Geostatistical and geoarchaeological study of Holocene floodplains and site distributions on the Sha-Ying River Basin, Central China

Hai Zhang1 | Wei Li2 | Andrew Bevan3 | Hui Wang4 | Fawei Liang5 | Yanpeng Cao5 | Yijie Zhuang3

1School of Archaeology and Museology, Peking University, Beijing, China
2Department of Cultural Heritage and Museology, Fudan University, Shanghai, China
3Institute of Archaeology, University College London, London, UK
4Institute of Archaeology, Chinese Academy of Social Sciences, Beijing, China
5Henan Provincial Institute of Culture Relics and Archaeology, Zhengzhou, China

Correspondence
Hai Zhang, School of Archaeology and Museology, Peking University, Beijing, China. Email: haizhang@pku.edu.cn
Yijie Zhuang, Institute of Archaeology, University College London, London, UK. Email: y.zhuang@ucl.ac.uk

Scientific editing by Michael Storozum

Funding information
National Nature Science Foundation of China, Grant/Award Number: 41971251; British Academy Newton Advanced Fellowship, Grant/Award Number: NAFR2180079

Abstract
Geostatistics has become a powerful method for investigating complex spatial variations of prehistoric settlements in floodplains and other geomorphological settings. A geoarchaeological drilling program that covers most of the Sha-Ying River Basin provides a rare opportunity with unusually detailed environmental data to contest and develop the geostatistics method, which proves to be essential, in combination with archaeological data, to understand long-term (9000–2500 B.P.) patterns of human inhabitation and adaption to volatile floodplain environments in eastern Central China. We analysed the variography and multivariate ordination of the borehole data and explored the complexities of landform evolution, with reference to sedimentation processes and soil development in the floodplain of the Sha-Ying River. The recurrent impact of river floods on regional landforms is manifested by spatial-autocorrelated properties over distances up to 10 km, sometimes with directional trends. We then developed a model of landform evolution through kriging and compared the model with detailed reconstruction of archaeological settlement patterns. Our results illustrate long-term socio-environmental dynamics by which human communities first populated and then adapted in diverse ways to the changing floodplain environments from the early to middle Holocene. This improved method will have far-reaching implications for future studies on similar geomorphological settings across vast floodplains of Central China and other global regions.

Keywords
geoarchaeology, geostatistics, Ordinary Kriging, paleo-floodplains, variograms

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. Geoarchaeology published by Wiley Periodicals LLC.
1 | INTRODUCTION

Machine drilling and systematic section sampling are amongst the most common methods applied in geoarchaeological field surveys. Alongside a reliable chronological framework, high-resolution geoarchaeological data acquired from such systematic surveys are vital to the study of environmental change, landscape evolution and human activities at multiple temporal and spatial scales (Branch, 2015; Canti, 2001; Luchsinger, 2008; Stafford, 1995; Weisler & Love, 2015). However, there are persisting methodological challenges with regard to how to better evaluate and integrate multiple environmental proxies that exhibit complex spatial-temporal variations for a more holistic understanding of long-term geomorphological processes and human-environmental dynamics (Boyer et al., 2006; Brown, 2002; Ferring, 2001; Stafford, 1995). It has been made clear that an uncritical use of geoarchaeological data derived from imprudent sampling strategies and limited sample sizes often leads to biased conclusions (Butzer, 2008). This problem is particularly relevant in the alluvial geoarchaeology of vast Holocene floodplains in eastern Central China, which saw the dramatic transformation and increasingly large-scale human occupation during the Holocene (Jing et al., 1997; Institute of Archaeology in Chinese Academy of Social Science IA-CASS, Peabody Museum of Archaeology and Ethnology in Harvard University PMAE-HU (2017); Wang et al., 2017; Zhang, Mark Pollard, et al., 2019).

Unlike the typical loess-alluvial landform around the Songshan region in the heartland of the Central Plains, which experienced multiple aggrading-incising cycles since the beginning of the Holocene (Lu et al., 2019, 2021), eastern Central China has been mainly a depressed floodplain with continuous alluvial aggradation along many tributary branches of the upper Huai River. The active alluvial history of the rivers in this region renders it extremely difficult, if not impossible, to obtain systematic geoarchaeological data as most archaeological sites and environmental sequences are covered very deep in later-period alluviums. This significantly hinders our understanding of prehistoric adaptations to such environments, which fostered the eventual emergence of visibly river-focused early Chinese Civilization.

