sustainability

MDPI

Article

Research on the Driving Factors and Prediction Model of Urban Underground Space Demand in China

Yansheng Deng [1,2,*], Jun Chen [1], Baoping Zou [1,2,*], Qizhi Chen [1,2], Jingyuan Ma [1,2] and Chenjie Shen [1]

[1] School of Civil Engineering and Architecture, Zhejiang University of Science and Technology, Hangzhou 310023, China
[2] Zhejiang-Singapore Joint Laboratory for Urban Renewal and Future City, Hangzhou 310023, China
* Correspondence: ysdeng@zust.edu.cn (Y.D.); zoubp@zust.edu.cn (B.Z.); Tel.: +86-0571-85070521 (Y.D.)

Abstract: The development and utilization of urban underground space (UUS) have emerged as critical strategies to address the challenges posed by urban population growth and land resource depletion. Accurate prediction of UUS demand serves as the cornerstone for scientifically planning underground space and promoting sustainable urban development. In this study, statistical analysis methods were used to investigate the relationship between potential driving factors and UUS demand based on collected data from 16 cities in China. The identification of primary driving factors involves correlation, path, and determination coefficient analyses. Subsequently, univariate regression, multiple linear regression, and LASSO regression methods are employed to construct prediction models for UUS demand. Additionally, the link between historical data and UUS demand in each city was studied separately. The findings reveal that GDP per km^2 and GDP per capita comprehensively capture the influence of urban population, economy, and transportation on UUS demand. Notably, GDP per km^2 makes the most significant contribution to the proposed regression models, followed by GDP per capita. The application of LASSO regression proves effective in selecting potential factors while maximizing data utilization, presenting itself as a valuable auxiliary tool for UUS planning.

Keywords: UUS; driving factor; UUS demand; prediction model

check for updates

Citation: Deng, Y.; Chen, J.; Zou, B.; Chen, Q.; Ma, J.; Shen, C. Research on the Driving Factors and Prediction Model of Urban Underground Space Demand in China. *Sustainability* **2024**, *16*, 3700. https://doi.org/10.3390/su16093700

Academic Editor: Boris A. Portnov

Received: 23 March 2024
Revised: 20 April 2024
Accepted: 24 April 2024
Published: 28 April 2024

1. Introduction

In recent decades, China's urbanization level has shown steady improvement, reaching 65.22% in 2022, which is an annual increase of 0.5% points, approximately eight years ahead of the earlier predictions by the National Bureau of Statistics [1]. Concurrently, the urban permanent population has surged to 920.71 million, marking a rise of 6.46 million compared to the end of 2021. The rapid pace of urbanization and the resultant growth in the urban population inevitably lead to the depletion of urban land, posing challenges to the sustainable development of Chinese cities [2–4]. Furthermore, this trend may exacerbate various urban issues, including traffic congestion, environmental pollution, land resource scarcity, and ecological degradation. As an invaluable spatial resource, urban underground space (UUS) is assuming an increasingly pivotal role in the development of Chinese cities, particularly in major urban centers [5,6]. The development of UUS not only expands the available urban land resources and alleviates traffic congestion but also contributes to enhancing the urban environment and mitigating the impact of urban disasters [7]. Recognizing these advantages, UUS is garnering more attention as a potential emerging resource for development and utilization [7–12]. Many international metropolises, including Shanghai, Hong Kong, Singapore, Tokyo, and Mexico City, are actively engaging in research and exploration related to the development and utilization of UUS.

Undoubtedly, the development of UUS brings about numerous advantages, catering to the urban demands for sustainable development. Simultaneously, it faces various challenges, such as the lack of scientific and reasonable planning and the difficulty in determining the desirable demand for UUS, among others [4,13]. Furthermore, it is challenging to renovate or rebuild once the UUS has been built for some specific function, which leads to strong irreversibility [3,14]. Therefore, the scientific development of UUS emerges as a crucial approach to resolving the contradiction between population growth and land resource depletion in the process of urban development. In comparison to developed countries, the development and utilization of UUS in China, although initiated relatively late, has experienced rapid growth. As of the end of 2020, the total UUS area in China has reached 2.4 billion square meters. Due to different social, economic, environmental, and policy backgrounds, the development speed and scale of UUS vary across different cities. Therefore, the research objective in this work is to analyze the driving factors of UUS demand for cities in China and build a prediction model to provide a reference for decision-makers.

By collecting comprehensive data on developed UUS demand and potential driving factors in different cities, this research employs descriptive statistics, correlation analysis, and path analysis to identify the inherent driving factors of UUS development demand. Subsequently, a series of regression models are proposed and analyzed, with primary driving factors as independent variables and UUS demand as the dependent variable, including univariate optimal models, multiple linear regression models, and LASSO regression models. Additionally, the correlation between UUS demand and primary driving factors in each city is investigated and discussed separately. Hence, this work presents a series of prediction models to forecast UUS demand, offering valuable insights for UUS planning and sustainable development decision making.

2. Literature Review

Integrating UUS into urban master planning is considered as an ideal approach to develop underground space resources and achieve sustainable development [15]. However, in the UUS planning process, determining the amount of UUS remains a critical data requirement, lacking a unified characterization and prediction method [4,16]. Typically, the functions of each segment of UUS are initially planned and their volumes are separately determined before being summed up as the overall demand for UUS [14,17]. While this method ensures more accurate UUS calculations, it can be relatively cumbersome. Therefore, proposing metrics such as the amount of UUS per capita and per land area (the ratio of UUS area to permanent population and land area, respectively) have been suggested to represent UUS demand. These metrics minimize potential biases in proportions and eliminate inconsistencies that may arise when employing multiple indices [5,14,17,18]. Based on the improved indicators to characterize the UUS demand, the calculation method of UUS demand can be categorized into two main approaches, each considering various factors. The first approach focuses on the relationship between UUS and aboveground development. Some scholars treat UUS as a supplement to aboveground space, determining UUS capacity based on future population demographics and the eco-city model [5,19]. However, maintaining synchronous and coordinated development between aboveground and underground space is crucial for ensuring sustainable urban development, often overlooked. Alternatively, other scholars view UUS as a driving force for urban development, introducing a ratio of underground to aboveground space development to calculate UUS demand. This is primarily estimated through case comparison analysis and the expert scoring method [20]. These two methods, while valuable, rely heavily on specific cases and expert experience, carrying a degree of one-sidedness and subjectivity.

