Habitat International 156 (2025) 103288

Contents lists available at ScienceDirect

Habitat International

journal homepage: www.elsevier.com/locate/habitatint

Toward 3D hedonic price model for vertically developed cities using street view images and machine learning methods

Yue Ying, Shaoqing Dai *, Mila Koeva, Monika Kuffer, Claudio Persello, Wen Zhou, Jaap Zevenbergen

Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Hallenweg 8, 7522 NH, Enschede, the Netherlands

ARTICLE INFO

Keywords:
Hedonic price model
Property valuation
3D built environment
Street view image
Machine learning
Random forest

ABSTRACT

The vertical developments in cities reshape the urban form and structure, and the influences on human liveability can be reflected by the variations in property values. The hedonic price model (HPM) is commonly employed in city-scale property valuation to unravel the hedonic values of different influential variables. In vertically developed cities, it necessitates the exploration of the hedonic value in the vertical dimension (3D), which was previously under-researched due to limited 3D data and the complexity of processing techniques. Recent studies use eye-level street view images (SVIs) for valuation, but the 3D perspective is still missing. This study proposes a novel 3D property valuation method using SVIs acquired from two angles, eye-level (pitch 0°) and sky-view (pitch 90°, upwards), and machine learning method to complete the 3D perspective and provide explainability of 3D HPM. We also compared different valuation models – namely Ordinary Least Square (OLS), Geographically Weighted Regression (GWR), and Random Forest (RF) – using model performance metrics. Our main findings include: 1) 3D variables are statistically significant, and adding them improves the model performance (R^2 from 0.580 to 0.636 in GWR); 2) In the sky-view angle, the proportion of sky has a positive correlation while the presence of buildings and trees are negatively correlated with property values; 3) RF outperforms OLS and GWR with the highest R^2 (0.768) and the least RMSE (1669.60 yuan/m^2), which demonstrates its robust explainability and applicability for valuation. This study enriches the property valuation literature on the significance of the 3D variables and provides references to guide fair taxation and equal land use policy in vertically developed cities.

1. Introduction

As the outcome of urbanisation, cities are increasingly dense and compact in most of the world. In addition to horizontal expansions (x- and y-axis: horizontal dimension, 2D), significant vertical developments are happening in urban areas due to limited land availability and growing population (Wen et al., 2019). It includes the process of rebuilding – replacing the shorter and smaller buildings with taller and bigger ones and infilling – constructing new buildings in the vacant areas (Frolking et al., 2024). The buildings of different heights, structures and densities have created unique three-dimensional cityscapes, especially in the vertical direction (z-axis: vertical dimension, making it 3D); they are considerably influential in urban climate, carbon emission, human health, and daily active travel behaviour (Alavipanah et al., 2018; Li et al., 2016; Wang et al., 2020). Such aspects related to local living quality vary not only on a horizontal basis but also at the vertical dimension, which influences the property values. Despite the significant impacts 3D have on the environment, up till now, property valuation has failed to adequately quantify and reflect these impacts as mass valuation methods and data are still 2D-based in practice (Ying et al., 2021, 2023). In vertical cities with generous amounts of buildings, the famous mantra in the real estate sector, "location, location, location", needs conceptual and technical re-interpretation from a 3D perspective.

Property valuation is a systematic process of determining the values of locations correlated to the property, which has a considerable influence on different stakeholders, including homeowners, taxation authorities, and policymakers. The value recognised by consumers in the property market reflects not only the value of the property itself but also the desirability and accessibility of the surrounding environment, which significantly contributes to their living comfort. The hedonic price model (HPM) has been extensively used in city-scale mass property valuation (Lancaster, 1966; Rosen, 1974). It decomposes the value of a

* Corresponding author. Faculty of Geo-information Science and Earth Observation (ITC), University of Twente, Enschede, the Netherlands.
E-mail address: shaoqing.dai@outlook.com (S. Dai).

https://doi.org/10.1016/j.habitatint.2025.103288
Received 12 August 2024; Received in revised form 26 November 2024; Accepted 4 January 2025

property to understand the economic impacts of variables under different attributes (e.g., physical, locational, and environmental). The fundamental valuation model is Ordinary Least Square (OLS), which assumes a linear relationship between the property value and the influential variables. Advanced models which can capture the spatial heterogeneity and temporal variations of the property market have been widely applied, such as Geographically Weighted Regression (GWR) and Geographically and Temporally Weighted Regression (Huang, Wu, & Barry, 2010). The applications of data-driven models (e.g., machine learning) in valuation have also grown, among which Random Forest (RF) has proven its robustness and explainability (Antipov et al., 2012).

Numerous studies have utilised HPM to examine the relationships between certain variables and property values with a 2D perspective (Wen et al., 2014; Ying et al., 2021), such as the proximity to and the density of certain public goods and services. From a theoretical perspective, HPM creates a framework to recognise the economic values of different attributes and variables in properties. In the context of vertically developed cities, the spatial changes happening in 3D make multi-facet impacts in view, visibility and sunlight exposure, which are considerably linked to living comfort and property values (Dai, Felsenstein, & Grinberger, 2023; Yamagata et al., 2016). 3D geoinformation can specify nuanced 3D spatial differences that 2D methods and data are not able to capture by nature and build a more accurate valuation system. Therefore, valuation needs to adapt to the vertical spatial changes in dense cities and reflect their impacts in the form of hedonic value. Recent studies have already called for attention in 3D from the perspectives of different stakeholders (Wu et al., 2024). However, academic interest and scientific understanding in valuing 3D are still scarce and segmented (Fleming et al., 2018; Ying et al., 2021; Yu et al., 2007), mainly due to the difficulty of 3D data acquisition and the complexity of processing techniques (Ying et al., 2023). Points of interest (POI) data, commonly employed in valuation studies, normally do not include 3D geoinformation. Satellite and aerial images-derived data are primarily used for extracting 2D variables (e.g., greenery coverage) in HPM rather than 3D. Light Detection and Ranging (LiDAR) enables the measurement of rich 3D geoinformation, but data collection poses challenges in large-scale acquisition due to the high costs. One study conducted in Austria used LiDAR to increase the prediction accuracy of HPM, and with only 48 flats selected still came with high-level manual efforts (Helbich et al., 2013).

Nevertheless, the emergence of street view images (SVIs) opens up opportunities to see cities with low costs and wide coverage from a 3D perspective (Tang et al., 2019). Artificial intelligence (AI) further advances the documentation of semantic streetscape and complex 3D analysis to address different urban challenges, such as bikeability (Dai et al., 2023b) and street crime (Yue et al., 2023). SVIs capture street-level features (e.g., building facades, greenery, and visible sky) from an eye-level angle, which are important variables in estimating property values (Dai, Felsenstein, & Grinberger, 2023; Kara et al., 2023; Zhang et al., 2018). Compared to SVIs, satellite and aerial images are typically acquired from a top-down perspective rather than a human viewpoint for environmental monitoring; the lack of semantic details along the z-axis makes POI ineligible to provide rich 3D geoinformation rather than an abstract point. Besides, the observation point of SVIs can be freely adjusted in different pitches (from 0°, eye level, to 90°, upwards). The eye-level (pitch 0°) angle offers rich street features close to the ground, and the sky-view angle (pitch 90°, upward) offers the opportunity to see upward as an observer. Thanks to AI advances, the semantic segmentation of the sheer amount of SVIs has become efficient. Deep learning is widely adopted to learn and extract features from raw image data without explicit feature selection (Persello et al., 2017). Various pre-trained models (e.g., SegNet (Badrinarayanan et al., 2017)) are available; high-resolution street scene open datasets (e.g., CityScapes (Cordts et al., 2016)) offer rich image samples across diverse semantic categories. There are already studies using SVIs to estimate the impact of eye-level street features on property values (Fu et al., 2019; Suzuki et al., 2022; Wang, 2022), but the sky-view angle (pitch 90°, upward) is overlooked without exception. This missing angle is an important measurement of street openness and the higher-up scenery along the z-axis, creating possibilities for estimating the hedonic value of 3D.

Drawing from the facts above, we expect to answer the following research questions: 1) How do 3D variables affect property values? 2) Are there any differences in the impacts of 3D variables from eye-level angle (pitch 0°) and sky-view angle (pitch 90°)? 3) Does machine learning provide better explainability than OLS or GWR? Under the umbrella of HPM, this study proposes a novel 3D property valuation method using SVIs in two angles to extend 2D HPM to 3D HPM and machine learning to provide explainability of 3D HPM. Based on the classic 2D HPM, we constructed semantic segmentation on SVIs and derived 3D variables correspondingly. After that, we built HPM with different modelling methods (OLS, GWR and RF), and we identified the influences of 3D variables and offered explainability to 3D HPM.

To the best of the authors' knowledge, the concept of bringing SVIs from two angles to form a full 3D perspective has not yet been applied to mass property valuation at a city scale. This study enriches the property valuation literature on the significance of 3D in HPM and provides references to guide fair property taxation and equal land use policy in dense cities regarding access to 3D aspects. It bridges the gap between the theoretical HPM framework and practical land use policy-making needs in vertical cities. The research results will be of interest to property valuers, researchers, and local government involved in urban sustainable development and housing policy-making.

2. Materials and methods

2.1. Study area

The fast urbanisation in China is characterised by the construction of generous amounts of high-rise buildings and increased building height (Li et al., 2020). The most significant changes happen in first- and second-tier cities (e.g., the provincial capitals) (Yang et al., 2022). Xi'an ranks 62nd among cities with the most high-rise buildings worldwide and 29th nationwide (https://www.skyscrapercenter.com/cities). It is the capital of Shaanxi Province and the political, economic and educational centre of northwest China (Fig. 1). It is located in the centre of the Guanzhong Plain, with a population of 12.99 million by 2022 and coverage of 10,180 km^2 (Shaanxi Provincial Bureau of Statistics & NBS Survey Office in Shaanxi, 2023). The main urban region (659.06 km^2) is the study area. We highlight that Xi'an is under housing market control to curb speculations and lower excessive price hikes (Central Committee of the Communist Party of China, 2017). This policy control covers most first- and second-tier cities in China and influences the first-hand and the second-hand market at the same time.

