Archaeological and Anthropological Sciences (2024) 16:137
https://doi.org/10.1007/s12520-024-02049-7

RESEARCH

Living in the Mountains. Settlement patterns in Northwestern Iberia during the Palaeolithic period

Mikel Díaz-Rodríguez[1,2]

Received: 13 February 2024 / Accepted: 18 July 2024 / Published online: 1 August 2024

Abstract
Despite the presence of a theoretical model describing the settlement patterns of Palaeolithic sites in Northwestern Iberia, it has not yet been empirically tested using statistical analysis. This study explores the settlement patterns of the Palaeolithic period in Northwestern Iberia within two regions that share similar chronology and research traditions: the Northern and Central Mountain ranges of Northwestern Iberia. Employing Geographic Information Systems (GIS) and spatial statistics, the methodology has provided robust empirical support for several aspects of the theoretical model. The study rigorously tested the theoretical model proposed in the existing literature using statistical analysis and a comprehensive dataset of 50 variables. The findings highlight significant regional distinctions in the settlement patterns of Palaeolithic sites within both areas of Northwestern Iberia. This research not only confirms certain hypotheses related to Palaeolithic site locations but also underscores the need for further examination and refinement of others, particularly considering the notable regional variations.

Keywords Settlement patterns · GIS · Spatial statistics · Palaeolithic · Northwestern Iberia

Introduction

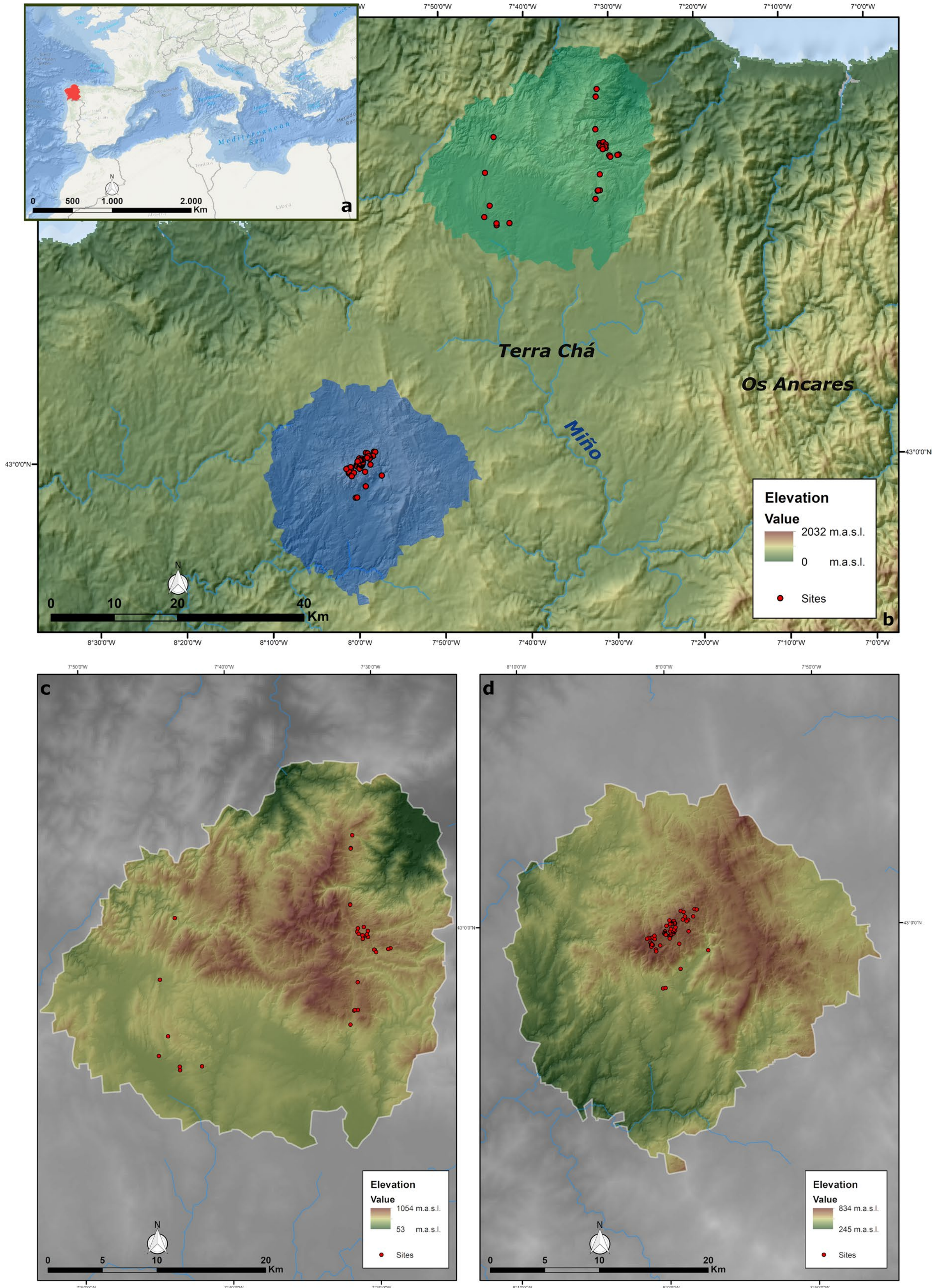
The study of Palaeolithic settlement patterns has garnered substantial scholarly attention in recent times owing to its capacity to yield copious insights into past societies (e.g. Turrero et al. 2013; Burke et al. 2014; 2017; 2021b; García Moreno and Fano Martínez 2014; Ludwig et al. 2018; Wren and Burke 2019). Indeed, there are tools available that enable us to reconstruct ancient landscapes, with the aim of quantifying environmental factors that may have influenced settlement patterns in Palaeolithic hunter-gatherer societies. Among these tools are Geographic Information Systems (GIS) and spatial statistics, invaluable for analysing and quantifying data, thereby enabling us to draw insightful conclusions regarding the priorities of these societies when they inhabited particular locations (Bevan et al. 2013). The use of GIS and various statistical tools has provided a robust framework for analysing the settlement patterns. These methodologies, combined with a comprehensive set of biotic and abiotic variables, offer significant insights into the prehistoric occupation of the region. The potential of these tools and the analytical results obtained in this study highlight the importance of applying such methods in other study areas and across different chronological periods. This approach can yield valuable comparative data, enhancing our understanding of settlement dynamics in various contexts. Nevertheless, investigations of this nature are conspicuously scarce when considering the Northwestern region of the Iberian Peninsula, save for a limited number of preliminary inquiries (de Lombera Hermida et al. 2015; Díaz Rodríguez 2017; Díaz Rodríguez and Carrero Pazos 2019; Díaz-Rodríguez et al. 2021, 2023; Díaz-Rodríguez and Fábregas-Valcarce 2022).

In recent decades, a series of investigations have unveiled archaeological sites in the Northwestern region of the Iberian Peninsula that had hitherto eluded scholarly detection. Particularly during the latter stages of the Palaeolithic period

✉ Mikel Díaz-Rodríguez
mikel.diaz@usc.es

1 Grupo de Estudos para a Prehistoria do Noroeste Ibérico, Arqueoloxía, Antigüidade e Territorio (GEPN-AAT), Dpto. de Historia, Universidade de Santiago de Compostela, Praza da Universidade, nº 1, 15782 Santiago de Compostela, Spain

2 CISPAC – Centro de Investigación Interuniversitario das Paisaxes Atlánticas, Edificio Fontán, Cidade da Cultura de Galicia, Monte Gaiás, S/N, 15707 Santiago de Compostela, Spain


a
0
500
1.000
2.000
Km
Terra Chá
Os Ancares
Miño
Elevation
Value
2032 m.a.s.l.
0 m.a.s.l.
Sites
b
0
10
20
40
Km
c
Elevation
Value
1054 m.a.s.l.
53 m.a.s.l.
Sites
0
5
10
20
Km
d
Elevation
Value
834 m.a.s.l.
245 m.a.s.l.
Sites
0
5
10
20
Km

◂**Fig. 1** **a**) General location of Galicia region (red polygon). **b**) Northern Mountain ranges (area in green) with Upper Palaeolithic and Epipalaeolithic sites (red dots) and Central Mountain ranges (area in blue) with Upper Palaeolithic and Epipalaeolithic sites (red dots). **c**) Detail of Northern Mountain ranges area with sites (red dots). **d**) Detail of Central Mountain ranges area with sites (red dots)

and within specific areas, such as the Central Mountain ranges (Criado Boado et al. 1988, 1989a, 1991a, b) and Northern Mountain ranges (Ramil Soneira and Vázquez Varela 1983; Llana Rodríguez et al. 1992; López Cordeiro 2003). A theoretical model has been formulated regarding the settlement patterns of archaeological sites based on the data acquired from various sites within the previously mentioned regions (Cerqueiro Landín 1989; Criado Boado and Cerqueiro Landín 1991). However, the validation of this theoretical model remains outstanding, as contemporary analytical tools have yet to be employed to ascertain its conformity.

This article delves into an exploration of various environmental variables that could have shaped the settlement patterns of Upper Pleistocene sites within two distinct regions of the Northwestern Iberian Peninsula. A thorough analysis of 50 locational variables has been conducted to ascertain whether unique patterns of settlement emerge within the aforementioned regions or if a single overarching pattern prevails. It is plausible that specific variables, as proposed within the theoretical framework, hold significance in site selection. The identification of these variables offers valuable insights into the intricacies of hunter-gatherer interactions with their environment. The variables have been meticulously modelled using GIS and the outcomes derived from sites within each region have undergone thorough analysis and subsequent statistical comparison against values generated under randomized conditions. This analytical process is complemented by a statistical comparison with values generated under randomized conditions, with the primary aim of elucidating whether the choice of archaeological sites may be influenced by, or bear any relationship to, these variables. The complete methodological procedure employed in this study, spanning from the inception of the variables to their subsequent statistical analysis, is entirely replicable, in strict accordance with the principles of reproducible research (Marwick 2017; Marwick et al. 2017; Karoune and Plomp 2022). Moreover, the code and data, essential for this reproducibility, are readily accessible and can be found in the Data Availability section.

Regional setting

This study is focused on two distinct regions situated in the Northwest of the Iberian Peninsula (Fig. 1a): the Northern Mountain ranges and the Central Mountain ranges (Fig. 1b). The Northern Mountain ranges form a natural corridor that separates the northeastern coastal sector of Galicia from the inland Terra Chá basin, extending southwards into the current province of Lugo (Fig. 1c and Fig. 19 in the Supplementary Material). This area has undergone intensive archaeological study since the mid-1970s to the 1990s, resulting in the discovery of over 50 archaeological sites, primarily from the Upper Palaeolithic and Epipalaeolithic periods (Ramil Rego and Ramil Soneira 1996).

The second region encompasses the Galician Central Mountain Ranges (Fig. 1d and Fig. 20 in the Supplementary Material), with a notable concentration of archaeological sites in the O Bocelo mountain range and along the Furelos River. A research project conducted in the 1980s in this specific area aimed to explore its occupation history from the Palaeolithic era to medieval times, employing Landscape Archaeology methods. As a result, more than 80 archaeological sites have been identified in this region, primarily attributed to the Upper Palaeolithic and Epipalaeolithic periods (Cerqueiro Landín 1989; Criado Boado and Cerqueiro Landín 1991).

Northern Mountain ranges

The study of the Upper Palaeolithic and Epipalaeolithic periods in Galicia finds its roots in the exploration of the Northern Mountain ranges. This era began to be systematically investigated during the 1970s and remained a focal point of research throughout the 1980s. These investigations introduced updated archaeological methodologies, systematic action plans and interdisciplinary collaboration aimed at addressing specific objectives and challenges. These efforts were primarily led by Medical Doctor J. Ramil Soneira, a passionate advocate for archaeological heritage, especially the Palaeolithic period. He collaborated with researchers associated with the University of Santiago de Compostela (USC) to initiate surveying activities in various municipalities of Lugo, an area that had hitherto been overlooked by scientific inquiry (Senín Fernández 1996, p. 34).

This early groundwork led to the formation of a research team in the mid-1980s. Their investigations extended across contemporary municipalities such as Muras, Vilalba, Xermade, the Serra do Xistral and the Arnela River Valley. The team devised comprehensive survey programs and archaeological excavation campaigns, supplemented by soil and palaeobotanical analyses (Llana Rodríguez et al. 1992; Martínez Cortizas and Moares Domínguez 1995; Ramil Rego and Fernández Rodríguez 1996).

The investigations carried out allowed the discovery of the first sites such as Pena Grande or Prado do Inferno (Alonso del Real and Vázquez Varela 1976) in the 1970s. The subsequent surge in archaeological activities revealed more sites, including Férvedes, Xestido 3 and the A Veiga and Piñeiro Flint Workshops, all unearthed in the early 1980s (Ramil Soneira and Vázquez Varela 1976). The initial phase of work in the region culminated with the excavation of the Férvedes II shelter (Ramil Soneira and Vázquez Varela 1983), where a stone pendant from the early Magdalenian period was uncovered. The pendant's decorative elements draw parallels with contemporaneous findings in the Cantabrian Region, such as La Paloma, Altamira, El Castillo and Balmori (Ramil Soneira and Vázquez Varela 1983; Villar Quinteiro 1997).

The mid-1980s marked the onset of a second phase of investigation, focused on reconstructing the paleoenvironment and characterizing human settlement in the Northern Mountain ranges. This phase involved an extensive survey and excavation initiative, resulting in the identification of over twenty archaeological sites (Ramil Rego 2014).

Research activities in the area were interrupted in 1994 (Ramil Rego and Ramil Soneira 1996), but investigations continued through large-scale public and private initiatives (López Cordeiro 2003). In the late 1990s, the implementation of the Galician Strategic Wind Plan posed a threat to the O Xistral Mountain range's integrity. Several companies commissioned projects to evaluate and mitigate the archaeological impact of wind farms, entrusting these tasks to GIARPA (Research Group within the Laboratory of Archaeology and Cultural Forms at USC). Unlike previous endeavours, this project sought not only to understand the Galician Palaeolithic in the area but also to minimize its impact.

Some companies contracted projects for the evaluation and correction of the archaeological impact of wind farms to GIARPA. The intervention strategy was carried out following an action program that included two action plans: an applied research plan and a basic research plan (López Cordeiro 2015).

In tandem with these archaeological investigations, typological studies of lithic industries recovered from specific sites in the region were conducted (Villar Quinteiro 1996, 1997, 2008). Valuable insights from these studies and findings unearthed during the 1970s to 1990s have recently come to light through the defence of two doctoral theses (Ramil Rego 2014; López Cordeiro 2015).

Central Mountain ranges

At the end of the 1980s, a research project was conducted in the O Bocelo mountain range and the Furelos River valley. This area was chosen for its favourable conditions for this work. Directed by the USC under the leadership of F. Criado Boado, this project embraced the principles of Landscape Archaeology. It aimed to achieve several objectives: illuminate poorly understood periods of Galician prehistory, define settlement characteristics from the final Upper Palaeolithic to the Middle Ages, ascertain the impact of the natural environment on human communities, reconstruct ecological changes in the natural environment throughout the Holocene, conduct a diachronic, historical and cultural analysis of the Galician landscape over the last 10000 years, identify key periods in the transformation of the Galician landscape, test a theoretical-interpretative model for spatial analysis and human positioning within it, perform an evaluative analysis of prehistoric periods in an area to assess the percentage and type of remains that were not located during extensive works and systematically study sites and cultural periods that were previously unknown to define their conditions and facilitate their discovery in future works.

The project's work plan spanned five years. The first two years (1987 and 1988) focused on fieldwork, including intensive surveying during the initial phase (from October 1987 to March 1988) and archaeological excavations in extension during the second phase (in July and August 1988). These activities targeted megalithic and Iron Age sites and involved surveys based on small pits in sites from the Palaeolithic/Epipalaeolithic, Chalcolithic and medieval periods. Subsequently, materials and data collected during these excavations were processed (from November 1988 to June 1989). The third phase (July and August 1989) extended the previous campaign, continuing with excavation in extension, surveys and intensive prospecting. Physical–chemical surveying methods were also incorporated to address the issue of archaeological sites without visible structures (Criado Boado et al. 1991a, b).

Since the project adopted a Landscape Archaeology perspective, the survey aimed to identify archaeological sites and included an environmental survey to gather geographical and ecological data complementing the archaeological information. Information from environmental documentation and physical–chemical samples of archaeological documentation were separated into distinct files, each specific to one of the two groups. During the survey, the concept of primary sites (referring to locations where remains occupy their original position, are identifiable and correspond to a specific site) and secondary sites (materials displaced from their primary position) was abandoned in favour of archaeological points (PA). PA referred to all locations in the workspace where archaeological material appeared, irrespective of whether it was found in situ or not. PAs had their dedicated files, but there was also a file for documenting environmental conditions (CA). In addition to the PA concept, the term 'dispersion area' (DISP) was introduced to describe specific sectors within a PA where concentrations of archaeological material or specific structures were abundant. Locating a PA inherently implied the existence of at least one DISP, even if more material concentrations were subsequently identified within that PA and were numbered and incorporated accordingly (Criado Boado et al. 1991a, b).

Specifically concerning the Palaeolithic period, the project aimed to define settlement patterns from the Late Glacial to the early Holocene (Criado Boado et al. 1988, 1989a, b; Cerqueiro Landín 1989; Criado Boado 1991; Criado Boado and Cerqueiro Landín 1991). To achieve this, a set of factors and criteria that could be controlled to study the landscape of these hunter-gatherer societies were established. These factors included absolute and relative altitude, proximity to wetland areas, visibility and proximity between sites, among others (Cerqueiro Landín 1989, pp. 50–57). For the first time in the Northwest of Iberia, variables that could influence the location of Palaeolithic sites were defined.

Upon completion of the project, it was concluded that a complex system of territorial organization revolved around wet and depressed areas. Sites were found to cluster around these areas (Cerqueiro Landín 1989, p. 48,85; Criado Boado et al. 1991a, b, p. 99). Three settlement patterns were proposed: 1) Sites situated in the lowest regions of the area, very close to wetland areas, with limited access. These sites had a visual domain over them but did not extend to more distant sites in the vicinity. 2) Another group was located at higher absolute and relative altitudes, farther from wetland areas, affording them a broad visual domain over a vast territory and distant sites, although not those in the immediate vicinity. 3) Finally, a smaller group comprised sites associated with streams that flowed into the valley. These sites had no direct relationship with wetland areas, given their distance and topographic location on the opposite side of the dividing lines that marked the depressed areas.

Material and methods

Data acquisition and software

The archaeological sites used in this study were sourced from the General Directorate of Cultural Heritage of the Xunta de Galicia (DXPCXG). This public entity maintains a comprehensive catalogue encompassing all the archaeological sites within the Galician autonomous community. Additionally, some site information was gleaned from publications arising from research projects associated with the aforementioned regions. It is important to note that the data employed for this study originates from diverse sources, lacking uniformity. Consequently, the information found within the DXPCXG catalogue and the research projects may differ based on their initial objectives. The DXPCXG catalogue primarily aims to preserve and protect archaeological evidence, while research projects focus on studying, understanding and disseminating this data.

This diversity in data sources has led to the identification of location inaccuracies in some coordinates within the DXPCXG catalogue. These inaccuracies result from the passage of information through various hands and the absence of clear criteria during data collection. Something as fundamental as establishing a universal coordinate system for all researchers to report their findings was not initially considered. This omission has necessitated additional work to reconcile the various coordinate systems used. As a prudent approach, it was decided to work with sites that had precise location data. Incorporating archaeological sites with coordinates varying by hundreds of meters could compromise the study's integrity. Therefore, sites present both in the DXPCXG catalogue and referenced in at least one publication derived from research and emergence projects in both study areas were selected.

While resolving the coordinate issue was relatively straightforward, the reduction in the number of usable sites posed a significant challenge. Additionally, various other challenges should be considered, although some remain cannot be controlled. Each mentioned project had distinct origins and objectives, influencing data collection methods. These projects were conducted in different decades when information and technology varied significantly. Therefore, the project in the Northern Mountain ranges, initiated as a personal endeavour and later joined by various entities, differs substantially from the project in the Central Mountain ranges. The latter project had specific objectives, a defined methodology and a predetermined duration. Moreover, preventive archaeology, driven by public works in Galicia, adapted well in the Central Mountain ranges, evolving into a multidisciplinary Landscape Archaeology project. Nevertheless, the primary focus in this region was not the Palaeolithic period. It was viewed as one element contributing to the understanding of landscape transformations by ancient societies. However, this diachronic perspective may have limited the information gathered, given the project's extensive chronological framework. It is logical to assume that in such a large-scale project, some information had to be sacrificed for the greater good. In this case, the methodology used for identification, while suitable for other periods, might not have been ideal for the Palaeolithic. Subsequently, researchers recognized this limitation when adapting the methodology to their subsequent work in Palaeolithic chronology within professional archaeology (López Cordeiro 2015).

Despite these differences, there were some similarities in data collection methodologies, particularly regarding surveying, which became essential for locating remains of hunter-gatherer communities. Unlike vestiges from other eras, these sites cannot be identified through other means. However, archaeological surveying posed challenges due to inaccessible areas and the tendency of lithic tools, the primary remains from this period, to be buried. Their appearance is often linked to areas where land was manipulated for agricultural purposes, further complicating their discovery. In any case, archaeological surveys in both areas began with a preliminary identification of potential sites based on the locations of shelters, which may have led to the settlement of prehistoric communities. Consequently, the presence of biases in data collection is acknowledged and should be taken into account. Nevertheless, it is assumed that the same bias exists in both areas since the methodology for locating sites was similar, centred on archaeological surveys. Therefore, this bias should not significantly affect the comparison of results between the two areas.

The identification of Palaeolithic sites is associated with the recognition of scattered lithic industry remains across the landscape. This presents challenges in determining the functional category of these sites because reaching the extent of the archaeological sites is often difficult. Material dispersion could result from sedimentological or edaphic effects, extensive occupation of space with unknown boundaries, or repeated occupation. While the position and orientation of the pieces provide insights into their origin, their extent remains uncertain unless extensive excavation is conducted. To address this issue, the project in the Central Mountain ranges established the categories of PA (archaeological site) and DISP (dispersion of materials). However, this criterion was not applied to the other study area. Consequently, the diverse origins of data from both projects could influence the analysis of settlement patterns. For instance, one PA may contain multiple DISPs. To mitigate this, an attempt has been made to unify the archaeological points in both areas, maintaining the site category. In cases where a PA contains several DISPs, each is treated as a distinct site, unless their proximity is so close that differentiation becomes impractical.

