Dr. Rodrigo Medel Subject: PCI Ecology MS#81

Dear Dr. Medel,

Thank you for your email enclosing the reviewers' comments on our MS "Environmental variables determining the distribution of an avian parasite: the case of the *Philornis torquans* complex (Diptera: Muscidae) in South America". We are truly grateful to the recommender and reviewers for their time and constructive comments; we are sure that they greatly improved our manuscript. We have implemented almost every comment and suggestion provided and our responses are given in a point-by-point manner below.

We are submitting a revised version of the manuscript and a line-by-line response to all reviewers' remarks. We hope the revised version is now suitable for recommendation and look forward to hearing from you in due course.

Sincerely,

Pablo F. Cuervo

Martín A. Quiroga

Reviewer 1

This study uses species distribution modeling to identify the environmental variables that determine the geographic distribution of *Philornis torquans*, a dipteran ectoparasite of neotropical birds. The results indicate that temperature and moisture both contribute to limiting the geographic range of *P. torquans*. The study clearly and accurately communicates the results of the analysis. I am not familiar with MaxEnt and the analytical approach used, though my brief review of recent MaxEnt best practices suggests that the authors have conducted this analysis appropriately. I found the result that *P. torquans* may be limited by winter temperature, at both extremes, to be an interesting finding worthy of further study.

• <u>Reviewer #1</u>: Regarding the broader contribution of the study: the manuscript lacks a strong motivation at the beginning and a strong take-home message. As a result, it's not clear what broader, exciting contribution this study makes to the existing literature.

<u>Authors' answer</u>: the suggestion was considered, the opening of the introduction was rephrased and a motivation message was included (lines 53-68). Additionally, we included a take-home message at the end of introduction (lines 88-90).

• <u>Reviewer #1</u>: The abstract contributes to this problem – it's hard to read and the key problem and contribution of the study are not clear. It would be better structured as a single paragraph (no bullets) or with structured bullets (problem, approach, results, take-home).

Authors' answer: the abstract was modified to be clearer and more compelling.

• <u>Reviewer #1</u>: The opening paragraph of the Discussion highlights the methodological care with which the study was conducted. It would be more valuable to start off this section with a brief summary of the major problem and the take-home message, unless the methodology is the key contribution of the study (I don't believe that's the intent, but I am not qualified to evaluate any methodological contributions). If the key goal is to contribution is the creation of a reference for conservation efforts, the Discussion should link the findings to considerations relevant to conservation.

Authors' answer: the opening paragraph of the discussion was modified (lines 326-331).

Reviewer's major comments

• <u>Reviewer #1</u>: Philornis flies are obligate parasites as larvae right? Then it would be helpful to have more information about the host range of *P. torquans*. I gather it's a generalist, but it seems to be strongly associated with certain host species. To what extent does the distribution of host species explain the geographic distribution of *P. torquans*? Are environmental factors acting directly on *P. torquans* distribution or indirectly, via bird hosts? It's necessary to present (even if to ultimately dismiss) the possibility that host availability contributes to the geographic range when studying an obligate parasite.

<u>Authors' answer</u>: the reviewer's concern is appropriate. The issue is briefly discussed in lines 418-423. Furthermore, a supplementary table with known hosts is provided.

• <u>Reviewer #1</u>: Sampling and sample size: line 117 mentions field surveys that were conducted – how were locations for these chosen? Likewise for the data from existing literature – a summary of how sites

were selected across studies would give some impression of the potential for biased sampling to impact the results of the analysis. The sample size is also fairly low – I see that this sample size is within the recommended range for minimum sample sizes from van Proosdij *et al.* (2016), but they also indicate that this range is highly sensitive to species and study area. Did the authors investigate the sensitivity of model accuracy to sample size?

<u>Authors' answer</u>: Additional information about our field sampling (lines 113-118) and literature review (lines 118-120) is now provided in the text. Regarding bias and selection of sampling sites, we are fully aware it is a pervasive fact when recovering data records from museum databases or literature, as we have done, where the purpose of the original study was other than dealing with niche modeling or spatial analyses. Indeed, this type of occurrence data is typically biased in favor of accessible areas (near towns, roads, protected natural areas, etc.), and thus some environmental conditions are more heavily sampled than others (Merow *et al.* 2013, *Ecography* 36:1058-69). Yet, as *Philornis* species are poorly known and the number of occurrence records available is quite limited, we are obliged to rely on these records for modeling.