2.2. Data source

2.2.1. Property value database

The geocoded property value database documents the actual transaction records with second-hand residential property at the household level and physical variables, such as built year, floor level, and building type. It was provided by Lianjia (https://bj.lianjia.com), one of China's largest real estate brokers. To ensure temporal consistency and minimise the impact of the COVID-19 pandemic, we selected the period dated from January 1st, 2020 to May 31st, 2020, inclusive. There are, in total, 8055 records after data cleaning. We decided to use transaction prices rather than list prices to be truly reflective of the market fluctuations. We focused on only apartments in residential buildings in the second-hand market as the local government put a price cap on the first-hand residential property market. The details are referred to in Table 1 under the category of physical variables.

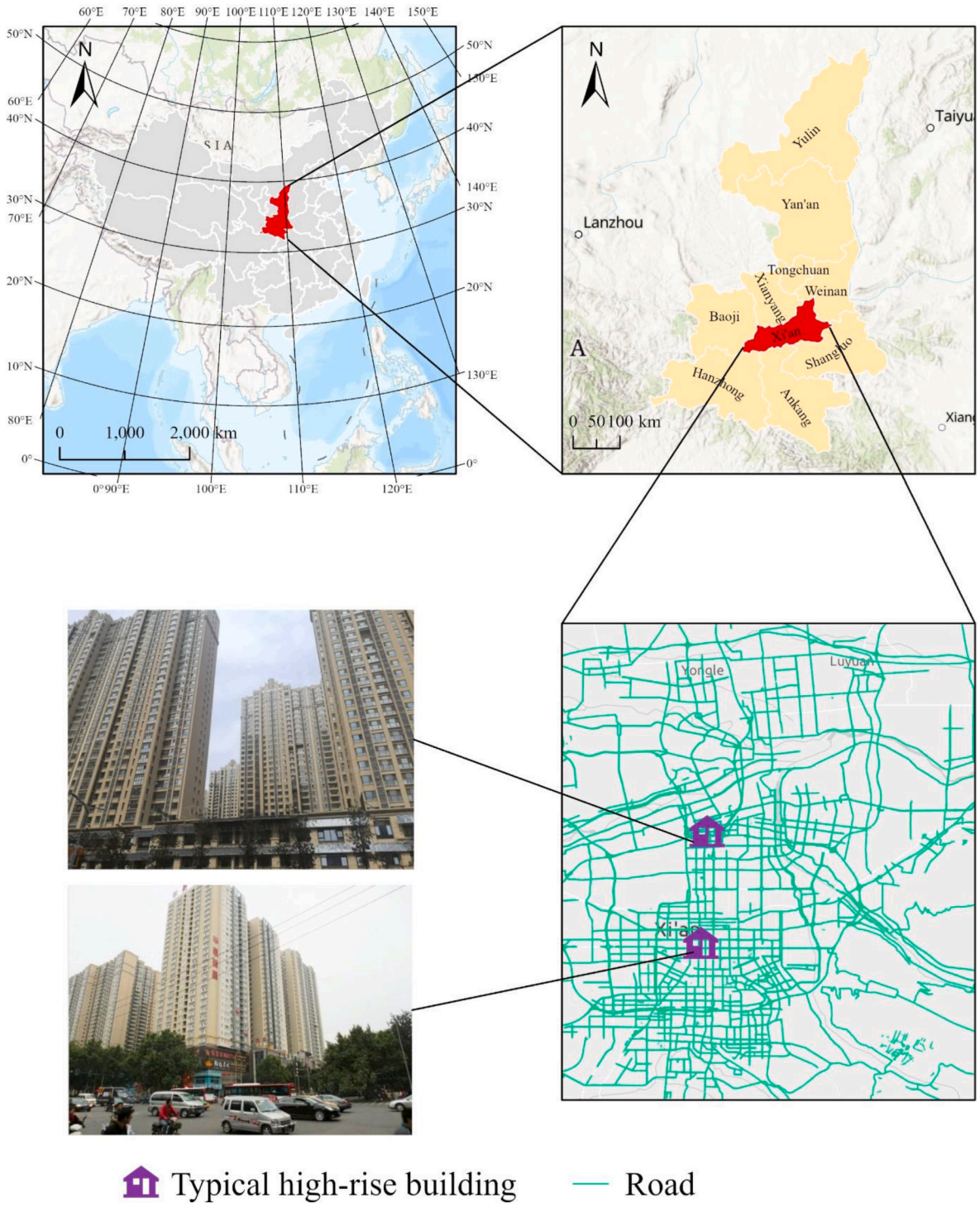


Fig. 1. The study area. The red polygons represent the adminstrative boundaries of Shaanxi Province (upper-left) and Xi'an city (upper-right); the purple houses (bottom-right) rerpresent two typcial high-rise buildings.

2.2.2. Point of interest (POI)

POI was acquired from AutoNavi Maps (https://lbs.amap.com/), which is one of the leading map service providers in China. The collection time is December 2020. There are 23 categories, including restaurants, subway stations, and hospitals. It was used to generate locational variables, such as the distance to the nearest shopping mall and the density of the convenience stores. The details are referred to in Table 1 under the category of locational variables.

2.2.3. Street view images (SVIs)

SVIs were collected from Baidu Maps, another China's leading map service provider (https://map.baidu.com). They were used to form 3D variables (Table 1) via semantic segmentation. The workflow of mass-downloading SVIs started with generating the sampling points with an interval of 10 m along the OpenStreetMap (https://www.openstreetmap.org) road networks (Fig. 2a). Second, we determined the heading direction and pitch, two parameters in image customisation (Fig. 2b and c). Heading direction defines the compass heading of the camera with the value from 0° to 360°. Four heading directions (0°, 90°, 180°,

Table 1
The descriptive statistics and explanations of variables.

Category	Variable name	Variable definition	Mean	Min	Max	Std
Dependent variable	Average price	The average property value (unit: yuan/m^2)	14,318	1196	47,384	3936.30
Physical variables	Elevator *(ELEV)*	Dummy variable that evaluates whether the property has an elevator service. 1 = it has an elevator, 0 = no elevator.	0.90	0	1	0.30
	Decoration *(DECO)*	Dummy variable that evaluates the decoration level of the property. 1 = it has fine decoration, 0 = it has no/not enough decoration.	0.54	0	1	0.50
	Relative floor *(REL_floor)*	Categorical variable that evaluates the relative location of the property in the building. 1 = low floor level, 2 = middle floor level, 3 = high floor level	2.08	1	3	0.79
	Orientation *(ORI)*	Dummy variable that evaluates whether the property has an orientation to south. 1 = it is oriented to south; 0 = it is oriented to north/east/west.	0.66	0	1	0.47
	Total floor *(TOTAL_floor)*	Numerical variable that measures the total floor number of the building where the property locates.	23.8	1	39	9.70
	Area number*(AREA_num)*	Numerical variable that measures the property size in the (unit: m^2)	91.47	13.02	457.12	34.70
	Room number *(ROOM_num)*	Numerical variable that measures the total room number of the property	2.29	1	8	0.81
	Built year *(BUILT_y)*	Numerical variable that documents when the property is built.	2011	1980	2020	4.91
Locational variables	Distance to CBD *(DIST_cbd)*	Euclidean distance from the property to the traditional Central Business District (CBD) -Bell Tower (*Zhong Lou*) in kilometre	7.78	0.33	17.54	2.67
	Distance to shopping mall *(DIST_shop)*	Euclidean distance from the property to the closest shopping mall in kilometre	0.95	0.01	3.99	0.60
	Distance to hospital *(DIST_hos)*	Euclidean distance from the property to the closet general hospital (*sanjia yiyuan*) in kilometre	2.54	0.01	9.27	1.64
	Distance to park *(DIST_par)*	Euclidean distance from the property to the closet park in kilometre	1.23	0.07	6.27	0.69
	Density of convenience shop *(DEN_con)*	The count of convenience shops within 1 km vicinity centered from the property	88.04	1.00	361.00	58.64
	Density o educational facility *(DEN_edu)*	The count of educational facilities (primary school, middle school, universities and colleges) within 1 km vicinity centered from the property	19.97	0.00	157.00	17.48
	Density of sports and recreation facility *(DEN_sar)*	The count of sports and recreation facilities within 1 km vicinity centered from the property	54.85	0.00	265.00	43.26
	Density of restaurants *(DEN_res)*	The count of restaurants within 1 km vicinity centered from the property	447.6	0.00	1977.00	320.95
	Density of bus stops *(DEN_bus)*	The count of bus stops within 1 km vicinity centered from the property	17.91	1.00	68.00	9.31
	Density of subway *(DEN_sub)*	The count of subway stations within 1 km vicinity centered from the property	0.87	0.00	4.00	0.84
3D variables	$3D_{ss}$	The proportion of sky in sky-view images (pitch 90°)	0.51	0.07	0.84	0.11
	$3D_{st}$	The proportion of tree in sky-view images (pitch 90°)	0.04	0.00	0.18	0.02
	$3D_{sb}$	The proportion of building in sky-view images (pitch 90°)	0.10	0.00	0.20	0.03
	$3D_{es}$	The proportion of sky in eye-level images (pitch 0°)	0.15	0.02	0.42	0.04
	$3D_{et}$	The proportion of tree (pitch 0°) in eye-level images	0.08	0.01	0.23	0.03
	$3D_{eg}$	The proportion of grass in eye-level images (pitch 0°)	0.01	0.00	0.05	0.01
	$3D_{er}$	The proportion of road in in eye-level images (pitch 0°)	0.23	0.12	0.33	0.04
	$3D_{ew}$	The proportion of sidewalk in eye-level images (pitch 0°)	0.03	0.00	0.07	0.01
	$3D_{eb}$	The proportion of building in eye-level images (pitch 0°)	0.25	0.04	0.41	0.06

270°) were selected for a full surrounding view. The pitch represents the up/down angle at which the street view is taken. Compared to other studies that adopt pitch 0° for eye-level only, our strategy incorporates two pitches (0° and 90°) to collect both eye-level images reflecting horizontal streetscape and sky-view images reflecting the vertical sky openness (Fig. 2d and e). Finally, we determined the acquisition time: 2019 or 2020. Our objective is to improve the quantity of SVIs by employing a compact sampling point density and a comparatively inclusive acquisition timeframe. Preliminary assessments indicated that not all sampling points have images, and annual updates are not guaranteed. Ideally, if images are available, eight images are collected per year per sampling point. Each image has 1024 * 512 pixels. In total, we generated 83,925 sampling points and downloaded 737,704 images.