Another challenge arises from the fact that some sites discovered during archaeological surveys have been categorized as archaeological sites or PAs with varying quantities of pieces, ranging from fewer than five to over 100. In the Central Mountain ranges, the average number of lithic pieces found at each site is approximately 10. Excavated sites, however, exhibited significantly larger quantities of lithic remains: PA 154.1 with 525 pieces, PA 149.1 with 78 remains, PA 74.1 with 350 lithic pieces and PA 69.1 with 200 remains. While exact data on piece counts for sites in the Northern Mountain ranges are unavailable, the overall pattern in both areas is similar. Non-excavated sites typically contain a handful to a few dozen pieces, while excavated sites feature more than a hundred pieces. It is important to note that an archaeological site with only three lithic tools does not hold the same significance as one with hundreds of artifacts. However, this determination can only be confirmed through archaeological intervention. Nevertheless, the mere presence of more lithic remains in one site compared to another may suggest different functionalities. While there was some uniformity in conducting archaeological surveys at sites with higher concentrations of pieces and suitable topography for archaeological level preservation, it should be noted that the same treatment was applied to the archaeological points during analysis. Without this approach, there would be very few archaeological sites to work with. In essence, establishing a threshold, such as more than one hundred pieces, would exclude most sites from the analysis, rendering it unfeasible. While this challenge lacks a feasible solution, it remains important to consider when analysing the data as a whole and within each study area.

In the Northern Mountain ranges, the study includes 34 archaeological sites, categorized into two types: sheltered sites (N = 23) and open-air sites (N = 11). These sites span various chronocultural phases, including Epipalaeolithic (N = 26), Lower-Middle Magdalenian (N = 5), Azilian (N = 2) and Magdalenian (N = 1) (see Table 1 in the Supplementary Material). In the Central Mountain ranges, 61 archaeological sites have been included, similarly divided into sheltered sites (N = 31) and open-air sites (N = 30). However, specific chronocultural information for these sites is lacking. According to the literature, these archaeological sites are broadly categorized under the Final Upper Palaeolithic/Epipalaeolithic period (refer to Table 2 in the Supplementary Material).

The chronological attributions of each site have been based on previous works and research projects conducted in both study areas. The concerns regarding the detailed characterization of the materials from these sites are acknowledged. However, comprehensive studies of the recovered materials do not exist, and it was not feasible to conduct a thorough review of these materials for this work. Furthermore, the primary aim of this study was not to reassess the material culture but to analyse settlement patterns. The inherent complexity and partial characterization of the archaeological record are acknowledged, and future research should aim to address these gaps by conducting detailed material analyses.

Spatial data analysis was conducted using various software applications designed for Geographic Information Systems (GIS). The chosen coordinate system was EPSG: 25829 (ETRS89 / UTM zone 29N). GRASS GIS has been used in versions 6.4.3, 7.0.2 and 7.0.4 (Grass Development Team 2020). Quantum GIS (versions 2.8.1 and 2.10.1) (QGIS.org 2021) and SAGA GIS (version 2.2.1) (Conrad et al. 2015) have also been used. The latest GIS software employed has been ArcGIS 10.3 (USC license) (Esri 2011). It is the only one that does not share the GNU-GPI license. Finally, to carry out the different analytical approaches, R version 4.0.5 was used, with the R Studio graphical interface (R Core Team 2021) and the requisite packages for conducting the various analyses (Table 1).

The Digital Elevation Model (DEM) has been used as a base map to elaborate the different locational variables analysed in this paper. This DEM has been obtained from the National Centre for Geographic Information (CNIG) and has a resolution of 25 m. It is cartography that collects the information obtained from the photogrammetric and LiDAR flights of the National Plan for Aerial Orthophotography (PNOA) (http://centrodedescargas.cnig.es/CentroDescargas/index.jsp). Additionally, the Geologic Map was accessed and retrieved from the Spanish Government's online repository (López Olmedo et al. 2022).

Spatial distribution of sites

The initial step in conducting spatial analysis of archaeological sites in both areas involved an assessment of whether Complete Spatial Randomness (CSR) was present. To achieve this, a random sample of points, equating to the same number of points as the archaeological dataset, was generated for the purpose of comparison. Subsequently, the distribution of both datasets was examined to determine if they could be considered as originating from the same population, signifying the absence of significant differences between the two datasets and thereby precluding the rejection of CSR. To conduct this analysis, it was used the UTM X variable, representing the x-coordinate of each point and performed normality tests using the Shapiro–Wilk test. It was also assessed whether both datasets belonged to the same population using the Kolmogorov–Smirnov (K-S) test. Furthermore, Ripley's K functions and their L and G variants were employed (Bivand et al. 2013). Both homogeneous and inhomogeneous K, L and G functions were computed (Baddeley et al. 2015), utilizing a confidence interval derived from Monte Carlo simulations ($n = 99$).

Definition of covariates

In preparation for the statistical analysis, a set of covariates was meticulously chosen, building upon insights gleaned from prior research conducted in the study areas and analogous regions. The total number of variables selected for inclusion stands at 50. These variables can be categorized into three primary classes of influencing factors: abiotic, biotic and other determinant factors. These, in turn, have been further grouped into overarching variables such as altitude, slope, hydrology, or geology, among others. Subsequently, the covariates utilized in this study will be described below. A more comprehensive explanation of the process employed to obtain each covariate is available in the Supplementary Material (SM) file.

Altitude

Altitude refers to the elevation calculated with reference to specific data points, often sea level for absolute altitude or, alternatively, the base of a valley for relative altitude. In the context of defining the occupation patterns of Galician Palaeolithic sites, altitude has been one of the key factors considered. The prevailing notion suggests that archaeological sites tend to be situated at higher elevations in the landscape (Ramil Rego 1989/1990, p. 194; Fábregas

Table 1 Synthesis of R packages used, authors and application details

Package	Author/s	Application Details
dismo	Hijmans et al. (2017)	Methods for species distribution modelling
foreach	Daniel et al. (2022)	Package for the foreach looping construct
geostatsp	Brown (2015)	Geostatistical Modelling with Likelihood and Bayes
GGally	Schloerke et al. (2021)	This package is a plotting system based on the grammar of graphics
ggplot2	Wickham et al. (2020)	Package for creating graphics
maps	Becker et al. (2021)	Display of maps
maptools	Bivand et al. (2020a)	Set of tools for manipulating geographic data
MASS	Ripley et al. (2020)	Functions and datasets to support Venables and Ripley
patchwork	Pedersen (2022)	Package for combining multiple plots
plyr	Wickham (2020)	Set of tools that solves problems relates with applying or combining data
purrrlyr	Henry (2022)	This package allows some functions at the intersection of “dplyr” and “purrr”
raster	Hijmans et al. (2020)	Reading, writing, manipulating, analysing and modelling of spatial data
readxl	Wickham et al. (2019a)	Package for read excel files
Rcmdr	Fox et al. (2023)	A platform-independent basic-statistics GUI (graphical user interface) for R
rgdal	Bivand et al. (2020a)	Provides bindings to the “GDAL” and “PROJ” library
rgeos	Bivand et al. (2020c)	Package for topology operations on geometries
sp	Pebesma et al. (2020)	Classes and methods for spatial data
spatstat	Baddeley et al. (2020)	Toolbox for analysing Spatial Point Patterns
tidyverse	Wickham et al. (2019a, b)	Data representations and API design

Valcarce et al. 2010, p. 267; de Lombera Hermida et al., 2015, p. 285). Absolute altitude (ALTA) represents a general variable, from which other derived variables such as ALTm (Table 2) have been computed. ALTm corresponds to the average elevation within the four adjacent cells surrounding each site. This variable is particularly valuable for gaining an overview of the site's environmental context, especially in the case of open-air sites where materials may not necessarily be in situ.

In addition to the absolute altitude, the relative altitude will be considered. On the one hand it is going to calculate the topographic prominence, which has been defined, in a GIS environment, by M. Llobera (2001) as "the function of the differential height between an individual and the environment as it is perceived from the point of view of the individual in question." The focus of this study builds upon prior research that underscores the significance of topographic prominence as a locational variable, particularly in contexts such as Early Prehistory, Protohistory and the Iron Age (Carrero Pazos 2017; Cazorla Martín et al. 2008; Cerrillo Cuenca 2011; De Reu 2012; De Reu et al. 2013, 2011; Fábrega Álvarez 2004; Parcero Oubiña and Fábrega Álvarez 2006; among others).

In the specialize literature, it has been considered that Palaeolithic sites would be acting as landmarks in the landscape. These would be reference points that would stand out from the surrounding terrain and would be visible at a certain distance (Fábregas Valcarce and de Lombera Hermida 2010). It has also been considered that there was intervisibility between these sites (López Cordeiro 2002, 2004a, 2015). However, as the contemporaneity of these archaeological sites cannot be confirmed, intervisibility has not been analysed into the present study.

For modelling topographic prominence, the Topographic Prominence Index (TPI) obtained from SAGA GIS software was employed. Following the recommendations of domain experts (Nakoinz and Knitter 2016) and based on a prior work in the Northern Mountain ranges (Díaz Rodríguez and Carrero Pazos 2019), TPI was calculated at three different radii (100, 500 and 1000 m) (TPI100, TPI500 and TPI1000). Additionally, TPI was computed considering the average value of the four adjacent cells for each site at different radii (TPI100m, TPI500m and TPI1000m).

Finally, the relative altitude has been considered. This variable is deemed crucial for the analysis of the strategic potential of Palaeolithic sites as it quantifies precise

Table 2 Covariates used related with altitude, conditioning type, variables, acronym, description and ID number

Conditioning type	Variables	Acronym	Description	ID Number
Abiotic	Absolute altitude	ALTA	Elevation at a given point, in meters, above sea level. It is calculated from a DEM	v1
Abiotic	Average altitude	ALTm	Average elevation in the four neighbouring cells adjacent to each site, in meters, above sea level	v2
Abiotic	Topographic Prominence Index	TPI100	Calculation of the TPI that consists of comparing the elevation of each of the cells of the DEM with the average of the surrounding elevations. Obtained for 100 m radii	v3
Abiotic	Average Topographic Prominence Index	TPI100m	From the TPI100, the average value of the four neighbouring cells adjacent to each site is calculated. Obtained for 100 m radii	v4
Abiotic	Topographic Prominence Index	TPI500	Calculation of the TPI that consists of comparing the elevation of each of the cells of the DEM with the average of the surrounding elevations. Obtained for 500 m radii	v5
Abiotic	Average Topographic Prominence Index	TPI500m	From the TPI500, the average value of the four neighbouring cells adjacent to each site is calculated. Obtained for 500 m radii	v6
Abiotic	Topographic Prominence Index	TPI1000	Calculation of the TPI that consists of comparing the elevation of each of the cells of the DEM with the average of the surrounding elevations. Obtained for 1000 m radii	v7
Abiotic	Average Topographic Prominence Index	TPI1000m	From the TPI1000, the average value of the four neighbouring cells adjacent to each site is calculated. Obtained for 1000 m radii	v8
Abiotic	Relative altitude	ALTrA	Maximum relative altitude index calculated within the 20-min isochrone. It is calculated from the DEM	v9
Abiotic	Relative altitude	ALTrB	Minimum relative altitude index calculated within the 20-min isochrone. It is calculated from the DEM	v10

positioning and elucidates whether a site occupies the highest terrain within a designated radius. While some studies have calculated the relative altitude index using a fixed radius of 1000 m from the archaeological site (Marcos Sáiz 2006, pp. 49–50), this paper adapted this analysis by introducing a limit determined by travel cost distance rather than Euclidean distance. For this purpose, a 20' isochrone was used, which corresponds to a radius of 1000 m. Two indices were computed: the maximum relative altitude index (ALTrA) and the minimum relative altitude index (ALTrB). Both variables are complementary. The value of the first variable indicates whether the site is located in the highest part of the established environment or not. The values range from 0 to 1, with 0 being the lowest area and 1 being the highest. The closer its value is to 1, the more it stands out in that territory. The combination of the two variables can indicate whether the site is located in a very homogeneous spatial area. It could be the case that the same archaeological site has the highest index in both variables and although it may seem contradictory, it would indicate that, within the defined area, it would be located in a completely flat area, where the absolute height of the site would be very similar to the maximum and minimum relative height. This could reveal whether there are significant differences or not.

Slope

Slope represents the maximum degree of elevation change at a specific location and is derived from the DEM. This covariate has been a subject of consideration in prior studies of the Galician Palaeolithic (de Lombera Hermida et al. 2015, p. 280) as well as in other regions of the Iberian Peninsula, such as the Cantabrian (García Moreno 2010), Asturian (Fernández Fernández 2010) and the Sierra de Atapuerca (Marcos Sáiz 2006).

Similar to altitude, slope serves as a general variable that can be employed to calculate various derived variables (refer to Table 3). In this study, the absolute slope (SLO) and the average slope of the terrain within the four cells adjacent to each site (SLOm) have been included. Calculating slope in neighbouring cells can be closely associated with site accessibility.

Moreover, a set of slope indices has been employed concerning the site's surroundings (refer to Table 3). These indices, originally defined by F. J. Marcos Sáiz (2006) and adapted to this study, include the Geomorphological slope area index (SLOga), the Theoretical slope index (SLOt), the Steepest true slope index (SLOst) and the Plateau index (SLOpi). Finally, another relevant variable linked to slope is accessibility. Accessibility, in this context, refers to the conditions provided by a surface for movement from a specific point, taking into account both distance and surface characteristics (Fábrega Álvarez 2004, p. 16). It has been considered to evaluate the proximity to resources and the defensive potential of archaeological sites (for further details on these variables, please refer to the SM file).

Aspect

Aspect has been recognized as a significant covariate for determining the location of archaeological sites (de Lombera Hermida et al. 2015, p. 289). In the Northern Mountain ranges, it has been observed that the majority of sites are oriented toward the second and third quadrants (Ramil Rego and Ramil Soneira 1996). Aspect data has been derived from the DEM and the aspect in the cell where each site is situated (ASP) has been considered, as well as the average aspect of the four cells surrounding each site (ASPm) (refer to Table 4).

Hydrology

Hydrology is another critical variable in this study. Previous research has suggested a strong association between Palaeolithic sites and watercourses (Ramil Rego 1989/1990, p. 193; Villar Quinteiro 1996; Fábregas Valcarce and de Lombera

Table 3 Covariates used related with slope, conditioning type, variables, acronym, description and ID number

Conditioning type	Variables	Acronym	Description	ID Number
Abiotic	Slope	SLO	Slope at a given point, in degrees. It is calculated from a DEM	v11
Abiotic	Average slope	SLOm	Average slope in the four neighbouring cells adjacent to each site, in degrees. It is calculated from a DEM	v12
Abiotic	Geomorphological slope area index	SLOga	Index of the slope of the geomorphological area at the 20' isochrone	v13
Abiotic	Theoretical slope index	SLOt	Theoretical slope index at the 10' isochrone	v14
Abiotic	Steepest true slope index	SLOst	Index of the steepest true slope at the 10' isochrone	v15
Abiotic	Plateau index	SLOpi	Plateau index at the 10' isochrone	v16
Abiotic	Accessibility	INCr15-45	Increase in the accessible surface (between the isochrone of 15' and that of 45')	v17

Hermida, 2010, p. 267). Additionally, a connection between archaeological sites and wetland areas, which could have served as attracting locations for hunting resources, has also been considered (Criado Boado et al. 1991a, b; de Lombera Hermida et al., 2015; López Cordeiro, 2002, p. 72).

However, the current hydrology map shows evidence of human-induced alterations. Therefore, this paper proposes to analyse a potential hydrographic network derived from theoretical points that would topographically be more likely to have served as water accumulation sites. This methodology has been used and explained in other works (García García 2015; Díaz-Rodríguez et al. 2023). The process of obtaining this potential hydrology map is detailed further in the SM file.

From the potential hydrology, various variables have been derived for analysis (refer to Table 5). Site proximity to potential hydrology has been measured from all points within each study area to the nearest watercourse, both in terms of distance (HYDROE) and displacement cost time (HYDROC). Although straight-line distance is an idealized measure, it has been utilized in the literature on settlement patterns. Therefore, it is included to determine its effectiveness as a variable. Proximity to hydrology has also been considered by examining the four cells surrounding each site in a straight line (HYDROEm) and displacement cost time (HYDROCm).

As mentioned earlier, previous studies have considered proximity to wetland areas and visual control over these areas (Criado Boado et al. 1991a, b; López Cordeiro 2002, p. 72; de Lombera Hermida et al. 2015; López Cordeiro 2015). Wetland areas can be defined as regions with accumulated water points. The SAGA GIS software was used to model this variable, particularly the *Topographic Wetness Index* (TWI), which indicates the topographic humidity index in each map cell. After obtaining the map and identifying areas with higher humidity values (considered most interesting), quartiles were calculated. This retained map cells with values above the third quartile. It is important to note that TWI values are substantially elevated in cells where rivers coincide, but these cells are not of interest. Thus, the hydrological map, created earlier, was subtracted from the TWI polygon map to retain higher humidity values in areas not intersected by rivers. Subsequently, displacement cost time was calculated from every point in the study area to the nearby wetland areas identified as points (WET). The cost of displacement to nearby wetland areas was also calculated by considering the average of the four neighbouring cells for each site (WETm).

Table 4 Covariates used related with aspect, conditioning type, variables, acronym, description and ID number

Conditioning type	Variables	Acronym	Description	ID Number
Abiotic	Aspect	ASP	Aspect at a given point, in degrees. It is calculated from a DEM	v18
Abiotic	Average aspect	ASPm	Average aspect in the four neighbouring cells adjacent to each site, in degrees. It is obtained from a DEM	v19

Table 5 Covariates used related with hydrology, ID number, conditioning type, variables, acronym and description

Conditioning type	Variables	Acronym	Description	ID Number
Abiotic	Distance to hydrology	HYDROE	Distance, in meters, of the closest river channel to each site. It is obtained from a DEM	v20
Abiotic	Average distance to hydrology	HYDROEm	Average distance in the four neighbouring cells adjacent to each site, in meters, from the nearest river channel. It is obtained from a DEM	v21
Abiotic	Cost to hydrology	HYDROC	Displacement cost, in time, of the closest river channel to each site. It is obtained from a DEM	v22
Abiotic	Average cost to hydrology	HYDROCm	Average displacement cost in the four neighbouring cells adjacent to each site, in time, from the nearest river channel. It is obtained from a DEM	v23
Abiotic	Visibility of hydrology	HYDROV	River courses visible from each site	v24
Abiotic	Cost to wetland areas	WET	Displacement cost, in time, of the closest wetland areas to each site. It is obtained from a DEM	v25
Abiotic	Average cost to wetland areas	WETm	Average displacement cost in the four neighbouring cells adjacent to each site, in time, from the closest wetland areas to each site. It is obtained from a DEM	v26
Abiotic	Visibility of wetland areas	WETv	Visibility of the wetland areas from each site. It is obtained from a DEM	v27

Lastly, a variable based on the visibility of wetland areas from each site (WETv) was included. This variable was obtained by identifying cells coinciding with wetland areas visible from each archaeological site. The visible surface area for each site was calculated in hectares (ha).

Geology

Another variable that was considered in the present study is potential geology. For the Palaeolithic hunter-gatherers it was important to have raw materials such as quartzite or quartz, since these are the raw materials used by these societies in this region of NW Iberia (Llana Rodríguez 1990; Villar Quinteiro 1997). These resources could be obtained from the fluvial courses or by going to collect raw materials from the veins where this material was found. Geology has been identified as an important variable in previous studies (Ramil Rego 1989/1990; Villar Quinteiro 1996; López Cordeiro, 2002; 2015; Fábregas Valcarce and de Lombera Hermida, 2010; de Lombera Hermida et al., 2012). The geology map has been obtained from the Mining Geological Institute (IGME). On this map, those areas that could contain raw materials of interest to Palaeolithic hunter-gatherers have been identified. The cells of interest have been selected and divided into points at established radii. Subsequently, the cost of moving, in time (GEOLC) and distance (GEOLE), from the rest of the cells in the study area to the closest potential geology points has been calculated (see SM file for further details).

Due to the acidic nature of Galician soils, faunistic remains that would provide taphonomic information are typically only preserved in the limestone cavities of north-eastern Galicia. The study areas analysed in this work consist solely of rock shelters and open-air sites, which do not preserve faunistic remains. Consequently, faunistic data are not available for these areas.

Additionally, paleoenvironmental information is very limited because few sites have been excavated and studied in detail. While the IGME geological map lacks detailed archaeological information, it has been used in this study as a proxy to consider potential geological resources. This map is referred to as potential geology throughout the manuscript and no definitive assumptions are made regarding the use of quartz or quartzite as raw material based solely on this data.

In addition to the maps of travel cost in time and distance, other variables related to geology have been obtained (Table 6). The average of the four cells adjacent to each archaeological site has been calculated for both variables (GEOLEm and GEOLCm). Finally, the visible potential geological surface, in ha, from each site (GEOLV) has been quantified.

CPFPC

Within the context of biotic conditioning, the variable related to travel costs to potential hunting areas was considered. The Central Place Foraging Prey Choice (CPFPC) model, originally proposed by M. Cannon (2003), which is rooted in foraging theory was applied. This model was previously utilized by Marín Arroyo to investigate mobility patterns and territorial control in the eastern Cantabrian region, focusing on deer and goats, the most prominent species in the Magdalenian diet (Marín Arroyo 2008, 2009).

In this study, potential goat hunting areas are analysed and a covariate was created based on travel cost in time from any point within the study area to these prospective goat and deer hunting zones. To maximize productivity, a maximum travel time threshold of 2.15 h was set for deer hunting, while a 1.2-h threshold was established for goat hunting. The 1.2-h and 2.15-h isochrones were calculated for each site and within these isochrones, the slope map was used to identify cells with slopes greater than 30º for goats and less than 30º for deer (refer to the SM file for further details on the calculation process).

To assess the significance of this conditional factor, the total area of each site in ha that represents potential

Table 6 Covariates used related with geology, ID number, conditioning type, variables, acronym and description

Conditioning type	Variables	Acronym	Description	ID Number
Abiotic	Distance to geology	GEOLE	Distance, in meters, of the closest cell with potential geology interest to each site. It is obtained from a DEM	v28
Abiotic	Average distance to geology	GEOLEm	Average distance in the four neighbouring cells adjacent to each site, in meters, from the nearest potential geology interest area. It is obtained from a DEM	v29
Abiotic	Cost to geology	GEOLC	Displacement cost, in time, of the closest cell with potential geology interest to each site. It is obtained from a DEM	v30
Abiotic	Average cost to geology	GEOLCm	Average displacement cost in the four neighbouring cells adjacent to each site, in time, from the nearest potential geology interest area. It is obtained from a DEM	v31
Abiotic	Visibility of geology	GEOLV	Cells, in m^2, with potential geology visible from each site	v32

hunting areas for goats (CPFPCGs) and deer (CPFPCDs) was calculated. Additionally, the slope map was used to determine the cost of travel from each site to reach the potential catchment areas for goats (CPFPCGc) and deer (CPFPCDc) (Table 7).

