

Article

Likelihood of Transformation to Green Infrastructure Using Ensemble Machine Learning Techniques in Jinan, China

Khansa Gulshad , Yicheng Wang ^{*}, Na Li, Jing Wang and Qian Yu 

State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research (IWHR), Beijing 100038, China; khansag2@gmail.com (K.G.); lina@iwhr.com (N.L.); wangjing@iwhr.com (J.W.); yuqian@iwhr.com (Q.Y.)

* Correspondence: wangych@iwhr.com

Abstract: Rapid urbanization influences green infrastructure (GI) development in cities. The government plans to optimize GI in urban areas, which requires understanding GI spatiotemporal trends in urban areas and driving forces influencing their pattern. Traditional GIS-based methods, used to determine the greening potential of vacant land in urban areas, are incapable of predicting future scenarios based on the past trend. Therefore, we propose a heterogeneous ensemble technique to determine the spatial pattern of GI development in Jinan, China, based on driving biophysical and socioeconomic factors. Data-driven artificial neural networks (ANN) and random forests (RF) are selected as base learners, while support vector machine (SVM) is used as a meta classifier. Results showed that the stacking model ANN-RF-SVM achieved the best test accuracy (AUC 0.941) compared to the individual ANN, RF, and SVM algorithms. Land surface temperature, distance to water bodies, population density, and rainfall are found to be the most influencing factors regarding vacant land conversion to GI in Jinan.



Citation: Gulshad, K.; Wang, Y.; Li, N.; Wang, J.; Yu, Q. Likelihood of Transformation to Green Infrastructure Using Ensemble Machine Learning Techniques in Jinan, China. *Land* **2022**, *11*, 317. <https://doi.org/10.3390/land11030317>

Academic Editor: Salvador García-Ayllón Veintimilla

Received: 25 January 2022

Accepted: 18 February 2022

Published: 22 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: green infrastructure; vacant land transformation; ensemble stacking; multifunctionality; machine learning

1. Introduction

Urbanization brings both opportunities and challenges, such as cities enhancing the quality of life while urbanization poses a threat to the environment, e.g., excessive pollution, changes to local hydrology, and biodiversity loss. The trend towards urbanization is dramatically increasing, and it is projected that by 2050, nearly 70.0% percent of people will live in urban areas [1].

Consequently, the demand for gray infrastructure (GY) is expected to increase, which is associated with enormous costs. Reliance on GY will be insufficient to meet the growing demand from urbanization, compounded by the effects of climate change and energy scarcity [2]. While cities are challenged to balance urban development and its impact on the environment, green infrastructure (GI) provides opportunities to enhance the resilience of socioecological systems in urban areas. Therefore, protecting, improving, re-establishing, and increasing urban and peri-urban green infrastructure is an important measure for sustainable development of urban areas [3].

GI offers multiple benefits such as provision of ecosystem services, reduction of the urban heat island, increase in biodiversity, and positive effects on human wellbeing [4]. Despite GI multifunctionality, green space area has been significantly reduced in cities, and the majority of vacant land in cities is being converted into gray infrastructure as it is economically viable [5]. At the same time, governments are beginning to realize the importance of urban green spaces. This has pushed urban planners to integrate GI in urban planning and policymaking, which may involve transforming vacant, abandoned, or bare land to GI to make more informed development choices. Therefore, it is important

to understand the trend of land conversion to GI or GY in cities and how changes in their spatiotemporal patterns are linked with the processes that drive these changes.

In recent years, several studies have been conducted for exploring the suitability of vacant land or derelict site conversion to GI or GY using GIS-based multicriteria decision analysis (MCDA) [6,7]. These researches have utilized the weighted linear methods or analytical hierarchy approaches to determine the importance of factors. However, these methods recommend what sites should be converted into GI or GY based on multiple benefits instead of understanding the spatiotemporal trend. Furthermore, these conventional methods lack in certain ways: (1) difficulty in selecting an appropriate method for assigning weights for factors, (2) weights are given based on an expert choice which can be subjective and biased, (3) varying results from different standardization methods, and (4) no standardized methods for results evaluation.

Apart from multicriteria-based methods, other methods have also been developed to simulate and predict future LULC scenarios, such as cellular automata (CA), Markov chain (MC), and agent-based modeling [8]. However, these methods are not capable of dealing with spatial heterogeneity in cities and are applied at large spatial scales [9]. Problems associated with these methods can be solved by using data-driven methods, i.e., the integration of artificial intelligence (AI) with GIS methods. AI-based methods use a mathematical framework to drive the relationship between the trend of conversion to GI/GY and their driving factors. Weights for factors are determined by the trends of the real-world conditions rather than the subjective choices [10]. AI-based methods are built upon large datasets, which enable us to combine information from multiple factors involved in GI development. Finally, data-driven methods are capable of predicting potential future scenarios.

Utilizing AI, Labib [10] predicted the likelihood of conversion to GI or GY along waterways and derelict sites in Manchester city council based on machine learning algorithms LR, ANN, and adaptive neuro-fuzzy inference system (ANFIS). However, these machine learning algorithms still face challenges due to several reasons, such as selection of models from a wide range of AI models available, distinct outcomes of each model [11], complex nonlinear relationship between influencing factors and sites [12], quality of model used [13], and the issue of accuracy.

Simultaneously, researchers have developed ensemble machine learning methods, which combine a set of classifiers working collectively for handling complex and high-dimensional data [13]. These methods provide results with better accuracy by expanding the hypothesis space of the fitting function [12]. Several studies used metalearning techniques with homogeneous models such as bagging [14], boosting [15], AdaBoost [16], and random forest [17]. These methods use the same classifiers multiple times to construct a single model, leading to duplicating a single classifier's shortcomings. To overcome this, heterogeneous ensemble methods are developed which combine different classifiers, such as CNN-RNN-SVM-LR [18], SVM-ANN-NB-LR [12], and ANN-SVM-RF-RS-bagging [11], and can capture the advantages of numerous classifiers.

However, these ensemble methods have not been used before in the domain of vacant land transformation likelihood into GI or GY. Therefore, to fill this knowledge gap, this study aims to (1) propose a new ensemble model based on the heterogeneous stacking algorithm ANN-RF-SVM and the analysis of individual models to understand the spatiotemporal trend of GI and GY development in Jinan, China. (2) Based on GI/GY change analysis within the time frame of 2000–2020, the locations are identified which were converted into GI and GY during these 20 years in Jinan. (3) The trend of land-use change (GI or GY) in Jinan along with driving forces (such as nearby green areas, population density, accessibility, pollution, etc.) is used to train the models, and (4) the trained models predicted the future locations for the transformation of GI or GY. (5) The results of the stacking model are then compared with the performance of conventional standalone models SVM, RF, and ANN. These findings will inform Jinan authorities and urban planners to make GI decisions about specific sites. It will give them information about the factors involved

in GI development and if they need to change the urban planning of Jinan to consider the neglected factors. This study will also be a guide for future studies focusing on such methods for GI development.

2. Materials and Methods

2.1. Study Area

Jinan, extending between $36^{\circ}01' N$ – $37^{\circ}32' N$ northern latitude and $116^{\circ}11' E$ – $117^{\circ}44' E$ eastern longitude, is the capital of Shandong province, occupying an area of about 8133.57 km². It is located in the northwestern part of Shandong province, spanning over six districts, one county-level city (Zhangqiu City), and three counties (Pingyin County, Jiyang County, and Shanghe County) [19] (Figure 1). In this study, only the main central urban area of Jinan is considered, which covers an area of 2083.87 km². Jinan has warm temperate and semihumid monsoon climate conditions with an average annual temperature of 14.6 °C, and an average annual rainfall of 500 mm to 600 mm. The rainfall is unevenly distributed during the year, with the most extreme events occurring between July and August [20].

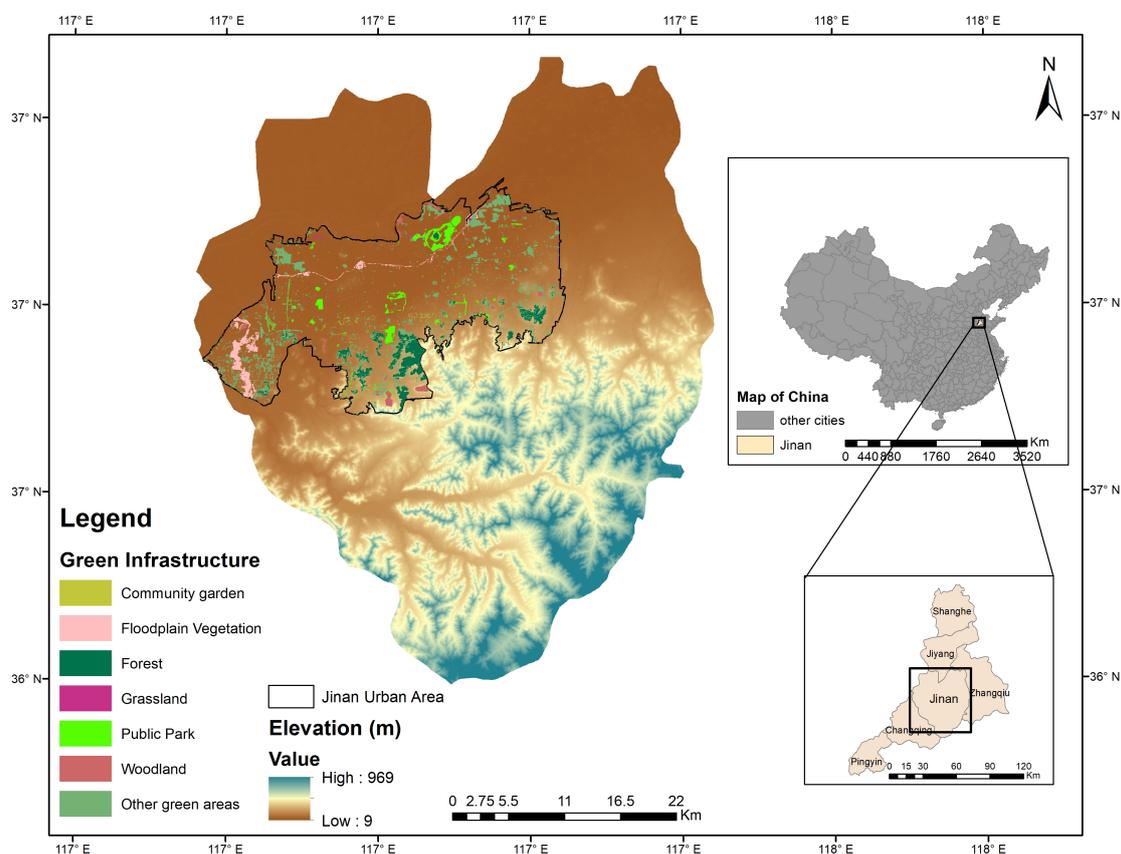


Figure 1. The geographical location of the study area: Location map of Jinan in China (top right), Jinan (bottom right), and Jinan urban area (top) showing multifunctional GI.