Regarding sample size and sensitivity, we recognize that model quality is clearly influenced by the number of records used in model building. Yet, as the number of occurrence records of any *Philornis* species is limited, we cannot investigate the sensitivity of model accuracy across artificial manipulations of sample size. On these regards, Maxent is among the algorithms with best predictive power across all sample sizes (as low as ten records) (see Hernandez *et al.* 2006, *Ecography* 29:773-85; and Wisz *et al.* 2008, *Diversity Distrib.* 14:763-73), having moderate sample size sensitivity combined with excellent predictive ability (Wisz *et al.* 2008, *Diversity Distrib.* 14:763-73). Maxent's strong and consistent performance across sample size manipulations may be explained by the way it uses regularization to avoid over-fitting (Wisz *et al.* 2008, *Diversity Distrib.* 14:763-73). The amount of regularization varies flexibly with sample size to ensure consistent performance (Warren & Seifert 2011, *Ecol. Appl.* 21:335-42; Scheglovitova & Anderson 2013, *Ecol. Model.* 269:9-17), which we have considered when evaluating a number a models across a range of regularization values and features classes (as suggested in Morales *et al.* 2017, *PeerJ* 5:e3093). Some of this is now discussed in the text (see lines 333-347).

• <u>Reviewer #1</u>: The Discussion is long. Areas that feel particularly long: the first paragraph on methodological sophistication; references to modeling immature parasitic stages as a limitation of the study (given that the immature stages are obligate parasites, and conservation of bird hosts is cited as a primary use of the study results, I think modeling the ecological niche from the parasitic stage is legitimate, or at least not a sufficient problem to merit this much discussion); paragraphs from 366-389 could be collapsed and shortened to make the point here more clearly; lines 431-444.

<u>Authors' answer</u>: The discussion was fully revised, and most of the reviewer's suggestions were considered.

• <u>Reviewer #1</u>: Figures: I found Figure S1 (even S2) far more helpful in making sense of the results than Figure 2. Figure 3c is difficult to interpret from the legend.

<u>Authors' answer</u>: As similar (not the same) information is provided in Table 2, Figure 2 is now provided as supplementary material. In addition, Figure S2 was moved to the main manuscript, to help make better sense of the results.

• <u>Reviewer #1</u>: There are several spots, particularly the Abstract and Introduction, where word choice or sentence structure need revision. I've highlighted a few examples below:

Line 57 – check sentence structure – a misplaced and?

Line 69 – an to and

- Line 79 mortality of nestlings
- Line 96 constrain in place of restrain
- Line 104 indistinguishable in place of undistinguishable
- Line 275: had a bell shape
- Line 280: had a sigmoidal shape
- Line 367 native to southern south America
- Line 382 remove "on the other hand"
- Line 403-404: clause structure "which lately" doesn't make sense
- Line 407 was the most influential factor
- Line 425 though the humidity and moisture are clearly relevant factors determining...
- Line 436 the latter makes sense
- Line 442 relatively in place of relative

<u>Authors' answer</u>: All the observations highlighted by the reviewer were considered and revised in the text. In addition, the whole manuscript was thoroughly revised.

Reviewer 2

Dear Editor I pleased to present you the review of manuscript entitled "Environmental variables determining the presence of an avian parasite: the case of the *Philornis torquans* complex (Diptera: Muscidae) in South America" of Cuervo *et al.* This study pretends provide a methodological framework to understand the potential distribution of *Philornis* complex, a parasitic fly group of birds. The manuscript is clear in their methodology (with exceptions, see general comments) and results and could be used as a powerful tool to predict the distribution of this flies along the distribution known of this complex and also, to conserve endangered birds. However, need a major review in some aspects. I recommend publish this manuscript when the questions are resolved by the authors.

• <u>Reviewer #2</u>: 1. In Methodology, the authors not mentioned as we obtained their dataset. This is very important because is the baseline for any modelling niche study. I necessary that the author mention how sites were surveyed? how many replicates (or pseudoreplicates) have any sites? Seasonality? How standardized the information? Taxonomic authority that classify the philornis? Methodology.... Etc, etc... is completely necessary read this information.

<u>Authors' answer</u>: Between lines 113 and 123, we gave some information regarding field sampling and literature review. All the same, we have revised and slightly modified the text, adding some more detail. Regarding the reviewer's concern about "replicates", field surveys were not part of an observational or experimental design requiring for replicates to verify consistency among findings. These field surveys were aimed to collect *Philornis* flies to analyze distribution and genetic differences among populations. As it is, these occurrence records are equally useful for this type of study.

• <u>Reviewer #2</u>: 2. The low amount of records not autocorrelated. This said me that all previous sampled were realized in closer sites. Although the authors adequately detail each step carried out through the recommendations of other manuscripts, I wish they could at least include in the discussion of the work somewhat more elaborated with respect to the predictions of the model with a larger number of data from the *Philornis torquans* complex.

<u>Authors' answer</u>: The reviewer's concern is well justified. Yet, as *Philornis* species are poorly known and the number of occurrence records available is quite limited, we are forced to rely on these records for modeling. Sampling bias is a pervasive fact when recovering data records from museum databases or literature, where the purpose of the original study was other than dealing with niche modeling or spatial analyses. Indeed, this type of occurrence data is typically biased in favor of accessible areas (near towns, roads, protected natural areas, etc.), and thus some environmental conditions are more heavily sampled than others (Merow *et al.* 2013, *Ecography* 36:1058-69). In general, it is advised against modeling with biased not filtered data, as these geographic clusters of localities artificially increase spatial autocorrelation, and such a situation can cause the model to overfit to environmental biases that correspond to these influences in geographical space (Boria *et al.* 2014, *Ecol. Model.* 275:73-7).