Visibility

Various types of conditioning factors have been considered in this study, including visibility, which has been a critical aspect in defining the occupation of archaeological sites in prior research (López Cordeiro, 2002; 2004; 2015; Rodríguez Álvarez et al. 2008; Fábregas Valcarce and de Lombera Hermida, 2010; de Lombera-Hermida et al. 2011; de Lombera Hermida et al., 2015). Different approaches related to visibility have been used in the present study. Firstly, it was calculated the number of cells visible from each archaeological site (VISC) and the number of cells in the study area from which each site is visible (VISZ). Additionally, it was explored visual prominence, which entails computing the points that are visible from specific locations. To achieve this, it was used an observer height of 1.75 m. The outcome of this calculation is a raster file in which cells are assigned values indicating the number of cells visible from each location (VISPR) (refer to the SM file for further methodological details). Furthermore, the average value of visibility from the four adjacent cells for each site (VISPRm) was obtained (Table 8).

Potential least cost path

Another significant conditional factor considered in this study is the calculation of potential least cost paths (LCP). These paths are a representation of the relationships between archaeological sites and movement within the surrounding landscape, essentially indicating the routes of least energy or time cost between two points.

In previous literature, natural transit routes, often referred to as *royal roads*, have been explored in relation to Palaeolithic sites. These routes have been identified as a key variable influencing the distribution of Palaeolithic sites in the Northwestern Iberia (Ramil Rego and Ramil Soneira 1996, p. 125; Fábregas Valcarce and de Lombera Hermida 2010, p. 267; de Lombera Hermida et al. 2015, p. 289; López Cordeiro 2015, p. 301; Díaz Rodríguez 2017).

In this study, it has been computed transit routes within a specific area without considering the archaeological sites themselves. To perform this analysis, it is necessary to have starting and ending points. This approach is inspired by a previous work that utilized the *From Everywhere to Everywhere* (FETE) methodology, which calculates optimal routes between all points in the landscape (White and Barber 2012). In FETE, every point in a grid serves as both a starting and ending point, enabling comprehensive route calculations across the entire territory. However, this method demands substantial computational power. As an alternative, a simplified model based on a methodology from another study was adopted (Rodríguez Rellán and Fábregas Valcarce 2015).

Table 7 Covariates used related with CPFPC, conditioning type, variables, acronym, description and ID number

Conditioning type	Variables	Acronym	Description	ID Number
Biotic	Surface of potential goat hunting	CPFPCGs	Total area, in m^2, of the potential catchment area for goat hunting. Based on the CPFPC	v33
Biotic	Cost to potential goat hunting areas	CPFPCGc	Displacement cost, in minutes, of the closest cell with potential goat hunting to each site. It is obtained from a DEM	v34
Biotic	Surface of potential deer hunting	CPFPCDs	Total area, in m^2, of the potential catchment area for deer hunting. Based on the CPFPC	v35
Biotic	Cost to potential deer hunting areas	CPFPCDc	Displacement cost, in minutes, of the closest cell with potential deer hunting to each site. It is obtained from a DEM	v36

Table 8 Covariates used related with visibility, conditioning type, variables, acronym, description and ID number

Conditioning type	Variables	Acronym	Description	ID Number
Other	Visibility	VISC	Number of cells visible from each site. It is obtained from a DEM	v37
Other	Visibilization	VISZ	Number of cells from which each site is observed. It is obtained from a DEM	v38
Other	Visual prominence	VISPR	Number of cells visible in a study area for each site in that bounded area. It is obtained from a DEM	v39
Other	Average visual prominence	VISPRm	Average of the visual prominence value in the four neighbouring cells adjacent to each site. It is obtained from a DEM	v40

This simplified approach has been described in greater detail elsewhere (Díaz Rodríguez 2017; Díaz-Rodríguez et al. 2021, 2023). In brief, it involves dividing the study area's boundary into points spaced at a predetermined radius and calculating LCPs between them. It was considered one point as the starting point and the others as stopping points, repeating this analysis with all points. Subsequently, the travel cost in minutes from each site to the nearest route (LCPC) was determined. Additionally, it was calculated the average value for the four neighbouring cells adjacent to each site (LCPCm) (Table 9).

Potential insolation

Insolation, or the exposure to sunlight, has been a consideration in the context of Northwestern Palaeolithic sites. It has been suggested that these sites were often situated on west-facing slopes to maximize the heat from the sun's rays (Ramil Rego and Ramil Soneira 1996, p. 125). While insolation has not been extensively discussed in the literature regarding this region, it has been a factor in the selection of occupation sites in other parts of the Iberian Peninsula. For instance, some studies have examined its influence on site selection in the Asón River Valley area of Cantabria (García Moreno 2008, 2015).

In this study, insolation data was obtained using SAGA GIS and the *Potential Incoming Solar Radiation* module (Conrad et al. 2015). Several insolation-related variables were calculated for each site, including the received potential solar radiation (TOTINS), the direct solar radiation received in the cell where each site is located (DIRINS) and the diffuse solar radiation received in the cell of each site (DIFINS). Additionally, values for these variables were obtained for the four neighbouring cells adjacent to each site (TOTINSm, DIRINSm and DIFINSm) (Table 10).

Potential wind

The final conditioning factor considered in the analysis is related to potential wind exposure. The literature suggests that protection from prevailing winds may also be a factor in the selection of occupation sites by hunter-gatherer societies (Villar Quinteiro 1996; García Moreno 2010; de Lombera Hermida et al. 2015, p. 290). To quantify this variable, the wind exposure index was obtained using the *Wind Effect Index* from the SAGA GIS module (Böhner and Antonić 2009; Conrad et al. 2015) for each site (WIND). Additionally, the average value of this index for the four neighbouring cells adjacent to each site was calculated (WINDm) (Table 11).

Table 9 Covariates used related with least cost path, conditioning type, variables, acronym, description and ID number

Conditioning type	Variables	Acronym	Description	ID Number
Other	Cost to least cost path	LCPC	Displacement cost, in minutes, of each cell to the least cost path. It is calculated from a DEM	v41
Other	Average to cost of the least cost path	LCPCm	Average of the displacement cos of each cell to the least cost path. It is calculated from a DEM.	v42

Table 10 Covariates used related with the potential insolation, conditioning type, variables, acronym, description and ID number

Conditioning type	Variables	Acronym	Description	ID Number
Other	Total insolation	TOTINS	Total potential solar radiation received in the cell in which each site is located. It is obtained from a DEM	v43
Other	Average total insolation	TOTINSm	Average value of the four neighbouring cells adjacent to each site for the total insolation. It is obtained from a DEM	v44
Other	Direct insolation	DIRINS	Direct potential solar radiation received in the cell in which each site is located. It is obtained from a DEM	v45
Other	Average direct insolation	DIRINSm	Average value of the four neighbouring cells adjacent to each site for the direct insolation. It is obtained from a DEM	v46
Other	Diffuse insolation	DIFINS	Diffuse potential solar radiation received in the cell in which each site is located. It is obtained from a DEM	v47
Other	Average diffuse insolation	DIFINSm	Average value of the four neighbouring cells adjacent to each site for the diffuse insolation. It is obtained from a DEM	v48

Results

Complete spatial randomness

With the aim of analysing Complete Spatial Randomness (CSR), it was first assessed whether the distribution of archaeological sites in both areas follows a normal distribution. This was done by choosing UTMX as the variable for this comparison between the actual site data and randomly generated points. The Shapiro–Wilk test was employed for this purpose and the results indicate a lack of normality in both areas. Specifically, the p-value was less than 0.05 in both the Northern Mountain ranges (W = 0.63203, p-value = 5.181^{e-06}) and the Central Mountain ranges (W = 0.95885, p-value = 0.03876) (Table 12).

Furthermore, when comparing the archaeological sample with a randomly generated sample, by performing a K-S Test, it revealed that the two samples belong to different populations in both the Northern Mountain ranges (D = 0.58824, p-value = 1.555^{e-05}) and the Central Mountain ranges (D = 0.42623, p-value = 3.077^{e-05}). In the K, L and G homogenous functions graphs (see Fig. 2 and Fig. 3), the black line does not closely align with the confidence interval, indicating that H_0 (the null hypothesis) can be rejected in favour of H_1 (the alternative hypothesis), as CSR is not supported. In the inhomogeneous K and L functions, it is evident that site clustering occurs up to 4 km in the case of sites in the Northern Mountain ranges and up to 6 km for sites in the Central Mountain ranges.

Evaluation of locational variables

This section includes an evaluation of the variables presented above. Some types of statistical analysis have been conducted based on the nature of each variable. In other words, for those variables that have been modelled, resulting in a raster map, they were analysed by comparing the values of archaeological sites with random values for each study area. For absolute variables, the values were compared with 999 random samples and resampling density plots were created, following previous approaches (Bocinsky 2017; Cascalheira et al. 2022).

The results for average variables obtained from the raster layers (mean value of the 4 cells adjacent to the site) are presented for each zone and compared with the results of the random points using box plots and statistical analysis. An equal number of random points was generated as archaeological sites in each of the study areas and the results were compared for each variable to statistically evaluate the mean variables defined in the Material and Methods section. In cases where normality exists, the Shapiro–Wilk test was applied to check normality and Fisher's F test was used to check homoscedasticity. The T-Student test was also used when the variances were equal and the Welch test when the variances were different. The Mann–Whitney-Wilcoxon test was used for those samples that did not present normality.

Subsequently, both zones are compared and for this purpose, all the variables are analysed, including the results of specific variables that have not been obtained from a raster file but through calculation. The aim is to identify trends by presenting the results for each area individually and then comparing the results of the two areas together.

Northern area

There are 19 variables associated with raster data that have been assessed both graphically and statistically. For the variable ALTA (Fig. 4a), it is evident that the spatial distribution of archaeological sites deviates from what would

Table 11 Covariates used related with the potential wind, ID number, conditioning type, variables, acronym and description

Conditioning type	Variables	Acronym	Description	ID Number
Other	Wind effect index	WIND	Potential wind effect index obtained in the cell in which each site is located. It is obtained from a DEM	v49
Other	Average wind effect index	WINDm	Average value of the four neighbouring cells adjacent to each site for the wind effect index. It is obtained from a DEM	v50

Table 12 Results for the statistical tests applied in both areas

	Shapiro–Wilk Test (sites)		Shapiro–Wilk Test (random sites)		K-S Test (sites vs random sites)	
Region	W	*p-value*	W	*p-value*	D	*p-value*
Northern Mountain ranges	0.63203	5.181^{e-08}	0.96294	0.2957	0.58824	1.555^{e-05}
Central Mountain ranges	0.95885	0.03876	0.96351	0.06598	0.42623	3.077^{e-05}

be expected under random conditions. This deviation is supported by the Mann–Whitney U test, which returned a p-value < 0.05 (Table 13). Other variables like TPI100, SLO and HYDROE (Fig. 4b-d) also show deviations in the values of archaeological sites from random expectations, though statistical analysis with the Mann–Whitney U test did not yield significant differences (Table 13).

Among the studied variables analysed using this method, some display visually distinct spatial distributions for archaeological sites compared to random conditions, yet they do not exhibit statistically significant differences (Table 13). This is observed in the cases of WET, GEOLE, GEOLC and VISPR (Fig. 5), as well as LCPC, WIND and CPFPCGc (Fig. 6).

Other variables studied showed neither visual nor statistical distinctions when comparing the archaeological sample to random simulations. This is correct for variables like TPI500, TPI1000 and HYDROC (Fig. 11 in the SM). Similarly, locational variables like TOTINS and DIRINS (Fig. 12a-b in the SM) demonstrated this absence of visual and statistical differences. The variable ASP, on the other hand, did not exhibit apparent visual differences (Fig. 11c in the SM) but did show statistical significance with a p-value of 0.046 (Table 13). Similar observations were made for the variable DIFINS, which displayed no clear visual differences (Fig. 12c in the SM) but did exhibit statistical significance with a p-value of 0.008 (Table 13). However, it is essential to note that the variable CPFPCDc presented significant differences both visually and statistically. However, caution is needed when interpreting this variable as all archaeological sites have a value of 0, rendering it unusable (Fig. 12d in the SM and Table 13).

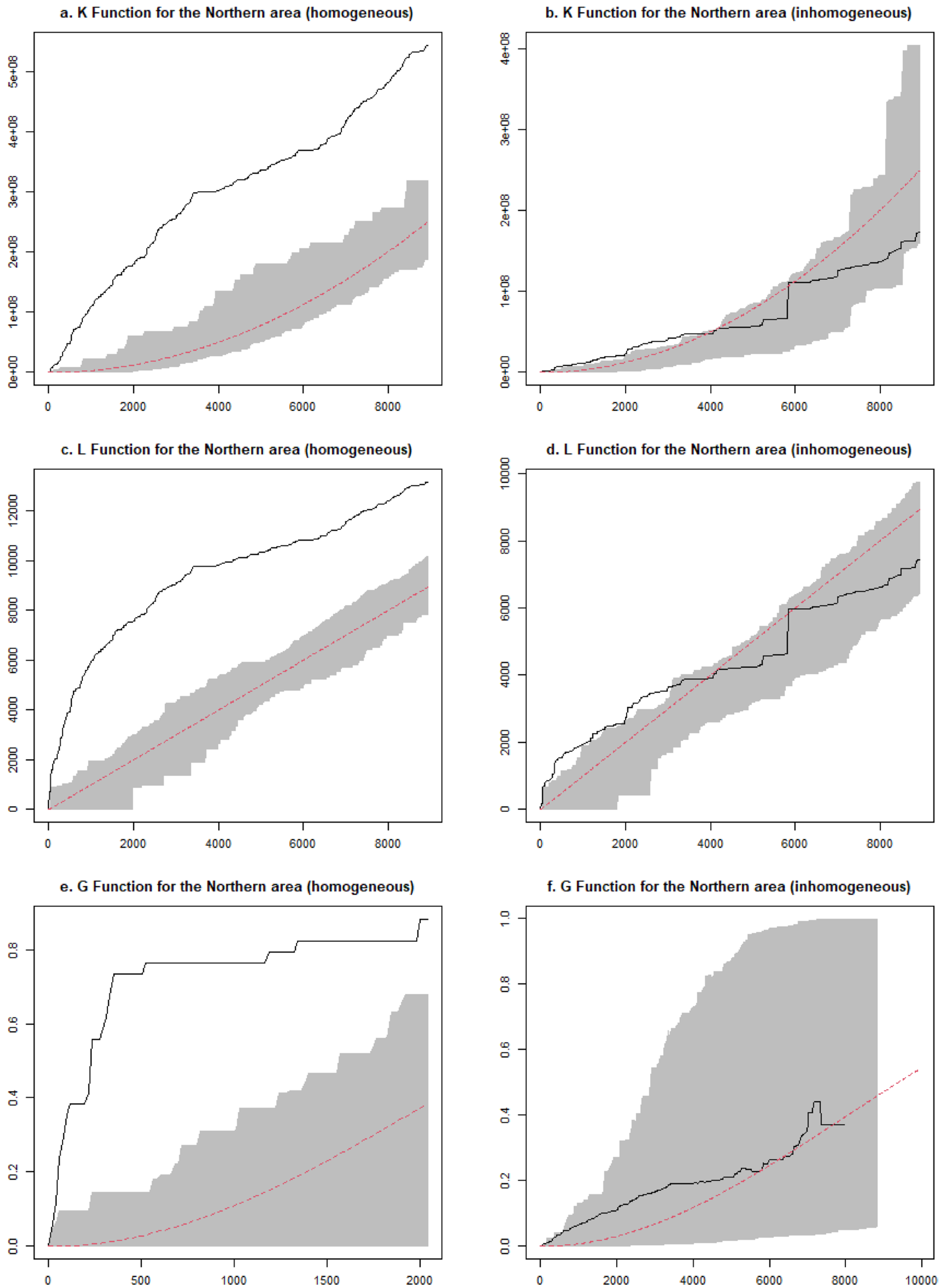
For the variable ALTm, the archaeological sites exhibit a minimum altitude of 451 m, while the random sites register a minimum of 261 m. The median altitude for the archaeological sites is 688.5 m and for the random sites, it is 527 m. The highest recorded altitude for archaeological sites is 758 m, whereas for random sites, it is 734 m. Notably, the majority of archaeological site altitudes cluster between 600 and 750 m, as depicted in Fig. 7a. The Shapiro–Wilk test for normality indicates that the distributions of both datasets are non-normal (p-value = 0.04185). Consequently, a non-parametric Mann–Whitney-Wilcoxon test was employed to examine homoscedasticity, which confirmed differing variances (p-value = 0.00001762). Subsequently, the T Welch test supported the assertion that both sets of data originate from distinct populations (p-value = 0.00001234). The results of the statistical analyses for this variable and for the following variables studied in this area can be checked in Table 3 and Table 4 in the SM file.

In the case of TPI100m, the minimum values for archaeological sites and random points are -4.32 and 9.5, respectively. The median TPI100m value for archaeological sites is 1.05, whereas for random points, it is 0.52. The maximum TPI100m value recorded for archaeological sites is 8.36 and for random points, it is 9.40. Notably, both datasets predominantly exhibit values clustered around 0, as represented in Fig. 7b. The Shapiro–Wilk normality test indicated non-normal distributions for both datasets (p-value = 0.008938). Regarding homoscedasticity, the Fisher's F-test suggested that the variances of both datasets differ insignificantly (p-value = 0.468). Consequently, a Student's t-test was conducted to evaluate differences in means, which concluded that both datasets are drawn from the same population, thereby not displaying statistically significant distinctions (p-value = 0.1943).

For the TPI500m variable, archaeological sites have a minimum value of -3.88, whereas random points show a minimum value of -47.89. The median TPI500m value for archaeological sites is 0.49, while for random points, it is -1.21. The maximum TPI500m values recorded are 10.55 for archaeological sites and 39.96 for random points. Both datasets exhibit a concentration of values around 0, as illustrated in Fig. 7c. The Shapiro–Wilk normality test confirmed normal distributions for both datasets (p-value = 0.09752). Consequently, the Fisher's F-test was utilized to assess homoscedasticity, indicating similar variances (p-value = 0.463). Subsequently, a Student's t-test was employed to compare means, revealing that both datasets belong to the same population, with no statistically significant differences (p-value = 0.5105).

In the case of the TPI1000m variable, archaeological sites display a minimum value of -3.88, while random points exhibit a minimum value of -111.29. The median TPI1000m value for archaeological sites is 0.49 and for random points, it is -3.68. The maximum TPI1000m values are 10.55 for archaeological sites and 57.96 for random points. Similar to the previous variable, both datasets demonstrate a concentration of values near 0, as presented in Fig. 7d. The Shapiro–Wilk test for normality confirms non-normal distributions for both datasets (p-value = 0.02265). The Mann–Whitney-Wilcoxon test, assessing homoscedasticity, concludes that both datasets have similar variances (p-value = 0.7839), while the Student's t-test establishes that they belong to the same population, without statistically significant differences (p-value = 0.5793).

Regarding the SLOm variable, the minimum slope value for archaeological sites is 0.07, whereas for random points, it is 0.22. The median slope value for archaeological sites is 9.96 and for random sites, it is 7.07. The maximum slope value for archaeological points is 25.25 and for random points, it is 27.10. The dispersion of values for archaeological sites spans between 5 and 15, which is slightly broader than the range for random points, concentrated between 4 and 11, as visualized in Fig. 7e. Statistical analysis indicates that both datasets do not adhere to a normal distribution


a. K Function for the Northern area (homogeneous)
b. K Function for the Northern area (inhomogeneous)
c. L Function for the Northern area (homogeneous)
d. L Function for the Northern area (inhomogeneous)
e. G Function for the Northern area (homogeneous)
f. G Function for the Northern area (inhomogeneous)
0e+00
1e+08
2e+08
3e+08
4e+08
5e+08
0
2000
4000
6000
8000
10000
12000
500
1000
1500
0.0
0.2
0.4
0.6
0.8
1.0

◂**Fig. 2** K, L and G Functions for the Northern Mountain ranges (**a**-**f**). The envelope in grey represents the confidence interval of the test, which is defined by the lowest (Kl0(r)) and highest (Khi(r)) values. The flashing red line shows the theoretical mean of the results of the 99 random simulations (Ktheo(r)). The black line (Kobs(r)) marks the mean of the results for the sample of archaeological sites

(p-value = 0.00005112). The Mann–Whitney-Wilcoxon test establishes that both datasets possess similar variances (p-value = 0.3567) and the Student's t-test affirms that they originate from the same population, without statistically significant differences (p-value = 0.3596).

The variable ASPm showcases archaeological sites with a minimum aspect value of 11.34, whereas random points exhibit a minimum value of 13.01. The median aspect value for archaeological sites is 229.45, while for random points, it is 159.44. The maximum aspect value recorded for archaeological sites is 330.94 and for random points, it is 351.37. The distribution of values for archaeological sites primarily ranges from 150 to 250, while random points cluster between 100 and 240, as depicted in Fig. 7f. The Shapiro–Wilk test confirms normality for the datasets (p-value = 0.1197). Regarding homoscedasticity, the Fisher's F-test indicates similar variances (p-value = 0.5768). Subsequently, the Student's t-test affirms that both datasets belong to the same population (p-value = 0.02261).

For the HYDROEm variable, the minimum value for archaeological sites is 34.42, whereas random points display a minimum value of 13.94. The median hydrological exposure value for archaeological sites is 687.89 and for random points, it is 627.03. The maximum value recorded for archaeological sites is 1783.49 and for random points, it is 2100.81. Both datasets illustrate a similar trend, oscillating between values of 250 and 1200, as visualized in Fig. 7g. Statistical analyses confirm that neither dataset conforms to a normal distribution (p-value = 0.003188). The Mann–Whitney-Wilcoxon test suggests that both datasets possess similar variances (p-value = 0.9079) and the Student's t-test establishes that they emanate from the same population, with no statistically significant distinctions (p-value = 0.9887).

