

Re-Evaluating the Premium Effect of Urban Green Space

Duogui Yang, Junyi Guo

yangdg@casisd.cn guojunyi241@mailsucas.ac.cn

Institute of Science and Development, CAS, China; School of Public Policy and Management, CAS, China;

Abstract

Urban green space (UGS) delivers ecological and livability benefits and creates housing price premiums. Quantifying UGS-related premiums and their spatial heterogeneity is critical for green-space layout and refined urban governance. Constrained by linear assumptions and flawed parameter tuning, existing research mostly confirms premium existence rather than quantifying scale-dependent premium discrepancies. Using 2591 residential-POI samples from Huangdao District (Qingdao), this study applies PSO-RF and WKmeans-SMOTE hybrid machine-learning models for empirical analysis. Results reveal the feature contribution order: greening rate > straight-line distance to green space > actual distance to green space. The premium exhibits notable scale heterogeneity: greening rate governs premiums for large-scale UGS, whereas accessibility matters more for small-scale UGS. Higher green-space effect intensity corresponds to larger housing premiums, alongside stronger price dispersion for large-scale UGS. The models overcome traditional methodological limits, offering empirical evidence for green-space planning and sophisticated real-estate regulation.

Keywords:

urban green space, housing price premium effect, spatial heterogeneity, machine learning

Existing studies employ hedonic pricing models and standalone machine-learning tools to investigate green-space-housing-price linkages [1,4,5]. The hedonic model is popular for preliminary premium testing yet constrained by linear and global-homogeneity assumptions, which hamper its capacity to capture market nonlinearity and spatial heterogeneity [1,7]. Though random-forest-type machine-learning approaches partly remedy such defects, three key shortcomings remain [2,5]. Conventional random forests using grid-search hyperparameter tuning risk local optima for high-dimensional housing datasets [6]. Few works conduct stratified quantification across green-space scales to compare premium differences. Moreover, unaddressed sample imbalance and unweighted clustering distort premium-effect grouping [2]. No integrated framework exists combining parameter tuning, sample balancing and heterogeneity measurement, creating gaps for refined premium assessment [3].

1. Introduction

Against the backdrop of high-quality urbanization, human-settlement quality boosts urban competitiveness and public well-being. As a scarce ecological resource, urban green space alleviates urban heat-island effects, improves air and landscape conditions, and acts as a core metric for urban livability. Quality green space creates housing-price premiums, an implicit driver of residential-price spatial divergence. Driven by intensive land use, green spaces show uneven distribution and scale differences, bringing heterogeneous value-added impacts on housing markets. Quantifying green-space premium magnitude and its spatial heterogeneity helps measure ecological economic value and supports balanced green-space deployment for coordinated urban ecological-economic-social development.

This paper investigates Huangdao District (Qingdao), which has an 83.52% urbanization rate and diverse green-space types ranging from large parks to pocket parks, offering typical conditions to study scale-related premium patterns. The dataset contains 2591 valid samples merging Lianjia residential transaction records (2016-2023) and Amap POI data. After ArcGIS coordinate correction and spatial matching, 31 location-, community- and architecture-related independent variables are retained. Samples are further split into large-scale and small-scale green-space subgroups for stratified empirical analysis.

2. Methodology

2.1 PSO-RF Model

Random Forest (RF) suits multi-variable large-sample housing-price analysis with strong nonlinear-fitting and anti-overfitting capacity, though conventional grid-search hyperparameter tuning is inefficient and prone to local optima. This paper applies Particle Swarm Optimization (PSO) to obtain globally optimal hyperparameters including tree quantity and split-relevant parameters through particle-population iteration and information sharing for robust fitting and feature-contribution outputs [6]. Using MSE as the fitness function, MAE, MSE, RMSE and R^2 are utilized to evaluate model performance and quantify feature-contribution gaps of greening rate and two distance indicators across green-space scales.

2.2 WKmeans-SMOTE Hybrid Clustering Model

Traditional Kmeans applies equal feature weights and poorly handles imbalanced samples, producing inaccurate grouping for green-space effect intensities [2]. This study builds the WKmeans-SMOTE hybrid model: weighted Kmeans assigns differentiated weights to core green-space features according to PSO-RF-derived feature importance; SMOTE oversampling balances minority-class samples to remove size-driven clustering bias; secondary weighted clustering finally classifies samples into strong, moderate and weak green-space-effect subgroups.

3. Empirical Results and Analysis

3.1 Model Performance Comparison

The fitting performance of PSO-RF and grid-search-based RF is compared. For large-scale green space samples, PSO-RF achieves ($R^2=0.92$) and MAE=697.57, while the traditional model obtains ($R^2=0.83$) and MAE=1173.09. For small-scale green space samples, PSO-RF presents ($R^2=0.86$) and MAE=258.04, in contrast to ($R^2=0.73$) and MAE=452.38 of the traditional model. PSO-RF exhibits lower errors and higher fitting accuracy,

effectively avoiding local optimal defects and ensuring reliable feature importance outputs.

