GIS-Based Precise Predictive Model of Mountain Beacon Towers: A First Application to Bacon Sites in Wenzhou, China

Lifeng Tan
Tianjin University
Bei Wu (wubei@tju.edu.cn)
Tianjin University
Yukun Zhang
Tianjin University
Shuaishuai Zhao
Tianjin University

Research Article

Keywords: Mountain beacon towers, predictive model, GIS, logistic regression, viewshed analysis

Posted Date: October 20th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-969243/v1

GIS-based precise predictive model of mountain beacon towers: a first application to bacon sites in Wenzhou, China

Lifeng Tan[1], Bei Wu[1, 2, *], Yukun Zhang[1] and Shuaishuai Zhao[1]

[1] Key Laboratory of Information Technology for Architectural Heritage Inheritance of the Ministry of Culture and Tourism of China, School of Architecture, Tianjin University, China.

[2] Department of Architectural Engineering & Technology, Faculty of Architecture and the Built Environment, Delft University of Technology, the Netherlands.

*Corresponding author: wubei@tju.edu.cn

Abstract

In ancient China, where was frequently troubled by invaders, the government set up many beacon towers for alerting and transmitting military information along the border and the coast. Many beacon sites still exist in some areas, which are generally located in dangerous places with high mountains and rough terrain, bringing great difficulties to archaeological discovery. Therefore, it is particularly important to develop a predictive model applicable to the distribution of mountain beacon sites. Taking 68 beacon sites found in Wenzhou as research samples, this study uses the superimposed method of logistic regression and GIS viewshed analysis for the first time, forming a high-precision, scientific and operational predictive model for the distribution of beacon sites, which is verified by the cross-validation method. The results show that the beacon site predictive model simulated in this study can reduce the probability scope of site location by 90% compared with the traditional predictive model, which greatly improves the accuracy and ability of site prediction. At the same time, it can also be used to understand the relationship between the environment and known sites to assist in decision-making about conservation and management.

Keywords

Mountain beacon towers, predictive model, GIS, logistic regression, viewshed analysis

1 Introduction

Cultural heritage, as a witness of history, is a precious treasure left to mankind by history. In the coastal areas of China, there are many beacon tower remains, which were high platforms used to light fires to convey messages in ancient times, built to prevent enemy invasion, and played a vital role in the struggle against Japanese pirates. Once the enemy appeared, the garrisons of the beacon towers would quickly light the fire to report important military information so that the message of invasion could be spread one by one at the fastest speed until the soldiers in the nearby forts found it. To obtain a good view and expand its coverage as much as possible, beacon towers were usually built in places with high altitudes and wide vision. Therefore, there are many beacon sites on the mountaintops and high hills in coastal areas. However, over the course of history for centuries, beacon sites have gradually faced extinction due to natural erosion and man-made vandalism. Thus, it is exceedingly significant to discover these beacon sites as soon as possible and implement scientific and effective conservation measures.

However, the discovery of mountain beacon sites is time consuming and labor consuming. Owing to their long history, remote location, vegetation coverage, dilapidated condition and small dimension, beacon sites are difficult to investigate so that archaeologists usually acquire site information by asking local residents, consulting materials and combining field investigations. However, in fact, few local residents truly know about the beacon sites, so the useful information that can be obtained through inquiry is very limited, which poses great difficulties to archaeologists. According to the local archaeologists, the mountainous terrain in Wenzhou was complex and rugged, with tight vegetation coverage. In most cases, they could not rely on transportation tools, and the only way to reach the mountaintops was by walking. In extreme weather conditions, there are also many potential dangers. Even after overcoming those physical difficulties, there was no guarantee that new beacon sites would be found in each survey, which greatly reduced the efficiency of archaeological work and was very detrimental to conservation work. Therefore, we hope to simulate a predictive method for the distribution probability of mountain beacon sites to solve these difficulties, reduce the research workload, and improve the efficiency of the site survey.

Over the years, the most commonly used method of site probability prediction has been to build a mathematical model in the research area and calculate the probability values that the site may exist in an unknown area by analyzing the data of known samples [1]. This research originated in the 1950s, when Willey first applied the predictive model to the regional settlements of Weilu Valley, providing insight into the relationship between site and environment through qualitative analysis [2]. In the 1970s, quantitative analysis emerged, and Green used multiple linear regression to predict the probability of the existence of prehistoric Mayan sites in Northern British Honduras, laying the groundwork for the subsequent application of site predictive models to cultural resource management [3]. Thanks to the development of computer science, the two disciplines of statistics and spatial analysis skillfully combined in the 1980s led to great advances in site predictive research [4-8]. Entering the 1990s, with the emergence of relevant computer software and the continuous promotion of GIS analysis methods in the field of archaeological research, the research methodologies gradually diversified, and the predictive means were more scientific and mature, which has vigorously developed site predictive research in terms of both depth and breadth [9-12].

Constructing a site predictive model is a complex task that requires effective methodology, appropriate parameters, and a scientific approach of validation [13]. Statistical regularities and distribution characteristics of known sites are identified first based on the analysis of environmental factors in the given area, such as altitude, slope, distance from transport system or hydrographic net; then, a multivariate discriminant function is adopted to assess the existence probability of other sites in this area, giving a potential site distribution map [14]. The application of quantitative analysis makes the study of human-land relationships more accurate and scientific, among which the most widely used is the logistic regression analysis method. This approach is frequently conducted as a useful tool for archaeological investigations, not only providing an important decision support system to obtain usable information but also saving time and money, especially in large areas [15].

The locational selection of mountain beacon towers was not random, but it was a result of logical decisions on the basis of geographical considerations and factors related to human activity. Most of them are found on the mountaintops along the traffic arteries between the forts or on the high hills of the coastal peninsula, where are easy to look out. Moreover, due to the requirement of visual factors, beacon towers were placed to ensure the accessibility of sightlines. Considering the two factors of environment and vision, we propose to adopt the superimposed method of logistic regression and GIS viewshed analysis to simulate beacon sites' location preferences. Theoretically, this method can effectively improve the prediction accuracy and further narrow the possible scope of unknown beacon sites compared to the single logistic regression predictive method commonly used in the past. The working flow pursued the following steps (Figure 1): 1) The historical construction characteristics and distribution features of known beacon sites were analyzed to screen out the natural and cultural factors affecting their positions in space; 2) logistic regression analysis was integrally conducted on the beacon sites, nonsites and multiple independent variables, establishing a GIS calculation model for predicting the discovery probability of unknown sites in the given area; 3) the accuracy of the result was verified by the cross-validation method; and 4) GIS viewshed analysis was superimposed on the previous result to obtain the final prediction result of unknown beacon sites.

