***Editor's comments:***

*~~I now have two completed reviews on your manuscript and based on those reports it may be reconsidered following major revision~~.*

*Reviewer #1 wanted a clearer objective, to which I may add ~~that EcoMod publishes process based models and SDMs are statistical by nature. Early ones wer pubilshed here as new methods and it is the novelty method (not the application) that has the most relevance for EcoMod.~~*

*Reviewer #2 pointed out some potential issues with the method or presence only data. ~~They have both provided you with extensive comments below. Please give a written description of the changes you make in your revision to addreess these comments.~~*  
  
**Response:**

**Although in general terms SDMs can be essentially classified as correlative methods, they have been found as fundamental parts of more mechanistic approaches. As such, it is important to create more robust SDMs to eventually develop more reliable process based models. And this is the main purpose of MinBAR and its approach behind, not to delve into the statistics of the algorithms but to give tools which help modellers to better understand the natural processes and to make better predictions. For that reason, with MinBAR, we address one of the less studied parts of only presence SDMs, and thus probably one of their main weaknesses, which is the accurate delineation of the background area. And because all of this, we believe that our manuscript is adequate to be published in Ecological Modelling.**

**The objectives have now………….**

***Reviewer #1:***

*Dear authors,  
~~This is a very interesting manuscript and very useful in the field~~. ~~After reading the entire manuscript, however, I'm not sure if the manuscript aims (1) to propose a method for choosing the minimum calibration area to pick background points (as the title suggests) or (2) evaluating the validity of building partial models for cases where modeling complete distributions are computationally demanding and, then, how to make the process faster. In my opinion and after reading several times the method, the second option could be more exact. But the important question is, what do you want to show?~~*

**Response:**

**Actually, we pursue both aims mentioned by Reviewer #1, as they are not mutually exclusive. Take two of the case studies as examples. On the one hand, case study1 uses the whole extent of the species and run the function. The result is a set of buffers from which the user can see which one is the minimum that gives sufficient performance according to his/her needs, not only by choosing the faster buffer, but especially by seeing from which one the Boyce indexes no longer increase substantially.**

**On the other hand, in the case study on islands (thus partial distributions/models), after running the function, the user can see in the plots if the curve has reached a sufficient level of the evaluator or, on the contrary, there is still room for improving the models if the extent of the background area goes further the limit of the islands.**

**Now the aims of the tool have been rewritten in lines 68-70 as follows: “*MinBAR* is an R package that aims at (1) defining the minimum background extent necessary to fit SDMs reliable enough to extract ecologically relevant conclusions from them and (2) optimizing the modelling process in terms of computation demands”.**

**In addition, we have included some further explanations in the outputs and case studies sections to make clearer the aims and potentialities of MinBAR:**

**- “… to let the user know which one is the buffer resulting in the model that best captures the environmental conditions of the species distribution”. Lines 148-149.**

**- “Therefore, this would corroborate our hypothesis that a part of the species distribution close to its geographical centre would produce models with the highest levels of evaluation metrics”. Lines 189-191.**

**- “In addition, they show how this tool can help the user to confirm that, even though the area of study might not be the entire distribution of the species, the models can still be reliable enough”. Lines 235-237.**

*I~~n addition, little or no discussion of other studies that have also attempted to propose modifications in the selection of the calibration area is included, so it is impossible to know in what state of knowledge we are now.~~*

**Response:**

**Unfortunately, as far as we know, there are not many studies addressing this issue. Most of them, all mentioned in the introduction, limit the delineation of the background area either by the availability of data, or by intrinsic characteristics of the species such as dispersal capacity, or even simply by political boundaries. *MinBAR* aims to be a tool which helps the modeller to objectively define this extent. We have mentioned now in lines 45-46 that “as far as we know, a methodology to objectively define it [the background extent] has not been developed so far”.**

*OTHER IMPORTANT SPECIFIC COMMENTS AND QUESTIONS ARE BELOW:  
  