The second approach seeks to build a prediction model for UUS demand based on the relationship between the amount of UUS and potential influencing factors. For instance, He et al. [14] considered population density, annual gross domestic product (GDP) per capita, and real estate prices as influencing factors, proposing a predictive model for UUS demand in various districts in Shanghai using the multivariate regression method. Other factors, such as spatial location, flow area ratio, car ownership per 100 people (100 times the ratio of the total number of registered vehicles in a city to its permanent population), land use type, and land price, have also been applied to predict UUS demand [20–26].

The factors and methods adopted for UUS development demand in previous research are summarized in Table 1. Despite the multitude of factors and methods related to UUS amounts, some are challenging to quantify for prediction, particularly those associated with energy conservation or sustainable development factors.

Table 1. Factors and methods for UUS demand in previous research.

Indicator	Method	References
Land price Land use type Flow area ratio Rail transit passenger flow Accessibility of rail transit	Regression model	[21]
Population density Annual GDP per capita Real estate price	Regression model	[14]
Land use type Ground building FAR[2]	Comparative analysis	[20]
Population density GDP per capita Spatial location vehicle possessive quantity per 100-person GDP per km^2 Land price	Regression model	[22]
GDP per capita Population density Real estate price	Regression model	[23]
GDP per capita Car ownership per 100 people Average price of land sold Tertiary industry's regional GDP	Regression model Grey neural network model	[24,25]
Accessibility Traffic congestion Land use type Location condition Building density House price level Population density Night light intensity	Geodetector model	[26]

3. Methodology

3.1. Potential Driving Factors

Building on the insights from the previous section, it is evident that numerous factors influence the amount of UUS. However, a unified factor to predict UUS demand remains elusive. Drawing from existing studies, the selection of influencing factors for building a prediction model of UUS demand is approached from three perspectives: correlation, representativeness, and data accessibility. Based on these considerations, nine factors are identified as potential driving factors for UUS demand. These factors include population size, population density, GDP per capita, GDP per km^2, vehicle possession quantity per 100 persons, land price, monthly mean temperature, monthly mean maximum temperature, and monthly mean minimum temperature. The potential driving factors and their symbol, respective units are detailed in Table 2.

Table 2. Potential driving factors and units.

Factor	Symbol	Units	Units Symbol
Population size	F1	Person	person
Population density	F2	Person per square kilometers	person/km^2
GDP per capita	F3	Yuan	Yuan
GDP per km^2	F4	Yuan per square kilometers	Yuan/km^2
Vehicle possessive quantity per 100 persons	F5	Vehicle	-
Land price	F6	Yuan per square meter	Yuan/m^2
Monthly mean temperature	F7	Centigrade	°C
Monthly mean maximum temperature	F8	Centigrade	°C
Monthly mean minimum temperature	F9	Centigrade	°C

Population size and density offer insights into the degree of land congestion in a city. GDP per capita, GDP per km^2, and land price reflect the level of urban economic development, serving as the foundational basis for UUS development. Vehicle possession quantity per 100 persons provides an indication of urban traffic conditions. With the rise of urban temperatures, the demand for energy increases, especially for electricity used in air conditioning. This requires urban planners and architects to adopt sustainable design principles such as green buildings, optimized urban planning, and improved energy efficiency to reduce energy consumption. Developing and utilizing urban underground space development can, to some extent, reduce energy consumption and promote sustainable development of the city. To comprehensively reflect the impact of temperature on UUS demand from a sustainable perspective, this study also considers monthly mean temperature, monthly mean maximum temperature, and monthly mean minimum temperature as potential factors.

3.2. Data Acquisition Sources and Methods

A comprehensive dataset comprising 69 sets of data was meticulously collected from 16 cities in China spanning the years 2002 to 2020. The cities included in the dataset are Beijing, Shanghai, Nanjing, Hangzhou, Xuzhou, Guangzhou, Shenyang, Ningbo, Zhengzhou, Nanchang, Wuxi, Qingdao, Wenzhou, Fuzhou, Chengdu, and Taizhou, as illustrated in Figure 1. The data sources for F1, F2, F3, and F5 are extracted from the statistical yearbooks of the corresponding cities and respective years. To eliminate biases stemming from variations in urban areas, F4 is introduced to characterize the impact of GDP per unit urban area, calculated as the GDP of a city divided by its total administrative area. Both GDP and total administrative area data are sourced from statistical yearbooks. F6 data are obtained from the China Land Price Information Service Platform (www.landvalue.com.cn (accessed on 4 February 2024)). Temperature data for F7, F8, and F9 are procured from the China Meteorological Data Service Centre (https://data.cma.cn/en (4 February 2024)). The UUS area data for different cities were extracted from the published literature, as detailed in the Appendix A.

Figure 1. Distribution of Chinese cities for data collection.

