Archeologia e Calcolatori
34.1, 2023, 225-236
doi 10.19282/ac.34.1.2023.25

BALANCING BETWEEN BIASES AND INTERPRETATION. A PREDICTIVE MODEL OF PREHISTORIC SCANIA, SWEDEN

1. Introduction

This contribution focuses on the distribution pattern of prehistoric sites in the county of Scania (*Skåne län*), Sweden (Fig. 1). The high number of archaeological investigations carried out in the past few years makes the area particularly suitable for predictive modelling. However, a closer look at the current situation shows that our knowledge is not uniform. In fact, some areas attracted more attention due to commercial archaeology or research interest. This is the case of the Malmö area, the West coast, the North-East and, in part, the Ystad area (Larsson *et al.* 1992; Artursson 2005, 2007; Sjögren 2006; Hadevik, Steineke 2009). The difference in research intensity introduces a set of possible biases in our knowledge, on top of the less controllable (but more predictable) recovery biases or information loss. Undoubtedly, our knowledge is heavily affected by post-depositional factors and modern activities such as agricultural activity or presence of infrastructures (Bevan 2012). The goal of this paper is to quantify and disentangle these different factors, as a first step before setting up interpretative models.

The dataset used for this paper is a subset of the Swedish National Heritage database (Riksantikvarieämbetet, RAA, https://pub.raa.se/). In particular, the study deals with Stone Age settlements and different types of prehistoric monuments dating between the Early-Middle Neolithic (EN-MN) and the Bronze Age (BA). These monuments should be analysed separately due to their nature but also because of the different histories of research underlying them. This is because some monuments (mounds, megaliths and cairns) are highly visible in the landscape and therefore known since a very long time even without thorough archaeological investigations (survey and excavations). Conversely, most of the prehistoric settlements were only known after the application of systematic surveys and modern excavations. Thus, they are the product of different processes and they are affected by different biases or to a different extent from each other. Considering them at once would increase the risk of interference or confounding.

The analyses presented in this paper are all carried within an R environment (v.4.2.1 R Core Team 2022), using a fully documented and reproducible approach. Unfortunately, despite soil and geological data could greatly contribute to this model, they were not included because they are not freely accessible. The paper is structured as follows: section 2 describes in detail the