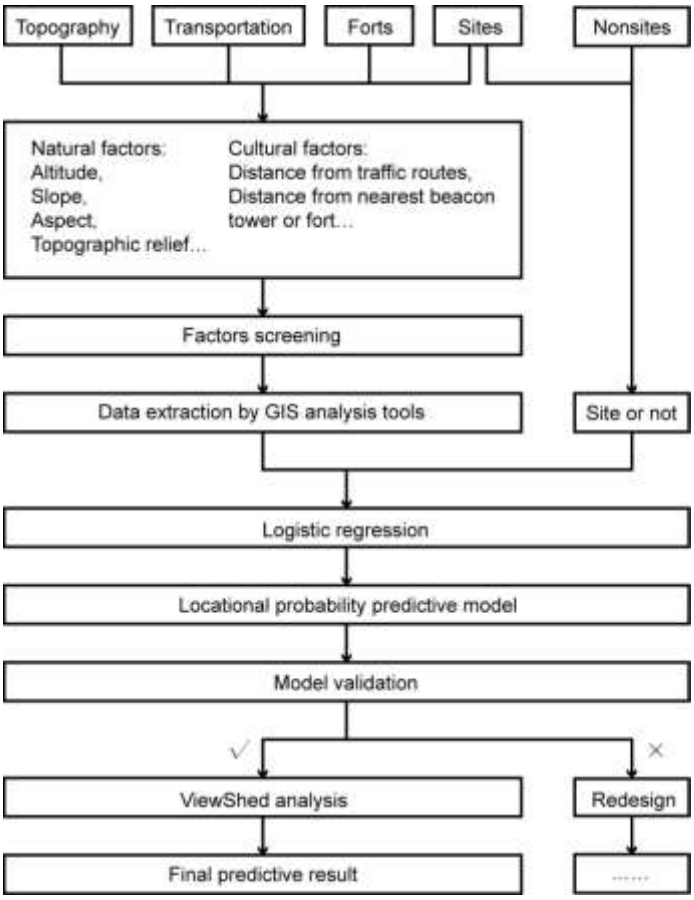

Figure 1 The working flow chart for the prediction of beacon sites

2 Study object and study area

Beacon tower (烽燧, Feng Sui), the fastest and most effective way to transmit military information in the cold weapon era, has a long history in China [16]. According to the *Records of the Grand Historian*, there were beacon towers as early as the Western Zhou Dynasty, and the beacon system was quite mature during the Han Dynasty and reached its peak during the Tang Dynasty, following which it was used throughout the Song to Qing Dynasties [17]. Before the emergence of modern communication technology, beacon towers had been valued by successive Chinese dynasties as ancient but effective military communication tools and were commonly placed in border areas. In the Ming Dynasty, the coastal areas were constantly invaded by remnant rebel forces and Japanese pirates, and the maritime situation was extremely severe. To anticipate the invasion of enemies and take timely precautions, the Ming government constructed beacon towers along the coast as indispensable coastal defense facilities to resist enemies. When enemies occurred, smoke was burnt during the day, called "Feng", and fire was lit at night, called "Sui", so that the beacon towers were connected to convey information [18]. Beacon towers, together with forts, played a vital role in the struggle against Japanese pirates on the coastline. Therefore, a systematic study on the distribution of beacon towers is a prerequisite and necessary process for the exploration of coastal defense systems.

In this study, the Wenzhou (温州) area is chosen as the study area, including the inner city of Wenzhou as well as the counties and villages under its jurisdiction, with a total area of approximately 12,110 km^2 [19]. Wenzhou is located in East China, southeast of Zhejiang Province, lower reaches of the Ou River, bordering the East China Sea to the east and Fujian Province to the south, which is one of the nationally famous historical and cultural cities (Figure 2-a). Being a coastal city, Wenzhou has been a key place for military defense throughout its history. Especially in the Ming Dynasty, Wenzhou was heavily infested by Japanese pirates with an extremely serious situation, so 58 beacons were built in Yueqing County, Rui'an County and Pingyang County along the coast of Wenzhou to prevent and resist enemies [20] and were increased in the late Ming and Qing Dynasties. By far, archaeologists have discovered 68 beacon sites in the Wenzhou area (Figure 2-b), most of which have fallen into disrepair. The materials used to build the beacon towers were all taken from nature, mostly made of clay and gravel, and rammed layer by layer; some were piled up with bricks and stones, with square, rectangular and round shapes. These beacon sites are historical evidence of Zhejiang coastal people's bravery in fighting against Japanese pirates. Based on the known data, this study will simulate a predictive method applicable to Wenzhou’s mountain beacon sites.

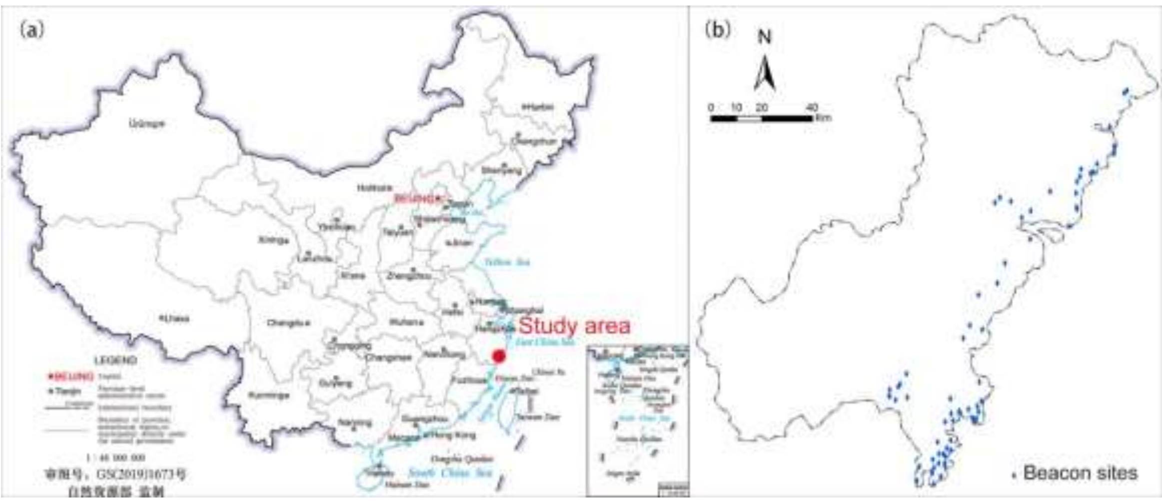

Figure 2 Study area: (a) location of Wenzhou (source: Ministry of Natural Resources of China, Approval number: GS(2019)1673); (b) distribution of beacon sites

