Vegueta. Anuario de la Facultad de Geografía e Historia
23 (1), 2023, 57-95
eISSN: 2341-1112
https://doi.org/10.51349/veg.2023.1.03

Modelling Bias and Environmental Preferences in Archaeological Spatial Analysis

Sesgo de modelado y preferencias ambientales en el análisis espacial arqueológico

Michael Kempf*
University of Kiel, Germany
Department of Geography, Physical Geography
- Landscape Ecology and Geoinformation
https://orcid.org/0000-0002-9474-4670
kempf@geographie.uni-kiel.de

Eli J. S. Weaverdyck
University of Freiburg, Germany
Beyond the Silk Road (BaSaR). Seminar für Alte Geschichte
https://orcid.org/0000-0002-9474-4670
eli.weaverdyck@geschichte.uni-freiburg.de

Recibido: 07/10/2022; Revisado: 20/12/2022; Aceptado: 22/12/2022

Abstract

Point pattern analysis (PPA) has gained momentum in archaeological research that models large-scale distributions of sites and explanatory covariates. As such, there has been increased interest in the bias of archaeological distributions, which mostly have an impact due to modern land-use change. These interactions, however, have not yet been fully explored. In order to better understand archaeological point patterns as functions of explanatory covariates, we offer three different approaches: (i) environmental preference modelling of archaeological records in different chronological phases; (ii) a custom bias surface that represents the variability of the regional landscape; (iii) an R-package (rbias) allowing the generation of a fuzzified bias surface based on Open Street Map (OSM) data.

Keywords: Quantitative Archaeology, Spatial Modelling, Environmental Archaeology, Landscape Archaeology, Roman.

* Corresponding author / *Autor de correspondencia.*

Resumen

El análisis de patrones de puntos (PPA) ha cobrado impulso en la investigación arqueológica en el modelado de las distribuciones a gran escala de sitios y las covariables explicativas. Se ha puesto más interés en el sesgo de las distribuciones arqueológicas, que en su mayoría impactan por el cambio moderno en el uso de la tierra. Estas interacciones, sin embargo, aún no se han explorado completamente. Presentamos tres enfoques diferentes para comprender los patrones de puntos arqueológicos como funciones de covariables explicativas: (i) modelado de preferencias ambientales de registros arqueológicos en diferentes fases cronológicas; (ii) una superficie de sesgo personalizada que representa la variabilidad del paisaje regional; (iii) un paquete R (rbias) que permite generar una superficie de polarización difusa basada en datos de Open Street Map (OSM).

Palabras clave: Arqueología cuantitativa, Modelización espacial, Arqueología ambiental, Arqueología del paisaje, Romano.

1. INTRODUCTION

Quantitative, digital statistics, and spatial analyses are common tools in archaeological research that focuses on the recognition of patterns in past societies' behavior and particularly settlement distribution and land-use (Brandolini and Carrer, 2021; Carrero-Pazos *et al.*, 2019; Gillings *et al* ., 2020; Kempf, 2020d, 2021; Kempf and Günther, 2023; Verhagen and Whitley, 2012). In this context, the use of point pattern analysis has become increasingly important to understand the spatio-temporal components behind the socio-cultural, political, and ecological driving factors that control individual and group decision-making processes (Bevan, 2012; Crema *et al.*, 2010). Eventually, the immediate landscape affordances, which comprise the potential and actual opportunities and propositions offered by the socio-ecological setting of the locale and the human being interacting with them in the moment of mutual confrontation, has entered the discussion (Gibson, 1979; Kempf, 2020b, 2020c; Knappett, 2004). If affordances emerge from a confrontation between a strategic human actor and an environmental feature with particular qualities, statistically identified spatial associations between archaeological sites and environmental features can be interpreted as evidence for the existence of such affordances, and therefore for the social context that shaped the human side of the human-environment confrontation.

However, a significant methodological problem hinders the use of this type of analysis for understanding the past: bias in the distribution of archaeological sites caused by formation processes. Formation processes include the processes that mediate the deposition of material traces by ancient humans, the transformation of those traces after deposition, and the discovery of those traces in the present (Schiffer, 1996). These processes all but guarantee that the sites we record in the present are not a random sample of the activity areas that constituted ancient landscapes. The problem is even more acute when we base our analyses on legacy data. When working with old survey data, one can consider the survey methodology and test for the resulting biases (Casarotto *et al.*, 2018; Purtill, 2022).

The problem becomes more complicated when analyzing data that was not collected as part of a single project. Many countries now maintain national databases of archaeological sites. This data has usually accumulated over many years in a piecemeal fashion with no consistent collection methodology (Cowley, 2016; Kempf, 2021; Kreiter, 2021; van Leusen, 1996). This poses a major challenge for researchers because it is very difficult to control for the biases that must necessarily shape the distribution of sites in these databases. At the same time, these databases often include huge numbers of sites over very large areas, far in excess of what could be obtained by a single survey. Although problematic, the scale of these datasets means that they have the potential to provide evidence for much larger groups of people and larger-scale social processes than any individual survey project ever could. To realize that potential, we must devise ways of identifying and controlling for the biases that plague them. This article offers a way of identifying and controlling for biases arising from one type of formation process: discovery. It is not a complete solution to the problem of bias in large, cumulative, legacy data sets, but we think that it proves a useful tool upon which others can build. We provide two different approaches to bias detection in a regional case study in eastern France (Alsace). Focusing on the transitions from pre-Roman to the Roman period and to post-Roman chronology, we assemble a large number of archaeological sites of various origin to measure environmental landscape transformation processes as well as the impact of modern land-use and landcover change. We build on the statistical methods described in Kempf (2020) and (2021) and Kempf and Günther (2023) and test the infrastructural bias by applying the recently developed R-package *rbias*, which ultimately is designed for this purpose (Günther *et al.*, 2022; Kempf and Günther, 2023).

2. MATERIAL AND METHODS

We begin from the premise that it is not actually possible to "unbias" the archaeological record through data manipulation (van Leusen, 1996). Instead of trying to manufacture an archaeological dataset that is a representative sample of the ancient landscape, we manufacture a comparative dataset that is biased in the same ways as the archaeological record. Statistical identification of associations between archaeological sites and environmental features relies on a null hypothesis that the archaeological sites are randomly distributed with respect to those features. But this is a proxy for the null hypothesis that scientists interested in past behavior actually want to test: that the (ancient) places that produced the (modern) archaeological sites were randomly distributed with respect to those features. Cultural heritage management, of course, is more focused on the archaeological sites than the ancient places that produced them. This mirrors the distinction between predictive modeling (for cultural heritage) and location analysis (for historical research) (Verhagen, 2007; Verhagen *et al.*, 2010; Verhagen, 2018; Verhagen and Whitley, 2012; Verhagen and Whitley, 2020). A biased comparison dataset represents a situation in which ancient places exhibited complete spatial randomness, but then went through the same set of formation processes that lie behind the actual archaeological data under investigation. Here we use one, very large comparison dataset because it makes it easier to compare variables with the environmental data, but many smaller comparison data sets could also be used. The key insight we propose that constructing a comparison dataset allows one to control for trends arising from the process of discovery and therefore more confidently identify trends arising from ancient behavior.

2.1. Pattern detection, scale, and inherent uncertainties

Pattern analysis becomes particularly useful if one aims at detecting development of human behavior in the landscape. One of the basic considerations in this approach is the *a priori* assumption of change over time, which can be attributed to changing environmental conditions, societal development, migration and mobility, or technological enhancement, innovation, and transformation of a socio-cultural system. A second pillar is the spatial extent of the study area and the scale of human-landscape interaction. The question of scale is twofold. First, the study area is limited by its subjective extent (e.g. a rectangle or modern administrative boundaries, mostly considered as *bounding box* or *window of operation*). which all too often ignores the large-scale ecological feedbacks and the supraregional administrative and political dynamics of cultural relations. Second, the individual scale of the human activity spheres, which determines the range of agricultural intensification or extensification, livestock breeding, market orientation, settlement dispersal, and general communication networks and exchange patterns.

The scale of human activity spheres depends on the life ways being practiced in general and on the particular activity under investigation. Subsistence strategies are particularly influential, and spatial analysis can provide evidence for which strategy was more common. Under an intensive agricultural regime, in which the goal was to maximize production through the investment of labor and other resources in cultivation, people should maximize their access to the environmental affordance of "fertile land". The distance between settlement and field constrains the amount of labor one is able to invest in cultivation. Through cross-cultural ethnographic comparison, Chisholm showed that the amount of labor invested in fields declines sharply somewhere between one and two kilometers from the settlement (Chisholm, 1979 (2007)). Therefore, maximizing access to fertile land means maximizing the amount of fertile land within one to two kilometers of the settlement. Different approaches, however, have recently been emphasized for the construction of complementary regions around settlements, using various distance-based relationships. Most of them, however, were building on plain surfaces instead of integrating permeability or accessibility into the cost-distance-expenditure calculation (Brönnimann *et al*., 2018; Depaermentier *et al*., 2020; Depaermentier *et al*., 2021; Kempf, 2020d). Under an extensive agricultural regime, the spatial relationship between the settlement and the field is less constraining, since less labor is invested. Extensive agriculturalists might still opt to live in places with fertile land nearby, but more distant places would be considered "close enough". Therefore, analysis of the immediate surroundings of settlements and how these change through time has the potential to yield evidence for agricultural intensification.

Disentangling the different types of spatial patterning requires the creation of a comparison dataset subject to the same biases as the archaeological record. This dataset consists of a weighted distribution of random points, with the likelihood of any given location receiving a point dependent on the biasing factors present. This comparison dataset represents the null hypothesis that behavior in the past was spatially random and that the traces left by that behavior have been subject to the formation processes affecting the observed archaeological record. Significant differences between this comparison dataset and a subset of the archaeological record (sites identified as Roman-period settlements, for example) are more likely to reflect real trends in ancient behavior than post-facto trends in formation processes.

There is a danger that certain behavioral trends will be obscured using this method when factors that affect the archaeological record as a whole also impacted ancient behavior. Forested areas, for example, might both obscure archaeological remains in the present and be less-densely populated in the past. Building an aversion to forested areas into the comparison dataset will make it impossible to identify a real aversion to these areas in the past. The results of analysis using this method will, necessarily, be incomplete. On the other hand, this could be seen as an advantage. When comparing chronologically and typologically specific subsets of the archaeological record (e.g., pre-Roman, Roman, and post-Roman rural settlements) to a dataset reflecting tendencies that are common to the entire archaeological record, trends specific to the subset under consideration will be emphasized, making it easier to track diachronic change. In any case, by distinguishing between spatial patterns attributable to formation processes and those that are not, we strengthen the empirical basis for our interpretations of spatial patterning in the archaeological record.

2.2. Environmental settings of the study area

The study area covers a section of the French Upper Rhine Area (URA) from its southern border to 48.5° N and stretches over the current administrative *Region Grand Est* (formerly *Region Alsace*, Département Haut-Rhin (68) and Département Bas-Rhin (67)) (Fig. 1). It measures approximately 120 km north to south and around 50 km east to west, covering nearly 490,000 ha. The URA is characterized by broad agricultural cropland including vineyards, orchards, and increasingly monoculture maize cultivation. The particular topographic situation between the Vosges mountain range and the Black Forest supports low precipitation rates and a high drought and flooding vulnerability of the low-lying floodplain of the river Rhine, which is built of porous aquifer, sandy Quaternary gravel, and clayey and silty interspersed alluvial deposits (Averbeck *et al*., 2019; Carbiener and Schnitzler, 1990; Erfurt *et al*., 2020; Kempf, 2018, 2019a, 2019b; Minářová *et al*., 2017b; Minářová *et al.*, 2017a; Preusser, 2008; Preusser *et al.*, 2016; Rentzel *et al.*, 2009; Schmitt *et al.*, 2007; Stork and Menzel, 2016). Large parts of the slightly elevated Mesozoic outcrops, the alluvial river terraces, and the foothill area are loess-covered, which led to the development of fertile soils. These soils experienced intense agricultural exploitation since the Neolithic period (Blöck, 2016; Brönnimann *et al.*, 2020; Faustmann, 2007; Kempf, 2018; Knipper *et al.*, 2018; Koehler *et al.*, 2013; Mischka, 2007). Parts of the Alsatian floodplain are characterized by various tributaries to the river Rhine, which drain the URA in the northern direction. The river L'Ill forms a secondary floodplain with low-drainage velocity and an accumulation regime, which traps fine-grained sediment from the headwater streams of the Vosges drainage system. Consequently, Gleysols and Alluvisols developed over clayey deposits with an increased waterlogging sensitivity. In combination with precipitation anomalies, a high aquifer and increased melt-water discharge can lead to rapid waterlogging and extensive flooding of broad parts of the floodplain (Giacona *et al.*, 2018; Himmelsbach *et al.*, 2015a, 2015b; Kempf, 2019a, 2019b, 2020c; Martin *et al.*, 2010). On the other hand, during increased summer temperature and precipitation decrease, the region suffers from long-term hot drought periods, which impacts the natural and human-made ecosystem (Blauhut *et al.*, 2015; Erfurt *et al.*, 2019; Kempf and Glaser, 2020).