~~LINES 14-15: "One of the crucial choices when modelling species distributions using pseudo-absences approaches is the delineation of the background area to fit the model."  
COMMENT: Because pseudo-absences and background -points- are not exactly the same, I suggest to use something like: "One of the crucial choices when modelling species distributions using pseudo-absences and background approaches is the delineation of the calibration area to fit the model."~~*

**Response: Done.**

*~~LINES 16-17: "… around the geographical centre of the species distribution that characterizes".  
COMMENT: What do you think about the environmental (not geographical) center of the species.  
I really think that given that niche models are built in environmental space, the background dataset should be chosen in the context of the environmental conditions in which species may occur or, even more complicated to define, within "the area that has been accessible to the species of interest over relevant time periods" (sensu Barve et al. 2012).~~*

**Response:**

**We agree with Reviewer #1 in that, when the only aim of the modelling process is to define the ecological niche of a species, the environmental space may be highly relevant. However, when the actual distribution of a species wants to be modelled, other information should be included in the model, such as species interactions or the mentioned accessibility. A highly interesting overview of how the environmental niche should be defined can be seen in Jiménez et al. (2019), published by this same journal. They state that the geographical component has to be included in the model. We believe that a way to include this information could be proxied by using the geographic centre of the species more than only the environmental centroid. However, given that the centre-periphery hypothesis in biogeography has been sometimes criticized in the literature (many times corroborated, though), we think important to develop in future *MinBAR* versions the possibility of using either the geographic or the environmental centre depending on user’s preferences. And this is now more explicitly mentioned in line 244.**

*~~LINES 21-22: "… models (SDMs) until a satisfactory model is reached."  
COMMENT: In terms of some metrics? Ecological realism? Please, specify.~~*

**Response:**

**Done. Now, “...models (SDMs) until a satisfactory model in terms of the included metrics is reached” (lines 20-21).**

*LINES 35-39: "However, regarding the public biodiversity data, they are often limited to only  
species presences and with a lack of occurrences in poorly sampled areas. These facts  
limit the use of some techniques or algorithms and force to make critical assumptions and  
choices, introducing different levels of uncertainty to model predictions (Jarnevich et al.,  
2017)."  
COMMENT: Although I agree with these ideas, I see no obvious connection to the purpose of this study. How do these facts directly affect the selection of the calibration area? Please, read my main concern at the top of the revision.*

**Response:**

**With this paragraph we wanted to introduce some usual sources of uncertainty to the species distribution modelling processes. For instance, the fact that usually the available biodiversity data is limited to presences forces to use only-presence based approaches (e.g. MaxEnt), which come with the intrinsic difficulty of defining an accurate calibration area. As mentioned before, its delineation often relies on subjective choices made by the modeller which lead to include some levels of uncertainty to the final conclusions. Therefore, the use of *MinBAR* functionalities may reduce such levels of uncertainty by helping to make more accurate decisions in a crucial step of the modelling process when using only-presence approaches.**

*~~LINES 40-42: "One of the crucial choices when using pseudo-absences approaches is the delineation of the background area to fit the model, also called "landscape of interest" or "study area"  
(Elith et al., 2011; Raes, 2012)."  
COMMENT: Please, note my comment in the Abstract about the use of pseudoabsences and background points. In addition, I think that "calibration area" is another common and probably more appropriate term in the field.~~*

**Response: Included.**

*LINES 54-56: "Any of these situations usually lead to fit models using only part of the species distribution (i.e. partial models), which might or might not imply a reduction of model performance (El-Gabbas and Dormann, 2018)."  
COMMENT: It does not seem convincing me that these limitations mentioned lead to having to build partial models. In fact, in my experience, I don't think it's a common thing to try to solve the problem building partial models. Do you have more examples of this being a common or suggested strategy?*