3 Materials and methods

3.1 Data sources and extraction

To build a predictive model for the beacon sites, a certain number of experimental samples in the survey area are needed to study the probability of the dependent variables, including both sites and nonsites. The data of beacon sites are collected from the survey data of the Wenzhou Institute of Cultural Relics and Archaeology, and nonsites are variables introduced specifically for the predictive model. Conceptually, all the locations where no sites are found can be used as nonsites, so the "Create Random Points" tool of ArcGIS software is used to generate 100 random points as nonsites of the experimental samples.

Next, the independent variables related to the construction of the predictive model need to be screened out. For the independent variables, natural and cultural factors that potentially influenced the spatial distribution of the sites can be integrated into the predictive model. The selection of independent variables was mainly determined by the mode and features of beacon transmission during the Ming and Qing dynasties, combined with the distribution characteristics of existing sites.

Wenzhou is mostly mountainous, and the terrain generally shows the features of high inland and low coastal areas; hence, the topographical parameters are mainly considered for natural variables. The survey data show that the discovered beacon sites are mainly located in the coastal area and on the top of the hill with a good view as a result of the consideration for facilitating the information transmission among beacon towers. In addition, taking into account the delivery of living supplies and smoke-burning materials, the locational selection of the beacon towers must also consider the ease of access. Therefore, the altitude, slope, aspect and topographic relief that may be linked to their locations will be taken as alternative independent variables for further screening.

Except for topographical factors, the distribution of beacon sites was also probably influenced by cultural factors related to human activity. For example, because the beacon tower relied on smoke to transmit, only the neighboring beacon tower or fort was set within its viewing range so that the information could be transmitted one by one until the soldiers of forts received it, so the distance from nearest beacon tower or fort was inevitably under consideration when constructing the beacon system. Similarly, due to the need for convenient access to materials and resources, the distance between beacon towers and main traffic routes might also be one of the factors to be measured. Therefore, the cultural variables considered in this study mainly include two aspects, namely, the distance from traffic routes and the distance from the nearest beacon tower or fort.

In general, six alternative independent variables are selected in this study for further screening, and the relevant data can be extracted using GIS spatial analysis tools (Table 1). Variables related to natural factors, including altitude, slope, aspect and topographic relief, can be extracted from the digital elevation model (DEM) of the Wenzhou area, provided by the Geospatial Data Cloud site, Computer Network Information Center, Chinese Academy of Sciences (http://www.gscloud.cn); variables linked to cultural factors, including the distance from traffic routes and nearest beacon tower or fort, can be derived from the historical literature using GIS spatial analysis tools. After extraction, the raster data of each alternative independent variable required in this study can be obtained (Figure 3).

Table 1 Alternative independent variables

Type	**Independent variables**	**Data sources**	**Acquisition method**
Natural factors	Altitude	DEM	/
	Slope	DEM	GIS→Slope
	Aspect	DEM	GIS→Aspect
	Topographic relief	DEM	GIS→Block Statistics
Cultural factors	Distance from traffic routes	Routes of urgent delivery stations in the Ming Dynasty [20]	GIS→Euclidean Distance
	Distance from nearest beacon tower or fort	Beacon sites data, military forts data in the Ming Dynasty [20]	GIS→Neighbor