Given the substantial variations in land area among cities in China, solely analyzing UUS demand may lead to biased results. Consequently, to account for the relatively small changes in each city, the UUS area per km^2 (Y) was introduced to characterize the demand for UUS. This can be expressed as

$$Y_{\mathrm{UUS}} = \frac{A_{\mathrm{UUS}}}{A_{\mathrm{LA}}} \tag{1}$$

where Y_{UUS} is the UUS area per km^2, A_{UUS} is the developed UUS area, m^2, and A_{LA} is the total administrative area of a city, m^2.

3.3. Flowchart of This Work and Analysis Method Adopted

To examine the relationship between potential driving factors and UUS demand, a series of statistical analysis methods were employed, encompassing descriptive statistics, correlation analysis, path analysis, and regressive analysis. In this work, the descriptive statistics are used to analyze the statistical characteristics of different indicator data, like maximum (Max), minimum (Min), mean, standard deviation (SD), and coefficient of variation (CV). Because constructing regression model requires independent variables to be uncorrelated, correlation analysis is performed to study the correlation between indicators. The path analysis is to find the main factors for the regression model. The flowchart of this analytical process is illustrated in Figure 2.

(1) Descriptive Statistics: Descriptive statistics were conducted on various potential influencing factors, encompassing Max, Min, mean, SD, and CV.

(2) Correlation Analysis: Correlation coefficients and significance levels (P) were computed and analyzed to explore the correlation between factors and UUS demand.

(3) Path Analysis: Through standardizing the regression coefficients, the path coefficient of each variable to the dependent variable can be obtained, as indicated by the following formula:

$$P_{y,x_i} = b_i \frac{\sigma_{x_i}}{\sigma_y} \tag{2}$$

where $P_{y,xi}$ is path coefficient of the i-th independent variable, b_i is the regression coefficient of the i-th independent variable, σ_{xi} is the SD of the i-th independent variable, and σ_y is the SD of the dependent variable.

Utilizing the correlation coefficient between independent variables and the path coefficient to the dependent variable, the indirect path coefficients of each independent variable to the dependent variable can be expressed as

$$P_{x_i x_j} = r_{ij} \frac{\sigma_{x_i}}{\sigma_y} P_{j,y} \ (i \neq j) \tag{3}$$

where P_{xixj} is indirect path coefficient and r_{ij} is the correlation coefficient between two factors. $P_{j,y}$ is the path coefficient.

(4) Regressive Analysis: The univariate optimal model, multiple linear regression model, and LASSO regression model were employed to establish the relationship between potential factors and UUS demand. A multiple regression model was formulated for UUS demand, incorporating influencing factors as independent variables and UUS demand as the dependent variable. This model can be expressed as

$$y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_i x_i \tag{4}$$

where y is the dependent variable (UUS demand), b_0 is the intercept, x_i is the independent variable, and b_i is regression coefficient.

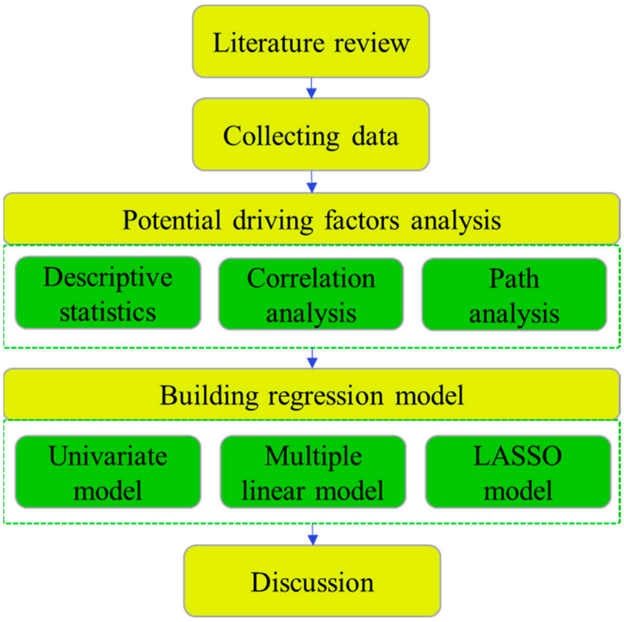

Figure 2. The flowchart of this work.

4. Results

4.1. Descriptive Statistics

The descriptive statistical results of UUS demand and influencing factors are presented in Table 3. Notably, F4 exhibits the highest coefficient of variation (CV), reaching 96.76%, whereas F8 has the minimum CV at 8.13%. The CV of UUS demand is recorded at 91.87% and, for all other factors, the CV exceeds 40%. It can be seen that, apart from temperature related factors, other factors have significant discreteness, among which the GDP per km^2 of different cities has the greatest discreteness. These results emphasize substantial variations in UUS demand and influencing factors due to differences in cities and years.

Table 3. Statistics of factors used for UUS.

Factors	Max	Min	Mean	SD	CV (%)
F1	2.48	0.51	1.31	0.64	48.81
F2	3.91	0.52	1.60	1.19	74.13
F3	15.86	0.60	8.50	3.56	41.93
F4	5.99	0.20	1.41	1.37	96.76
F5	30.75	4.20	17.59	7.07	40.19
F6	3.37	0.20	1.03	0.83	80.02
F7	22.53	7.43	16.69	2.02	12.07
F8	26.64	13.18	21.06	1.71	8.13
F9	19.63	1.92	13.09	2.53	19.33
Y	1.88	0.02	0.47	0.44	91.87

