Journal of Archaeological Science 70 (2016) 91–101

Contents lists available at ScienceDirect

Journal of Archaeological Science

journal homepage: http://www.elsevier.com/locate/jas

Geostatistical modelling of chemical residues on archaeological floors in the presence of barriers

Joan Negre [a, b, *], Facundo Muñoz [c], Carla Lancelotti [d]

[a] Centro Austral de Investigaciones Científicas, Consejo Nacional de Investigaciones Científicas y Técnicas, B. Houssay, 200 E-9410 Ushuaia, Tierra del Fuego, Argentina

[b] Laboratori d'Arqueologia Quantitativa, Departament de Prehistòria, Universitat Autònoma de Barcelona, Edifici B — Fac. Filosofia i Llletres, 08193 Bellaterra, Barcelona, Spain

[c] UR0588 Unité Amélioration Génétique et Physiologie Forestières, INRA Val de Loire Orléans, 2163 Avenue de la Pomme de Pin, CS 40001 Ardon, Orléans, Cedex 2 F-45075, France

[d] CaSEs Research Group Department of Humanities, Universitat Pompeu Fabra, C/Trias Fargas 25-27, 08005 Barcelona, Spain

A R T I C L E I N F O

Article history:
Received 6 November 2015
Received in revised form
8 March 2016
Accepted 15 April 2016

Keywords:
Non-euclidean geostatistics
Cost-based distance
Spatial heterogeneity
Kriging
Housefloors

A B S T R A C T

Maps representing the distribution of chemical residues over anthropogenic floors are the main diagnostic tools used by archaeologists for addressing the identification of geochemical signatures of past actions. Geostatistics allows producing these maps from a sample of locations by modelling the spatial autocorrelation structure of these kind of phenomena. However, the homogeneity of the prediction regions is a strong assumption in the model. The presence of barriers, such as the inner walls of domestic units, introduces discontinuities in prediction areas. In this paper, we investigate how to incorporate information of a geographical nature into the process of geostatistical prediction. We propose the use of cost-based distances to quantify the correlation between locations, a solution which has proved to be a practical alternative approach for archaeological intrasite analysis. The cost-based approach produces more reliable results avoiding the unrealistic assumption of the homogeneity of the study area. As a working example, a case study of the distribution of two specific chemical signatures in domestic floors is presented within a controlled ethnographical context in Northern Gujarat (India). On a broad disciplinary scale, the benefits of using our approach include improved estimates in regions with complex geometry and lower uncertainty in the kriging predictions.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

The analysis of chemical soil composition in archaeological domestic floors is becoming increasingly considered as an important topic for historical research. Mapping the distribution of certain combinations of chemical elements allows us to understand the activities that were developed in the areas under study. This is based on the idea that different social actions of production, consumption or distribution are the cause of the variations observed in the material consequences detected through fieldwork. In this case, the variability of chemical soil composition is considered to be a reliable marker in order to detect, identify and analyse different activities in domestic contexts (Rondelli et al., 2014; Salisbury, 2013; Middleton et al., 2010). The model connecting the concentration of particular residues (proxies) with the specific activities inferred from different information sources (archaeological experimentation, ethnoarchaeological reasoning, etc.) is defined as an anthropic activity marker. Nonetheless, the reading of chemical differential concentrations in archaeological floors is not exempt of critical reflections about its limitations (Lancelotti and Madella, 2012; Vyncke et al., 2011; Dore and López Varela, 2010; Wells, 2010; Terry et al., 2004).

Geostatistical methods are increasingly used to model the results of geochemical analyses, hence facilitating the interpretation. These techniques provide a set of statistical tools specifically designed for spatial problems, in which predictions of missing values are required over a region of interest where some observations have been taken. Predictions are based on an underlying

* Corresponding author. Centro Austral de Investigaciones Científicas, Consejo Nacional de Investigaciones Científicas y Técnicas, B. Houssay, 200 E-9410 Ushuaia, Tierra del Fuego, Argentina.

E-mail addresses: negreperez@gmail.com (J. Negre), facundo.munoz@orleans.inra.fr (F. Muñoz), carla.lancelotti@upf.edu (C. Lancelotti).

http://dx.doi.org/10.1016/j.jas.2016.04.016
0305-4403/© 2016 Elsevier Ltd. All rights reserved.

statistical model that can take additional information into account as explanatory variables. In addition, the prediction error can be estimated based on propagation of uncertainty. One of the main limits of the use of geostatistical methods for this purpose is the assumption of a homogeneous, unrestricted space of analysis (López-Quílez and Muñoz, 2009). This premise fails when we consider the spatial demarcations and topography that affect the distribution of our phenomena over the study region. The analysis of chemical residues in a domestic unit floor is a classic example of this kind of situation, where the walls of the house affect the distribution of the chemical elements.

In this work, we propose to overcome this problem by using cost-based distances to quantify the correlation between sampling locations (López-Quílez and Muñoz, 2009). Thus, we present a case study on the distribution of chemical elements in domestic floors within a controlled ethnographical context in North Gujarat (India) (Rondelli et al., 2014). This paper explores the relative spatial variability of residues, taking into consideration spatial demarcations, to provide a method for the detection and interpretation of specific areas of activity. Our technique, therefore, can substantially improve the identification of both clustering patterns and different processes of floor maintenance and postdepositional dynamics considered as background noise (Rondelli et al., 2014; Pecci et al., 2013; Barba, 2007; Lloyd and Atkinson, 2004).