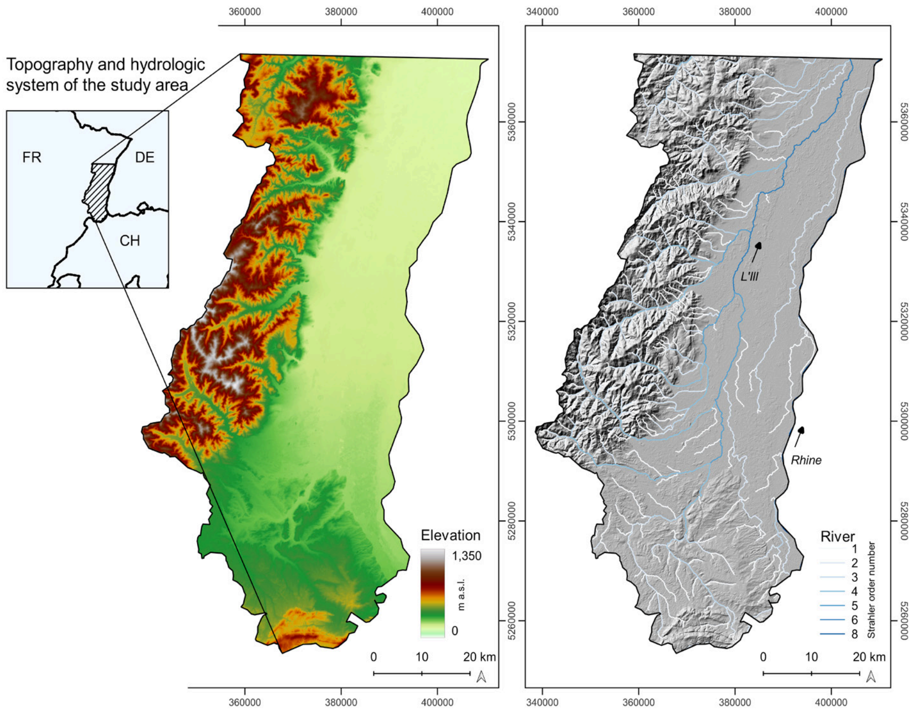

Figure 1. Topographic and hydrologic settings of the study area in Central Europe. The Alsace (shaded black area) is located at the eastern French border facing western Germany

and northern Switzerland and aligns with the river Rhine at the eastern margins. The hydrologic system is controlled by the river Rhine and the river L'Ill run-off systems draining the Upper Rhine Area to the north.

2.3. Archaeological data

Archaeological data is drawn from the French national archaeological database, *Patriarche*, specifically, the datasets *patriarche 67* and *patriarche 68*, which were uploaded to ArkeoGIS in 2016 (www.arkeogis.org; Dr. Loup Bernard, University of Strasbourg). Together, these datasets consist of 8136 records in 2396 unique locations, with information on location, chronology, typology, and research history. From these records, we extracted 799 archaeological sites that represent the locations of ancient and medieval rural settlements (Fig. 2). The study area is eventually determined by the extent of the archaeological coverage of the database. That means that our analyses are limited and biased *a priori* by the subjective delineations of the archaeological distribution and political boundaries. However, to integrate larger environmental feedbacks into the analyses, we included groundwater variability as a proxy for superregional climatic and environmental feedbacks. These dynamics mirror the broader ecological drivers behind settlement and land-use dynamics in the URA.

2.3.1 Archaeological data subsets

Assigning chronological classifications to these settlements was not straightforward. Chronological data are recorded in two columns: starting period and ending period, both of which are populated by a wide chronological range. Therefore, we were forced to use an extremely coarse chronological scheme: Pre-Roman, Roman, and Post-Roman. Pre-Roman settlements are those with a starting period prior to the turn of the era. None of these had an ending period that stopped before 26 BCE. Roman settlements are those with starting or ending periods that encompass the first five centuries CE, including the so-called Migration Period, which is basically a historical construction and does not reflect local rural settlement development (Brather, 2008). Post-Roman settlements have either a starting or ending period of 450-999 or 450-1491 CE. With such vague chronological data, only the broadest trends will be visible. However, such imprecision allows us to include every rural settlement in the database. If we were to restrict our focus to well-dated settlements, not only would our sample size be radically decreased, but we would run the risk of biasing our sample. Settlements whose inhabitants consumed more durable goods, especially if these goods were imported or of high quality, are more likely to yield datable archaeological evidence.

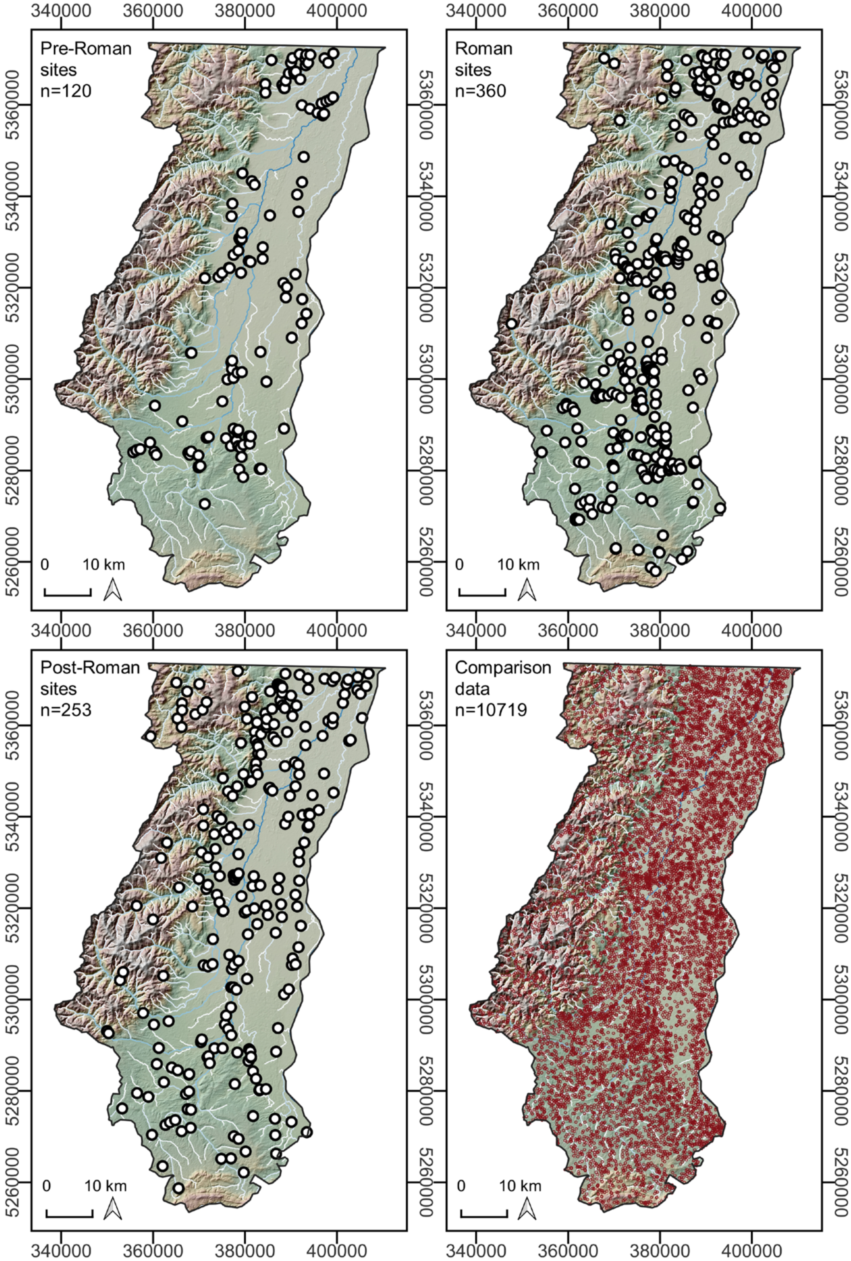

Figure 2. Site distribution of archaeological subsets and comparison dataset in the study area. Background is visualized by topography and hydrologic system.

2.3.2. Comparison bias surface and dataset

Spatial associations between archaeological sites and features in the landscape can provide valuable evidence about ancient societies, but to be a sound basis for inference, it is necessary to test any observed association against the null hypothesis that the association is the product of random chance. This can be done by creating a comparison dataset of random points and comparing the association between these random points and environmental features with the association between archaeological sites and the same features. However, due to the influence of formation processes affecting archaeological preservation and discovery, even a random distribution of activity areas in the past would not produce a random distribution of archaeological sites in the present. Therefore, the comparison dataset should be distributed to reflect the same factors that affect the distribution of all archaeological sites.

In this study, we focus on formation processes that affect discovery rates. We begin by examining the impact of modern land cover, since different types of landcover can preferentially conceal or expose archaeological remains to modern observers. Of course, landcover is not static, and the study area experienced considerable landcover change, deforestation and built-up change during the past decades (Kempf, 2020c). These areas have been subjected to intensive human attention, usually involving clearing and digging operations and, especially since 1992, commercial archaeology, which is an essential part of the Alsatian archaeology since 1941, when the Carcopino Law (confirmed in 1945) introduced the authorization and supervision of excavations by the French government and made the reporting of finds obligatory. The *Association pour les fouilles archéologiques nationales* (Afan) was founded in 1973. In particular, it administered the funds of the Ministry of Culture for planned and rescue excavations and carried out the measures. From the beginning, Afan established itself as an unavoidable intermediary of the government. *Antea Archéologie* was founded in 1998 and was the first French private company for preventive archaeology, licensed in 2005. The Law on Preventive Archaeology, approved on January 17, 2001, introduced a fee to finance preventive archaeology diagnoses and excavations. It has its legal basis in the European Convention for the Protection of the Archaeological Heritage, signed in Malta on January 16, 1992. The law established the *Institut national de recherches archéologiques préventives* (Inrap). which was created on February 1, 2002. It is a public administrative institution that replaced Afan. The Archéologie Alsace (AA, formerly known as PAIR) exists since 2006 and was created by the desire to have a preventive archaeology for the whole region Alsace (including Haut-Rhin and Bas-Rhin). AA has a scientific, cultural, and didactic function to save, study, protect, and promote the cultural heritage (Antea, 2022; Archéologie Alsace, 2022; Inrap, 2016). Human interventions such as digging are likely to reveal archaeological sites, and these institutions ensure that practically all archaeological sites discovered in this way appear in the national archaeological database. Therefore, in addition to landcover, we investigated the biasing impact of changes in forest cover and building activities (Kempf, 2020a).

To test the biasing influence of these factors, we broke each one into discreet categorical variables, shown in table 1. For each variable, we used the chi-square test to determine if archaeological sites were over- or under-represented in areas characterized by that variable as compared to the rest of the study area. Variables with a p-value less than 0.05 were considered significant. The biasing influence of these variables was quantified by dividing the observed number of sites by the number expected under conditions of complete spatial randomness. Variables with a p-value greater than 0.05 were given a weight of 1. All weights were then rescaled to fall between 0 and 1 to create a bias surface that could be used to generate a biased set of comparison points (Tab. 3).