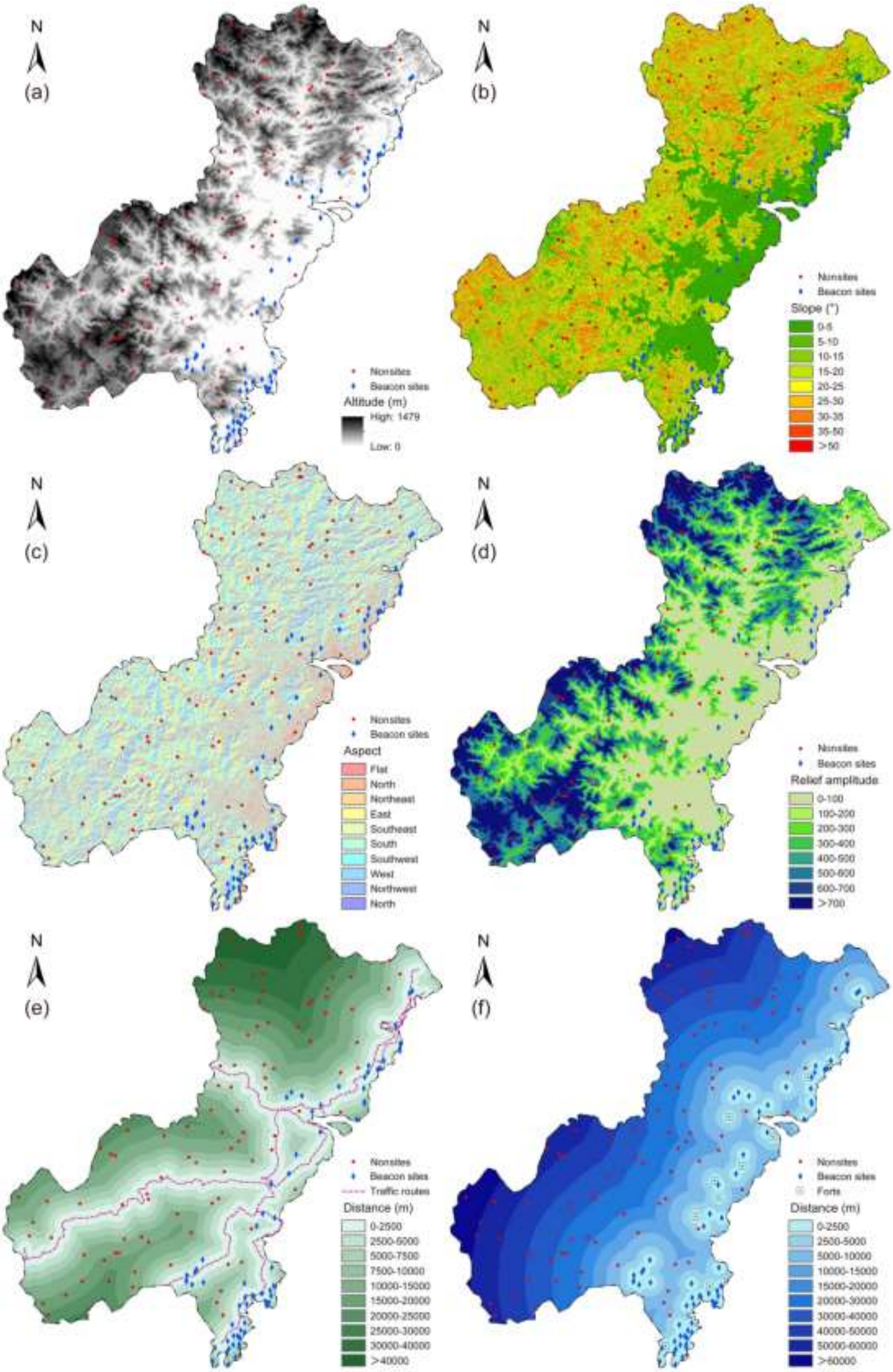

Figure 3 Raster data of alternative independent variables: (a) altitude; (b) slope; (c) aspect; (d) topographic relief; (e) distance from traffic routes; (f) distance from nearest beacon tower or fort

3.2 Independent variables screening

In the second stage, the alternative independent variables need to be screened by statistically analyzing their correlation with the beacon sites, and those that are significantly correlated with the locational choice of beacon sites will be considered as the independent variables in this study. For the five factors of altitude, slope, aspect, topographic relief and distance from traffic routes, superimposing the raster data with the beacon site distribution map respectively, the attribute values of the 68 beacon sites were extracted using the "Extract Multi Values to Points" tool of ArcGIS software, while the distance from the nearest beacon tower or fort was calculated by the "Near Neighbor" tool of ArcGIS software (Supplementary Table S1). The statistical results are shown in Table 2.

Table 2 Beacon sites distribution statistics

Factors	Interval partition	Number of sites	Percent
Altitude	0-100 m	21	30.88%
	100-200 m	23	33.82%
	200-300 m	15	22.06%
	300-400 m	6	8.82%
	400-500 m	2	2.94%
	500-600 m	0	0
	600-700 m	1	1.47%
	＞700 m	0	0
Slope	0-5°	11	16.18%
	5-10°	27	39.71%
	10-15°	20	29.41%
	15-20°	5	7.35%
	20-25°	2	2.94%
	25-30°	3	4.41%
	＞30°	0	0
Aspect	Flat	0	0
	North	4	5.88%
	Northeast	11	16.18%
	East	9	13.24%
	Southeast	10	14.71%
	South	9	13.24%
	Southwest	9	13.24%
	West	10	14.71%
	Northwest	6	8.82%
Topographic relief	0-100	28	41.18%
	100-200	17	25%
	200-300	14	20.59%
	300-400	6	8.82%
	400-500	2	2.94%
	500-600	0	0

	600-700	1	1.47%
	＞700	0	0
Distance from traffic routes	0-1000 m	19	27.94%
	1000-2000 m	11	16.18%
	2000-3000 m	13	19.12%
	3000-4000 m	12	17.65%
	4000-5000 m	4	5.88%
	5000-6000 m	3	4.41%
	6000-7000 m	3	4.41%
	＞7000 m	3	4.41%
Distance from nearest beacon tower or fort	0-1000 m	10	14.71%
	1000-2000 m	20	29.41%
	2000-3000 m	18	26.47%
	3000-4000 m	7	10.29%
	4000-5000 m	8	11.76%
	5000-6000 m	1	1.47%
	6000-7000 m	3	4.41%
	＞7000 m	1	1.47%

Statistical results show that the locations of beacon sites are preferred in terrain areas with altitude 0-300 m, slope 0-15°, and topographic relief 0-300, whereas showing no obvious preference for aspect; the distances from traffic routes are relatively close, mostly reachable within 4000 m; and the distribution of beacon sites is relatively concentrated, mostly within 5000 m apart. This suggests that the locations of beacon sites present a greater correlation with altitude, slope, topographic relief, distance from traffic routes and nearest beacon tower or fort, while there is no obvious correlation with aspect. Therefore, in this study, altitude, slope, aspect, topographic relief, distance from traffic routes and distance from nearest beacon tower or fort will be taken as independent variables in the predictive model of beacon sites.

3.3 Logistic regression and modeling

After identifying the independent and dependent variables, regression analysis and modeling can be carried out. Logistic regression is a nonlinear categorical statistical method used in regression analysis of qualitative variables, originated by the biomathematician P. F. Verhult in 1838, and later popularized and widely used in population statistics and prediction [21]. According to the differences in variables, logistic regression can be divided into binary logistic regression and multiple logistic regression. Binary logistic regression, in which the dependent variable can only take 0 and 1, has the advantage of dealing with multiple types of independent variables such as classification, sequencing and distance fixing at the same time, which is suitable for analyzing complex variables and the interactions among them [22]. Based on that, the binary logistic regression method is adopted in this study to quantitatively analyze the locational preferences of beacon sites in the Wenzhou area.

According to the rule of binary logistic regression modeling, P is set as the probability of the existence of beacon sites, with a value in the range [1,0], and 1-P is the probability of the nonexistence of beacon sites. Taking the natural logarithm of ratio $\frac{P}{1-P}$ to obtain $\ln\frac{P}{1-P}$, namely, logit transformation of P, and denoted as logit P, then the value range of logit P is (-∞, +∞). With logit P as the dependent variable, the linear regression equation is established as follows.

$$\text{logit } P = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \quad (1)$$

The logistic regression equation is obtained from Equation (1), as follows.

$$P = \frac{\exp(\alpha+\beta_1 x_1+\beta_2 x_2+\cdots.+\beta_n x_n)}{1+\exp(\alpha+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n)} \quad (2)$$

where α is a constant term, indicating the natural logarithm of the ratio when the independent variable takes the value of 0; x_n is the independent variable of factors affecting the distribution of sites; and β_n is the partial regression coefficient of the logistic regression.