**Response:**

**On the one hand, in the cited paper by Melo-Merino et al. (2020), they reported that 48% of the analysed studies (n = 328), based on marine species and mostly using correlative models, were at local scales (i.e. covering the territory of one country or less). Of course some of them could be focused on species with very limited range, but probably not all of them (at least one example is reported along the text). On the other hand, for instance, El-Gabbas and Dormann (2018) and Titeux et al. (2017) also addresses the issue of using large-scale vs regional data when modelling species distributions, indicating the common use of the latter.**

**In addition, just some examples found in a quick search of recently published works, based in MaxEnt, which use local/regional data for species with larger ranges (we do not think they have to be cited in the manuscript):**

**- Yan, Huyong and Feng, Lei and Zhao, Yufei and Feng, Li and Wu, Di and Zhu, Chaoping. 2019. Prediction of the spatial distribution of Alternanthera philoxeroides in China based on ArcGIS and MaxEnt. Global Ecology and Conservation 21: e00856**

**- Zhang, J, Jiang, F, Li, G, et al. 2019. Maxent modeling for predicting the spatial distribution of three raptors in the Sanjiangyuan National Park, China. Ecology and Evolution 9: 6643– 6654.** [**https://doi.org/10.1002/ece3.5243**](https://doi.org/10.1002/ece3.5243)

*~~LINES 60-61: "…both quality of the model and execution time."  
COMMENT: Please clarify what you mean by model quality.~~*

**Response: Now “… in terms of quality of the model (i.e. giving highest, but not overestimated, levels of model’s evaluation metrics)… ” (L64-66)**

*~~LINE 67: "PROBLEM"  
COMMENT: I am confused with the main problem. Is the problem of not being able to know which is the minimal area to extract the background points (which applies to any species) or the inherent problem when building partial models in the case of species with a very wide geographical distribution? I think both are different things and therefore the solutions might be different.~~*

**Response:**

**Yes, we agree in that they are two different problems. But as explained before in the response to Reviewer’s question about the aims, both can be solved depending on the use given to MinBAR. On the one hand, the user can choose the narrowest buffer when the curve of the resulting metrics (Boyce index) stabilizes (i.e. the minimal area to extract the background points from without loosing modelling performance). On the other hand, the user can be able to see how the performance of the model diminishes, but so the computation resources (i.e. execution time), when using narrower areas for modelling (i.e. partial models instead the whole distribution), so that he or she can decide which is the best extent in terms of the compromise between quality (according to Boyce index estimates) and execution time.**

**In order to be clearer, we have modified these two paragraphs (Lines 72-81)**

*~~LINES 91-93: On the other hand, Boyce Index Total (BI\_tot) assesses predictions beyond the training area, across the whole distribution of the species (i.e. model transferability).  
COMMENT: Do you mean that in this case evaluation includes the ability of the "partial model" to predict records in other parts of the distribution not included in the calibration? If so, then I think a quantification of non-analog conditions should be included in the analyses, since the environmental conditions within the calibration area could be more restricted than in the transfer area.~~*

**Response:**

**A multivariate environmental similarity surface analysis (MESS) is now included in the package to evaluate the differences between the environmental conditions used to train the models and the transfer area. It is now mentioned in the manuscript in lines 99-101 and 149-150, respectively as:**

**“… in order to compare the environmental variables used for training the model with those across the projected area, a multivariate environmental similarity surface (MESS) analysis (Elith et al., 2010) is run.”**

**“… In addition, other ancillary data is generated during the process for each model (e.g. MESS maps).”**

*~~LINES 98-100: "…variables, the function calculates the geographical centre of the species distribution, the most distant occurrence and the buffers. "  
COMMENT: I'd add the assumptions to use the geographical center of the distribution to be used as the main factor to define the calibration area. Please, add a short sentence about this contentious issue.~~*