According to the Jinan Yearbook (2003–2020), the population of Jinan reached 9.20 million in 2020 compared with 5.82 million people in 2003. In 17 years, Jinan administrative area increased by 1.25 times from 8177 km² to 10,244 km², and built-up area increased by 4.2 times from 200 km² to 839.7 km² [21]. The urbanization in Jinan is constrained by mountains in the south and the Yellow River in the north, hence causing urban sprawl in only two directions [22]. In response to the undesirable impact of urban sprawl, Jinan Municipal Government announced planning for the Ecological Garden City for the years 2010–2020, and in 2021 Jinan Green Regulations were proposed, which will be implemented in March 2022 [23,24].

2.2. Land Use Land Cover (LULC) Change Analysis (2000–2020)

In this work, remote sensing satellite data of Landsat-8 Thematic Mapper (TM) (2020) and Landsat-7 Enhance Thematic Mapper Plus (ETM+) (2010 and 2000) were acquired from the United States Geological Survey (USGS) Earth Explorer [25]. Images were pre-processed in Arcmap 10.8 and radially calibrated and atmospherically corrected in Envi 5.3. The LULC maps for 2000, 2010, and 2020 were derived using supervised maximum likelihood classification in Envi 5.3. The quality of classified images was estimated using Kappa statistics.

2.3. Existing Green Infrastructure Identification

The existing GI in Jinan is located at residential communities, road networks, green belts, parks, river, and lake sites. However, to maximize the benefits of GI, only publicly accessible multifunctional GI (parks, grassland, outdoor areas with rain gardens and wetlands) (Figure 1) data are collected through the LULC change analysis and Baidu maps.

GI with high multifunctionality is coded as 1, while the GY is coded as 0. The GI and GY are divided into training (70%) and test datasets (30%).

2.4. Driving Factors for Transformation

As reviewed by previous studies, the major driving factors that influence GI development are natural, socioeconomic, and policy [6,10]. Hence, considering the existing literature, we selected biophysical factors, i.e., rainfall, air quality, land surface temperature (LST), and distance from river and water bodies. The socioeconomic factors under consideration are the population density, density of vacant land, roads density (accessibility), and built-up area density (surrounding land use). The spatial database of these factors is prepared in ArcGIS-10.8 as a matrix of 1856 columns, 2244 rows with a spatial resolution of $27\text{ m} \times 27\text{ m}$.

2.4.1. Biophysical Factors

Urban areas that are at the risk of flooding, or contribute to downstream flooding, may potentially influence the GI planning decision-making. Therefore, to identify areas producing overflow, the total rainfall of Jinan is taken into account. Jinan's average rainfall per month (2000–2021) data was downloaded from NASA Giovanni for TRMM at $0.25^\circ \times 0.25^\circ$ [26] (Figure 2).

The GI development decision could be affected by air pollution in the surroundings. The air quality index point data for different pollutants such as CO, NO₂, O₃, SO₂, and PM were obtained from Harvard Dataverse for China AQI [27] to prepare the air quality map for Jinan using the IDW interpolation method (Figure 2).

Urban areas having a higher temperature than surrounding areas could be the driving force for land transformation into GI. Hence, to consider the land surface temperature (LST) of Jinan, MODIS/Terra Land Surface Temperature imagery was downloaded from NASA Earthdata DAAC database [28] with a spatial resolution of 30 arcsec. It represents the daytime temperature which ranged between 21 °C and 33 °C for the Jinan.

Vacant land beside rivers or floodplains and water bodies could influence their conversion decision into wetlands or riparian vegetation due to their ecological benefits. For this purpose, buffer maps for distance from water bodies and rivers were prepared for Jinan. The buffer of <50 m, 100 m, 200 m, 300 m, 500, and >500 m was prepared using the spatial river and water bodies data of the city from the OSM network Figure 2.

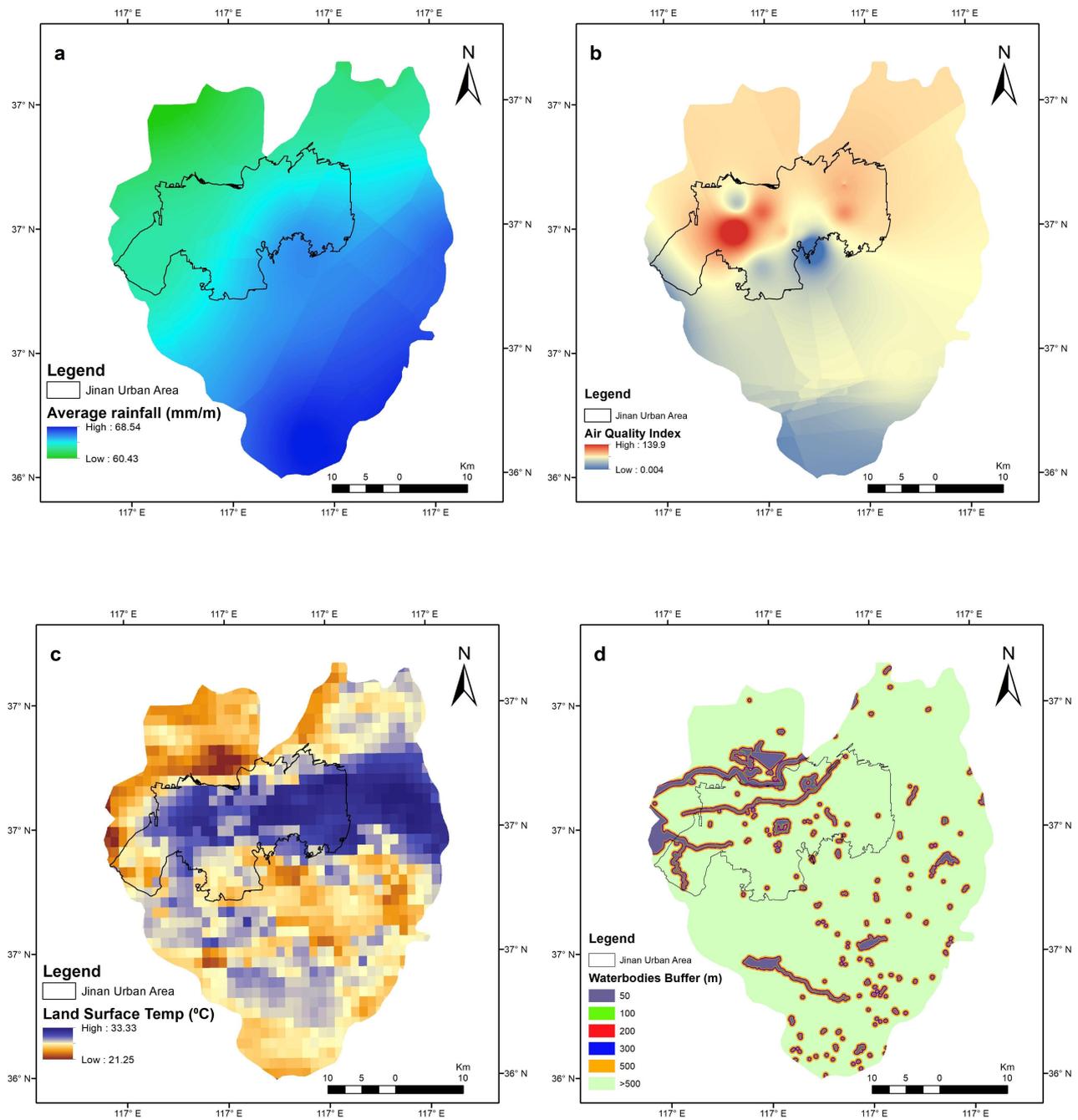


Figure 2. Maps of biophysical factors influencing GI development: (a) rainfall, (b) air pollution, (c) LST, and (d) water bodies buffer.

2.4.2. Socioeconomic Factors

Population density can also be a driving factor for transforming vacant land into GI to improve lifestyle and health. The population density at 30 arcsec resolution [29] for Jinan was obtained from NASA SEDAC gridded population of the world for the year 2020. (Figure 3).

The probability of vacant land conversion into GI depends on the accessibility to the vacant sites, because it means the public can easily access these areas. Hence, a road buffer map was prepared using the spatial data of the road network obtained from the OSM network.

Vacant land with no GI nearby can be preferable for GI development; therefore, the existing green areas can influence GI investment decisions. Thus, a buffer map of distance to vacant land was prepared [30]. A buffer of <50 m, 100 m, 200 m, 300 m, 500, and >500 m was prepared in Arc-GIS 10.8 (Figure 3).

An increase in the built-up area can cause the occupation of green spaces. However, to a certain extent, urbanization can increase green spaces attached to the residential areas. Therefore, a built-up density map for Jinan was prepared by downloading built-up areas from the open street map (OSM) [30]. It was converted into a raster to prepare a kernel density map.

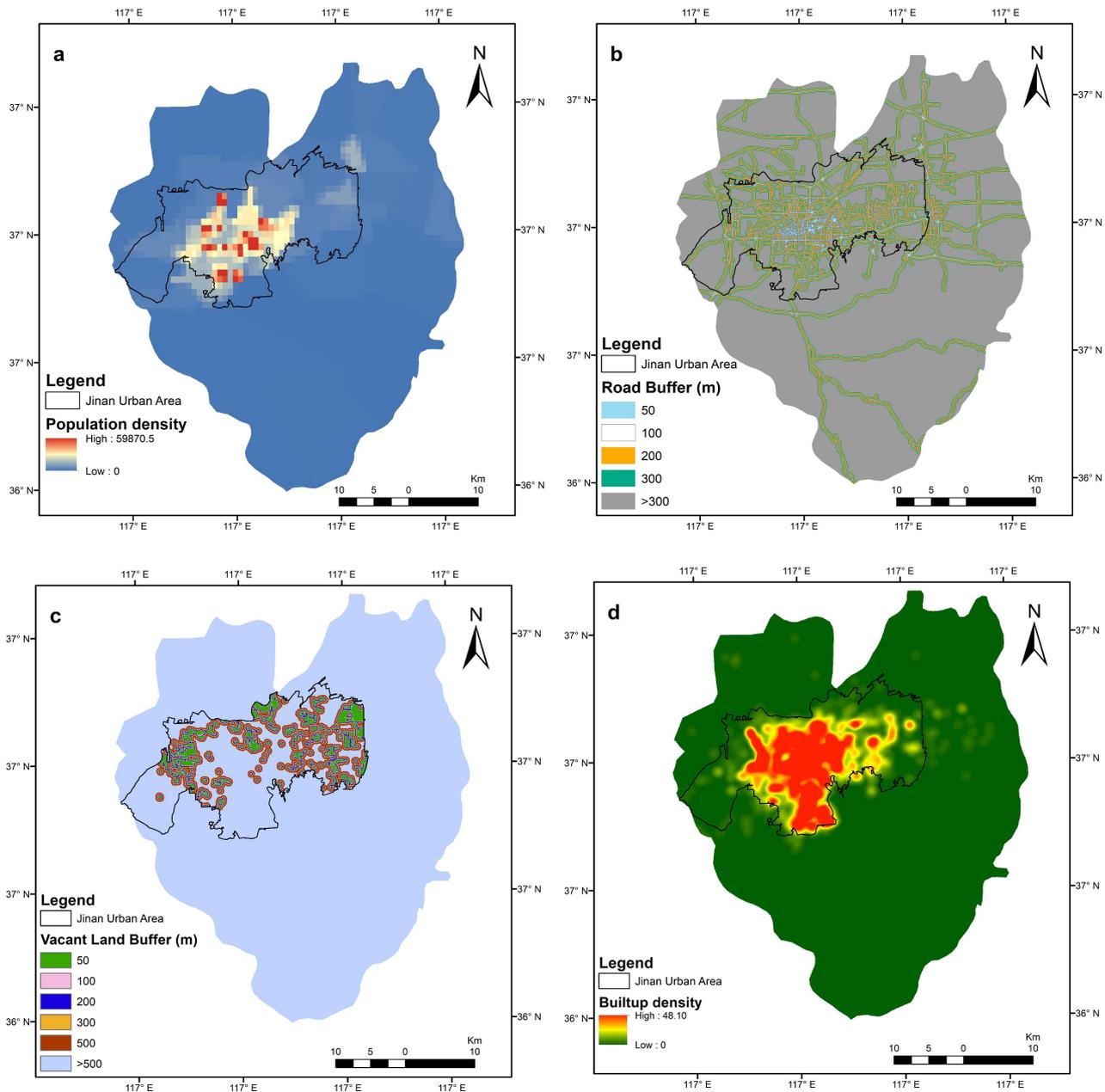


Figure 3. Maps of socioeconomic factors influencing GI development: (a) population density, (b) road buffer, (c) distance from vacant land, and (d) built-up density.