Through the combination of statistical and environmental sciences, geostatistics has proved to be a powerful subfield of spatial analysis, which has further benefitted from the rapid development of GIS and other user-friendly software (for an overview of the application of geostatistics in archaeology, see Lloyd & Atkinson, 2020). Specific geostatistical methods such as variogram modelling and kriging interpolation are well-suited to evaluate the potential structure of data arising from sparse or potentially biased sampling strategies in complex and often compromised working environments in geoarchaeological studies (Kuzyakova et al., 2001; Poizot et al., 2006; Trangmar et al., 1986; Xie et al., 2008). Floodplains, due to their exceptionally productive biomass, have been attractive places for human inhabitation. However, as one of the most active fluvial landforms that are profoundly affected by volatile hydrodynamic and sedimentation regimes, floodplains also produce highly complex environmental records (Boyer et al., 2006; A. Brown, 1997; A. G. Brown, 2002; T. Brown, 2002; Ferring, 2001; Howard & Macklin, 1999; Li et al., 2021; Nanson & Croke, 1992; Storozum et al., 2020). Unpacking the relationship between the evolution of floodplain environments and the history of human occupation therefore faces enormous challenges resulting from the intrinsic complexity of environmental processes and data abstraction procedures. Such a challenge also provides, however, an opportunity to demonstrate the methodological robustness and sophistication of geostatistics in alluvial geoarchaeology.

The Holocene landscape along the Sha-Ying River Plain (SYRP) and nearby regions of eastern Central China is distinctive for its record of prolonged alluvial aggradation, which has made it notoriously difficult to achieve a clear understanding of long-term environmental–human dynamics in this deeply buried alluvial landscape. In this article, we synthesize geological and pedological data from a systematic geoarchaeological drilling project (Zhang, Li et al., 2019) that aimed to reconstruct the early to middle Holocene (ca. 9000–2500 B.P., according to the classification of Holocene in China based on calibrated dating by Fang et al., 2012, p. 109) environment on the SYRP. Both variograms and kriging were used to examine and reconstruct the spatial and temporal patterning of these environmental records. The results were then further integrated with detailed archaeological data on long-term changes in regional settlement patterns to further illustrate the effectiveness of the geostatistics method and help to establish the SYRP as an important region to understand dynamic civilizational discourses in prehistoric Central Plains of China.

2 | GEOSTATISTICAL ASSESSMENT OF THE EARLY-MIDDLE HOLOCENE SYRP LANDFORMS

2.1 | Summary of geostatistical methods

Geostatistics is one of the most popular methods for assessing datasets that arise from point-based sampling and the technique typically combines a first step involving variogram modelling and a second implementing a kriging interpolation (Cressie, 1990; Matheron, 1963; Webster & Oliver, 2007). Lloyd & Atkinson (2004, 2020, see also Conolly & Lake, 2006, chapter 6) presented a detailed review of the applications of geostatistics in archaeology, which include (1) estimating artefact or site density in field survey areas (Ebert, 2002; Markofsky & Bevan, 2012); (2) interpolating spatial data from environmental archaeology or geoarchaeology, such as pollens/phytoliths, soil phosphates or tephras (Athanasas et al., 2018; Buckley et al., 1988; Lancelotti et al., 2012; Oliver et al., 1997) and (3) evaluating the spatial structure of chronological data such as radiocarbon dates and/or the terminus ante quem of ancient architecture such as the Mayan monuments (Bocquet-Appel & Demars, 2000; Racimo et al., 2020; Whitley & Clark, 1985).

In variogram analysis, the experimental variogram can be estimated by calculating a semivariance, \( \gamma(h) \), as the half of the
expected squared difference between pairs of observations with the specified separation, which is usually termed a lag in distance:

\[ y(h) = \frac{1}{2p(h)} \sum_{n=1}^{p(h)} [Z(X_n) - Z(X_0 + h)]^2, \]

where there are \( p(h) \) paired observations of \( Z(X_n) \) and \( Z(X_0 + h) \).

The identification of directional variation can be calculated as

\[ SS(h) = \frac{y(h)}{\gamma_{\max}}, \]

where the \( \gamma_{\max} \) is the largest value of the variogram and the ratio for anisotropy can be termed as

\[ k(h) = \frac{y(h, \theta_1)}{y(h, \theta_2)}, \]

where the \( y(h, \theta_1), y(h, \theta_2) \) is the variogram in the direction of \( \theta_1, \theta_2 \).

Mathematical models (e.g., most commonly spherical, exponential, Gaussian or Matérn models) can be fitted to the experimental variogram and typically share three parameters, nugget, sill and range that can also be used further to evaluate the spatial structure of the observed evidence. Another useful parameter for evaluating spatial structures of these environmental data is the ratio of the nugget to the sill, \( C_0/(C + C_0) \). The nugget \( C_0 \) is a modelled variance at zero intersample distance and hence represents the importance of stochastic factors arising from sampling error or other random elements caused by river floods, waterlogged conditions and biological or human activities in limited areas. A large value of \( C_0 \) can also imply that there are some potential spatial patterns at very small distances that is inadequately modelled. The partial sill \( C \) represents deterministic factors in the model such as the structuring effect of the main river course, the lakes or bogs with large water areas, inundation of large-scale flooding and so forth. Therefore, the ratio of nugget \( C_0 \) to sill \( C + C_0 \) can be used to describe the degree of spatial variation, where a high ratio implies more random than deterministic factors and vice versa (Lloyd & Atkinson, 2004).