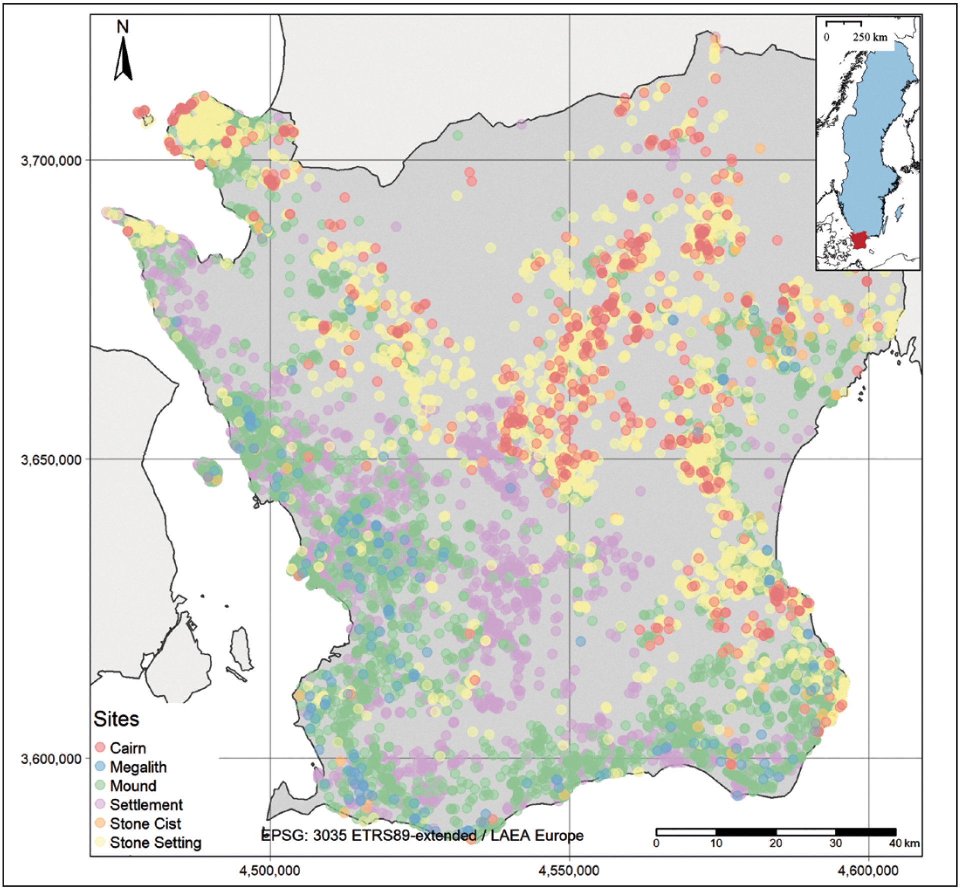

Fig. 1 – Sites and study area. On the top right corner the study area is shown on a map of Sweden.

starting data, the workflow followed and the analytical tools used; in section 3 the results are presented and section 4 discusses them with some additional remarks about future outcomes and caveats.

2. Data and methods

The data used for the analysis consist of sites and covariates. In order to be able to use them to feed the model they needed some filtering and editing. This section deals with the methodological and critical aspects of data and their handling.

2.1 *Data*

2.1.1 Archaeological data

The archaeological sites used in this study come from the Swedish Heritage database which can be freely accessed from Riksantikvarieämbetets öppna data server. In particular, data come from the *Fornlämningar och övriga kulturhistoriska lämningar* (Ancient monuments and other cultural-historical remains). As a whole, the dataset includes over sites from each (pre-)historical period. Here, only prehistoric sites belonging to the following typologies were considered: Stone Age settlements (*Stenåldersboplats*), megalithic tombs, burial mounds, cairns, stone settings. Stone Age settlements include all sites dating between the Mesolithic and the Neolithic. Unfortunately, most of them consist of stone scatters and cannot be precisely dated. These were found during field surveys and often contain mixed, un-datable or fragmented material. A refined chronology can be obtained only at the expenses of many sites. For this reason, it was decided to use them as they were, selecting the ones with a Stone Age dating. In total, 2162 sites belong to this category.

Megaliths are a special class of monuments that was built in some parts of Northern and Western Europe (Blank 2021, 20ss.). During the Funnel Beaker period (ca. 4000 BC-2800 BC in the area) two main types of megaliths were built in the study area: Dolmens and Passage graves. However, the period of megalithic construction can be restricted to the EN II-MN A II (3500-3000 BC), during the so called Klimax period (Persson, Sjögren 1995; Sjögren 2003, chap. 1). In the following degenerative period and up until the Late Neolithic (after 2300 BC) monuments were still in use but not built (other types are used, e.g. individual flat graves or gallery graves (Sjögren 2003; Blank 2021, chap. 7). In the study area 193 monuments are recorded.

Mounds, cairns and stone settings represent three different types of monuments, all interpreted as memorials and/or landmarks. Stone settings are the less impressive of the three types, but they are assumed to serve a similar function, especially when forming clusters or associated with nearby mounds (Nord 2009). What is relevant here is the difference in their construction and how this affects their preservation. Stone settings and cairns are both mainly made in stone, the former flat (or slightly domed) and the latter distinctly domed (Nord 2009,106). Mounds are earth made and have a grassy cover, although they often conceal a stone structure[1]. Most of the sites were built between periods I and IV of the Bronze Age (ca. 1700-900 BC), with

[1] It has been noted that the distinction is sometimes fuzzy (Nord 2009, 106), but this can only have an impact on single sites or a very local scale, which is not the focus here.

a peak in periods III and IV (1300-900 BC). In the Late Bronze Age fewer mounds were built and of a smaller size, whilst some older ones were used for secondary burials (Nord 2009, 103-104). From a simple distribution map we can already see some differences in the pattern, with mounds fairly widespread, with a predominance in the S and W of Scania, whereas cairns and stone settings are dominant in central and northeastern Scania (Fig. 1).