2.5. Importance Analysis of Driving Factors

The evaluation of determinants of GI development is vital before the training and validation of models, as some factors might generate noise or have less predictive power and do not contribute much to achieve better accuracy of the model for the target variable [11,31].

Two multicollinearity tests, i.e., variance inflation factors (VIF) and Pearson's correlation coefficient, are utilized to measure the correlation of factors. VIF for the estimated regression coefficient is influenced or inflated when the independent variables are not linearly related [11]. VIF value of 1 indicates that variables are not related, while a value greater than 5 suggests that collinearity exists [13]. The VIF is determined using Equation (1):

$$VIF = \frac{1}{1 - R^2} \quad (1)$$

where R^2 is a coefficient of determination for regressing an independent variable on the other variables. The high value of R^2 means that the GI development factor is correlated with other factors.

Pearson's correlation is also used to identify the linear dependence between two variables, X (GI development driving independent factor) and Y (dependent factor GI, GY). Its value ranges between 1 and -1 . It is obtained through the following Equation (2):

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \quad (2)$$

The information gain ratio (InGR) calculates the reduction in entropy from the transformation of the dataset. It can be used for factors selection by evaluating the information gain of each factor in the context of the target factor (GI, GY) [14]. The higher the value of InGR, the greater the significance of the factor is. It is calculated using Equation (3):

$$\text{Gain ration}(x, Z) = \frac{\text{Entropy}(Z) - \sum_1^n \sum_{i=1}^n \frac{|Z_i|}{|Z|} \text{Entropy}(Z_i)}{- \sum_{i=1}^n \frac{|Z_i|}{|Z|} \log \frac{|Z_i|}{|Z|}} \quad (3)$$

where Z is the training dataset with $Z_i = 1, 2, 3, \dots, n$ subset of data.

2.6. The Theoretical Background of Methods Used

The workflow of the proposed methodology for future transformation to GI sites is given in Figure 4. After screening factors, the values of the selected factors for the whole study area and the GI and GY sites were extracted in ArcGIS. These values were then exported in the .csv format. All factor values were normalized before calibrating the models. Next, the SVM, RF, ANN, and stacking model was constructed in the Python environment under two open frameworks, Keras [32] and Scikit-Learn [33], to predict the likelihood of future GI sites. These models were then evaluated based on different statistical measures.

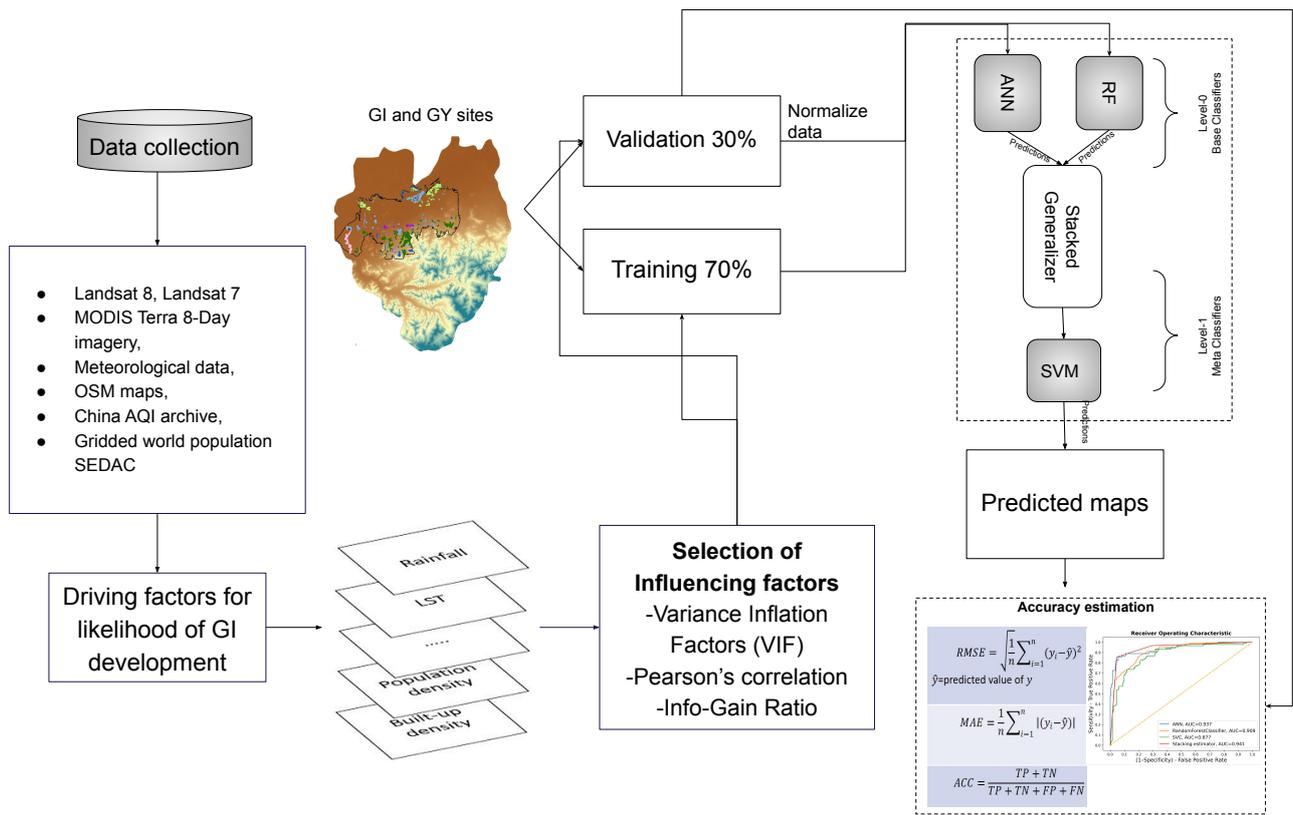


Figure 4. Flowchart of the methodology for sites development into GI/GY in Jinan: Data collection, factors selection, modeling, site prediction, and verification.

2.6.1. Support Vector Machine (SVM)

SVM is a supervised machine learning model that is based on the structural risk minimization concept [34]. In the current study, the SVM model used the GI training datasets to calculate the optimum hyperplane to maximize the distance between data points of two classes, i.e., GI and GY. In this way, the dataset is classified into two classes for the binary classification problem, GI and GY. A separating hyperplane (4) can be defined as

$$\frac{1}{2} \|w\|^2 + C \sum_i \xi_i \cdot s \cdot t \cdot y_i (w \cdot x_i + b) \geq 1 - \xi_i \tag{4}$$

where $\|w\|$ is the norm of the hyperplane, b is the offset/bias of hyperplane, w is the weight vector, ξ_i is the positive slack variables, C is the penalty parameter, and y_i is the classified variable. The kernel function performs data transformations to separate the data based on the defined labels, which is computed as (5):

$$k(x_i, x_j) = (\phi(x_i), \phi(x_j)) \tag{5}$$

where $k(\cdot, \cdot)$ is a kernel function, and x_i is the training vectors. In this study, five-fold cross-validation was used to find optimal parameters for the SVM model, as given in Table 1. The RBF kernel was found to be better than other kernels, as in nonlinear classification, RBF-SVM provides better results [35,36].

2.6.2. Random Forest (RF)

Random forest, introduced by Breiman, is a supervised ensemble learning algorithm that builds a forest from an ensemble of decision trees and is trained by the bagging method for both classification and regression [37]. The RF involves Breiman’s bagging and Ho’s random selection of features. Bagging is an ensemble machine learning method

that improves the accuracy of a weak classifier by creating the number of subset classifiers. The selected parameters of random forest in the current study are given in Table 1. Factors information is passed through each tree of random forest to obtain predictions. First, it resamples the GI training data several times, and then at each resample step, it chooses different random features of the GI. Given the resample and random features' subset, it then estimates a decision tree [14,38]. Finally, it combines the set of estimated decision trees into a single decision tree [38]. Among the subset of randomly selected features, RF uses the Gini-impurity index (6) as a feature selection measure, and also to set the threshold for best splitting. It is given as follows:

$$\sum_{j \neq i} \sum \frac{f(Y_i, T)}{|T|} \frac{f(Y_j, T)}{|T|} \tag{6}$$

where T is the training dataset of driving factors for GI development, and $\frac{f(Y_i, T)}{|T|}$ is the probability of each tree for each class Y_i , either GI or GY.

Table 1. Parameters of machine learning algorithms used for prediction of likelihood of GI transformation.

Model Name	Description of Parameters
SVM	Complexity parameter = 10, gamma = "scale", kernel = radial basis function, tolerance parameter = 0.001, probability = True
RF	n_ = 100, max_depth = 9, split criterion = gini
ANN	Keras sequential model, hidden layers = 6, nodes for each layer = 80, 70, 20, 40, 40, 1, learning rate = 0.004, activation function = relu, sigmoid, optimizer = Adam, kernel initializer = uniform, epochs = 50, validation threshold-20
Stacking model	SVM = meta classifier, ANN-RF = base learners, stacking method = predict_proba, decision function or predict

2.6.3. Artificial Neural Network (ANN)

Artificial neural networks are computing systems inspired by the human neurological network. In this work, we use a multilayer neural network. The basic structure of an artificial neural network consists of three layers; the input layer, a hidden layer, and an output layer [39]. In this study, after model hyperparameter tuning, ANN with six layers was trained with the backpropagation algorithm for GI transformation prediction using the input data of driving factors and ground truth labels of GI and GY datasets. The input layer size is eight, equal to the number of driving factors. Hidden layers information processing yields the labels for GI development likelihood, which is passed to the output layer. The output layer consists of one node, which is 1 in the case of the GI, and 0 for the GY.