2.3. Methods

As shown in Fig. 3, this study is conducted within the HPM framework, which consists of three steps. We first generate physical, locational variables, extract 3D variables from SVIs (2.3.1, step 1 and 2), and then construct HPM in different valuation models, namely OLS, GWR and RF (2.3.2, step 3).

2.3.1. Physical, locational and 3D variables generation

The physical variables were interpreted from texts into numerical values from the property value database. The locational variables were calculated using geo-locations of properties and POI (2.2.2). The variable selection was based on existing studies, and we aimed to have a classic list in Table 1 because our paper focuses on 3D variables (Wu et al., 2020; Zhang et al., 2020).

3D variables were generated via SVIs semantic segmentation. Semantic segmentation for imagerys is time-consuming and computation-intensive due to the pre-labelling demands of imagery collection and the pre-training and optimising the algorithm (Persello et al., 2014). We applied a ready-to-use segmentation package, which has been proven to have low segmentation errors and high efficiency (Yao et al., 2019). It employs a Fully Convolutional Network pre-trained by ADE20K dataset (https://groups.csail.mit.edu/vision/datasets/ADE20K) to semantically segment different categories in each image. The output is the percentage of the recognised category. ADE20K dataset has over 20,000 multiple scene-centric images and includes 150 semantic categories. We selected 3D variables which have already been empirically proven important in living comfort in urban studies and are considered meaningful to our study (Wen et al., 2014; Zheng et al., 2023): sky, tree and building for sky-view images to form $3D_{ss}$, $3D_{st}$, and $3D_{sb}$; sky, greenery (tree and grass), building, and paved areas (road and sidewalk) for eye-level images to form $3D_{es}$, $3D_{er}$, $3D_{eg}$, $3D_{eb}$, $3D_{er}$, and $3D_{ew}$. Fig. 4 shows an example of segmentation. After the semantic segmentation, the properties and SVIs were linked based on their geographical locations. First, a buffer centered on the property with a radius of 800 m was created to involve images falling within. Then, the values of the specific categories of all the images were averaged to form the corresponding 3D variables. In this case, we did not distinguish years (e.g., if one sampling point has images in 2019 and 2020, both years are included). We tested the radius from 400 m to 1 km with an interval of 100 m. 800 m was selected with

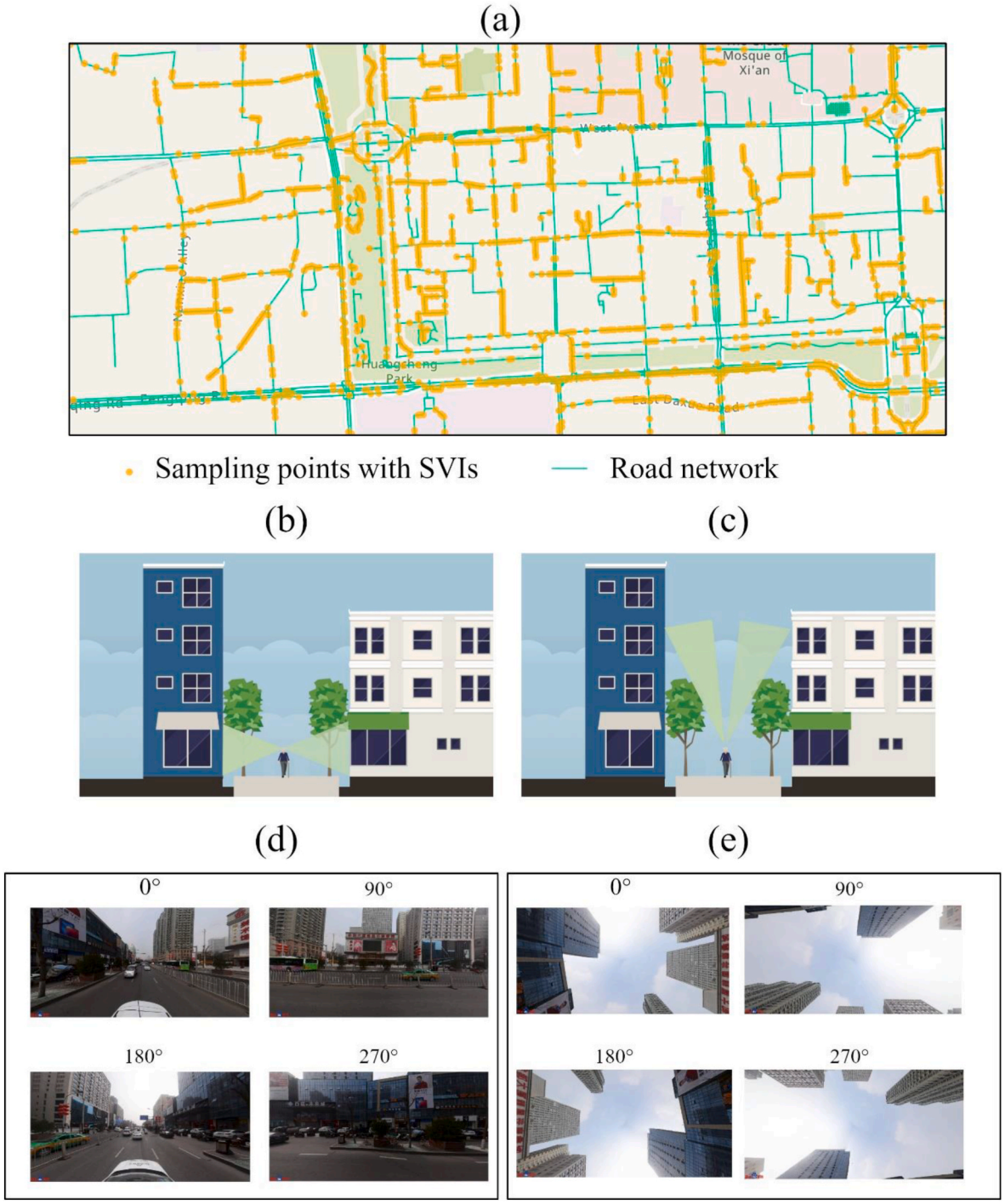


Fig. 2. Distribution of sampling points along the road network (a), the abstract simulations of eye-level angle (b) and sky-view angles (c), and illustrations of SVI in eye-level angle (d) and sky-view angle (e).

the best model fitting.

2.3.2. HPM in OLS, GWR and RF

HPM is widely adopted in property valuation to quantify the hedonic value of different variables on property values (Rosen, 1974). The property value is an aggregation of a bundle of intrinsic and extrinsic hedonic characteristics, among which physical, locational and environmental variables are commonly involved. The physical attributes refer to the internal structural characteristics of the property, such as plot size and the number of bedrooms. The locational attributes include characteristics such as the proximity to the nearest shopping mall and the density of hospitals. The environmental attributes refer to the variables such as the air quality level and the presence of greenery. The function of HPM can be written as follows:

$$V = f(P, L, E_{3d})$$

Where $f(...)$ indicates HPM, V stands for the property value, P for the physical variable, L for the locational variable, and E_{3d} for the 3D variable. In this study, the 3D variables replace the environmental variable in other studies, which can better describe the surrounding environment with a 3D perspective.

