-base number encodes view depth from the depth test of the OpenGL rendering, mapped to 0 - 6,554 m with 0.1 m resolution (Li et al. 2025). Scene pixels outside 3D CIM coverage are rendered white (255, 255, 255), indicating sky (infinite distance).

The second step automatically captures photorealistic and type-depth views citywide (see Figures 3c and d) in the rendered 3D geospatial scene. Window locations (*lon*, *lat*, *altitude*) and orientations (*heading*, *pitch*, *roll*) can be derived from multiple types of datasets, e.g., BIMs, building footprints, floorplan drawings, and photorealistic 3D CIMs. The sampling strategy is tailored to cost and accuracy needs. For example, since neighborhood windows often share similar views, façade-level sampling can cost-effectively support building-level window view assessment without much accuracy loss (Li et al. 2022). Alternatively, a virtual camera can be placed at every window to render both photorealistic and type-depth views. For each window, the camera is positioned at (*lon*, *lat*, *altitude*) and oriented by *heading*, *pitch*, and *roll* to match the window's location and outward orientation, respectively. The Field of View (FoV) of the camera is typically set to 55 or 60 degrees to simulate the major scene that a normal person views. Last, citywide photorealistic and type-depth window views (see Figure 3d) are stored with geocoded locations, orientations, and building metadata (e.g., building code, floor No., and apartment No.).

The last step is dataset cleansing. We automatically detect and remove incompletely represented views, particularly for windows blended with nearby trees or building surfaces (see Figure 3e). The incomplete representations occur because photorealistic 3D CIMs only model exterior surfaces. When the virtual camera clips a surface, the missing interior modeling yields such images, as shown in Figure 3f. In addition, inconsistencies in spatial accuracy and acquisition time among multi-source input datasets (e.g., building footprints and 3D CIMs) may also cause misalignment between window sites and 3D building models, occasionally resulting in incompletely represented views. We apply computer vision models (e.g., DeepLab V3+) to extract their high-dimensional features, then train ML classifiers to detect recurring artifact patterns and exclude incompletely represented window view images finally.

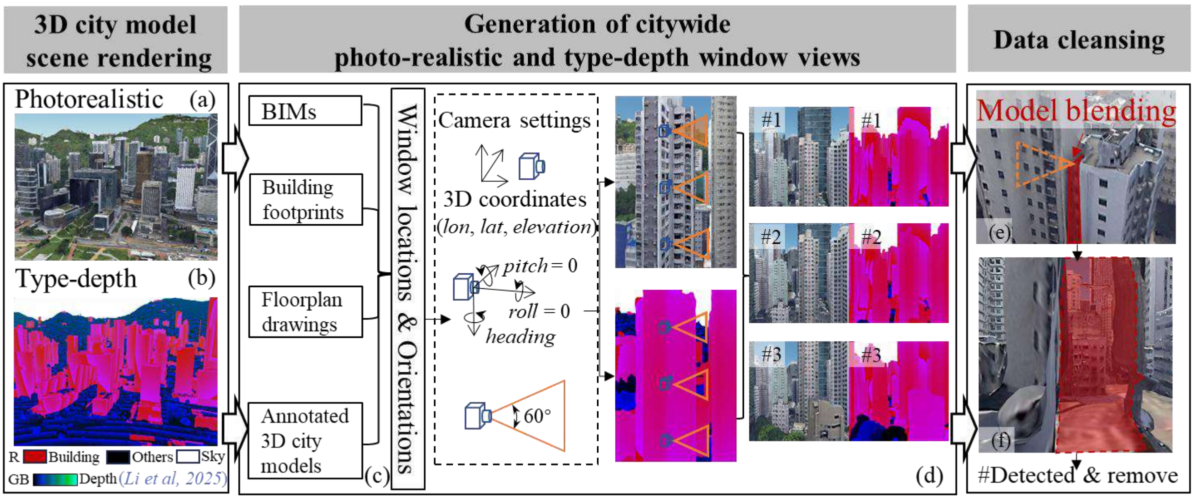

Figure 3. Automatic generation of citywide photorealistic and type-depth window views. (a)-(b) Photorealistic and type-depth scenes of photorealistic 3D CIMs, (c)-(d) generation of citywide window views in the rendered geospatial scene, and (e)-(f) data cleansing using a priori rule.

3.2.2 Collection of representative window views with quantified objective descriptors and perceived openness

The objective descriptors are first computed from citywide type-depth views (Figures 4a and 4b) for the subsequent organization of photorealistic window views. Specifically, f_{sky} and f_{bldg} are calculated using the ratio of white pixels (RGB 255, 255, 255) and pixels with R = 200 (see Equations 2 and 3). f_{dist} is obtained by converting each pixel's G and B values into a two-digit base-256 number (Eq. 4) to recover per-pixel view distance. To support the linear ML models in Section 3.2.3, the function set f for f_{sky} and f_{dist} includes a natural-logarithmic transform, while the function set for f_{dist} includes the average, maximum, minimum, sum, and standard deviation to derive multiple landscape distance indexes.

Then, we collect representatives of the citywide window views as shown in Figures 4c-e. Different from big data-driven analytical methods, e.g., each image of the MIT Place Pulse dataset (Salesses et al. 2013) rated by different persons anywhere online, following the classic psychophysics paradigm (Stevens 1957), we carefully selected stimuli through a control experiment and invited respondents to rate all images within the same room settings. The functional relationship is finally achieved between limited but controlled stimuli and perceptual responses. Regarding the similar distributions of descriptor values between the carefully selected representatives and the whole city's samples, the constructed ML model could be applied to estimate the human-perceived openness of the citywide window views. Specifically, stratified sampling is first applied in a three-dimensional descriptor space defined as $f_{\text{sky}} = ln\,(sky + 1)$, $f_{\text{bldg}} = bldg$, and $f_{\text{dist}} = ln(avg(dist))$ (see Figures 4b and d). These descriptors are selected regarding the observed linear relationships with perceived openness after natural logarithmic transformation and high correlations between variants of f_{dist} (Stamps III 2005; Weitkamp et al. 2014). We place equally spaced sampling points along each axis, x, y, and z, forming a grid of β^3 junctions. Rather than sampling entire cubic cells, we randomly select γ window view images in the vicinity of each junction to preserve spacing among samples. The values of β and γ are adjustable by researchers and practitioners to achieve the desired sample size.

Last, respondents' ratings on the perceived openness of window views are collected and cleansed for window view representatives using classic survey methods (see Figures 4e and f). For instance, respondents can assess the perceived openness of window views using a five-point scale (1 = Very low, 2 = Low, 3 = Average, 4 = High, and 5 = Very high), a continuous scale ranging from 0 to 100, or other rating scales. To validate reliability, we add a duplicate image as a gold-standard item and compare its ratings with the original's, discarding extremely inconsistent responses. Outlier ratings are also identified and removed using box-plot thresholds. In the end, the remaining ratings for each photorealistic window view are then averaged and normalized to generate the ground truth of HPWVOI.

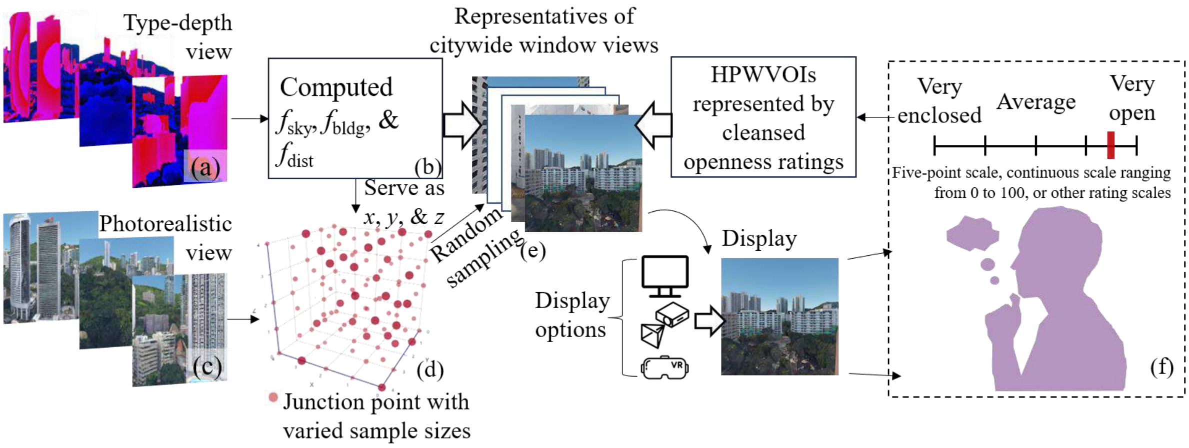

Figure 4. Collection of representatives of citywide window views and quantified objective and perceived view openness. (a) Citywide type-depth views together with (b) three quantified objective view descriptors, (c) photorealistic views organized in (d) 3D space constructed by (b), (e) collection of perceived openness as HPWVOI, and (f) outputs: quantified objective and perceived openness descriptors.