2.1.2 Biases and landscape variables

Sample bias is one of the main problems when it comes to large datasets. Another significant problem relies in research biases, with some areas being in the centre of more intense activities than others. The final picture risks to be incomplete and a simple distribution map largely meaningless. Nevertheless, if properly accounted, there is still a large potential to obtain good results when analysing them (as shown in Bevan 2012).

In this study, distribution of infrastructures (urban agglomerations, roads and railways) and farmland were used as bias variables (Fig. 2). The variables can be downloaded from Open Street Map Geofabrik landuse server, either manually or using a dedicated function from the *rbias* package (Günther *et al.* 2022). For the landscape-based modelling the following variables were considered:

– elevation, with a 30 m resolution from the EU-DEM v1.1;
– slope and the SAGA Wetness Index, calculated from the elevation using SAGA GIS (Conrad *et al.* 2015) through R, using the package Rsagacmd (Pawley 2022);
– distance from the coastline and from rivers. River data derive from Open Street Map and are also freely accessible. The resolution of the rasters was set to 30 m.

2.2 *Methods*

This study investigates the differences in presence/absence of different classes of sites tested against modern land-use (infrastructures and agricultural) and then against geomorphological variables (elevation, distance from resources, soil, etc.). The workflow is very simple and mostly based on the functions available in the *rbias* package (Günther *et al.* 2022):

– select the desired land-use data (they can be downloaded using the download_geofabrik_data function) and stack them into a single vector using the dissolve_osmdata function;
– create a distance raster using the bias_surface function. This step was also used to compute distance from the coastline and rivers;
– use the function bias_influence to create a fuzzy raster (based on Knitter, Hamer 2022) with values scaled between 0 to 1 (low to high bias). In short,

