Archaeological Predictive Modeling Based on RS and GIS in Taihu Lake Basin

YU Li-jun[1, 2], NIE Yue-ping[1, 2], YAO Yue-yin [2,3], LIU Fang[1, 2] ,Zhang Yan[4]

1. *Institute of Remote Sensing Applications of Chinese Academy of Sciences and Beijing Normal University, Beijing* 100101*, China;*
2. *Joint Laboratory of Remote Sensing Archaeology, Beijing* 100101*, China;*
3.*National Museum of China*
4.*Chongqing Cybercity Sci-tech co.Ltd*

Abstract: China has a long and glamorous history and many sites in China are of great value in archaeological research. The ancients heavily depended on location and environments due to the low productivity. The research analyzed the relationships between the archaeological sites and the environments. The 231 archaeologically important sites and 162 non-sites were taken as important references to deduce spectral and geomorphic characteristics such as geomorphology, water bodies and slope from multi-source data. These characteristics and terrain factors were correlated with the distribution of archaeological sites on which the predictive model was built. Then the predictive model correlated site locations with multiple environmental variables by logistic regression, and it was also used for Luoyang Basin where the model was proved to have high prediction accuracy. Especially it is likely to increase the efficiency and precision of predictive modeling by improvement of sample data collection and more collection of more thematic layers. The results indicate that altitude, slope and the distance to river were the mainly environmental parameters that influenced the distribution of archaeological sites. By comparing the predictive result with the location of known archaeological sites, the accuracy of model in Taihu Lake Basin can reach 87.5%. Future work includes site surveys in areas that have high potential if conditions allow. This predicting technology based on RS and GIS is cost-effective, flexible and timely, and more importantly it can be applied in vast regions, and reduce the disturbance and limit the damage to archaeological site as much as possible.

Key words: Remote Sensing, Archaeology, Taihu Lake Basin, Predictive Modeling,

---

Received: 空着 ; **Accepted**: 空着

Foundation: 1025 national science and technology support projects “Tracing Source Project of Chinese Civilization and Related Technology Research about the Protection of Cultural Relics” （2010BAK67B07）

First author biography: YU Lijun (1982-), female, Research assistant (Graduate school: the Institute of Remote Sensing application and researched. Research Area: remote sensing archaeology)

E-mail: yulj@irsa.ac.cn

Logistic regression analysis

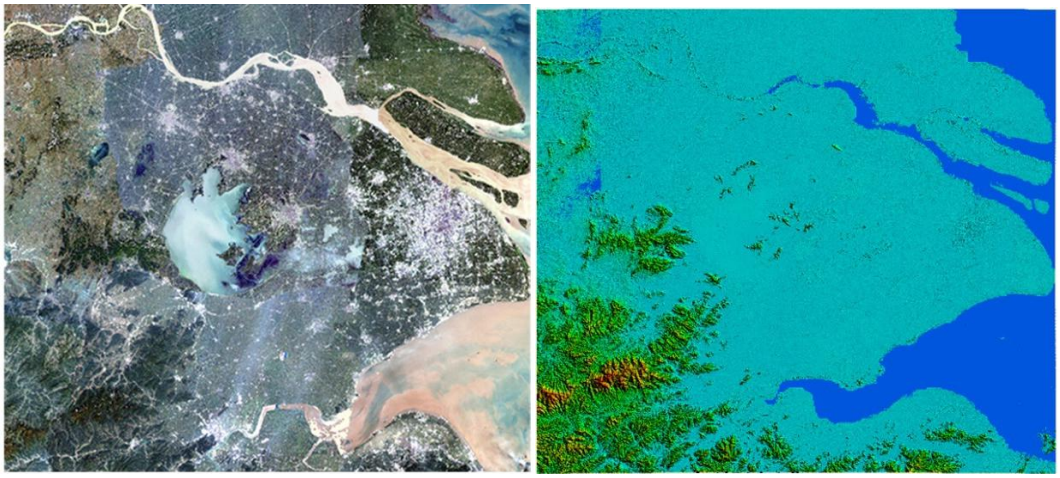
CLC number: （需查询中图分类号） **Document code:** A

1 INTRODUCTION

The objective of archaeological research has been gradually developed from simple description of artifacts to the study of human activities. The production and living activities of ancients has a close relationship with the natural environment, and the mode of living and distribution of settlements also has different characteristics in different geomorphological environment(Guo W M, 2008). The archaeological predictive modeling can be traced back to 1960s and the relationships between the development of human culture and environmental change have been one of the research hotspots in the environmental archaeology area. The development of sites is constrainted and impacted by nature and human factors to varying degrees. It's necessary to understand the specific natural and cultural environment around sites to study the distribution of sites. These factors plays different roles in sites(Yan W M, 1997 ; Williamsburg, Virginia,2009). Remote sensing images are of great value in documentation, which are the true reflection of geomorphology then. It's convenient to detect the buried sites from the images(Tang Z W, 2004). How to interpretate the available images properly is the key and prerequisite to find the location of sites. The relationship analysis between the archaeological target and environment comprehensively can provide a basis for the discovery and exploration of sites(Hans Kamermans, 2005). Logistic Regression can be used to identify areas of high archaeological potential(Sallie Vaughn,Tom Crawford, 2009; Jay F.Custer, 1986). The ancients settlement is affected by the natural environment and physiognomy, the predictive modelling can predict the presence of unknown archaeological sites according to the location characteristics of known sites in the survey area. Taihu Lake Basin is an important cultural center of middle and lower reaches of the Yangtze River basin. The paper employed the archaeological data and effective factors to build the archaeological predictive model to get the possibility map of the sites.