TABLE 1
Variables containing all the bias factors

Land cover class	Area (ha)	Area (%)	Sites	Sites(%)
Arable land	136941	27.95%	884	36.91%
Artificial surfaces	61382	12.53%	838	34.99%
Forests	203726	41.58%	347	14.49%
Heterogeneous agricultural areas	40498	8.26%	198	8.27%
Open spaces with little or no vegetation	148	0.03%	0	0.00%
Pastures	18754	3.83%	53	2.21%
Permanent crops	15167	3.10%	62	2.59%
Scrub and/or herbaceous vegetation associations	7922	1.62%	6	0.25%
Water bodies	5156	1.05%	7	0.29%
Wetlands	317	0.06%	0	0.00%

Forest class	Area (ha)	Area (%)	Sites	Sites(%)
afforested	5174	1.06%	21	0.88%
deforested	101054	20.62%	435	18.17%
forest	198552	40.52%	326	13.62%
not forest	185229	37.80%	1612	67.34%

built class	Area (ha)	Area (%)	Sites	Sites(%)
cleared	15029	3.07%	93	3.88%
never built	413596	84.41%	1464	61.13%
new built	48625	9.92%	464	19.37%
still built	12757	2.60%	374	15.62%

Our analysis of modern landcover relies on Corine Land Cover (CLC, https://land.copernicus.eu/pan-european/corine-land-cover/clc2018, last accessed 18th of December 2022) data from 2018 (100 m resolution) with the addition of roads from Open Street map data (OSM, Geofabrik GmbH, last accessed 24th of August 2020). CLC data are described using a three-level hierarchy of labels. For most land cover classes, we used the most general description, but "Agricultural areas" and "Forest and semi natural areas" included second-level categories with very different implications for site discovery. "Arable land," for example, is plowed every year and is therefore more likely to produce surface sherds than "Permanent crops." Similarly, "Forests" are more likely to obscure remains than "Scrub and/or herbaceous vegetation associations." For these areas, then, the second-level label was used. The CLC data does not include roads, so we added these by calculating a 2 m buffer around lines representing roads and combining the resulting polygon with the "Artificial surfaces" polygon of the CLC. To understand the role of deforestation, we compared the forested areas represented in our modified CLC data to forest cover from 1972, which were taken from previously processed landcover data (Kempf, 2020c), classifying the landscape into areas that were consistently forest, never forested, deforested, and afforested. Changes in the built-up area were analyzed by comparing the artificial surfaces represented in our modified CLC data to the built-up areas represented on historical maps from 1866 augmented with a 100 m buffer (see Tab. 1 for bias factors).

In order to achieve a holistic understanding of the relationship between modern land cover, changes in forest cover, and changes in the built-up areas, we combined all three factors into a single, multivariate bias surface with 39 composite variables that represented unique combinations of variables from each factor (see Tab. 2 for variables included in the bias surface). Many of these covered too little area to be statistically analyzed, so these were combined with closely related variables to arrive at 22 composite variables. Many of these small variables included land that had been cleared of buildings. These were combined with similar areas that had never been built up. In addition, areas of wetland were combined with water bodies and open areas were combined with areas of scrub.

Table 2
Variables included in the bias surface

<table>
<tr><th>Original composite variable</th><th>Ori. Var. Area (ha)</th><th>New composite variable</th><th>New area (ha)</th><th>New area (%)</th><th>Sites</th><th>Sites (%)</th></tr>
<tr><td>Arable land-cleared_built-deforested</td><td>552</td><td rowspan="2">Arable land-never_built-deforested</td><td>19202</td><td>3.92%</td><td>71</td><td>2.95%</td></tr>
<tr><td>Arable land-never_built-deforested</td><td>18651</td><td>19202</td><td>3.92%</td><td>71</td><td>2.95%</td></tr>
</table>

Arable land-cleared_built-not forest	3219	Arable land-cleared_built-not forest	3219	0.66%	35	1.46%
Arable land-never_built-not forest	114515	Arable land-never_built-not forest	114515	23.39%	778	32.38%
Artificial surfaces-new_built-deforested	19643	Artificial surfaces-new_built-deforested	19643	4.01%	124	5.16%
Artificial surfaces-new_built-not forest	28981	Artificial surfaces-new_built-not forest	28981	5.92%	340	14.15%
Artificial surfaces-still_built-deforested	4775	Artificial surfaces-still_built-deforested	4775	0.98%	81	3.37%
Artificial surfaces-still_built-not forest	7982	Artificial surfaces-still_built-not forest	7982	1.63%	293	12.19%
Forests-cleared_built-afforested	141	Forests-never_built-afforested	5166	1.06%	21	0.87%
Forests-never_built-afforested	5025					
Forests-cleared_built-forest	5129	Forests-cleared_built-forest	5129	1.05%	15	0.62%
Forests-never_built-forest	193411	Forests-never_built-forest	193411	39.51%	311	12.94%
Heterogeneous agricultural areas-cleared_built-deforested	1851	Heterogeneous agricultural areas-cleared_built-deforested	1851	0.38%	11	0.46%
Heterogeneous agricultural areas-cleared_built-not forest	757	Heterogeneous agricultural areas-never_built-not forest	17116	3.50%	113	4.70%
Heterogeneous agricultural areas-never_built-not forest	16359					
Heterogeneous agricultural areas-never_built-deforested	21528	Heterogeneous agricultural areas-never_built-deforested	1851	0.38%	11	0.46%

Open spaces with little or no vegetation-cleared_built-deforested	0	Scrub and/ or herbaceous vegetation associations-never_built-deforested	7146	1.46%	5	0.21%
Open spaces with little or no vegetation-never_built-deforested	134					
Open spaces with little or no vegetation-never_built-not forest	14					
Scrub and/ or herbaceous vegetation associations-cleared_built-deforested	450					
Scrub and/ or herbaceous vegetation associations-never_built-deforested	6561					
Pastures-cleared_built-deforested	1548	Pastures-cleared_built-deforested	1548	0.32%	8	0.33%
Pastures-cleared_built-not forest	196	Pastures-never_built-not forest	4062	0.83%	15	0.62%
Pastures-never_built-not forest	3866					
Pastures-never_built-deforested	13142	Pastures-never_built-deforested	13142	2.68%	30	1.25%
Permanent crops-cleared_built-deforested	664	Permanent crops- never_built-deforested	8520	1.74%	29	1.21%
Permanent crops-never_built-deforested	7857					
Permanent crops-cleared_built-not forest	395	Permanent crops-never_built-not forest	6646	1.36%	33	1.37%
Permanent crops-never_built-not forest	6251					

Scrub and/ or herbaceous vegetation associations-cleared_built-not forest	26	Scrub and/ or herbaceous vegetation associations-never_built-not forest	922	0.19%	1	0.04%
Scrub and/ or herbaceous vegetation associations-never_built-not forest	882					
Wetlands-cleared_built-deforested	17	Water bodies-never_built-deforested	3691	0.75%	2	0.08%
Wetlands-never_built-deforested	162					
Water bodies-cleared_built-deforested	53					
Water bodies-never_built-deforested	3459					
Wetlands-cleared_built-not forest	2	Water bodies-never_built-not forest	1775	0.36%	5	0.21%
Wetlands-never_built-not forest	135					
Water bodies-cleared_built-not forest	26					
Water bodies-never_built-not forest	1612					

The results of the chi-square tests and the weights of bias are shown in table 3. These values were used to create a bias surface raster with a resolution of 5 m. We then used the "Create Spatially Balanced Points" tool in ArcGIS Pro 2.5 to generate 10000 biased random points, or one point for every 49 ha in the study area. Because they are not evenly distributed the average nearest neighbor distance is 340 m. Our basic unit of analysis is a circle with a 1500 m radius (see below). so given the scale of analysis, these points are a good representation of the background environment. The maps were partly created in QGIS 3.22.4 and R software 4.2.1 (Fig. 3).

TABLE 3

Results of the chi-square test and weights used for the bias surface

Composite variable	P-value	Observed / expected	Weight	Rescaled weight
Water bodies-never_built-deforested	0.00015	0.11085	0.11085	0.01476
Scrub and/or herbaceous vegetation associations-never_built-deforested	0.00000	0.14300	0.14315	0.01906
Forests-never_built-forest	0.00000	0.32900	0.32896	0.04381
Pastures-never_built-deforested	0.00001	0.46700	0.46702	0.06219
Forests-cleared_built-forest	0.04320	0.59833	0.59833	0.07968
Permanent crops-never_built-deforested	0.04803	0.69632	0.69632	0.09273
Heterogeneous agricultural areas-never_built-deforested	0.00185	0.70321	0.70321	0.09364
Arable land-never_built-deforested	0.01607	0.75644	0.75644	0.10073
Scrub and/or herbaceous vegetation associations-never_built-not forest	0.09831	0.22194	1.00000	0.13317
Water bodies-never_built-not forest	0.21096	0.57614	1.00000	0.13317
Pastures-never_built-not forest	0.27377	0.75542	1.00000	0.13317
Forests-never_built-afforested	0.39497	0.83162	1.00000	0.13317
Permanent crops-never_built-not forest	0.92773	1.01581	1.00000	0.13317
Pastures-cleared_built-deforested	0.87504	1.05708	1.00000	0.13317
Heterogeneous agricultural areas-cleared_built-deforested	0.51610	1.21545	1.00000	0.13317
Artificial surfaces-new_built-deforested	0.00356	1.29144	1.29144	0.17198
Heterogeneous agricultural areas-never_built-not forest	0.00110	1.35063	1.35063	0.17986
Arable land-never_built-not forest	0.00000	1.39000	1.38990	0.18509
Arable land-cleared_built-not forest	0.00000	2.22000	2.22415	0.29618
Artificial surfaces-new_built-not forest	0.00000	2.40000	2.40011	0.31962
Artificial surfaces-still_built-deforested	0.00000	3.47000	3.47046	0.46215
Artificial surfaces-still_built-not forest	0.00000	7.50937	7.50937	1.00000

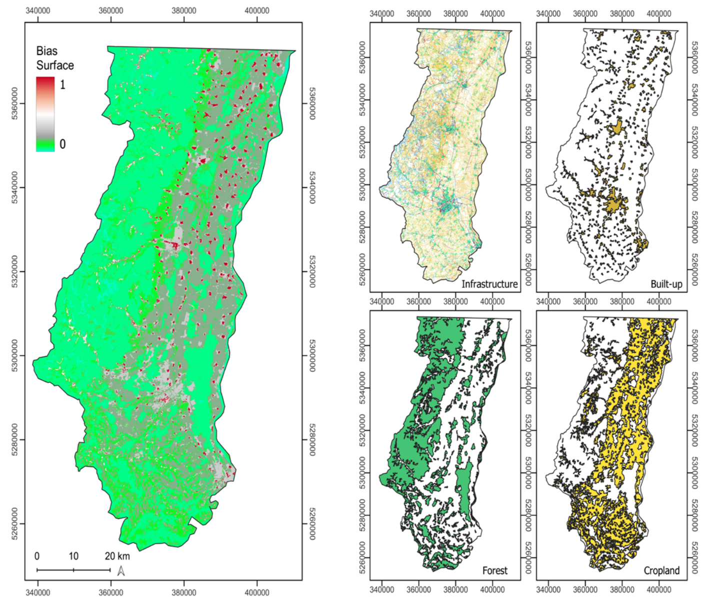

Figure 3. Accumulated bias surface, modern infrastructure, built-up, forest and cropland in the study area (CLC 2018).

2.4. Environmental data components and preprocessing

A broad set of environmental covariates was chosen for the analysis, mostly based on physical parameters. We determined elevation and slope, the hydrologic system, and soil properties to best represent a suitability package of land-use strategies. This set-up is, however, subjective and can be defined differently – depending on the geographical location of the study area or the desired outcome of the analysis. Site preferences in flat areas, for example, would not be controlled by slope gradients but rather by total elevation, groundwater depth, and flooding vulnerability, whereas mountainous regions would rather include water access, slope gradient, and erosion potential. Eventually, latitude and mean elevation of the study area control climate feedbacks and hence suitability for agriculture. In this case study and due to the location in Central Europe and a moderate temperate climate zone (Cfb after Koeppen and Geiger). we conclude that water access and topography in combination with soil properties best predict landscape preferences of premodern cultural groups.

With the exception of distance to water (1000 m). we measure the relationship of settlements to environmental variables by quantifying the prevalence of each variable within a fixed neighborhood around the settlement. Based on the work of Chisholm, we have chosen 1500 m as the radius for our neighborhood (Chisholm, 1979 (2007)). These neighborhoods frequently overlap, but for our purposes, this is appropriate. The 1.5 km neighborhood represents the land that could have been intensively cultivated from a settlement. Given the coarseness of chronological data, contemporaneity of settlements is impossible to assume, and even when settlements were contemporary, the practice of short-term tenancy and other systems of land allocation mean that the same field could have been cultivated sequentially by people living in different settlements.

2.4.1. Elevation data

Elevation data (digital elevation model, DEM) come from the Shuttle Radar Topography Mission (SRTM) 1 arc-second global dataset. We chose SRTM data to facilitate replicability of the analysis in different regions using different point patterns. The DEM provides information about absolute elevation and was used to calculate slope. The data was resampled to a 100 m grid cell resolution and fuzzified using a circular neighborhood of r=15 cells (that is a total of 31 cells, including the central cell, which equals 1500 m radius of the analysis). We focus on absolute height, impacting temperature and precipitation regime as well as slope gradient, which is decisive for settlement and cropland development. Furthermore, we estimated the aspect within the catchment due to the fact that the general aspect of the Vosges mountains and the foothill area is facing east (Fig. 1).