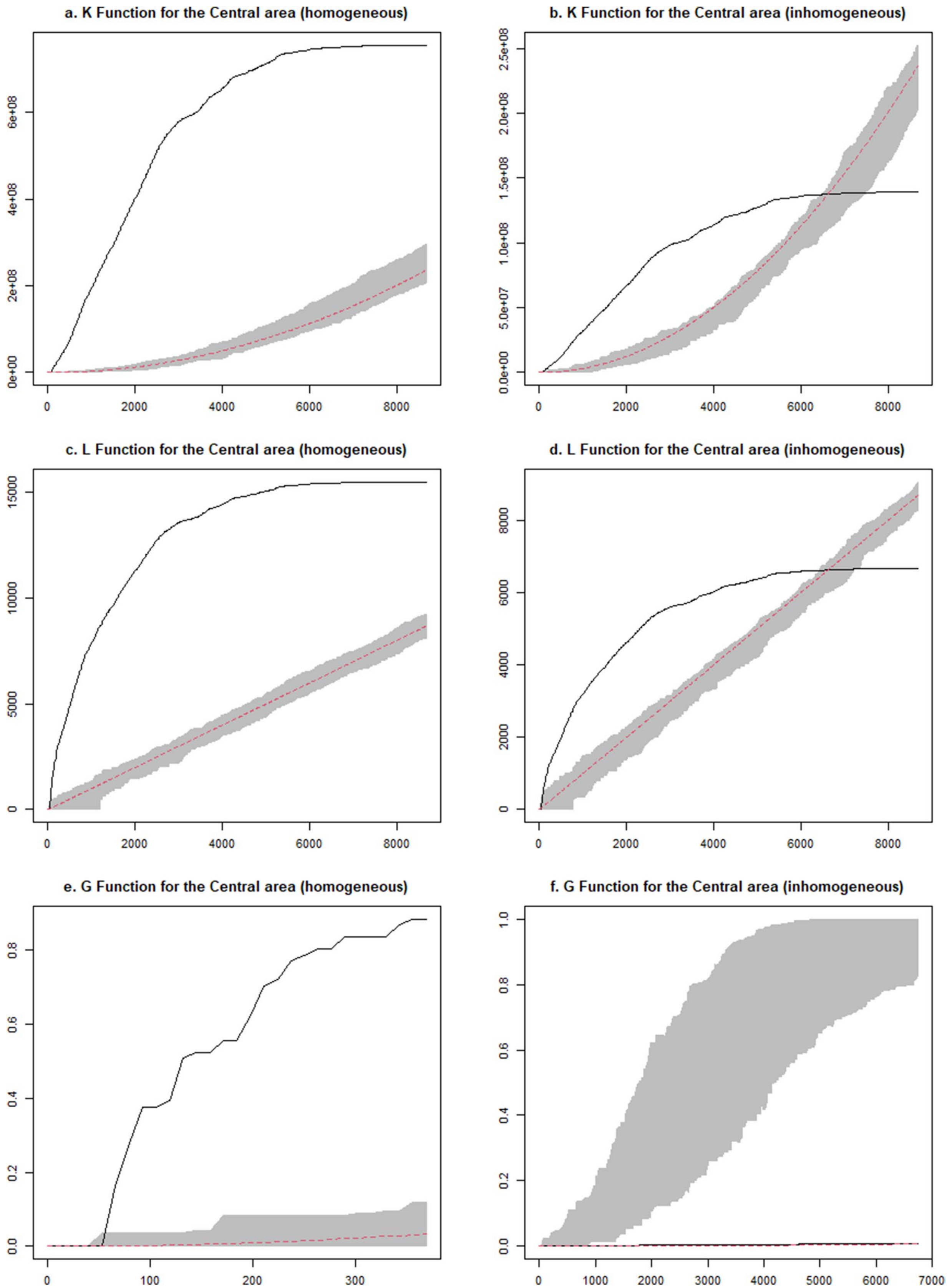
Another variable analysed is HYDROCm, for which the minimum values for both samples are 3.36 for archaeological sites and 3.27 for random points. As for the median, archaeological sites have 42.86, while random points have 44.21. Finally, the maximum value for archaeological sites is 99.44 and for random points, it is 118. Similar to the previous variable, the trend in both datasets is comparable, with values ranging from 15 to 70, as represented in Fig. 7h. The Shapiro–Wilk test indicated that there is no normality among the data (p-value = 0.002521). The Mann–Whitney-Wilcoxon test showed that both samples have the same variance (p-value = 0.8407) and the Student's t-test confirmed that both belong to the same population (p-value = 0.8503).

The WETm variable showed a minimum value of 0.11 for archaeological sites and 0 for random points. The median for archaeological sites is 13.61 and for random points, it is 18.73. Finally, the maximum value for archaeological sites is 44.04 and for random points, it is 85.16. Both datasets showcase a similar dispersion of values, concentrating between 5 and 20, as visualized in Fig. 7i. Regarding the statistical analyses conducted, it was confirmed that there is no normality (p-value = 0.000000006833), that the data has the same variance (p-value = 0.538) and that there are no significant differences according to the Student's t-test (p-value = 0.1682).

Another variable studied is GEOLEm. In this case, the minimum value for the archaeological sample and the random sample is 0. The median for archaeological sites is 802.42 and for random points, it is 1526.5. Finally, the maximum value for archaeological sites is 2975.79 and for random points, it is 4861.4. The trend for archaeological points oscillates between 0 and 1500, while random points exhibit a wider concentration, ranging from 500 to 2300, as shown in Fig. 8a. The Shapiro–Wilk statistical test confirmed that there is no normality (p-value = 0.000006409) and the Student's t-test indicated that the variables from both samples are different (p-value = 0.01154). Finally, the T Welch test confirmed that they belong to different populations and therefore exhibit statistically significant differences (p-value = 0).

In the case of the GEOLCm variable, the minimum values of both samples are found at 0. However, the median for archaeological sites is at 25.03, whereas for random points, it is at 61.01. The maximum value for archaeological sites is 152.54 and for random points, it is 240.52. The data trends for both samples are similar to what was observed in the previous variable, but in this case, the values for archaeological sites range between 0 and 80, while for random points, they are concentrated between 25 and 150, as depicted in Fig. 8b. Regarding the statistical analyses of both samples, it was verified that there is no normality (p-value = 0.000007255), both samples have different variances (p-value = 0.01469) and they exhibit significant differences (p-value = 0.01332).

Another variable studied is VISPRm, for which a minimum value of 0 was identified for archaeological sites and 3 for random points. The median for the former is located at 20.09 and for the latter, it is at 19.65. As for the maximum values, it is 75 for archaeological sites and 66 for random points. The value distribution shows that for archaeological points, it concentrates between 8 and 30, whereas for random points, it ranges from 9 to 22, as seen in Fig. 8c. The Shapiro–Wilk statistical test indicated that there is no normality (p-value = 0.000001251). The variances of the samples are similar, as demonstrated by the Mann–Whitney-Wilcoxon test (p-value = 0.8443) and it was confirmed that


a. K Function for the Central area (homogeneous)
b. K Function for the Central area (inhomogeneous)
c. L Function for the Central area (homogeneous)
d. L Function for the Central area (inhomogeneous)
e. G Function for the Central area (homogeneous)
f. G Function for the Central area (inhomogeneous)
0e+00
2e+08
4e+08
6e+08
0.0e+00
5.0e+07
1.0e+08
1.5e+08
2.0e+08
2.5e+08
0
2000
4000
6000
8000
5000
10000
15000
0.0
0.2
0.4
0.6
0.8
1.0
100
200
300
1000
3000
5000
7000

◂**Fig. 3** K, L and G Functions for the Central Mountain ranges (**a**-**f**). The envelope in grey represents the confidence interval of the test, which is defined by the lowest (Kl0(r)) and highest (Khi(r)) values. The flashing red line shows the theoretical mean of the results of the 99 random simulations (Ktheo(r)). The black line (Kobs(r)) marks the mean of the results for the sample of archaeological sites

there are no significant differences between both samples (p-value = 0.9161).

The LCPCm variable yielded a minimum value of 0.26 for archaeological sites and 0.50 for random points. The median for archaeological sites is 20.63, while for random points, it is 2.42. As for the maximum value for each sample, it is 99.41 for archaeological sites and 23.68 for random points. The trend in the dispersion of values for both samples is practically the same, as values are grouped between 2 and 4, as displayed in Fig. 8d. The Shapiro–Wilk test indicated that there is no normality (p-value = 0.0000002778) and the variances of both samples are equal (p-value = 0.7839). Finally, the T Student test confirmed that there are no significant differences between both samples (p-value = 0.2453).

Another variable that has been analysed is TOTINSm, for which a minimum value of 3.51 was identified for archaeological sites and 1.86 for random points. The median for archaeological sites is 4.57 and for random points, it is 4.42. As for the maximum value for the former, it is located at 5.26 and for the latter, it is at 5.40. The trend in the concentration of values for each sample is similar in both cases, oscillating between values of 4 and 5, as presented in Fig. 8e. The statistical analysis by Shapiro–Wilk indicated that there is no normality (p-value = 0.0000001963) and both samples have the same variance according to the T Student test (p-value = 0.2146). Finally, it was confirmed that both samples belong to the same population (p-value = 0.1367).

Regarding the DIRINSm variable, the minimum value of the archaeological sample is 0.84 and for the random sample, it is 1.01. Their respective medians are 3.70 for the first group and 3.53 for the second group. The maximum value for archaeological sites is 4.42 and for random sites, it is 4.53. The concentration of values for both samples is similar, as they are found between values of 3 and 4, as depicted in Fig. 8f. Both samples do not exhibit normality (p-value = 0.0000003559) and their variances are similar (p-value = 0.2797). The T Student test confirmed that there are no significant differences between both samples (p-value = 0.1618).

Another variable related to insolation is DIFINSm. In this case, archaeological sites present a minimum value of 0.84 and random points have a minimum value of 0.88. The median for archaeological sites is 0.91 and for random points, it is 0.85. The maximum value for both samples is located at 0.93. The concentration trends of values are different in both samples, as the values for archaeological points range between 0.89 and 0.92, while for random points, they are concentrated between 0.87 and 0.89, as seen in Fig. 8g. Both samples do not exhibit normality (p-value = 0.04847) and their variances are different (p-value = 0.005377). The T Welch test confirmed that they belong to different populations and therefore exhibit statistically significant differences (p-value = 0.003053).

For the WINDm variable, archaeological sites have a minimum value of 0.78 and random points have a minimum value of 0.77. The median for both samples is 0.97. The maximum value for archaeological sites is located at 1.23, while for random points, it is at 1.31. Their value concentration trends are similar, although for archaeological sites, it ranges between values of 0.9 and 1.1. For random points, it concentrates between 0.85 and 1.1, as displayed in Fig. 8h. The Shapiro–Wilk test confirmed that there is no normality (p-value = 0.006604) and the Mann–Whitney-Wilcoxon test showed that their variances are similar (p-value = 0.9854). Finally, the T Student test confirmed that both samples belong to the same population (p-value = 0.8687).

Central area

In the Central Mountain ranges area, the first variable under scrutiny was the locational variable ALTA. Visual examination did not reveal any significant differences in the distribution of archaeological points (refer to Fig. 9a); however, statistical analysis unveiled disparities between this variable and the random background points (see Table 14). Similar observations applied to the variable TPI1000, the sole one associated with Topographic Prominence analysis that displayed noteworthy deviations from the expected conditions under random sampling (p-value = 0.01, Table 14), although it did not manifest visually (see Fig. 9d). Conversely, the variables TPI100 and TPI500 showed no substantial evidence (refer to Fig. 9b-c and Table 14).

Variables pertaining to hydrology, such as HYDROE and HYDROC, exhibited marked distinctions from the expected patterns under random conditions. These distinctions were perceptible through visual analysis (see Fig. 10a-b) and statistically significant (p-value < 0.05 in Table 14). Another variable demonstrating this phenomenon was WIND, where the distribution of archaeological sites deviated from the distribution generated in 999 random simulations (see Fig. 10d). The Mann–Whitney U test confirmed statistically significant differences (refer to Table 14). In the case of the LCPC variable, visual observations suggested deviations in the distribution of archaeological points from what was expected under random conditions, with a notable concentration near potential transit routes (see Fig. 10c). However, statistical confirmation through the Mann–Whitney U test remained inconclusive (p-value = 0.54, Table 14).

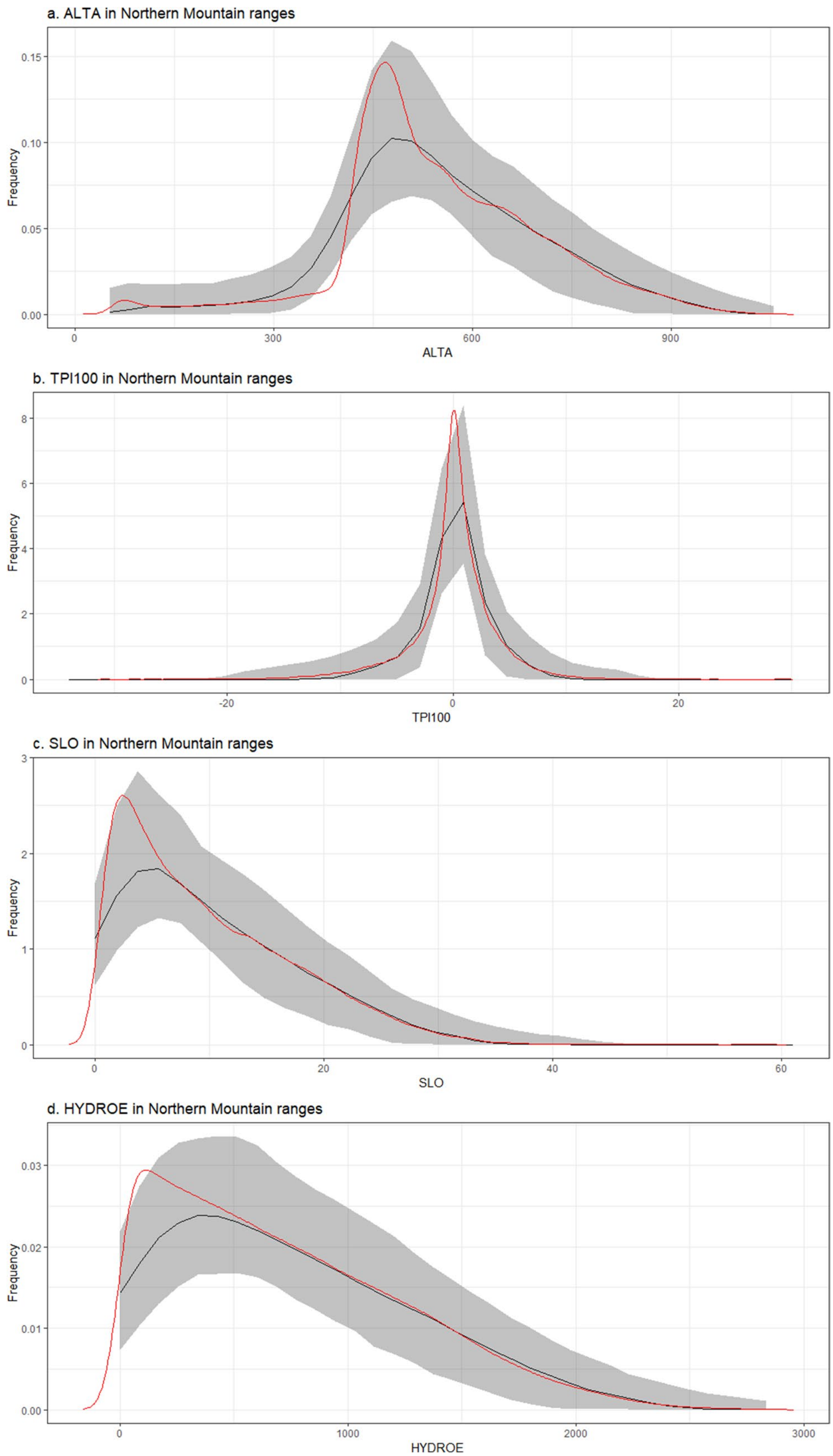


Fig. 4 Kernel density plots comparing archaeological sites (red line) and randomly resampled background points (black line) for different variables. The grey areas represent the 95% confidence intervals calculated from the resampling of 999 random samples

Fig. 5 Kernel density plots comparing archaeological sites (red line) and randomly resampled background points (black line) for different variables. The grey areas represent the 95% confidence intervals calculated from the resampling of 999 random samples

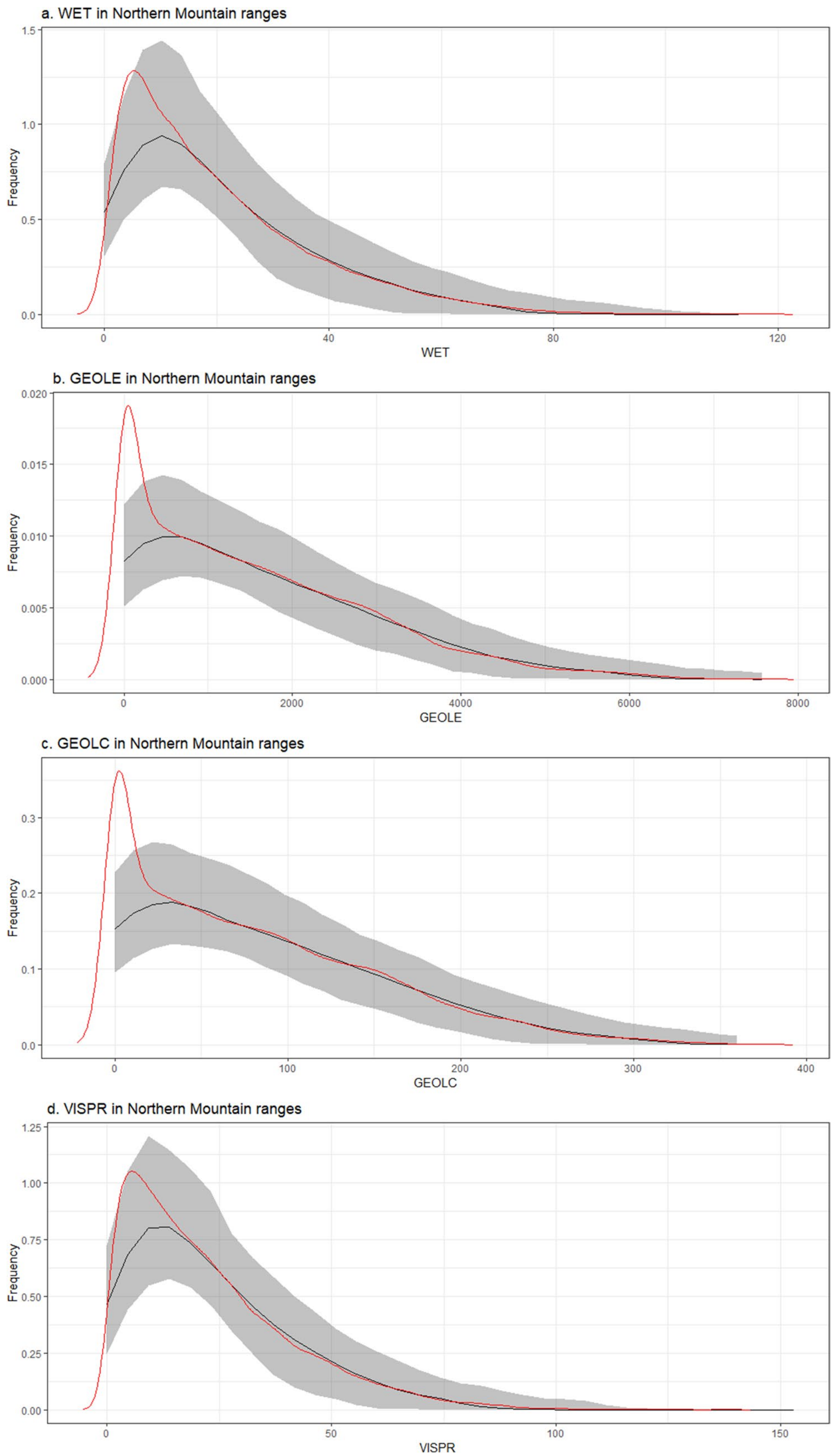


Table 13 Mann–Whitney U test results for the comparison between archaeological sites and randomly resampled background points in Northern area

Variable	Statistic	Lower CI	Median	Upper CI
ALTA	U statistic	728.475	850	958.025
	p-value	0.000	0	0.033
TPI100	U statistic	579.450	696.000	797.050
	p-value	0.004	0.075	0.495
TPI500	U statistic	502.950	640.000	759.100
	p-value	0.013	0.225	0.823
TPI1000	U statistic	465.950	581.000	695.000
	p-value	0.077	0.488	0.916
SLO	U statistic	492.950	602.500	719.500
	p-value	0.042	0.384	0.853
ASP	U statistic	590.375	715.500	840.000
	p-value	0.001	0.046	0.442
HYDROE	U statistic	478.45	596.500	705.05
	p-value	0.06	0.413	0.89
HYDROC	U statistic	448.950	571.000	685.100
	p-value	0.096	0.537	0.944
WET	U statistic	334.975	461.500	590.500
	p-value	0.441	0.924	0.999
GEOLE	U statistic	282.425	393.000	517.025
	p-value	0.776	0.989	1.000
GEOLC	U statistic	283.850	398.000	516.05
	p-value	0.779	0.987	1.00
VISPR	U statistic	415.950	516.500	630.575
	p-value	0.261	0.777	0.977
LCPC	U statistic	455.425	562.500	673.525
	p-value	0.122	0.578	0.934
TOTINS	U statistic	565.950	682.000	800.050
	p-value	0.003	0.102	0.561
DIRINS	U statistic	558.900	677.000	790.050
	p-value	0.005	0.113	0.595
DIFINS	U statistic	678.5	775.500	860.525
	p-value	0.0	0.008	0.110
WIND	U statistic	480.875	610.00	732.100
	p-value	0.030	0.35	0.884
CPFPCGc	U statistic	382.850	517.000	660.050
	p-value	0.159	0.775	0.992
CPFPCDc	U statistic	1156	1156	1156
	p-value	0	0	0

The distribution of archaeological points for the DIFINS variable displayed a distinct pattern compared to the median of the background derived from 999 random samples, though it consistently remained within the envelope generated by them (see Fig. 11a). Nevertheless, statistically significant differences emerged upon applying the Mann–Whitney U test ($p\text{-value} < 0.05$, Table 14). In the case of the CPFPCDc and CPFPCGc variables, the former exhibited non-normal values, as all random points possessed values of 0 and could not be deemed representative (see Fig. 11b). For the latter variable, no significant differences were observed either visually or statistically (see Fig. 11c and Table 14).

With regard to the variables SLO, ASP and WET, no significant differences were detected visually (refer to Fig. 13a-c in the SM) and statistical analysis corroborated these findings (see Table 14). Conversely, variables linked to geological proximity displayed a different pattern. For GEOLE, statistical significance was confirmed after conducting the Mann–Whitney U test ($p\text{-value} < 0.05$, Table 14). However, this distinction remained indiscernible through visual analysis (see Fig. 13d in the SM). The same applied to the variable GEOLC, which exhibited a p-value of < 0.05 following the Mann–Whitney U test (refer to Fig. 14a in the SM and Table 14 in the text). In contrast, the variables VISPR, TOTINS and DIRINS did not exhibit statistically significant differences and these distinctions were not visually evident through the probability density estimation graph for archaeological sites (see Fig. 14b-d in the SM and Table 14 in the text).

The analysis proceeds to present results for variables lacking raster data but obtained from cell values proximal to archaeological sites, offering insights into the immediate environment. This analysis involves a comparison of archaeological point values with an equal number of randomly generated points. The results of the statistical analyses for the following variables studied in this area can be checked in Table 3 and Table 5 in the SM file.