In practice, the geostatistical method has many advantages in the archaeology of palaeo-floodplains through expanding geological drilling data from isolated sampling points to a continuous surface represented by multiple sampling points and evaluating the spatial structures of landform evolution in terms of varied distances and directions. However, some disadvantages are also evident. First, the temporal variations can hardly be elucidated in the variogram models. Thus, sampling from typical sections with precise dating is indispensable. Second, the variables and parameters used in the models are sensitive. Careful pretreatment and evaluation prior to data analysis are therefore essential.

2.2 | Geoarchaeological fieldwork

The SYRP is located in the east of China’s Central Plains. The region sits to the south of the Songshan Mountains and east of the Funiu Mountains. Belonging to the upper Huai River system, the main streams of the Sha-Ying River meander their way across the region after flowing out of the mountain valleys (Figure 1). These alluvial processes have created a vast floodplain region with thick alluvial deposits and many small-sized lakes or bogs. The SYRP’s hydrological environment has been highly sensitive to climate change and subject to frequent fluctuations. The construction of embankments along the main river courses in recent historical times marked the continuous human effort to reduce the risk of river flood (Luohe Editorial Committee, 1999, p. 94). To better understand the Holocene history of human adaptations to the changing alluvial environment in the SYRP, a geoarchaeological drilling project was carried out, which covered an area of about 300 km² immediately east of modern Luohu City. A total of 361 boreholes were obtained between 2017 and 2018. Two hundred and twenty-seven of them were in a 1 × 1 km grid, whilst the remaining 134 boreholes were obtained at locations around Neolithic/Bronze Age sites or paleochannels (Figure 2). Most of the boreholes were machine-drilled by percussion rigs with a relatively small diameter (ca. 5 cm). These boreholes mostly reached 8 m deep below the modern ground level. But the drilling of some boreholes was stopped by pebbles or shallow bed rocks. Several AMS 14C and OSL dates from the bottoms of some boreholes show that this coring depth generally covers the entire Holocene epoch (Figure 3).

Each borehole was subject to careful observation and detailed recording to obtain three main types of information: sediment facies, degree of pedogenesis and anthropogenic inclusions. Sediment facies refer to colour, texture, grain size and sorting. Pedogenesis covers information that reflects post-depositional alterations of sediments, including clay content, organic carbon, Fe-Mn compounds or nodules, calcium carbonate, porosity and other information on plant rooting and/or bioturbation. Anthropogenic inclusions include charcoal, coprolites, bone fragment, shells and cultural remains such as potsherds, which can be considered indicators of land use activities. Based on such multivariate observations, each layer from each borehole can be categorized into one of the five types: a topsoil, an anthropogenic deposit, an alluvial deposit, a lacustrine deposit or a loess (Supporting Information: Table 1).

The reconstructed sedimentation sequences from these boreholes, combined with AMS 14C or OSL dates, demonstrate that early-to-middle Holocene archaeological sites are generally situated on four types of landforms in the region (Figure 3): (1) Loess terraces, which consist of wind-blown loess that was deposited during the Late-Pleistocene and subsequently incised by Holocene rivers into scattered hillocks within or on the edges of alluvial plains; (2) Drainage zones, which include those stable alluvial terraces consisting of alluvial deposits of brown or yellowish sandy or silty clays that were subject to post-depositional alterations (e.g., oxidation) on slightly higher elevations within the floodplain. Some parts on the tops of the cores have undergone a good degree of soil formation; (3) Seasonally flooded zones, which refer to the unstable alluvial terraces that experienced periodic changes in the sedimentation regime. In such locations, the pedogenic process was noticeably interrupted by some flood events; (4) Permanently inundated zones, which mainly refer to those lowlands that were more regularly waterlogged. Lacustrine deposits and black brownish or dark greyish
**FIGURE 1** Sha Ying River Plain (SYRP) in central China and the drilling area. (a) Central Plain in China; (b) SYRP in Central Plain of China; (c) drilling area in SYRP Jiahu sites: XLZ, Xielaozhuang; Yangshao sites: TCW, Tuchengwang; Longshan sites: BJT, Banjieta; FZ, Fuzhuang; MM, Mengmiao; Longshan to Erlitou sites: DX, Dengxiang; HJT, Haojiatai; SP, Shangpo, GL, Gouli).