2.4.2. Hydrologic system

Access to fresh water is critical for biological sustenance, but the actual influence of water on people's behavior varies depending on its availability and characteristics. We measured access to water by calculating Euclidean distance to streams, as represented in the European Environment Agency's Catchments and Rivers Network System (ECRINS, https://www.eea.europa.eu/data-and-maps/data/european-catchments-and-rivers-network, last accessed, 16th of February 2022). These stream segments were generated using hydrological modeling of a DEM with a resolution of 100 m. The data are, therefore, rather coarse. However, given the extensive hydrological interventions experienced by the Alsace in modern times, they are probably a closer representation of ancient hydrology than data representing the current hydrological system.

The dataset has another advantage built-in, which is the Strahler order number. Using this categorization, we subset each stream by rank in the network with 1 = smallest headwater in the catchment and 8 = river Rhine. For this reason, we buffered the bounding box by 500 m to integrate the river Rhine in its full extent (the river is the political border between France and Germany). Then, each Strahler order number river segment was buffered individually to create river network polygons. The following classification was chosen to best fit the physiological conditions: 1=2 m, 2=3 m, 3=7.5 m, 4=10 m, 5=12.5 m, 6=15 m, 7 and 8 = 100 m buffer radius. This accounts for the changing run-off characteristics in the different parts of the catchment. Eventually, all polygons were merged, and a binary streamflow raster was produced with watercourse =1 and no water = 0 value. Raster grid resolution was set to moderate resolution (ncols = 7k, nrows = 7k, which results in 10.44086 m x 17.31143 m) to preserve the small-scale river matrix and to reduce computational time. Eventually, the raster was fuzzified using the above described method and a radius of 1000 m. The computational time increases rapidly with increasing resolution and we used the *terra* package for this operation due to advantage in computational speed (Hijmans, 2022). The focal approach has been conducted using the *focalWeight* function in the *raster* package and *focal* in *terra*. We chose a radius of 1000 m for the water access calculation to emphasize the stronger dependencies on running freshwater. On the other hand, the distance to river can be biased by just implying superficial run-off. Hence, we integrated aquifer depth into the data analysis, which in turn provides information about flooding potential and waterlogged soil conditions. The depth of the aquifer was modeled by interpolating the results of 327 cores using raw data from Aprona (https://www.aprona.net/, last accessed 25th of August 2020) (Kempf, 2020c). Groundwater table data was then interpolated across the study area using IDW interpolation and a raster grid cell of 100 m. Eventually, the raster was cropped to the extent of the study area and fuzzified using a radius of 1000 m (Fig. 4).

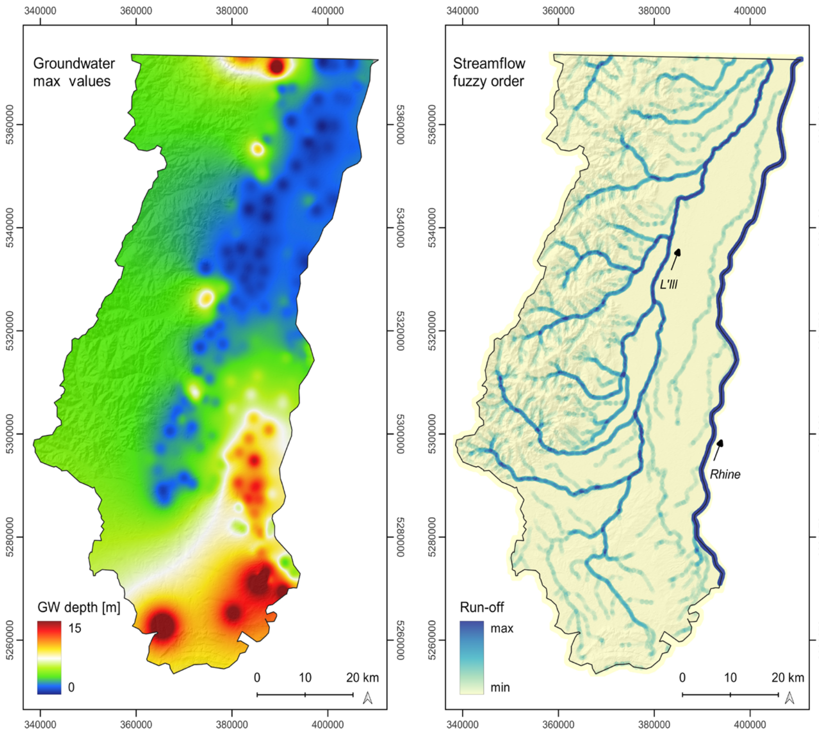

Figure 4. Groundwater table model for the study area (left) and visualization of the buffered and fuzzified river system using the Strahler order numbers 1-8 (right).

2.4.3. Soil data

Soil data, with a scale of 1:250,000 come from *l'Association pour la Relance Agronomique en Alsace* (ARAA) (https://www.datagrandest.fr/geoserver/araa/wfs, last accessed 04th of October 2022). From this dataset, we could identify 9 different classes that build the basis for soil characteristics, such as soils developed on crystalline rocks or alluvial sediments. The vector data was subset for each soil class, aggregates, unified, rasterized and eventually fuzzified with 1500 m radius. Consequently, every cell is the average value of a circle with 1500 m around the cell and thus represents soil quality within the catchment for each archaeological site (Fig. 5). The soil characteristics and reclassifications are described in Tab. 4. The soils were classified based on water storage capacity of the underlying geological units (e.g., loess soil (high) vs. Quaternary gravel (low)). Chemical attributes related to soil quality of the weathering material are further considered (e.g., sandstones rather leading to acidic and shallow soils). Crystalline and metamorphic bedrock of the Vosges mountain ranges are producing acidic and less fertile soils and are thus considered less favorable for extensive crop production – despite the fact that a generalization across long-standing human occupation is critical (Koerner *et al.*, 1997).

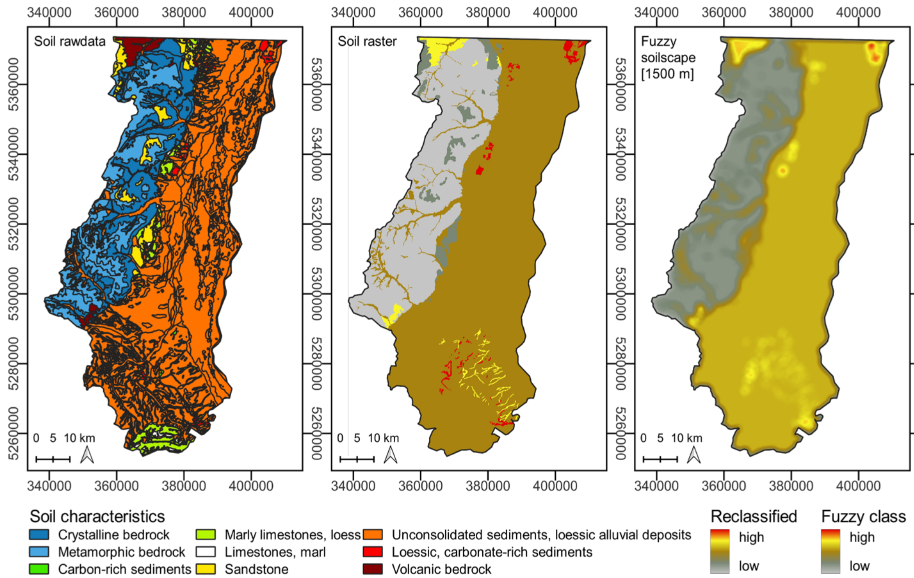

Figure 5. Soil characteristics in the study area (left). reclassified raster based on soil suitability (class ranges from 1 = low to 5 = high) (center). and focal/fuzzy soilscape with r=1500 m (right).

Table 4

Soil characteristics and reclassification according to potential productivity (1= low, 5=high). Acidic soils of the crystalline bedrock and soils developing over sandstones can be considered of poor productivity. Alluvial soils and soils on loess deposits show high productivity

Subset	Characteristic	reclass
sub1	Crystalline bedrock	1
sub2	Metamorphic bedrock	1
sub3	Carbon-rich	4
sub4	Marly limestones, loess	3
sub5	Limestones, marl	3
sub6	Sandstone	2
sub7	Unconsolidated sediments	3
sub8	Loessic, carbonate-rich sediments	5
sub9	Volcanic bedrock	4

2.5. Formal analyses

All formal analyses in this paper have been conducted using R software. The codes underlying the analysis of this paper were adapted from Kempf and Günther (2023) and are available from https://zenodo.org/record/7307543 (Kempf and Günther, 2023). The *rbias* package can be downloaded from https://zenodo.org/record/7071418 (Günther *et al.*, 2022).

2.5.1. Complete spatial randomness

Complete Spatial Randomness (CSR) and clustering processes were checked using Ripley's inhomogeneous K-Function (Ripley, 1977). This function allows for the detection of clustered points, regular point pattern, or spatial randomness (Bevan and Conolly, 2006; Crema *et al.*, 2010). With this function the spatial properties of a point pattern can be analysed. The function in a Monte-Carlo-Simulation evaluates whether the observed point pattern is different from a theoretical distribution (tested many times against a random comparison dataset) or if they are drawn from the same (random) sample. We included all archaeological sites and the comparison (random) data into the analyses to visualize the effects (Fig. 6). The output is a graph with a so-called Poisson distribution, in which the grey area (envelope) describes the Min/Max of a potential random distribution, a black line that represents the observed distribution, and a red line with the potential mean distribution. If the black line is located above the envelope, clustered behavior is evident, if it falls into the envelope, CSR is established (Nakoinz and Knitter, 2016).

Eventually, a Kernel Density Estimation (KDE) can be derived, which represents site intensity as a function of the underlying distribution at certain radius of a kernel placed on top of the sites locale (O'Sullivan and Unwin, 2010). All calculations were performed using the *spatstat* package (version 2.3-4) in R (https://cran.r-project.org/web/packages/spatstat/index.html, last accessed 03rd of October 2022) (Baddeley *et al.*, 2016; Baddeley and Turner, 2005) (Fig. 6, Fig 7).

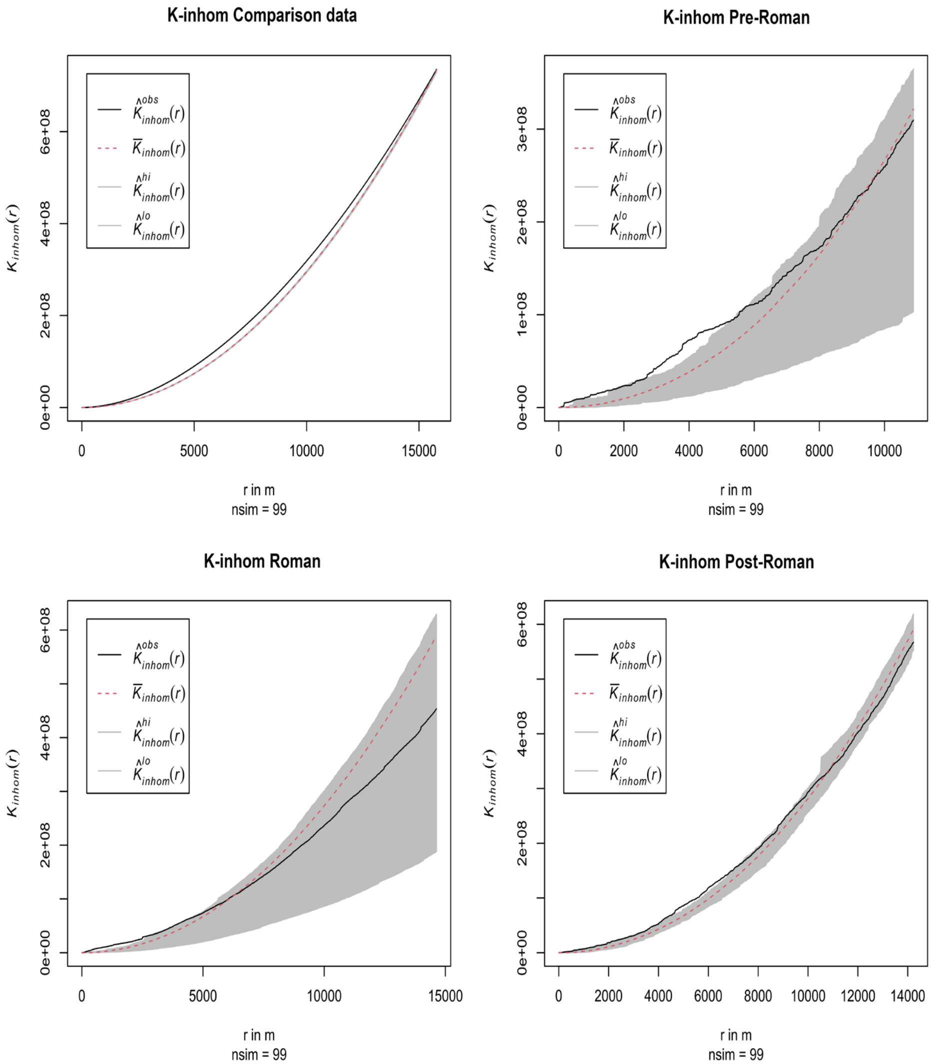

Figure 6. Results from the K-function for all four point patterns. The outcome of the random point pattern visualizes a random distribution of the points. The archaeological sites, however, show clustered behavior.