OLS, GWR, and RF are used to construct HPM. OLS is the baseline

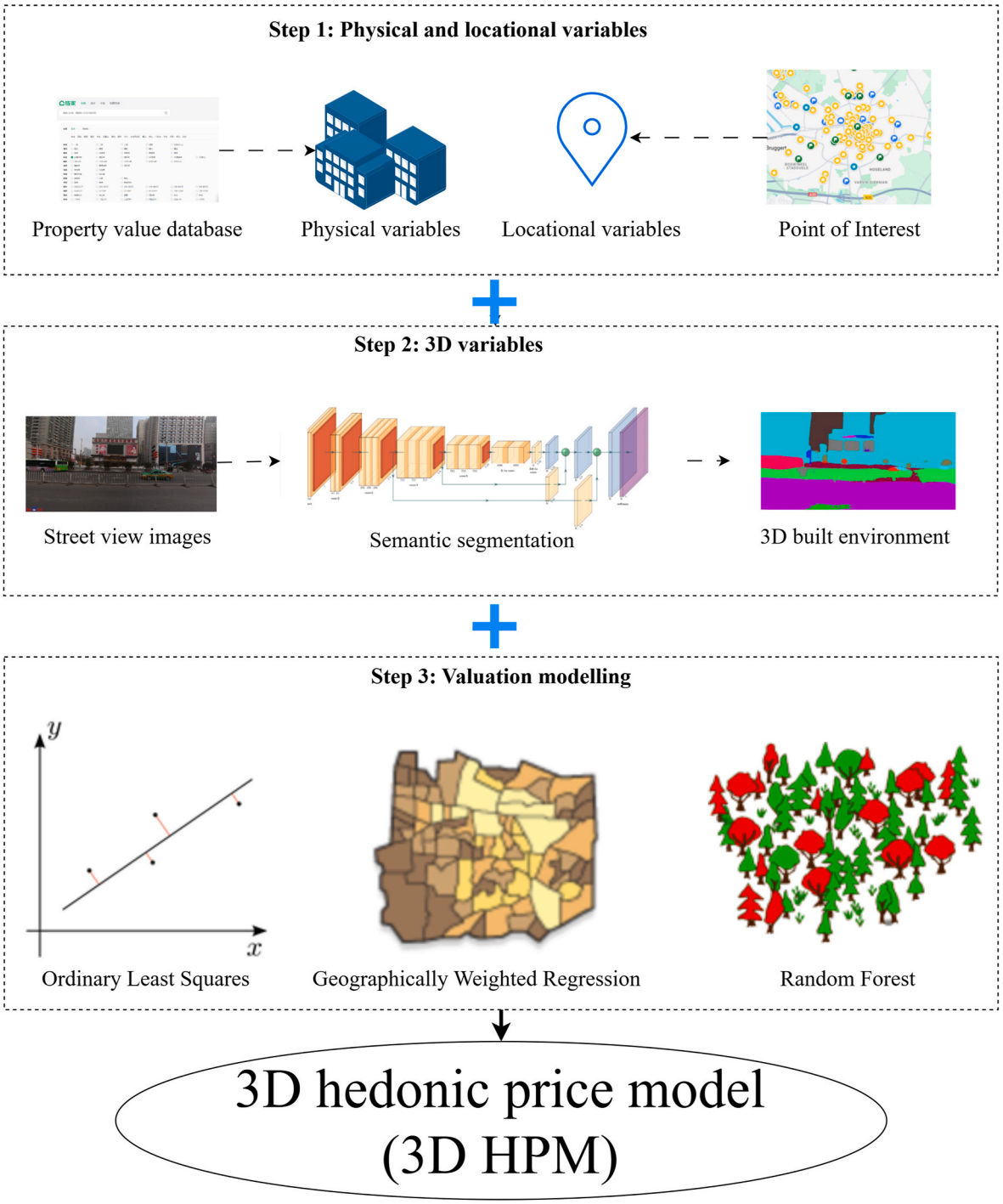
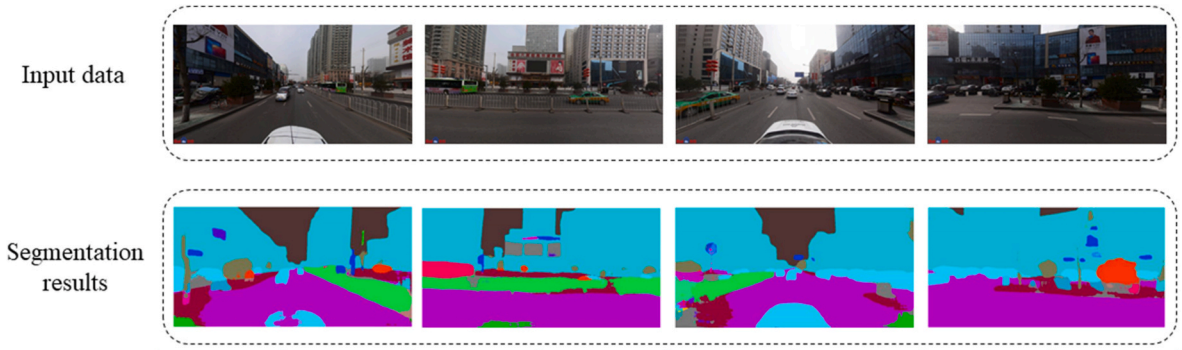


Fig. 3. The methodology design and the workflow of 3D hedonic price model (3D HPM).

model widely used in valuation studies (Qiu et al., 2022). It assumes a linear and global relationship between the dependent and independent variables but neglects spatial heterogeneity and spatial autocorrelation commonly found in geographic data (Wu et al., 2022). Moran's I and variance inflation factor (VIF) are used to assess the spatial autocorrelation in the dataset and the multi-collinearity between independent variables. Moran's I value ranges from −1 to 1. A positive value close to 1 indicates positive spatial autocorrelation, suggesting that similar values are clustered in space. GWR is an appropriate alternative to deal with the continuous spatial change and uneven distribution in property value data. It conducts a localised regression of variables to determine coefficients and identify spatial heterogeneity; therefore, GWR has been frequently adopted to explore the spatial non-stationarity between property values and the influencing variables (Qiu et al., 2022). The coefficient of determination (R^2) is commonly used to show the predictable portion of the dependent variable, i.e., property values, in this study. It measures how well property values are replicated by the model (Hong et al., 2020). The value is between 0 and 1. If $R^2 = 0.5$, it means 50% of the variance of the property values is explained by independent variables, i.e., physical, locational and 3D variables, in our case.

RF is a machine learning algorithm applied for classification and regression tasks that work by using predictions derived from a combination of decision trees and using the average to improve prediction accuracy and avoid over-fitting issues (Breiman, 2001). It has no



Wall	0.004	Wall	0.000	Wall	0.000	Wall	0.000
Building	0.352	Building	0.432	Building	0.400	Building	0.484
Sky	0.136	Sky	0.056	Sky	0.103	Sky	0.011
Floor	0.001	Floor	0.001	Floor	0.001	Floor	0
Tree	0.030	Tree	0.006	Tree	0.015	Tree	0.014
Road	0.253	Road	0.285	Road	0.226	Road	0.315
Sidewalk	0.062	Sidewalk	0.036	Sidewalk	0.048	Sidewalk	0.047
Window	0	Window	0	Window	0	Window	0
Grass	0	Grass	0	Grass	0.008	Grass	0

Fig. 4. The example of semantic segmentation.

requirement on data distribution or type and is robust against noises and outliers (Wu et al., 2022). It is applicable and effective in valuation studies because it can handle categorical variables, such as the building type, and gives interpretations of the influences from different variables on property values (Antipov et al., 2012; Wu et al., 2022). Comparatively, many machine learning algorithms have black box issues – i.e., only giving the outcomes without explaining why. RF can generate explainable diagrams such as relative feature importance and partial dependence plots (PDPs). The relative feature importance table provides a measure of the contribution of each variable in the model to the overall prediction accuracy. The PDPs represent how one variable influences the model estimation. It can be interpreted similarly to the coefficients in linear or regression models while it is able to capture more complex patterns in an advanced model. The prediction performance can be measured by the root mean square error (RMSE), mean absolute error (MAE) and mean absolute percent error (MAPE). RMSE represents the square root of the average of the squared differences between predicted and actual values. MAE is the average of the absolute differences between predicted and actual values. MAPE is the average of the absolute percentage differences between predicted and actual values. We randomly split the dataset 80% as training and 20% as a test to calculate the R^2, MAE, MAPE, and RMSE in RF.

3. Empirical results and analysis

3.1. The significance of 3D variables: evidence from model comparison

We built groups composed of different 3D variables to evaluate the contributions of eye-level and sky-view images separately. Group 1 does not include 3D variables. Groups 2 and 3 contain variables from eye-level ($3D_{es}$, $3D_{er}$, $3D_{eg}$, $3D_{eb}$, $3D_{er}$, and $3D_{ew}$) and sky-view images ($3D_{ss}$, $3D_{st}$, and $3D_{sb}$) separately. Group 4 include 3D variables from images in two angles. Table 2 gives an overview of R^2 in different models and groups. In terms of variable selection, group 4 has the best R^2 among all, so it serves as the basis for interpretations in the following sections. In OLS, we see a normal increase in R^2 after adding $3D_s$ (0.273) from sky-view images, a higher increase after adding $3D_e$ (0.370) from eye-level images, and a remarkable increase after adding both (0.376). GWR and RF share the same tendency. In terms of model comparison, GWR performs better than OLS, with an increase of R^2 from 0.376 to 0.636 in group 4. This means GWR can capture 63.6% of the influence of the variables, while OLS explains 37.6%. The improvement of R^2 in GWR confirms the spatial heterogeneity and the clustered feature of property values. RF performs the best among all three models. R^2 increases from 0.716 to 0.768 by adding 3D variables. In addition to R^2 values, RMSE, MAE, and MAPE are also used to evaluate the model accuracy, as shown in Table 3. The lowest RMSE in RF is 1669.60 yuan/m^2. Taking the average property value (14,318 yuan/m^2), the MAPE of estimation in RF is 11.66%. In summary, OLS has the basic performance, GWR has improvement, and RF has the best explainability and the least errors in

Table 2
The R^2 in different models and groups.

Group	OLS	GWR	RF
(1) BASIC	0.267	0.580	0.716
(2) BASIC + $3D_e$	0.370	0.622	0.763
(3) BASIC + $3D_s$	0.273	0.598	0.735
(4) BASIC + $3D_{s+e}$	0.376	0.636	0.768

Note: 3D, 3D variables (including eye-level and sky-view); BASIC, basic hedonic price model without 3D variables; GWR, Geographically Weighted Regression; OLS, Ordinary Least Squares; RF, Random Forest.

Table 3
The comparison of R^2, RMSE, MAE and MAPE among different models (based on group 4). The unit of RMSE and MAE: yuan/$m.^2$.

Indicator	OLS	GWR	RF
R^2	0.376	0.636	0.768
RMSE	3116.65	2331.41	1669.60
MAE	2199.91	1582.27	1006.63
MAPE	0.16	0.11	0.08

Note, GWR, Geographically Weighted Regression; MAE, mean absoulte error; MAPE, mean absolute percent error; OLS, Ordinary Least Squares; RF, Random Forest; RMSE, root mean square error.

different regards.