The first one is ALTm, for which a minimum value of 470 is observed, unlike the random points that show a minimum value of 307 m. As for the medians of both, the archaeological sites have a value of 725 and the random ones have 492 m. Finally, the maximum value of the former is 758 and the latter is 734 m. Observing the distribution of both samples, the archaeological sites have a concentration around 700 m, while the random points are grouped between 450 and 550 m (refer to Fig. 12a). The statistical analysis using the Shapiro–Wilk test has shown that both samples do not follow a normal distribution ($p\text{-value} = 0.000000007778$), which led to the use of a non-parametric test, the Mann–Whitney-Wilcoxon test, to check for homoscedasticity. This test confirmed that both samples have different variances ($p\text{-value} = < 2.2\text{e-}16$). Finally, the T Welch test confirmed that both samples have statistically significant differences ($p\text{-value} = < 2.2\text{e-}16$).

The next variable analysed is TPI100m, for which a minimum value of -5.63 is observed for archaeological sites and -5.40 for random points. The median for archaeological points is 0.22 and for random sites, it is 5.09. As for the maximum values, archaeological sample have a value of 10.11 and random sites have 4.10. Both samples are grouped around values of -1.5 and 1.5 (see Fig. 12b), so no significant

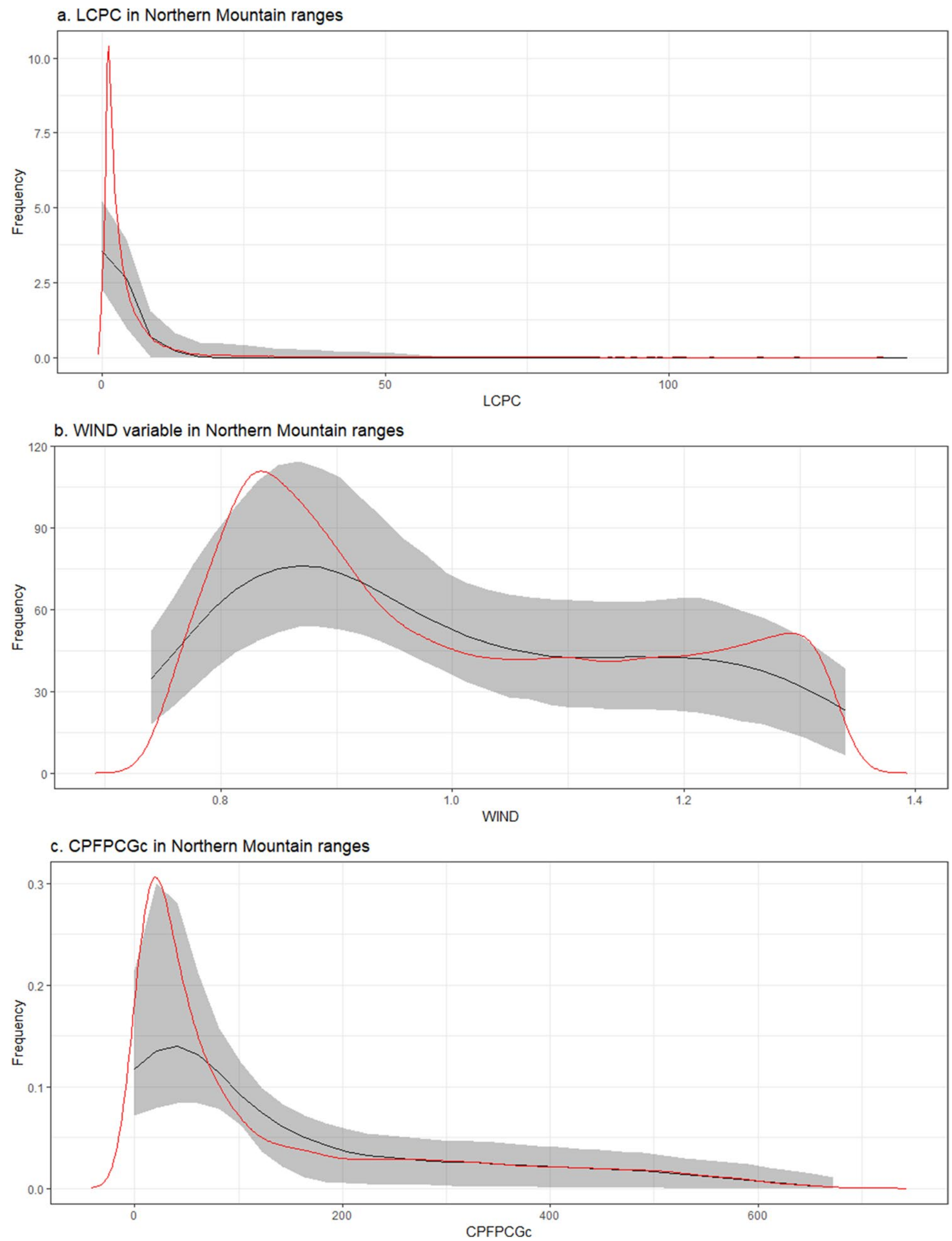


Fig. 6 Kernel density plots comparing archaeological sites (red line) and randomly resampled background points (black line) for different variables. The grey areas represent the 95% confidence intervals calculated from the resampling of 999 random samples

differences can be seen visually. The statistical analysis using the Shapiro–Wilk test shows that there is normality (p-value = 0.1544). This result led to the use of the F Test to test homoscedasticity, which confirmed the presence of the same variance in both samples (p-value = 0.3478). Since they have the same variance, the T Student Test was used and it was found that there are no significant differences between both samples (p-value = 0.7598).

Another topographic prominence-related variable is TPI500m. Archaeological sites display a minimum value of -6.08, whereas random points plummet to -19. Regarding the medians, the value for archaeological points is -0.19 and for random ones, it is -2.14. As for the maximum value, for archaeological sites, it is located at 5.52 and for random sites, it is at 18.94. The grouping of values in both samples is similar (refer to Fig. 12c). The statistical test indicated that there is no normality (p-value = 0.2957). The F Test allowed verifying that the variances of both samples are different (p-value = 0.02653). The T Welch test confirmed that both samples do not have statistically significant differences (p-value = 0.945).

The last variable related to topographic prominence is TPI1000m. The minimum value for archaeological sites is -6.08 and for random points, it is -31.38. The median for first ones is -0.19 and for random ones, it is 2.79. The maximum value for the archaeological sample is 5.52

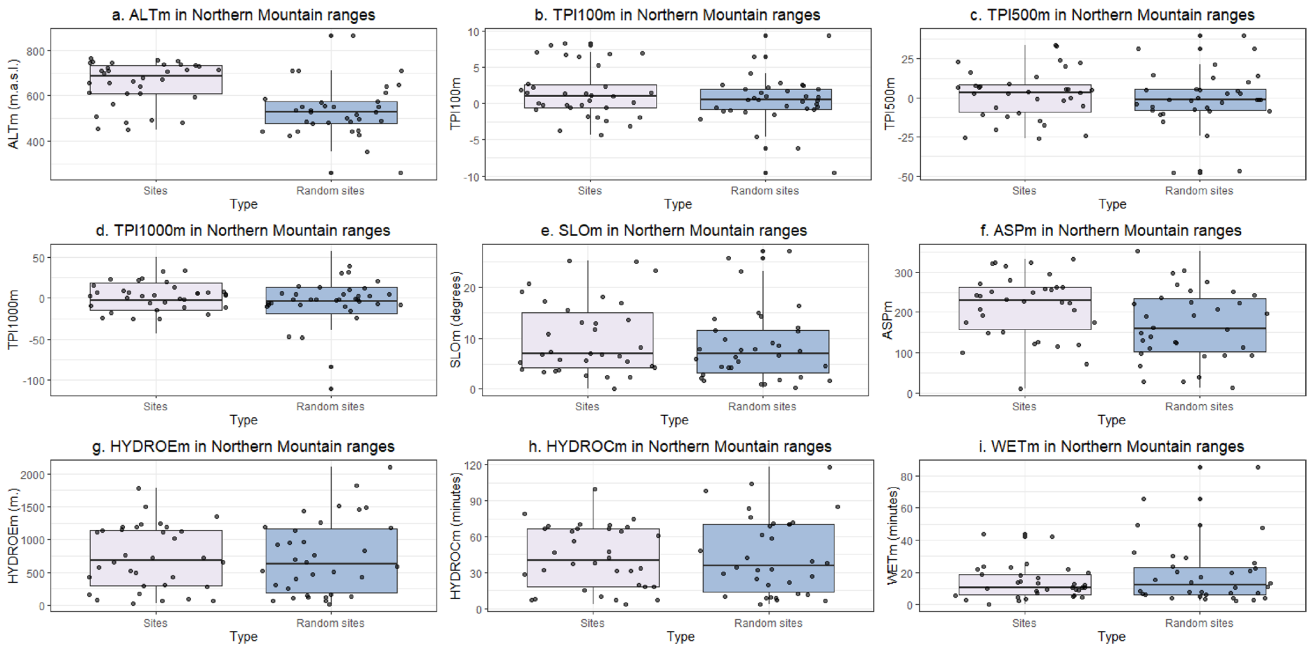


Fig. 7 Boxplots comparing the results of archaeological sites with random sites for different variables in the Northern area

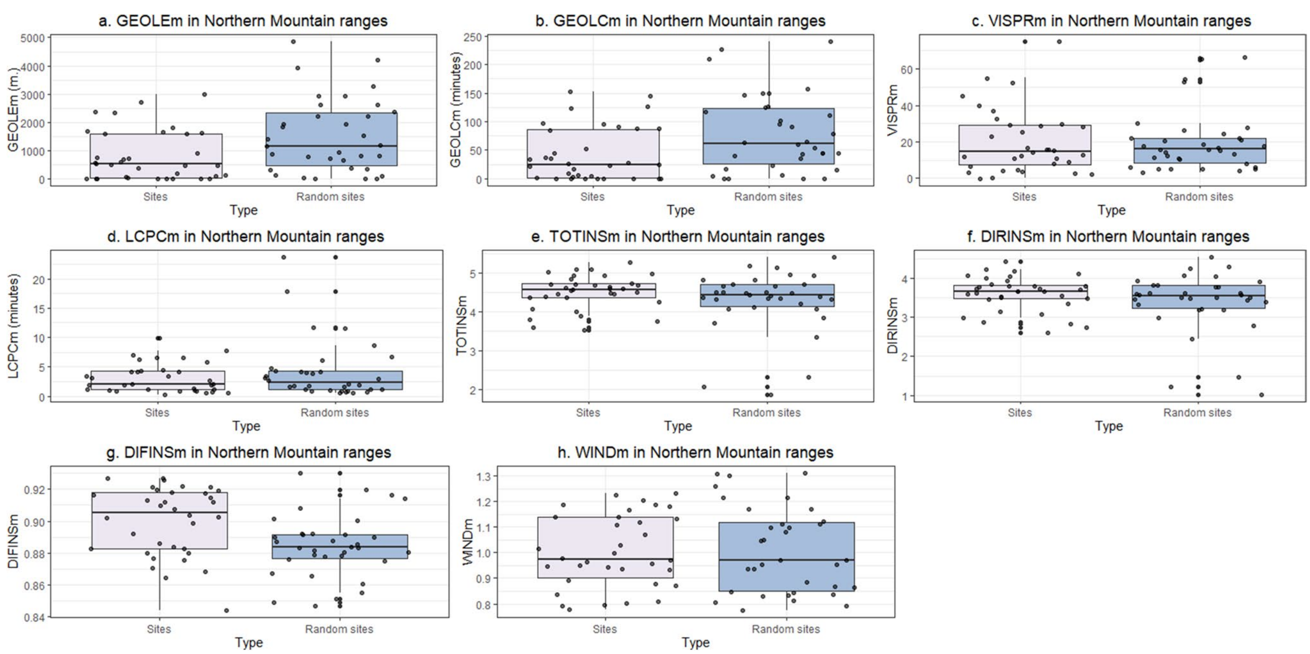


Fig. 8 Boxplots comparing the results of archaeological sites with random sites for different variables in the Northern area

and for random sample, it is 28.94. Their distributions, as depicted in Fig. 12d, appear strikingly alike. The Shapiro–Wilk test demonstrated that there is no normality (p-value = 0.0131). The Mann–Whitney-Wilcoxon test showed that both samples do not have the same variance (p-value = 0.03906). The T Welch test demonstrated that there are significant differences in both samples and therefore, they would not statistically belong to the same population (p-value = 0.03107).

The minimum value of the SLOm variable for archaeological sites is 0.85 and for random points, it is 0.48. Regarding the medians, for archaeological points, it is 5.14 and for random ones, it is 5.90. Finally, the maximum value for the archaeological sample is 15.19 and for random sites, it

Table 14 Mann–Whitney U test results for the comparison between archaeological sites and randomly resampled background points in Central area

Variable	Statistic	Lower CI	Median	Upper CI
ALTA	U statistic	3313.5	3467.5	3568.025
	p-value	0.0	0.0	0.000
TPI100	U statistic	1490.500	1718.500	1978.700
	p-value	0.273	0.767	0.971
TPI500	U statistic	1780.70	2046.000	2319.00
	p-value	0.01	0.172	0.66
TPI1000	U statistic	2016	2313.00	2573.100
	p-value	0	0.01	0.214
SLO	U statistic	1505.050	1793.000	2109.100
	p-value	0.102	0.636	0.966
ASP	U statistic	1463.400	1760.000	2042.525
	p-value	0.176	0.697	0.979
HYDROE	U statistic	2691.325	2977	3219.05
	p-value	0.000	0	0.00
HYDROC	U statistic	2689	2975	3212.1
	p-value	0	0	0.0
WET	U statistic	1306.950	1596.000	1913.550
	p-value	0.394	0.913	0.998
GEOLE	U statistic	2309.7	2714	3069.050
	p-value	0.0	0	0.011
GEOLC	U statistic	2288.75	2696	3089.400
	p-value	0.00	0	0.014
VISPR	U statistic	849.375	1078	1341.825
	p-value	0.996	1	1.000
LCPC	U statistic	1557.975	1841.50	2161.00
	p-value	0.062	0.54	0.94
TOTINS	U statistic	1882.875	2169.500	2429.200
	p-value	0.002	0.057	0.455
DIRINS	U statistic	1821.975	2100.50	2343.200
	p-value	0.007	0.11	0.579
DIFINS	U statistic	3318.9	3466	3565.025
	p-value	0.0	0	0.000
WIND	U statistic	2360.075	2635	2906.100
	p-value	0.000	0	0.005
CPFPCGc	U statistic	942.950	1345.000	1739.05
	p-value	0.734	0.996	1.00
CPFPCDc	U statistic	1830.000	1860.5	1860.5
	p-value	0.845	1.0	1.0

is 18.43. The distribution of both samples is similar (see Fig. 12e). The Shapiro–Wilk test shows that there is no normality (p-value = 0.000009964). Both samples have the same variance, as shown by the Mann–Whitney-Wilcoxon test (p-value = 0.3702). Finally, it has been verified that both samples are not statistically significant (p-value 0.2006).

Regarding the variable ASPm, the minimum value for archaeological sites is 11.59 and for random points 22.82. The median is 165.57 for archaeological ones and 202.74 for random sites. While the maximum value for the archaeological sample is 346.40, it is 335.70 for random points. Their distribution patterns, as depicted in the boxplot (Fig. 12f), suggest comparable distributions. The statistical analysis of both samples indicates that there is no normality (p-value = 0.04993). Archaeological sites and random points have the same variance, as demonstrated by the Mann–Whitney-Wilcoxon test (p-value = 0.5155). The T Student test showed that both samples do not have significant differences (p-value = 0.6325).

For the HYDROEm variable, the minimum value for archaeological sites is 33.25 and for random points is 14.98 m. The median for the archaeological sample is 1247.73 and for random ones, it is 589.14 m. The maximum value for the former sample is 2018.44 and for the latter, it is 1959.42 m. The distribution of both samples is visually different, as archaeological points are concentrated around 1000 and 1500 m. Visual examination reveals contrasting distributions: archaeological points tend to cluster around 1000 and 1500 m, while random points gravitate toward 200 and 800 m (refer to Fig. 12g). The statistical analysis has shown that there is no normality (p-value = 0.003927). The Mann–Whitney-Wilcoxon test indicates that both samples have different variances (p-value = 0.00000003019). The T Welch test shows that there are significant differences between the sample of archaeological sites and random points (0.000000002273).

In the case of HYDROCm, the minimum value for archaeological sites is 5.27 and the minimum value for random points is 1.95 min. The median for the former sample is 68.53 and for the latter, it is 33.85 min. The maximum value for archaeological points is 108.14 and for random points, it is 115.53 min. Their distributions are strikingly different, with archaeological sites concentrating between 60 and 80 min and random points ranging from 15 to 45 min (Fig. 12h). Statistically, both samples do not follow a normal distribution (p-value = 0.01832). The Mann–Whitney-Wilcoxon test indicates that both samples have different variances (p-value = 0.00000002122). Finally, it is observed that both samples come from different populations and exhibit statistically significant differences (p-value = 0.000000003962).

For the WETm variable, based on the cost of travel to potential wetland areas, a minimum value of 1.03 min has been identified for archaeological sites and 0.98 min for random points. The median for the archaeological sample is 12.25 min and for the random sample, it is 12.14 min. As for the maximum values, archaeological sites have a value of 45.93 min, while random sites have a value of 82.75 min. The distribution of values for each of the samples is concentrated between 10 and 20 min (Fig. 12i). The Shapiro–Wilk test indicates that there is no normality

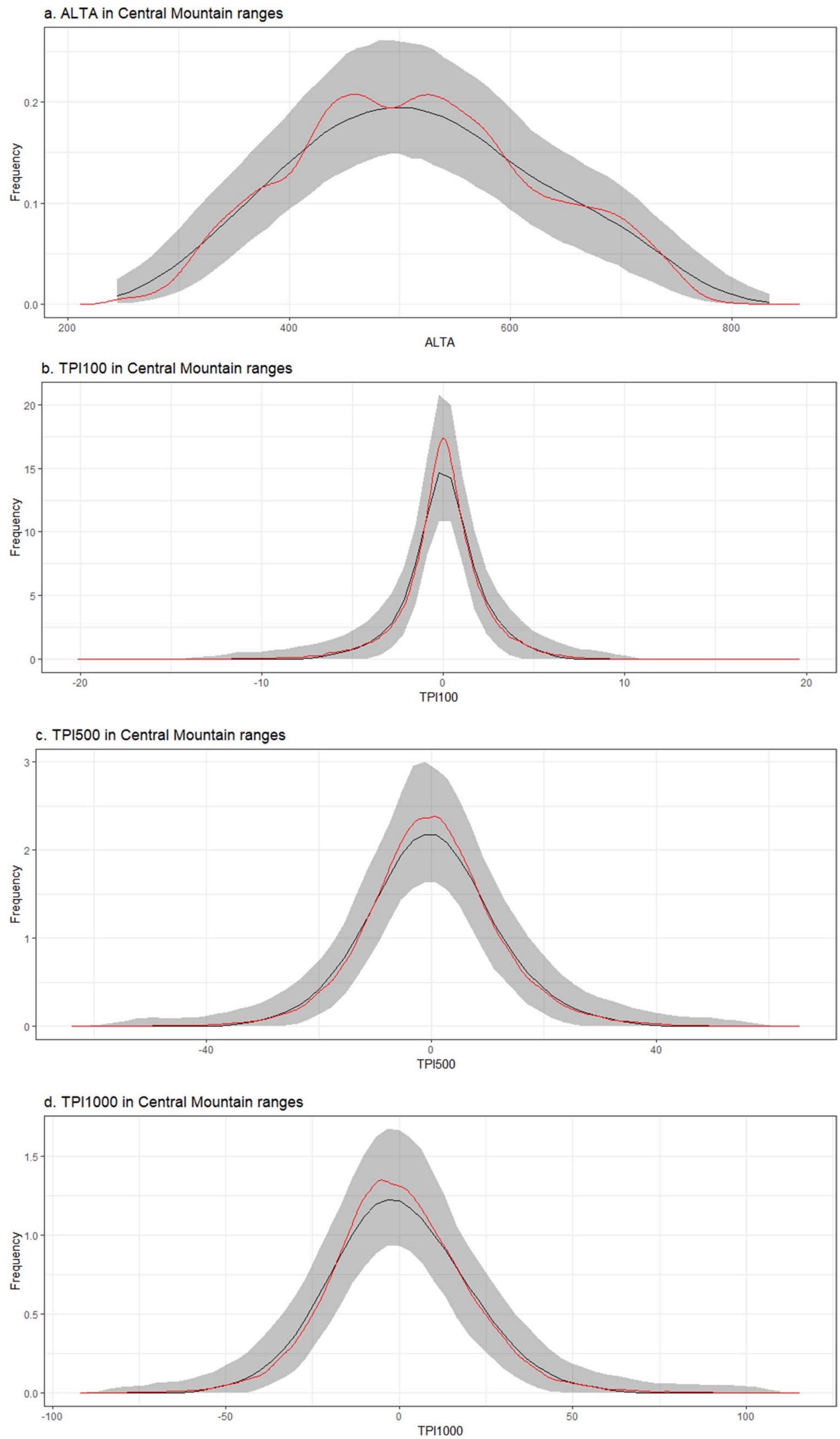
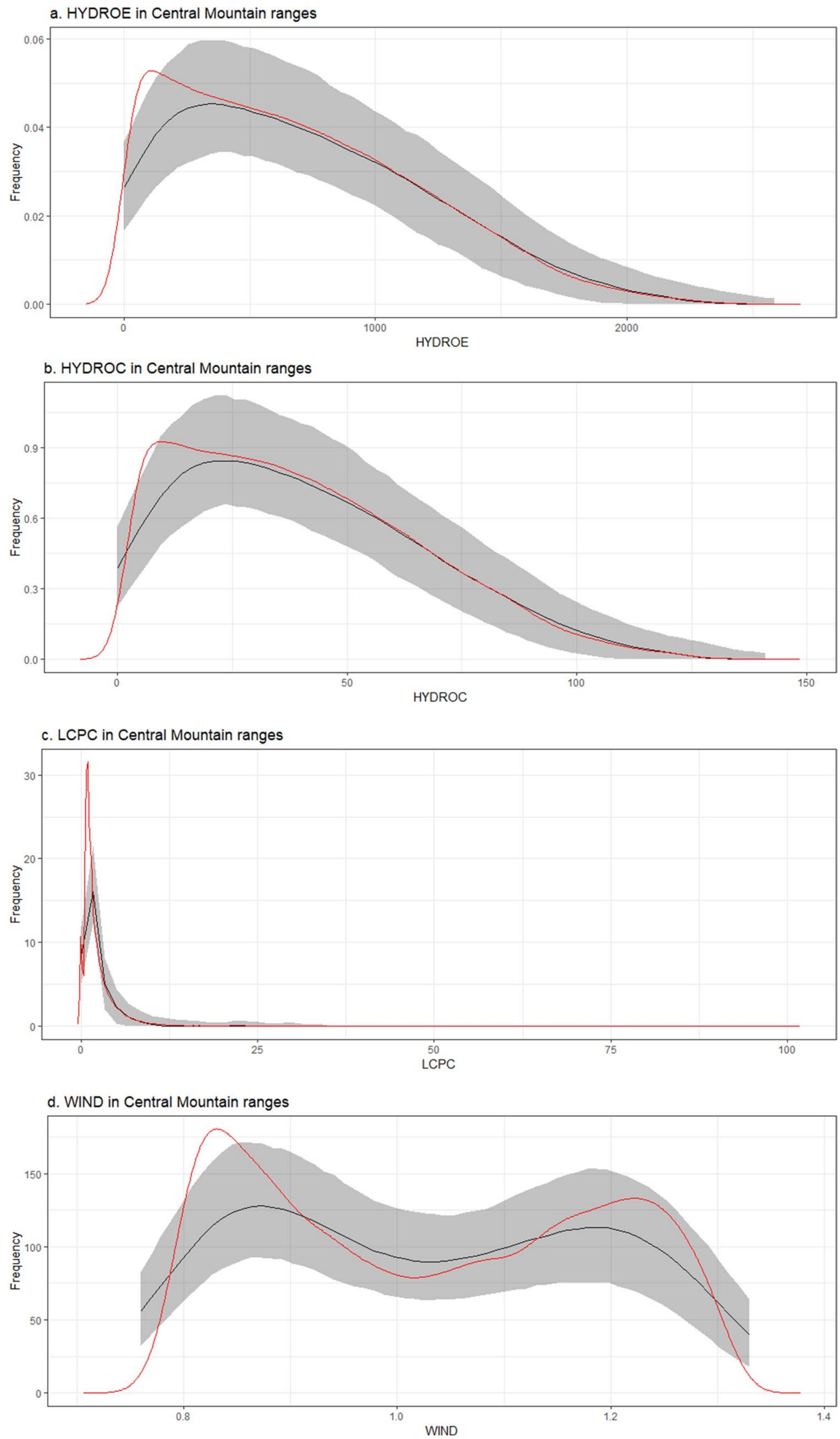


