Dear,

We have uploaded two versions of the revision – a clean copy without mark up and a tracked changes copy.
Reviewer's comments are in blue bold.

---

by Reuven Yeshurun, 09 Apr 2023 14:57
Manuscript: **https://zenodo.org/record/7674699** version 1

Minor revision

This is a high-quality preprint that should be useful for a wide archaeological community. Based on the three obtained reviews and my own reading, the preprint should be ready for recommendation following a minor revision. Please pay attention to the reviewers' suggestions and in particular to two points:

1. The preprint demonstrates the application of SEAHORS by using data from a specific Paleolithic site (Cassenade, France). Would the authors consider adding a few sentences on what was achieved in that specific case-study by using SEAHORS? This is certainly not essential for presenting SEAHORS, but may assist the readers who are not familiar with the published studies of the cave to appreciate the contribution of this application in a specific archaeological study (see also R1's recommendation).

2. Overall, the English is very good but the paper could use another round of editing. In the appended file, I marked several grammatical mistakes and phrasings that can be improved using the track changes tool (see also R3's suggestions).

Download recommender's annotations

Thank you very much. We have integrated most of English corrections and added few others.

Reviews

Reviewed by Frédéric Santos, 13 Mar 2023 14:49
The paper is very well written and clear, and the R shiny application is a useful contribution to the field. The set of features offered by the app (ranging from data wrangling and cleaning to adequate and interactive visualizations) is exhaustive and really well-thought. It should also be noted that the app is well documented, so that its use will be straightforward for everyone, and especially non-programmers.

I have tested the (real life) example files provided by the authors, and have been able to reproduce easily the expected outputs presented in the paper.

Overall, I have very few comments, and consider that this article is already almost ready for public dissemination.

Minor suggestions:

- Include a short note about the GPL license in the main text of the manuscript.

  We have added the following sentence: "This application is distributed under the General Public License (GPL) license v3.0.**"**

- Would the authors consider (maybe later) turning the current R script into a proper R package for local use? I acknowledge that this represents a substantial amount of additional work (especially because an R package and the scripts required for the shinyapps platforms would require a separate maintainance), and maybe not a huge improvement for the final user in practice. However, if the app evolves quite regularly, it would make easier the handling of different versions, and the citation of a precise version in research articles. (But this is really only a suggestion and absolutely not a request!)

  We have transformed the script as an R package. The previous version of script is now available at https://github.com/AurelienRoyer/SEAHORS-raw-script.

  The R package is available on the CRAN (https://cran.r-project.org/package=SEAHORS) and on Github (https://github.com/AurelienRoyer/SEAHORS). It is now the version 1.7. We have also modified the text about the installation and the launch of the app.

Minor spelling remarks:

- Line 98: the sentence "it allows to set up the four..." might be a grammar mistake, and in this case, shoule be rephrased.

  We have corrected the sentence as follows: "It allows the user to set the four basic data to import**…"**

- Several instances in the manuscript: replace "a R shiny app" by "an R shiny app", which is more common and sounds better.

  Corrected

Additionnally, other minor typos and suggestions for the app itself (both in the user interface and the markdown documentation file) have been submitted as pull requests in the Github repository online.

All the commits have been accepted on Github.

Thanks to the authors for their work!

Thank you very much for your review.

Reviewed by Jacqueline Meier, 05 Apr 2023 01:49

This article presents a valuable new tool for spatial analysis that is specifically designed for archaeological research. The authors have developed a useful avenue for analyzing spatial relationships among archaeological remains through a R Shiny, user-friendly web application. I also commend this effort and the related archeoviz application that they reference in this paper. These web applications have great potential to speed up traditional spatial analysis and the visualization of refits. They also make it easier to share clear visualizations of the results that other users can filter and explore.

One important point made by the authors is that a clean database is essential to use this application. While it will detect errors in the uniqueness of the IDs and make new concatenated columns of the data attributes, there are potential issues with using SEAHORS to interpret associations and densities of inaccurately entered data points that may be overlooked by the user. Pairing this application with "tidyr" or perhaps a different R package might ensure that the users of SEAHORS begin with clean datasets. Much like the data importer in SEAHORS, a specific R package for cleaning other common errors in archaeological datasets would perhaps be another great R Shiny contribution, much like SEAHORS, if it does not already exist.

The cleanliness of the database is one of the major aspects to be used then for spatial distribution or others analyses. For SEAHORS, we limited ourselves to a few cleaning and checking functionalities, such as the uniqueness, which is mandatory to concatenate several database. But, for the control of the data and its quality control, this is a larger topic that involves others database issues. We can indeed think that script(s) could be developed to control at minima the data and basic errors, but we also think that should be done on a case by case basis before using SEAHORS or others similar applications. There are so many different possible types of errors, depending of many different factors as how the data were acquired, by who …, that we also think that specific human controls are absolutely necessary.

I hope that the authors will continue to build this web application to incorporate further types of spatial data. As many archaeological sites record the x-y-z coordinates of some finds and only the excavation square/unit for other finds, such as those found in the screen or after floatation, it would be helpful if SEAHORS could combine the more and less specific locations of finds. Perhaps this could be achieved through a combination of imported point and polygon (entire excavation square) data.