Moran' I in this study is 0.78, which means a significant spatial autocorrelation, a common feature shared by property value data across different urban contexts (Dou et al., 2023). The multi-collinearity level can be accepted if the VIF value is less than 10 (Wu et al., 2022). We noticed that within 3D variables, only $3D_{eb}$ and $3D_{es}$ have strong linear correlation issues in group 4, with VIF values of 17.588 and 17.054, respectively. It means that the building and sky in eye-level images have a multi-collinear relationship, which is explainable because they are the main recognised semantic categories in one image and have offset effects on each other. The reasons that the two variables are kept in the models are as follows: 1) 3D is the main focus of this study; 2) we conducted the major part of the analysis based on RF, which tolerates the multi-collinearity among variables (Breiman, 2001).

3.2. Spatial effect of 3D variables: evidence from GWR

Table 4 lists the estimated coefficients of variables in GWR, including the mean, minimum (Min), median, and maximum (Max), in combination with the coefficients in OLS. The results of physical and locational variables on property values are consistent with existing valuation studies (Wen et al., 2014; Zheng et al., 2023), so we would focus the interpretations on 3D variables. Among the physical variables, *REL_floor* = 3 and *TOTAL_floor* have negative standardised coefficients of −0.138 and −0.142 ($p < 0.01$), which reflect buyers' preferences for low storey level and fewer storeys of a high-rise residential building.

All OLS coefficients and the GWR median have the same directions. Only $3D_{es}$ is statistically insignificant; all the other variables are statistically significant at 0.01 level. As mentioned before, it has a solid multi-collinear relationship with $3D_{eb}$. Instead, $3D_{eb}$ can be used for explainability. On the positive side, there are $3D_{eb}$, $3D_{et}$, $3D_{eg}$, $3D_{er}$, $3D_{ss}$, which are building, tree, grass, and road in the eye-level images, and sky in the sky-view images. Grass and tree ($3D_{et}$, $3D_{eg}$) share the top two largest values of the coefficients (1.041 and 1.026). Building ($3D_{eb}$) and road ($3D_{er}$) follow as the third and fourth with coefficients of 0.595 and 0.269, respectively. $3D_{ss}$ has a positive coefficient of 0.231. The negative variables are $3D_{ew}$, $3D_{sb}$, $3D_{st}$, which are sidewalk in eye-level images, and building and tree in sky-view images. $3D_{ew}$ holds the most significant absolute coefficient value of −1.102, reflecting the sidewalk's considerable negative impact. The coefficients of $3D_{sb}$ and $3D_{st}$ are −0.024 and −0.068.

Fig. 5 gives an overview of $3D_{ss}$ in GWR. The distribution of $3D_{ss}$ has a trend of expansion from the centre towards the periphery (Fig. 5c and d). The city centre of Xi'an has narrow street width and densely distributed buildings. The north and west peripheries have improved urban design with broader roads and more sky exposure. The distribution of its coefficients has a similar tendency (Fig. 5a). In the centre, $3D_{ss}$ has negative values. The fact shows that the presence of sky is not favoured for higher property values. The negativity reduces and eventually becomes positive as it gets farther from the centre. Fig.5b shows the distribution of the p value of $3D_{ss}$, indicating the geographical coverage when $3D_{ss}$ is statistical significant.

3.3. The contribution of 3D variables: evidence from RF

Fig. 6 displays the relative importance of all the variables. Among the 3D variables, road, tree and grass from eye-level images ($3D_{er}$, $3D_{et}$, $3D_{eg}$) are the top three influential variables with values of 1, 0.619 and 0.541. Sky from sky-view images ($3D_{ss}$) ranks fourth (0.432) in the total and first in sky-view variables. Building from sky-view images ($3D_{sb}$) closely follows with a value of 0.411. The sequence of other 3D variables is as follows: $3D_{eb}$ (0.369), $3D_{es}$ (0.361), $3D_{ew}$ (0.327), standing for building, sky, and sidewalk in eye-level images, separately. Tree in sky-view images ($3D_{st}$), has the least value (0.26) among 3D variables, but it is still ahead of several physical and locational variables.

Fig. 7 illustrates the PDPs of all 3D variables. $3D_{et}$, $3D_{eg}$, $3D_{er}$, $3D_{ss}$, and $3D_{sb}$ share a similar pattern. In the early stage, there is hardly an impact on property values; then, the influence becomes obvious after a

Table 4
The regression results of group 4 in GWR.

Variable	HPM		GWR (bandwidth = 6671.479)			
	Coefficients	Standardised Coefficients	Mean	Min	Max	Median
Room number	1203.804***	0.159***	1143.0	323.3	1973.7	1194.0
REL_floor = 2	22.913	0.143	−111.94	−314.54	70.96	−109.22
REL_floor = 3	−316.676***	−0.138***				
TOTAL_floor	−15.527***	−0.142***	−24.53	−93.53	84.54	−23.54
AREA_num	−13.160***	−0.142***	−18.91	−57.59	8.05	−18.25
ORI	359.452***	0.147***	161.27	−564.11	1137.33	86.48
BUILT_y	180.628***	0.145***	277.7	−211.94	277.74	207.21
DECO	1062.988***	0.157***	795.4	137.7	1678.9	784.83
ELEV = 1	1716.392***	0.166***	1522.70	−25.04	4012.4	1441.14
DIST_cbd	−246.576***	−0.139***	−360.97	−17,208.09	12,362.59	−330.59
DIST_shop	−734.632***	−0.132***	−594.4	−4308.1	3141.6	−499.4
Dist_hos	220.003***	0.146***	−39.5	−19,420.7	12,286.3	249.0
DIST_par	−114.664**	−0.141**	66.87	−1977.37	20,495.49	20.86
DEN_con	−10.847***	−0.142***	−19.57	−824.68	77.01	−8.26
DEN_edu	−10.460***	−0.142***	−9.74	−1614.89	369.82	−12.69
DEN_sar	0.238	0.143	−1.77	−108.05	588.52	4.98
DEN_sub	2.086***	0.143***	1.92	−147.93	129.90	−0.58
DEN_bus	−55.331***	−0.142***	12.77	−2324.95	212.06	8.98
DEN_sub	200.268***	0.145***	115.62	−6493.17	3807.67	109.41
$3D_{eb}$	32,824.630***	0.595***	21,547	−85,055	167,169	20,984
$3D_{es}$	447.816	0.149	19,417	−279,849	738,261	1944
$3D_{et}$	65,194.840***	1.041***	50,074	−307,557	445,101	44,086
$3D_{er}$	9200.529***	0.269***	13,047	−158,162	87,571	12,753
$3D_{eg}$	64,165.820***	1.026***	49,330	−742,635	3,381,151	40,723
$3D_{ew}$	−90,355.550***	−1.102***	−17,209	−225,127	943,558	−14,720
$3D_{sb}$	−12,112.580***	−0.024***	−210.5	−272,292.5	209,371.4	−2412.0
$3D_{ss}$	6447.751***	0.231***	−182.1	−340,716.1	109,583.2	1946.5
$3D_{st}$	−15,285.010***	−0.068***	−28,856	−660,197	193,147	−28,138
Constant	−364,088.400***	−4.872***	−382,278	−551,634	41,221	−417,452

Note: *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$, REL_floor and ELEV are regarded as numerical variable in GWR.

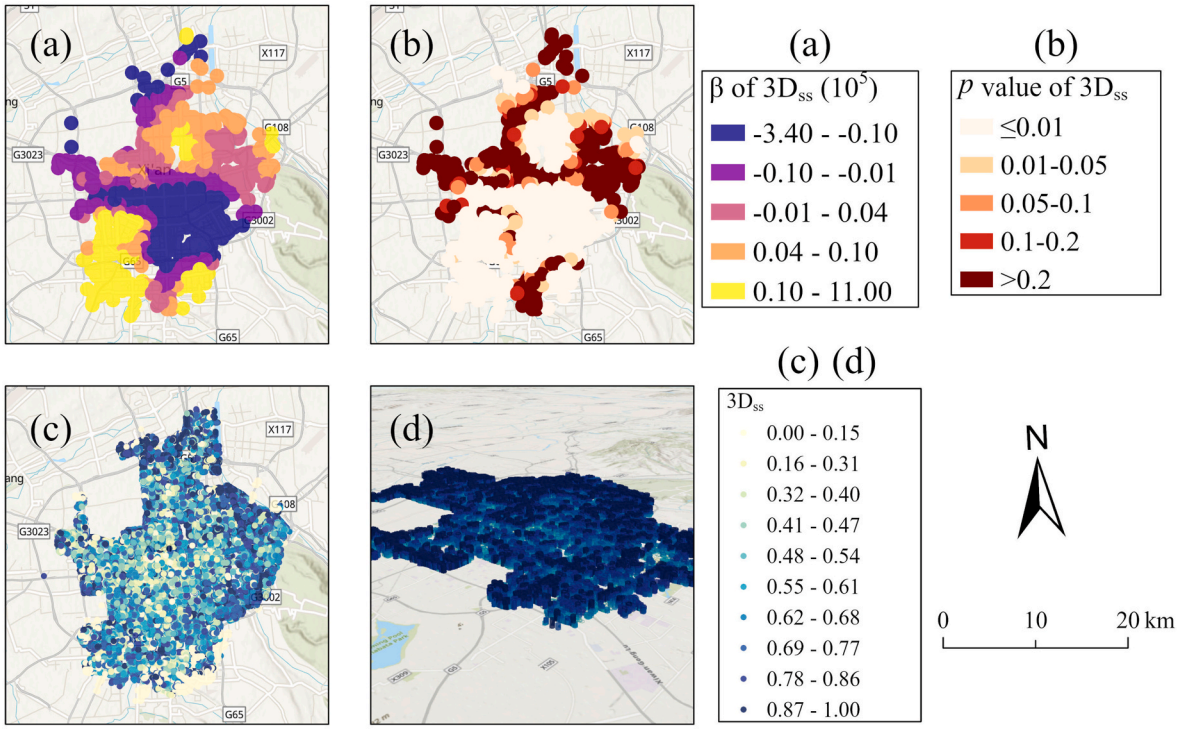


Fig. 5. The distribution of $3D_{ss}$ in GWR model: coefficient (a), *p* value (b), variable itself (c), and 3D visualisation (d).