The input to each neuron j in the hidden layer is the sum of a weighted input signal x_i ($\sum w_{ji}x_i = net_j$), while w_{ji} is the weight between neuron j in the hidden layer and neuron i in the input layer. The output y_i by neuron passes by activation function $f = \frac{1}{1+e^{-net_j}}$ which is given (7) as

$$y_i = f(\sum w_{ji}x_i) = \frac{1}{1 + e^{-net_j}} \tag{7}$$

Training of ANN with the backpropagation algorithm begins by randomly initializing the network weights. Then, the comparison between predicted values of the outputs and the actual values is performed, and the difference between them is termed an error. The weights w_{ji} and w_{jk} , which connect input layers to hidden layers and hidden layers to output layers, are adjusted by backpropagating the error calculated through the error function [40] (8):

$$E_k = \frac{1}{2}(z_k - t_k)^2 \tag{8}$$

where z_k is the layer's output while t_k is the target value. The differences between predicted and actual values are minimized by updating the layers' weights through a backpropagation algorithm. Weights are adjusted using gradient descent algorithm (9) as

$$\Delta w_{ji} = -\eta \frac{\partial E_k}{\partial w_{ji}} \quad \Delta w_{jk} = -\eta \frac{\partial E_k}{\partial w_{jk}} \quad (9)$$

where η is a learning rate parameter. Finally, the weights (10) are updated as

$$w_{ji} = w_{ji} + \Delta w_{ji} \quad w_{jk} = w_{jk} + \Delta w_{jk} \quad (10)$$

2.6.4. Stacking Model

The stacking ensemble model was first introduced by Wolpert [41]. The ensemble method combines different models to improve the model's performance compared to the individual model. Compared to the other ensemble methods, the stacking method uses metalearning, which combines different types of algorithms [42]. There are two levels in stacking structure, level-0 and level-1. In this study, at level-0 heterogeneous models, RF and ANN were trained, and each of these base learners gave different predictions, which were then forwarded to level-1. On the other hand, SVM is used at level-1 and is called a metalearner because simple models work well for metaclassifier [18,43]. The output of base learners becomes the input for a metalearner as training data.

The stacking method uses a similar idea to k-fold cross-validation to create out-of-sample predictions and capture the distinct regions where each model performs the best. The dataset D consists of $d_i = (x_i, y_i)$ with GI transformation driving factors with x_i and y_i as the final classifications of either GI or GY. The base learner models are denoted as $L_t (t = 1, 2)$, i.e., RF and ANN. The dataset D is divided into two subsets. One of them is used to train the base learners to generate level = 0 classifiers [12] (11):

$$h_t^i = L_t(D - d_i) \forall i = 1, 2, \dots, N; \forall t = 1, 2 \quad (11)$$

$i \in [1, N]$, where N is the total number of datasets. A remaining subset is used on x_i to predict z_{it} (12):

$$z_{it} = f_t^i(x_i) \quad (12)$$

and f_t^i is the generalizer function that is responsible for combining the different model's predictions. f_t^i can either be a generic function, such as average, or a model algorithm. The predictions from level = 0 along with their true classification generate a new dataset $D' = ((z_{it}, z_{it}, \dots, z_{it}), y_i)$. These are then given to a level-1 meta classifier, i.e., SVM in this study. The metaclassifier then combines all predictions from base learners to give final predictions (13) as

$$Y_x = SVM(f_1(x), f_2(x), f_3(x)) \quad (13)$$

2.7. Evaluation and Assessment of Models

The results for the likelihood of transformation to GI are evaluated using several statistical metrics for training and test data. We selected root mean squared error (RMSE), mean absolute error (MAE), accuracy (ACC), sensitivity, specificity, and area under the receiver operating characteristic ROC curve (AUC), and their details are given in Table 2. ROC is a widely known standard criterion for model validation. The ROC curve is constructed by plotting sensitivity on the y -axis and specificity on the x -axis. Sensitivity represents the proportion of correctly classified GI cells, while specificity is the proportion of GY cells correctly classified as GY sites [11]. The value of area under the ROC curve, called AUC, is used to predict the reliability of models. Its value ranges between 0 and 1, where values

nearer to 1 show an efficient performance of the model, while values nearer to 0 show that the model is noninformative [44]. *AUC* (14), *sensitivity*, and *specificity* (15) are given as

$$AUC = \frac{\sum TP + \sum TN}{P + N} \tag{14}$$

$$Sensitivity = \frac{TP}{TP + FN} \qquad Specificity = \frac{TN}{FP + TN} \tag{15}$$

where *P* is the total number of GI pixels, and *N* is the total GY pixels. *TP* and *TN* are true positives and true negatives correctly classified as GI and GY cells. *FP* and *FN* are false positives and false negatives incorrectly classified as GI and GY cells. The ROC curve was used for both training and test datasets to validate the final prediction maps.

Table 2. Statistical index-based evaluation measures used for the assessment of models.

Measure	Equation	Explanation
Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y})^2}$ <i>y</i> is a predicted value of <i>y</i>	Measures the average error performed by the model for predicting the output values.
Mean Absolute Error (MAE)	$MAE = \frac{1}{n} \sum_{i=1}^n y_i - \hat{y} $	Measures the average error in a set of predictions from actual values.
Accuracy (ACC)	$ACC = \frac{TP+TN}{TP+TN+FP+FN}$	The proportion of correctly classified GI and GY pixels.

3. Results and Discussion

3.1. GI and GY Distribution (2000–2010–2020)

The spatial distributions of LULC classes for the years 2000, 2010, and 2020 are shown in Figure 5. There are significant changes in all LULC classes; in this research paper, only GI and GY changes in the central urban area of Jinan are discussed.

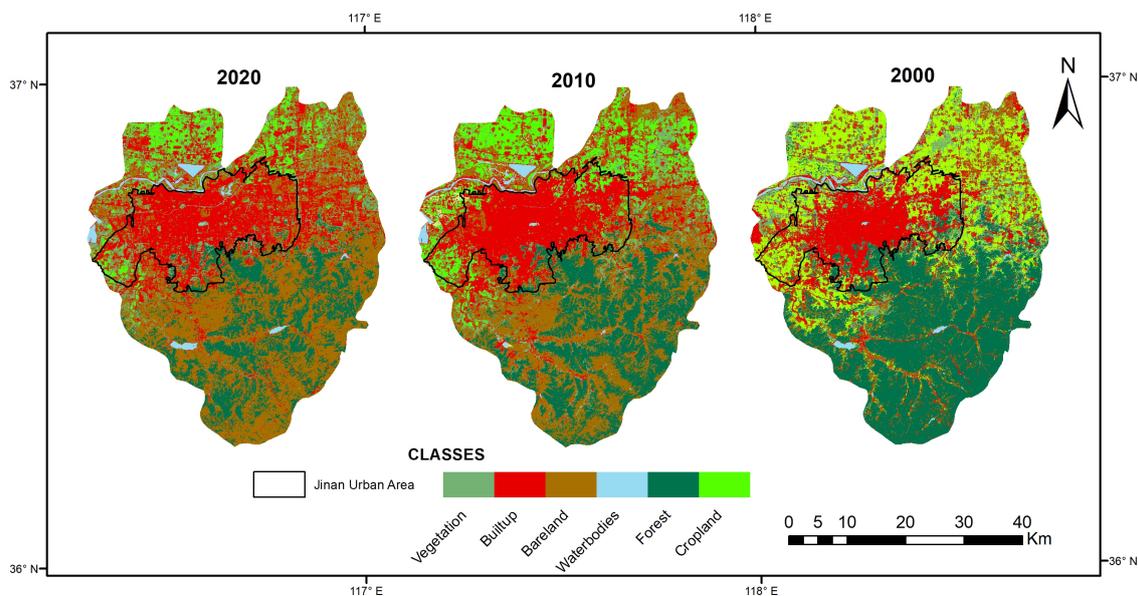


Figure 5. Classified LULC maps of Jinan for year 2000, 2010, and 2020.

For 2000, GY covered an area of 230.41 km², which increased to 267.62 km² in 2020. The major contributor to the GY increase is the cropland in the east and west of the urban area, which decreased significantly. As the south of Jinan is mountains while the north is the Yellow River, urban expansion is mainly toward the east and west. On the other hand,

GI covered an area of 61.36 km² in 2000, which decreased to 35.58 km² in 2010 and then increased to 51.90 km² in 2020. The major contributor to the decrease in GI from 2000 to 2010 is an increase in built-up area and bare land. On the contrary, an increase in GI from 2010 to 2020 is attributed to bare land and cropland in Jinan.

The increase in GI during 2010–2020 is due to the Ecological Garden City Initiative by Jinan Municipal Government from 2010 to 2020. Under this initiative, green corridors were introduced along roads and river networks. Furthermore, tree planting, biodiversity protection planning, and public communication were adopted. Jinan was also a pilot city for The Sponge City Project by the Chinese Government in 2015, which involved the construction of sunken greenbelts, and rain gardens in communities [45].

3.2. GI Network and Vacant Land

Mapped GI of Jinan shows that it is unevenly distributed in the urban area. About 51.90 km² GI is mapped, primarily located in the suburban part of the city. Only a small part of the fragmented GI lies in the central urban area. In this study, only multifunctional GI is considered for further analysis, which covers an area of about 23.35 km² (Figure 6).

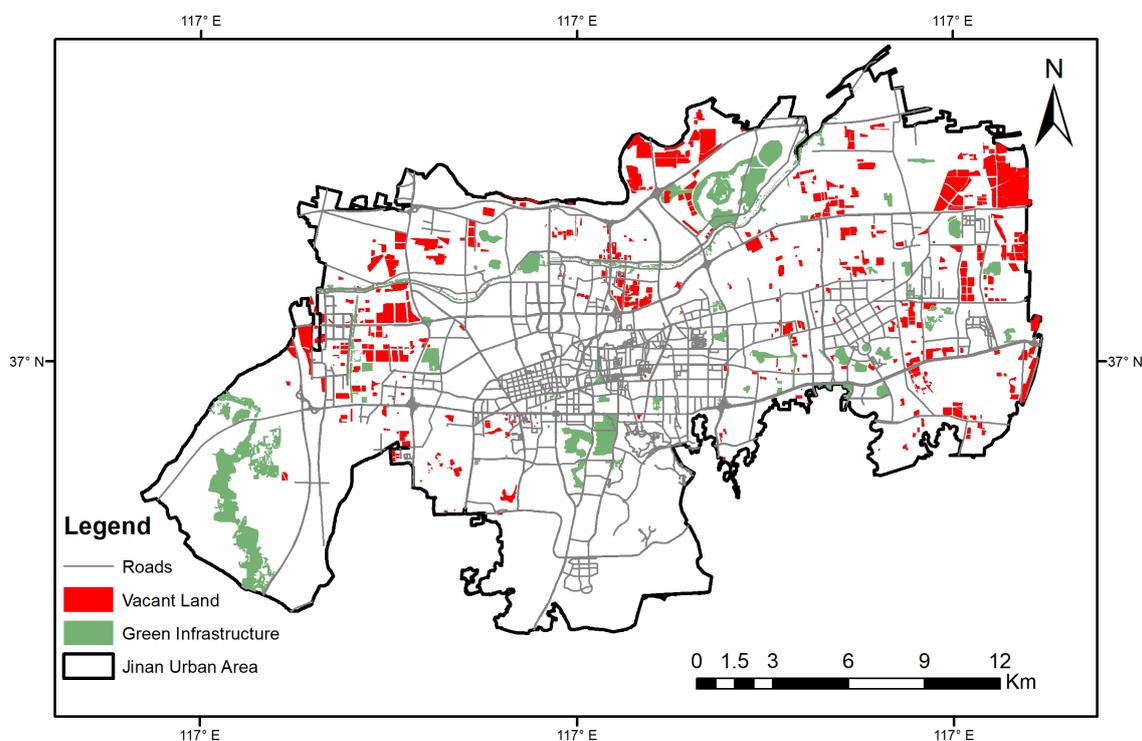


Figure 6. Map of GI and vacant land in Jinan.