**FIGURE 2** Layout of drilling boreholes in the Sha-Ying River Plain. (a) drilling points and archaeological sites; (b) drilling points classified by landforms.
horizons indicative of reducing environments were often encountered. Seasonal floods can also influence these areas and lead to the sudden and complete inundation of lakes and bogs and their surrounding areas; (5) Paleochannels that were abandoned permanently usually form slightly high-elevation natural levees. As to be illustrated below, geostatistical analysis demonstrates tremendous potential in uncovering the complex spatial-temporal variations of these sedimentary types and how were these related to the history of the Holocene human occupation.

2.3 | Data processing and validating the geostatistical method

The borehole data from the SYRP documented a wide range of highly variable temporal and spatial information on Holocene sedimentation history, and different statistical approaches might be applied to simplify this variability. As a first step, we summarized the observational information described in Section 2.2 into a more limited set of six landform properties for each borehole (Supporting Information: Table 2): (1) a flooding parameter (each of these parameters is bold, same below) summarizes the observed intensity of flooding events based on sedimentation cycles in the alluvial sediments. The sedimentation cycle refers to the time over which one complete cycle of alluvial accretion and incision occurs. The number of sedimentation cycles represents the frequency of river flooding at that location; (2) a sediment variable is calculated as the occurrence of characteristic sedimentary horizons weighted by the total depth of the drilling hole. The sedimentation sequence could be documented according to grain size change from coarser to finer. The number of sedimentary horizons for each borehole represents the degree of stability of the local sedimentation environment, which is influenced by both natural processes and human activities; (3) Loess is windblown sediments that are scattered within the floodplain. The loess in the SYRP was mainly deposited in the late Pleistocene and was eroded and re-deposited by rivers during the Holocene. The loess index is calculated as the total thickness of paleosol horizons divided by the number of layers; (4) Paleosol is summarized as an index for the thickness of paleosol horizons multiplied by the number of layers buried beneath the ground surface. Such paleosols were
normally considered to be formed in the past during periods that were warmer and wetter than today, which promoted vegetation growth; (5) **Channel** refers to the location where the paleo-river courses cut across the terrain. The sediments of these abandoned and buried channels could be easily identified in the field by the presence of the characteristic coarser sands, sandy silts and even pebbles. Channel index is the product of the thickness of channel deposits and the number of channels in one borehole; (6) **Lake** variable is the index of lacustrine deposits, which represent Holocene lakes or bogs in SYRP. Lake index also includes the calculation of total sediment thickness and the number of lacustrine horizons.

The above six numerical variables have been simplified from the original field records to identify multiple factors and agencies in floodplain evolution, such as overbank flooding, river diversion, erosion of loess deposits, expansion and extinction of lakes and bogs, vegetation conditions affected by both climate changes and biological activities, which can effectively represent the heterogeneity of geomorphological landforms of the Holocene SYRP. A cross-validation check further reveals no or only weak pairwise correlations between these variables (Figure 4, largest $R^2$ of 0.36 between Sediment and Flooding, and all others $R^2 < 0.2$) and therefore renders independent input variables. However, the latter has been log-transformed where necessary to avoid the problems of higher nugget and sill values, increased estimation errors, and nonstationary effects sometimes noted for non-normal geostatistical variables (Manchuk et al., 2009).

An additional methodological concern is with regard to an acceptable lag size. Because in the practice of plotting the semivariogram graph, the distances of paired points are not calculated individually, but instead grouped into bins, the choice of lag size (aka bin range) and therefore has significant effects on the empirical semivariogram. A very large lag size would mask the short-range spatial autocorrelation, whilst a very small lag size would reduce paired-point numbers in the bins and increase the local variations in modellings. While some rule-of-thumb approaches to choosing lag size have been suggested (e.g., in ArcGIS, the product of lag size and the lag number is half of the largest distance of all paired points), in this research, the sampling strategy presents a complicated case. At first, regular sampling of 1-km intervals was adopted. Then, to gain more insight into the geomorphological processes near archaeological sites and paleo-rivers, more densely distributed cores were obtained around these special locations (Figure 2). Because of this, it is difficult to use just one single lag size for this mixture of regular and nonregular sampling. To solve this problem, we

![Figure 4](image-url)  
**Figure 4** Cross-correlation plot of parameters on landform evolution in the Sha Ying River Plain.
introduced a series of lag sizes from 50 to 1000 by 50 m spacing for multiple semivariograms. This would also help us evaluate the spatial autocorrelation of the variables at multiple scales.

3 | RESULTS

In applying the analytical procedures described above, we first focused on exploring the empirical semivariogram for each of the six variables separately. This method allowed us to evaluate the intrinsic spatial, scalar and directional variations of landform evolution in the studied region. We anticipated a priori that there should be both deterministic and stochastic factors in landform evolution on the floodplain and we tested these factors with our results.