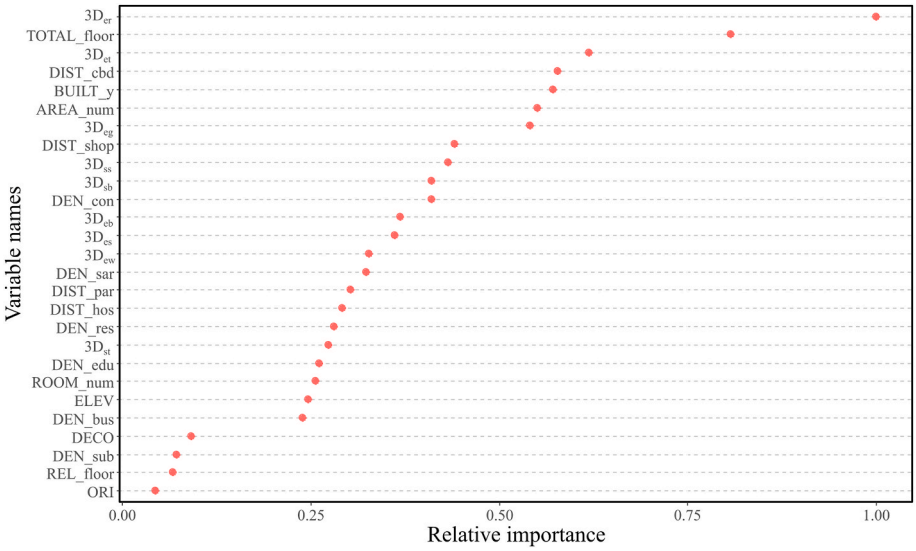


Fig. 6. The relative importance of variables in Random Forest (RF) model.

certain value. After reaching another value, the impact again becomes insignificant. Certain values vary among different variables. In eye-level images, tree ($3D_{et}$) starts at 0.10 and ends at 0.15; grass ($3D_{eg}$) starts at 0.08 and ends at 0.22, road ($3D_{er}$) starts at 0.23 and ends at 0.26. In sky-view images, sky ($3D_{ss}$) starts at 0.30 and ends at 0.65, building ($3D_{sb}$) starts at 0.06 and ends at 0.12.

Building and sky in eye-level images share a pattern. They firstly have a positive contribution, but the positivity drops after a certain value and starts a negative impact. Building ($3D_{eb}$) starts to drive up property values at 0.12 considerably, and after 0.18, the effect becomes negative until it reaches 0.3, and the curve gets stable. Sky ($3D_{es}$) has a positive influence when the value ranges from 0.1 to 0.18, and the influence becomes insignificant after 0.24.

Sidewalk in eye-level images ($3D_{ew}$) and tree in sky-view images ($3D_{st}$) have another pattern: it impacts negatively at the beginning, and after a certain value, it becomes a positive variable. $3D_{ew}$ gradually declines from the start until 0.04, reaching its bottom and then regaining growth. This means that $3D_{ew}$ first reduces the property value with its increase. After 0.04, it has a positive influence, and after 0.046, the influence becomes insignificant. Regarding $3D_{st}$, the decrease tendency stops at 0.05. The impact becomes inapparent after 0.08.

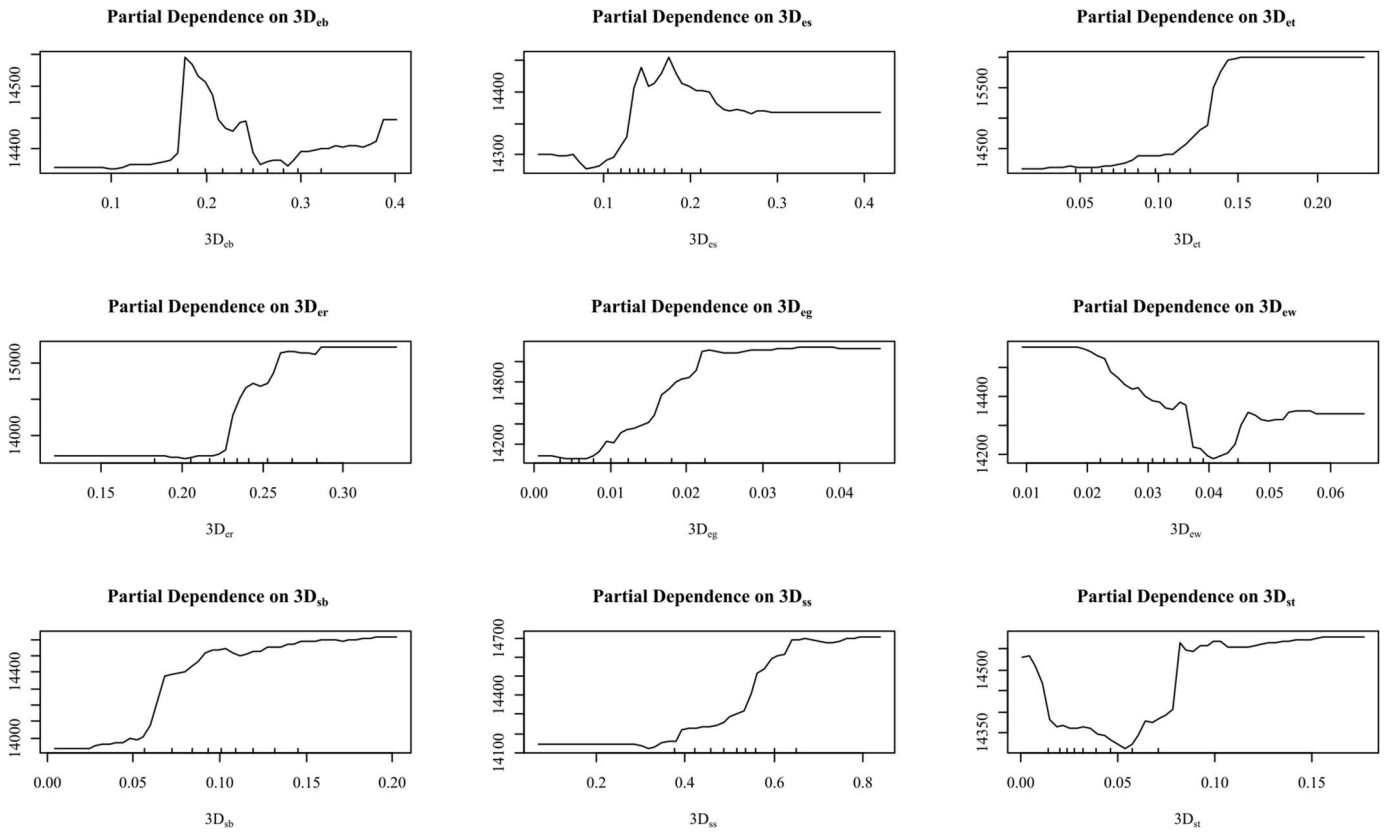


Fig. 7. The partial dependence plots (PDPs) of 3D variables.

4. Discussions

4.1. 3D variables in eye-level images

In GWR, eye-level greenery (tree and grass) has the greatest positive influence on property values, which is consistent with the existing knowledge (Li et al., 2015; Zhang et al., 2018). Building ($3D_{eb}$) ranks the third. More building presence in images may stand for a more centralised location, an important indicator always appreciated by buyers (Ying et al., 2022). There is conflict in the paved area: sidewalk ($3D_{ew}$) has a negative influence, while road ($3D_{er}$) has a positive influence. From a macroscopic urban perspective, this is evidenced by the current car-oriented urban transportation in China (Quan et al., 2020). People prefer a wide road as an indicator of good transportation access, while a wide sidewalk is not advantageous. Another local reason is the urban infrastructure. Renowned as an ancient capital, the downtown areas of Xi'an are characterised by old infrastructure – narrow sidewalks – while in the newly developed districts, which have comparatively lower property values and are remote from CBD, the quality and width of the roads and sidewalk are improved.

The presence of road ($3D_{er}$) receives the most significant relative importance, which can be related to how SVIs are collected. SVIs are typically captured by custom-made cameras mounted on top of the vehicles and, thus by nature, contain a large share of roads in the images (Li et al., 2015). The ranking of greenery ($3D_{et}$, 2nd and $3D_{eg}$, 3rd) robustly proves its important role. The PDPs show that tree, grass, road ($3D_{et}$, $3D_{eg}$ and $3D_{er}$) are always positive variables. Building, sidewalk and sky ($3D_{eb}$ and $3D_{ew}$) show a mixed effect. This can be because property values increase when there are more built-up elements in the downtown areas, but the trend reverses in suburban areas.

PDPs visualise when variables start or stop to impact property values. These values set up important references for urban planners that the presence of specific objects has an optimal value range for price premiums. If it is not in the range, the variables may have an impact in opposite directions or barely have an impact. For example, the optimal range of the tree ($3D_{et}$) is between 0.10 and 0.15. This means that if the presence of trees is under 10%, it hardly makes an impact; if the presence is over 15%, the impact becomes insignificant. Urban planners can optimise their planning strategy according to these values to avoid wasting resources and investment. This rule applies to both angles.