With the aim of verifying the accuracy of the predictive model, this paper divides beacon sites into two parts, one part for validation samples and one part for experimental samples, as detailed in section 3.4. Extracting data on altitude (X1), slope (X2), topographic relief (X3), distance from traffic routes (X4) and distance from nearest beacon tower or fort (X5) of the 58 beacon sites and 100 nonsites in the experimental samples, a binary logistic regression analysis was conducted (Supplementary Table S2). However, when the number of alternative independent variables is large, the variables may have a high correlativity, and some of them may not have a significant effect on the dependent variable. Hence, the target variables need to be screened for significance in the logistic regression analysis. The "Backward: LR" algorithm of SPSS software can automatically screen and eliminate variables with low significance relationships with the model and carry out iterative calculations to arrive at the ideal model parameters, as shown in Table 3.

Table 3 Predictive model parameters of beacon sites

	Variables	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1	X1	0.338945	0.111998	9.158720	1	0.002475	1.403466
	X2	-0.013722	0.042333	0.105077	1	0.745820	0.986371
	X3	-0.337555	0.111967	9.088920	1	0.002572	0.713512
	X4	-0.000467	0.000184	6.478197	1	0.010921	0.999533
	X5	-0.000615	0.000179	11.735614	1	0.000613	0.999385
	Constant	4.858216	1.456475	11.126213	1	0.000851	128.794264
Step 2	X1	0.337295	0.110906	9.249337	1	0.002356	1.401152
	X3	-0.336198	0.110917	9.187388	1	0.002437	0.714482
	X4	-0.000463	0.000182	6.457170	1	0.011051	0.999537
	X5	-0.000615	0.000179	11.749648	1	0.000609	0.999386
	Constant	4.715911	1.374997	11.763261	1	0.000604	111.710527

The calculation result indicates that although slope is one of the influencing factors of beacon tower distribution, the significance is not high in the model. Therefore, the final predictive model contains four independent variables, namely, altitude (X1), topographic relief (X3), distance from traffic routes (X4) and distance from nearest beacon tower or fort (X5). By substituting the parameters in Table 3 into Equation (2), the equation for the locational probability of beacon sites in the Wenzhou area can be derived as follows.

$$P = \frac{\exp(4.715911+0.337295x_1-0.336198x_3\ -0.000463x_4-0.000615x_5)}{1+\exp(4.715911+0.337295x_1-0.336198x_3\ -0.000463x_4-0.000615x_5)} \tag{3}$$

With that, a predictive model of the beacon sites can be generated using the "Raster Calculator" tool of ArcGIS software. The raster calculator is a tool that conducts mathematical operations on multiple raster layers, thus generating a new raster layer and recording the operation result. By substituting each raster layer (Figure 3) into Equation (3), a locational probability map of the beacon sites in the Wenzhou area can be obtained (Figure 4), with probability value P ranging from 0 to 1.

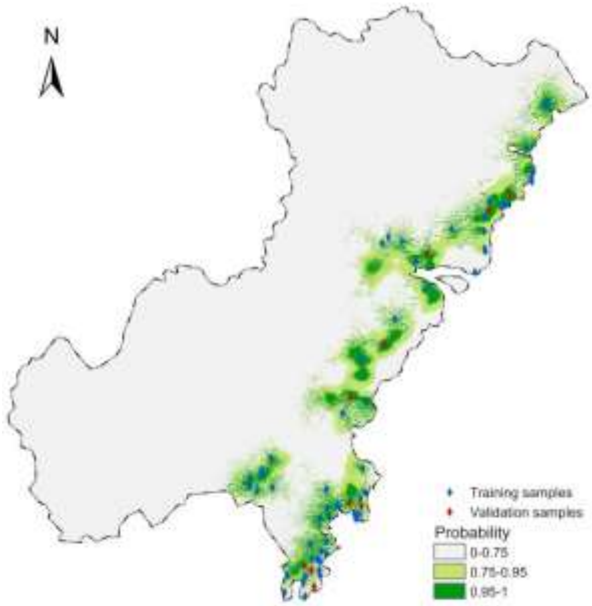

Figure 4 Locational probability map of beacon sites in the Wenzhou area

3.4 Model validation

After establishing the predictive model, the validity and accuracy of the predictive model need to be evaluated. To meet this requirement, this study uses the cross-validation method to evaluate the accuracy of the predictive model. Cross-validation is a statistical method based on the idea of data slicing, in which the whole data set is split into several data sets according to the actual needs. In this study, we divide the beacon site dataset into two parts, one of which is used as the training sample and the remaining as the validation sample. Usually, the training samples are processed first, and then the validation samples are used to test the model obtained from the training sample dataset to evaluate the accuracy.

To ensure the number of training samples, 10 sites are taken as validation samples, and the remaining 58 sites are used as training samples. Additionally, considering the homogeneity of the samples and further calculations, the sites in the uniformly distributed areas are selected as validation samples (Figure 4). Extracting the probability values corresponding to the 10 validation samples, the results are shown in Table 4. The data show that there are 9 out of 10 validation samples with probability values above 0.95, which are in the very high probability zone, indicating that the predictive model of beacon sites established in this study has a high efficiency and can be used for further research.

Table 4 Probability values of the validation samples

Validation samples	1	2	3	4	5	6	7	8	9	10
Probability values	0.984	0.999	0.998	0.995	0.995	0.982	0.943	0.999	0.999	0.999

3.5 Viewshed analysis

Having confirmed the validity of the predictive model, a further analysis can then be performed on this basis to precisely determine the prediction range. Due to the influence of visual factors, the setting of beacon towers must be within the effective scope of vision to form one or more coherent information flows to serve as a qualified enemy alerting system. Thus, the beacon towers must be located in places that not only meet the requirements of terrain and resources but also ensure the safety and unobstruction of the sightline. Based on that, the "Viewshed" tool of ArcGIS software can be used for superimposed analysis. The fundamental principle is as follows: assuming there are beacons A and B in a certain area and the sightline is obstructed, it can be concluded that there may be a beacon C between A and B that has not yet been discovered. Then, by calculating the viewshed of beacons A and B and superimposing the overlapping area with the high probability area of the sites, we can obtain the possible existence area of beacon C.