Fig. 9 Kernel density plots comparing archaeological sites (red line) and randomly resampled background points (black line) for different variables. The grey areas represent the 95% confidence intervals calculated from the resampling of 999 random samples

Fig. 10 Kernel density plots comparing archaeological sites (red line) and randomly resampled background points (black line) for different variables. The grey areas represent the 95% confidence intervals calculated from the resampling of 999 random samples



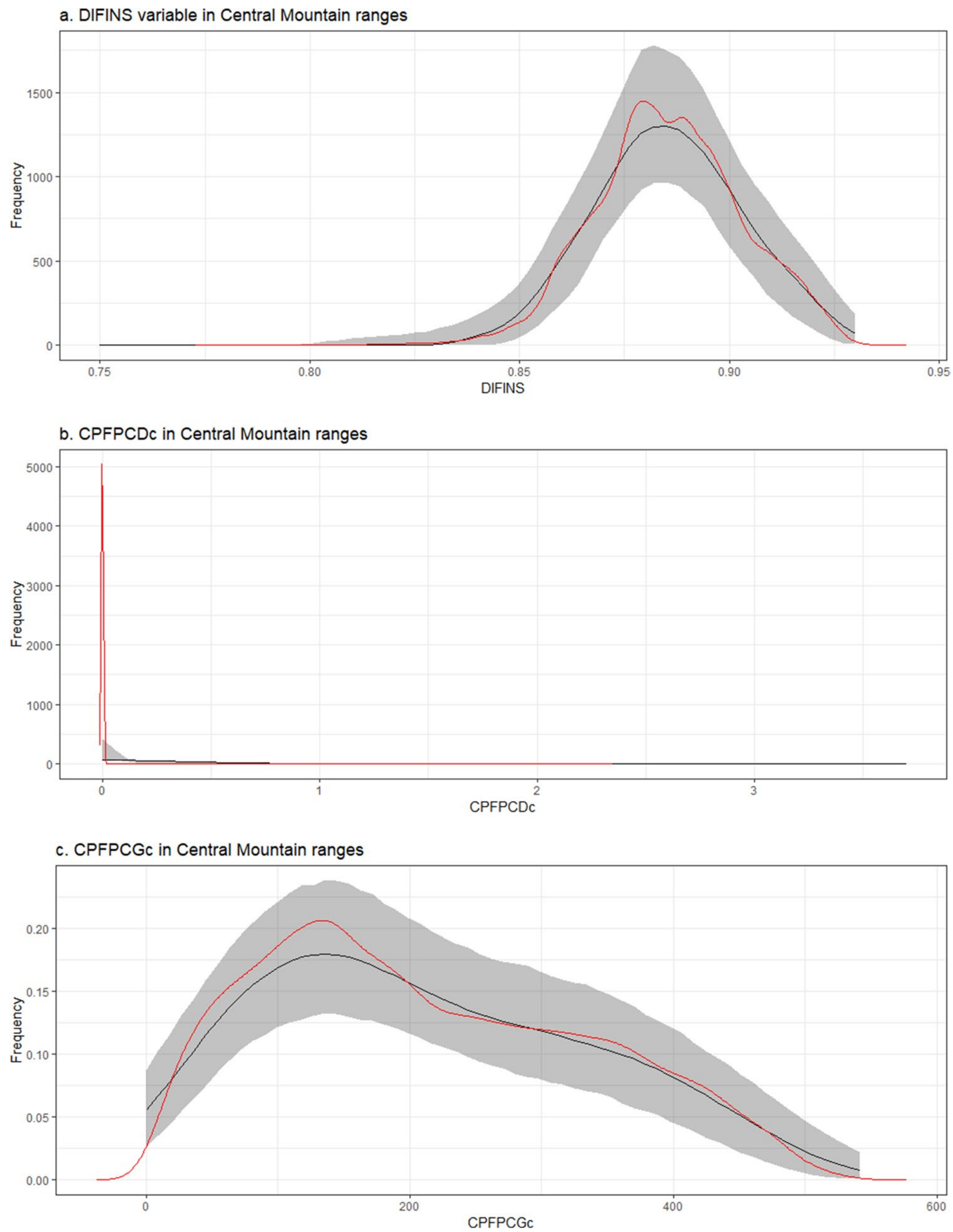


Fig. 11 Kernel density plots comparing archaeological sites (red line) and randomly resampled background points (black line) for different variables. The grey areas represent the 95% confidence intervals calculated from the resampling of 999 random samples

(p-value = 2.332e-10). The Mann–Whitney-Wilcoxon test shows that both samples have the same variance (p-value = 0.8417). Ultimately, the T Student test demonstrates that there are no statistically significant differences between these two samples (p-value = 0.4089).

The distance in meters to potential geology has been calculated using the variable GEOLEm. The minimum values for the archaeological sample and the random sample are 5367 and 0 m, respectively. The median value for archaeological points is located at 8575 m and for random points, it is 5204 m. The maximum value for archaeological sites is 9522, while for random points, it is 16,127 m. Visual representation of the data (see Fig. 13a) indicates that the values in the archaeological sample cluster between 7500 and 8000 m, whereas random values are more widely dispersed. It has been confirmed that there is no normality after applying the Shapiro–Wilk test (p-value = 0.003927). Variances have been checked to be different (p-value = 0.00000003019). Finally, the T Welch test has confirmed that there are statistically significant differences between both samples (p-value = 0.000000002273).

The other variable related to potential geology but in terms of travel cost in time is GEOLCm. The minimum value for archaeological sites for this variable is 259.3 min and 0 for random points. The median for the former sample is 431.6 min and for the latter, it is 253.9 min. The maximum value for archaeological sites is 474.5 and for random points, it is 801.7 min. As indicated in Fig. 13b,

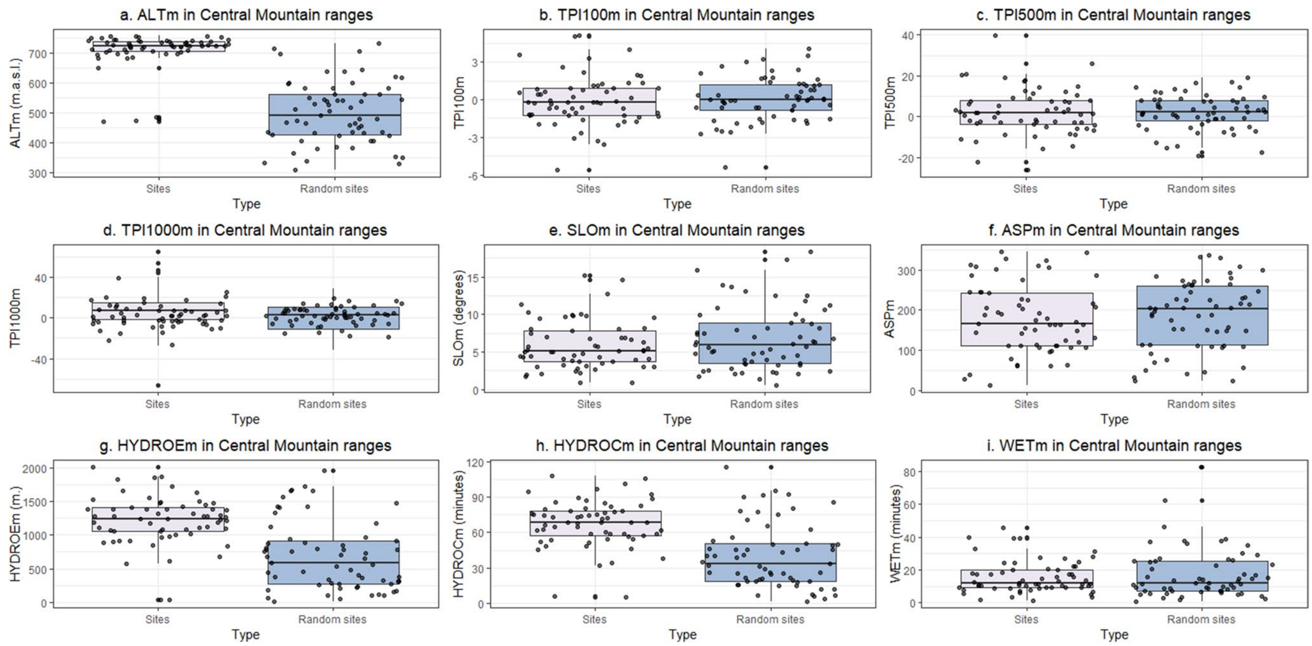


Fig. 12 Boxplots comparing the results of archaeological sites with random sites for different variables in the Central area

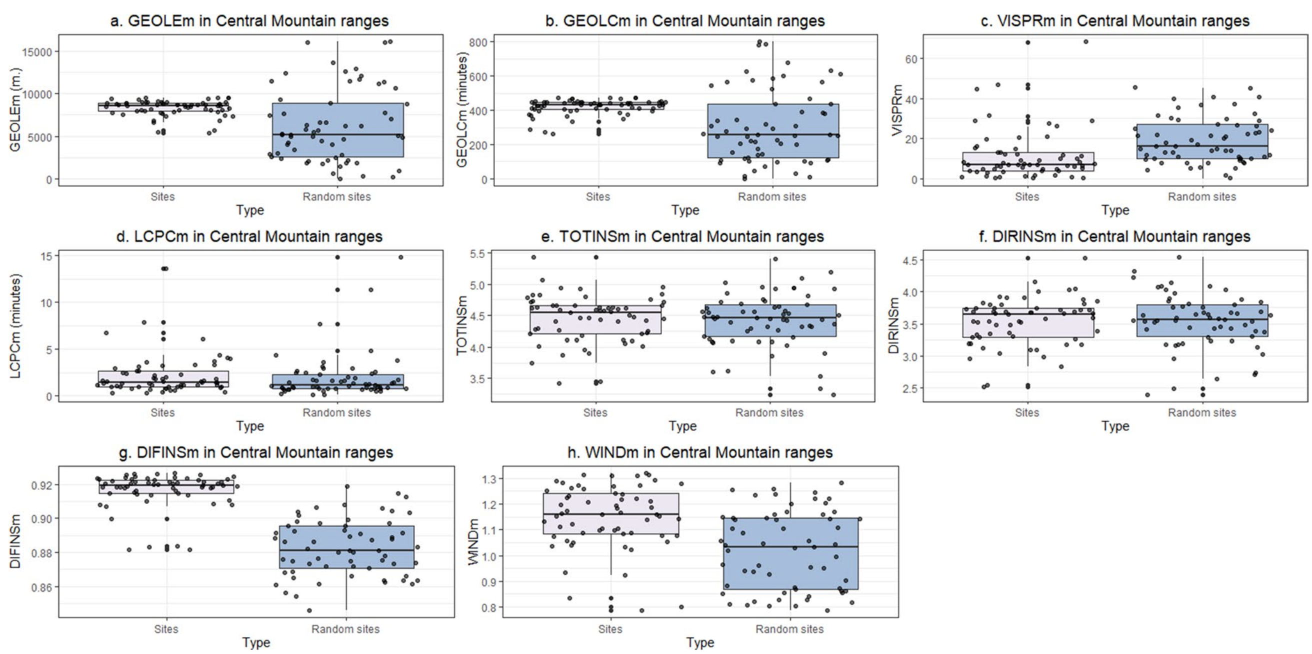


Fig. 13 Boxplots comparing the results of archaeological sites with random sites for different variables in the Central area

the data shows that archaeological sample travel times, to potential geology, cluster between 400 and 500 min, while random sample travel times exhibit greater variability. The Shapiro–Wilk test has confirmed that there is no normality (p-value = 0.01832). It has also been verified that both samples have different variances (p-value = 0.0000002122). Similar to the previous variable, it confirms the existence of significant differences between archaeological sites and random points (p-value = 0.000000003962).

For the VISPRm variable, there is a minimum value of 0 for archaeological sites and random sites. The median value for the former sample is 7 and for the latter, it is 19. The maximum value for the archaeological sample is 68 and for the random sample, it is 45. The data suggests that archaeological sites are primarily concentrated between values of 5 and 15, while random points cover a wider range of values from 10 to 25, as illustrated in Fig. 13c. It has been confirmed that there is no normality with the Shapiro–Wilk test (p-value = 0.0000000288). The variance of both samples is different (p-value = 0.000002453). The T Welch test has allowed confirming that there are significant differences between both samples (p-value = 0.0004423).

The LCPCm variable has a minimum value of 0.3 min and a minimum value of 0.11 min for archaeological sites and random sites, respectively. Looking at the median of both samples, the median for archaeological sites is 1.46 min and for random sites, it is 1.16. The maximum value for archaeological sites is 13.63 and for random points, it is 14.85 min. The distribution of values for both samples is similar, with values clustered between 0 and 2.5 min (see Fig. 13d). The Shapiro–Wilk test has confirmed that there is no normality (p-value = 3.324e-16). Furthermore, the variances of both samples are similar according to the Mann–Whitney-Wilcoxon test (p-value = 0.09815). Finally, it has been demonstrated that both samples do not possess statistically significant differences (p-value = 0.6297).

The next variable analysed is TOTINSm, related to insolation. The minimum value for archaeological sites is 3.43 and for random points, it is 3.24. The median for the archaeological sample is 4.55 and for the random representation 4.46. The maximum value for the former sample is 5.43 and for the latter, it is 5.41. Data distribution is consistent between both samples, with values ranging from 4.2 to 4.7 in Fig. 13e. It has been confirmed that there is no normality (p-value = 0.0178). The Mann–Whitney-Wilcoxon test has yielded a result that confirms both samples have similar variances (p-value = 0.6486). Ultimately, the T Student test substantiates the absence of statistically significant differences between archaeological sites and random points (p-value = 0.6973).

Another variable related to insolation is DIRINSm. The minimum value for the archaeological sample for this variable is 2.52 and for the random sample, it is 2.4. The medians for both are 3.64 and 3.57, respectively. The maximum value for archaeological sites is 4.53 and for random points, it is 4.54. A visual examination of the data (refer to Fig. 13f) suggests that the distribution in both samples is alike. The Shapiro–Wilk test has confirmed that there is no normality (p-value = 0.02601). Both samples have the same variance according to the Mann–Whitney-Wilcoxon test (p-value = 0.9918). Finally, it has been verified that there are no statistically significant differences between archaeological sites and random sites (p-value = 0.9409).

The last of the variables related to insolation is DIFINSm. For this variable, the minimum value for archaeological sites is 0.88 and for random points, it is 0.85. The median for the archaeological sample is 0.92 and for the random sample, it is 0.88. Regarding the maximum values, for archaeological sites, it is 0.93 and for random points, it is 0.92. Visually, the archaeological sample is clustered at higher values, approximately around 0.92. In contrast, the values in the random sample are notably lower (see Fig. 13g). Statistically, it has been confirmed that there is no normality (p-value = 0.000000192). The variances of both samples are different (p-value = < 2.2e-16). It has also been verified that both samples have statistically significant differences based on the T Welch test (p-value = < 2.2e-16).

For the variable WINDm, there is a minimum value of 0.79 for archaeological sites and 0.79 for random points. The median for the archaeological sample is 1.16 and for the random sample, it is 1.03. The maximum value for the former is 1.32 and for the latter, it is 1.28. Notably, the distribution of both samples varies, with archaeological sites having higher values than random points (see Fig. 13h). It has been confirmed that there is no normality (p-value = 0.00002514). There are no similarities in terms of variances between both samples either (p-value = 0.000007451). Lastly, the T Welch test has confirmed that there is statistical significance between archaeological sites and random points (p-value = 0.000001081).

Northern area vs Central area

In this section, the comparison of various variables for both study areas will be presented. The objective is to augment the preceding analysis, which compared the variables of each area with expectations under random conditions. This comparison aims to shed light on the settlement patterns in both areas.

For both the Central and Northern areas, the results for ALTA and ALTm are similar. There are no discernible differences between the values obtained at specific points and the average values in the surrounding cells. In the Central area, archaeological sites are predominantly clustered between 700 and 750 m, while in the Northern area, they span from 600 to 750 m (see Fig. 15a-b in the SM).

In terms of topographic prominence, the comparison reveals some differences. In TPI100 and TPI100m, sites in the Central area exhibit a tighter cluster between -2 and 2, while in the Northern area, they range from -2 to 3 (refer to Fig. 15c-d in the SM). In TPI500, the archaeological sites in the Northern area are situated in more prominent areas and a similar distribution is observed when considering TPI500m (Fig. 15e-f in the SM). The variables TPI1000 and TPI1000m follow a similar pattern, where sites in the Central area are found in less prominent areas compared to the Northern area. However, for TPI1000m, the values are more equalized and even the median of the Central area points is slightly higher (Fig. 15g-h in the SM).

The last two variables related to altitude are ALTrA and ALTrB, both complementary. When comparing ALTrA in the two areas, it is observed that there are no major differences between them and the sites are located at similar indices because the median of the areas is the same (0.95) and is very close to 1, which indicates that the sites are in prominent areas. In the Northern area, the ALTrB index ranges from 1.01 to 1.16 and the median is located at 1.08. Meanwhile, in the Central area, it ranges from 0.99 to 1.14 and the median is located at 1.05 (see Fig. 14a-b).

For the archaeological sites in the Northern area, the variables SLO and SLOm have slightly higher values than in the Central area (Fig. 15i in the SM). The sites are located on moderate slopes, whereas in the Central area, the slopes are lower and their distribution is concentrated between values of 4 and 8 (Fig. 16a in the SM). The slope of the geomorphological area (SLOga) has low values in the Central Mountain ranges, while in the Northern Mountain ranges, they are slightly higher. These can be considered average values that make occupation less favourable (Fig. 14c).

Regarding the Theoretical Slope (SLOt), the values in the Northern area are concentrated very close to zero, while in the Central area, they are grouped around 10 (Fig. 14d). Another variable for analysing the surroundings of the sites is the steepest real slope (SLOst), which is related to the theoretical slope. The coincidence of both indices, in some of the sites, would indicate that in both cases, these areas are completely uniform and do not show differences between the highest and lowest slope values (Fig. 14e). Based on this, in the Central area, 29.51% of the sample shows coincidences in both variables, while in the Northern area, it accounts for 32.35% of the total (Table 15).

The last two variables related to slope, which also approach the study of accessibility, are the plateau slope index (SLOpi) and the increase in the 15' to 45' isochrone (INCr15-45). For the SLOpi variable, values close to 0 indicate more difficult accessibility, while values close to 100 indicate complete accessibility. In both areas, it has been observed that the sites, being located in mountainous areas, have poor accessibility, as their indices are low and close to 0 (see Fig. 14f). As for the INCr15-45 variable, it involves comparing the increase in the 15' isochrone to the 45' isochrone. If the index obtained is less than 9, accessibility is good and if it is greater than 9, it is poor. This has resulted in the Northern area having a higher percentage of sites with good accessibility (67.64%), while in the Central area, sites with poor accessibility predominate (60.66%) (see Fig. 14g and Table 16).

The two variables used for calculating accessibility are complementary and as can be observed, they yield different results. While the plateau slope analyses an environment close to the site (based on a maximum cost of 10'), the

Table 15 Total number and percentage of sites where the SLOt and SLOst indexes are coincident

Area	Number of sites	Total	Percentage
Northern	11	34	32.35%
Central	18	61	29.51%

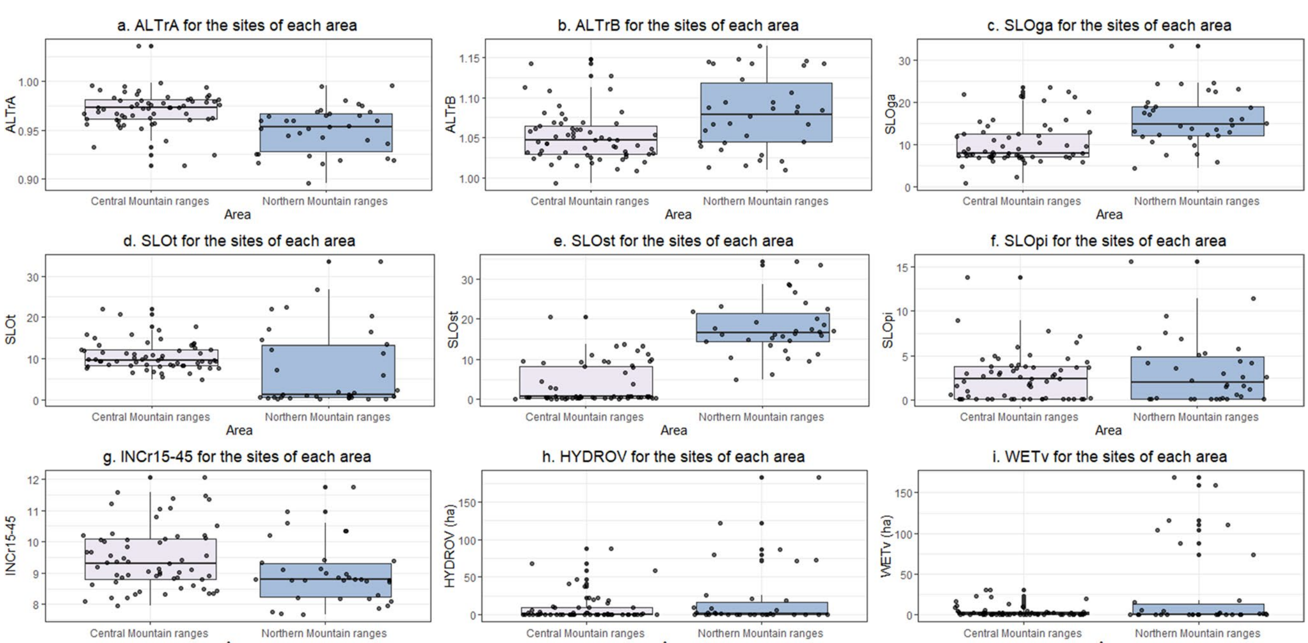


Fig. 14 Boxplots comparing the results of archaeological sites in Central area and Northern area for some variables

Table 16 Number and percentage of sites with poor and good accessibility for both study areas

Area	Sites with value < 9	Sites with value > 9	Percentage < 9	Percentage > 9
Northern	23	11	67.65%	32.35%
Central	24	37	39.34%	60.66%

isochrone increase is based on data from the surrounding area (from about 15' to 45'). The results show that in the immediate vicinity, the terrain relief conditions accessibility. In mountainous areas, accessibility is more challenging than in flatter or lower areas. On the other hand, when this analysis is extended to a medium-range environment, this model no longer holds and the nearby geomorphology of the area does not matter. The Central area has poorer accessibility in a medium-range environment than the Northern area.