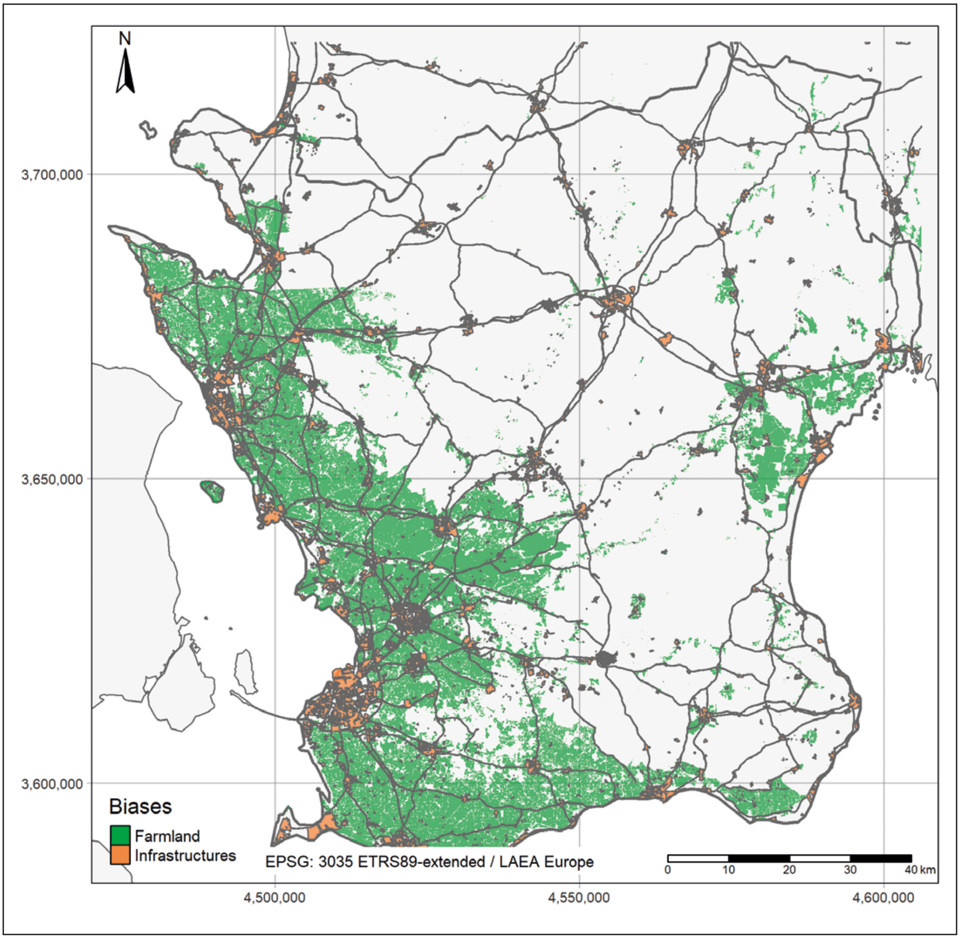

Fig. 2 – Maps of the biases used in this study (data/Maps copyright 2018 Geofabrik GmbH and OpenStreetMap Contributors).

it uses a decay function to attribute a value to every raster cell. Here, a triangular transformation with a decreasing bias from 0 m (inside the area) to 250 m was used. Of course, these values could be easily modified in order to fit different requirements. Distance from rivers and coastline were used as they were, without fuzzification;

– test the density of observations against the background using the function sites_vs_background. The function creates a frequency plot for each covariate at sampling locations, simulating the background as random sampling process with a 95% confidence interval (based on Bocinsky 2017);

– when a bias is identified, a new subset of the area is cropped according to the bias surface and only the spatial patterns of the archaeological sites within this smaller study area are tested against the landscape variables. The main idea is to obtain a region where the bias is assumed to be uniform.

3. Results

Settlements and megaliths are analysed separately while the other monuments are analysed together because of their similar chronology and function. The first part is dedicated to the identification of possible biases, followed by an analysis of covariate influence based on the results of the first step as described in Section 2.2.

3.1 *Stone Age settlements*

Settlements are highly correlated with the presence of farmland (Fig. 3a). The most likely explanation is that most of the recorded settlements were found during surveys, more likely to be carried on accessible land and more successful on ploughed soils. The other parameters also show some correlation, but this is generally close to the confidence interval. The lower impact of cities at very short distance (approaching 1, i.e. inside the city) could indicate that settlements are under-represented there, possibly due to their destruction during urban expansion in times when archaeological investigations were not routine. It is possible that if we remove larger cities from the analysis the result will be similar to Bevan (2012).

In order to have a more homogeneous study area for the subsequent analysis, only the settlements found on farmland were considered, assuming a uniform bias. The results are instructive, showing that the background influence is very strong. Nevertheless, we can observe that settlements occur more frequently than expected at elevation between 30 and 60 m (Fig. 3c). This is a good result if we compare the same analysis carried on the entire study area. In terms of distance from the sea or rivers, the behaviour shows slight preference for proximity to water sources (Fig. 3d-e).

3.2 *Funnel Beaker megaliths*

As for the settlements, farmland seems to have a strong impact in the distribution of these monuments, with higher than expected numbers in it or its proximity and less at increasing distances. A very similar pattern can be observed also for cities and infrastructures[2] (Fig. 4a-b). In general,

[2] Inside urban areas (values tending to 1) they are not found more frequently than expected by chance alone.