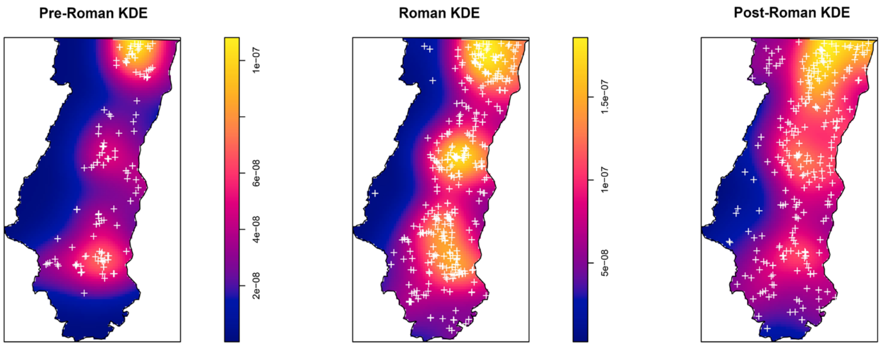

Figure 7. KDE for the archaeological sites using a sigma of 7500 m within the boundary of the study area. The KDE identifies three major clusters for each time period. The site intensity changes between the clusters due to different sample size.

2.5.2. Environmental preference model

There are a great many approaches to model landscape or environmental feature preferences using archaeological point data, ranging from site-based and 2-dimensional approaches to catchment and complementary region analyses (Brandolini and Carrer, 2021; Kempf, 2021; Laabs and Knitter, 2021). In general, the site-based or point-based approach can be regarded as less meaningful because it only provides information about the very locale of a site – a concept, which can be considered to be of methodological uncertainty. A site rather represents a polygon with fuzzy margins than a two-dimensional entity. Hence, the integration of the environmental or socio-cultural complexity within the complementary area to a site offers far more valuable information regarding settlement or agricultural potential of a region. However, what best describes the complementary region is related to the size of the settlement or the dwelling, the population density, network and market integration, political or administrative centrality, and also cultural and religious significance. Eventually, the radius of a proper catchment area to a site is subject to the research goals and approaches and cannot be generalized.

Here, we deploy the *rhohat* function, implemented in the *spatstat* package in R (Baddeley and Turner, 2005). The function itself does not use a specific radius as complementary area around one site. However, we use a focal approach to integrate the spatial components in a fixed neighborhood to gain knowledge about the environmental parameters within a circular surrounding. For this reason, we use a nearest neighbor analysis where every cell of an underlying explanatory raster (e.g., a digital elevation model) is the average of the surrounding cells within a circle of 1500 m. This allows to understand site preferences that take into account the catchment compositions instead of just point-based data information

(Carrero-Pazos *et al.*, 2019; Kempf, 2021).

Eventually, *rhohat* calculates site intensity as a function of the pre-processed focal raster data. With this approach, we can visualize the effect of attraction or repulsion given by a specific parameter – the environmental raster. The environmental covariate can further be interpreted as an explanatory variable for site location preferences and avoidance regarding the catchment composition in the custom neighbourhood. According to Baddeley *et al*. (2016, 180), "the plot method generates a plot of the estimated function *ρρ* (*zz*) against covariate values z, together with 95% confidence bands assuming an inhomogeneous Poisson point process" (Baddeley *et al*., 2016). This means that we can estimate the correlation between site intensity and variable in a 95% confidence envelope (Baddeley *et al*., 2012).

2.5.3 Bias model using rbias

To test the performance of the bias surface, the recently developed R-package *rbias* (Günther *et al*., 2022) has been used to compare the results of the Rhohat function within *spatstat* and the output of *rbias*. Basically, the package uses data from OpenStreetMap (OSM) and features characterized by the classes "residential", "industrial", "commercial", and "retail" (modern land-use) as well as features tagged with the classes "motorway", "motorway_link", "primary", "primary_link", "secondary", "secondary_link" to map the major road network, and features marked as "rails" to represent railways. To smooth the harsh boundaries of the OSM features, the package uses fuzzy variables and the R package *FuzzyLandscapes* (Hamer and Knitter, 2018). Consequently, the impact of modern infrastructure on the archaeological sites can be modelled at different ranges. We use ranges of 1000 m for our analysis. A range of 1000 m implies that a modern object's impact is limited to a radius of 1000 m. Thus, archaeological sites located at the same location as the object are potentially heavily biased. Therefore, they are assigned a membership degree of 1. The bias or the membership degree decreases with an increasing distance until it drops to 0 at distances of 1000 m or more.

3. RESULTS AND DISCUSSION

In the following the results from the *rhohat* functions will be presented, followed by the bias surface model and a comparison to the recently developed *rbias* package. From the *rhohat* plots, a certain preference for particular environmental components can be observed. The x-axis represents the observed sites (black tickmarks) at particular values of the explanatory raster and the y-axis is the site intensity. The envelope characterizes 95% confidence level.

3.1. Landscape preferences

3.1.1. Topography

First, the data has been analyzed for topographic preferences, including elevation, slope, and aspect (Fig. 8). From the plots, there is not much difference between all archaeological chronologies, indicating a preference for low-lying locations up to 400 m (pre-Roman). up to more than 500 m (Roman). and up to 700 m (post-Roman). However, the site intensity is largest at low values with a narrow confidence interval. No sites are located at very low values, which is due to the general absolute height above sea level of the study area. The comparison dataset shows equally distributed sites with a focus on 200-300 m. This is in contrast to the post-Roman sites, which show a concentration below 200 m, followed by a steep drop-off as elevation values rise. The use of biased comparison data allows us to conclude that the initial spike is likely a result of biasing factors, but that the drop-off indicates a real aversion to lower elevations in the post-Roman period. This preference is a combination of multiple factors of the study area, which are composed of high soil quality of the lowlands and foothill area, climate suitability, and forest cover. The slight increase in elevation during the post-Roman period can probably best be explained by increasing land-use in the mountains, including forest management and mining activities as well as spread of Christianity and the establishments of monasteries.

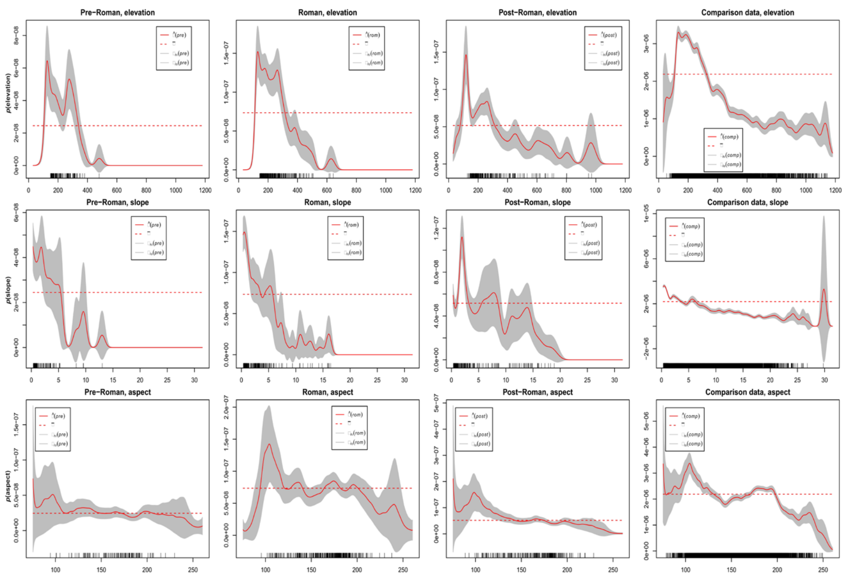

Figure 8. Topographic preference model of Pre-Roman, Roman, and Post-Roman sites

compared to a biased comparison dataset. From top: absolute elevation, slope, and aspect. The input DEM has been resampled to 100 m grid size and was fuzzified using a 1500 m radius around each cell. The x-axis represents the observed sites (black tickmarks) at particular values of the explanatory raster and the y-axis is the site intensity. The envelope characterizes 95% confidence level.

The picture changes when considering slope into the analysis. Pre-Roman sites are mostly located in flat areas or areas including gentle slopes. Roman and particularly post-Roman sites show increasingly sloping landscapes in their complementary regions of 1500 m. Potentially, this is again reflecting different land-use approaches and particularly landscape availability at earlier periods and more intensification in later periods. Comparison data shows equally distributed sites with no preferences. The aspect further plays a decisive role in pre-Roman site location. Although the plot in figure 8 shows relatively even preferences for most aspects, with a slight preference for east facing slopes, the comparison data reveal that this is not what one would expect from a random distribution. The comparison data show that the topography of the study area, when archaeological bias is taken into account, has more southeast facing slopes and fewer west facing slopes. Therefore, the apparent pre-Roman indifference to these aspects is actually evidence for a non-random response. Roman sites are equally facing south and south-west, with a broader spectrum of distribution. Post-Roman sites in principle follow that distribution. The patterns are most likely generated by the general south-west facing of the foothill areas, which are nowadays intensely used for growing high quality wines.

Compared to the biased random point distribution of the comparison data, all archaeological sites show particular site location patterns, which emphasized that they are not randomly distributed in the landscape. Low-lying, south-west facing slopes and flat areas were preferred and high altitudes and steep slopes generally avoided, which makes physiological sense, considering agricultural crop production in the very fertile region of the URA.

3.1.2. Hydrologic system

Two explanatory rasters were included into the analysis, the groundwater model and the Strahler order river network (Fig. 9). The groundwater model reveals insights into the use of floodplain areas of the river Rhine tributaries. The comparison data shows that many locations are characterized by a very shallow water table and possibly significant number of locations are characterized by a deep water table (the peak to the right of the graph). with a fairly even distribution of locations with moderate depths. In contrast, the pre-Roman sites show no peak at the right side of the graph, indicating the avoidance of areas with a very deep water table, unlike the Roman and Post-Roman sites. Upwelling groundwater of the aquifer and heavy rainfall can lead to persistent flooding of the so-called

'Ried' in the Alsace (Kempf, 2019b, 2020c), but the similarities observed between archaeological sites and the comparison data at low values of groundwater depth means that we cannot observe any archaeological response to this danger.

We do observe differences between archaeological sites and comparison data in terms of river water availability. The comparison data is evenly distributed, but archaeological sites from all periods are concentrated at low values, indicating greater access to river water.

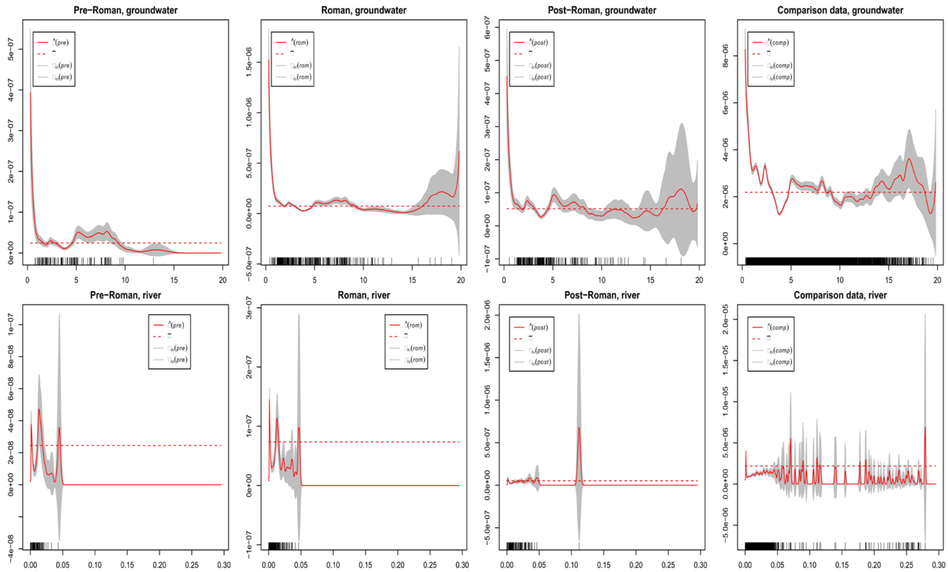

Figure 9. Results from the groundwater model (upper part) and Strahler order river network (lower part). The x-axis represents the observed sites (black tickmarks) at particular values of the explanatory raster and the y-axis is the site intensity. The envelope characterizes 95% confidence level.