2. On the use of non-euclidean distances in geostatistics

Geostatistics is a branch of statistics that encompasses the techniques that apply to geographical analysis. We owe its origins to the works of Krige (1951) and Matheron (1963) in the central decades of the twentieth century. There are several applications of geostatistical methods in a wide range of disciplines that share the problem of modelling a stochastic process over a continuous spatial region from a partial group of observations. This process of interpolation is commonly assumed to be Gaussian, isotropic and intrinsically stationary (Cressie, 1993). Geostatistical modelling is based on the principle of spatial dependence, which states that near events are more related than distant ones. Nevertheless, what does *near* mean and how do we calculate it?

Interpolation techniques assume that the correlation between the elements of a group of observations is a function of the Euclidean distance between them. In other words, stationarity is often accepted to mean that the spatial point process has constant intensity and uniform correlation depending only on the lag vector between pairs of points (Møller and Toftaker, 2012). Considering the inherent irregularity of geographical terrain, either the presence of barriers or the difficulty to cross a region are presented as a major problem for this technical requirement. Imagine two locations at a given (Euclidean) distance such that they are significatively correlated, because of underlying relevant factors affecting both of them. Now put a barrier between them that blocks or absorbs the effect of the underlying factors. This obviously pulls the correlation down. Therefore, when some kind of barriers exist, the correlation depends on something other than the simple euclidean distance between two points, which therefore cannot account for the correlation by itself.

There are more general situations where barriers are not absolute, but regions that are either harder or easier to cross depending on a series of relevant factors. For example, microtopography of the study region, soil texture and composition or the relationship of different anthropic activities between them are important partial restrictions that should be taken into consideration. All kind of heterogeneities in the surface in which chemical elements spread might be modelled with a cost surface, representing how hard it is to cross a given portion of area. And accordingly, the correlation between two locations should be associated with the minimum-cost path connecting them. A cost surface presenting every relevant factor affecting correlation is, therefore, an efficient tool to deal with the distribution of chemical signatures in all kind of surfaces. In this framework, the standard geostatistical assumptions of a homogeneous region is a particular case where the Cost surface is a constant 1-valued surface. Therefore, the minimum-cost path between two given locations is the straight line connecting them; hence, the Cost-Based distance equals the Euclidean distance. Also, the more general situation with barriers in the working region is another particular case where the Cost surface takes the value 1 over non-barrier areas and the value ∞ over barrier areas, therefore the Cost-Based distance equals the minimum distance needing to be traveled without crossing any barriers, as was required (López-Quílez and Muñoz, 2009).

Methodologically, the first step in classical geostatistical processing is to fit the data and its empirical semivariogram function to a known parametric model. There is a variety of methods for estimating this correlation (Cressie, 1993). Our approach here is to use maximum likelihood methods that fit the mean value and the parameters of the semivariogram function. Once fitted, the main analytical interest lies in obtaining spatial prediction. Kriging assumes that the distance or direction between sample points reflects a spatial correlation that can be used to explain variation in the surface. This technique is one of the most used approaches to this problem, in which a weighted average of the sample values is applied to generate the prediction. That is, sample points near the prediction's location are given larger weights than those far away. The general formula for the interpolator is formed as a weighted sum of the data:

$$\widehat{Z}(s_0) = \sum_{i=1}^{N} \lambda_i Z(s_i)$$

where $Z(s_i)$ is the measured value at the i th location, λ_i an unknown weight for the measured value at the i th location, s_0 the prediction location and N the number of measured values.

Kriging determines these weights calculating them according to the value of the semivariogram, which is a function of the Euclidean distance (López-Quílez and Muñoz, 2009). That seems to incur into the above mentioned error of assuming the validity of the spatial homogeneity premise. Thus, in certain cases, alternative measurements to Euclidean metrics, such as cost-based or pseudo-Euclidean ones, represent the distance argument r of the semivariogram function more naturally.

2.1. *Cost-based distances*

Alternative measures to Euclidean distances have been largely tested in several disciplines. A multidimensional-scaled reconfiguration of the spatial distribution has proved to be very useful in some cases, allowing to create a pseudo-Euclidean framework on which the analysis can be performed (Løland and Høst, 2003; Negre, 2015). A fast Fourier Transform has also been explored for integrating moving-average functions that may be used to generate a large class of valid, flexible variogram models. This transform allows to both compute the cross-variogram on a set of discrete lags and to interpolate it for any continuous lag (Ver Hoef et al., 2004). In this same direction, recent works also propose the use of Riemannian metrics associated to cost-based distances and Banach algebra using Kuratowski immersion (Muñoz, 2012: 118). For its relative simple implementation, the use of cost-based distances directly into the covariance matrix of the kriging, has proved to be a practical and competitive option for our research topic.

Download English Version:

https://daneshyari.com/en/article/7441395

Download Persian Version:

https://daneshyari.com/article/7441395

Daneshyari.com