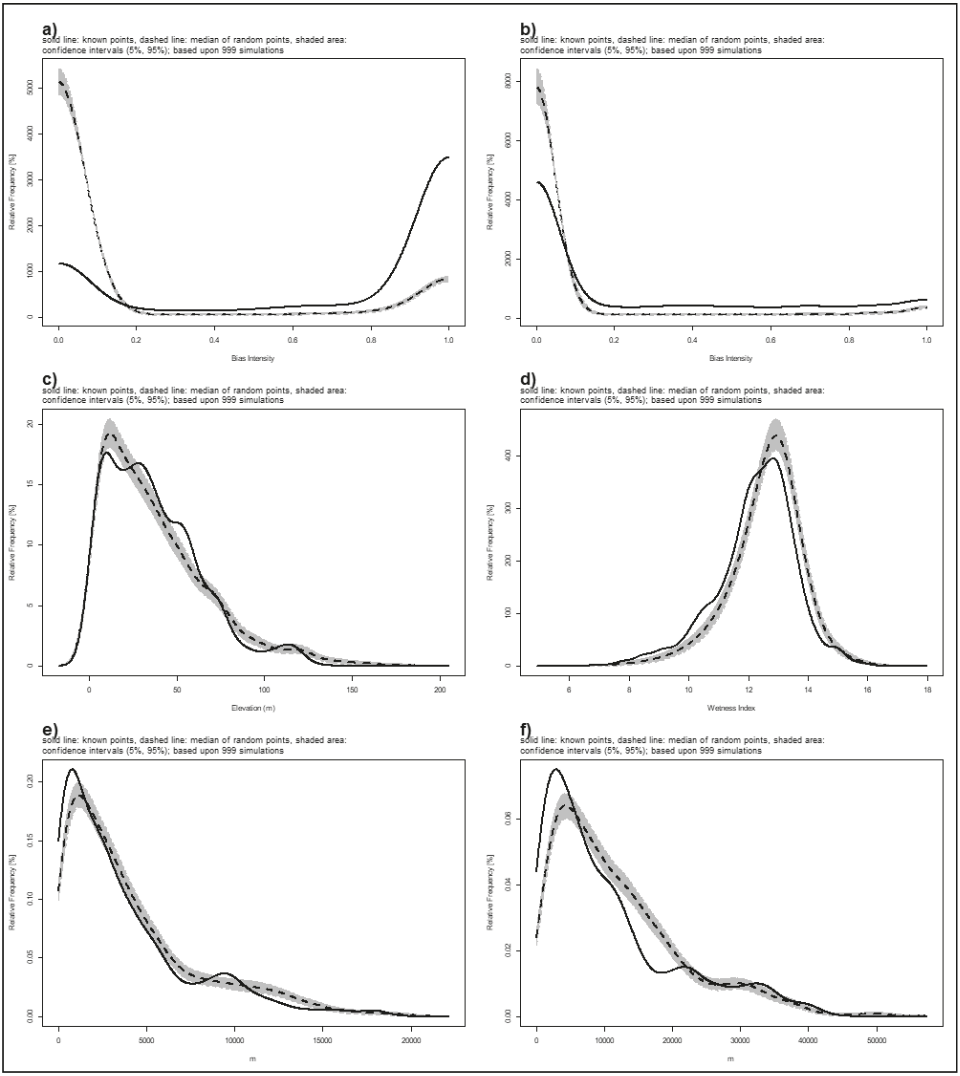

Fig. 3 – Settlement frequency plots: a) farmland as covariate; b) infrastructures as covariates; c) elevation; d) SAGA Wetness Index; e) distance from rivers and from the sea as covariates using the reduced study area as background.

their pattern does not differ greatly from what we observed for settlements. However, if we reduce the study area to the farmland we observe that proximity to the coast becomes a very good predictor, with a peak of site frequency at around 5 km and a higher presence of monuments compared