3.2.3 Automatic machine learning-based calibration of objective window view descriptors for HPWVOI

Multiple linear ML models serving as the function F in Eq. 1 are applied to calibrate three window view openness descriptors, i.e., f_{sky}, f_{bldg}, and f_{dist}, for HPWVOI, as shown in Figure 5a. For practitioners, e.g., urban planners, designers, and property agents, the HPWVOI's subjective nature requires both high accuracy and interpretability of the ML models for public acceptance. Thus, we test four types of linear models, i.e., Ordinary Linear Regression (OLR), Ridge Regression (RR), Lasso Regression (LR), and Elastic Net Regression (ER) with explicit equations. We control the number and function f of the three objective descriptors, i.e., f_{sky}, f_{bldg}, and f_{dist}, for the training of four types of ML models, as shown in Figure 5b. Specifically, we generate all possible combinations of the three types of descriptors for the performance comparison in Section 4.3. To avoid duplication, only one variant of a descriptor is applied in each combination. Thereafter, hyperparameters of each ML model are finetuned through automatic grid search and five-fold cross-validation to achieve unbiased optimal performance. Specifically, the grid search tests a list of values of α for RR and LR, and both α and L1_ratio for ER, to identify the optimal model. Last, we select the model with the optimal performance to estimate HPWVOIs of unseen photorealistic window views citywide.

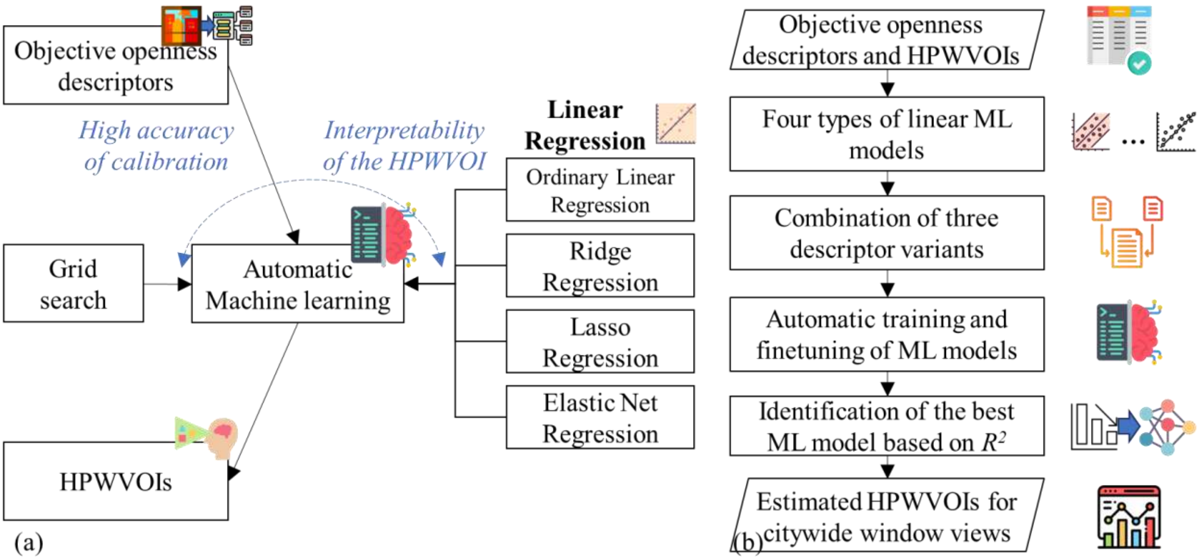

Figure 5. ML methods. (a) Four types of ML models for both high accuracy of calibration and interpretability of the HPWVOI, and (b) workflow of linear ML training and finetuning.

3.3 2/3D visualization of HPWVOIs for urban applications

We visualize HPWVOIs at the apartment and building levels using building footprints with height and floor data. For apartment-level 3D visualization, the building footprint is partitioned into apartment polygons based on floor layouts. Then, the apartment polygons are copied and placed at multiple floor positions. Thereafter, we extrude the apartment polygons according to the floor height. Last, we color the extruded apartment polygons using a weighted sum of HPWVOIs for major rooms, i.e., dining room, living room, and bedroom. For building-level 2D mapping, the entire footprint is colored by the mean HPWVOI of all sampled windows. The building-level HPWVOIs visualized in the Geographic Information Systems represent the overall perceived openness of the surrounding environment. Example visualizations are shown in Figures 11 and 12. Apartment-level HPWVOIs can guide buyers and renters to properties with more open views for living comfort, while building-level HPWVOI maps can highlight aging buildings and city blocks with limited view openness for prioritized improvements.

4. Experimental settings

We tested our method in eight districts of Hong Kong, i.e., Kowloon City, Kwun Tong, Sham Shui Po, Wong Tai Sin, Yau Tsim Mong, Central & Western, and Eastern, as shown in Figures 6a and b. These areas span the two highest-density zones defined by the Hong Kong Planning Standards and Guidelines (HKPlanD 2018). Mean and 75th-percentile building heights were 29.88 and 38.80 m, respectively, and the maximum permitted plot ratio reached 10 (HKPlanD 2018). The combination of high-rise development and hilly and waterfront terrain yields substantial variation in window view openness across floors and orientations. We collected the photorealistic 3D CIMs (HKPlanD 2019; HKLandsD 2019) in the form of 3D Tiles (.b3dm) (see Figure 6c), and the building datasets (HKLandsD 2023; HKLandsD 2025) in the format of .shp with height and floor numbers (see Figures 6d and e). The deep learning model, Stratified Transformer, was trained following Li's method (2025) and then used for 3D semantic segmentation of the photorealistic models.

Since neighborhood windows often share similar views (Li et al. 2022), we sampled window sites at ground, middle, and upper floors across 24,438 buildings (see Figure 12a), generating 933,561 photorealistic images and an equal number of type-depth counterparts (60° FOV; 900 × 900 pixels). Based on 500 randomly selected samples, we validated the alignment between sampled viewpoints and matched window centers (median distance = 1.27 m; 95.4% and 99.2% within 5 m and 20 m, respectively), indicating acceptable spatial accuracy for window-view quantification (Li et al. 2022). An 800-image audit of the cleansing pipeline using DeepLab V3+ embeddings and an ensemble ML achieved satisfactory accuracy, with an Area Under the Precision-Recall Curve (AUPRC) of 0.90. After filtering, 767,717 photorealistic (and matched type-depth) images remained. Following the classical psychophysical paradigm (Stevens 1957), we tried to build the regression relationships between three objective openness descriptors and human perceived window view openness through carefully controlled stimuli. We selected the representatives as stimuli via a stratified sampling. Specifically, we set $\beta = 5$ (points per axis) and defined grid points for $ln(sky + 1)$, $bldg$, and $ln(avg(dist))$ at the x, y, and z axes as {0, 0.10, 0.21, 0.31, 0.41}, {0, 0.25, 0.5, 0.75, 1}, and {2.97, 4.02, 5.07, 6.12, 7.17}, respectively. The grid junction buffers were set to 0.03 for $ln(sky + 1)$, 0.03 for $bldg$, and 0.03 for $ln(avg(dist))$. To reduce respondent burden, we set $\gamma = 1$ (image per junction) to curate 50 quality-enhanced photorealistic window view images (see Sections 4.1.1 and 4.1.3). To showcase HPWVOI's applications in housing marketing and built environment improvement, we annotated window locations and averaged room-level HPWVOIs (living, dining, and bedroom) to generate a floor-level 3D map (see Figure 11). HPWVOIs across sample building floors were averaged for a citywide building-level 2D mapping (see Figure 12). For automated linear ML model tuning through a five-fold cross validation, we determined α ranged from 0.0001 to 1 and L1_ratio from 0 to 1 following Hastie's (2009) statistical learning practices. Each hyperparameter was discretized into 20 values, regarding a balance between resolution to capture performance variations and computational efficiency during grid search.

