**WebApp Texts**

**Name:**

Robusta – An Application for Quantifying Uncertainties in Spatial Analysis in Archaeology

**Introduction:**

Based on the paper by Herrera Malatesta and de Valeriola (2024), this web app seeks to allow interested researchers, specialized and non-specialized in computational archaeology, to reproduce the paper’s results and, most importantly, to be able to quantify the uncertainty on their own spatial dataset. Following the aim of the paper of contributing to the growing literature on data incompleteness and uncertainty quantification in archaeology this web app is the direct expression of formalizing practices and procedures into a framework.

Robusta aims to help colleagues who desire to better understand their models by quantifying their uncertainties. The framework we propose here focuses on assessing the robustness and uncertainties of the conclusions drawn from applying point pattern analysis to, mostly, non-systematic regional data in archaeology. To achieve this aim, we have articulated the discussion on the reconstruction of past landscapes using computational methods around three key aspects: the use of point pattern analysis (PPA) in archaeology, the quantification of uncertainties, and the consideration of robustness.

The framework that will be presented is designed to aid archaeologists -working with datasets that are known to contain sources of uncertainty- to be able to apply spatial statistical methods and achieve a higher understanding of the uncertainties of the resulting models.

**Method:**

The framework we proposed in our 2024 paper, and that is used in this web app follows three simple steps of what we called the ‘robustness assessing framework.’ The first step is what we defined as the “observable” which is a point clustering metric and the quantity whose changes we measure when considering deviations. In this case, this is the model resulting from a Pair Correlation Function (PCF) with a Monte Carlo simulation envelope based on the 100% of the dataset used as input (the dataset should be load as a ‘.shp’ file and consists of a series of georeferenced points, in addition a georeferenced polygon of the research area needs to be also loaded). The web app will provide the result of the ‘observable’ and this will be used as a reference value for the second step.

The second step is ‘the experiment’ where regular intervals of data will be sample from the loaded dataset. More precisely, the app will deduct 10%, 20%, 30%, 40%, and 50% of the database’s sites and perform again the PCF with a Monte Carlo simulation envelope for each of the sampled groups. The resulting models are what we have called the ‘robustness scenarios.’ Note that in our paper we used two sampling methods, one using a uniform distribution and another one using an inhomogeneous distribution. To make the web app works optimally, we decided only to provide here the sampling via the uniform distribution. If a more experienced researcher would like to access the code to also consider the an inhomogeneous distribution, please check our code here (link: I do not know if we will paste here the link to the OSF or Niels GitHub).

The third step consist of the ‘comparison tools,’ which are methods to assess the frequencies and interval midpoint densities. The comparison tools is the step that will allow the analyst to assess the robustness and quantify the uncertainty of the spatial models created based on their dataset. More information about this step on the “Results” section.

**Results:**

The results will provide two figures that can be downloaded as a PDF. The figure of the first comparison tool will present the percentages of sites that are kept in each robustness scenarios against the percentage of robustness scenarios in which the conclusion is similar to the observable. This figure provides a direct percentage of the probability that, by extracting a particular percentage of the dataset, the results of the robustness scenario will be similar to the ‘observable.’ The second comparison tool goes deeper in the understanding of the patterns and provide an insight of what can be further observed from the pattern in term of its clustering.

At this point, it is important to clarify that the web app is only analyzing the statistically significant patterns on the PCF, and not the regular ones. Again, if an analysis would like to include this in their analysis, they need to modify the source code used for this web app.

**Creators:**

This web app was created by Niels Aalund Krogsgaard ([nakro@cas.au.dk](mailto:nakro@cas.au.dk)) from the Centre for Humanities Computing (Aarhus University) based on the research and codes by Eduardo Herrera Malatesta and Sébastien de Valeriola.

**Contact:**

For more information about this web app or contact to its creators please contact

Eduardo Herrera Malatesta - [ehmalatesta@yahoo.com](mailto:ehmalatesta@yahoo.com)

Sébastien de Valeriola - [sebastien.de.valeriola@ulb.be](mailto:sebastien.de.valeriola@ulb.be)

Ross Deans Kristensen-McLachlan, Centre for Humanities Computing (Aarhus University) - [rdkm@cas.au.dk](mailto:rdkm@cas.au.dk)

**Funding Information:**

This research has received funding from Horizon Europe, HORIZON-MSCA-2021-PF-01-01, Marie Skłodowska Curie Action, Grant agreement n˚ 101062882. Granted to Eduardo Herrera Malatesta.