As one of us (A.R.) mainly analyses remains from splits (i.e. small vertebrates), we would like to incorporate this type of analysis of material that is not coordinated. We are keeping this idea of spatial analysis of material from both types of provenances for future developments of the application. However, at present, we do not know how, because 1) unlike coordinated materials, the recording of this type of material is highly variable from excavation to excavation, using either a single point, all the square data, or only Z information, leading to some difficulties to reconstruct the polygon; 2) there are many types of graphical options that could be included such as pie charts, bubbles or random coordinated points.

Some other R Shiny web applications allow users to try out the tool with built-in demo files. While the demonstration video of SEAHORS is helpful and the github page includes links to example data files, embedding the demo dataset directly in the application would likely allow new users to try it out with greater ease as they watch the video walk-through. In my opinion,

making it more directly available in this manner would make users more likely to use SEAHORS to explore their own data. A good example of one with directly available example data, the gmshiny web application includes test data that is not perfect and a walk-through video that shows users how to clean up the data (similarly to the SEAHORS video).

We have added to the first slide (overview slide) a button, which allows to directly download from github the example of Cassenade coordinates file and to load it. Furthermore, we have also put also the file named "example dataset" with the Cassenade data on Github, so that is accessible.

Another point that could draw more researchers to use this tool would be to expand the video to apply SEAHORS to test an example hypothesis related to the example dataset. A highly detailed, published dataset is used to demonstrate the capabilities of the tool in the video demo. This demo could go one step further and demonstrate how it can test one of the hypotheses presented in the Discamps et al. 2019 paper that is the source of the example data. While the methods are clearly presented for broad applications by researchers with different spatial questions, the authors could present how the tool addresses at least one question to fully convince future researchers to use it and give them ideas of how to do so. This would also make the tool applications more obvious to archaeology students who may use it to begin to think about how to explore spatial questions.

We have added the next following paragraph just before the conclusion. This paragraph summarises the main result obtained by Discamps et al on Cassenadre by using the projections of coordinated material:

"Example of spatial distribution at Cassenade

As interpreted by Discamps et al. (2019), the use of spatial distribution involving field observations, faunal and lithic data was crucial to understand Cassenade sequence and disentangling the different archaeological layers. Two distinct assemblages were identified in a single lithostratigraphic layer on the basis of projections of the coordinated material. These assemblages were impossible to distinguish during the excavation. Lithic artefacts and charcoal are much more abundant in the upper part of the deposits, while faunal remains were found throughout all the stratigraphy. The concentration of lithic artifacts visible in the upper assemblage shows a steep slope towards the cave entrance in the north, which is reflected in the strongly oriented distribution pattern of refitted artifacts. Cave bear remains and bones with a brown patina are more abundant in the lower assemblage, whereas anthropic modifications of bones (cut-marks, percussion marks, burnt bones) are concentrated in the upper part of the deposits. A nearly sterile band between upper and lower assemblages is observable in some parts of the site, but not everywhere. The exploration of the spatial distribution of the archaeological materials allowed, in the case of Cassenade, to correctly delimit two distinct assemblages, to characterize them and to discuss the Châtelperronian occupations of the site. All the projections used for this study, originally performed with QGIS (Discamps et al., 2019), can be reproduced using SEAHORS in a much more practical and faster way.

Overall, the authors present an impressive web application that will extend the spatial analysis of archaeological finds to many researchers who may otherwise feel prohibited by the task of learning to use R or the high costs of complex mapping programs. I will look forward to using SEAHORS and to the future web applications for archaeology by these authors.

Thank you for your comments and suggestions.

Reviewed by Maayan Lev, 23 Mar 2023 12:08

The paper presents a clear and concise description of the SEAHORS application. The application overview is thorough and well-detailed, includes all the relevant information, and enables easy use of the application.

While some basic knowledge of R programming software is needed, SEAHORS provides a useful tool for intra-site spatial analysis for researchers lacking expertise in other software (e.g., ArchGIS) and enables the easy production of 2D and 3D plots. My personal attempt to use the application with my data was successful. If plausible further graphical options would be appreciated (axis labels font and size, legend properties, etc.).

Thank you for your comments and suggestions. We have added a fifth sub-panel in the sidebar panel, which offers the possibility to change the size of axis labels and tick marks, change the name of axis legends, change the space between tick marks, and add a additional *minor* grid (i.e. minor tick breaks). For density plot and "simple 2D plot", it is possible to select one of the basic ggplot theme. All these modifications have been added in the article. We cannot change easily the policy fonts, so no modification has been done for this point.

Besides all these changes, we have added new types of plots: "2d plot" tab is divided into two subtabs: one named 'advanced 2 plot" and the second named "simple 2D plot". The second is an addition, which is obtained using *ggplot2* package. This new subtab is quicker and allows to create and download plots with a 1:1 ratio (which is easily modified by changing the value of the ratio just below the figure). Nonetheless, this plot do not allow to show information while hovering over a point. A similar mode is displayed with 2D slice plot, just by unchecking the box "Advanced plot".

Furthermore we have also made a few English minor edits, and modify the acknowledgements section.