Since the distribution of beacon towers in the northern area of Wenzhou is relatively simple, this study selects the beacon towers within the jurisdiction of Jinxiang Wei in the southern area as a case for analysis. Jinxiang Wei was the highest-ranking fort in the southern area of Wenzhou, and the military mobilization and war command of other forts within this area were under its administration, which was the final collection point for all kinds of military information. There are 7 known experimental beacon sites and 1 fort site in the study area. Taking each site point as the observation point, the visible conditions among the sites are calculated, and the statistical results are shown in Table 5.

Table 5 Viewshed analysis results statistics of experimental sites

Sites	Jinxiang Wei	Fenghuang	Beishan	Lingtou	Nanshan	Ma’anshan	Xiaoyandun	Dayandun
Jinxiang Wei	•		•					

Fenghuang		•	•					
Beishan	•	•	•					
Lingtou				•	•	•		
Nanshan				•	•	•		
Ma’anshan				•	•	•		
Xiaoyandun							•	•
Dayandun							•	•

(Note: “ • ” means visible)

Through the preliminary viewshed analysis, it is found that the distance between the two beacon towers connected by sight is within 4 km, transmitting information to Jinxiang Wei from the southwest and southeast directions separately, which created a precise intelligence network. Among them, the Beishan beacon is within the visual area of Jinxiang Wei, and the Fenghuang beacon can also indirectly transmit information to Jinxiangwei via the Beishan beacon. However, the transmission route between the Dayandun beacon and Jinxiang Wei and the transmission routes between the Lingtou beacon, Nanshan beacon and Beishan beacon are blocked so that the information from the Xiaoyandun beacon and Ma'anshan beacon cannot be transmitted to Jinxiang Wei. Therefore, it is judged that other beacon towers that are not yet found may exist in areas A and B (Figure 5).

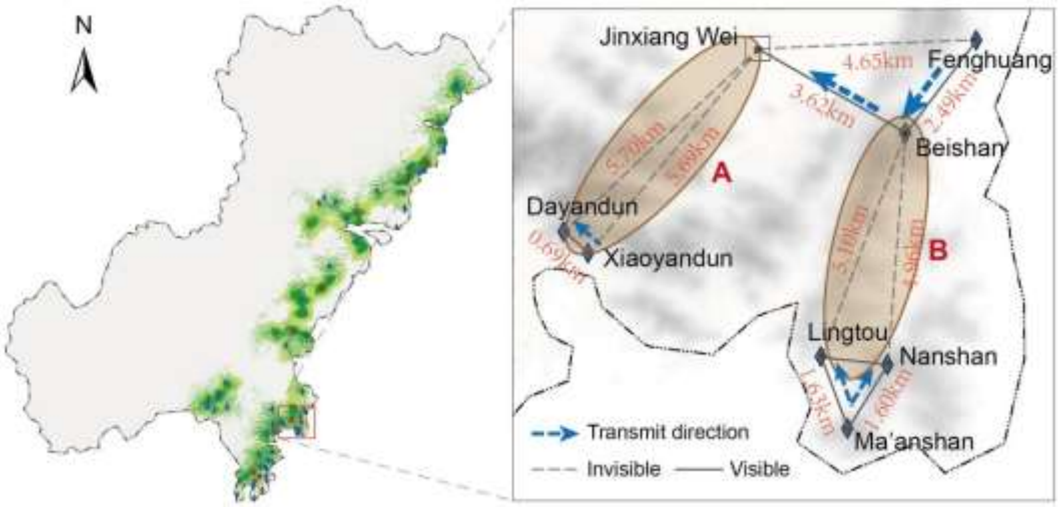

Figure 5 Viewshed analysis area

Next, the viewshed areas of Jinxiang Wei and Dayandon Beacon (Figure 6-a) and the viewshed areas of Nanshan Beacon, Lingtou Beacon and Beishan Beacon (Figure 6-b) are superimposed to obtain the possible existence areas of beacon sites in terms of visual factors.

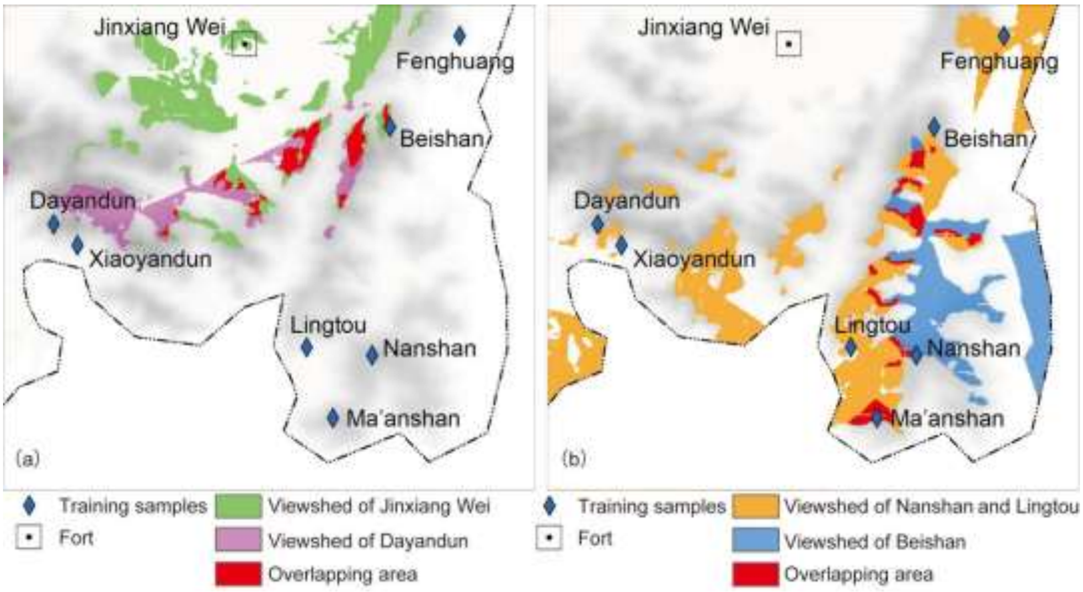

Figure 6 Viewshed analysis results: (a) the viewshed area of Jinxiang Wei and Dayandon Beacon; (b) the viewshed area of Nanshan Beacon, Lingtou Beacon and Beishan Beacon

4 Results and discussion

By superimposing the viewshed analysis result on the distribution probability map of the beacon sites, the final prediction result of beacon sites in the Jinxiang Wei area can be obtained (Figure 7-a). To verify the accuracy of the final prediction result, we project the validation samples of beacon sites onto the final probability map (Figure 7-b), and it can be seen that the actual beacon site locations basically match the final prediction result, which demonstrates the scientificity and practicability of the predictive method constructed in this study.