For a preliminary test, we ran an in-house survey in May 2024 with 70 respondents aged 18-40, all university students or staff with at least one year of residence in Hong Kong. The survey was approved by the Human Research Ethics Committee before commencement, with the reference number as EA210471. As part of the eligibility screening, all respondents confirmed that they were in good health, had normal cognitive function, and had no vision-related impairments (e.g., partial blindness or color blindness). Individuals who did not meet these criteria were excluded before completing the online registration. Participants were notified of their rights via a signed informed consent. Respondents rated the openness of 50 selected window views through a continuous scale ranging from 0 to 100 on the same-size screens of assigned computers in a classroom setting. Before the formal rating task, participants were provided with standardized instructions on how to navigate through the slides and rate the openness of the window-view images. During the rating process, both ratings and demographics (e.g., age, gender, residential location, and educational background) were recorded. A duplicate image was included among the 50 images as a gold-standard quality control item to evaluate the consistency of individual openness ratings but was excluded from the regression analysis due to the violation of the assumption of independent observations. Participants whose ratings of the duplicated image showed poor consistency were further reviewed, and those deemed to have provided unreliable responses were excluded from subsequent analyses. Last, we evaluated the duplicate-item consistency, inter-rater reliability, and the uncertainty of the image-level mean ratings from the remaining 70 respondents.

The 70 samples met the sample-size requirement corresponding to a 90% confidence level and a margin of error below 10%, and their 1:1.19 male-to-female ratio closely matched the gender distribution of Hong Kong's 18–40 population. The self-rated openness of the windows most frequently used at participants' homes or workplaces was distributed across three categories, i.e., high, average, and low, in a ratio of 25:24:21, towards a balanced perception representation across different everyday window-view environments. We removed outlier ratings per image using methods in Section 3.2.2 and used the normalized mean ratings as the HPWVOI ground truth. Last, to evaluate the generalizability of the proposed linear model, an independent validation dataset comprising 125 window-view images was collected through hierarchical sampling by *sky*, *bldg*, and *dist*, each divided into five intervals. These images were independent of the original 49 images and were not involved in feature engineering, model development, or parameter tuning. Ten new participants independently rated the perceived openness of the images in a laboratory survey, and the original regression equation was directly applied without recalibration.

The experiments were run on a Windows 10 (64-bit) workstation with an Intel i9-11900K CPU (16 cores), 64 GB RAM, and an Nvidia GeForce RTX 3090 graphics card (24 GB). Photorealistic and type-depth window views were generated on Cesium (ver. 1.99) using Firefox Browser (ver. 139.0.1). Descriptor quantification (f_{sky}, f_{bldg}, and f_{dist}) from type-depth window view images was performed through a customized Python (ver. 3.7.11) program. Last, linear ML regressions mapping three descriptors to HPWVOI were implemented with the Python library, scikit-learn (ver. 1.4.2). The 3D and 2D maps were visualized in ArcGIS Pro (ver. 2.9.0).

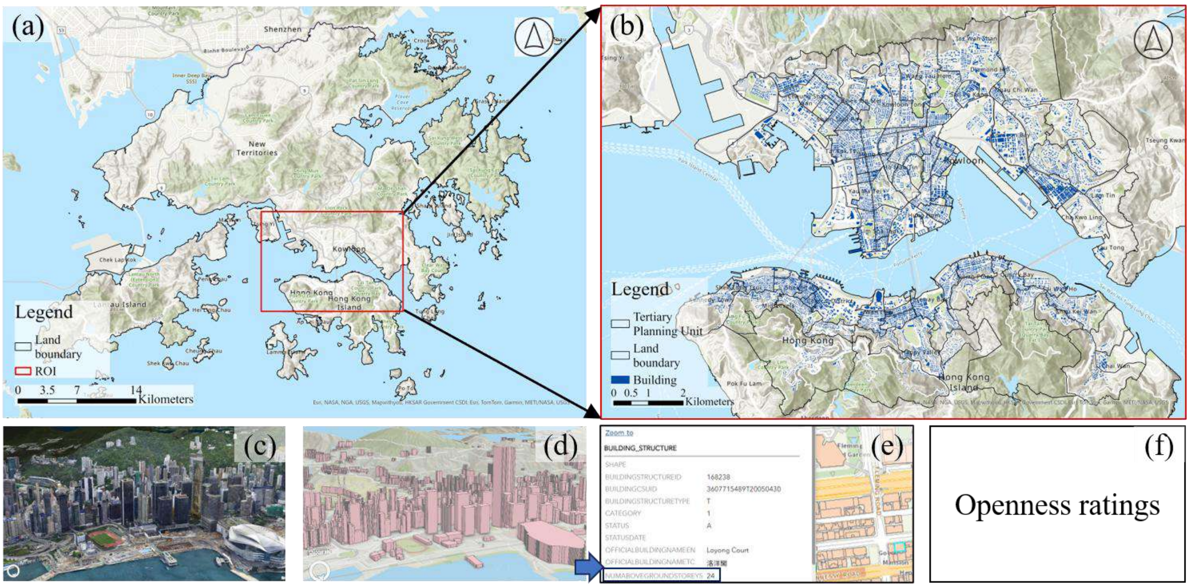

Figure 6. Study area and input data. (a)-(b) High-density urban area, eight districts of Hong Kong, (c)

photorealistic CIMs, (d)-(e) building footprints with attributes of height and floor No., and (f) openness ratings.

4.1 Results

4.1.1 Reliability of openness ratings and representativeness of selected window views

The two duplicated images demonstrated good test–retest consistency, with a mean absolute difference of 6.09 points (on a 100-point scale) and a Pearson's r at 0.803. Across the 50 images, the reliability of individual ratings was moderate (ICC (2,1) = 0.654), whereas the reliability of mean ratings aggregated across all 70 respondents was excellent (ICC (2,70) = 0.992), which also indicated the importance of averaging for substantially reduced individual rating variability. Besides, the uncertainty of the image-level mean openness ratings was also small, with a mean standard error of 1.88 points and an average 95% confidence interval width of 7.51 points (range: 3.34–10.35 points), suggesting that the mean openness ratings achieved satisfactory precision.

Figure 7 shows the distribution of 50 representative window view images within a 3D descriptor space parameterized by $ln(sky + 1)$, $bldg$, and $ln(avg(dist))$. Figure 7a shows 8,649 photorealistic window view samples distributed around 47 grid junctions. Junction sample sizes, ranging from 1 to 4,937, were grouped into five categories via a log10 transform. Close-range building views dominated the dataset, while there existed few distant views with high proportions of sky elements. Figure 7b shows the selected 50 photorealistic window view images spanning all categories, demonstrating the selected images' representation of the diversity of citywide window views. Two additional randomly collected images were from the categories ($ln(sky + 1) = 0.0$, $bldg = 1.0$, $ln(avg(dist)) = 2.97$), and ($ln(sky + 1) = 0.0$, $bldg = 1.0$, $ln(avg(dist)) = 4.02$). The last image was the duplicate one as the Gold Standard to validate the survey results. Generally, window view images with higher values of $ln(sky + 1)$ and $ln(avg(dist))$, and lower building proportions ($bldg$), received higher averaged openness ratings. The ratings ranged from 4.82 to 96.2 (on a scale from 0 to 100) and were approximately normally distributed according to the Shapiro–Wilk test ($W = 0.980$, $p = 0.550$).

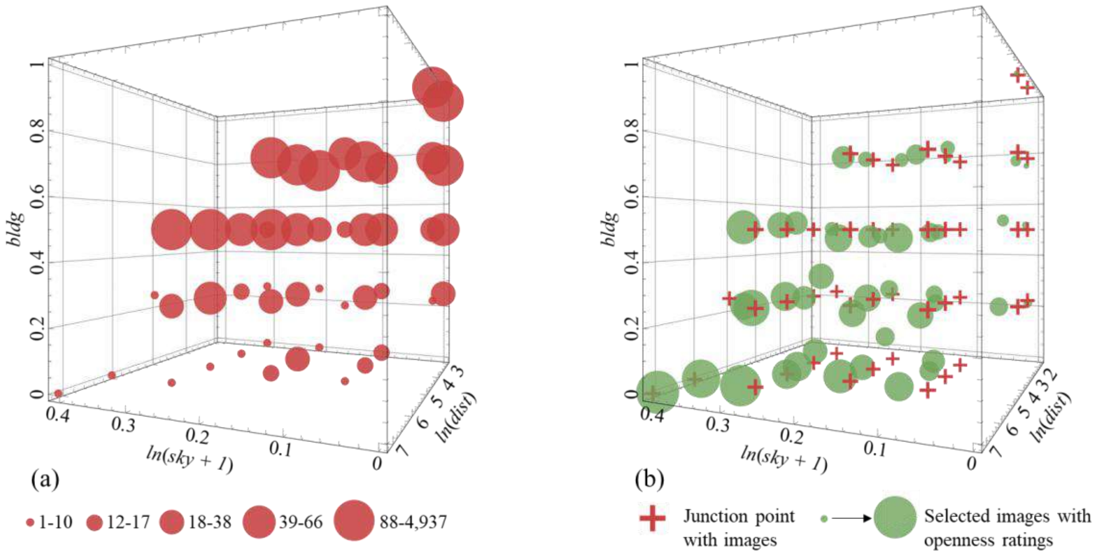

Figure 7. Representativeness of window views. (a) Distribution of window views around grid junction points by $ln(sky + 1)$, $bldg$, and $ln(dist)$, and (b) distribution of selected images spanning all categories.