2 STUDY AREA AND DATA

Though there are thousands of important sites discovered in Taihu Lake basin, many more probably are undetected on the ground even buried underground due to natural calamities, war and other natural or human factors. Taihu Lake basin was selected as the typical research area. The most important is to obtain quality and quantity of archaeological data. The feature of sites distribution was extracted from ETM+, DEM data, river classification and geomorphology. The derivatives of DEM such as terrain, altitude, slope and aspect were available at ASTER GDEM with 30m resolution. According to the spatial distribution of the known sites in this area, archaeological predictive model was built on the basis of quantitative relationship between the sites and environmental factors.



(a) ETM of Taihu Lake Basin (b) DEM of Taihu Lake Basin

Fig. 1 Data Source

3 VARIABLE SELECTION

The natural environment is the most direct basis of settlement formation and development. The establishment of predictive model is based on the assumption that settlement location was influenced merely by environmental factors. Site location is greatly dependent on the natural environmental factors, such as terrain, slope, erosion, water, food and soil. The life style and adaptation to the environment can be hypothesized by understanding the interaction of morphological and ecological environment of the ancients. According to the available data, model establishment and the geographical features of Taihu lake basin, the following factors were chosen as main research factors: terrain, altitude, slope, aspect and the distance to the river. The independent variables were derived from remote sensing data including DEM data from ASTER imagery. The cell size of raster is 50m × 50m. There are 231 cells with site and 163 non-site cells among these grids which were selected as the samples. Variable values corresponding to these cells were obtained from the DEM and imagery.

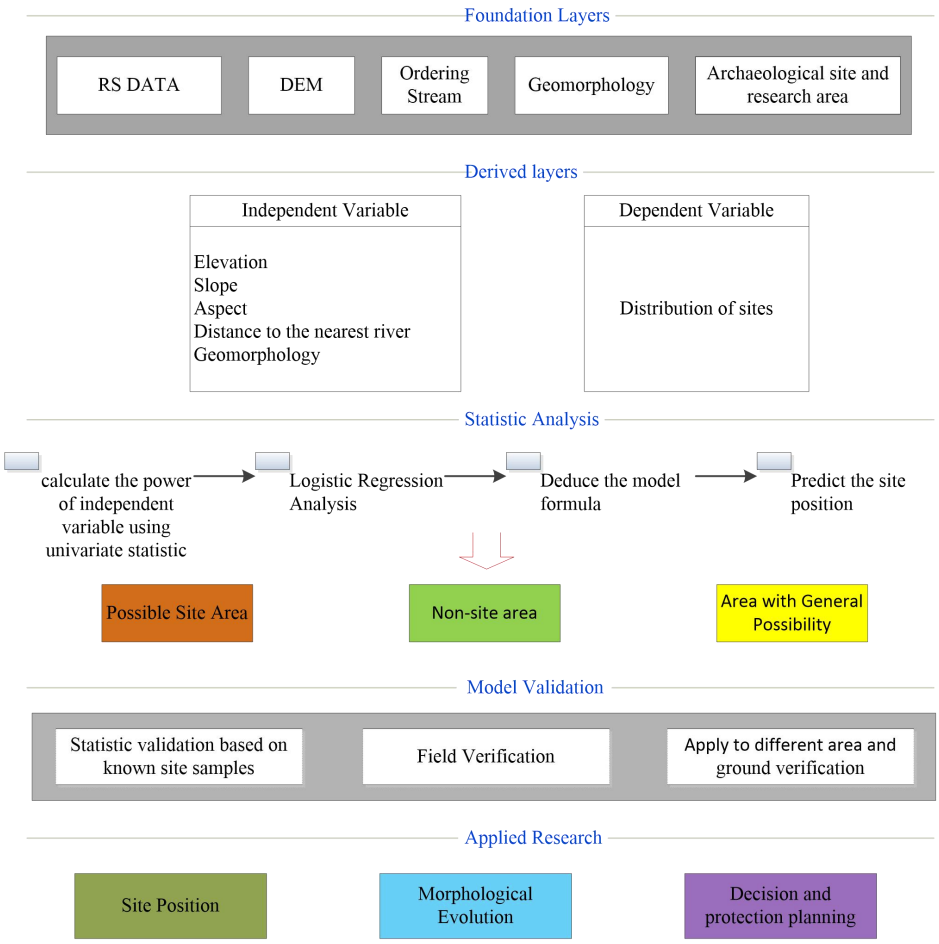


Fig. 2 Flow Chart

4 LOGISTIC REGRESSION ANALYSIS

Logistic regression is a powerful predictive method, which has been shown to be more robust than many other methods (kvamme,1990).The Binary Logistic Regression takes the following form (Sallie Vaughn,Tom Crawford,2009):

$$Probability(y) = 1/[1 + e^{-(b_0+b_1X_1+b_2X_2+b_3X_3.....-b_pX_p)}]$$

Where exp is the base of natural log system, b0 is the constant term; each b_1 to b_p represents the effect of the associated predictive variable. The probability of site presence ranges from 0-1.