3.1 | Basic variography

To explore the relative structure of different variables, we swept different model types (spherical, exponential, Gaussian and Matérn) and lag sizes, before choosing the model with the best fit (smallest sum of squared errors or highest R²). In the semivariogram plot, the fitted range parameter can be regarded as expressing the extent of spatial autocorrelation in that particular variable (Supporting Information: Tables 3 and 4). Figure 5 illustrates the ranges of best-fitted models under multi-lag size spacing by 50 m, which is the cell size of the kriging interpolation raster in the following analysis. The plots above and below the dashed lines are either off the setting extent of 20 times of lag sizes or shorter than the lag size itself, which should be omitted. It is evident that lag size has a significant impact on the model structure and multi-scalar spatial autocorrelation can be detected on three variables, flooding, sediment and loess. The ranges of flooding are concentrated on three levels from the small (<2 km, lag size < 500 m) to the middle (5–6 km, lag size: 500–800 m) and the large (10 km, lag size: 800–1000 m) scales, implying that flooding as an agent of environmental change is constantly active on the floodplain over fairly large areas. Sediment has two levels of structure in the space of intermediate to large ranges (2–4 km, lag size 250–650 m; 7 km, lag size: 700–1000 m) but the range is a little shorter than that of flooding, which implies that the sedimentation process in the floodplain has been affected by more factors than just the flooding. The spatial structure of loess is mainly fitted at ranges of 3–4 km with a lag size of 600–1000 m, which exhibits a different pattern compared to the similar middle-ranged sediment parameter with shorter lag sizes. A long-distance range of 10 km for loess can be occasionally observed on the lag of 700 m, which was due to the two mass blocks of loess terraces located on the north and south of the study region.

In contrast, the other three variables of paleosol, channel and lake only show short ranges of less than 2 km despite different lag sizes, indicating that correlated variability in soil development, waterlogged conditions and channel shifting (mainly in wetlands) does exist in short distances, but that spatial changes are fairly frequent over larger areas. It should be noted that the typical spacing distance between archaeological sites is less than 3 km, suggesting that variations in paleosol distribution and groundwater conditions do exist within any single site catchment of the region.

In the analysed data, the channel shows a very small nugget–sill ratio (0.3%), indicating some strong spatial autocorrelation among over-flooding and channel shifting factors as a deterministic component but only in the extremely short range (with a range of 463 m × 50 m lag size). Flooding has a relatively smaller nugget (ratio <50%) from short to long spatial autocorrelation (0.2–10 km), which indicates that 10 km is the maximum extent through which river floods could reach. However, sediment on shorter spatial autocorrelation exhibits a larger nugget–sill ratio (>50%), where stochastic components account for more than half of the total variability, implying that more localized sedimentary conditions, perhaps

FIGURE 5 Lag-Range plot of best-fitted models for isotropic variograms for landform evolution parameters in the Sha Ying River Plain.
including geological, biological, pedological and anthropogenic factors, operate within a range between 2 and 4 km. It is noticed that a large nugget effect (the ratio is 70.9%) exists on paleosol but with short-range distance of 1.4 km. The large lag size (800 m) indicates that it was a measurement error.

3.2 | Anisotropic variography

Turning to the additional insight that might be gained when we switch from omni-directional to anisotropic variograms in our analysis (Supporting Information: Table 4), it is evident from Figure 6a that the flooding process forms multiscale patterns from short to long distances but in different directions: the short ranges less than 2 km are on N-S (0); the middle ranges between 3 and 5 km occur on the directions of NE-SW (45) and E-W (90). This clockwise increase of spatial autocorrelation might be related to the directions of the Holocene channel and overbank flows, which affected the landform evolution in the SYRP. Since spatial variation is more frequent in N-S (0) and NW-SE (135) directions, these are probably the prevailing directions of the overbank floods. In contrast, longer distance spatial autocorrelation in the NE-SW (45) and E-W (90) directions helped us to identify the main river channel in the study area. Figure 6b shows the sediment parameter that demonstrates similar spatial patterns with flooding but with much shorter ranges for each direction. Ranges reaching as long as 10 km can be occasionally detected in E-W (90), which is close to that of the flooding. This means that the sedimentation process is mainly controlled by the agent of flooding but may also be affected by some other local factors, such as vegetation and human activities.

With the flooding and sediment factors being the stable and basic factors underpinning landform evolution on the floodplain, other geological, biological, pedological and anthropogenic factors operate in smaller-scale and secondary ways. The directional variation of loess (Figure 6c) is roughly in the north–south direction in the middle-scaled ranges. The channel (Figure 6e) shows similar directional variation with the short-ranged flooding (0 and 135) along the cross-sectional erosion of the main rivers, indicating that the surface flow is the main agent among the complicated hydrodynamic factors in the floodplain. E-W (90) is obviously the prevailing direction of the lake (Figure 6f) at the intermediate range scale between 3 and 4 km. This is also the main direction of river water flow, which indicates that the evolution of Holocene lakes, including oxbow lakes and back swamps, in the floodplain is mainly controlled by rivers. For the parameter of paleosol (Figure 6d), it has much greater spatial continuity in the E-W direction (90) than NW-SE (135), which means a more stable state of soil development.