Further comparison of clustering algorithms based on SW, CH, and DB indicators shows that WKmeans-SMOTE achieves higher CH values and lower DB values for both large and small samples. The SMOTE oversampling effectively resolves sample imbalance and improves intra-cluster compactness and inter-cluster separation, making the model suitable for classifying strong, moderate, and weak green space effect groups.

3.2 Heterogeneity of Feature Importance Under Different Green Space Scales

The feature importance ranking remains consistent for both sample groups: greening rate > straight-line distance to green space > actual distance to green space, while significant numerical differences exist across scales. For large-scale green spaces, the importance values of greening rate, straight-line distance, and actual distance are 0.045, 0.009, and 0.007, respectively. For small-scale green spaces, the corresponding values are 0.031, 0.021, and 0.018.

Sufficient ecological supply of large-scale green spaces makes greening rate the dominant premium factor, whereas the limited radiation capacity of small-scale green spaces significantly increases the weight of spatial accessibility indicators [1,4]. In addition, architectural attributes (building area, house age) and public facility supporting variables show higher importance than green space features, indicating that green space premiums belong to superimposed value appreciation based on fundamental housing and public service attributes.

3.3 Spatial Heterogeneity of Green Space Premium Effects

After Min-Max standardization, WKmeans-SMOTE clustering results and Gaussian KDE peak values show that the housing price peaks of strong, moderate, and weak effect groups for large-scale green spaces are 15000 CNY, 13000 CNY, and 11000 CNY, while those for small-scale green spaces are 13000 CNY, 10000 CNY, and 9500 CNY, respectively.

Empirical results verify significant spatial heterogeneity. First, under equivalent effect intensity, large-scale green spaces present higher price peaks due to more comprehensive ecological service functions and stronger value spillover effects. Second, regardless of green space scale, housing prices follow the gradient of strong effect > moderate effect > weak effect, fully confirming the positive premium effect of urban green space and verifying Hypothesis H1. Third, large-scale green spaces have wider price distribution intervals and higher discreteness due to internal spatial differentiation of landscape and facility quality, while small-scale green spaces feature homogeneous functions and concentrated price distributions. The premium effect is significantly regulated by green space scale, which validates Hypothesis H2.

4. Conclusions and Practical Implications

4.1 Main Conclusions

Based on residential and POI data of Huangdao District, this study adopts PSO-RF and WKmeans-SMOTE hybrid machine learning models to quantitatively analyze UGS premium effects and spatial heterogeneity. The core conclusions are summarized as follows. First, urban green space generates significant positive housing price premiums. Greening rate, straight-line distance, and actual distance constitute core influencing features with stable contribution rankings, dominating the spatial differentiation of residential prices. Second, green space premiums present prominent scale heterogeneity. The premium of large-scale green spaces is driven mainly by greening rate, while small-scale green space premiums rely more on spatial accessibility, forming differentiated value formation mechanisms. Third, green space premium intensity is positively correlated with housing prices. Large-scale green spaces have stronger premium capabilities and more significant spatial price differentiation, whereas small-scale green spaces show weaker premiums and more balanced price distributions.

4.2 Practical Implications

The findings provide practical references for refined urban green space planning and real estate market regulation. First, differentiated green space planning strategies should be implemented. For large-scale urban green spaces, priority should be given to

improving green coverage and optimizing ecological service functions to maximize ecological value spillover. For small-scale community green spaces and pocket parks, spatial site selection and accessibility optimization should be emphasized to compensate for insufficient ecological radiation capacity. Second, urban renewal and real estate value evaluation should fully consider the scale and intensity differences of green space effects. A differentiated ecological premium evaluation system should be established instead of unified criteria to promote equitable allocation of urban spatial resources and refined urban governance [3,7].

References

- [1] Chen H, Wu C. Spatial heterogeneity of urban green space premium on housing prices[J]. *Progress in Geography*, 2024, 43(2): 289-298.
- [2] Li S, Liu Y. Research on influencing factors of urban housing prices based on machine learning[J]. *Statistics & Decision*, 2023, 39(15): 18-23.
- [3] Xu M, Zhang J. Quantitative evaluation and premium mechanism analysis of urban ecological resources[J]. *Ecological Economy*, 2025, 41(1): 112-118.
- [4] Iban M. Spatial heterogeneity of urban green space premium on housing prices[J]. *Sustainable Cities and Society*, 2022, 87: 104236.
- [5] Wu C, Wang L. Research on influencing factors of urban green space housing premium based on random forest[J]. *Urban Problems*, 2022(8): 36-44.
- [6] Zhao Y, Chen X. Application of particle swarm optimization in machine learning hyperparameter tuning[J]. *Computer Engineering and Applications*, 2024, 60(3): 134-140.
- [7] Hu L, Zhou K. Equitable allocation of urban green space and its correlation with housing price premium[J]. *Urban Planning Review*, 2019, 43(7): 68-75.

Duogui Yang

Professor

Institute of Science and Development. CAS

School of Public Policy and Management. CAS

yangdg@casid.cn

Junyi Guo

Master candidate

Institute of Science and Development. CAS

School of Public Policy and Management. CAS

guojunyi241@mailsucas.ac.cn