**Response:**

**Included as follows (Lines 107-114): ”… variables, the function calculates the geographical centre of the species distribution, the most distant occurrence and the buffers. The geographical centre corresponds to the mean location for longitude/latitude coordinates dealing with angularity (presences points and variables are previously “reprojected” to WGS84 if needed). It can be found in the literature some criticisms about the geographical component of the centre-periphery hypothesis (e.g. Pironon et al., 2017). This is why in future versions of MinBAR the user will be able to choose among the geographical and the ecological centre of the species.”**

*~~LINES 183-184: "… distribution (environmental suitability raster) by means of ENMTools (Warren et al., 2017) using I and D metrics."  
COMMENT: Please, add a description of those metrics, how they are calculated and what they indicate. How are they interpreted?~~*

**Response:**

**Now both metrics description and interpretation are included in lines 205-209, adding also the main reference for further details, as follows:**

**“Both metrics are defined to measure similarities between two environmental niche models projected to the geographic space. And both range from 0 to 1, meaning no overlap at all and perfectly identical niches respectively. See Warren et al. (2008) for further details on how they are calculated and their similarities and differences.”** *~~LINES 188-190: "… of the virtual species (Table 2) gave an average of 0.69 (SD = 0.16) and 0.87 (SD = 0.11) for D and I respectively."  
COMMENT: This result (and Table 2) is not discussed further below. What do those similarity values imply?~~*

**Response: Now the result is discussed in lines 215-217 as follows:**

**“These results showed how, on average, the potential distribution of the best buffers resulted on a good overlap with the actual distribution of the species, meaning that *MinBAR* is able to capture the real distribution in a notable degree.”** *~~LINES 218-219: "Furthermore, the implementation of other algorithms and modelling techniques would be highly convenient  
COMMENT: I am also thinking about the resolution (cell size) of the study. Is there any exercise using a finer resolution?~~*

**Response:**

**The tool has been tested so far at two different scales, as presented in the case studies, being case 3 the one using the finest one (0.5 arc-min). However, MinBAR is scale-independent, although, obviously, the finer the scale is, the more sense makes using it, as the computing resources increase with the number of pixels of the background area. Now this fact is mentioned in lines 171-173 (“… and, although MinBAR is scale-independent, it was tested with two different resolutions depending on the study case”).**  
  
*~~TABLE 1.  
COMMENT: Species names must be specified in the first column or in the legend. They must also specify the unit of measure in the rest of the columns.~~*

**Response: Done. Now it is Table 2. The same improvements have been applied to Table 3 (now Table 4).** *~~TABLE 2.  
COMMENT: In the legend, what do those values indicate? How are they interpreted?~~*

**Response:**

**Now the table caption (now Table 3) has been rewritten as follows: “Degree of similarity between the probability of presence of the best buffer calculated by MinBAR with the actual distribution of the same (virtual) species. The number in 'BestBuffer' represents the buffer number, being 1 the closest to the centre and 10 the furthest. Metrics 'D' and 'I' represent the level of similarity, being 0 no similarity at all and 1 identical probability of presence”.**

**In addition, as shown above, now both metrics definition and their interpretation are better explained along the text.**

***Reviewer #2:***

*~~Many thanks for inviting me to review the manuscript behind the MinBAR package. I enjoyed reading the text. It is well written with a clear storyline to follow. The codes behind the package are all available on GitHub and I smoothly installed the library from CRAN. The authors are targeting an important issue in modelling the distribution of species which is defining the study area to sample the pseudo-absents. I believe the manuscript offers a valuable contribution to the community of species distribution modellers and nicely fits the scope of the journal. There are, however a few points that I would like to discuss with the authors to be considered in the next revision of the manuscript.~~*

**Response: We thank Reviewer #2 for his/her kind words and valuable comments below.** *~~MinBAR uses the Boyce Index as an evaluation metric. Boyce Index only requires presences and thus appropriate metric in the case of presence-only models. It demonstrated predicted-to-expected ratio across the suitability gradient. My concerns are; a) whether the selected evaluation metric would be appropriate for other presence-absence SDMs too?~~*