3.3 | Kriging interpolation

Based on the best-fit models described above, we then used ordinary kriging (OK, Cressie 1990) to interpolate landform properties across the study area. Figure 7 shows the interpolated surfaces for the six parameters. Two kinds of lag-sized models with different ranges were chosen for each parameter to interpolate on both local and global spatial

**FIGURE 6** Lag-Range plot of best-fitted models for anisotropic variograms for landform evolution parameters in the Sha-Ying River Plain. (a) flooding; (b) sediment; (c) loess; (d) paleosol; (e) channel; (f) lake.
structures (Supporting Information: Figure 1). Prediction variance maps were calculated at the same time, which shows that areas lacking sampling points such as the mid-west of the maps have large prediction errors and will be erased in the following analysis. Figure 8 provides the density plots of the interpolated surfaces containing all the raster values. In general, the two surfaces interpolated from the same parameter present similar density curves in spite of some minor differences. This indicates that the overall landform structure of the floodplain in the SRYP can be estimated by kriging effectively regardless of the scales of measurement. If getting insight, it is notable that flooding shows three peak density values, which might be considered as three levels of hydrodynamic agents in floodplain formation, which is consistent with the three groups of range distance shown in the best-fit models of semivariograms. These three hydrological levels in the Holocene SRYP warrant further scrutiny with the boreholes and sampling sections with typical geoarchaeological methods aided with more accurate dating in the future. In contrast, the plot of sediment presents a relatively even distribution. Since the sediment mainly represents the changing frequency of sedimentation layers, it can imply the diversity of the depositional process within the floodplain. It is not surprising that the channel has a skewed distribution in that it only covers limited areas in the floodplain. The other three parameters, paleosol, lake and loess, exhibit a roughly normal distribution, suggesting the widespread soil development, waterlogged environments of lakes/bogs and loess terraces in the SRYP.

To summarize, the six parameters demonstrate spatial autocorrelation at different scales in the evolution of diverse landforms on the floodplain of the SRYP, with high-frequency variation across the floodplain associated with river flooding, changes in waterlogged condition and soil formation and inevitably leading to a rather complex, ecologically diversified environment in which human settlement and subsistence activities might take place. Furthermore, the kriging interpolation makes it possible for us to evaluate the geographical coverage of the six factors across the floodplain and thereby further reveals spatial heterogeneity in the evolution of the floodplain over the course of the early-middle Holocene.

4 | DISCUSSION

The evolution of alluvial floodplains had a profound impact on prehistoric human inhabitation worldwide, stimulating diverse cultural adaptations to dynamic floodplain environments. Interactions between the evolving landforms and prehistoric occupations on floodplains are complex and multidirectional. Previous archaeological surveys, including the third nationwide census of cultural relics and
many recent rescue excavations, have documented 26 early to middle Holocene sites in the SYRP region (School of Archaeology and Museology, Peking University SAMPU, Henan Provincial Institute of Cultural Relics and Archaeology HPI CRA and Luohe Municipal Cultural Affairs Bureau LMCAB (2017); Zhang, Li et al., 2019). Our own drilling survey re-examined these data carefully. Through a combination of pottery typologies and new radiocarbon dates from samples acquired from several archaeological sites, the archaeological remains from these sites can be dated to different periods, as shown in Table 1.

There are some gaps in the Holocene occupation of the SYRP, notably for the 7500–6000 B.P. and 3000–2700 B.P. periods. Beyond these gaps, demographic fluctuations in the region can be inferred from the changes in settlement numbers. For example, the beginning of the Longshan Period at ~4500 B.P. saw a pronounced increase in settlement number and size on the floodplains, followed by rapid population decline in the earlier Bronze Age (Table 1).

Another settlement boom occurred in the late Bronze Age during the East Zhou period (2700–2500 B.P.). The abandonment of sites and decrease in the population on the floodplains of the region were related to climate change and population movement (Fan et al., 2011; Mao et al., 2016; Mo et al., 2003; Nesje & Dahl 1993; Ren et al., 2006; Shi et al., 1993; C. Zhang, 2011). The temporal fluctuation of settlements on the floodplains indicates that human occupation at different times is likely to be spatially heterogeneous, which can be examined by considering the site distribution and its relationship with landform evolution, which we further investigate below.