4.2. 3D variables in sky-view images

In the group comparison, the increase of R^2 of the model with sky-view images is smaller than in eye-level images (0.019–0.047 in RF; see Table 2). Nevertheless, it does not make sky-view images less important than eye-level ones. The fact aligns with the inherent geo-information richness. The eye-level images capture richer geo-information about the streetscape, while the sky-view angle mainly documents sky, tree and building. In GWR, the distribution of $3D_{ss}$ is uneven and spatially heterogeneous (Fig. 5c). One of the important findings is the mixed effect of $3D_{ss}$ as its coefficients are negative in the centre but become positive as it grows out of the city. In practice, it means that in the downtown area, people pay more attention to the 2D locations while other influencing variables become insignificant, which is reversed in the suburban areas. It reminds policymakers to adopt fit-for-purpose land use policies in different city regions and highlights the need for adequate sunlight in vertical cities (Wu et al., 2024). The negativity imposed by $3D_{sb}$ and $3D_{st}$ are reasonable. More buildings and trees in sky-view images stand for less sky exposure and narrower street openness enjoyed by the residents. The street canyons with high-rise and closely-packed buildings will influence the urban micro-climate and residential quality (Chen et al., 2010), which deserves to be a primary concern for urban planning. It also aligns with the abovementioned negative effect of *TOTAL_floor*, that too many high-rise buildings in the viewshed may reduce the property value.

The ranking of $3D_{ss}$ and $3D_{sb}$ (sky, 4th and building, 5th) in relative feature importance (Fig. 6) reveals the important role of sky-view angle. The ranking of building needs particular attention. With an average of 10% presence in the image, the building reaches an importance approximately equal to the sky (average presence of 51%). Only buildings high enough can be captured in the sky-view angle, so the construction of high-rise buildings should be carefully considered – as it impacts property values negatively. $3D_{st}$ does not receive high importance; the low presence of trees in the image (average 4%) can be accountable for that.

PDPs reveal that the sky and building are always positive, while the tree shows a mixed effect – first decrease, then increase. This means that only trees within a certain amount (between 0.08 and 0.15) can improve the values, while too few or too many do not exert positivity. It has a different direction from the $3D_{et}$, which is always positive. The difference between the same object at two angles means positive eye-level greenery can sometimes decrease the property value in the sky-view angle. It means the quality and the shape of trees, such as canopy structure, also deserve focus. The findings in Japan support our insights with similar logic that not only poor green views but also excessive green views have a negative impact on values (Yamagata et al., 2016). Similarly, the building also has different impact directions from two angles. The mismatch of the influence directions of the same object indicates the effect of tree and building is non-linear, further proving the importance of studying the sky-view angle for landscape design optimisation.

4.3. Summary of findings

The overall significance of 3D is confirmed by the following facts: 1) the R^2 increase in groups 2, 3 and 4 by adding 3D variables; 2) the statistical significance ($p < 0.05$) of $3D_{ss}$, $3D_{st}$, $3D_{sb}$ from sky-view images; and 3) the average relative importance of 3D variables is 0.481, which outperforms the average value of physical and locational variables (0.323). In addition, we also found the influence of 3D in physical variables, with the negative coefficients of *REL_floor* and *TOTAL_floor,* the relative location of the property and the total floor numbers of a building. This means the residents do not favour either the relatively high-storey level or the absolute tall building height. The reasons could be the long commuting and waiting time between the property and the ground, the negative externalities from wind and traffic noise (Ying et al., 2021), safety concerns (e.g., vertical fire evacuation) (Ding et al., 2021), and a better living experience brought by fewer residents in one building (Ying et al., 2022).

In HPM, RF shows better model performance than GWR and OLS. It also has strong explainability supported by relative feature importance and PDPs (Figs. 6 and 7). They are easy to understand and help researchers explore how and when 3D variables influence property values. Explainability should not only refer to being more accessible to explain by the researchers who create the model but also being more understandable to users (Gevaert, 2022). Algorithms with higher explainability and predictive accuracy with user-friendly visualisation are preferable in valuation because the homeowners can dispute their valuation or the property tax based on it (and even appeal to the courts). The taxation authorities then need to be able to give a proper explanation of how the value is reached.

4.4. Limitations and ways forward

This study has the following limitations: first, the sky-view angle can only simulate but not fully restore the view from a high-storey level because SVIs are taken on the ground. SVIs do not represent the exact view from the windows of an apartment or the internal neighbourhood environment; they reflect the visible streetscape alongside the main roads to where vehicles have access. By its nature, the perspective of SVIs is opposed to the apartments themselves and from the standpoint of pedestrians, and it provides rich semantic information about the surrounding environment (Dai et al., 2024). City-scale photo datasets showing views from apartments are unavailable at present, and simulating such views is computationally intensive and highly complex (Ito, Quintana, Han, Zimmermann, & Biljecki, 2024). These limitations restrict the scalability and transferability of studies in other urban contexts relying on views from the apartments. Our results further prove that SVIs can be a high-quality and cost-effective data alternative to provide 3D geoinformation when the large-scale 3D data of the study area is not available, which makes it a pragmatic tool for researchers and policymakers in rapidly urbanised cities globally. Second, we extracted semantic information from 2D-based SVIs in two different angles to simulate a 3D perspective due to the lack of city-scale 3D data; more specifically, the images in sky-view angle are used to simulate the vertical angle. In this case, the 3D variables serve as the proxies of the actual 3D built environment. Third, we also assumed there were no time-related influences on the property values in the half-year period, such as inflation or currency depreciation.

This study provides insights into future works aiming for reconstructing 3D built environment. Building height information can be a promising 3D data type in the data aspect, as it is a key attribute for reflecting urban form, human activities, and human-environment interactions (Wu et al., 2023; Yang et al., 2022). High-resolution drone image is also a promising complementary source to provide property-level data, which is still absent at a large scale (Tan et al., 2021). The potential of SVIs can be extended from static semantics to dynamic information (e.g., noise, sunlight, and ventilation) to quantify the economic value of living comfort (Diener et al., 1997). Besides, combined with socioeconomic data, it is also meaningful to track if spatio-temporal changes in the 3D built environment can be linked to segregation of different socioeconomic groups or accessibility of specific ecosystem services (Csomós et al., 2024; Luo et al., 2024; Tang et al., 2025). Regarding the technology, multi-modal deep learning models support diverse data input, which can be a promising research direction (e.g., combining text data with images). In traditional valuation, texts (e.g., the names and descriptions of properties and public facilities) are usually excluded from statistical models as they do not accommodate textual data. Therefore, the messages from texts remain yet unknown. Transformer is an exemplary case, a type of neural network architecture known for its attention mechanism. The multi-modal fusion of texts and images is proven valid for building semantic labelling, which facilitates a thorough understanding of the urban environment (Zhou et al., 2023). Differentiating what kind of building facades add or decrease environmental aesthetics is an inspiring topic. The challenges would lie in the efficiency of model training with massive data while keeping AI explainability. It is also worth highlighting that we primarily focus on using machine learning, i.e., RF in this study, to provide explainability and interpretations on 3D variables. We keep the parameter setting in a default mode without extensive adjustments. In future works, hyperparameter tuning (e.g., number of trees, maximum depth, and feature selection strategies) could be explored further to improve the model performance in a computationally efficient way. To move a step beyond data and technology, the importance of 3D in property valuation calls for more formal recognition to foster equal land use policy and fair property taxation in vertically developed cities. For example, it is reasonable to raise the property tax for brighter and broader views. However, this aspect cannot be reflected in the current taxation criteria. The first step can be an extension to existing standards, such as the Land Administration Domain Model (LADM) (Unger et al., 2021), to provide structured documentation and rules for 3D valuation data as a foundation for legalisation.

5. Conclusion

With the spatial changes happening in the vertical direction in cities because of the construction of high-rise buildings, it is necessary to conceptually and technically identify the hedonic value of 3D and add

3D variables in HPM for property valuation. Our empirical results show: 1) 3D variables have made significant contributions to HPM, and the sky-view angle cannot be overlooked. 2) In the sky-view angle, sky is positively correlated while the presence of buildings and trees are negatively correlated with property values. The influence direction of building and tree are different in eye-level and sky-view images, which further reminds the importance of sky-view angle. 3) RF outperforms OLS and GWR regarding a higher model fitting and stronger explainability power to interpret the influences of 3D variables. The study confirms the important role of 3D in HPM in the context of vertically developed cities with generous amounts of high-rise buildings. Our findings are expected to provide consultancy value for housing policy-making, such as fair property taxation and equal land use policy. In conclusion, this study provides solid evidences that SVIs and AI can quantify the 3D built environment at a low cost on a mass scale (*technically feasible*), and the sky-view angle is meaningful in valuation practices in vertical cities (*practically significant*), laying a foundation for future 3D property valuation.

CRediT authorship contribution statement

Yue Ying: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Data curation, Conceptualization. **Shaoqing Dai:** Writing – review & editing, Visualization, Software, Methodology, Data curation. **Mila Koeva:** Writing – review & editing, Supervision, Methodology. **Monika Kuffer:** Writing – review & editing, Supervision, Methodology. **Claudio Persello:** Writing – review & editing, Methodology. **Wen Zhou:** Writing – review & editing, Methodology. **Jaap Zevenbergen:** Writing – review & editing, Supervision.

Declaration of competing interest

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.

Acknowledgement

Yue Ying is funded by China Scholarship Council (CSC) under the grant number [201906560015]. Shaoqing Dai is funded by China Scholarship Council (CSC) under the grant number [201904910428]. Wen Zhou is funded by China Scholarship Council (CSC) under the grant number [201906400036].