Bare land with no vegetation or pioneer vegetation and streams or riverbanks without vegetation is considered as vacant land. A total of 526 vacant land parcels (26.97 km²) were mapped in Jinan (Figure 6) using the LULC change analysis and Baidu map, which were further analyzed for their potential to convert into GI or GY.

3.3. Data Extraction and Factor Importance

The factor values for GI and GY sites were extracted to train, test, and predict the future sites. Rainfall per month for Jinan is given in Figure 2. In the south, with high elevation, and southeastern intensive construction parts of Jinan, rainfall is high (68.54 mm/m), whereas the urban center and northwest of Jinan have comparatively low rainfall (60.43 mm/m). A similar rainfall pattern was observed in other studies [46,47], and it is attributed to the local elevation and urbanization of Jinan. AQI values are high in the urban center compared to the rest of the city (Figure 2). Jinan temperature ranged between 21 °C and 33 °C with

the high temperature in urban center, industrial clusters, and farmland on the eastern side. In contrast, a low temperature was observed in the south of Jinan. A similar trend of temperature was observed in previous studies conducted in Jinan [48,49]. Similar to the AQI, population density, built-up density, and road density are also high in the city center (Figure 3). Due to high built-up density in the city center, vacant land parcels are high in number towards the suburban part of Jinan (Figure 3).

Multicollinearity and InGR were used to determine the correlation and importance of each factor for GI transformation. Results for multicollinearity showed that the minimum VIF value (1.06) was obtained for water bodies buffer, and the maximum value (1.66) was obtained for LST. The VIF values for all factors are less than 5, which ensures the independence of all the factors for potential GI sites prediction variables. The VIF values for all the factors are given in Table 3.

Table 3. Multicollinearity analysis showing VIF values less than five for all the driving factors of GI transformation.

Variables	VIF	Variables	VIF
Population density	1.36	Rainfall	1.53
Vacant land buffer	1.30	Road buffer	1.19
Air quality	1.62	Built-up density	1.61
LST	1.66		
Water buffer	1.06		

Correlation between GI driving factors is also estimated through Pearson’s correlation. There is no significant correlation found between all factors except built-up density and population density, with a correlation of 0.51 (Table 4). Hence, it depicts that all factors can be used for GI transformation analysis.

Table 4. Pearson’s correlation coefficient pairs of factors.

	Rainfall	AQI	LST	Water Buff.	Pop. Density	Road Buff.	Vacant Land Buff.	Built-Up Density
Rainfall	1							
AQI	−0.34	1						
LST	0.28	0.38	1					
Water buff.	0.17	−0.021	0.061	1				
Pop. den.	0.027	0.11	0.1	−0.094	1			
Road buff.	−0.071	−0.13	−0.21	0.064	−0.23	1		
Vacant land buff.	−0.087	−0.24	−0.38	−0.14	0.086	0.015	1	
Built-up density	0.13	0.097	0.13	−0.11	0.51	−0.36	0.21	1

InGR results showed (Figure 7) that the highest value was 0.193, for LST, followed by water buffer (0.191), population density (0.161), and rainfall (0.104). Vacant land buffer (0.066), AQI (0.046), built-up density (0.026), and road buffer (0.018) indicate that they had no significant role in vacant land to GI likelihood transformation.

This result indicates that the LST, distance from water bodies, population density, and rainfall contributes to sites transformation to GI in Jinan. Hence, high temperature areas, sites near water bodies, high population density, and high rainfall sites will likely be converted into GI. Previous studies also recognized the importance of population density, sites near water bodies, and flood zones as the driving force of conversion to GI sites, while built-up density and road density had low importance in GI transformation [6,10].

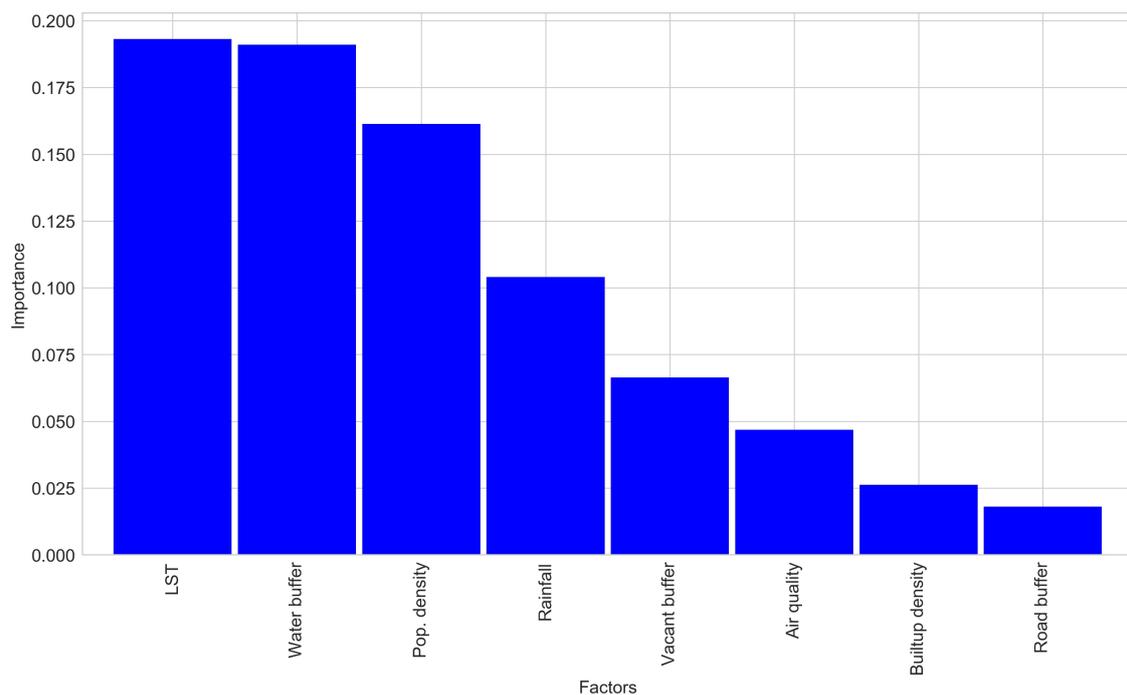


Figure 7. Prediction power of each driving factor influencing GI development, using InGR.

According to the InGR result, LST is the main driving factor behind the GI development in Jinan. Distance from water bodies is the next important factor. There is a strong interdependence between urban GI and urban water bodies for water management and treatment. That is why distance from water bodies is a significant factor for developing GI. Several multifunctional GI sites in Jinan are around water bodies, such as Daming Lake, Spring City Square, Baotu Spring, Five Dragons Pool, and GI along riverbanks. Population density is another factor that contributes to GI development. Jinan's population density is 898 people per square kilometer; the east–west expansion of Jinan also influences its population distribution. A study conducted by Ma [50] found that in the core center of Jinan, population size is positively correlated to the GI present within a 1.2 km distance. However, at local (>1.2 km) and city scale, population size is negatively correlated to the GI distance.

Rainfall distribution also influenced GI development. The average rainfall in the central urban area is 60.43 mm/m, in the low-lying region. Therefore, apart from local rainfall, runoff from surrounding high-elevation regions moves toward lower elevations of the urban center and becomes stagnant points in the case of intense rainfall. That is why many GI projects were initiated previously in the core urban area under The Sponge City project to retain and reduce stormwater in flood risk areas.

Distance of vacant land to existing GI played no significant role in the decision-making of GI development. The central urban area of Jinan is a low-lying area, so it is difficult for air pollutants to transport and diffuse, whereas industries lie in the southwest and northeast part, which are upwind [51]. Therefore, air pollutants enter the city and aggravate the air pollution. However, air pollution was not a driving factor for GI development in Jinan. Likewise, built-up density has less influence on GI investment decisions, indicating that GI development is more likely to occur in open areas with vacant land rather than congested built-up areas. Furthermore, sites in urban centers are more likely to be converted into GY due to economic value [10,52]. In literature, road density and accessibility to green areas are identified as important factors [53]. However, in this study, distance to road network was not a significant driving force to develop GI, which is also found in a study conducted in Jinan [50], indicating that for >1.5 km, accessibility to GI is dispersed.

3.4. Models for GI Likelihood Site Prediction

Four models, i.e., SVM, RF, ANN, and a stacking model, were employed to predict the likelihood of vacant land conversion to GI or GY in Jinan. Using the geometrical interval classification in ArcGIS, the maps were divided into five classes: very high, high, moderate, low, and very low (Figures 8–11). Higher values give a higher likelihood of being transformed into GI sites.

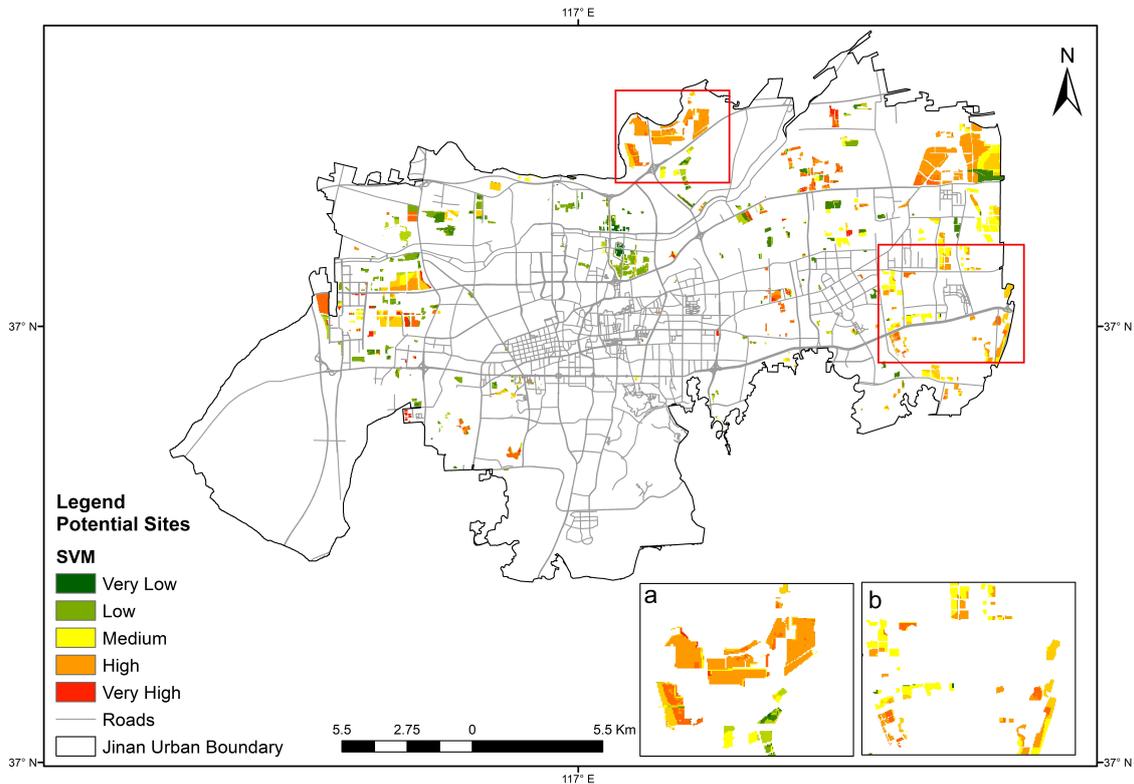


Figure 8. GI development potential sites mapping by SVM model: (bottom right) (a) North and (b) East of Jinan likely to be converted into GI.

