

Editorial decision

Dear authors, we have received two reviews of your article, both very positive. However, the reviewers also posted some interesting comments that should be addressed before the manuscript can be accepted and recommended. We therefore strongly encourage you to include these suggestions in a new article and resubmit the manuscript. See the details of the reviews below. We look forward to the new version of the article.

We thank the reviewers and editor for the constructive and thoughtful comments. They have been very helpful in revising the manuscript, and we hope that the revised version addresses all concerns. We have lightly edited the comments by reviewers for clarity, by removing general comments that did not call for a response, as well as a the review checklist.

While working on this revision, we have also added several functionalities to the package, which are now similarly documented and covered in the supplementary notebooks.

Reviews

Reviewer 1

There's lots of good info in Elith and Leathwick (2009) but they missed how the work of the BIOCLIM group provided the basis for the most widely used SDM variables and data (for more see the following 2025 paper <https://doi.org/10.3390/earth6010012>). You might consider citing the 2014 BIOCLIM review (which has been cited almost 1000 times in Google Scholar) along with Elith and Leathwick. The 2014 review explains the variable/data issue and also cites the first SDM climate change papers which were also missed by E&L and are still important for reasons explained in the 2025 paper.

We thank the reveiwer for bringing these papers to our attention. We had read them, but forgot to cite them in this manuscript. This omission has been corrected.

You mention mean temp of the wettest quarter in your Figure 5. Though it may not be a problem here it's worth knowing that there are some problems with this surface -see open access 2022 paper on 'Checking bioclimatic variables....'.

We had checked the results to ensure that they did not reflect the well-documented biases with this variable. We now mention this point (and a more general recommendation about checking the partial responses for each variable in space) in the main text.

Reviewer 2

Congratulations to the authors - this package looks to be the product of an enormous amount of careful work. I'm very happy to see more ecology-adjacent Julia packages become available, as I'm interested in doing more work with Julia going forward.

We thank the reviewer for their comments. We are aware of the challenge of introducing a tool in a language that is not the most commonly used, and we have made changes to the text with this specific context in mind.

As I note below, I'd like to see the authors make a stronger case in the conclusion for why their package should be adopted over existing (mainly R) options - this is a very cool package, sell us on the cool stuff it does in the conclusion!

The reviewer asked many relevant questions that, hopefully, helped us make a stronger case for what the package brings.

There were no line numbers in the version of the paper I reviewed so I had to use quotes below.

We apologize for the oversight, and have added line numbers.

“The ability to link data to these steps is central to support the correct interpretation of these models”

I don’t understand what this means, can you clarify what it means to “link” data to a step?

We have clarified this sentence to emphasize that we care about having data in the right format for each of these different steps.

Should be “In recent years”, not “In the recent years”

Indeed - fixed.

“Independently” not “independentl”

Corrected.

“20% of publications” Not sure how you are getting this number, based on Kellner 2025 it should be 12% (17/141 papers that had code) or 3% (17/497 of all papers)

In Kellner 2025, 17 out 97 failures to reproduced in the presence of code were due to missing packages. This is about 20% of the papers that can be evaluated for this criteria. We have clarified the sentence.

“SDT uses the built-in Julia package manager...”

It is unclear to me what specific advantage the monorepo→ multiple component packages approach has for **users**. Presumably when someone updates or newly installs SDT, they’ll pull the code from the entire repository via the package manager (?).

You are correct about the installation process. The Julia package manager handles dependencies automatically regardless of structure. We made no changes to the manuscript in response to this point, but wanted to establish this as the baseline moving forward.

So what specific advantage is there for a **user** to think about SDT as multiple component packages in a monorepo vs just a single overarching package SDT? I understand for developers it makes sense to break things up into parts like this for organizational reasons etc. But for average users it might just be confusing to conceptualize it this way.