4.2. Correlation Analysis

Figure 3 plots the correlation and significance relationships between various dependent and independent variables. The dependent variable y is positively correlated with all independent variables, exhibiting the highest correlation coefficient with F4 (0.93) and the lowest correlation coefficient with F8 (0.065). The correlation between y and F1, F2, F3, F4, and F6 is extremely significant ($p < 0.01$), significant with F5 ($p < 0.05$), and not significant with F7, F8, and F9. The order of the correlation coefficients between each factor and y is F4 > F6 > F2 > F3 > F1 > F5 > F7 > F9 > F8. With the exception of F1 and F5, F1 and F8, and F2 and F5, all other factors show a positive correlation. There is a high correlation between the monthly mean, mean maximum, and mean minimum temperature. The correlation coefficient between F7 and F9 is the highest (0.99), followed by F7 and F8 (0.96) and F8 and F9 (0.93). Additionally, the correlation coefficients of F1 and F2, F2 and F4, and F3 and F5 are all greater than 0.8, while F1 and F4 as well as F3 and F6 are all greater than 0.7. Thus, GDP per km^2 (F4) has a strong correlation with population size (F1) and population density (F2) and GDP per capita (F3) has an extremely strong correlation with Vehicle possessive quantity per 100 persons (F5), along with a strong correlation with F6. In this way, the GDP per km^2 of a city has a strong correlation with population factors. It is evident that the nine potential factors exhibit a fairly complex correlation.

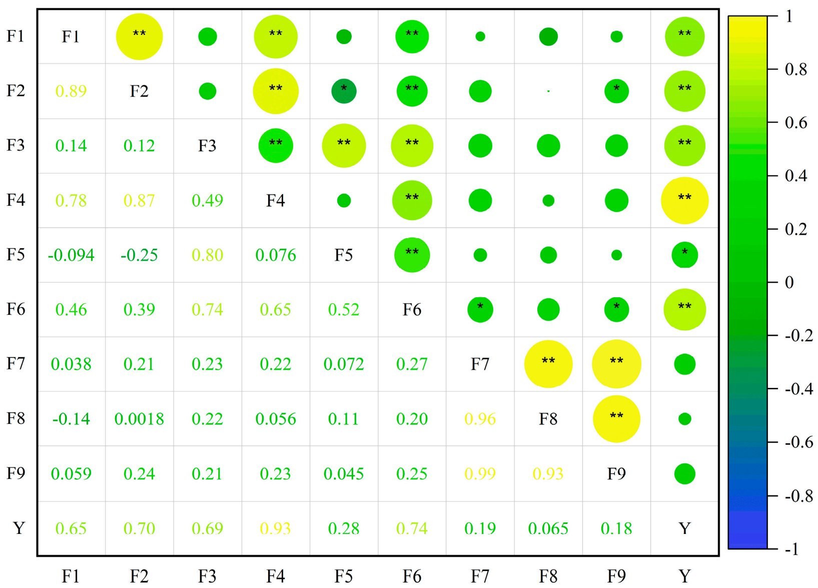

Figure 3. Correlation analysis of various factors and UUS demand (* means significantly different ($p < 0.05$) and ** means very significantly different ($p < 0.01$)).

Due to the complex correlation between potential factors, it is necessary to calculate the variance inflation factor (*VIF*) to check multicollinearity, which can be expressed as

$$VIF = \frac{1}{1 - R^2} \tag{5}$$

where R^2 is the determination coefficients. In Figure 4, the *VIF* values of potential factors are depicted. It is evident that the temperature-related factors (F7, F8, and F9) exhibit intense multicollinearity, followed by population density (F2) and GDP per km^2 (F4). Consequently, feature selection becomes imperative for constructing a more accurate prediction model.

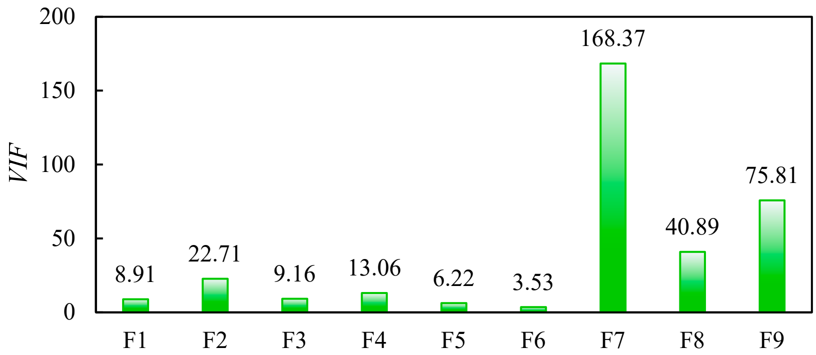

Figure 4. *VIF* value of potential factors.

4.3. Path Analysis

Considering both significance and multicollinearity, F3, F4, and F7, all with significant levels, are selected for path analysis, as illustrated in Figure 5. The corresponding results are presented in Table 4.

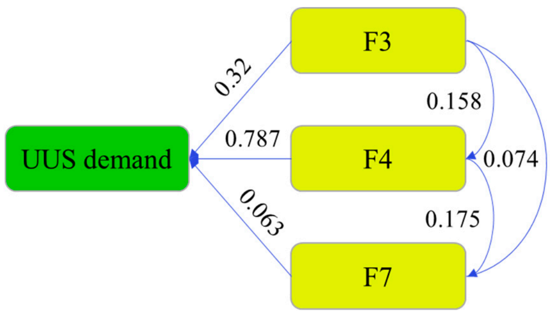

Figure 5. Path analysis of UUS demand.

Table 4. Direct and indirect path coefficients.

Factors	Correlation Coefficient	Direct Path Coefficient	Indirect Path Coefficient			
			F3	F4	F7	Total
F3	0.695	0.320 **	–	0.158	0.074	0.232
F4	0.931	0.787 **	0.390	–	0.175	0.564
F7	0.186	0.063 *	0.015	0.014	–	0.029

Note: * means significantly different ($p < 0.05$) and ** means very significantly different ($p < 0.01$)

Table 4 reveals that the direct path coefficients of F3, F4, and F7 on UUS demand are all positive. Both F3 and F4 reach extremely significant levels, while F7 is significant. F4 possesses the highest direct path coefficient for UUS demand (0.787), signifying the greatest direct effect. Following closely is F3 (0.32), while F7 exhibits the smallest (0.063).