Results obtained with SVM are shown in Figure 8. It was observed that only 1.64% of the area falls under very high class, and 9.93% under high class for becoming GI sites. With the RF model, 2.14% of the area fell under very high class to be converted into GI (Figure 9), and 3.58% of the study area for high class. ANN showed a 5.27% area (Figure 10) for the very high class and 1.11% for the high class. Finally, the stacking model gave 5.12% (Figure 11) for the very high class and 0.05% for the high likelihood class.

Three significant regions were identified for all maps, which will likely be converted into GI, and these were grouped as A (north), B (east), and C (west of central urban area) (Figures 8–11). These regions lie within the second ring road of Jinan, which comes under the past and future City Master Plans by Jinan Municipal Government. Furthermore, these regions depict high LST, are near water bodies, have high population density, and high rainfall. Region A is along the Yellow River, which comprises a large area of vacant land to be transformed into riparian vegetation. This region is a pilot area for an ecological corridor beside the Yellow River. Ninety-four projects have been proposed for this region, and 24 have been completed. Region B encompasses the under-construction areas, i.e., the industrial south road and the industrial north road. This area is rapidly being transformed due to urban expansion in this direction. Some of the vacant land identified in this region is likely to be transformed into GI, which will improve the ecological and social services in intensive construction areas. It is practically significant to incorporate GI at the beginning of the project. Region C also falls under the ongoing urbanization such as vacant land near Jing Shi Lu road, Jinan West Railway Station, and Jinan West Expressway.

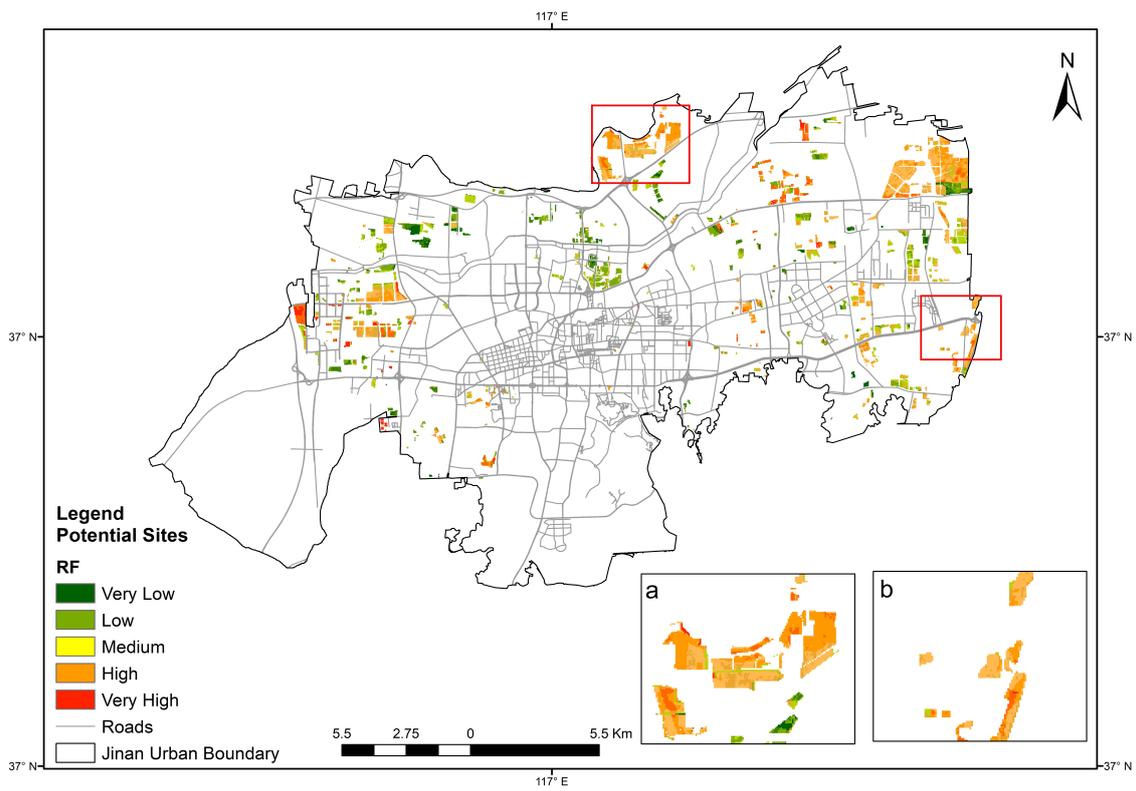


Figure 9. GI development potential sites mapping by RF model: (bottom right) (a) North and (b) East of Jinan likely to be converted into GI.

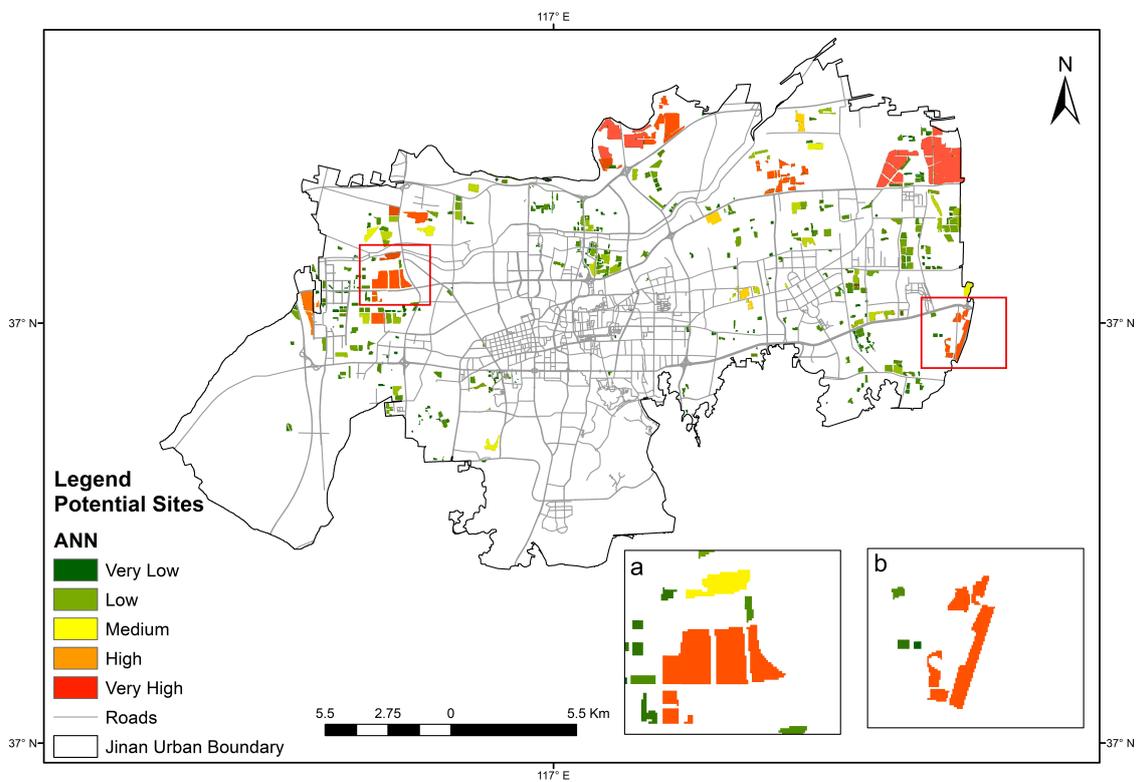


Figure 10. GI development potential sites mapping by ANN model : (bottom right) (a) West and (b) East of Jinan likely to be converted into GI.

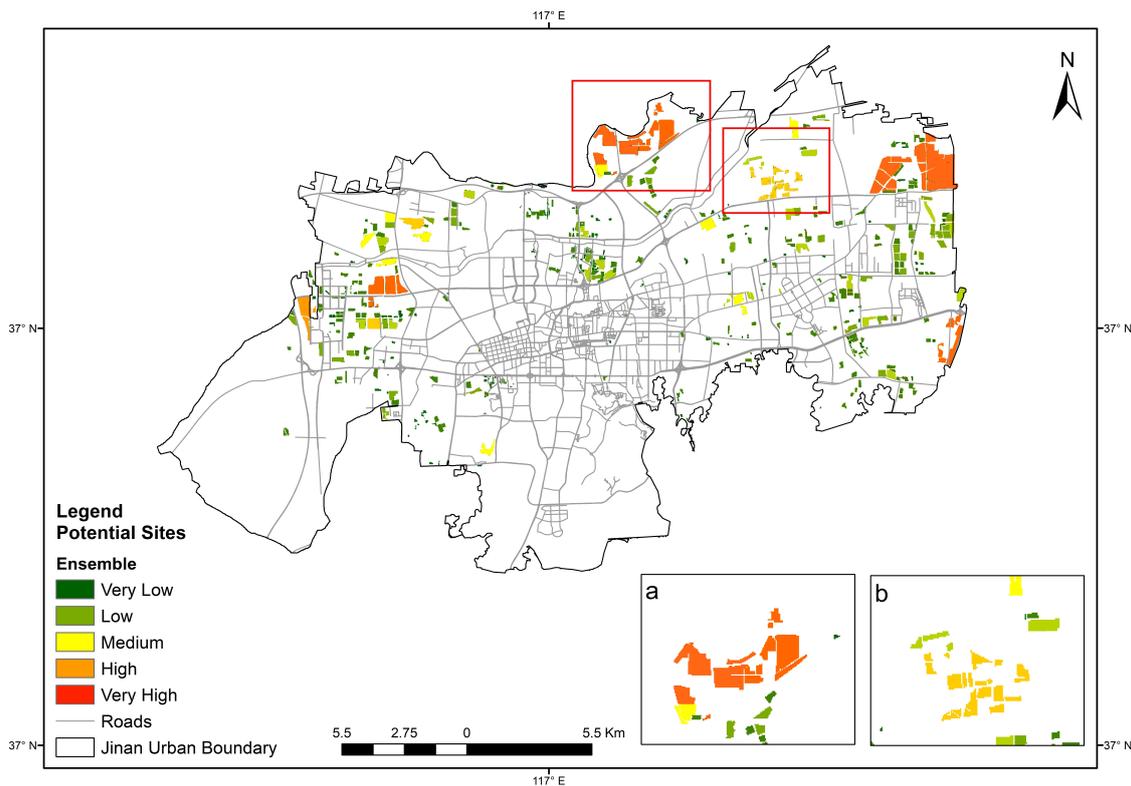


Figure 11. GI development potential sites mapping by stacking model: (bottom right) (a) (b) North of Jinan likely to be converted into GI.

On the other hand, the identified vacant land in the central core area of Jinan is under a very low potential class to be converted into GI because of the densely built-up area having a high demand for GY with a small area of public green belt. Hence, the areas predicted by models are logically reasonable for conversion into GI or GY sites due to the combination of biophysical and socioeconomic conditions in these regions.