Compared with the controlled variables (i.e., sky proportion, building proportion, and landscape distance), Figure 8 presents the distributions of the uncontrolled scene attributes, including greenery, water, ground proportions, average brightness, and visual complexity, for both the selected stimuli and the citywide window-view dataset. Overall, the distributions of the selected stimuli closely resembled those of the citywide dataset (Pearson's $r = 0.859$–0.995), with similar right-skewed patterns for greenery, water, ground, and visual complexity (see Figures 8a-c and e-f), and a comparable unimodal distribution for brightness (see Figure 8d). Moreover, the selected stimuli provided coverage across nearly all value bins, with at least one representative image in almost every interval. These results suggest that the selected stimuli remained broadly representative of citywide window-view images, without substantial bias toward the three controlled scene characteristics.

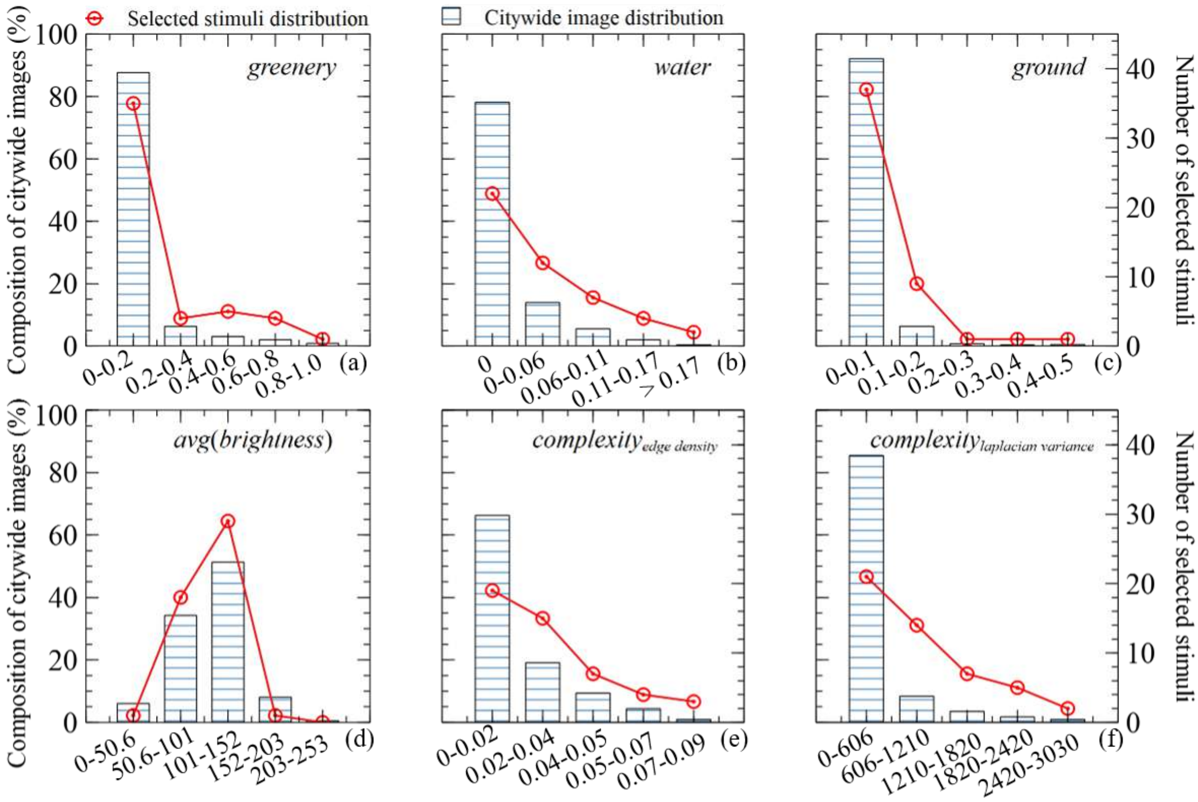

Figure 8. Distributions of (a) greenery, (b) waterbody, and (c) ground view proportions, (d) brightness, and (e)-(f) visual complexity between stimulus and citywide images

4.1.2 Efficiency and accuracy of the calibration

The generation of each photorealistic view and type-depth view consumed 3.13 s and 0.99 s, respectively. Computing the 13 variables for the three objective descriptors (using f defined in Equations 2, 3, and 4) costs 0.02s from a type-depth view image. The linear ML-based estimation of the HPWVOI consumed less than 0.001 s for each unseen photorealistic window view. The StratifiedTransformer achieved satisfactory performance on detecting building elements from the photorealistic 3D CIMs (per-class IoU = 0.93). For quantification of sky elements in type-depth views, counting white pixels (RGB 255, 255, 255) introduced no semantic segmentation errors. The error in the estimated window view distance resulted from scene depth decoding ($\leq$ 0.1 m) and the spatial resolution of the photorealistic 3D CIMs (0.3 m horizontally and 0.5 m vertically).

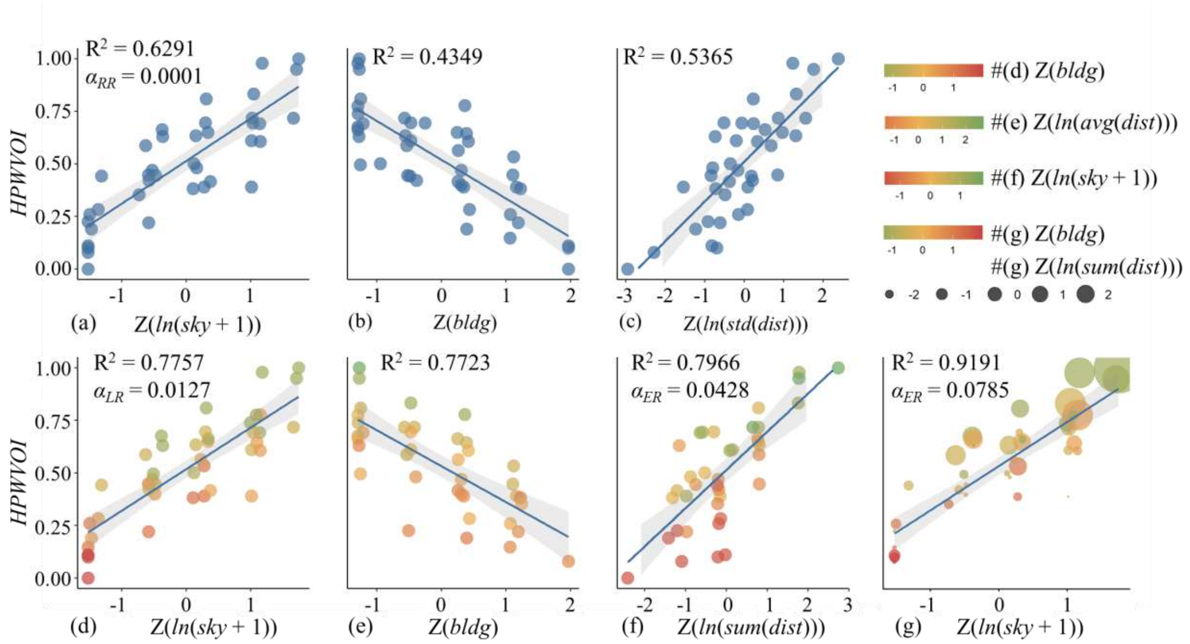

Figure 9. Calibration of single, dual, and triple objective view descriptors for HPWVOI via linear machine learning.