References

Alavipanah, S., et al. (2018). The effect of multi-dimensional indicators on urban thermal conditions. *Journal of Cleaner Production, 177*, 115–123.
Antipov, E. A., & Pokryshevskaya, E. B. (2012). Mass appraisal of residential apartments: An application of Random forest for valuation and a CART-based approach for model diagnostics. *Expert Systems with Applications, 39*(2), 1772–1778.
Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*(12), 2481–2495.
Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32.
Central Committee of the Communist Party of China. (2017). *The Report of the 19th national Congress of the Communist party of China*. Beijing.
Chen, L., et al. (2010). Sky view factor analysis of street canyons and its implications for daytime intra-urban air temperature differentials in high-rise, high-density urban areas of Hong Kong: A GIS-based simulation approach. *International Journal of Climatology, 32*(1), 121–136.
Cordts, M., et al. (2016). The cityscapes dataset for semantic urban scene understanding. In *Proceedings of the IEEE Conference on computer vision and pattern recognition (CVPR)*.
Csomós, G., Farkas, J. Z., & Kovács, Z. (2024). A GIS-based assessment of different income groups' access to multiple types of green areas in Budapest, Hungary. *Habitat International, 146*, Article 103054.
Dai, S., et al. (2023b). Assessing spatiotemporal bikeability using multi-source geospatial big data: A case study of Xiamen, China. *International Journal of Applied Earth Observation and Geoinformation, 125*, Article 103539.
Dai, S., et al. (2024). Street view imagery-based built environment auditing tools: a systematic review. *International Journal of Geographical Information Science, 38*(6), 1136–1157.
Dai, X., Felsenstein, D., & Grinberger, A. Y. (2023). Viewshed effects and house prices: Identifying the visibility value of the natural landscape. *Landscape and Urban Planning, 238*, Article 104818.
Diener, E., & Suh, E. (1997). Measuring quality of life: economic, social and subjective indicators. *Social Indicators Research, 40*(1), 189–216.
Ding, N., et al. (2021). State-of-the-art high-rise building emergency evacuation behavior. *Physica A: Statistical Mechanics and Its Applications, 561*, Article 125168.
Dou, M., Gu, Y., & Fan, H. (2023). Incorporating neighborhoods with explainable artificial intelligence for modeling fine-scale housing prices. *Applied Geography, 158*, Article 103032.
Fleming, D., et al. (2018). Valuing sunshine. *Regional Science and Urban Economics, 68*, 268–276.
Frolking, S., et al. (2024). Global urban structural growth shows a profound shift from spreading out to building up. *Nature Cities, 1*(9), 555–566.
Fu, X., et al. (2019). Do street-level scene perceptions affect housing prices in Chinese megacities? An analysis using open access datasets and deep learning. *PLoS One, 14*(5), Article e0217505.
Gevaert, C. M. (2022). Explainable AI for earth observation: A review including societal and regulatory perspectives. *International Journal of Applied Earth Observation and Geoinformation, 112*, Article 102869.
Helbich, M., et al. (2013). Boosting the predictive accuracy of urban hedonic house price models through airborne laser scanning. *Computers, Environment and Urban Systems, 39*, 81–92.
Hong, J., Choi, H., & Kim, W.-s. (2020). A house price valuation based on the random forest approach: The mass appraisal of residential property in South Korea. *International Journal of Strategic Property Management, 24*(3), 140–152.
Huang, B., Wu, B., & Barry, M. (2010). Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *International journal of geographical information science, 24*(3), 383–401.
Ito, K., Quintana, M., Han, X., Zimmermann, R., & Biljecki, F. (2024). Translating street view imagery to correct perspectives to enhance bikeability and walkability studies. *International Journal of Geographical Information Science, 38*(12), 2514–2544.
Kara, A., et al. (2023). Visualisation and dissemination of 3D valuation units and groups – an LADM valuation information compliant prototype. *Land Use Policy, 132*, Article 106829.
Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy, 74*(2), 132–157.
Li, X., et al. (2015). Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forestry and Urban Greening, 14*(3), 675–685.
Li, X., et al. (2016). Urbanization and health in China, thinking at the national, local and individual levels. *Environmental Health, 15*(Suppl 1), S32.
Li, M., et al. (2020). Continental-scale mapping and analysis of 3D building structure. *Remote Sensing of Environment, 245*, Article 111859.
Luo, M., Zhang, S., & Deng, W. (2024). Has urban expansion alleviated working-residential spaces segregation across inner-outer cities? A multi-scale study with location-based social bigdata. *Habitat International, 153*, Article 103183.
Persello, C., & Bruzzone, L. (2014). Active and semisupervised learning for the classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing, 52*(11), 6937–6956.
Persello, C., & Stein, A. (2017). Deep fully convolutional networks for the detection of informal settlements in VHR images. *IEEE Geoscience and Remote Sensing Letters, 14*(12), 2325–2329.
Qiu, W., et al. (2022). Subjective or objective measures of street environment, which are more effective in explaining housing prices? *Landscape and Urban Planning, 221*, Article 104358.
Quan, B., & Xu, D. (2020). China's urban transport policy: From car-oriented to people-oriented cities. In H. P. Chia-Lin Chen, Q. Shen, & J. J. Wang (Eds.), *Handbook on Transport and urban Transformation in China* (pp. 390–403). Edward Elgar Publishing Limited.
Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy, 82*(1), 34–55.
Shaanxi Provincial Bureau of Statistics & NBS Survey Office in Shaanxi. (2023). *Shaanxi statistical Yearbook 2022*.
Suzuki, M., et al. (2022). The economic value of urban landscapes in a suburban city of Tokyo, Japan: A semantic segmentation approach using google street view images. *Journal of Asian Architecture and Building Engineering, 22*(3), 1110–1125.
Tan, L. K. L., et al. (2021). Public acceptance of drone applications in a highly urbanized environment. *Technology in Society, 64*, Article 101462.
Tang, J., & Long, Y. (2019). Measuring visual quality of street space and its temporal variation: Methodology and its application in the Hutong area in Beijing. *Landscape and Urban Planning, 191*, Article 103436.
Tang, Y., Xiao, W., & Yuan, F. (2025). Evaluating objective and perceived ecosystem service in urban context: An indirect method based on housing market. *Landscape and Urban Planning, 254*.
Unger, E.-M., et al. (2021). LADM for sustainable development: An exploratory study on the application of domain-specific data models to support the SDGs. *Land Use Policy, 108*, Article 105499.
Wang, G. Y. (2022). The effect of environment on housing prices: Evidence from the Google Street View. *Journal of Forecasting, 42*(2), 288–311.
Wang, M., & Vermeulen, F. (2020). Life between buildings from a street view image: What do big data analytics reveal about neighbourhood organisational vitality? *Urban Studies, 58*(15), 3118–3139.

Wen, D. W., et al. (2019). Monitoring 3D building change and urban redevelopment patterns in inner city areas of Chinese megacities using multi-view satellite imagery. *Remote Sensing, 11*(7).
Wen, H., Zhang, Y., & Zhang, L. (2014). Do educational facilities affect housing price? An empirical study in hangzhou, China. *Habitat International, 42*, 155–163.
Wu, C., et al. (2022). Does visual contact with green space impact housing prices? an integrated approach of machine learning and hedonic modeling based on the perception of green space. *Land Use Policy, 115*, Article 106048.
Wu, W.-B., et al. (2023). A first Chinese building height estimate at 10 m resolution (CNBH-10 m) using multi-source earth observations and machine learning. *Remote Sensing of Environment, 291*, Article 113578.
Wu, S., et al. (2024). The interplay of cloud cover and 3D urban structures reduces human access to sunlight. *Nature Cities, 1*(10), 686–694.
Wu, Y. Y., Wei, Y. H. D., & Li, H. (2020). Analyzing spatial heterogeneity of housing prices using large datasets. *Applied Spatial Analysis and Policy, 13*(1), 223–256.
Yamagata, Y., et al. (2016). Value of urban views in a bay city: Hedonic analysis with the spatial multilevel additive regression (SMAR) model. *Landscape and Urban Planning, 151*, 89–102.
Yang, C., & Zhao, S. (2022). A building height dataset across China in 2017 estimated by the spatially-informed approach. *Scientific Data, 9*(1), 76.
Yao, Y., et al. (2019). A human-machine adversarial scoring framework for urban perception assessment using street-view images. *International Journal of Geographical Information Science, 33*(12), 2363–2384.
Ying, Y., et al. (2021). Making the third dimension (3D) explicit in hedonic price modelling: A case study of Xi'an, China. *Land, 10*(1), 24.
Ying, Y., et al. (2022). The perception of the vertical dimension (3D) through the lens of different stakeholders in the property market of China. *Land, 11*(2), 312.
Ying, Y., et al. (2023). Toward 3D property valuation—a review of urban 3D modelling methods for digital twin creation. *ISPRS International Journal of Geo-Information, 12*(1), 2.
Yu, S. M., Han, S. S., & Chai, C. H. (2007). Modeling the value of view in high-rise apartments: A 3D GIS approach. *Environment and Planning B: Planning and Design, 34*(1), 139–153.
Yue, H., Liu, L., & Xiao, L. (2023). Investigating the effect of people on the street and streetscape physical environment on the location choice of street theft crime offenders using street view images and a discrete spatial choice model. *Applied Geography, 157*, Article 103025.
Zhang, Y., & Dong, R. (2018). Impacts of street-visible greenery on housing prices: Evidence from a hedonic price model and a massive street view image dataset in beijing. *ISPRS International Journal of Geo-Information, 7*(3).
Zhang, L., Zhou, J., & Hui, E. C.m. (2020). Which types of shopping malls affect housing prices? From the perspective of spatial accessibility. *Habitat International, 96*, Article 102118.
Zheng, W., & Wang, M. (2023). Three-dimensional land-use configuration and property prices: A spatially filtered multi-level modelling perspective. *Environment and Planning B: Urban Analytics and City Science, 51*(2), 438–455.
Zhou, W., et al. (2023). Building use and mixed-use classification with a transformer-based network fusing satellite images and geospatial textual information. *Remote Sensing of Environment, 297*, Article 113767.