Similarly, the indirect path coefficients of the three factors on UUS demand are all positive, with F4 still exerting the greatest indirect effect, while F7 has the smallest. This suggests that urban temperature is not a major driving factor for the development of UUS, aligning with previous research [27]. Furthermore, the direct path coefficients of the three factors are greater than the indirect coefficients, indicating that these three factors primarily influence UUS demand through direct effects.

4.4. Determination Coefficient Analysis

Table 5 provides the determination coefficients of the three factors on UUS demand. The diagonal showcases the coefficient of determination for F3, F4, and F7, while the upper right corner of the diagonal represents the joint coefficient of determination for two factors on UUS demand.

Table 5. Determination coefficients of three factors to the UUS demand.

Factors	**Coefficient of Determination**			**Total**	**Surplus Factor**
	F3	**F4**	**F7**		
F3	0.482	0.939	0.483	0.943	0.333
F4	–	0.867	0.867		
F7	–	–	0.035		

The coefficient of determination for UUS demand by F4 is notably high at 0.867, surpassing other factors, followed by F3 (0.482). F7 exhibits the smallest coefficient of determination, standing at 0.035. The joint determination coefficient of F3 and F4 is the highest, reaching 0.939, while the value for F3 and F7 is the lowest, at only 0.483. The total coefficient of determination for UUS demand considering all three factors is 0.943, indicating that UUS demand is predominantly determined by the combined efforts of the three mentioned factors.

Based on this, the surplus factor (e) can be calculated using Equation (6), resulting in a value of 0.333. Therefore, F3, F4, and F7 can effectively describe UUS demand, although there might be other factors not considered in this analysis.

$$e = \sqrt{1 - R^2} \tag{6}$$

4.5. Regression Analysis

Considering the findings from Sections 4.3 and 4.4, where the correlation coefficient, path coefficient, and determination coefficient are all relatively small, F3 and F4 are chosen as the primary factors for constructing both the univariable optimal model and the multiple linear regression model.

- Univariable optimal model

Single-variable curve fitting was conducted with F3 and F4 as independent variables and Y as the dependent variable. Five types of curve models were introduced, including the linear model, logarithmic model, quadratic function, power function, and exponential function. The optimal model was determined based on R^2 and F, as presented in Table 6.

Table 6. The optimal model of UUS demand.

Factors	**Optimal Model**	**Model Summary**			**Parameter Evaluation**	
		R^2	***F*-Stat**	P	**Constant**	b_1
F3	Exponential model	0.58	92.496	0	0.044	0.227
F4	Linear model	0.867	435.657	0	0.055	0.297

The optimal model for F3 is the exponential function, while for F4, it is the linear function model. The F4 model exhibits the highest R^2 and F, with values of 0.867 and 435.657. Following closely is the F3 model, with R^2 and F values of 0.58 and 92.496. Moreover, only the models of F3 and F4 demonstrated extremely significant levels.

- Multiple linear regression

Similarly, using F3 and F4 as independent variables and Y as the dependent variable, stepwise regression was employed for regression analysis. The stepwise regression process is outlined in Table 7. It is evident that the multiple correlation coefficient, R^2, and adjusted R^2 gradually increase, while the standard error decreases with the inclusion of predictive variables. Based on regression coefficients and analysis of variance, the partial regression coefficients of the two independent variables and the regression relationships demonstrated a highly significant level ($p < 0.01$). In the end, a multiple linear regression equation can be constructed as

$$y = 0.038F_3 + 0.248F_4 - 0.198 \tag{7}$$

Table 7. The stepwise regression processes.

Model	Multiple Correlation Coefficient	R^2	Adjusted R^2	Standard Error	Predictive Variables
1	0.931	0.867	0.865	0.16	constant, F4
2	0.939	0.939	0.937	0.109	constant, F4, F3

- LASSO regression

LASSO regression is an effective method for handling multicollinear data, utilizing an L1-norm penalty term to shrink linear regression coefficients, achieving both feature selection and penalty regulation [28]. LASSO aims to minimize the following loss function:

$$J(\beta) = \frac{1}{N}\sum_{i=1}^{N}(y_i - \sum_{j=1}^{P} x_{ij}\beta_j)^2 + \lambda \parallel \beta \parallel_1 \tag{8}$$

where β is the coefficient matrix, N is the number of groups of data, y_i is the independent variable, x_{ij} is the dependent variable, P is the number of independent variables, and λ determines the strength of the penalty, influencing the degree of sparsity in the coefficients. $| \; | \; | \; |_1$ is the L1-norm. The 10-fold cross-validation method was employed to determine regression model coefficients, following the following steps:

(1) Randomly shuffle the sample data;
(2) Divide the randomly shuffled sample data into 10 segments;
(3) Select any nine segments of data to train the model and use the remaining one segment of data to validate the model;
(4) Repeat step (3) and determine the lambda value with the smallest mean-squared error (MSE).