Figure 9g shows that linear ML models trained on three standardized features, i.e., Z(*ln*(*sky* + 1)), Z(*bldg*), and Z(*ln*(*sum*(*dist*))), achieved the best cross-validated performance ($R^2 = 0.9191$). This likely reflected that total landscape distance, together with the sky view proportion, captured the extent of visible landscape, while the building proportion captured vertical obstruction. Compared with Z(*ln*(*avg*(*dist*))), the use of Z(*ln*(*sum*(*dist*))) resulted in a slightly higher predictive performance ($\Delta R^2 = 0.0032$). Although the difference was small, we retained Z(*ln*(*sum*(*dist*))) in the final model for more accurate estimation. By contrast, models trained on any two descriptors showed inferior performance ($R^2 = 0.7757$; 0.7723; 0.7966), and single-descriptor models, i.e., Z(*ln*(*sky* + 1)), Z(*bldg*), or Z(*ln*(*std*(*dist*))), the poorest (R^2 ranged between 0.4349 and 0.6291). Note that, the summed distance, *sum*(*dist*), is not equivalent to the complement of sky proportion, *sky*, regarding its addition incorporation of the distance of each visible landscape pixel. For example, two views with identical sky proportions may have substantially different *sum*(*dist*) if their visible landscapes occur at different viewing distances. Example window views are shown in Figures 10a, f, and g.

4.1.3 Interpretability of linear machine learning models and HPWVOIs

The linear models with explicit equations were interpretable, supporting practitioners' decision-making and public awareness. From the highest-performing ER model (see Figure 9g), the three descriptors had comparable influence on respondents' perceived openness: *ln*(*sky* + 1) ranked the highest ($\beta = 0.11$, N = 49), then *bldg* ($\beta = -0.10$, N = 49), and last the *ln*(*sum*(*dist*)) ($\beta = 0.07$, N = 49). Instead of setting the weight ratios empirically, assigning the weight ratios for Z(*ln*(*sky* + 1)), Z(*bldg*), and Z(*ln*(*sum*(*dist*))) as 0.11, – 0.10, and 0.07 in Eq. 5 was reasonable for citywide estimation of HPWVOI, given quantitative evidence from human openness ratings on representative window views. For *ln*(*sky* + 1), *bldg*, and *ln*(*sum*(*dist*)) without standardization, the constant value was -0.55 while the weight ratios were 0.77, -0.33, and 0.08, respectively.

$$HPWVOI = 0.52 + 0.11 \times Z(ln(sky + 1)) - 0.10 \times Z(bldg) + 0.07 \times Z(ln(sum(dist))) \qquad (5)$$

External validation was performed by directly applying Eq. (5) to the independent dataset comprising 125 window-view images. The image-level mean openness ratings demonstrated excellent reliability (ICC (2,10) = 0.957) and low estimation uncertainty, with a mean standard error of 4.15 points on the 100-point scale. The predicted HPWVOI values showed strong agreement with the corresponding mean openness rating scores ($R^2 = 0.804$, Pearson's $r = 0.932$, RMSE = 0.114). These findings further suggest that the proposed model had a satisfactory generalization to previously unseen data in the study area.

Group #1

	(a)	(b)	(c)	(d)	(e)
sky	0.0000	0.1299	0.2569	0.3719	0.4948
bldg	1.0000	0.4899	0.4710	0.2507	0.0060
sum(dist)	50,845	484,375	1,193,101	2,911,267	6,737,648
HPWVOI	0.0000	0.4179	0.6020	0.8325	1.0000

Group #2

	(f)	(g)	(h)	(i)	(j)
sky	0.0000	0.0030	0.1021	0.1228	0.2217
bldg	0.7222	0.7232	0.7774	0.7608	0.7719
sum(dist)	225,355	556,958	493,954	212,035	129,602
HPWVOI	0.1602	0.2370	0.3014	0.2558	0.2943

Figure 10. Example photorealistic window view images. Group #1: window views (a)-(e) with increasing sky view proportions and overall landscape distances and decreasing building view proportions, and Group #2: window views (f)-(j) with three indicators, none of which dominate the others.

To demonstrate the effectiveness of the proposed HPWVOIs, Figure 10 shows two groups of photorealistic window view images with estimated HPWVOIs using the ER model (g) (see Figure 9). First, HPWVOIs quantified and differentiated perceived openness. Examples are increasing HPWVOIs of five photorealistic window view images with varied sky and building proportions, and landscape distances in Group #1. Particularly, HPWVOIs resolved trade-offs where no single descriptor dominates (see Figures 10f-j). For instance, in Group #2, view (j) had the most sky but the shortest landscape distance; View (g) had the largest landscape distance but almost no sky. Views (h) and (i) had building proportions similar to (j) but greater distances and less sky; and view (f) had the fewest building obstructions, but no visible sky, and the second-shortest distance. Aligned with respondents' perceptions, HPWVOI ranked view (h) highest due to its combination of generally large landscape distances and sky view proportions. Although view (h) had a smaller sky view proportion and a slightly higher building proportion than views (j) and (i), which had the second and third highest HPWVOIs, the overall landscape distance of view (h) was decisive, exceeding that of (j) by more than threefold and (i) by twofold.

4.1.4 3D and 2D visualization results

Figures 11 and 12 show the 3D visualizations of HPWVOIs at the apartment and building levels, respectively. Apartment-level HPWVOIs for individual buildings can support housing marketing. For example, Figure 11 shows HPWVOIs of a 63-story building with eight units in the Central and Western District, Hong Kong. Generally, most apartments of Units #7, #8, and #2 consistently had high HPWVOIs due to fewer visual obstructions (see Figures 11 b and e). Such apartments can be candidates for purchasers and renters seeking high perceived openness. By contrast, apartments on the middle and bottom floors of Units #1, #3, #4, #5, and #6 had low HPWVOIs because of obstruction by adjacent buildings. These quantitative disparities in HPWVOIs could potentially be incorporated into property valuation systems as supplementary information for housing valuation and pricing.

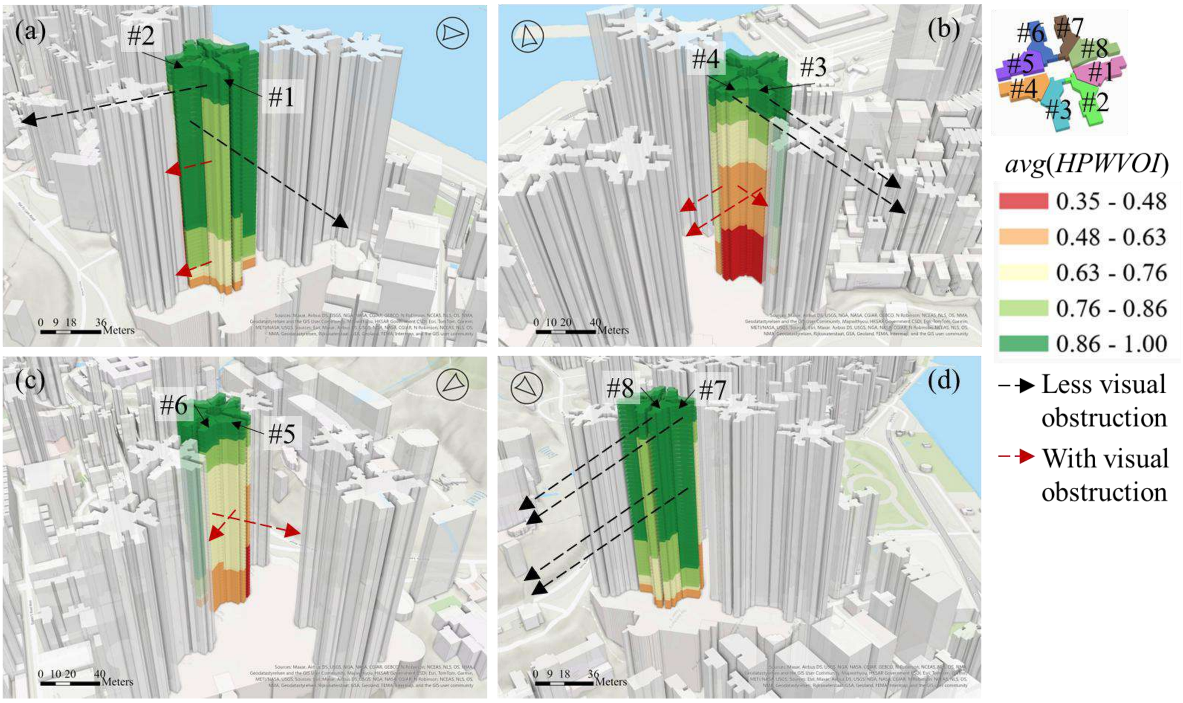

Figure 11. 3D visualization of apartment-level HPWVOIs for (a) Units 1-2, (b) 3-4, (c) 5-6, and (d) 7-8.