A monorepo offers a significant advantage: it provides a single source of truth for documentation and issue tracking. In other words, users do not need to know specifically what component package is responsible for a bug in order to open an issue, and everything is centrally managed. The same logic goes for documentation: although each component package has its own technical documentation, the manual that users are likely to consult is centralized and provides examples that naturally integrate functionalities from all component packages. In addition, a monorepo creates a unified codebase where all functionalities are easily discoverable - users do not need to navigate multiple repositories to understand how components interact. We have expanded our explanation in the manuscript to highlight these advantages.

I see later you note you can install the component packages separately, so maybe worth making that point earlier in the paper.

We have clarified this point earlier in the “Application description” section.

“Following a constructive cost model analysis...”

This is interesting, though I know nothing about this method. But what is the importance of the \$325k? I.e., what would you actually use this information for?

This information was mostly relevant for us, to assess the value of the software contribution for e.g. grant applications. We have removed it from the manuscript.

“The GBIF package”

One challenge I have found when doing SDMs is with GBIF data is that they want you to register the datasets you download to specific DOIs when you use them in a publication (<https://www.gbif.org/citation-guidelines>). With access to a nice API like this, some users may never think to check the proper way to cite these data. Is there a way to make this clear in your package? Or perhaps this is already done?

This is indeed already implemented. We have updated the code in the SDM notebook to show how this can be done (`GBIF.download(doi)`), and added a few sentences explaining that this is best practice. We have also added a sentence to the main text to clarify this (with a citation to the dataset).

“Fauxcurrences”

This looks really cool, glad it is included.

Thank you!

“Finally SDeMo provides tools for training and education”

What do you mean by “education” here? Education in the sense of users learning how to use these models? If so what specific features are there for users to learn things? I would not use the word “education” when you really mean “model training” or “learning”.

We do indeed mean “education” in the didactic sense. We have clarified that **SDeMo** is a high-level interface to train SDMs, and outline at the end of the paragraph why some design choices make it useful when teaching. We use this package for both workshops and a graduate level class in predictive modeling.

Can you provide a citation for the BIOCLIM classifier? I have never heard of this, I think of BIOCLIM as being the set of bioclimatic variables, not a statistical analysis tool.

We have added a reference to a more recent paper outlining the development of the model in relationship to work on the bioclimatic variables themselves.

Are there plans for additional classifiers? That’s likely to be the thing users want the most, I would guess, considering the popularity of other models like GAM, Maxent, etc. for SDMs

The structure of **SDeMo** makes it easy to extend. There are no particular plans to add more built-in classifiers (with maybe the exception of boosting functions), but contributions by users are welcome. We have clarified this point.

I realize you can’t provide every detail in a paper like this (vs offloading some of it to the notebooks), but as someone who sometimes builds SDMs, I’d be specifically looking for details on what approaches are available for dividing training/testing data and what statistics/methods are available for evaluation. So a short summary of this info would be helpful.

We have clarified that the package will always at least maintain class balance, but that users can provide their own splits for cross-validation, together with a description of future plans for more stratified methods.

“In practice, flexible (and more performant)...”

Can you use the recommended MLJ via your package? It's unclear from the way this is worded.

We have clarified this section of the main text. MLJ can be integrated with SDT, but this is a technical step that goes beyond the scope of this paper. SDT can be extended to use any arbitrary function for classification and data transformation, which includes using code from MLJ.

As you note in the intro, right now, most SDM development is done in R. Do you have specific reasons for why you think scientists should consider switching to Julia to use your software? Are there specific advantages over what's available for R already, such as biomod2? In general I think the software looks really cool, so the authors ought to make it a bit more clear why someone would use SDT in Julia over the other existing, long term options. Maybe highlight some features not found in any R packages etc.?

We have added a few explanations at the end of the text.

I looked through the provided Jupyter notebooks and both the explanations and the code seemed clear to me. I am just a beginner Julia user though, so I can't provide detailed comments on them

Thank you - we have further clarified some of the steps during the revisions.