Regarding the aspect, understood as the slope orientation, two variables have been analysed. The first is ASP and the second is ASPm. The medians of the sites in the Northern area are located between values of 200 and 250 for both variables. Meanwhile, for the Central area, the median is situated between 150 and 200 (Fig. 16b-c in the SM). This translates to the archaeological sites in the Northern area tending to be oriented to the Southwest, while the points in the Central area tend to be oriented to the Southeast.

In the analysis of potential hydrology, various variables have been studied, categorized into two groups: those related to watercourses and those related to wetland areas. For the first group, the Euclidean distance between potential river courses and sites in each area (HYDROE) has been calculated. After this analysis, it can be observed that in the Northern Mountain ranges, archaeological sites are located at short distances, while in the Central Mountain ranges, they are situated at medium distances (see Fig. 16d in the SM).

It is in the Central area that the trend of the archaeological sample with respect to randomness varies, as shown in the previous section. This trend is confirmed by the variable that analyses the Euclidean distance to watercourses in the immediate cells of each site (HYDROEm), which is different in each area (refer to Fig. 16e in the SM). Therefore, the Euclidean distance to river courses appears to be a relevant factor in the Central area, not because the sites in this area are closer than usual but because they are located at medium distances.

The proximity to hydrological courses calculated in travel time cost (HYDROC) has allowed verifying that the same thing occurs in the 2 study areas as for the previous variable (Fig. 16f in the SM). Similarly, with the travel cost analysis in nearby cells (HYDROCm), a close pattern is observed as for HYDROEm (Fig. 16 g in the SM). In the Central Mountain ranges, sites are at medium time distances and do not coincide with the expected randomness conditions, whereas in the Northern area, archaeological sites are located in the immediate vicinity, but their pattern matches when comparing the data with random samples, as seen previously.

Water is a fundamental resource for life and our ancestors must have had areas where they could regularly obtain water, which also served for animals to drink and be hunted. Based on this premise, the surface of potential visible watercourses (in ha) from each of the archaeological sites has been analysed using the variable HYDROV. When comparing the data from the 2 areas, it has been found that there is very limited or practically no visual control for both mountain areas (refer to Fig. 14h).

The sites in both areas are located nearby, just a few minutes travelling time from potential wetland areas modelled with the variable WET. Sites in the Central Mountain ranges are slightly more dispersed and at greater distances (see Fig. 16 h in the SM). Analysing the mean value of the cells adjacent to the sites, using WETm and comparing these values in both areas, the same trend as the previous variable is observed (Fig. 16i in the SM). Regarding the visual control of wetland areas obtained with WETv, the same trend as with the visual control of potential hydrology (HYDROV) has been identified. In both areas, there is limited visual control (refer to Fig. 14i), but this visual scarcity is more notable in the Central area (Table 17).

Potential geology emerges as a variable to be considered since hunter-gatherer communities crafted their tools from stone and needed areas to procure raw materials. Some raw materials might have come from more distant areas, probably through trade with other groups. While local materials were used, the selection was also influenced by the need for high-quality materials suitable for laminar and microlaminar techniques, which were not always available locally (Llana Rodríguez 1990). The variables related to potential geology selected and how the areas susceptible to exploitation were chosen to have already been explained in the methodological section. However, it should be noted that this is a very basic approach and should be further investigated in the future to obtain more robust results. In any case, it has been observed that in the Northern Mountain ranges, archaeological sites are close to areas with geological potential (GEOLE), at

Table 17 Statistical summaries of the values in each study area for the WETv variable

Area	Minimum (ha.)	Median (ha.)	Mean (ha.)	Maximum (ha.)
Northern	0	0.60	25.64	168.58
Central	0	1.02	4	30.78

distances ranging from 0 to 1 km. On the other hand, in the Central Mountain ranges, sites are located between 7 and 10 km (Fig. 17a in the SM). Considering the mean of the cells adjacent to each site, using the variable GEOLEm, the same pattern is observed (Fig. 17b in the SM). It is also identified that sites in the Central area are farther away than those in the Northern area, concentrated between 5 and 10 km.

Regarding the proximity to these areas with geological potential calculated in travel time cost in minutes using the GEOLC variable, it shows a similar picture to what was mentioned in the previous paragraph. In the Northern area, sites are concentrated between 0 and 100 min (see Fig. 17c in the SM). For archaeological sites in the Central area, the bulk of the sample is between 400 and 500 min. When calculating this value for the cells surrounding the sites, based on GEOLCm, the same trend persists (Fig. 17d in the SM). Just as it was done with watercourses and potential wetland areas, the visible surface with geological potential has also been calculated using the GEOLV variable. These surfaces are again larger in the Northern area compared to archaeological sites in the Central area (refer to Fig. 15a and Table 18).

After completing the abiotic factors, the results for the biotic factors will be analysed. In this group, there are only two major factors summarized in the areas of land suitable for hunting goats and hunting deer. Within each of these, travel costs in minutes (CPFPCGc and CPFPCDc) and exploitable surface area in ha (CPFPCGs and CPFPCDs) have been calculated. The travel cost for the CPFPCGc variable shows that in the Central area, sites are mainly around 1 min away, while in the Northern area, they concentrate between 0 and 30 min. In the latter case, values are lower and more concentrated. In the case of CPFPCDc, it is challenging to make a comparison because the values are extremely low. Almost all sites are at values of 0 (see Fig. 17e-f in the SM).

Regarding the exploitable surfaces for the CPFPCGs variable, it was found that in the Central area, there is a trend towards low surfaces, close to zero. However, for sites in the Northern area, values are higher and there are larger exploitable areas for hunting these herbivores. For the CPFPCDs variable, there are larger areas. But the trend is different from the previous variable. It is observed that in the Central area, sites have a greater extent than in the Northern area (see Fig. 15b-c).

Once the abiotic and biotic constraints have been studied, the group of other conditioning factors will be compared. Within this group are the variables related to visibility, insolation, proximity to least cost paths and potential wind. The first of these is the variable related to the visual catchment area from each of the sites in each zone (VISC, calculated in ha). This variable can provide an idea of visual control from the sites and can be complemented by visibility from the rest of the terrain to each of the sites through VISZ variable. For both zones, it is observed that sites have low values for the VISC variable (refer to Fig. 15d).

Regarding VISZ and based on the surface from which each of the sites can be seen (in ha), it is very similar to the previous variable (Fig. 15e). There are archaeological sites that see more surface area and are seen from less surface

Table 18 Statistical summaries of the values in each study area for the GEOLV variable

Area	Minimum (ha.)	Median (ha.)	Mean (ha.)	Maximum (ha.)
Northern	0	18.70	179.65	1063.06
Central	0	0	9.67	73.77

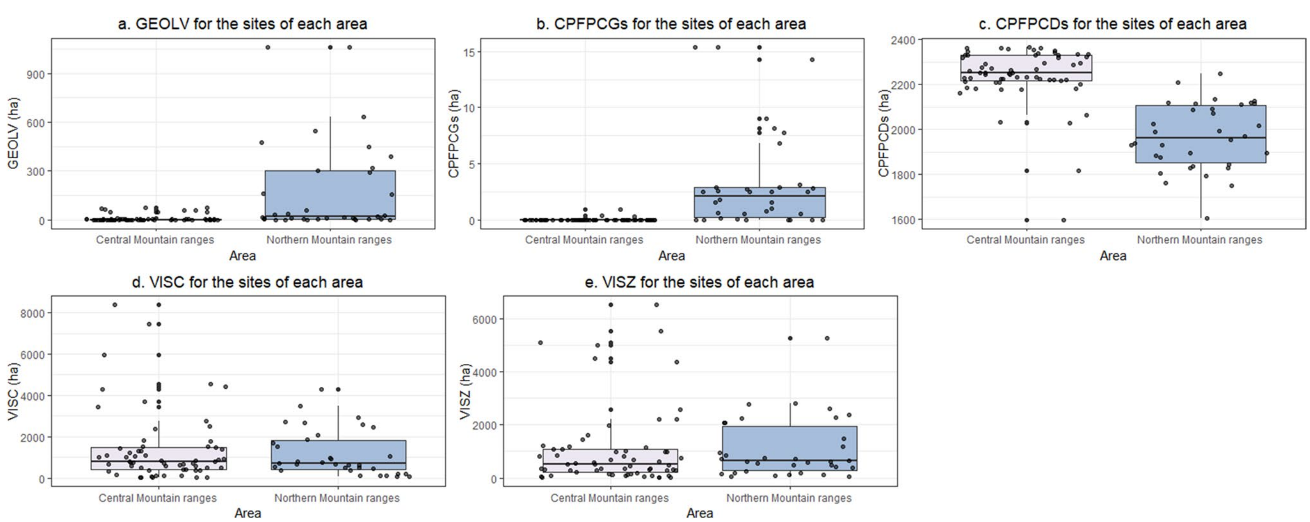


Fig. 15 Boxplots comparing the results of archaeological sites in Central area and Northern area for some variables

area and vice versa, which specific data could indicate if some of these sites have a clear visual control over certain areas or, on the contrary, are located in areas with the aim of being seen from different parts of the territory and acting as reference points in the landscape. This study has not addressed the individual analysis of each of the sites because the aim is to study the general trend of each area.

The next variable to be evaluated is the analysis of visual prominence through VISPR. The higher the value of this variable, the more visible the archaeological site is. Taking this into account, it has been found that in neither of the zones are the sites associated with high values but, quite the opposite, tend to have low values. Nevertheless, those with higher values are in the Northern area compared to the Central area. The mean visual prominence of the cells adjacent to each of the sites using the VISPRm variable was also considered, confirming practically the same trend as the previous variable (Fig. 17 g-h in the SM).

The sites in both study zones are located near potential transit routes, analysed with the LCPC variable (refer to Fig. 17i in the SM). When calculating the same variable for cells adjacent to the sites, through LCPCm, similar results were found (Fig. 18a in the SM). These results seem to indicate that there is a relationship of proximity between archaeological sites in both areas and potential transit routes.

The next variables to be studied are those related to potential insolation. Among these is TOTINS, which in broad terms, is composed of direct and diffuse insolation, which will be analysed later. The results show values located in medium and medium–high indices. It is observed that the distribution for each area is very similar. In the calculation of the average of the cells adjacent to the sites, using TOTINSm, practically the same distribution pattern was confirmed (see Fig. 18b-c in the SM). Another variable related to potential insolation is direct insolation, whose variable has been named DIRINS. It has been found that the sites in both zones tend to have medium or medium–high indices. The result of the calculation of direct insolation in the cells near each site, through DIRINSm, yields similar results. Although the archaeological sites in the Northern area have slightly higher values (Fig. 18d-e in the SM). The last of the variables related to insolation is DIFINS, where higher values are found in the sites of the Central area. This data is reinforced when calculating differential insolation in the cells near the archaeological sites using DIFINSm (Fig. 18f-g in the SM).

The final set of variables examined pertains to the influence of wind exposure. Higher values in these variables indicate greater susceptibility to wind exposure, whereas lower values suggest reduced exposure. For the WIND variable, a noteworthy pattern has emerged. Archaeological sites in the Northern area display moderate to low values in comparison to what would typically be expected in their geographic region. Conversely, in the Central area, the sites exhibit higher values (refer to Fig. 18 h in the SM). This suggests that these sites in the Central area may be located in spaces that are more exposed to prevailing winds, possibly on slopes or in positions directly facing these winds. Furthermore, it is important to note that both mountainous regions have wind exposure indices that align closely with what would be anticipated in random conditions and within the respective areas. This alignment is substantiated when comparing the archaeological samples in each zone with random samples, as seen previously in this paper. An additional variable, WINDm, which calculates the mean of the dominant wind patterns in the cells surrounding the sites, reaffirms the same trends observed with the preceding variable (see Fig. 18i in the SM) In summary, these results suggest that Palaeolithic hunter-gatherer groups may not have prioritized protection from prevailing winds when selecting the locations for their settlements in those areas.

Discussion

The investigation of Palaeolithic archaeological sites in Northwestern Iberia has provided valuable insights into the factors influencing their location. Absolute altitude emerged as a pivotal variable in determining site placement in both the Northern Mountain ranges and the Central Mountain ranges. This revelation underscores the importance of considering altitude in future research. However, it is crucial to approach this finding with caution. While altitude is evidently a fundamental factor, it does not imply a universal preference for high, medium, or low-altitude sites. Instead, the pattern of altitude-based site distribution is closely aligned with the specific characteristics of each region, a pattern consistent in both study areas. Furthermore, it is essential to mention that in colder periods, higher altitudes might have been less habitable (Viana-Soto and Pérez-Alberti 2019).

Another significant factor shaping site locations is the cost of traveling to potential wetland areas, primarily in the Northern area and the cost of reaching potential watercourses, which holds importance in the Central area. These hydrology-related variables prove instrumental in predicting the archaeological site's locations. As with altitude, proximity does not consistently define the site distribution pattern, as sites can also be situated in intermediate zones. This study not only reaffirms the importance of factors such as altitude, wetland areas and watercourses as mentioned in prior literature, but it also highlights other variables with statistically significant differences worth considering. These variables include potential geology and diffuse insolation, particularly in the Northern area and aspect, specifically in the Central area. The analysis reveals general trends within the samples of archaeological sites in each region, but it is essential to recognize that examining individual sites might uncover distinctive characteristics related to their chronology and functionality.

The primary focus of this study has been the examination of general trends in the location of Palaeolithic sites across the landscape, emphasizing what can be categorized as primary factors or first-order effects. Notably, central emphasis has not been placed on the investigation of what can be termed secondary factors, particularly those related to cultural and social considerations. These called second-order effects, often associated with the spatial relationships between sites (Nakoinz and Knitter 2016) and their possible cultural or social implications (Bevan et al. 2013), have not been the central theme of the current analysis. However, recognizing the importance of these secondary factors in shaping site distribution, future research will necessitate a more comprehensive examination of the spatial relationships among archaeological sites, delving into the potential cultural and social influences that play a role in their placement.

In the Northern area, a recurring model of site occupation revolved around the notion that shelters, near which Palaeolithic sites were located, functioned as landmarks in the landscape. These sites might serve as prominent zones in their immediate vicinity, within approximately 0 to 100 m. Beyond this range, these characteristics tend to fade. In this sense, sites in the Cantabrian region have been identified as landmarks and interpreted as places that allowed them to take part in their social construction (García-Moreno 2013).

The study explored accessibility to sites within both the immediate and medium-range environments. Accessibility has been a major factor to explain settlement patterns in other regions of Iberia (Mas et al. 2018; Fano et al. 2022). The present study has found that in the immediate environment, the findings suggest a model of relatively poor accessibility, potentially adapted to the mountainous terrain. However, the overall trend in both study regions, considering regional characteristics, indicates good accessibility. In the medium-range environment (15–45 min), two models emerged: one revealing better accessibility in the Northern area and the other indicating poorer accessibility in the Central area.

Notably, zones for potential catchment, such as goats and deer, appeared statistically less significant in site location choices in both study areas. However, their relevance varied between the Northern area, where proximity to potential goat capture zones seemed significant and the Central area, where proximity to deer exploitation areas took precedence due to larger exploitable surfaces.

The relationship between Palaeolithic archaeological sites and transit zones, a notable criterion in previous research (Ramil Rego and Ramil Soneira 1996; López Cordeiro 2015), was found to be less significant than expected when compared to random conditions. This suggests the necessity for a more focused investigation in subsequent research, in agreement with prior findings (Díaz Rodríguez 2017; Díaz-Rodríguez et al. 2023).

With regard to insolation, most sites tended to be located in areas with medium to high levels of insolation, with few significant deviations from randomness. One exception was diffuse insolation in the Northern Mountain ranges, where this variable effectively predicted site locations, implying that in regions with greater topographical variability, insolation plays a more crucial role. Comparing the results obtained with other region, in the case of the Late Magdalenian and Cantabrian Azilian, the analysis of potential insolation at archaeological sites has revealed that it does not appear to have been a determining factor in site selection. However, it is possible that during specific seasons, hunter-gatherers may have favoured sites with high insolation among those meeting other essential criteria (García-Moreno 2015).

An additional variable that emerged in the literature was the shelter from dominant winds (Ramil Rego 1989/1990). The analysis indicated that sites in the Northern Mountain ranges, were typically located in areas with medium to low wind shelter values, suggesting a preference for more exposed settings. However, it is the contrary on Central Mountain ranges.

Regarding the selected variables, it has been found that some of them provided similar results, often because their creation was based on the same underlying principles. For example, variables derived from Euclidean distance and travel time costs demonstrated statistically similar results in their comparisons. For this reason, it is necessary to take this into account when including them in further analysis.

In the Northern and Central Mountain ranges of Northwestern Iberia, landscape dynamics such as river valley erosion and sediment deposition have significantly influenced the archaeological record. Understanding these processes is essential for interpreting the distribution and preservation of Palaeolithic sites. Therefore, considering landscape dynamics is crucial when studying settlement patterns in this area. Previous studies, such as the work of Dimuccio et al. (2023) in Leiria, Portugal, have provided valuable insights into the preservation potential of Palaeolithic sites. This study offers a comprehensive perspective on the relationship between landscape evolution and site preservation, which can be similarly applied to the study areas in Northwestern Iberia.

When comparing the analyses of this study with other neighbouring areas, such as the Cantabrian region or Western Iberia, notable differences become apparent. In a region adjacent to this study area, Eastern Cantabria, it was documented that sites inhabited during the Upper-Final Magdalenian were located either on predominantly flat terrain or in rocky areas (García Moreno 2010). In the case of other areas in Western Iberia studied within the context of the Middle Palaeolithic, archaeological sites have been found at lower elevations and closer distances to riverbanks. This pattern appears to influence the availability and exploitation of lithic raw materials (Cascalheira et al. 2022). Finally, the results presented at the Monforte de Lemos basin area, in Northwestern Iberia, show that the main predictor variables are elevation, slope, cost to potential hydrology, the cost to wetland areas and visual prominence (Díaz-Rodríguez et al. 2023). These findings underscore the importance of considering the specific characteristics of each region when interpreting past prehistoric settlement patterns.

Such comparative insights not only elucidate regional variations in site selection but also contribute to a broader understanding of the adaptive strategies employed by Palaeolithic populations across different environmental contexts. The observed similarities and differences between regions may reflect a complex interplay of environmental factors, such as topography, hydrology and resource availability, which influenced the subsistence strategies and mobility patterns of prehistoric groups. Additionally, cultural factors, including technological innovations, social organization and symbolic behaviours, likely played a role in shaping the spatial distribution of settlements. For example, variations in site location could indicate specialized adaptations to local ecological niches or the transmission of cultural practices over time.

In summary, these findings provide valuable insights into the multifaceted factors that shape the location of Palaeolithic archaeological sites in Northwestern Iberia. These results highlight the need for region-specific approaches when examining occupation patterns and underscore the significance of a nuanced understanding of the intricate interplay of these factors.

Conclusion

In summary, the theoretical model proposed in the existing literature had been rigorously tested, demonstrating the utility of statistical analysis in evaluating a comprehensive set of 50 variables within two regions that exhibit some orographic similarities and are hypothesized to have been occupied during overlapping broad chronological periods, including the Upper Palaeolithic and Epipaleolithic. However, it is important to note that the lack of absolute dating makes precise chronological comparisons challenging and this hypothesis is based on relative dating and stratigraphic correlations. The findings offer a rich tapestry of insights that not only illuminate intricate details of site placement but also underscore the regional distinctions that challenge the notion of a uniform occupation pattern.

In brief, the sites within the Central Mountain ranges are predominantly situated in high-altitude areas with gentle slopes. They are positioned near wetland zones, in intermediate settings concerning potential hydrology, with relatively low visual prominence. These sites typically lack wind protection, orient towards the Southeast, lie in proximity to potential transit routes and are easily accessible from nearby areas. However, access from medium-range environments is comparatively challenging. These sites benefit from moderate to high insolation levels and possess a moderate topographical prominence index. Additionally, they offer extensive nearby areas suitable for deer hunting. Furthermore, they are located at medium distances from geological potential zones (refer to Table 6 in the SM).

Conversely, the sites in the Northern Mountain ranges are primarily found in mid-altitude regions with gentle slopes. They are located close to wetland areas, in intermediate potential hydrology environments, with relatively low visual prominence. These sites exhibit limited visibility but are sheltered from dominant winds. They are oriented to the Southwest, situated near potential transit routes and readily accessible. Furthermore, they enjoy medium to high levels of insolation, with differential insolation becoming a significant variable in areas with higher values. They feature a moderate topographical prominence index and offer extensive nearby areas suitable for goat hunting. These sites are positioned in proximity to geological potential zones (refer to Table 6 in the SM).

Their findings enable them to both validate and challenge hypotheses grounded in the theoretical model used to explain the location of Palaeolithic sites in Northwestern Iberia. While some hypotheses have found support, others warrant further examination and refinement. This methodology, integrating Geographic Information Systems (GIS) and spatial statistics, has proven effective in empirically testing these initial hypotheses, facilitating the transformation of conjecture into scientifically verified results. Although researchers can never definitively ascertain the thoughts of ancient hunter-gatherer communities, this work serves as a foundational cornerstone for future research into Palaeolithic settlement patterns.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12520-024-02049-7.

