Dear Dr. Cook,

Thank you for giving us the opportunity to improve and resubmit our paper. We appreciate your time and consideration of our submission.

Reviewer,

Thank you for all your comments and constructive feedback. We responded to each of your comments in this document and addressed these accordingly through meticulous changes in the manuscript. We think that your revision has substantially improved our paper. We greatly appreciate your time, and we look forward to your response regarding the changes that we made.

Original reviewer comments in bold, our response follows in plain text.

Overview:

The authors describe the rassta package, which provides several functionalities to work with landscape datasets. The topic is of relevance and the package could be useful for a wide range of R users coming from different areas of expertise. Therefore, I think that the paper can be considered for publication once the authors consider/address the comments below.

Article:

The authors should emphasize sometimes that a function belongs to the rassta package. For instance, when the som_gap() and som_pam() functions are mentioned, I was unsure if these were functions from rassta or the kohonen package.

We have rephrased the introductory paragraph for the functions som_gap() and som_pam() to prevent any confusion about which functions belong to the rassta package (page 4). In addition, please note that Figure 1 shows the name and category of each function in rassta.

I have had to load the package Rcpp to run som_gap(). If required, it should be included in the supplementary code, or even in the paper.

Thanks for pointing out this issue. We have included a note in the first paragraph of page 2 about the requirement of Rcpp in order to handle the classes used by rassta. Also, we have included a line of code in the manuscript and its associated R file to load Rcpp to the session.

I think it might be convenient to warn the reader about the computational cost of som_gap() and som_pam().

We agree. We have mentioned the computational cost of som_gap() for the corresponding example and briefly discussed how this cost increases in relation to the function's arguments. Please note that the function som_pam() should not require a significant amount of time since it only creates raster versions of som_gap() outputs. The new content can be found on the second paragraph of page 4.

The authors comment that for creating classification units the user might prefer to use some method other than som_gap or som_pam. How would this be 'adapted' to rassta objects?

Any set of classification units can be used with rassta if it can be represented through a raster layer. If this is the case, the user will only need to import the raster layer of classification units into R as a SpatRaster class. This ensures total compatibility between the object and the methods provided by rassta. We consider this an important point and we made sure to include it in the paper (page 5, second paragraph).

I think that the section 'Landscape stratification' needs some subsections to better make the distinction between classification and stratification.

We agree. We have divided the content of classification and stratification units into two subsections: Classification units (page 4) and Stratification units (page 5) for improved clarity.

I think that the section 'Metrics of landscape correspondence' might be unnecessary. Could the authors merge this section into the following related sections?

We have rewritten 'Metrics of landscape correspondence' as the main section, and 'Spatial signature of classification units' and 'Landscape Similarity to stratification units' as its subsections. This reorganization has improved the flow of the section and can be better noted in the HTML format of the manuscript.

The spatial signature concept needs greater clarification. Some specific formulas would be welcomed, in my opinion.

We have added a pseudocode to the manuscript that shows the computational process (including the formula for LOESS fitting) to calculate the spatial signature of a classification unit (Figure 8, page 7). We hope that the pseudocode and Figure 7 above it will allow a better understanding of the spatial signature concept.

The idea of spatial representativeness in the context of sampling could be also outlined.

We are not completely sure that we understand this point. Does the reviewer refer to a spatially-balanced coverage of samples (e.g., non-clustered observations)? Or to the representativeness of landscape/environmental conditions and response phenomena across geographic space? The latter directly relates to rassta's sampling approach since it allows the use of landscape similarity layers and stratification units to constrain the sampling process (i.e., stratified non-probability sampling). This stratified, non-probability sampling is based on the assumption that strata and corresponding similarity layers adequately represent the spatial variability of landscape conditions, and that these conditions exert control on response phenomena. Accordingly, the adequate representation of these conditions in the sampling process will result in a sample with similar statistical characteristics to the population of the response phenomena (see Figure 14, first and second paragraph of page 11). Please note that users also have the option to conduct a random stratified sampling with rassta.

The section 'Predictive modeling' is unclear to me (including Figure 15). I think that it needs greater clarification, as it does not seem like the typical modeling scheme. Besides, in the Introduction the authors state that 'The rassta package is not intended as a drop-in replacement for statistically-robust environmental modeling approaches'. Hence, is the function 'engine' necessary? Should users follow a more general modeling framework for prediction? Some comments in this regard might be convenient.

Thanks, this is a very important aspect and we have added more detail to the manuscript about the modeling process in regard to scenarios where its application is appropriate, similarity with other modeling frameworks, and its nature overall (last paragraph of page 12/first paragraph of page 13, and third paragraph of page 13). Overall, the predictive modeling with engine() is based on a weighted average for continuous responses and a 'weighted' mode for categorical responses. The weights are the landscape similarity values at a raster cell, and the actual values to average/select the mode from are the representative response values of the stratification units (with the units being represented by the landscape similarity layers; i.e., one similarity layer per unit). If a raster cell has a greater landscape similarity value to *x* stratification unit, the modeled response value at the raster cell will be more similar to the representative response value of *x*. Modeling with weighted averages/modal values across space has been widely applied in the geographic information systems (GIS) field and for modeling landscape attributes with limited data. The difference, and perhaps what is somewhat confusing about engine(), is that engine() allows to select for each cell a subset from the complete set of landscape

similarity layers. For instance, the user may only want to consider the 3 similarity layers (and associated representative response values) with the greatest values at the raster cell in question, rather than using the full set of similarity layers. We have made sure to expand more on these aspects in the manuscript, including relevant citations.

Regarding the need for engine as a modeling framework instead of using more general ones (e.g., statistical modeling/learning), we have added references to previous studies that have proposed the use of similarity layers and/or weighted averages for modeling environmental phenomena (last paragraph of page 12/first paragraph of page 13). In summary, these studies suggest applying modeling frameworks similar to that proposed by engine() in scenarios where observed/measured response data for model fitting is limited, initial modeling for survey design is necessary, expert knowledge and conceptual models are available, and ecologically sound representations of spatial variability are desired. Although rassta is not intended to replace general, quantitatively-robust statistical modeling, we believe that its use is appropriate in such scenarios, at least to derive geospatial information and initial models of spatial variability for their use to improve inference with more general, statistically-oriented modeling frameworks.

I find the notation XY for a point rather unusual. Is it necessary?

We agree. This notation is unusual in disciplines other than geospatial analysis, where XY is often used to denote positions in physical space (e.g., XY coordinates). This notation is not necessary and we have replaced it with the term *raster cell* for improved clarity and to avoid overloading notation. We think this replacement is more appropriate since most of the methods in rassta use the raster cell as the minimum unit of analysis/operation.

Package:

I think that the package is easy-to-use and well-documented.

Thanks.