**Response:**

**As the reviewer pointed out, the Boyce index is adequate for models derived with only-presence approaches, which are the main target of the MinBAR tool. However, in case MinBAR would like to be adapted to work with presence-absence modelling techniques, and according to Hirzel, et al. (2006), the continuous Boyce index would be a good complement for their evaluation together with other metrics which probably perform better than this (e.g. those based on sensitivity and specificity). Actually, we already mentioned in the conclusions of the manuscript that the inclusion of other metrics based on sensitivity/specificity is something that we will absolutely address in future versions of MinBAR.**

*~~b) considering that the "background" in MaxEnt is different from the pseudo-absences in other SDMs, do authors expect to get the same results with the POC plots (~~*[*~~https://doi.org/10.1890/09-07~~*](https://doi.org/10.1890/09-07)*~~)?~~*

**Response:**

**Di Cola, et al. (2017), who developed the tool used by MinBAR to calculate the Boyce index, already mentioned that this index is the quantitative equivalent of the POC plot described by Phillips & Elith (2010). Therefore, we expect that the POT plot would corroborate the MinBAR results obtained by the Boyce index. Furthermore, the POC plot, as its name indicates, is a graphical tool which might help the user to visually improve the presence-only models. However, we think that for MinBAR it is better to use the quantitative value as it is more useful to compare, for instance, the results for the different buffers, and also to be used in further automatic calculations and generation of results (e.g. for several species, etc).** *~~Please add the complete list of parameters (i.e. arguments) used in running MaxEnt.~~*

**Response: Done. Now it is Table 1 in the manuscript.**

*~~Considering 30% test dataset, 3 model replicates (even with crossvalidation) per buffer seems insufficient, doesn't it~~?*

**Response:**

**We used 3 replicates both to make the calculations a bit faster due to the high number of species to run in the case studies and because it is the minimum number of occurrences to run MaxEnt according to van Proosdij et al (2016), but it is true that it might be a bit short to get averaged models to converge when increasing the number of occurrences. Now 15 is the default, as it is closer to the minimum number of occurrences in the real data study of the same article.**

*~~Authors presented 3 case studies with presence-only species data. I was expecting to see a proof of concept with absence-presence species data too. To use the presence occurrences to run the MaxEnt and use the absence-presence data to report a series of evaluation metrics not only on the discrimination capacity but also on the goodness-of-fit (i.e. calibration) of the models.~~*

**Response:**

**On the one hand, the delineation of a background area is an inherent issue of the only-presence modelling approaches, which are among the main scope of MinBAR. Presence-absence species models do not need to define a theoretic space, based on assumptions, where to extract absences from, but the area of study is defined by where fieldwork has been conducted with either positive or negative results (Guillera-Arroita, et al., 2015; Valavi, et al., 2019). On the other hand, this version of MinBAR is only implemented for MaxEnt, thus only for models using pseudo-absences. These are the main reasons why we did not present case studies with real absences.**

**In addition, regarding the evaluation metrics issue mentioned by the Reviewer, actually the implemented continuous Boyce index is already able to capture the goodness-of-fit of the models (Di Cola, et al., 2017; Manzoor, et al., 2018), which is more useful in many cases. However, as mentioned in the manuscript, the discrimination accuracy of the model is also reported by MinBAR as ancillary information.**

*~~Considering the algorithm of MaxEnt it is expected that the broader study area will return models with higher discrimination capacity (this is also visible in supplementary material 3). So the question is that except the technical benefits (i.e. time and computational resource), is there any conceptual benefits in applying MinBAR for users? Authors briefly discuss this in lines 46- 51, but it would be important to elaborate further on the application of the package in selecting the study area not by delimiting but also to set up the minimum extent.~~*