3.1.3. Soil units and characteristics

The focal soils were reclassified into 5 groups, regarding their general suitability for agriculture and/or settlement spot (Fig. 10). The soil development in the study area is strongly connected to the elevation, thus representing low-quality soils over crystalline bedrock of the central mountain range and rather fertile and highly productive soils in the plain and on top the foothill zone. The comparison data shows a fairly even distribution of locations across soil quality values, which matches the distribution of Roman and post-Roman sites, but Pre-Roman sites are concentrated on soils of higher quality. Therefore, we can conclude that Pre-Roman sites were located to take advantage of the most fertile soils, but Roman and post-Roman sites were not.

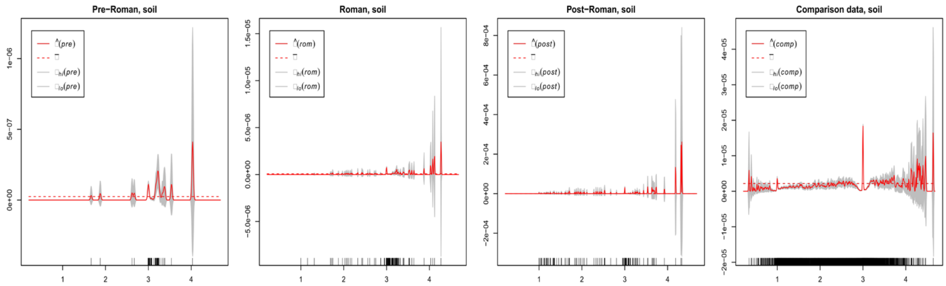

Figure 10. Results from the soil quality model for each archaeological site and comparison data. Higher values indicate high quality (x-axis). The x-axis represents the observed sites (black tickmarks) at particular values of the explanatory raster and the y-axis is the site intensity. The envelope characterizes 95% confidence level.

3.1.4. Bias model 1: bias surface

We used the bias surface to estimate the impact of modern land-use, deforestation, infrastructure, and built-up change on the archaeological record. This comprehensive bias surface takes into account the chronological development of the study area during the past decades and thus represents a robust method to evaluate the influence of "findability" caused by increasing human permeability on the archaeological record in general. We resample the very detailed bias surface to a 10 m resolution and used a focal approach of r=250 m to best predict the influence of modern land-use. Eventually, the plots (Fig. 11) show that pre-Roman sites are less biased than Roman sites and both are less biased than post-Roman sites. With younger age, the impact of modern land-use gets stronger – which is reasonable in a cultural landscape that shows multiple high medieval and medieval village cores, a long-standing Roman history, and strong modern reshaping related to building activity and agriculture.

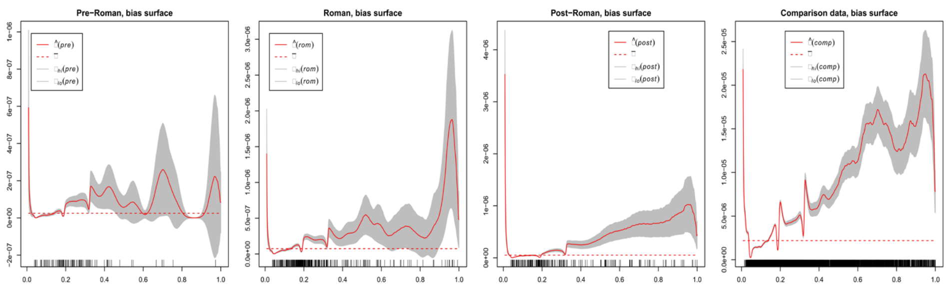

Figure 11. Bias surface *rhohats* for all archaeological sites and comparison data. The x-axis represents the observed sites (black tickmarks) at particular values of the explanatory raster and the y-axis is the site intensity. The envelope characterizes 95% confidence level.

3.1.5. Bias model 2: rbias

The R package *rbias* was used to compare the results from the custom bias surface to an automatically produced bias surface. The package can be customized with different input variables, according to cell size and extent of the bias surface. Here, the fuzzy bias influence range was set to range = c(1e-2, 1, 2.6e-1, 0) with a cell size of 500 m and a xyrange = c(0, 1000). This produces the following plots, in which the archaeological site distribution was plotted against a n=999 simulation of random points. We can see that for pre-Roman sites, the random distribution lies above the observed site frequency for very small values, which means that we should expect more random sites with no bias than actually apparent. At high values, the observed sites lie over the simulated distribution, which means that more sites are biased within 1000 m than expected in the simulation. That is true for all sites, including the comparison data (that is not random but equally biased). Hence, we can detect bias for all archaeological features in the study area (Fig. 12).

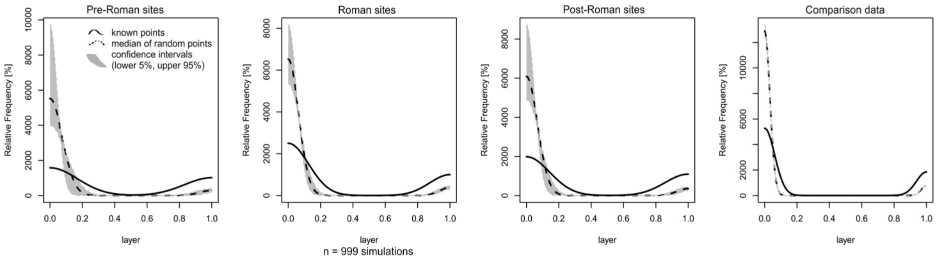

Figure 12. *rbias* covariates characteristcs for all archaeological sites and comparison data. Solid line: known points; dashed line: median of random points; shaded area: confidence intervals (5%, 95%).

3.2 Landscape development and impact on site distribution

Our study area encompasses a region of the Upper Rhine that formed part of the Roman frontier zone from the end of the first century BCE through the fifth century CE. The presence of thousands of Roman soldiers and political integration with the Mediterranean set this period apart from the pre-Roman and post-Roman periods. In particular, military demand is thought to have increased agricultural production. There are, essentially, two ways to increase agricultural production: extensification, the cultivation of more land, or intensification, the investment of more labor and resources in each unit of land. In the Roman Rhine frontier zone, there is evidence for both. Paleobotanical remains of new plants and faunal remains showing livestock improvement have been interpreted as evidence for intensification, while the harvesting of spelt with famous “Gallic reaper” only makes sense in the context of extensive agriculture. Eventually, soil quality and availability are primary control factors for the development of prosperous land-use.

Soil texture and chemical and physical soil composition varies significantly in the Upper Rhine Area and particularly from the Palaeozoic Vosges mountains and the Mesozoic foreland to the Quaternary floodplain, which is mostly characterized by gravel and sandy alluvial deposits. The soil development during the Holocene is tied to the underlying geological conditions, hydrological erosion and accumulation, and eventually climate fluctuations during the past 11,500 years. The very fragmented geomorphology in the study area has led to mosaic soil conditions with patchy surface characteristics and locally heterogeneous soil thickness, composition, grain size, and drainage potential. Hence, the extensive agricultural exploitation, which dominates modern land-use, cannot be assumed for prehistoric periods without considering local differences in soil quality and potential yield quantities. In this context, loess coverage is often assumed to be the first choice for agricultural exploitation (Kempf, 2018). In the study area, broad parts of the Quaternary sediments and the Mesozoic forelands are covered with loess or secondary relocated loess deposits, which derived from Pleistocene aeolian transport and subsequent sedimentation (Antoine *et al*., 2001; Antoine *et al*., 2016; Froehlicher *et al*., 2016; Lehmkuhl *et al*., 2016). The primary loess-covered areas, however, feature further location parameters, which are strongly interrelated. Most of the aeolian sediment is trapped by vegetation and topographic obstacles. Hence, landscape roughness of the floodplain plays a major role in the deposition processes of loess in the study area. The slightly elevated Mesozoic outcrops and plateaus are further situated outside the flood prone areas of the prehistoric meandering and anastomozing river Rhine hydrologic system (Preusser, 2008), which is characterized by frequent channel shifts during extreme run-off values and the input from the various tributaries that drain the Vosges mountain range (Rentzel *et al*., 2009). Heavy precipitation can cause rapid changes in headwater run-off patterns, which results in high-velocity washloads and extensive flooding events in the lower parts of the river Rhine and river L'Ill floodplain. The elevated areas of the floodplain are thus not only favorable in terms of flood security but also with respect to soil condition and drainage, which prevents harvest loss during early summer run-off maximum.

The results from the spatial model reveal a strong preference of pre-Roman and Roman occupation of silty soils (loess, high suitability) and clear avoidance of sandy and clayey soils (low suitability). Sandy soils in the area are characterized by high drainage potential and low water storage capacity with rather acidic and semi-fertile soil composition. Clayey soils show strong waterlogging potential and heavy soil compositions, which are rather unsuitable without deep-plowing agricultural techniques. Furthermore, the sedimentation of fine-grained material is often accompanied with upwelling groundwater and local flooding events, which increases the risk of harvest loss (see Fig. 4, groundwater model). Scattered population dispersal in the pre-Roman periods would thus be linked to the favorable parts of the landscape. The subsequent Roman occupation shows a significant increase in settlement numbers, which are more homogeneously distributed in the URA (Fig. 2). During the Roman period, soil preferences change slightly. This can be seen as a general trend in extensification of agricultural exploitation of the landscape during increased population dynamics and a general trend towards human presence in the URA - mostly related to military fortifications and the rural development of the local people during the 1st to the 3rd century AD. In this context, the introduction of new agricultural crops like spelt and the technological enhancements through communication and transfer has led to an increase in yield production in the Roman administrative realm. A combination of stable climatic conditions during the early Roman period and the potentially non-extensive pre-Roman land-use strategies have further enabled the expansion of crop cultivation and eventually the utilization of most of the favorable and even the semi-favorable soils with the goal to maximize yield production. The continuous declining loess preferences during post-Roman land-use aligns with the technological development of the turning plough in the Early Middle Ages, which allows for plowing deeper soils of the floodplain and the cultivation of rye - maybe as a reaction to disturbances in climate stability and a trend towards more humid conditions (Brombacher and Hecker, 2015; Büntgen *et al* l., 2016; McCormick *et al*., 2012).

In this context, flooding susceptibility has been intensely discussed and a broad number of studies have pointed out the high flooding sensitivity of the URA (Giacona *et al*., 2018; Glaser *et al*., 2010; Martin *et al*., 2017; Wetter *et al*., 2011). Extensive flooding is not only caused by a heavy precipitation events in the floodplain due to convective conditions during summer but also by long-term precipitation in the mountain range, which causes drainage collapse of the soils and extreme discharge of the headwaters. The high groundwater table in the Alsace, which is to a large extent dependent on Alpine surface discharge and aquifer, locally amplifies the water saturation of the surface-near soil layers. The spiral-like amplification of fine-grained clayey and waterlogged deposits, high groundwater level, the hydrologic network, and increased precipitation during summer impacts the drainage potential of the soils and locally triggers soil textures and units, which cannot be utilized without heavy drainage activity. Furthermore, the massive channelization of the river Rhine and parts of the tributaries led to a general groundwater drop in the Upper Rhine aquifer. Taking into account the local differentiations of groundwater height and response to modern anthropogenic impact, the overall floodplain dynamics during the pre-Roman to post-Roman period can be assumed to be much more affected by periodical flooding and generally more humid conditions. This is not only affecting potential agriculture and settlement dynamics but also the accessibility, permeability and availability of local- to regional-scale landscape patches and resource exploitation like pastures or timber respectively. In addition, life-quality close to the marshy and swampy parts of the floodplain must have decreased rapidly due to potential malaria and fever hot spots, elimination of which was one of the side-effects of the floodplain correction during the 18th and 19th century.