Acknowledgements This work has been prepared during a postdoctoral research stay at the Aarhus University within the framework of the CLIOARCH project which has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 817564); and the NeanderEDGE project which has received funding from the Independent Research Fund Denmark (case number 9062-00027B). The analyses of this work have been carried out in the statistical environment R (R Core Team 2021). I would like to extend my sincere gratitude to the Editor for their dedicated work and diligence throughout the review process. I also want to express my appreciation to the anonymous reviewers whose thoughtful comments and constructive feedback have significantly contributed to the improvement of this manuscript. Their meticulous review has undoubtedly elevated the quality of this work. I am truly thankful for their time and effort in helping me refine this research.

Authors contribution M.D-R.: Conceptualization, Methodology, Software, Validation, Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization.

Funding M. Díaz-Rodríguez is granted by Margarita Salas Fellowship from the European Union—NextGeneration EU by the Programme for the requalification, international mobility and attraction of talent through the University of Santiago de Compostela (Spain) and the Ministry of Universities (Spain) (UP2021–042).

Data availability The data used in this paper and the code executed in R are available at https://github.com/mikeldiazrodriguez/mountain_ranges and archived on zenodo (https://doi.org/10.5281/zenodo.12629750).

Declarations

Competing interest The authors declare no competing interests.

References

Alonso del Real C, Vázquez Varela JM (1976) Excavaciones en el abrigo Vidal I, en Prado do Inferno. Excavaciones en Pena Grande. Excavaciones en Os Penedos de O Carrizo. Not Arqueol Hispánico 5:53–61

Baddeley A, Rubak E, Turner R (2015) Spatial point patterns: methodology and applications with R. Chapman and Hall/CRC Press

Baddeley A, Turner R, Rubak E (2020) spatstat: Spatial Point Pattern Analysis, Model-Fitting, Simulation, Tests

Becker RA, Wilks AR, Brownrigg R, et al (2021) maps: Draw Geographical Maps

Bevan A, Crema ER, Li X, Palmisano A (2013) Intensities, Interactions, and Uncertainties: Some New Approaches to Archaeological Distributions. In: Bevan A, Lake M (eds) Computational Approaches to Archaeological Spaces. pp 27–52

Bivand R, Keitt T, Rowlingson B, et al (2020a) rgdal: Bindings for the 'Geospatial' Data Abstraction Library

Bivand R, Lewin-Koh N, Pebesma E, et al (2020b) maptools: Tools for Handling Spatial Objects

Bivand R, Rundel C, Pebesma E, et al (2020c) rgeos: Interface to Geometry Engine - Open Source ('GEOS')

Bivand RS, Pebesma E, Gómez-Rubio V (2013) Applied Spatial Data Analysis with R. Springer, New York

Bocinsky K (2017) Landscape-based Hypothesis Testing in R. In: Marwick B (ed) How To Do Archaeological Science Using R. Vancouver

Böhner J, Antonić O (2009) Land-Surface Parameters Specific to Topo-Climatology. In: Hengl T, Reuter HI (eds) Geomorphometry: Concepts, Software, Applications. Elsevier, Amsterdam (The Netherlands), pp 195–226

Brown PE (2015) Model-Based Geostatistics the Easy Way. J Stat Softw 63:1–24. https://doi.org/10.18637/jss.v063.i12

Burke A, Levavasseur G, James PMA et al (2014) Exploring the impact of climate variability during the Last Glacial Maximum on the pattern of human occupation of Iberia. J Hum Evol 73:35–46. https://doi.org/10.1016/j.jhevol.2014.06.003

Burke A, Kageyama M, Latombe G et al (2017) Risky business: The impact of climate and climate variability on human population dynamics in Western Europe during the Last Glacial Maximum. Quat Sci Rev 164:217–229. https://doi.org/10.1016/j.quascirev.2017.04.001

Burke A, Peros MC, Wren CD et al (2021) The archaeology of climate change: The case for cultural diversity. Proc Natl Acad Sci U S A 118:1–10. https://doi.org/10.1073/pnas.2108537118

Cannon MD (2003) A model of central place forager prey choice and an application to faunal remains from the Mimbres Valley, New Mexico. J Anthropol Archaeol 22:1–25

Carrero Pazos M (2017) El Fenómeno tumular y megalítico en Galicia. Aportaciones desde los Sistemas de Información Geográfica y la estadística espacial para el estudio de los patrones de localización. Tomo I. Universidade de Santiago de Compostela

Cascalheira J, Gonçalves C, Maio D (2022) The spatial patterning of Middle Palaeolithic human settlement in westernmost Iberia. J Quat Sci 37:291–299. https://doi.org/10.1002/jqs.3286

Cazorla Martín R, Cerrilo Cuenca E, Mayoral Herrera V, Salgado Carmona JA (2008) El sistema defensivo de Villasviejas de Tamuja (Botija, Cáceres): una aplicación de software libre en el análisis de paisajes arqueológicos. In: II Jornadas de SIG libre

Cerqueiro Landín D (1989) La utilización del espacio en el Paleolítico. Universidade de Santiago de Compostela, Patrón de asentamiento en la Sierra de O Bocelo

Cerrillo Cuenca E (2011) Recorriendo un territorio desaparecido: restitución fotogramétrica y análisis del paisaje de la necrópolis prehistórica del vado de Alconétar. In: Tecnologías de Información Geográfica y análisis arqueológico del territorio. Actas del V Simposio Internacional de arqueología de Mérida. pp 145–159

Conrad O, Bechtel B, Bock M, et al (2015) System for Automated Geoscientific Analyses (SAGA) v.2.1.4. Geosci Model Dev 1991–2007. https://doi.org/10.5194/gmd-8-1991-2015

Criado Boado F (1991a) Arqueoloxía da paisaxe: estudio da Serra do Bocelo e do val do río Furelos (Toques e Melide, A Coruña). In: Arqueoloxía. Informes. Campaña 1988. Consellería de Cultura e Xuventude, Xunta de Galicia, Santiago de Compostela (A Coruña), pp 103–106

Criado Boado F (1991) Arqueoloxía da paisaxe: estudio da Serra do Bocelo e do val do río Furelos (A Coruña). Larouco 1:99–105. https://doi.org/10.1590/S1516-18462008000300012

Criado Boado F, Bonilla Rodríguez A, Cerqueiro Landín D, et al (1991) Arqueología del paisaje, el área Bocelo-Furelos entre los tiempos paleolíticos y medievales (campañas de 1987,1988 y 1989). Dirección Xeral do Patrimonio Cultural, Santiago de Compostela (A Coruña)

Criado Boado F, Bonilla Rodríguez A, Cerqueiro Landín D, et al (1988) Proyecto Bocelo-Furelos: Arqueologia del Paisaje y Prospección intensiva en Galicia. Trab Antropol e Etnol XXVIII:241–250

Criado Boado F, Bonilla Rodríguez A, Cerqueiro Landín D, et al (1989a) Prospección Intensiva da Serra do Bocelo e val do río Furelos. In: Arqueoloxía. Informes. Campaña 1987. Xunta de Galicia / Consellería de Cultura e Xuventude, A Coruña, pp 128–134

Criado Boado F, Bonilla Rodríguez A, Cerqueiro Landín D, et al (1989b) Resultados de la prospección intensiva en la Sierra de O Bocelo y en el Valle del Furelos (A Coruña). Cuad Estud Gall XXXVIII:25–50. https://doi.org/10.1590/S1516-18462008000300012

Criado Boado F, Cerqueiro Landín D (1991) Paleolítico: actuacións en xacementos con industria lítica. In: Arqueoloxía. Informes. Campaña 1988. Consellería de Cultura e Xuventude, Xunta de Galicia, Santiago de Compostela (A Coruña), pp 107–110

Daniel F, Ooi H, Calaway R, et al (2022) foreach: Provides Foreach Looping Construct de Lombera Hermida A, Rodríguez Álvarez XP, Fábregas Valcarce R, Lazuén Fernández T (2011) The Paleolithic Settlement of the Monforte Basin (Lugo, Galicia). In: Lombera Hermida A, Fábregas Valcarce R (eds) To the West of Spanish Cantabria. The Paleolithic Settlement of Galicia. BAR International Series 2283, Oxford, pp 93–110

de Lombera Hermida A, Rodríguez Álvarez XP, Rabuñal Gayo X et al (2012) El yacimiento de Valverde (Monforte de Lemos, Lugo, Galicia, España) y las primeras evidencias de poblamiento en el pleniglacial del NO de la Península Ibérica. Espac Tiempo y Forma Ser i, Prehisoria y Arqueol 5:363–382. https://doi.org/10.5944/etf

de Lombera Hermida A, Díaz Rodríguez M, Pérez Alberti A et al (2015) Evolución de los patrones de asentamiento de los yacimientos paleolíticos de la Depresión de Monforte de Lemos (Lugo, Galicia). Sémata Ciencias Sociais e Humanidades 27:267–297

De Reu J (2012) The Land of the Dead. Universiteit Gent

De Reu J, Bourgeois J, De Smedt P et al (2011) Measuring the relative topographic position of archaeological sites in the landscape, a case study on the Bronze Age barrows in northwest Belgium. J Archaeol Sci 38:3435–3446. https://doi.org/10.1016/j.jas.2011.08.005

De Reu J, Bourgeois J, Bats M et al (2013) Application of the topographic position index to heterogeneous landscapes. Geomorphology 186:39–49. https://doi.org/10.1016/j.geomorph.2012.12.015

Díaz Rodríguez M (2017) Patrones de asentamiento en el Paleolítico superior y el Epipaleolítico en el Noroeste Peninsular. Cálculo de rutas óptimas y movilidad. In: Vázquez Martínez A, Cordeiro Macenlle R, Carrero Pazos M, et al. (eds) (Re)escribindo a Historia. Achegas dos novos investigadores en Arqueoloxía e Ciencias da Antigüidade. Santiago de Compostela (A Coruña), pp 25–46

Díaz Rodríguez M, Carrero Pazos M (2019) La Prominencia Topográfica de los yacimientos en abrigo en las sierras septentrionales (Lugo, Galicia). Una aproximación de estudio con SIG. In: Los tiempos cambian, de la piedra al teclado. X Jornadas de Jóvenes en Investigación Arqueológica. Burgos, pp 363–373

Díaz-Rodríguez M, Fábregas-Valcarce R, de Lombera-Hermida A et al (2021) Análisis locacional del yacimiento Paleolítico de Cova Eirós (Triacastela, Lugo). Estud do Quaternário 21:14–32. https://doi.org/10.30893/eq.v0i20.199

Díaz-Rodríguez M, Fábregas-Valcarce R (2022) Evaluating the effectiveness of three spatial cluster analysis methods on Palaeolithic site location distributions in Galicia, NW Iberian Peninsula. J Archaeol Sci Reports 41. https://doi.org/10.1016/j.jasrep.2021.103323

Díaz-Rodríguez M, Fábregas-Valcarce R, Pérez-Alberti A (2023) A predictive model for Palaeolithic sites: A case study of Monforte de Lemos basin, NW Iberian Peninsula. J Archaeol Sci Reports 49:104012. https://doi.org/10.1016/j.jasrep.2023.104012

Dimuccio LA, Ferreira R, Batista A et al (2023) Predictive spatial analysis for a critical assessment of the preservation potential of Palaeolithic record in the Leiria region (central Portugal). Quat Int 668:44–62

Esri (2011) ArcGIS Desktop

Fábrega Álvarez P (2004) Poblamiento y Territorio de la Cultura Castreña en la comarca de Ortegal. Santiago de Compostela (A Coruña)

Fábregas Valcarce R, de Lombera HA (2010) El paleolítico superior en Galicia a la luz de las últimas investigaciones. El Paleolítico superior peninsular. Novedades del siglo XXI. Homenaje al Profesor Javier Fortea. Monografíes SERP. Universitat de Barcelona, Barcelona, pp 255–270

Fábregas Valcarce R, Alonso Fernández S, Ameijenda Iglesias A et al (2010) Completando o mapa. Novas datacións absolutas para o Paleolítico e Mesolítico do interior galego. Gallaecia 29:5–28

Fano MÁ, García-Moreno A, Portillo F, et al (2022) Beyond images: the external archaeological context of Palaeolithic rock art (EAC). A case study from the Nalón river basin (northern Iberia). Archaeol Anthropol Sci 14:. https://doi.org/10.1007/s12520-022-01649-5

Fernández Fernández J (2010) Una aportación desde la arqueología del paisaje al conocimiento del primer poblamiento humano del Valle del Trubia. Universidad de Oviedo, Oviedo, Estudio geoarqueológico y análisis SIG del territorio

Fox J, Marquez MM, Bouchet-Valat M (2023) Rcmdr: R Commander

García Moreno A (2008) Insolación y hábitat Paleolítico en el Valle del Asón (Cantabria, España). Cuaternario y Geomorfol 22:93–105

García Moreno A (2010) Patrones de asentamiento y ocupación del territorio en el Cantábrico Oriental al final del Pleistoceno: Una aproximación mediante SIG. Universidad de Cantabria

García-Moreno A (2013) To see or to be seen... is that the question? An evaluation of palaeolithic sites' visual presence and their role in social organization. J Anthropol Archaeol 32:647–658. https://doi.org/10.1016/j.jaa.2013.03.003

García Moreno A (2015) The Incidence of Potential Insolation on Settlement Dynamics and Site Location Preferences: A Case Study From the Cantabrian Late Palaeolithic. J Archaeol Sci Reports 3:90–99. https://doi.org/10.1016/j.jasrep.2015.05.020

García-Moreno A (2015) The Incidence of Potential Insolation on Settlement Dynamics and Site Location Preferences: A Case Study From the Cantabrian Late Palaeolithic. J Archaeol Sci Reports 3:90–99. https://doi.org/10.1016/j.jasrep.2015.05.020

García García M (2015) Poblamiento y sociedad en el valle medio del Duero durante la edad de los metales (ca. 2800–450 cal. ANE). Un estudio de arqueología del paisaje. Universidad de Valladolid

García Moreno A, Fano Martínez MÁ (2014) Paleolithic sites beyond the archaeological deposits. In: García A, García J, Maximiano A, Ríos-Garaizar J (eds) Debating Spatial Archaeology: Proceedings of the International Workshop on Landscape and Spatial Analysis in Archaeology (Santander, June 8th-9th, 2012). Santander, pp 231–241

Grass Development Team (2020) Geographic Resources Analysis Support System (GRASS) Software

Henry L (2022) Tools at the Intersection of 'purrr' and 'dplyr'

Hijmans RJ, Phillips S, Leathwick J, Elith J (2017) dismo: Species Distribution Modeling

Hijmans RJ, van Etten J, Sumner M, et al (2020) raster: Geographic Data Analysis and Modeling

Karoune E, Plomp E (2022) Removing Barriers to Reproducible Research in Archaeology. Peer Community Archaeol. 10.5281/zenodo.7320029

Llana Rodríguez JC (1990) Algunas consideraciones económicas del Paleolítico superior a través de los cuarzos y cuarcitas de grano grueso. Gallaecia 12:29–38. https://doi.org/10.1590/S1516-18462008000300012

Llana Rodríguez JC, Martínez Cortizas A, Ramil Rego P (1992) Algunas consideraciones acerca de la estratigrafia y del marco temporal para los yacimientos al aire libre del Paleolítico final-Epipaleolítico de Galicia. Zephyrus 44–45:155–166

Llobera M (2001) Building Past Landscape Perception With GIS: Understanding Topographic Prominence. J Archaeol Sci 28:1005–1014. https://doi.org/10.1006/jasc.2001.0720

López Cordeiro MDM (2002) Ocupación de espacios graníticos en la Sierra de Xistral: el área arqueológica de Nordés (Valadouro, Lugo). Gallaecia 21:61–86

López Cordeiro MDM (2003) El yacimiento epipaleolítico de Chan da Cruz (Valadouro, Lugo): Síntesis de los primeros resultados. Complutum 14:39–54

López Cordeiro MDM (2004a) La prospección en el paleolítico: aplicaciones en el registro paleolítico gallego. Ser Arqueol Espac Del Semin Arqueol y Etnol Turol 24–25:33–55
López Cordeiro MDM (2004b) Resultados obtenidos en la construcción de una 'Nueva Arqueología' del paleolítico gallego. Cuad Estud Gall LI:133–151
López Cordeiro MDM (2015) Industrias Líticas y Ocupación del espacio en el Paleolítico de Galicia. Universidade de Santiago de Compostela
López Olmedo F, Palacio Suárez JP, Dávila Ruiz MDT, et al (2022) Mapa Geológico Digital continuo E. 1: 50.000, Zona Ibérica (Zona-1700). In: GEODE. Mapa Geológico Digital continuo de España
Ludwig P, Shao Y, Kehl M, Weniger GC (2018) The Last Glacial Maximum and Heinrich event I on the Iberian Peninsula: A regional climate modelling study for understanding human settlement patterns. Glob Planet Change 170:34–47. https://doi.org/10.1016/j.gloplacha.2018.08.006
Marcos Sáiz FJ (2006) La Sierra de Atapuerca y el Valle del Arlanzón. Patrones de asentamiento prehistóricos. Burgos
Marín Arroyo AB (2008) Patrones de movilidad y control del territorio en el Cantábrico oriental durante el Tardiglaciar. Trab Prehist 65:29–45
Marín Arroyo AB (2009) Assessing what lies beneath the spatial distribution of a zooarchaeological record: The use of GIS and spatial correlations at El Mirón Cave (Spain). Archaeometry 51:506–524. https://doi.org/10.1111/j.1475-4754.2008.00411.x
Martínez Cortizas A, Moares Domínguez MC (1995) Edafología y arqueología: aplicaciones al estudio de yacimientos arqueológicos al aire libre en Galicia. Consellería de Cultura, Santiago de Compostela (A Coruña)
Marwick B (2017) Computational Reproducibility in Archaeological Research: Basic Principles and a Case Study of Their Implementation. J Archaeol Method Theory 24:424–450. https://doi.org/10.1007/s10816-015-9272-9
Marwick B, d'Alpoim GJ, Barton CM et al (2017) Open science in archaeology. SAA Archaeol Rec 17:8–14
Mas B, Allué E, Sánchez de la Torre M et al (2018) Settlement patterns during the Magdalenian in the south-eastern Pyrenees, Iberian Peninsula. A territorial study based on GIS. J Archaeol Sci Reports 22:237–247. https://doi.org/10.1016/j.jasrep.2018.10.002
Nakoinz O, Knitter D (2016) Modelling Human Behaviour in Landscapes. Quantitative Archaeology and Archaeological Modelling. Springer, Londres (Reino Unido)
Parcero Oubiña C, Fábrega Álvarez P (2006) Diseño metodológico para el análisis locacional de asentamientos a través de un SIG de base 'Raster'. In: Grau Mira I (ed) La aplicación de los SIG en la arqueología del paisaje. Universidad de Alicante, pp 69–90
Pebesma E, Bivand R, Rowlingson B, et al (2020) sp: Classes and Methods for Spatial Data
Pedersen TL (2022) patchwork: The Composer of Plots
QGIS.org (2021) QGIS Geographic Information System
R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing
Ramil Rego E (2014) Prehistoria antigua del occidente cantábrico: las primeras poblaciones. Universidade de Santiago de Compostela
Ramil Rego E (1989/1990) Habitabilidad cavernícola: elección de asentamientos. Brigantium 6:191–197
Ramil Rego E, Ramil Soneira JM (1996) El fin de los tiempos glaciares en Galicia. Magdaleniense y Epipaleolítico. In: Fábregas Valcarce R (ed) Os primeiros poboadores de Galicia: o Paleolítico. Ediciós do Castro, Sada (A Coruña), pp 117–146
Ramil Rego P, Fernández Rodríguez C (1996) Marco cronológico y paleoambiental de la ocupación paleolítica en el NW Ibérico. In: Fábregas Valcarce R (ed) Os primeiros poboadores de Galicia: o Paleolítico. Ediciós do Castro, Sada (A Coruña), pp 165–191
Ramil Soneira JM, Vázquez Varela JM (1976) Cuatro yacimientos prehistóricos con industrias líticas en la provincia de Lugo. Gallaecia 2:47–54
Ramil Soneira JM, Vázquez Varela JM (1983) Primer hallazgo de arte mueble paleolítico en Galicia. Ars Praehist 2:191–193
Ripley B, Venables B, Bates DM, et al (2020) MASS: Support Functions and Datasets for Venables and Ripley's MASS
Rodríguez Álvarez XP, Fábregas Valcarce R, Lazuén Fernández T et al (2008) Nuevos yacimientos paleolíticos en la Depresión de Monforte de Lemos (Lugo, Galicia, España). Cuaternario y Geomorfol 22:71–92. https://doi.org/10.1016/S0034-7094(11)70016-2
Rodríguez Rellán C, Fábregas Valcarce R (2015) Arte rupestre galaica: unha achega dende a estatística espacial e os SIX. Semata Ciencias Sociais e Humanidades 27:9–34
Schloerke B, Cook D, Larmarange J, et al (2021) GGally: Extension to 'ggplot2'
Senín Fernández IJ (1996) Historia da investigación do Paleolítico galego. In: Fábregas Valcarce R (ed) Os primeiros poboadores de Galicia: o Paleolítico. Ediciós do Castro, pp 25–45
Turrero P, Domínguez-Cuesta MJ, Jiménez-Sánchez M, García-Vázquez E (2013) The spatial distribution of Palaeolithic human settlements and its influence on palaeoecological studies: A case from Northern Iberia. J Archaeol Sci 40:4127–4138. https://doi.org/10.1016/j.jas.2013.06.003
Viana-Soto A, Pérez-Alberti A (2019) Periglacial deposits as indicators of paleotemperatures. A case study in the Iberian Peninsula: The mountains of Galicia. In: Permafrost and Periglacial Processes. John Wiley and Sons Ltd, pp 374–388
Villar Quinteiro R (1996) El Paleolítico superior y el Epipaleolítico en Galicia. Universidade de Santiago de Compostela
Villar Quinteiro R (1997) El Paleolítico Superior y Epipaleolítico en Galicia. Zephyrus 50:71–106
Villar Quinteiro R (2008) Secuencia do Paleolítico Superior en Galicia. In: Méndez Quintas E (ed) Homenaxe a Xosé María Álvarez Blázquez II. Estudos sobre Paleolítico. Asociación Cultural Instituto de Estudos Miñoranos, Gondomar, pp 347–372
White DA, Barber SB (2012) Geospatial modeling of pedestrian transportation networks: A case study from precolumbian Oaxaca, Mexico. J Archaeol Sci 39:2684–2696. https://doi.org/10.1016/j.jas.2012.04.017
Wickham H (2020) plyr: Tools for Splitting, Applying and Combining Data
Wickham H, Averick M, Bryan J et al (2019) Welcome to the tidyverse. J Open Source Softw 4:1686. https://doi.org/10.21105/joss.01686
Wickham H, Bryan J, Kalicinski M, et al (2019b) readxl: Read Excel Files
Wickham H, Chang W, Henry L, et al (2020) ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics
Wren CD, Burke A (2019) Habitat suitability and the genetic structure of human populations during the Last Glacial Maximum (LGM) in Western Europe. PLoS ONE 14:1–22. https://doi.org/10.1371/journal.pone.0217996

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.