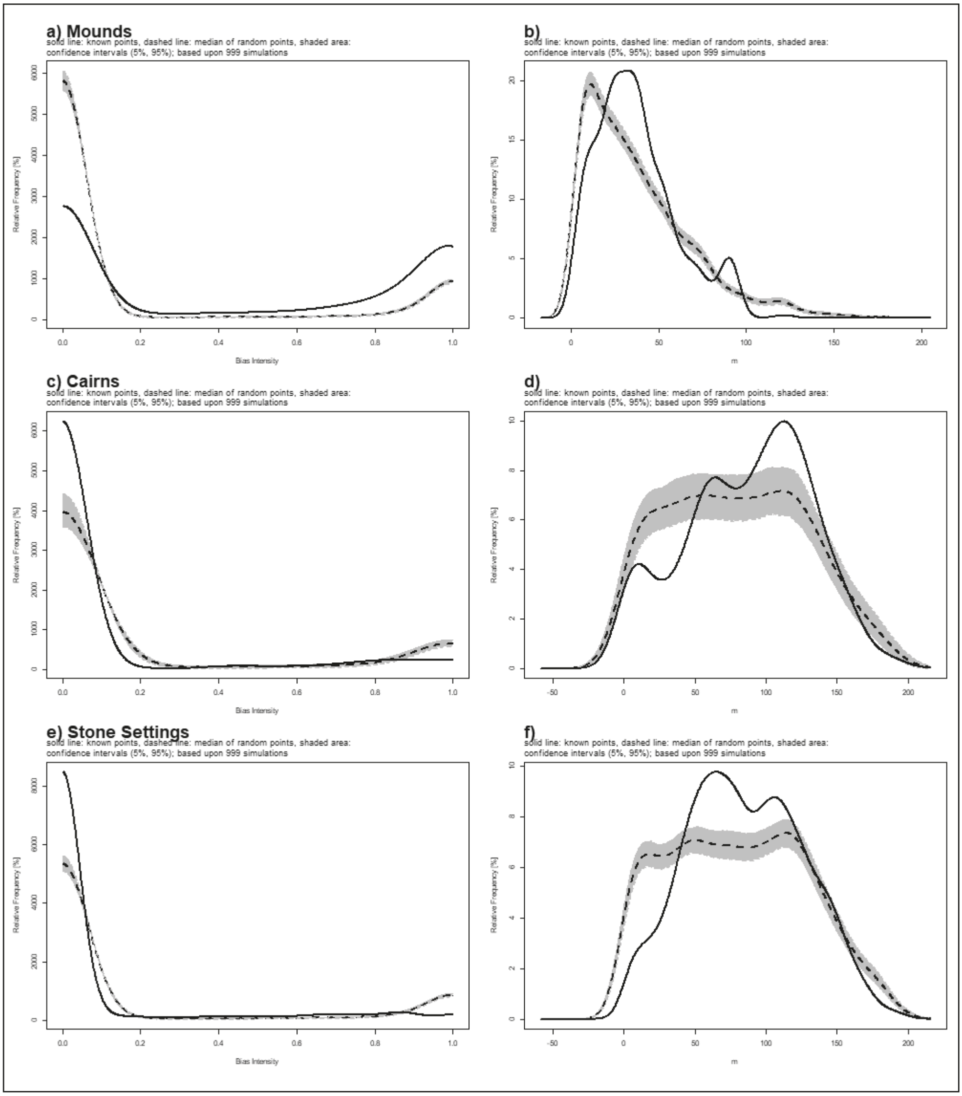

Fig. 4 – Megalith frequency plots: a) farmland and b) infrastructures as covariates; c) elevation and d) distance from the sea as covariates using the reduced study area as background.

to a completely random distribution (Fig. 4d). Elevation, which was a good predictor also before sub-setting the area, does not change, with more sites than expected by chance alone below 40 m asl ca. and fewer above 50 m asl (Fig. 4c).