These general considerations of landscape development are visible in the land-use development from the pre-Roman Iron Age to the Roman Period (Fig. 4). A moderate aquifer was preferred by pre-Roman land-use activity, which can be linked to climate minimum and a general increase in precipitation at that time. In turn, artificial irrigation would not have played a major role in crop cultivation on loess-covered on flat and gentle slopes. In the Roman period, the availability of agriculturally utilizable soil patches on gentle slopes and in flat areas was still high, which allowed for continuous use and extensification of crop production during the first 4th to 5th centuries AD. The extensification towards lower parts of the floodplain is visible in the increasing density of sites in areas with a generally lower average elevation and particularly in the significant increase of sites in areas dominated by quaternary gravel and alluvial deposits and a decrease of site occupation on loess-covered, silty areas. Consequently, the drainage potential of the soils (not the geology) decreases rapidly from areas with low aquifer and silty aeolian deposits towards sandy and more clayey soil textures of the floodplain, which are furthermore characterized by higher groundwater availability. This can be a sign for technological development and the introduction of broad varieties of cereals, like spelt, or an increasing demand of agricultural cropland caused by massive increase of population and the presence of military activities at the borders of the Roman administration. The latter would explain the growing pressure of Roman land-use and population dynamics during the first centuries in the URA. In addition, Roman technological development led to the construction and maintenance of regional to supraregional infrastructural networks in the URA, which in turn supported the establishment of Roman *villae*, settlements, and market-oriented production units in close distance to accessible and stable routes and roads (Weaverdyck, 2019). Further amplified by the military presence of the Roman army, these pull-factors have caused a strong transformation of the landscape into a Roman cultural activity area, which is not only mirrored by geomorphological proxies like colluvial development (Lang *et al*., 2003; Mäckel *et al*., 2002; Mäckel *et al*., 2003), but also by vegetation change through clearing activity, mineral resource exploitation, and socio-political development of the local, peripheral society - on both side of the river Rhine.

4. CONCLUSION

Spatial analysis in archaeological research covers a large variety of theoretical and methodical approaches and has established an individual debate. In this paper, we focus on the application of Point Pattern Analysis (PPA) to describe patterns in archaeological records across different chronological periods. We apply quantitative statistics and spatial modelling to understanding not only the environmental explanatory covariates that control settlement and land-use strategies, but also to evaluate the impact by modern infrastructural development in eastern France. Modelling pre-Roman, Roman, and post-Roman sites, we find that the modern impact by land-use and built-up change accounts for a certain bias in the distribution of sites. A custom-built bias surface that includes also historical surface development probably shows that Roman and post-Roman sites are more biased than pre-Roman site locations – pointing towards continuous land-use and settlement development in the landscape since Roman times. The new R-package *rbias*, that generates a fuzzified bias surface using OSM data, however, points towards more biased archaeological records in general. On the other side, this is most likely caused by input variables and resolution of the underlying explanatory raster data. Eventually, we evaluated site preferences in archaeological data using the *spatstat* package in R and the implemented function *rhohat*. This function allows to understand site intensity as a function of a covariate. We used elevation, slope, aspect, groundwater table and hydrology as well as different reclassified soil characteristics to detect preferences in the landscape. We find that low lying areas with flat or gentle slopes on fertile soils were increasingly dominant during pre-Roman and Roman times. This changed at the transition to the Early Middle Ages and the Middle Ages in general, where different soil types and remote areas started to gain attraction. Most likely, this is due to technological developments in the region.

The use of biased comparison data has allowed us to distinguish locational trends in the distribution of archaeological sites that are the result of actual ancient preferences from those that could be the product of discovery bias. There are multiple ways to identify and quantify this bias, and in the future, archaeologists should experiment with several in order to bolster the empirical basis of the conclusions they draw from point pattern analysis.

5. ACKNOWLEDGEMENTS

The authors would like to thank Alfredo Cortell-Nicolau from the McDonald Institute for Archaeological Research at the University of Cambridge for the invitation to contribute to this Special Issue. MK's research received funding from the German Research Foundation and the CRC1266 'Scales of Transformation' at Kiel University under grant number 290391021. EJSW's research at Freiburg received funding from the ERC-project 'BaSaR – Beyond the Silk Road' under grant number 742645.

6. REFERENCES

Antea (2022): "*Archéologie préventive*", available at: http://www.antea-archeologie.com/antea/poles-de-competences.html (accessed 18 December 2022).

Antoine, P.; Coutard, S.; Guerin, G.; Deschodt, L.; Goval, E.; Locht, J.-L.; Paris, C. (2016): "Upper Pleistocene loess-palaeosol records from Northern France in the European context: Environmental background and dating of the Middle Palaeolithic", *Quaternary International*, Vol. 411, pp. 4-24.

Antoine, P.; Rousseau, D.-D.; Zöller, L.; Lang, A.; Munaut, A.-V.; Hatté, C.; Fontugne, M. (2001): "High-resolution record of the last Interglacial-glacial cycle in the Nussloch loess-palaeosol sequences, Upper Rhine Area, Germany", *Quaternary International*, 76-77, pp. 211-229.

Archéologie Alsace (2022): "*Présentation*", available at: https://www.archeologie.alsace/fr/le-service/presentation.

Averbeck, P.; Frör, O.; Gartiser, N.; Lützel, N.; Rudolf, F. (2019): "Climate change preparedness of enterprises in the Upper Rhine region from a business perspective. A multidisciplinary, transboundary analysis", *NachhaltigkeitsManagementForum | Sustainability Management Forum*, 27 (2): 83-93.

Baddeley, A.; Chang, Y.-M.; Song, Y.; Turner, R. (2012): "Nonparametric estimation of the dependence of a spatial point process on spatial covariates", *Statistics and its interface*, 5: 221-236.

Baddeley, A.; Rubak, E.; Turner, R. (2016): *Spatial point patterns: Methodology and applications with R*, A Chapman & Hall book, CRC Press Taylor & Francis Group, Boca Raton, London, New York.

Baddeley, A.; Turner, R. (2005): "spatstat An R Package for Analyzing Spatial Point Patterns", *Journal of Statistical Software*, 12 (6).

Bevan, A. (2012): "Spatial methods for analysing large-scale artefact inventories", *Antiquity*, 86 (332): 492-506.

Bevan, A.; Conolly, J. (2006): Multiscalar approaches to settlement pattern analysis", in G. Lock and B. Molyneaux (Eds.), *Confronting Scale in Archaeology: Issues of Theory and Practice*, 1st ed., Springer, Boston, MA: 217-234.

Blauhut, V.; Stahl, K.; Stagge, J.H.; Tallaksen, L.M.; Stefano, L. de; Vogt, J. (2015): "Estimating drought risk across Europe from reported drought impacts, hazard indicators and vulnerability factors", *Hydrology and Earth System Sciences Discussions*, 12 (12): 12515-12566.

Blöck, L.A. (2016): *Die römerzeitliche Besiedlung im rechten südlichen Oberrheingebiet, Dissertation, Forschungen und Berichte zur Archäologie in Baden-Württemberg*, Band 1, Dr. Ludwig Reichert Verlag, Wiesbaden.

Brandolini, F.; Carrer, F. (2021): "Terra, Silva et Paludes Assessing the Role of Alluvial Geomorphology for Late-Holocene Settlement Strategies (Po Plain – N Italy) Through Point Pattern Analysis", *Environmental Archaeology*, 26 (5): 511-525.

Brather, S. (2008): *Zwischen Spätantike und Frühmittelalter: Archäologie des 4. bis 7. Jahrhunderts im Westen, Ergänzungsbände zum Reallexikon der germanischen Altertumskunde*, Bd. 57, Walter de Gruyter, Berlin, New York.

Brombacher, C.; Hecker, D. (2015): "Agriculture, food and environment during Merovingian times: plant remains from three early medieval sites in northwestern Switzerland", *Vegetation History and Archaeobotany*, 24 (2): 331-342.

Brönnimann, D.; Knipper, C.; Pichler, S.L.; Röder, B.; Rissanen, H.; Stopp, B.; Rosner, M.; Blank, M.; Warnberg, O.; Alt, K.W.; Lassau, G.; Rentzel, P. (2018): "The lay of land: Strontium isotope variability in the dietary catchment of the Late Iron Age proto-urban settlement of Basel-Gasfabrik, Switzerland", *Journal of Archaeological Science: Reports*, 17: 279-292.

Brönnimann, D.; Wimmer, J.; Müller-Kissing, M.; Stopp, B.; Rissanen, H.; Spichtig,

N. (2020): “One man’s trash is another man’s treasure. Interdisciplinary examination of taphonomic aspects of ceramic sherds, animal bones and sediments from the La Tène period settlement at Basel-Gasfabrik”, *PloS one*, 15 (7), e0236272.

Büntgen, U.; Myglan, V.S.; Ljungqvist, F.C.; McCormick, M.; Di Cosmo, N.; Sigl, M.; Jungclaus, J.; Wagner, S.; Krusic, P.J.; Esper, J.; Kaplan, J.O.; Vaan, M.A.C.; de, Luterbacher, J.; Wacker, L.; Tegel, W.; Kirdyanov, A.V. (2016): “Cooling and societal change during the Late Antique Little Ice Age from 536 to around 660 AD”, *Nature Geoscience*, 9 (3): 231-236.

Carbiener, R.; Schnitzler, A. (1990): “Evolution of major pattern models and processes of alluvial forest of the Rhine in the rift valley (France/Germany)”, *Vegetatio,* 88:115-129.

Carrero-Pazos, M.; Bevan, A.; Lake, M.W. (2019): “The spatial structure of Galician megalithic landscapes (NW iberia): A case study from the Monte Penide region”, *Journal of Archaeological Science*, 108 (3): 104968.

Casarotto, A.; Stek, T.D.; Pelgrom, J.; van Otterloo, R.H.; Sevink, J. (2018): “Assessing visibility and geomorphological biases in regional field surveys: The case of Roman Aesernia”, *Geoarchaeology*, 33 (2): 177-192.

Chisholm, M. (1979) [2007]: *Rural Settlement and Land Use*, Taylor and Francis, Somerset.

Cowley, D.C. (2016): “What Do the Patterns Mean? Archaeological Distributions and Bias in Survey Data”, in M. Forte and S. Campana (Eds.), *Digital Methods and Remote Sensing in Archaeology, Quantitative Methods in the Humanities and Social Sciences*, Springer International Publishing, Cham: 147-170.

Crema, E.R.; Bevan, A.; Lake, M.W. (2010): “A probabilistic framework for assessing spatio-temporal point patterns in the archaeological record”, *Journal of Archaeological Science*, 37 (5): 1118-1130.

Depaermentier, M.L.; Kempf, M.; Bánffy, E.; Alt, K.W. (2021): “Modelling a scale-based strontium isotope baseline for Hungary”, *Journal of Archaeological Science*, 135 (1-2): 105489.

Depaermentier, M.L.C.; Kempf, M.; Bánffy, E.; Alt, K.W. (2020): “Tracing mobility patterns through the 6th-5th millennia BC in the Carpathian Basin with strontium and oxygen stable isotope analyses”, *PloS one*, 15 (12): e0242745.

Erfurt, M.; Glaser, R.; Blauhut, V. (2019): “Changing impacts and societal responses to drought in southwestern Germany since 1800”, *Regional Environmental Change*, 19 (8):2311-2323.

Erfurt, M.; Skiadaresis, G.; Tijdeman, E.; Blauhut, V.; Bauhus, J.; Glaser, R.; Schwarz, J.; Tegel, W.; Stahl, K. (2020): “Exploring the added value of a long-term multidisciplinary dataset in drought research – a drought catalogue for southwestern Germany dating back to 1801”, *Natural Hazards and Earth System Sciences Discussions*.

Faustmann, A.C. (2007): Besiedlungswandel im südlichen Oberrheingebiet von der Römerzeit bis zum Mittelalter, Zugl.: Freiburg (Breisgau) Univ., Diss., 2005, *Freiburger Beiträge zur Archäologie und Geschichte des ersten Jahrtausends*, Vol. 10, Leidorf, Rahden/Westf.

Froehlicher, L.; Schwartz, D.; Ertlen, D.; Trautmann, M. (2016): "Hedges, colluvium and lynchets along a reference toposequence (Habsheim, Alsace, France): history of erosion in a loess area", *Quaternaire,* 27 (2): 173-185.

Giacona, F.; Martin, B.; Furst, B.; Glaser, R.; Eckert, N.; Himmelsbach, I.; Edelblutte, C.; Christiany, L. (2018): "Improving the understanding of flood risk in the Alsatian region by knowledge capitalization: the ORRION participative observatory", *Natural Hazards and Earth System Sciences Discussions*: 1-49.

Gibson, J.J. (1979): *The Ecological Approach to Visual Perception*, Houghton Mifflin, Boston, MA.