The 2D map of building-level HPWVOIs (see Figure 12) illustrated the disparity of perceived openness in the study area. Hilltop and seaside buildings generally had high HPWVOIs due to fewer visual obstructions. By contrast, five clusters of low-HPWVOI buildings were concentrated inland. Particularly, aging buildings in four Tertiary Planning Units (TPUs) on Hong Kong Island and five TPUs in the Kowloon Peninsula with consistently low HPWVOIs (polygons with black points in Figure 13f) may represent potential priority areas for district-level urban renewal, where interventions could include planning new open spaces to offset low perceived view openness or district redevelopment to improve overall visual openness.

4.2 Comparative analysis between objective and perceived openness indexes

We compared the assessed HPWVOIs with two objective openness measures—*sky* and visible volume—commonly used in street view and urban landscape analysis (see Section 2.2). We applied a constructed visible volume using *sum*(*dist*) and set the distance to sky at 1,200 m (Weitkamp et al. 2011). In the hot spot analysis, Figures 13a-c show that HPWVOIs differed from the *sky* in five city blocks, and from the visible volume in two city blocks (#1 and #4). The observed major reason was that the objective openness measures omitted key descriptors of human-perceived openness: *bldg* and *dist* for *sky*, and *bldg* for visible volume. For instance, blocks #2, #3, and #5 showed different identifications between *sky* and HPWVOIs because the *sky* metric failed to represent varied landscape distances in blocks #2 and #5 and the significantly short landscape distances in #3 (see Figure 13b). Additionally, Figures 13b and c show differences in blocks #1 and #4 between HPWVOIs and the visible volume. According to the calculated *bldg* and *sum*(*dist*) of HPWVOIs, block #4 had high HPWVOIs owing to low building proportions, while block #1 had varied HPWVOIs due to varied *sum*(*dist*).

At the TPU level, the proposed HPWVOIs also produced different local patterns. Compared to HPWVOIs as shown in Figure 13f, rectangles in Figures 13d and e showed five additional TPUs identified by *sky* and visible volume, and one TPU unidentified by *sky*. There existed four TPUs in the Kowloon Peninsula and four TPUs in the Hong Kong Island consistently identified by three measures (see black polygons in Figures 13d-f). Regarding human perception, the HPWVOIs offered additional quantitative evidence to prioritize planning interventions in the eight city blocks.

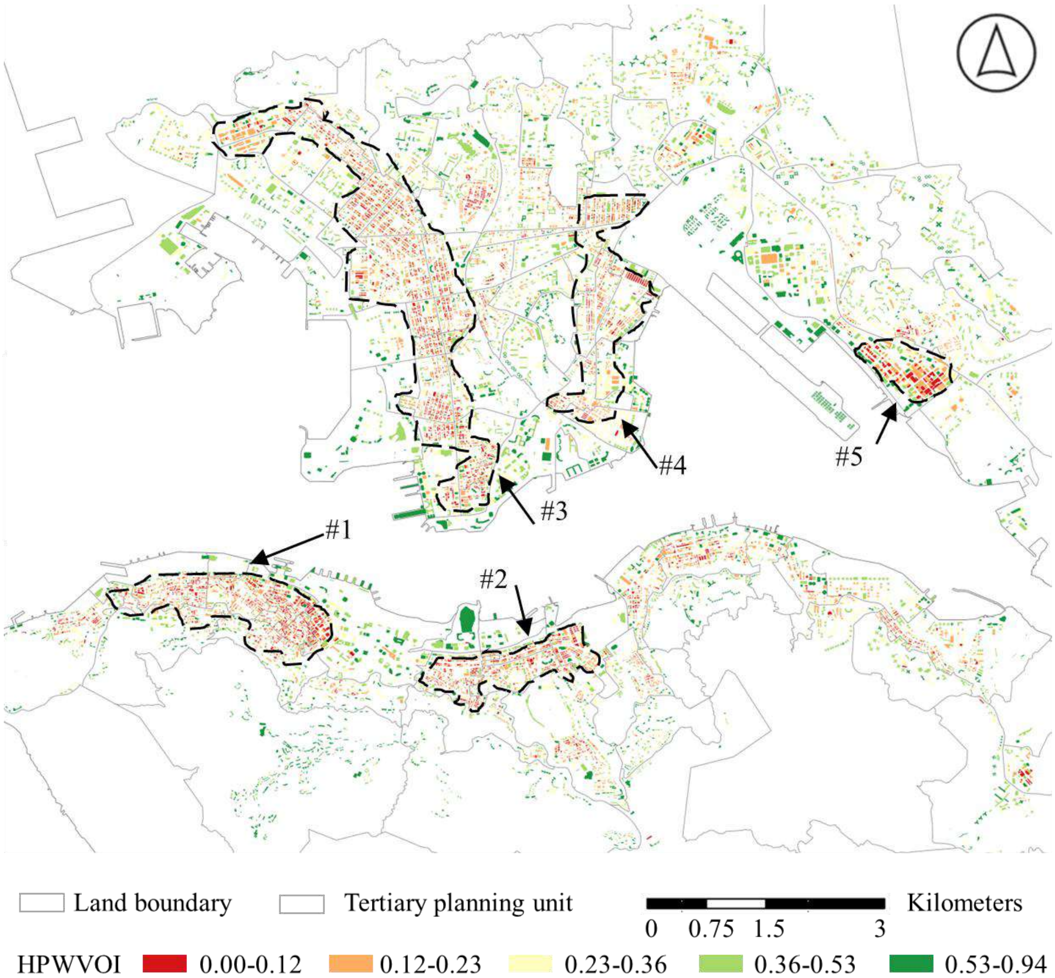

Figure 12. 2D mapping of building-level HPWVOIs: Five clusters with low HPWVOIs.

5. Discussion

5.1 Significance

High-density cities offer efficient access to services and amenities but can also engender feelings of crowdedness, enclosure, and stress. However, traditional objective measures of view openness, e.g., sky view proportions and landscape distance, failed to represent human perceptions of window view openness. Assessments on subjective view openness have been confined to small areas due to the high-cost collection of citywide realistic window views. Towards human-centered building and environmental applications, e.g., building mass design, housing, and urban planning, citywide assessment of human-perceived window view openness has been challenging but significant to examine the disparity of 3D visual density and inform the prioritization of local built environment improvement, benefiting billions of urban dwellers.

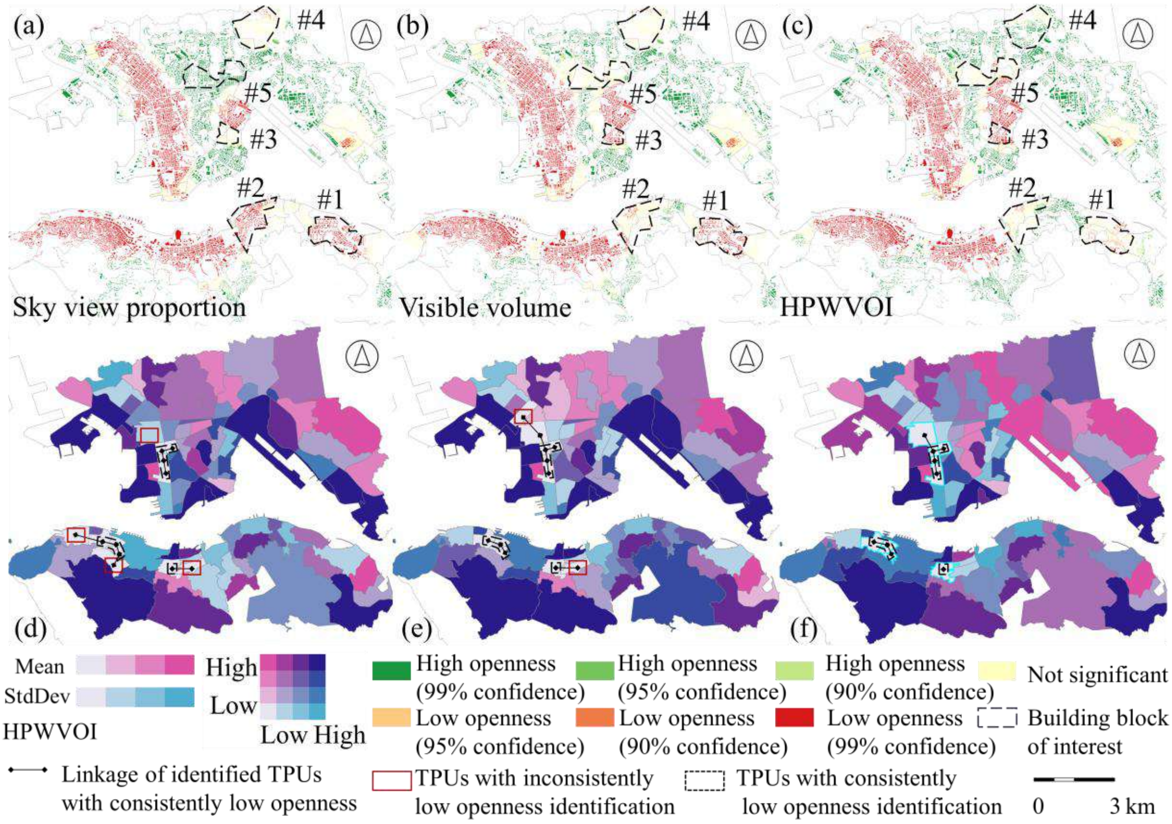

Figure 13. Comparison of HPWVOIs with two typical objective openness metrics at the building and TPU levels.