3.5. Models Validation

Models were evaluated using statistical measures and ROC curve for training and testing data. Table 5 presents the statistical results of the models.

Table 5. Performance evaluation of models for training and validation datasets using statistical measures.

Methods	RF		SVM		ANN		Stacking	
	Training	Validation	Training	Validation	Training	Validation	Training	Validation
RMSE	0.269	0.314	0.360	0.428	0.187	0.280	0.158	0.273
MAE	0.073	0.099	0.130	0.183	0.059	0.096	0.053	0.013
Accuracy	0.935	0.892	0.871	0.818	0.967	0.913	0.975	0.923
AUC	0.918	0.909	0.879	0.877	0.920	0.937	0.975	0.941
Sensitivity	0.978	0.975	0.989	0.897	0.973	0.964	0.998	0.964
Specificity	0.934	0.847	0.938	0.737	0.898	0.912	0.934	0.891

The RMSE and MAE values in training datasets of the stacking model are (0.158, 0.053), respectively, ANN (0.187, 0.059), RF (0.269, 0.073), and SVM (0.36, 0.13). Results showed that the lowest errors are obtained for the stacking model, then ANN, RF, and SVM. Similar results are obtained for the validation data as well, i.e., stacking model (0.273, 0.013), ANN (0.28, 0.096), RF (0.314, 0.099), and SVM (0.428, 0.183). In terms of accuracy for training and validation data, the stacking model showed the ACC of 0.975 and 0.923, respectively,

indicating that 97% of pixels are correctly classified as GI and GY, while the second-best model is ANN (0.967, 0.913), followed by RF (0.935, 0.892) and SVM (0.871, 0.818).

For the validation of GI site maps produced by four models, AUROC was obtained for both training and validation datasets. The success rate curve of training data is shown in (Figure 12). The AUC value (0.975) for the stacking model is well-fitted compared to other models, followed by ANN (AUC = 0.920), RF (AUC = 0.918), and SVM (AUC = 0.879). The predictive capability of models for the validation dataset is obtained by prediction rate curve (Figure 12), which showed similar results to training data. The stacking model outperformed, with an AUC of (0.941), followed by ANN (AUC = 0.937), RF (AUC = 0.909), and SVM (AUC = 0.877), the same as validation data.

Different statistical measures (RMSE, MAE, ACC, sensitivity, specificity) and ROC curve AUC showed that the newly proposed heterogeneous ensemble stacking model performed the best, followed by the ANN model, RF, and SVM. The stacking model increased the accuracy and AUC values of metaclassifier and base learners in an ensemble compared to individual models. This difference in performance between the stacking ensemble model and individual models is also observed in several studies [11,12,16,18,54,55], where they provided better results than the conventional single machine learning models. However, all the works mentioned above focus on natural disaster prediction using stacking models, while this research uses the stacking model to predict the likelihood of transformation to GI.

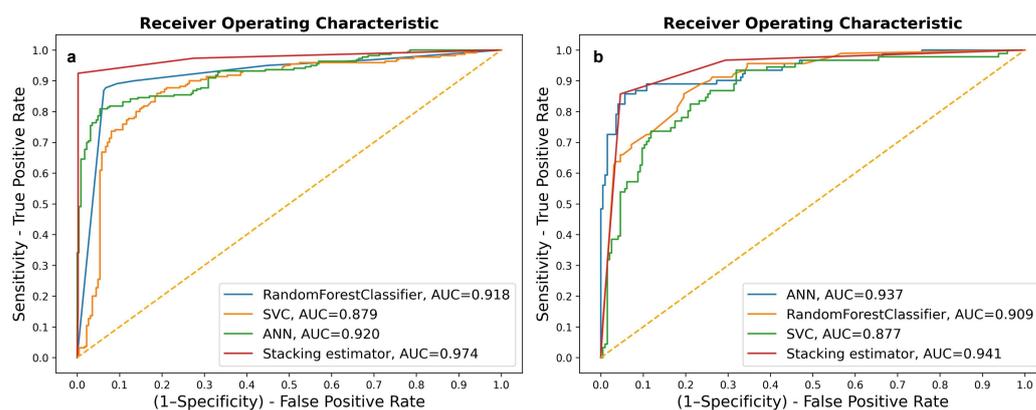


Figure 12. ROC analysis for three models and stacking estimator with AUC values: (a) training dataset; (b) validation dataset.

Due to the availability of large and multisource data and the advancement in AI, the major concern is improving the accuracy of prediction maps. Ensemble approaches proved to be more accurate than the individual models as the two-level framework learns better and the metaclassifiers reduce the bias generated at the first level [12]. The main strength of the stacking model is that it improves the predictive power of the classifiers, and therefore it is necessary to consider the models carefully before selecting them as base learners or metaclassifiers [13]. A powerful model adds positive properties to stacking and other subordinate models, i.e., the overall performance of the stacking model depends on individual classifiers [56]. Consequently, ANN was used as one of the base learners in this study, which increases other models' differentiation index, whereas RF is an ensemble technique with multiple trees. SVM and ANN proved to be the best combination in ensemble model studies conducted by [12,56,57]. An ensemble can also empower the models with lower or moderate performances, such as SVM in this study. SVM does encounter problems in the case of high-dimensional data and complex classes, but with RBF kernel and hyperparameter tuning, it can overcome this problem [58].

Among the applied models (SVM, RF, ANN, stacking) in the current study, only the ANN was used previously for the GI development likelihood study. To the best of the authors' knowledge, only one study was carried out to predict vacant land transformation into GI/GY along waterways and derelict sites in Manchester City Council, utilizing

machine learning models. Their model's verification indicated that ANN performed better (90.3%, 89.3% training, 91.8%, 89% validation) than the ANFIS and LR model [10].

Consequently, this research provides a valuable contribution because it extends the previous research into another study area with different topographical and meteorological conditions and policies. In Jinan, the transformation of vacant land into GI/GY has not been assessed before. Hence, this study will update policymakers and urban planners on what driving factors are responsible for GI/GY development in Jinan and what factors are not involved in the transformation. Similarly, results can also guide where the focus should be shifted to consider the neglected factors. Moreover, it will provide information about vacant sites which are more likely to be converted into GI and GY. Thus, they can make decisions related to those specific sites. On the other hand, understanding the spatial pattern of GI and vacant land in Jinan could help assess and create future planning scenarios.

4. Conclusions

The previous trend of vacant land conversion into GI/GY in Jinan, China, from 2000 to 2020 was examined. Based on this, a new stacking ensemble model (ANN-RF-SVM) and the individual models were used to predict future sites having the higher potential of conversion to GI while considering the ecological and socioeconomic factors. From this study, the following conclusions are drawn:

- Mapped GI from 2000–2010 indicates a decrease in the area, i.e., 61.36 km² to 35.58 km², whereas it increased from 2010–2020, i.e., 51.90 km², due to Jinan Municipal Government policies and projects.
- According to VIF and Pearson correlation values, all factors are independent of each other. InGR values depict the most contributing factors for GI development in Jinan as LST, water bodies distance, population density, and rainfall.
- A heterogeneous ensemble stacking model is designed with ANN-RF as base learners and SVM as metaclassifier along with individual models of SVM, RF, and ANN. The output of models showed that the stacking model outperforms (AUC 0.97428, 0.94069) the base classifier ANN, RF, and SVM, because an ensemble stacking model reduces overfitting and bias and increases the efficiency of base classifiers. Hence, stacking is a promising and reliable way to predict the sites which will likely be converted into GI/GY in any area.
- Due to the black-box nature of machine learning models, current research results are more predictive than explanatory. However, to overcome this, factors influencing the outcome of models are determined, including their degree of influence on prediction.
- Few of the driving factors, such as site size, land price, ownership, and green equity index, are not integrated into this research due to the unavailability of data that could play a significant role in GI development. Therefore, we suggest that economic factors should be included in the decision-making process by urban planners. Furthermore, it is also recommended to add economic factors to the modeling process for future studies to improve model predictions further.