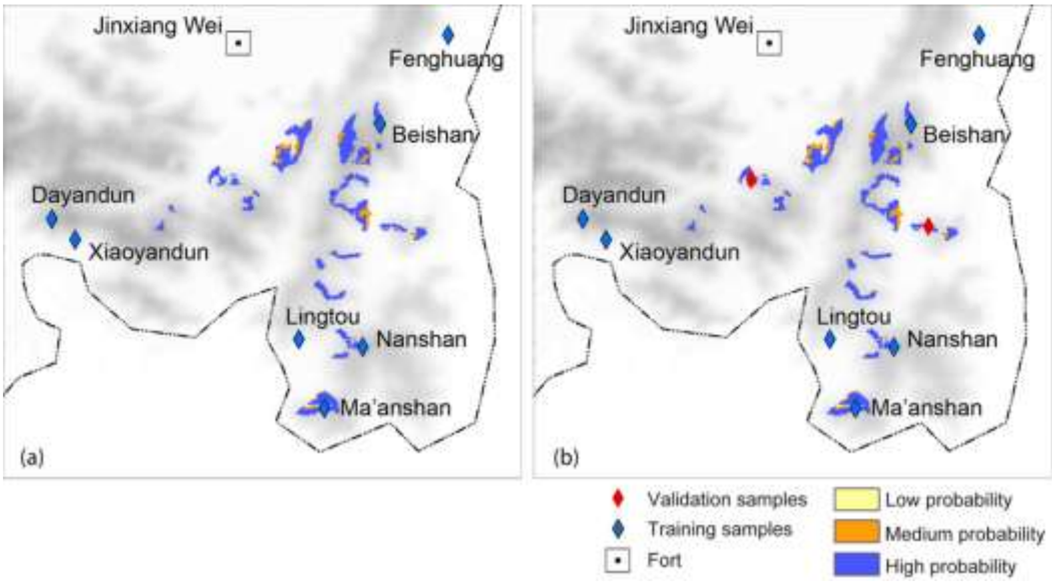

Figure 7 Final prediction result in the Jinxiang Wei area: (a) the prediction result; (b) validation result

In addition, the following conclusions can be drawn from this study.

1) By comparing the final prediction result obtained after superimposing viewshed analysis with

the traditional logistic prediction result, it is found that the predictive scope is reduced by more than 90%, which greatly improves the accuracy and predictive ability of site prediction and can provide decision-making guidance for archaeological excavation work to a certain extent, thus effectively saving manpower and material resources, reducing the blindness of archaeological work, and yielding twice the result with half the effort.

2) During the research process, the independent variables that have a significant influence on the establishment of the predictive model can be screened out through data extraction and statistical analysis, and all those variables have been regarded as important environmental factors that influence ancient humans' residential choices by researchers in archaeological research.

3) Through GIS spatial analysis and logistic regression analysis, it is found that each natural and cultural independent variable affects the distribution of the sites quantitatively. The results suggest that the locational selection of beacon sites in Wenzhou is significantly influenced by traffic routes, followed by topographic relief and altitude, while slope does not have strong effects on site distribution. Moreover, the rules of siting for beacon towers in ancient times can also be inferred: the ancient people attached great importance to building beacon towers in areas with wide vision and small topographic relief, not only could get a good view, easy to look out but also convenient for construction; additionally, they also preferred locations near transportation system, where enemy activities were frequent, to monitor their movements and to quickly transmit alerts to military forts for timely decision making and counterattack organization.

Of course, the shortcomings of the method used in this study must be recognized.

1) The independent variable factors selected in this study are limited. The site selection of beacon towers might be closely related to military, politics, Fengshui, climate and other factors, which are difficult to extract quantitatively, so they are not considered influencing factors in this study.

2) There are some constraints in data acquisition. The research object was in ancient times, but ancient data could not be obtained, so modern data were used as a substitute; moreover, there were some human errors in data vectorization. If it is possible to obtain more accurate and reliable data, the prediction result will be more scientific and can further improve the reliability of the research.

3) The nonsites in the experimental samples come from the random points generated by ArcGIS software, among which are likely to be beacon sites waiting for excavation. When they act as nonsites in the predictive model, it will have a weak impact on the accuracy of the final prediction results. In addition, there might be some inevitable subjective factors when classifying the training and validation samples.

4) There are certain limitations of the viewshed analysis method, which is only applicable to the case in which the beginning and ending beacon sites of a transmission route are known and the intermediate connection beacon site is missing. In contrast, it is impossible to make inferences by analyzing the viewshed of known beacon sites.

Nevertheless, the use of GIS and statistical methods has transformed archaeological research from qualitative to quantitative analysis, and the quantification of data analysis has provided new perspectives for settlement archaeological research. The predictive method of Wenzhou’s beacon sites proposed in this study has a huge breakthrough in prediction accuracy and ability compared with previous predictive methods, which can narrow the possible area of the beacon sites to a smaller scope and establish a predictive model with more practical predictive ability.

5 Conclusions

In this study, 68 discovered beacon sites and 100 random nonsites in the Wenzhou area were taken as experimental samples, which were divided into two parts: training samples and validation samples. Next, with the help of GIS spatial analysis tools, data extraction, statistical analysis and screening of each independent variable were carried out, and the distribution predictive model of beacon sites was established through logistic regression. On this basis, the viewshed analysis of the beacon sites was conducted according to their visual attributes to further precisely determine the predictive scope, improve the accuracy and predictive ability of traditional prediction methods, and reduce the difficulty of archaeological work to the greatest extent possible. The results show that the predictive method of beacon sites proposed in this study reduces the possible location scope by more than 90% and is proven to be scientific and reliable.