Figure 6 presents the cross-validation for the LASSO regression model. The red dashed line represents the lambda value (λ = 0.00281) with the minimum MSE in Figure 6a and the corresponding coefficients for the LASSO model in Figure 6b. From Figure 6b, F3, F4, and F6 were the only selected independent variables with coefficients of 0.033, 0.236, and 0.003, respectively. The coefficients of all remaining potential factors are zero. Therefore, the forecast model using the LASSO variable selection method can be expressed as

$$y = 0.033F_3 + 0.236F_4 + 0.003F_6 - 0.143 \tag{9}$$

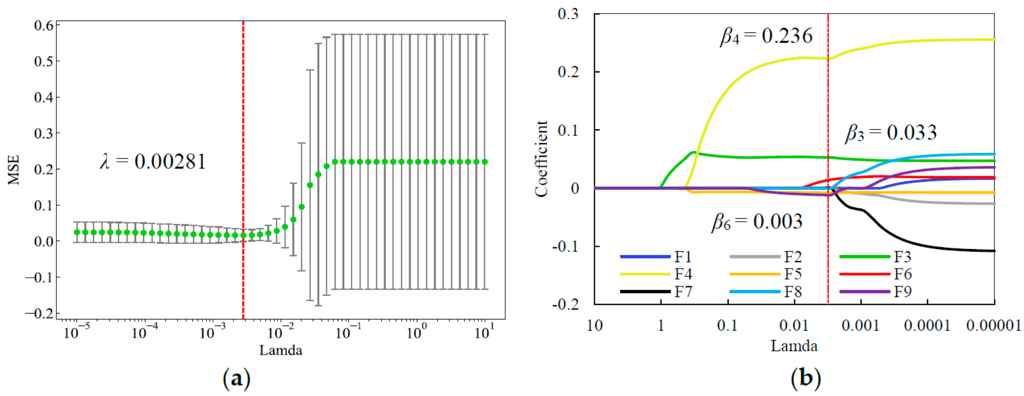

Figure 6. The cross-validation for the Lasso model. (**a**) MSE in dependence on lambda. (**b**) Coefficient paths in dependence on lambda.

5. Discussion and Limitation

5.1. *Influence of F3 and F4 on Y in Each City*

The relationship between F3, F4, and Y is depicted in Figure 7. Upon observing Figure 7a,b, it is apparent that the relationship between F3, F4, and Y appears to be discrete, especially between F3 and Y. The R^2 for the linear model between F4 and Y is 0.867, which is promising and significantly better than F3 and Y (0.482). However, upon marking the data of different cities with different colors (excluding cities with less than three sets of data), Figure 7c,d reveals a strong positive correlation between F3 and Y as well as F4 and Y in each city. Due to significant differences in economic levels among different cities, it is worthwhile to study the relationship between Y, F3, and F4 in each city separately.

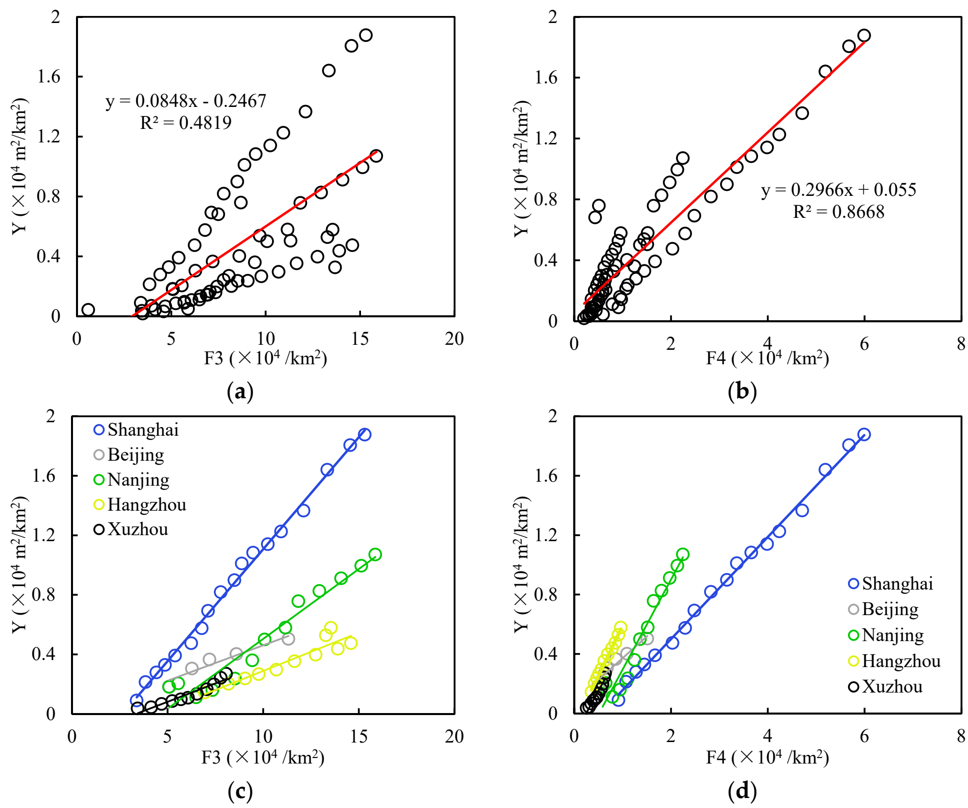

Figure 7. Variation relationship between F3, F4, and Y. (**a**) F3 (all data), (**b**) F4 (all data), (**c**) F3 (partial data), and (**d**) F4 (partial data).

The linear model was employed to fit the relationship between Y, F3, and F4 for each city. The R^2 and coefficients of the model for each city are presented in Table 8. Except for the model of F3 and Y in Hangzhou, the R^2 values of all other models are greater than 0.9, with some cities even reaching 0.99, like Shanghai and Hangzhou. Therefore, it may be more effective, simple, and accurate to make full use of the relationship between existing F3 or F4 and Y.

Table 8. The linear model of Y in each city.