According to the formula and the relationships between the sites and surrounding environmental variables, the archaeological potential model was established using the SPSS statistical software.

$$P = 1/[1 + e^{-(7.508-0.018X_1-0.464X_2-0.245X_3-0.01X_4)}]$$

Where x_1 to x_4 represent the environmental variables, which is altitude, slope, relief amplitude and distance to the river; P means the probability. And the probability map can be obtained according to this formula.

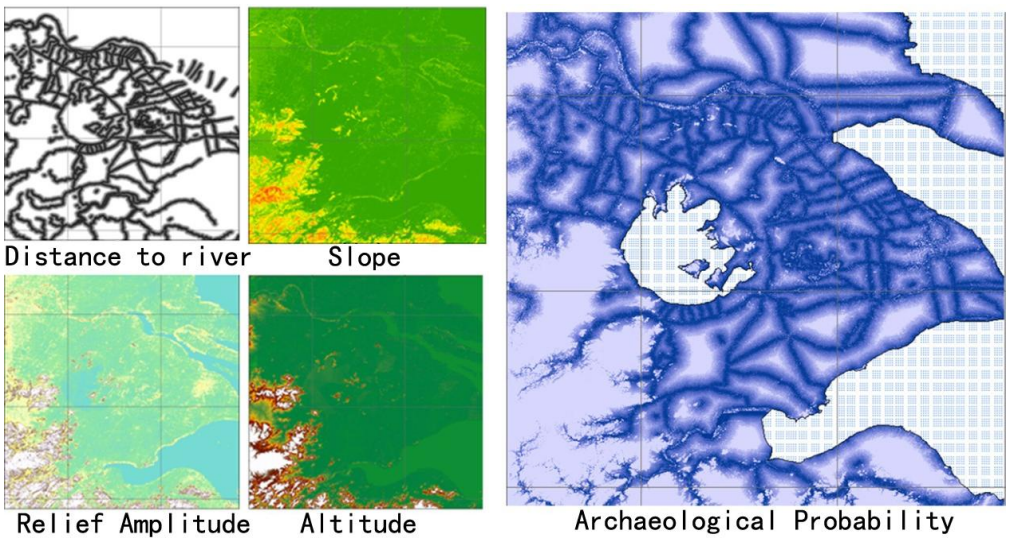


Fig. 3 Probability Map in Taihu Lake Basin

According to Fig.3, it can be seen that the distribution of sites has a specific direction. Important activities of human are subject to environmental resources. The closest environmental variable to the probability is water. The sites are often found in the terrace along the river bank where there is enough water without flood threat. By comparing the model results with the location of known archaeological sites, the accuracy of model in Taihu Lake Basin can reach 87.5% and conducted site surveys in areas with high probability.

5 Conclusion

The sites were widespread in Taihu Lake basin and some site areas overlapped in different periods. The predictive modeling involves remote sensing data, GIS and statistical analysis. The predictive modeling can identify the areas with high archaeological probability, improve efficiency and accuracy and give full play to RS and GIS in archaeology to a great extent. It will become an efficient tool for future field surveys, providing the scientific evidence for the site protection and urban development.

REFERENCES

Chen Z Y, Wang Z H.1998. Neolithic period in Taihu Basin: water and the rise and decline of the culture. Science, 50(2):29-32.陈中原,王张华.1998.太湖流域新石器时期：水与文化兴衰.科学,50(2):29-32

Guo W M. Huaxia Archaeology 郭伟民. 论聚落考古中的空间分析方法. 华夏考古,2008(4):142-150

Yan WM. The settlement archaeology and research of prehistory. Cultural Relics,1997,(6):29-37. 严文明. 聚落考古与史前社会研究.文物,1997,(6):29-37

汤卓炜. 环境考古学. 北京:科学出版社,2004. 277-278

Sallie Vaughn,Tom Crawford.2009.A predictive model of archaeological potential: An example from northwestern Belize. Applied Geography 29:542-555

Williamsburg, Virginia.2009. Archaeological predictive modeling in East Anglia and Norfolk. Computer Applications to Archaeology

Hans Kamermans.2005. Searching for tools to predict the past: the application of predictive modeling in archaeology. International Symposium 24.Kyoto.35-45

Jay F.C, Timothy E, Vytautas K, Ian W.1986. Application of Landsat data and synoptic remote sensing to predictive models for prehistoric archaeological sites: an example from the Delaware Coastal Plain. American Antiquity, 51(3):572-588.