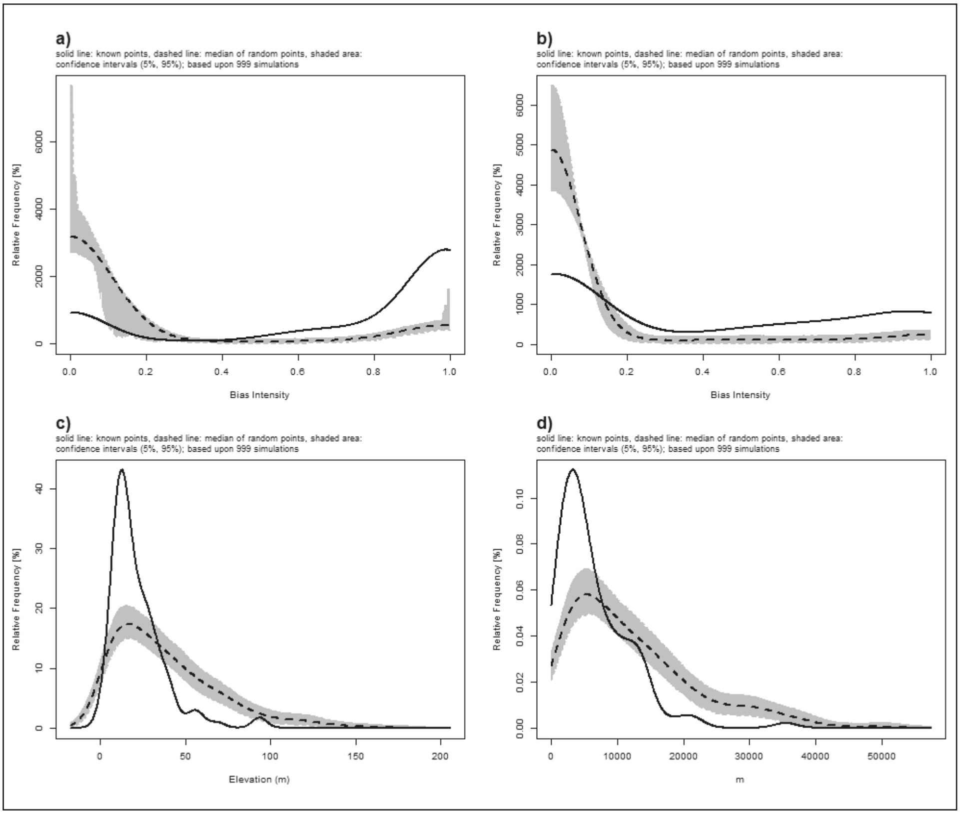

Fig. 5 – BA monuments frequency plots. Mounds: a) farmland as covariate; b) elevation (restricted area); cairns: c) farmland as covariate; d) elevation (restricted area). Stone settings: e) farmland as covariate; f) elevation (restricted area).

3.3 *Mounds, cairns and stone settings*

Mounds have a strong bias due to modern farmland but they also show a strong positive correlation with elevation. In fact, burial mounds tend to occur at low-intermediate elevation, with a peak around 40-50 m asl ca., while the farmland peaks at elevations closer to the sea level (Fig. 5a-b). In addition, mounds tend to occur more likely near the sea (peak below 5 km, and mostly below 10 km) and less than expected by chance alone at higher distances.

Cairns and stone settings tend to appear outside farmland, which is therefore excluded from the subsequent analysis. Again, elevation seems to be a very good predictor, with more stone settings than expected by chance alone between 50-100 m and cairns at 100-150 m and fewer than expected monuments below 50 m asl (Fig. 5c-f). In addition, cairns seem to occur with more frequency at higher distances from the coastline (> 35 km) and on drier soils (their Wetness Index, WI, is substantially lower than the surroundings). Stone settings have a similar behaviour in terms of WI, while they have a more composite behaviour in terms of distance from the sea, with two distinct peaks, one below 10 km and the other at 30-50 km. It is not possible to exclude that agricultural field clearances destroyed many of these monuments at different locations, but their distribution pattern seems to be complementary to the one of mounds, reinforcing an archaeological interpretation of the observed pattern. In addition, their behaviour does not substantially change when we use the entire study area or a subset.

This is an important result because it diverges from Neolithic monuments and settlements, giving us insights in different population patterns between the EN-MN and the Bronze Age.

4. Conclusions

Although a strong bias seems evident for many classes of sites, their impact is variable. When it is very strong, as for settlements, it is more challenging to suggest a reliable model but some hypothesis are nonetheless possible, given the fact that the distribution of megalithic burials partially reflects settlement pattern. However, only a more sophisticated model (on which I am currently working) could shed more light over population patterns during the Stone Age. This is not the case for the Bronze Age. In this case, the different classes of monuments have a complementary behaviour, strengthening our confidence in the model and reducing uncertainty already at this stage. Certainly, this paper only represents the first step for a more thorough analysis of prehistoric population patterns, which requires the inclusion of more variables and an organic combination into a more sophisticated environmental and social model.

Giacomo Bilotti
University of Kiel
gbilotti@sfb1266.uni-kiel.de

Acknowledgments

This paper is the product of my PhD research, supported by the CRC1266 "Scales of Transformations" and the German research Foundation (DFG) under grant number 2901391021-SFB 1266. I wish to thank my colleague Roberto Ragno for his indications and comments.

REFERENCES

Artursson M. 2005, *Byggnadstradition och bebyggelsestruktur under senneolitikum och bronsalder. Västra Skane i ett skandinaviskt* perspektiv, Lund, Riksantikvarieämbetet.