Figure 9 shows a pair correlation function (PCF) of early-to-middle Holocene archaeological sites across the floodplains, along with a critical envelope for significance levels based on random simulation (for the technique, see Baddeley et al., 2016; and for its application in archaeology, see Bevan, 2020). In general, despite the obvious potential first-order heterogeneity in the landscape, these sites are distributed in a random pattern when treated as a single

<table>
<thead>
<tr>
<th>Sequence</th>
<th>C14 date (B.P.)</th>
<th>Culture name</th>
<th>Archaeological Phase</th>
<th>Site numbers</th>
<th>Total site areas (m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9000–7500</td>
<td>Jiahu</td>
<td>Early to middle Neolithic</td>
<td>2</td>
<td>125,000</td>
</tr>
<tr>
<td>2</td>
<td>6000–4500</td>
<td>Yangshao</td>
<td>Late Neolithic</td>
<td>4</td>
<td>220,000</td>
</tr>
<tr>
<td>3</td>
<td>4500–3800</td>
<td>Longshan</td>
<td>Chalcolithic Age</td>
<td>19</td>
<td>923,500</td>
</tr>
<tr>
<td>4</td>
<td>3800–3600</td>
<td>Erlitou</td>
<td>Early Bronze Age</td>
<td>7</td>
<td>346,500</td>
</tr>
<tr>
<td>5</td>
<td>3600–3000</td>
<td>Shang</td>
<td>Middle Bronze Age</td>
<td>9</td>
<td>393,500</td>
</tr>
<tr>
<td>6</td>
<td>2700–2500</td>
<td>Eastern Zhou</td>
<td>Late Bronze Age</td>
<td>11</td>
<td>549,000</td>
</tr>
</tbody>
</table>

FIGURE 8 Density plot of OK prediction map values of the landform evolution parameters in the Sha-Ying River Plain.

TABLE 1 Archaeological sites of the Mid-Holocene period in the Sha-Ying River Plain.
chronologically undifferentiated group, which implies that there are probably no clear spots with rich ecological resources attractive to prehistoric occupation within the study area. However, the typical spacing between contemporary sites is usually ca. 1–3 km in the region, which is the same or a little smaller than the smaller-scale variation exhibited by landforms. This suggests an unstable and frequently changing floodplain landscape and hence considerable diversity in settlement strategies in the region. Figure 10 shows the boxplots of kriging values for archaeological sites within 1 h catchment areas (by r.walk in GRASS GIS, Supporting Information: Figure 2), grouped by chronological periods. The basic raster interpolated maps are also summarized and plotted as background. In this way, the exploitation of the floodplain in different periods can be scrutinized.

The earliest known occupation on the floodplain comprises two Jiahu period settlements (9000–7500 B.P.), and these fall on locations with lakes and bogs and relatively stable hydrological conditions far away from the seasonal flooding zone. Loess terraces were obviously not the preferred locations for the Jiahu period sites. While the sample size is too small to allow for any further statistical support of this suggestion, it nonetheless raises the possibility that Jiahu communities were prioritizing particular resources in limited areas, such as around the lakes or bogs on the broad floodplain environment for their early cultivation of rice and hunting–gathering–fishing activities (Wu et al., 2015; J. Zhang et al., 2018; Zhao & Zhang, 2009).

After a gap of more than a thousand years, new Yangshao settlements (6000–4500 B.P.) appeared across a wider range of alluvial landforms in the region (Cheng, 2016): some are found on the edges of alluvial plains and loess terraces, which were unlikely to be flooded. Although the Yangshao settlements seem to continue to prioritize places with stable hydrological regimes like Jiahu, more of them settled on the loess terraces and higher positions in the region. This new settlement strategy was consistent with the emphasis on millet cultivation among Yangshao culture communities (Larson et al., 2014; Lu et al., 2019).

In the following Longshan period (4500–3800 B.P.), there was a sharp increase in settlement numbers, and it appeared that the entire

![Figure 9](image)

**FIGURE 9** Pair correlation function analysis on the distribution of archaeological sites in the Sha-Ying River Plain.

![Figure 10](image)

**FIGURE 10** Density plot of ordinary kriging prediction map values within 1 h catchment of archaeological sites.
region, especially the low-lying floodplains, had been almost fully exploited, even including some places that had more dramatic hydrological fluctuations as indicated by higher flooding conditions and where the degree of soil development was very low. Millet cultivation and hunting–gathering–fishing activities were adopted by some Longshan period sites (Li et al., 2021; Zhou 2017). This change might reflect the fact that under a much drier and cooler climate regime around 4200 B.P., widely recognized as the so-called 4.2 kaB.P. event (He et al., 2022; Railback et al., 2018; Ran & Chen, 2019; Ren et al., 2006; Weiss, 2016), the Longshan communities began to explore a diverse range of habitats, including many low-lying floodplain environments, which would have also provided much-needed water and natural resources despite the risk of being flooded periodically. This regional settlement pattern continued into the early Bronze Age (3800–3600 B.P.) but with a noticeable decline in settlement numbers, not least on the loess tablelands in the region, which may be due to political reasons that triggered population migration to the Luoyang basin some 170 km to the northwest with the rise of the Erlitou centre in Luoyang (H. Zhang, 2021; Zhang, Li, et al., 2019). It is also important to note that some other external drivers might have also contributed to this process, such as the introduction of domesticated cattle and sheep into the Central Plain of China during the Longshan period (Yuan, 2019) as new agencies for the expanding exploitation of new environments and resources in the floodplain. The cultivation of wheat and barley, which began soon after the Longshan period, would have also required different ecological settings, especially irrigation, and thus responsible for the aforementioned settlement change.