Author Contributions: Conceptualization, Y.W., N.L., J.W. and Q.Y.; methodology, K.G.; formal analysis, K.G.; investigation, K.G.; data curation, K.G.; writing—original draft preparation, K.G.; writing—review and editing, K.G., Y.W., N.L., J.W. and Q.Y.; visualization, K.G.; supervision, Y.W.; project administration, Y.W.; funding acquisition, Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This study is funded by the National Key Research and Development Program (Grant No 2017YFC1502701) during the 13th Five-year Plan, Ministry of Science and Technology, PRC.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Dreiseitl, H.; Wanschura, B. *Strengthening Blue-Green Infrastructure in Our Cities*; Liveable Cities Lab, Ramboll: Copenhagen, Denmark, 2016.
- Elmqvist, T.; Fragkias, M.; Goodness, J.; Güneralp, B.; Marcotullio, P.J.; McDonald, R.I.; Parnell, S.; Schewenius, M.; Sendstad, M.; Seto, K.C.; et al. *Urbanization, Biodiversity and Ecosystem Services: Challenges and Opportunities: A Global Assessment*; Springer Nature: Berlin/Heidelberg, Germany, 2013.
- Kabisch, N.; Korn, H.; Stadler, J.; Bonn, A. *Nature-Based Solutions to Climate Change Adaptation in Urban Areas: Linkages between Science, Policy and Practice*; Springer Nature: Berlin/Heidelberg, Germany, 2017.
- McFarland, A.R.; Larsen, L.; Yeshitela, K.; Engida, A.N.; Love, N.G. Guide for using green infrastructure in urban environments for stormwater management. *Environ. Sci. Water Res. Technol.* **2019**, *5*, 643–659. [[CrossRef](#)]
- Kim, G. Reimagining Vacant Urban Land as Green Infrastructure: Assessing Vacant Urban Land Ecosystem Services and Planning Strategies for the City of Roanoke, Virginia. Ph.D. Thesis, Virginia Tech, Blacksburg, VA, USA, 2015.
- Sanches, P.M.; Pellegrino, P.R.M. Greening potential of derelict and vacant lands in urban areas. *Urban For. Urban Green.* **2016**, *19*, 128–139. [[CrossRef](#)]
- Abebe, M.T.; Megento, T.L. Urban green space development using GIS-based multi-criteria analysis in Addis Ababa metropolis. *Appl. Geomat.* **2017**, *9*, 247–261. [[CrossRef](#)]
- Mishra, V.N.; Rai, P.K. A remote sensing aided multi-layer perceptron-Markov chain analysis for land use and land cover change prediction in Patna district (Bihar), India. *Arab. J. Geosci.* **2016**, *9*, 249. [[CrossRef](#)]
- Triantakoustantis, D.; Mountrakis, G. Urban growth prediction: A review of computational models and human perceptions. *J. Geogr. Inf. Syst.* **2012**, *4*, 26323. [[CrossRef](#)]
- Labib, S. Investigation of the likelihood of green infrastructure (GI) enhancement along linear waterways or on derelict sites (DS) using machine learning. *Environ. Model. Softw.* **2019**, *118*, 146–165. [[CrossRef](#)]
- Islam, A.R.M.T.; Talukdar, S.; Mahato, S.; Kundu, S.; Eibek, K.U.; Pham, Q.B.; Kuriqi, A.; Linh, N.T.T. Flood susceptibility modelling using advanced ensemble machine learning models. *Geosci. Front.* **2021**, *12*, 101075. [[CrossRef](#)]
- Hu, X.; Zhang, H.; Mei, H.; Xiao, D.; Li, Y.; Li, M. Landslide susceptibility mapping using the stacking ensemble machine learning method in Lushui, Southwest China. *Appl. Sci.* **2020**, *10*, 4016. [[CrossRef](#)]
- Dou, J.; Yunus, A.P.; Bui, D.T.; Merghadi, A.; Sahana, M.; Zhu, Z.; Chen, C.W.; Han, Z.; Pham, B.T. Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. *Landslides* **2020**, *17*, 641–658. [[CrossRef](#)]
- Chapi, K.; Singh, V.P.; Shirzadi, A.; Shahabi, H.; Bui, D.T.; Pham, B.T.; Khosravi, K. A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environ. Model. Softw.* **2017**, *95*, 229–245. [[CrossRef](#)]
- Wu, Z.; Zhou, Y.; Wang, H.; Jiang, Z. Depth prediction of urban flood under different rainfall return periods based on deep learning and data warehouse. *Sci. Total Environ.* **2020**, *716*, 137077. [[CrossRef](#)] [[PubMed](#)]
- Arabameri, A.; Saha, S.; Mukherjee, K.; Blaschke, T.; Chen, W.; Ngo, P.T.T.; Band, S.S. Modeling spatial flood using novel ensemble artificial intelligence approaches in northern Iran. *Remote Sens.* **2020**, *12*, 3423. [[CrossRef](#)]
- Kranjčić, N.; Medak, D.; Župan, R.; Rezo, M. Machine learning methods for classification of the green infrastructure in city areas. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 463. [[CrossRef](#)]
- Fang, Z.; Wang, Y.; Peng, L.; Hong, H. A comparative study of heterogeneous ensemble-learning techniques for landslide susceptibility mapping. *Int. J. Geogr. Inf. Sci.* **2021**, *35*, 321–347. [[CrossRef](#)]
- Helian, L.; Shilong, W.; Guanglei, J.; Ling, Z. Changes in land use and ecosystem service values in Jinan, China. *Energy Procedia* **2011**, *5*, 1109–1115. [[CrossRef](#)]
- Cheng, T.; Xu, Z.; Hong, S.; Song, S. Flood risk zoning by using 2D hydrodynamic modeling: A case study in Jinan City. *Math. Probl. Eng.* **2017**, *2017*, 5659197. [[CrossRef](#)]
- Qilu Evening News. Available online: <http://www.cqyy.net/csnews/2021/0702/68903.html> (accessed on 9 December 2021).
- Kong, F.; Nakagoshi, N. Spatial-temporal gradient analysis of urban green spaces in Jinan, China. *Landsc. Urban Plan.* **2006**, *78*, 147–164. [[CrossRef](#)]
- Jinan Municipal Government. Announcement of Jinan Urban Green Space Planning Program (2010–2020). 2015. Available online: http://www.jinan.gov.cn/art/2015/8/22/art_24749_1751284.html (accessed on 29 November 2021).
- Qilu Net-Shandong Radio and TV News. Announcement of Jinan Greening Regulations, Effective on 1 March 2022. 2021. Available online: <https://baijiahao.baidu.com/s?id=1718113634939859799&wfr=spider&for=pc> (accessed on 9 December 2021).
- NASA JPL. NASA Shuttle Radar Topography Mission Global 1 Arc Second [SRTM1N36E116V3, SRTM1N36E117V3]. 2013. Available online: <https://doi.org/10.5067/MEaSURES/SRTM/SRTMGL1.003> (accessed on 28 December 2020).
- TRMM. TRMM (TMPA/3B43) Rainfall Estimate L3 1 Month 0.25 Degree × 0.25 Degree V7. 2011. Available online: <https://doi.org/10.5067/TRMM/TMPA/MONTH/7> (accessed on 30 December 2020).
- Berman, L. China AQI Archive (Feb 2014–Feb 2016). 2011. Available online: <https://doi.org/10.7910/DVN/GHOXXO> (accessed on 30 December 2020).
- Zhengming, W.; Simon, H.; Hulley, G. MOD11A2 MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 1km SIN Grid V006. 2015. Available online: <https://doi.org/10.5067/MODIS/MOD11A2.006> (accessed on 29 December 2020).

29. NASA SEDAC. Gridded Population of the World, Version 4 (GPWv4): Administrative Unit Center Points with Population Estimates, Revision 11. 2018. Available online: <https://doi.org/10.7927/H4BC3WMT> (accessed on 25 December 2020).
30. OpenStreetMap. Jinan LULC 2021. Available online: <https://www.openstreetmap.org> (accessed on 12 February 2021).
31. Sameen, M.I.; Sarkar, R.; Pradhan, B.; Drukpa, D.; Alamri, A.M.; Park, H.J. Landslide spatial modelling using unsupervised factor optimisation and regularised greedy forests. *Comput. Geosci.* **2020**, *134*, 104336. [[CrossRef](#)]
32. O'Malley, T.; Bursztein, E.; Long, J.; Chollet, F.; Jin, H.; Invernizzi, L. Keras Tuner. 2019. Available online: <https://github.com/keras-team/keras-tuner> (accessed on 18 February 2021).
33. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
34. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
35. Pourghasemi, H.R.; Jirandeh, A.G.; Pradhan, B.; Xu, C.; Gokceoglu, C. Landslide susceptibility mapping using support vector machine and GIS at the Golestan Province, Iran. *J. Earth Syst. Sci.* **2013**, *122*, 349–369. [[CrossRef](#)]
36. Pham, B.T.; Pradhan, B.; Bui, D.T.; Prakash, I.; Dholakia, M. A comparative study of different machine learning methods for landslide susceptibility assessment: A case study of Uttarakhand area (India). *Environ. Model. Softw.* **2016**, *84*, 240–250. [[CrossRef](#)]
37. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
38. Dou, J.; Yunus, A.P.; Bui, D.T.; Merghadi, A.; Sahana, M.; Zhu, Z.; Chen, C.W.; Khosravi, K.; Yang, Y.; Pham, B.T. Assessment of advanced random forest and decision tree algorithms for modeling rainfall-induced landslide susceptibility in the Izu-Oshima Volcanic Island, Japan. *Sci. Total Environ.* **2019**, *662*, 332–346. [[CrossRef](#)]
39. Ngo, P.T.T.; Hoang, N.D.; Pradhan, B.; Nguyen, Q.K.; Tran, X.T.; Nguyen, Q.M.; Nguyen, V.N.; Samui, P.; Tien Bui, D. A novel hybrid swarm optimized multilayer neural network for spatial prediction of flash floods in tropical areas using sentinel-1 SAR imagery and geospatial data. *Sensors* **2018**, *18*, 3704. [[CrossRef](#)] [[PubMed](#)]
40. Paul, A.; Das, P. Flood prediction model using artificial neural network. *Int. J. Comput. Appl. Technol. Res.* **2014**, *3*, 473–478. [[CrossRef](#)]
41. Wolpert, D.H. Stacked generalization. *Neural Netw.* **1992**, *5*, 241–259. [[CrossRef](#)]
42. Zhou, Z.H. *Ensemble Methods: Foundations and Algorithms*; CRC Press: Boca Raton, FL, USA, 2012.
43. Witten, I.H.; Frank, E.; Hall, M.A. *Data Mining Practical Machine Learning Tools and Techniques Third Edition*; Morgan Kaufmann: Burlington, MA, USA, 2017.
44. Shafizadeh-Moghadam, H.; Valavi, R.; Shahabi, H.; Chapi, K.; Shirzadi, A. Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping. *J. Environ. Manag.* **2018**, *217*, 1–11. [[CrossRef](#)] [[PubMed](#)]
45. Song, J.; Wang, J.; Xi, G.; Lin, H. Evaluation of stormwater runoff quantity integral management via sponge city construction: A pilot case study of Jinan. *Urban Water J.* **2021**, *18*, 151–162. [[CrossRef](#)]
46. Zhao, Y.; Xia, J.; Xu, Z.; Zou, L.; Qiao, Y.; Li, P. Impact of urban expansion on rain island effect in Jinan city, north China. *Remote Sens.* **2021**, *13*, 2989. [[CrossRef](#)]
47. Chang, X.; Xu, Z.; Zhao, G.; Cheng, T.; Song, S. Spatial and temporal variations of precipitation during 1979–2015 in Jinan City, China. *J. Water Clim. Chang.* **2018**, *9*, 540–554. [[CrossRef](#)]
48. Dong, F.; Chen, J.; Yang, F. A study of land surface temperature retrieval and thermal environment distribution based on landsat-8 in Jinan City. *IOP Conf. Ser. Earth Environ. Sci.* **2018**, *108*, 042008. [[CrossRef](#)]
49. Zhou, X.; Gao, Z. Landscape pattern change analyses of land surface radiation during the city expansion in Jinan City. *Remote Sensing and Modeling of Ecosystems for Sustainability IV. Int. Soc. Opt. Photonics* **2007**, *6679*, 667919.
50. Ma, F. Spatial equity analysis of urban green space based on spatial design network analysis (sDNA): A case study of central Jinan, China. *Sustain. Cities Soc.* **2020**, *60*, 102256. [[CrossRef](#)]
51. Zhang, W.; Zhang, X.; Li, L.; Zhang, Z. Urban forest in Jinan City: Distribution, classification and ecological significance. *Catena* **2007**, *69*, 44–50. [[CrossRef](#)]
52. Longo, A.; Campbell, D. The determinants of brownfields redevelopment in England. *Environ. Resour. Econ.* **2017**, *67*, 261–283. [[CrossRef](#)] [[PubMed](#)]
53. Cetin, M. Using GIS analysis to assess urban green space in terms of accessibility: Case study in Kutahya. *Int. J. Sustain. Dev. World Ecol.* **2015**, *22*, 420–424. [[CrossRef](#)]
54. Choubin, B.; Moradi, E.; Golshan, M.; Adamowski, J.; Sajedi-Hosseini, F.; Mosavi, A. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. *Sci. Total Environ.* **2019**, *651*, 2087–2096. [[CrossRef](#)]
55. Mojaddadi, H.; Pradhan, B.; Nampak, H.; Ahmad, N.; Ghazali, A.H.b. Ensemble machine-learning-based geospatial approach for flood risk assessment using multi-sensor remote-sensing data and GIS. *Geomat. Nat. Hazards Risk* **2017**, *8*, 1080–1102. [[CrossRef](#)]
56. Pourghasemi, H.R.; Yousefi, S.; Kornejady, A.; Cerdà, A. Performance assessment of individual and ensemble data-mining techniques for gully erosion modeling. *Sci. Total Environ.* **2017**, *609*, 764–775. [[CrossRef](#)]
57. Chen, W.; Pourghasemi, H.R.; Kornejady, A.; Zhang, N. Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques. *Geoderma* **2017**, *305*, 314–327. [[CrossRef](#)]
58. Svoray, T.; Michailov, E.; Cohen, A.; Rokah, L.; Sturm, A. Predicting gully initiation: Comparing data mining techniques, analytical hierarchy processes and the topographic threshold. *Earth Surf. Process. Landforms* **2012**, *37*, 607–619. [[CrossRef](#)]