Artursson M. (ed.) 2007, *Vägar till Vætland: en bronsåldersbygd i nordöstra Skåne 2300-500 f. Kr*, Lund, Riksantikvarieämbetet.

Bevan A. 2012, *Spatial methods for analysing large-scale artefact inventories*, «Antiquity», 86, 492-506 (https://doi.org/10.1017/S0003598X0006289X).

Blank M. 2021, *Mobility, Subsistence and Mortuary Practices. An Interdisciplinary Study of Neolithic and Early Bronze Age Megalithic Populations of Southwestern Sweden*, Doctoral thesis, University of Gothenburg.

Bocinsky K. 2017, *Landscape-based hypothesis testing in R*, in B. Marwick (ed.), *How To Do Archaeological Science Using R* (https://benmarwick.github.io/How-To-Do-Archaeological-Science-Using-R/).

Conrad O., Bechtel B., Bock M., Dietrich H., Fischer E., Gerlitz L., Wehberg J., Wichmann V., Böhner J. 2015, *System for Automated Geoscientific Analyses (SAGA) v. 2.1.4*, «Geoscientific Model Development», 8 (https://doi.org/10.5194/gmd-8-1991-2015).

Günther G., Kempf M., Knitter D. 2022, *Rbias. R Package version 0.1* (https://doi.org/10.5281/zenodo.7071418).

Hadevik C., Steineke M. (eds.) 2009, *Tematisk rapportering av Citytunnelprojektet, Rapport*, Malmö museer, Arkeologienheten.

Knitter D., Hamer W. 2022, *Fuzzy Landscapes: Fuzzy Analyses with a Focus on Raster Data*, R package version 1.0 (https://doi.org/10.5281/zenodo.1747005).

Larsson L., Callmer J., Stjernquist B. (eds.) 1992, *The Archaeology of the Cultural Landscape: Field Work and Research in a South Swedish Rural Region*, Acta Archaeologica Lundensia, Stockholm, Almqvist&Wiksell.

Nord J. 2009, *Changing Landscapes and Persistent Places: An Exploration of the Bjäre Peninsula*, Acta Archaeologica Lundensia, Lund, University Lund.

Pawley S. 2022, *Rsagacmd: Linking R with the open-source SAGA-GIS software*, R package version 0.2.0 (https://stevenpawley.github.io/Rsagacmd/).

Persson P., Sjögren K.-G. 1995, *Radiocarbon and the chronology of Scandinavian megalithic graves*, «Journal of European Archaeology», 3, 59-88 (https://doi.org/10.1179/096576695800703694).

R Core Team 2022, *R: A Language and Environment for Statistical Computing*, Wien, R Foundation for Statistical Computing.

Sjögren K.-G. 2003, *Mångfalldige uhrminnes grafvar: megalitgravar och samhälle i Västsverige*, Göteborg, Göteborgs universitet, Institutionen för arkeologi.

Sjögren K.-G. (ed.) 2006, *Ecology and Economy in Stone Age and Bronze Age Scania (Skånska spar, Arkeologi Langs Vastkusbanan)*, Lund, National Heritage Board.

ABSTRACT

Southern Sweden, and especially the area around Malmo in southwestern Scania, is perhaps one of the most archaeologically investigated areas in the world. Our knowledge of the local Prehistory has greatly increased in the past decades although it is also the product of centuries of agricultural practices, urban expansion and a relatively early (18th-19th c.) interest for prehistoric monuments (e.g. burial mounds and megaliths). However, despite the deluging amount of available information (over 50,000 ancient sites recorded in Scania), their distribution is not homogeneous and archaeologists are restlessly trying to explain this pattern and its underlying causes. In addition, post-depositional factors

(infrastructure works, agricultural practices, etc.) heavily affect site distribution and preservation, blurring the global interpretation. The aim of this paper is to reduce the impact of post-depositional factors on our interpretations on site distribution. In addition, the results can be used as a starting point for further and more elaborate analyses (spatial statistics and simulations). All the models presented here were computed in a reproducible way, relying on FOSS and open data only, in order to allow anyone interested to replicate the model and adapt it to their own purposes and study regions.