Interestingly, during the Shang period (3600–3000 B.P.), many sites started to dwell directly on the floodplains. Most of the Shang period sites were situated in the southern part of the region, but the reasons for this change are still unclear. The Western Zhou period (3000–2700 B.P.) marked another gap in the history of floodplain occupation in the region, but such a vacuum was quickly filled up in the succeeding Eastern Zhou (2700–2500 B.P.) with a pronounced expansion onto even the high loess terrace. Indeed, the largest site, Shaoling Gucheng, which is thought to have been a centre within a vassal state, was built on the top of the largest loess terrace (Shaoling Gang).

5 | CONCLUDING REMARKS

By synthesizing the multivariate evidence obtained from a geoarchaeological drilling survey and archaeological excavations in the SYRP of eastern Central China, we identified six main variables that reflect predominantly conditions of flooding, sedimentation, formation of lakes/bogs, river channels, loess terraces and soil development. A geostatistical approach was developed to identify the spatial structure of these variables across the floodplain environment. Indeed, on a floodplain such as the SYRP that has experienced continuous and sometimes rapid accretion during the Holocene, a geostatistical analysis based on a systematic sampling strategy is crucial for a more reliable reconstruction of the paleo-environment and more holistic understanding of the environmental–human dynamics. The spatial heterogeneity of landform evolution on the floodplain can be represented via two aspects: distance and direction. The different ranges of the best-fitted models in variograms indicate that the spatial autocorrelation of flooding, sediment and loess variables can be detected at large and medium scales, while those of the paleosol, channel and lake variables were smaller. The difference suggests that slightly higher-ground areas on the floodplain underwent more frequent spatial variations due to more dramatic hydrological fluctuations (e.g., seasonal overbank flow) and their evident impact on soil development, while by contrast, the environment on the loess terraces was more stable. The anisotropy within the floodplain was mainly reflected in the flooding variable, which may indicate the prevailing direction of the main river course and its overbank flooding.

The interpolation of these variables into continuous surfaces via ordinary kriging made it possible to evaluate broader patterns and trends of landform evolution within the diverse floodplain environments of the region. These enabled us to better integrate the environmental and archaeological data for a deeper and more detailed reconstruction of Holocene human adaptations to the region’s volatile floodplain environments. As revealed in our analysis, in the SYRP region, the Holocene environment was dominated by prolonged alluvial aggradation with frequently changing sedimentation regimes and volatile hydrological conditions. Compared to the heartland of the Central Plains, such as the Songshan region, which enjoyed an overall stable Holocene environment and a celebrated continuity of cultural developments, the SYRP region remained a ‘peripheral region’ of the Central Plain civilization for a long time, characterized distinctively by two pronounced hiatuses in the regional occupation history. The prehistoric settlements did not display an evident regional stratification before the late Neolithic period, and their diverse subsistence strategies had been closely related to changing environmental and ecological conditions. It was not until the Eastern Zhou that, as discussed above, a relatively large urban centre occurred in the region. Such volatility in both regional environmental and settlement distributions profoundly defined the model of prehistoric adaptation of the region.

Our geostatistical analysis of the SYRP region contributes an important case to understand the intra- and inter-regional differences and diverse discourses of social development on the Central Plain of China. This geostatistical method has great potential for the geoarchaeological investigation of paleo-floodplains in other regions. However, as shown in this research, since there are also some pitfalls in the pure mathematic models with sensitive parameters such as lag size, careful spatial designs of drilling and recording are essential in successful geostatistical studies. Equally, accurate dating and a robust sampling strategy on representative locations and points during geoarchaeological surveys remain fundamental and irreplaceable.
ACKNOWLEDGEMENTS

The authors are grateful to the Luohu Municipal Institute of Cultural Relics and Archaeology and the Zhumadian Municipal Institute of Cultural Heritage and Archaeology. Financial support for the project has come from the National Natural Science Foundation of China (41971251) and a British Academy Newton Advanced Fellowship (NAFR2180079).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

ORCID

Hai Zhang http://orcid.org/0000-0001-8237-5057
Yijie Zhuang http://orcid.org/0000-0001-5546-0870

REFERENCES


**SUPPORTING INFORMATION**

Additional supporting information can be found online in the Supporting Information section at the end of this article.