City	F3 and Y			F4 and Y		
	R^2	Constant	b_1	R^2	Constant	b_1
Shanghai	0.995	−0.3977	0.1056	0.9959	−0.1853	0.3437
Beijing	0.9248	−0.0141	0.0476	0.9303	0.0818	0.2915
Nanjing	0.9437	−0.4344	0.0939	0.9455	−0.3106	0.6069
Hangzhou	0.8767	−0.2241	0.0513	0.9943	−0.0868	0.6806
Xuzhou	0.9195	−0.1657	0.0494	0.9316	−0.1455	0.5951

5.2. Limitation

While the proposed model effectively describes the relationship between factors and UUS demand in cities with available data, it is crucial to validate it in other cities to ensure its applicability. The surplus factor indicates the existence of other potential factors influencing UUS demand that have not been considered, such as policy and geological conditions. Additionally, the relationship between factors and UUS demand established in this study is valid within the scope of the collected data and, with the continuous expansion of developed UUS capacity, the demand for UUS may change. Therefore, these limitations should be taken into account when applying the models for UUS demand forecasting.

6. Conclusions

This study, based on collected data, investigates the primary driving factors for UUS demand in China and proposes a series of regression models. The findings can be summarized as follows:

(1) UUS demand has a strong correlation with urban economy and population factors, like GDP per km^2, population density, and land price. GDP per km^2 and GDP per capita are identified as primary driving factors and can comprehensively characterize the impact of urban population, economy, and transportation on UUS demand, while the influence of temperature-related indicators is small;

(2) Both multiple linear regression and LASSO models effectively describe the relationship between influencing factors and UUS demand, in which GDP per km^2 contributes the most, followed by GDP per capita, suggesting their potential as auxiliary tools for UUS development planning;

(3) The LASSO model stands out for selecting primary factors for UUS demand and constructing regression models without multicollinearity, showcasing its utility;

(4) Historical GDP per km^2 and GDP per capita of a city demonstrate a strong correlation with its UUS demand, serving as a valuable supplement for predicting UUS demand.

Author Contributions: Conceptualization, Y.D. and J.C.; methodology, B.Z.; software, J.C.; validation, J.C., Y.D. and B.Z.; formal analysis, J.M.; investigation, C.S.; resources, J.C.; data curation, Q.C.; writing—original draft preparation, J.C. and Y.D.; writing—review and editing, J.C. and J.M.; supervision, B.Z., Y.D., Q.C. and J.M.; project administration, Y.D.; funding acquisition, Y.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Science and Technology Program of Zhejiang Province (2022C35026); Zhejiang University of Science and Technology (2023JLYB001); and Youth Natural Science Foundation of Zhejiang University of Science and Technology (2023QN092).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

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

Appendix A

The data sources of UUS demand in different cities are shown in Table A1.

Table A1. Data sources of UUS demand.

City	Year	Data Sources	City	Year	Data Sources
Beijing	2006	[29]	Xuzhou	2010–2020	[30]
	2008, 2010, 2012	[25]	Guangzhou	2010	[25]
	2015	[18]	Shenyang	2012	[22,25]
Shanghai	2002–2005, 2012, 2014–2019	[31]	Ningbo	2015	[25]
	2006	[29,31]	Zhengzhou	2015	[18]
	2007–2011	[25,31]	Nanchang	2011	[18]
	2013	[25]	Wuxi	2005	[29]
Nanjing	2008, 2009	[30]	Qingdao	2004	[29]
	2010–2016	[25]	Wenzhou	2012	[32]
	2017–2020	[24,25]	Fuzhou	2014	[25]
Hangzhou	2010	[33,34]	Chengdu	2010	[25]
	2011	[34]	Taizhou	2014	Taizhou Municipal People's Government (https://www.zjtz.gov.cn (accessed on 4 February 2024))
	2012–2018	[25,34]	-	-	-
	2019–2020	[24,25]	-	-	-

References

1. China Urban Development Report Editorial Board. *China Urban Development Report 2010*; China City Press: Beijing, China, 2011. (In Chinese)
2. Qiao, Y.K.; Peng, F.L.; Sabri, S.; Rajabifard, A. Socio-environmental costs of underground space use for urban sustainability. *Sustain. Cities Soc.* **2019**, *51*, 101757. [CrossRef]
3. Xi, Y.; Li, X.; Zhu, H.; Zhang, W.; Zhao, S.; Xu, W. Three-dimensional high-precision assessment of mountainous urban underground space resources: A case study in Chongqing, China. *Tunn. Undergr. Space Technol.* **2022**, *123*, 104439. [CrossRef]
4. Xu, Z.; Zhou, S.; Zhang, C.; Yang, M.; Jiang, M. A Bayesian network model for suitability evaluation of underground space development in urban areas: The case of Changsha, China. *J. Clean. Prod.* **2023**, *418*, 138135. [CrossRef]
5. Chen, Z.; Wang, Y. *The Urban Underground Space Planning*; Southeast University Press: Nanjing, China, 2005. (In Chinese)
6. Xia, H.; Lin, C.; Liu, X.; Liu, Z. Urban underground space capacity demand forecasting based on sustainable concept: A review. *Energy Build.* **2022**, *255*, 111656. [CrossRef]
7. Lin, Q.; Zou, B.; Deng, Y. Evaluation of the development potential of urban underground space using set pair analysis and the comprehensive weight method. *Sustainability* **2023**, *15*, 6255. [CrossRef]
8. Zhao, J.W.; Peng, F.L.; Wang, T.Q.; Zhang, X.Y.; Jiang, B.N. Advances in master planning of urban underground space (UUS) in China. *Tunn. Undergr. Space Technol.* **2016**, *55*, 290–307. [CrossRef]
9. Yuan, H.; He, Y.; Wu, Y. A comparative study on urban underground space planning system between China and Japan. *Sustain. Cities Soc.* **2019**, *48*, 101541. [CrossRef]
10. Frisk, E.L.; Volchko, Y.; Sandstrom, O.T.; Soderqvist, T.; Ericsson, L.O.; Mossmark, F.; Lindhe, A.; Blom, G.; Lång, L.O.; Carlsson, C.; et al. The geosystem services concept—What is it and can it support subsurface planning? *Ecosyst. Serv.* **2022**, *58*, 101493. [CrossRef]
11. Qiao, Y.K.; Peng, F.L.; Luan, Y.P.; Wu, X.L. Rethinking underground land value and pricing: A sustainability perspective. *Tunn. Undergr. Space Technol.* **2022**, *127*, 104573. [CrossRef]