**Response:**

**As explained before, in comments to Reviewer’s #1, now the aims of the tool are clearer expounded in the sense of which are the benefits of using MinBAR besides the “technical ones” mentioned by Reviewer #2. In this regard, now we think we are much clearer along the manuscript in giving more importance to the potentiality of MinBAR to find a best solution for the background area (i.e. a minimum extent), than to minimise the computational resources.**

**However, following the comment of Reviewer #2, we have also included some sentences reinforcing the idea of why is it important to minimise the background area extent:**

**- “... given that increasing such area too far away from the presences might lead to overestimate the accuracy of the SDMs (Lobo et al., 2008)”. Lines 42-44**

**- “Thus, fitting the SDM within this minimum geographical area should be the best solution in terms of quality of the model (i.e. giving highest, but not overestimated, levels of model’s evaluation metrics)…”. Lines 63-66.**

*~~Please add the method used to calculate the "geographical centre fo species distribution" and would be great to have it as a line in Figure 1 and supplementary material 4.~~*

**Response:**

**Explanation included in lines 109-111 (“... The geographical centre corresponds to the mean location for longitude/latitude coordinates dealing with angularity (presences points and variables are previously "reprojected" to WGS84 if needed”…), as well as in Figure 1 (now Figure 2) and Supplementary materials.**

*~~Please add a map to demonstrate the spacial representation of the concept. For example in the case of Fabus sylvatica (fig 1), the presence points, the geographical centre fo species distribution, and the buffers.~~*

**Response: Done. Now this is Figure 1, and it is mentioned in lines 182-183.**

**BIBLIOGRAPHY**

Di Cola, V., Broennimann, O., Petitpierre, B., Breiner, F.T., D’Amen, M., Randin, C., Engler, R., Pottier, J., Pio, D., Dubuis, A., Pellissier, L., Mateo, R.G., Hordijk, W., Salamin, N., Guisan, A. (2017). Ecospat: An r package to support spatial analyses and modeling of species niches and distributions. Ecography 40, 774–787

El-Gabbas, A. and Dormann, C.F. (2018) Wrong, but useful: regional species distribution models may not be improved by range‐wide data under biased sampling. Ecology and Evolution. 8(4):2196-2206

Guillera‐Arroita, G., Lahoz‐Monfort, J.J., Elith, J., Gordon, A., Kujala, H., Lentini, P.E., McCarthy, M.A., Tingley, R. and Wintle, B.A. (2015), Matching distribution models to applications. Global Ecology and Biogeography, 24: 276-292

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C., Guisan, A. (2006). Evaluating the ability of habitat suitability models to predict species presences. Ecological Modelling 199: 142–152

Laura Jiménez, Jorge Soberón, J. Andrés Christen, Desireé Soto. (2019). On the problem of modeling a fundamental niche from occurrence data. Ecological Modelling 397: 74–83

Manzoor, S.A., Griffiths, G. & Lukac, M. (2018) Species distribution model transferability and model grain size – finer may not always be better. Scientific Reports 8: 7168

Phillips, S. J. and Elith, J. (2010). POC plots: calibrating species distribution models with presence-only data. – Ecology 91: 2476–2484

Titeux, N., Maes, D., van Daele, T., Onkelinx, T., Heikkinen, R. K., Romo, H., ... Luoto, M. (2017). The need for large-scale distribution data to estimate regional changes in species richness under future climate change. Diversity and Distributions, 23(12), 1393–1407. https://doi. org/10.1111/ddi.12634

Valavi, R, Elith, J, Lahoz‐Monfort, JJ, Guillera‐Arroita, G. blockCV: An r package for generating spatially or environmentally separated folds for k‐fold cross‐validation of species distribution models. Methods Ecol Evol. 2019; 10: 225– 232

van Proosdij, A.S.J., Sosef, M.S.M., Wieringa, J.J. and Raes, N. (2016), Minimum required number of specimen records to develop accurate species distribution models. Ecography, 39: 542-552