From the perspective of historical research, the predictive model can also be used to investigate the relationship between the beacon sites and various natural and cultural environmental parameters, revealing the deployment characteristics of mountain beacon towers and deepening the understanding of military heritage. In addition, for heritage management, the site predictive model has great potential for application. China has numerous cultural heritages, and even though three national relic surveys have been conducted, a large number of sites are still missing. In the face of frequent infrastructure construction, finding ways to improve efficiency, save costs, and protect the sites that have not yet been discovered and registered is a serious problem. The site predictive model can be used to predict the probability of site discovery in unknown locations and assist in protection and management decisions. Moreover, regional systematic investigation for settlement morphology research has been widely carried out nationwide, which provides the basic conditions for the establishment of archaeological site predictive models in various places. Therefore, the application and promotion of archaeological predictive models in heritage conservation and infrastructure construction are not only necessary but also practical.

References

[1] Danese, M., et al. "Predictive modeling for preventive Archaeology: overview and case study." Central European Journal of Geosciences 6.1:42-55. https://doi.org/10.2478/s13533-012-0160-5 (2014).

[2] Willey, Gordon R. Prehistoric settlement patterns in the Virú; Valley, Peru. Bureau of American

Ethnology Bulletin. 155:1–453 (1953).
[3] Green, Ernestene L. Location Analysis of Prehistoric Maya Sites in Northern British Honduras. American Antiquity, Society for American Archaeology 38.3: 279–93. https://doi.org/10.2307/279715 (1973).
[4] Kvamme, K. Computer processing techniques for regional modeling of archaeological site locations. Advances in Computer Archaeology. 1.1: 26-52 (1983).
[5] Kvamme, K. Development and testing of quantitative models in *Quantifying the Present and Predicting the Past: Theory, Methods, and Applications of Archaeological Predictive Modeling*. 325-428 (Denver: US Department of Interior Bureau of Land Management, 1988).
[6] Kohler T A. 1988. Predictive locational modelling: History and current practice in *Quantifying the Present and Predicting the Past: Theory, Method and Application of Archaeological Predictive Modelling*. 19-60 (Denver: US Department of Interior Bureau of Land Management, 1988).
[7] Judge W J, Sebastian L. *Quantifying the present and predicting the past: theory, method, and application of archaeological predictive modeling*. (Denver: US Department of Interior Bureau of Land Management, 1988).
[8] Kohler, T. A., and S. C. Parker. Predictive Models for Archaeological Resource Location. Advances in Archaeological Method and Theory 9.2:397-452 (1986).
[9] Märker, M., & Heydari-Guran, S. Application of data-mining technologies to predict Paleolithic site locations in the Zagros Mountains of Iran. Geography. doi:10.1165/ajrcmb.25.6.4616 (2009).
[10] Balla, A., et al. Locating Macedonian tombs using predictive modeling. Journal of Cultural Heritage 14(5):403-410 (2013).
[11] Graves, D. The use of predictive modelling to target Neolithic settlement and occupation activity in mainland Scotland. Journal of Archaeological Science 38.3:633-656 (2011).
[12] Hou, G. L., C. J. Xu, and J. Y. Xiao. Comparative analysis of prehistoric sites distribution around 4kaB.P. in Gansu-Qinghai region based on GIS. Scientia Geographica Sinica. 32.1 (2012).
[13] Yan L, Lu P, Chen P, Danese M, Li X, Masini N, Wang X, Guo L, Zhao D. Toward an Operative Predictive Model for the Songshan Area during the Yangshao Period. ISPRS International Journal of Geo-Information. 10.4:217. https://doi.org/10.3390/ijgi10040217 (2021).
[14] 高立兵 [Libing Gao]. 时空解释新手段——欧美考古 GIS 研究的历史、现状和未来 [A new means of spatiotemporal interpretation -- the history, present situation and future of archaeological GIS research in Europe and America]. Archaeology. 07: 89-95 (1997).
[15] 郭飞 [Fei Guo]. 基于 Logistic 回归最优化方法的遗址预测分布研究 [Study of Archaeological Sites Predictive Distribution Based on Logistic Regression Optimization Method —— A Cast Study of Fenhe River Basin]. Diss. University of Chinese Academy of Sciences (2018).
[16] 尹泽凯，张玉坤，谭立峰 [Zekai Yin, Yukun Zhang, Lifeng Tan]. 明代海防层次和聚落体系研究 [Studies On The Levels And Forts Of The Coastal Defense System Of The Ming Dynasty]. Architecture and Culture. 001:104-105 (2016).

[17] 石一民 [Yimin Shi]. 风雨时光远，沧桑烽燧近 [Wind and rain are far away, beacons of vicissitudes are near]. Zhoushan Daily. 04-03:005 (2018).
[18] 刘建军, 张玉坤, 曹迎春 [Jianjun Liu, Yukun Zhang, Yingchun Cao]. 基于可达域分析的明长城防御体系研究 [Study on the defense system of Ming's Great Wall based on the reachable domain]. Journal of Architecture S1:108-111 (2013).
[19] Baidu Baike. *Wenzhou*. https://baike.baidu.com/item
[20] Zhang Y, Wu B, Tan L, Liu J. Quantitative research on the efficiency of ancient information transmission system: A case study of Wenzhou in the Ming Dynasty. PLoS ONE 16.4: e0250622. https://doi.org/10.1371/journal.pone.0250622 (2021).
[21] Guanghui Jiang, Fengrong Zhang, Junwei Chen, et al. Analysis of the driving forces of change of rural residential areas in Beijing mountainous areas based on Logistic regression model. Transactions of the CSAE. 23.5:81-87 (2007).
[22] 张海 [Hai Zhang]. *GIS 与考古学空间分析* [GIS and archaeological spatial analysis]. (Peking: Peking University Press, 2014).

Data availability statement

All data generated or analyzed during this study are included in this published article (and its Supplementary Information files).

Acknowledgments

This work was supported by the National Natural Science Foundation of China [grant numbers 52078324, 51778400]; Major Research on Philosophy and Social Sciences of the Ministry of Education of China [grant numbers 19JZD056, 18JZD059]; and the China Scholarship Council.

Author contributions statements

The idea and methods were proposed by L.T. The data were collected and analyzed by B.W. The paper was written and structured by B.W. and reviewed by S. Z. The work was discussed and supervised by Y.Z.

Competing interests

The authors declare no competing interests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTableS1.xlsx
- SupplementaryTableS2.xlsx