The proposed 3D GIS and ML-based photorealistic 3D City Information Modeling approach enables citywide assessment of human-perceived window view openness. The defined HPWVOI integrates objective view descriptors for a multi-dimensional representation for perceived openness. Citywide generation of photorealistic and type-depth window views on 3D CIMs supports identification of representative window views and quantification of three objective view descriptors. Generation of photorealistic window views in computers, for both openness ratings and more fidelity of view descriptor quantification, addresses both the scalability limitations of traditional psychological studies and the oversimplifications inherent in 3D GIS–based cityscape modeling for estimating citywide human-perceived view openness. Given the inherent subjectivity of human perception, we adopt linear ML algorithms for better interpretability and public acceptance in planning decision-making. The automatic linear ML regression calibrates these three descriptors to estimate HPWVOIs with interpretable equations and linear weightings. Compared with the possible non-linear machine learning regression methods shown in Supplementary Table 1, we argue that the proposed linear regression model effectively captured the observed linear relationships (see Figures 9a–c) between the three transformed descriptors and HPWVOI, achieving even better performance, with a maximum improvement of $R^2 = 0.1137$. Experimental results in eight districts of Hong Kong showed that the proposed approach was automatic to generate citywide window views (≤ 3.13 s per image), their representatives, three types of objective descriptors (≤ 0.02 s per image), and the estimated HPWVOIs (≤ 0.001 s per image). The proposed ML model was robust and accurate ($R^2 = 0.9191/0.8040$ in cross/independent validation) in representing perceived openness using the three objective view descriptors. Compared to two commonly used objective openness metrics in street view and landscape analytics, the proposed HPWVOI can effectively identify buildings and blocks with low human-perceived openness, providing additional quantitative evidence for human-centered built environment improvement.

Beyond the applications illustrated in Figures 11 and 12 and Section 1, e.g., window view-guided housing valuation and transaction and built environment improvement, floor-level HPWVOIs from multi-story windows also provide new datasets for urban health analytics, with potential to shed light on urban health issues, e.g., depression, stress, and mortality, particularly for individuals who spend substantial time indoors, e.g., the disabled and elderly.

5.2 Limitations and future work

There exist observed limitations in the study. Temporal inconsistency among the multi-source spatial datasets, e.g., 3D CIMs and building footprints, may introduce outdated or defective view representations, particularly in rapidly developing areas, through geometric mismatches. Although this effect is expected to be limited in our study area (e.g., 98.63% of buildings built before 2019 (HKBD 2026)), future work should incorporate regularly updated 3D CIMs and building footprints within digital twin city frameworks to improve the robustness of large-scale window-view assessment. Despite the high resolution of selected photorealistic window views, there still exists room to increase their quality towards the real-world ones to improve possible biases of human perception. Future work includes the beautification of photorealistic window views via Generative Artificial Intelligence for 3D CIM-generated window view-based surveys and analytics in disciplines, e.g., psychology, architectural design, landscape architecture, and urban health. Another common concern in survey-based research includes the representativeness of respondents and the collected openness ratings, as well as potential biases in respondents' understanding and the limited number of photorealistic window views. In this study, we aimed to provide a preliminary demonstration of the workflow for assessing the perceived openness of window views citywide. Thus, we acknowledge that the population may limit the generalization of the assessment results. Future work includes a more extensive survey, involving more respondents and photorealistic window view images as stimuli. Last, we only calibrated typical objective view openness descriptors to estimate HPWVOIs. Future work will focus on understanding human-perceived window view openness by exploring and calibrating high-dimensional visual features. These include view structure, captured through both local descriptors (e.g., gaps between vertical obstructions such as buildings and trees) and holistic descriptors (e.g., graph-based measures of continuity), as well as deep image features (Guo et al. 2025). Graph learning and foundation computer vision models will be employed to model these structural and visual representations, respectively.

6. Conclusion

Window view openness is valued by urban dwellers, especially in high-density cities. A distant landscape view without any visual obstruction offers urban dwellers' feelings of openness for stress relief in the crowded urban context. Assessing human-perceived window view openness citywide provides new quantitative evidence to inform housing marketing, architectural design, and urban planning for local visual density improvement. However, current assessments of objective view openness often miss human subjective perception, while subjective openness evaluations remain small-scale because of the difficulty in collecting citywide window views and their representatives. To address this research gap, we propose a Perceived Window View Openness Index (HPWVOI) and a citywide assessment approach using photorealistic 3D CIMs and automatic linear Machine Learning (ML).

The HPWVOI is first defined on three objective window view openness descriptors, i.e., the sky view proportion, building view proportion, and landscape distance, using human perception as calibration. Then, we introduce a citywide assessment approach using photorealistic 3D CIMs and automatic linear ML models. The approach comprises three steps: i) automatic generation of citywide photorealistic and type-depth window views, ii) collection of representatives of citywide window views with quantified objective and perceived view openness, and iii) automatic ML-based calibration of objective window view descriptors for HPWVOIs. Last, apartment-level and building-level HPWVOIs are visualized on 3D and 2D maps for urban applications in housing marketing and built environment improvement. Experimental results confirmed that the proposed approach was automatic to generate citywide window views ($\leq$ 3.13 s per image), their representative set, three types of objective descriptors ($\leq$ 0.02 s per image), and the estimated HPWVOIs ($\leq$ 0.001 s per image). The linear ML models built on openness ratings of the representative set and three objective descriptors were accurate (R^2 = 0.9191/0.8040 in cross/independent validation) and scalable to citywide estimation. Compared to two typical objective openness metrics in street view and landscape analytics, the proposed HPWVOI identifies eight city blocks with low perceived openness, supporting human-centered built environment improvement.

The introduced HPWVOI and 3D CIM-based citywide assessment approach offer new solutions for potential urban applications in housing marketing, architectural design, and built environment improvement. For example, the HPWVOI can serve as a new indicator for housing price estimation. The ML model trained in the proposed approach can inform architects to finetune the designs for maximized window view openness. The citywide HPWVOIs support urban planners to identify aging buildings with high local crowdedness for prioritized improvement. Future works include enhancement of up-to-date photorealistic window views via Generative Artificial Intelligence and understanding human-perceived window view openness by calibrating high-dimensional features, e.g., view structure and deep image features, using graph learning and foundation computer vision models, respectively.

Acknowledgement

The work presented in this study was supported in part by the Department of Science and Technology of Guangdong Province (GDST) (2020B1212030009) and the Guangdong Provincial Talent Program (2025GA12X007). The study was an extension of a PhD thesis supported by the University of Hong Kong (HKU). We would like to thank Ms. Siyuan Meng, Ms. Qianyun Zhou, and other members of HKU for their valuable assistance with the survey and data collection.

Ethical Approval

The ethics approval was valid at the time of data collection. All data were collected within the approved period and in accordance with the approved protocol (EA210471) from the Human Research Ethics Committee, the University of Hong Kong. The approval covered data retention and future academic analysis, and all participants provided informed consent accordingly.

References

Abd-Alhamid, F., Kent, M. & Wu, Y. (2023). Quantifying window view quality: A review on view perception assessment and representation methods. *Building and Environment, 227*, 109742. doi:10.1016/j.buildenv.2022.109742

Daniel, T. C. & Meitner, M. M. (2001). Representational validity of landscape visualizations: the effects of graphical realism on perceived scenic beauty of forest vistas. *Journal of Environmental Psychology, 21*(1), 61-72. doi:10.1006/jevp.2000.0182

Google LLC. (2023, 9 1). *Photorealistic 3D Tiles*. Retrieved from https://developers.google.com/maps/documentation/tile/3d-tiles

Guo, S., Jang, K. M., Duarte, F., Kang, Y. & Ratti, C. (2025). Urban visual uniqueness: A landmark-free framework to quantify city's identity and distinctiveness from everyday scenes. *Computers, Environment and Urban Systems, 122*, 102351. doi:10.1016/j.compenvurbsys.2025.102351

Hastie, T., Tibshirani, R. & Friedman, J. (2009). *The Elements of Statistical Learning.* New York: Springer New York. doi:10.1007/978-0-387-84858-7

HKBD. (2026). *Building Information and Age Records.* Hong Kong: Buildings Department, Government of Hong Kong SAR. Retrieved from https://portal.csdi.gov.hk/geoportal/?lang=en&datasetId=bd_rcd_1631167534872_19764

HKLandsD. (2019). *3D Visualisation Map (Tile-based models).* Hong Kong: Lands Department, Government of Hong Kong SAR.