12. Qiao, Y.K.; Peng, F.L.; Wu, X.L.; Luan, Y.P. Visualization and spatial analysis of socio-environmental externalities of urban underground space use: Part 2 negative externalities. *Tunn. Undergr. Space Technol.* **2022**, *121*, 104326. [CrossRef]
13. Zhao, P.; Li, X.; Zhang, W.; Xiong, Z.; Xu, W.; Dai, J.; Deng, L. System dynamics: A new approach for the evaluation of urban underground resource integrated development. *Tunn. Undergr. Space Technol.* **2022**, *119*, 104213. [CrossRef]
14. He, L.; Song, Y.; Dai, S.; Durbak, K. Quantitative research on the capacity of urban underground space—The case of Shanghai, China. *Tunn. Undergr. Space Technol.* **2012**, *32*, 168–179. [CrossRef]
15. Bobylev, N. Mainstreaming sustainable development into a city's master plan: A case of urban underground space use. *Land Use Pol.* **2009**, *26*, 1128–1137. [CrossRef]
16. Wang, X.; Zhen, F.; Huang, X.; Zhang, M.; Liu, Z. Factors influencing the development potential of urban underground space: Structural equation model approach. *Tunn. Undergr. Space Technol.* **2013**, *38*, 235–243. [CrossRef]
17. Bobylev, N. Underground space in the Alexanderplatz Area, Berlin: Research into the quantification of urban underground space use. *Tunn. Undergr. Space Technol.* **2010**, *25*, 495–507. [CrossRef]
18. Chen, Z.; Chen, J.; Liu, H.; Zhang, Z. Present status and development trends of underground space in Chinese cities: Evaluation and analysis. *Tunn. Undergr. Space Technol.* **2018**, *71*, 253–270. [CrossRef]
19. Chen, L.; Zhu, X. *The Theory and Practice of Urban Underground Space Planning*; Tongji University Press: Shanghai, China, 1997. (In Chinese)
20. Li, X.; Xu, H.; Li, C.; Sun, L.; Wang, R. Study on the demand and driving factors of urban underground space use. *Tunn. Undergr. Space Technol.* **2016**, *55*, 52–58. [CrossRef]
21. Liu, J. An Empirical and Methodological Study on Demand Forecasting for Urban Underground Space. Master's Degree, Tsinghua University, Beijing, China, 2009. (In Chinese)
22. Gu, Q. A Comparative Study on Demand Forecasting of Underground Space in Different Cities. Master's Degree, Nanjing University, Nanjing, China, 2019. (In Chinese)
23. Lin, X.; Du, Z. Research on the forecast of urban underground space demand. *E3S Web Conf.* **2021**, *251*, 02094. [CrossRef]
24. Cui, J. Demand Analysis of Urban Underground Space Resources Development and Utilization Based on Grey Neural Network–Taking Nanjing and Hangzhou as Examples. Master's Degree, Nanjing University of Science & Technology, Nanjing, China, 2022. (In Chinese)
25. Xu, S.; Xu, S.; Cui, J. Development and utilization demand prediction of urban underground space resources. *J. Nanjing Univ. Sci. Technol.* **2023**, *36*, 26–36. (In Chinese)
26. Ge, R.; Li, X.; Yuan, F.; Jowitt, S.M.; Dou, F.; Xiong, Y.; Li, X. Demand evaluation of urban underground space through geospatial big data. *J. Urban Plan. Dev.* **2024**, *150*, 04023057. [CrossRef]
27. Cui, J.; Allan, A.; Lin, D. Analysis of motivations of developing underground pedestrian systems–decisive effect of weather conditions. In Proceedings of the 11th International WALK21 Conference & 23rd International Workshop of the International Co-Operation on Theories and Concepts in Traffic Safety, Hague, The Netherland, 16–19 November 2010.
28. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B* **1996**, *58*, 267–288. [CrossRef]
29. Qian, Q. Current status evaluation and prospects of urban underground space development and utilization in China. *Life Disaster* **2006**, *S1*, 1–5. (In Chinese)
30. Jiangsu Underground Space Society. *Report on the Development of Underground Space in Jiangsu Province 2022*; Jiangsu Phoenix Science Press: Nanjing, China, 2023. (In Chinese)
31. Xu, G.; Wang, B.; He, Y.; Qian, H. Study on accelerating the development and utilization of underground space resources in Shanghai. *Sci. Dev.* **2023**, *174*, 87–94. (In Chinese)
32. He, L.; Zhang, L.; Pan, N. Research on underground space planning, Wenzhou. *Planners* **2014**, *30*, 227–231. (In Chinese)
33. Sun, L.; Li, X.; Zhou, D.; Wang, R.; Liu, C. Study on the correlation between underground space development and socio-economic factors. *Chin. J. Undergr. Space Eng.* **2018**, *14*, 859–868+880. (In Chinese)
34. Xue, Y. Current situation of urban underground space data resources sharing and suggestions: Take Hangzhou as an example. *Value Eng.* **2020**, *39*, 258–261. (In Chinese)