Gillings, M.; Hacigüzeller, P.; Lock, G. (Eds.) (2020): *Archaeological spatial analysis: A methodological guide*.

Glaser, R.; Riemann, D.; Schönbein, J.; Barriendos, M.; Brázdil, R.; Bertolin, C.; Camuffo, D.; Deutsch, M.; Dobrovolný, P.; van Engelen, A.; Enzi, S.; Halíčková, M.; Koenig, S.J.; Kotyza, O.; Limanówka, D.; Macková, J.; Sghedoni, M.; Martin, B.; Himmelsbach, I. (2010): "The variability of European floods since AD 1500", *Climatic Change*, 101 (1-2): 235-256.

Günther, G.; Kempf, M.; Knitter, D. (2022): rbias, Zenodo. 10.5281/zenodo.7071418

Hamer, W.; Knitter, D. (2018): FuzzyLandscapes -- Fuzzy analyses with a focus on raster data, Zenodo. 10.5281/zenodo.1747005

Hijmans, R.J. (2022): terra: Spatial Data Analysis.

Himmelsbach, I.; Glaser, R.; Schoenbein, J.; Riemann, D.; Martin, B. (2015a): "Flood risk along the upper Rhine since AD 1480", *Hydrology and Earth System Sciences Discussions*, 12 (1): 177-211.

Himmelsbach, I.; Glaser, R.; Schoenbein, J.; Riemann, D.; Martin, B. (2015b): "Reconstruction of flood events based on documentary data and transnational flood risk analysis of the Upper Rhine and its French and German tributaries since AD 1480", *Hydrology and Earth System Sciences*, 19 (10): 4149-4164.

Inrap (2016): "De l'archéologie de sauvetage à l'archéologie préventive", available at: https://www.inrap.fr/de-l-archeologie-de-sauvetage-l-archeologie-preventive-9724

Kempf, M. (2018): "Migration or landscape fragmentation in Early Medieval eastern France? A case study from Niedernai", *Journal of Archaeological Science: Reports*, 21: 593-605.

Kempf, M. (2019a): "Paradigm and pragmatism: GIS-based spatial analyses of Roman infrastructure networks and land-use concepts in the Upper Rhine Valley", *Geoarchaeology*, 74 (285): 1-12.

Kempf, M. (2019b): "The application of GIS and satellite imagery in archaeological land-use reconstruction: A predictive model?", *Journal of Archaeological Science: Reports*, 25: 116-128.

Kempf, M. (2020a): "Fables of the past: landscape (re-)constructions and the bias in the data", *Documenta Praehistorica*, 47: 476-492.

Kempf, M. (2020b): "From landscape affordances to landscape connectivity: contextualizing an archaeology of human ecology", *Archaeological and Anthropological Sciences*, 12 (8): 310.

Kempf, M. (2020c): "Modelling multivariate landscape affordances and functional ecosystem connectivity in landscape archaeology", *Archaeological and Anthropological Sciences*, 12: 159.

Kempf, M. (2020d): "Neolithic land-use, landscape development, and environmental dynamics in the Carpathian Basin", *Journal of Archaeological Science: Reports*, 34: 102637.

Kempf, M. (2021): "Take a seed! Revealing Neolithic landscape and agricultural development in the Carpathian Basin through multivariate statistics and environmental modelling", *PloS one*, 16 (10): e0258206.

Kempf, M.; Glaser, R. (2020): "Tracing Real-Time Transnational Hydrologic Sensitivity and Crop Irrigation in the Upper Rhine Area over the Exceptional Drought Episode 2018-2020 Using Open Source Sentinel-2 Data", *Water*, 12 (12): 3298.

Kempf, M.; Günther, G. (under review): "A methodological framework for spatial analyses using archaeological and environmental data – a case study from the Carpathian Basin", *Journal of Archaeological Science: Reports*.

Kempf, M.; Günther, G. (2023): "Point pattern and spatial analyses using archaeological and environmental data – A case study from the Neolithic Carpathian Basin", *Journal of Archaeological Science: Reports*, 47: 103747.

Knappett, C. (2004): "The Affordances of Things: a Post-Gibsonian Perspective on the Relationality of Mind and Matter", in E. DeMarrais, C. Gosden and C. Renfrew (Eds.), *Rethinking materiality: The engagement of mind with the material world*, McDonald Institute monographs, McDonald Institute for Archaeological Research, Cambridge: 43-51.

Knipper, C.; Pichler, S.L.; Brönnimann, D.; Rissanen, H.; Rosner, M.; Spichtig, N.; Stopp, B.; Rentzel, P.; Röder, B.; Schibler, J.; Lassau, G.; Alt, K.W. (2018): "A knot in a network: Residential mobility at the Late Iron Age proto-urban centre of Basel-Gasfabrik (Switzerland) revealed by isotope analyses", *Journal of Archaeological Science: Reports*, 17: 735-753.

Koehler, H.; Angevin, R.; Bignon-Lau, O.; Griselin, S. (2013): "Découverte de plusieurs occupations du Paléolithique supérieur récent dans le Sud de l'Alsace", *Bulletin de la Société préhistorique française*, 110 (2): 356-359.

Koerner, W.; Dupouey, J.L.; Dambrine, E.; Benoit, M. (1997): "Influence of Past Land Use on the Vegetation and Soils of Present Day Forest in the Vosges Mountains, France", *Journal of Ecology*, 85 (3): 351-358.

Kreiter, A. (2021): "The Hungarian Archaeology Database", *Internet Archaeology*.

Laabs, J.; Knitter, D. (2021): "How Much Is Enough? First Steps to a Social Ecology of the Pergamon Microregion", *Land*, 10 (5): 479.

Lang, A.; Bork, H.-R.; Mäckel, R.; Preston, N.; Wunderlich, J.; Dikau, R. (2003): "Changes in sediment flux and storage within a fluvial system: some examples from the Rhine catchment", *Hydrological Processes*, 17 (16): 3321-3334.

Lehmkuhl, F.; Zens, J.; Krauss, L.; Schulte, P.; Kels, H. (2016): "Loess-paleosol sequences at the northern European loess belt in Germany: Distribution, geomorphology and stratigraphy", *Quaternary Science Reviews*, 153: 11-30.

Mäckel, R.; Schneider, R.; Friedmann, A.; Seidel, J. (2002): "Environmental changes and human impact on the relief development in the Upper Rhine valley and

Black Forest (South-West germany) during the Holocene", *Zeitschrift für Geomorphologie N.F.*, 128: 31-45.

Mäckel, R.; Schneider, R.; Seidel, J. (2003): "Anthropogenic Impact on the Landscape of Southern Badenia (Germany) During the Holocene - Documented by Colluvial and Alluvial Sediments", *Archaeometry*, 45 (3): 487-501.

Martin, B.; Giacona, F.; Furst, B.; Edelblutte, C.; Holleville, N.; With, L.; Heitz, C.; Glaser, R.; Himmelsbach, I.; Schönbein, J.; Bösmeier, A. (2017): "La variabilité spatio-temporelle des inondations dans le Fossé rhénan à la lumière de l'évolution de la vulnérabilité", *VertigO La revue électronique en sciences de l'environnement*, 17 (1).

Martin, B.; Himmelsbach, I.; Glaser, R.; With, L.; Guerrouah, O. (2010): "French – German flood risk geohistory in the Rhine Graben", *International Disaster Risk Conference*, Davos 2010, hal-00530788.

McCormick, M.; Büntgen, U.; Cane, M.A.; Cook, E.R.; Harper, K.; Huybers, P.; Litt, T.; Manning, S.W.; Mayewski, P.A.; More, A.F.M.; Nicolussi, K.; Tegel, W. (2012): "Climate Change during and after the Roman Empire: Reconstructing the Past from Scientific and Historical Evidence", *The Journal of Interdisciplinary History*, 43 (2): 169-220.

Minářová, J.; Müller, M.; Clappier, A. (2017a): "Seasonality of mean and heavy precipitation in the area of the Vosges Mountains: dependence on the selection criterion", *International Journal of Climatology*, 37 (5): 2654-2666.

Minářová, J.; Müller, M.; Clappier, A.; Kašpar, M. (2017b): "Characteristics of extreme precipitation in the Vosges Mountains region (north-eastern France)", *International Journal of Climatology*, 37 (13): 4529-4542.

Mischka, D. (2007): Methodische Aspekte zur Rekonstruktion prähistorischer Siedlungsmuster: Landschaftsgenese vom Ende des Neolithikums bis zur Eisenzeit im Gebiet des südlichen Oberrheins, Zugl.: Freiburg (Breisgau). Univ., Diss., 2004, *Freiburger archäologische Studien*, Vol. 5, Leidorf, Rahden, Westf.

Nakoinz, O.; Knitter, D. (2016): *Modelling Human Behaviour in Landscapes*, Springer International Publishing, Cham.

O'Sullivan, D.; Unwin, D.J. (2010): *Geographic Information Analysis*, John Wiley & Sons, Inc, Hoboken, NJ, USA.

Preusser, F. (2008): "Characterisation and evolution of the River Rhine system", *Netherlands Journal of Geosciences - Geologie en Mijnbouw*, 87 (1): 7-19.

Preusser, F.; May, J.-H.; Eschbach, D.; Trauerstein, M.; Schmitt, L. (2016): "Infrared stimulated luminescence dating of 19th century fluvial deposits from the upper Rhine River", *Geochronometria*, 43 (1): 131-142.

Purtill, M.P. (2022): "More than Just Clovis: the Broad Impact of Sampling Bias on Archaeological Site Distributions", *Journal of Archaeological Method and Theory*.

Rentzel, P.; Preusser, F.; Pümpin, C.; Wolf, J.-J. (2009): "Loess and palaeosols on the High Terrace at Sierentz (France). and implications for the chronology of terrace formation in the Upper Rhine Graben", *Swiss Journal of Geosciences*,

102 (3): 387-401.
Ripley, B.D. (1977): "Modelling Spatial Patterns", *Journal of the Royal Statistical Society: Series B (Methodological)*, 39 (2): 172-192.
Schiffer, M.B. (1996): *Formation Processes of the Archaeological Record*, University of Utah.
Schmitt, L.; Maire, G.; Nobelis, P.; Humbert, J. (2007): "Quantitative morphodynamic typology of rivers: a methodological study based on the French Upper Rhine basin", *Earth Surface Processes and Landforms*, 32 (11): 1726-1746.
Stork, M.; Menzel, L. (2016): "Analysis and simulation of the water and energy balance of intense agriculture in the Upper Rhine valley, south-west Germany", *Environmental Earth Sciences*, 75 (16): 57.
van Leusen, P.M. (1996): "Unbiasing the Archaeological Record", *Archeologia e Calcolatori*, 7: 129-136.
Verhagen, J.W.H.P. (2007): Case studies in archaeological predictive modelling, *Archaeological studies Leiden University*, Vol. 14, Leiden Univ. Press, Leiden.
Verhagen, P. (2018): "Predictive Modeling", in S.L. López Varela (Ed.), *The encyclopedia of archaeological sciences*, Vol. 19, Wiley-Blackwell, Malden, MA: 1-3.
Verhagen, P.; Kamermans, H.; van Leusen, M.; Deeben, J.; Hallewas, D.; Zoetbrood, P. (2010): "First Thoughts on the Incorporation of Cultural Variables into Predictive Modelling", in F. Niccolucci and S. Hermon (Eds.), *Beyond te Artifact: Digital Interpretation of the Past*, Proceedings of CAA2004, Archaeolingua, Budapest: 307-311.
Verhagen, P.; Whitley, T.G. (2012): "Integrating Archaeological Theory and Predictive Modeling: a Live Report from the Scene", *Journal of Archaeological Method and Theory*, 19 (1): 49-100.
Verhagen, P.; Whitley, T.G. (2020): "Predictive spatial modelling", in M. Gillings, P. Hacigüzeller and G. Lock (Eds.), *Archaeological spatial analysis: A methodological guide*, Routledge, New York: 231-246.
Weaverdyck, E.J.S. (2019): "The Role of Forts in the Local Market System in the Lower Rhine: Towards a Method of Multiple Hypothesis Testing Through Comparative Modelling", in P. Verhagen, J. Joyce and M.R. Groenhuijzen (eds.), *Finding the Limits of the Limes*, Springer International Publishing, Cham: 165-190.
Wetter, O.; Pfister, C.; Weingartner, R.; Luterbacher, J.; Reist, T.; Trösch, J. (2011): "The largest floods in the High Rhine basin since 1268 assessed from documentary and instrumental evidence", *Hydrological Sciences Journal*, 56 (5): 733-758.