HKLandsD. (2023). *iB1000 Digital Topographic Map.* Hong Kong: Lands Department, Government of Hong Kong SAR.

HKLandsD. (2025). *Building.* Hong Kong: Lands Department, Government of Hong Kong SAR. Retrieved from https://static.csdi.gov.hk/csdi-webpage/view/51d63757e2675874af80eef94afb6a35/sds_view

HKPlanD. (2018). *Hong Kong Planning Standards and Guidelines.* Hong Kong: Planning department, Hong Kong SAR. doi:https://www.pland.gov.hk/pland_en/tech_doc/hkpsg/full/pdf/ch2.pdf

HKPlanD. (2019). *3D Photo-realistic Model.* Hong Kong: Planning Department, Government of Hong Kong SAR. Retrieved from https://www.pland.gov.hk/pland_en/info_serv/3D_models/3D_Photo_realistic_Model_Specification.pdf

Jordan, M. I. & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255-260. doi:10.1126/science.aaa8415

Kaplan, R., Kaplan, S. & Brown, T. (1989). Environmental preference: A comparison of four domains of predictors. *Environment and Behavior, 21*(5), 509-530. doi:10.1177/0013916589215001

Kim, J., Kent, M., Kral, K. & Dogan, T. (2022). Seemo: A new tool for early design window view satisfaction evaluation in residential buildings. *Building and Environment, 214*, 108909. doi:10.1016/j.buildenv.2022.108909

Krukar, J., Manivannan, C., Bhatt, M. & Schultz, C. (2021). Embodied 3D isovists: A method to model the visual perception of space. *Environment and Planning B: Urban Analytics and City Science, 48*(8), 2307-2325. doi:10.1177/2399808320974533

Lange, E. (2001). The limits of realism: perceptions of virtual landscapes. *Landscape and Urban Planning, 54*(1-4), 163-182. doi:10.1016/S0169-2046(01)00134-7

Laovisutthichai, V., Li, M., Xue, F., Lu, W., Tam, K. & Yeh, A. G. (2021). CIM-enabled quantitative view assessment in architectural design and space planning. *38th International Symposium on Automation and Robotics in Construction (ISARC 2021)* (pp. 65-72). Oulu: IAARC.

Law, S., Horinouchi, M., Suel, E., Stalder, S., Volpi, M., Shen, Y., Li, M. & Takizawa, A. (2026). Estimating and mapping Tokyo Window View Desirability using 3D model imagery and the hedonic price approach. *Applied Geography, 194*, 104110. doi:10.1016/j.apgeog.2026.104110

Li, J., Xie, X., Li, X., Liu, C., Fukuda, H. & Gao, W. (2026a). Physical vs. Virtual Environments: A Comparative Psychophysiological Study on the Restorative Effects of Window Views. *Building and Environment, 299*, 114675. doi:10.1016/j.buildenv.2026.114675

Li, M., Guo, S., Duarte, F., Kumar, A., Kobori, N., Xue, F., Zhuang, W., Yeh, A. & Ratti, C. (2026b). Influence of objective and perceived exposures to urban nature on people's happiness. *npj Urban Sustainability, 6*, 6. doi:10.1038/s42949-025-00306-9

Li, M., Xue, F. & Yeh, A. G. (2023). 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. & 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., Xue, F., Wu, Y. & Yeh, A. G. (2022). A room with a view: Automated assessment of window views for high-rise high-density areas using City Information Models and transfer deep learning. *Landscape and Urban Planning, 226*, 104505. doi:10.1016/j.landurbplan.2022.104505

Li, M., Xue, F., Yeh, A. G. & Lu, W. (2021). Classification of photo-realistic 3D window views in a high-density city: The case of Hong Kong. In X. Lu, Z. Zhang, W. Lu & Y. Peng, *Proceedings of the 25th International Symposium on Advancement of Construction Management and Real Estate* (pp. 1339-1350). Singapore: Springer.

Li, W. & Samuelson, H. (2020). A new method for visualizing and evaluating views in architectural design. *Developments in the Built Environment, 1*, 100005. doi:10.1016/j.dibe.2020.100005

Liang, D. & Xue, F. (2023). Integrating automated machine learning and interpretability analysis in architecture, engineering and construction industry: A case of identifying failure modes of reinforced concrete shear walls. *Computers in Industry, 147*, 103883. doi:10.1016/j.compind.2023.103883

Lin, T.-Y., Le, A.-V. & Chan, Y.-C. (2022). Evaluation of window view preference using quantitative and qualitative factors of window view content. *Building and Environment, 213*, 108886. doi:10.1016/j.buildenv.2022.108886

Ling, Y. (2022). Estimating coastal premiums for apartment prices: Towards a new multilevel modelling approach. *Environment and Planning B: Urban Analytics and City Science, 49*(1), 188-205. doi:10.1177/23998083211000343

Liu, D., Wang, R., Grekousis, G., Liu, Y. & Lu, Y. (2023). Detecting older pedestrians and aging-friendly walkability using computer vision technology and street view imagery. *Computers, Environment and Urban Systems, 105*, 102027. doi:10.1016/j.compenvurbsys.2023.102027

Masoudinejad, S. & Hartig, T. (2020). Window view to the sky as a restorative resource for residents in densely populated cities. *Environment and Behavior, 52*(4), 401-436. doi:10.1177/0013916518807274

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*, 1-24. doi:10.1080/13658816.2026.2667807

Ode, Å., Tveit, M. S. & Fry, G. (2008). Capturing landscape visual character using indicators: touching

base with landscape aesthetic theory. *Landscape Research, 33*(1), 89-117. doi:10.1080/01426390701773854
Park, K. & Ergan, S. (2026). Window view optimization in building façade layout design (ViewOpt): A Human-centric 3D simulation and computer vision framework. *Building and Environment, 296*, 114538. doi:10.1016/j.buildenv.2026.114538
Pyka, K., Piskorski, R. & 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
Salesses, P., Schechtner, K. & Hidalgo, C. A. (2013). The collaborative image of the city: mapping the inequality of urban perception. *PloS One, 8*(7), e68400. doi:10.1371/journal.pone.0119352
Stamps III, A. E. (2005). Enclosure and safety in urbanscapes. *Environment and Behavior, 37*(1), 102-133. doi:10.1177/0013916504266806
Stevens, S. S. (1957). On the psychophysical law. *Psychological review, 64*(3), 153. doi:10.1037/h0046162
UN. (2018). *World Urbanization Prospects 2018.* New York: United Nations.
Wang, F. & Munakata, J. (2023). Simulated window views from different building floors: How roof types of surrounding buildings associate with subjective restoration. *Urban Forestry & Urban Greening, 89*, 128096. doi:10.1016/j.ufug.2023.128096
Wang, F. & Munakata, J. (2024). Assessing the impact of facade characteristics and visual elements on perceived oppressiveness in high-rise window views through virtual reality. *Building and Environment, 266*, 112043. doi:10.1016/j.buildenv.2024.112043
Weitkamp, G., Bregt, A. & Van Lammeren, R. (2011). Measuring visible space to assess landscape openness. *Landscape Research, 36*(2), 127-150. doi:10.1080/01426397.2010.549219
Weitkamp, G., Lammeren, R. v. & Bregt, A. (2014). Validation of isovist variables as predictors of perceived landscape openness. *Landscape and Urban Planning, 125* , 140-145. doi:10.1016/j.landurbplan.2014.02.021
Xia, Y., Yabuki, N. & Fukuda, T. (2021). Sky view factor estimation from street view images based on semantic segmentation. *Urban Climate, 40*, 100999. doi:10.1016/j.uclim.2021.100999
Zhu, H., Yang, X., Zhou, X., Zhang, H., Yang, L. & Tang, G. (2024). A 3D visible space index for evaluating urban openness based on the digital urban model. *International Journal of Geographical Information Science*, 1-26. doi:10.1080/13658816.2024.2397434