• <u>Reviewer #2</u>: 3. In the methodology, the rationale for using this fly complex is because it affects only one threatened bird species. The authors could provide in the supplementary material and then in the discussion a potential or actual list of birds affected by this fly complex.

<u>Authors' answer</u>: The reviewer's observation is not completely true. The rationale behind the selection of this *Philornis* complex was detailed between lines 101-111. Reasons were: i) availability of previous

knowledge concerning its biology; ii) availability of occurrence records; and finally, iii) endangered bird species potentially affected. Nevertheless, we have included other endangered bird species suspected to be parasitized by *Philornis torquans* complex, a supplementary table with known hosts, and a short paragraph about this regard in Discussion (see lines 418-423).

• <u>Reviewer #2</u>: 4. Shorten and rephrase part of the discussion. Please, take into account some of the suggestions written here take. In particular, I have many other recommendations.

<u>Authors' answer</u>: The discussion was fully revised, and most of the reviewer's suggestions were considered.

• <u>Reviewer #2</u>: L56-58: Reorder! First taxonomy and then, reference! L62. Please, clarify this. In the first sentences you tell me that there three genera, including *Philornis* generating Myasis. But then, you tell me that the larvae of *Philornis* are coprophagous, semihaematophagous and subcutaneous. So, what type of feeding is myasis?

<u>Authors' answer</u>: We removed references to other genera producing myasis and to *Philornis* species with coprophagous larvae for a clearer reading.

• <u>Reviewer #2</u>: L70-71: How depend? It is obvious. Change the magnitude? You said this in the preceding sentence. Change the intensity? You don't said the prevalence. Please clarify or remove and reinforce the previous sentence. L72: Reference after "...negligible". L84-89: Please separate in two phrases.

<u>Authors' answer</u>: All these comments were considered.

• <u>Reviewer #2</u>: L113: There is a problem here. The complex is choose because affect to yellow cardinal only? Or affect other endangered birds? Please add new examples or number of bird affected with respective references.

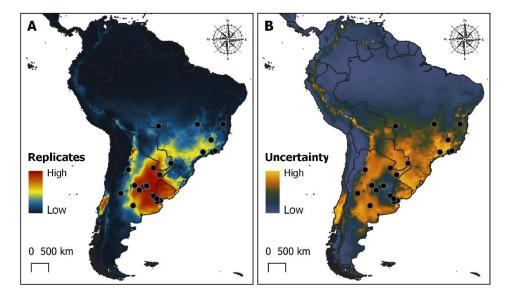
<u>Authors' answer</u>: this observation was treated in detail in a previous comment. We made this clear in text (see Lines 101-111), provided a supplementary table with known hosts, and a short paragraph about this regard in Discussion (see lines 418-423).

• <u>Reviewer #2</u>: L118-121: I need more information. How field surveys? Where? How many replicates por site? In what season your sampled? What literature you consulted? Number of references? How many nest were reviewed in each site? ... please, provided ALL information that support the obtained dataset. L119-120: who determined the larvae? The authors used taxonomic key? Molecular depositories? Please provide these data.

<u>Authors' answer</u>: as stated previously, we gave additional information regarding field sampling and literature review between lines 113 and 123, covering most of the reviewer's comments.

• <u>Reviewer #2</u>: L136-138: I am not convinced by using only 18 data to model the presence of Philornis. When I have worked in modeling, we have always been asked in journals for a number not less than 40 records and especially records that are not autocorrelated. I understand that sampling generates a bias. Questions: did you test the model with the 80 initial records, regardless of whether there is autocorrelation? Did you generate the model once you only considered 34 sites? My idea is that you test these models to see how substantive is the change between the initial model versus the clean version. <u>Authors' answer</u>: The reviewer's suggestion of modeling with occurrence data without removal of duplicates (80 initial records, most of them concerning different host species in the same site) or correcting for sampling bias (34 records) goes against commonly advised (see Merow *et al.* 2013, *Ecography* 36: 1058-69; Feng *et al.* 2019, *Nat. Ecol. Evol.* 3:1382-95), as such biases can lead to environmental bias as well, resulting in an over-representation of environmental conditions associated with regions of higher sampling (Aiello-Lammens *et al.* 2015, *Ecography* 38:541-5).

However, as requested by the reviewer, we did perform a model with the subset of 34 records, prior to spatial filtering (no sense in doing so with the full 80 records, as Maxent removes duplicated sites by default). As expected, in contrast with the prediction from filtered occurrences, the resulting prediction seems overfitted and centered around the geographical clusters of occurrences (regions with higher sampling).



Despite our interpretation of this prediction, it should be considered that Maxent models are fit assuming a uniform sampling (i.e. that all locations on the landscape are equally likely to be sampled) (Merow *et al.* 2013, *Ecography* 36:1058-69), so modeling with biased occurrence records will certainly yield biased predictions.

Regarding the final number of occurrences used for modeling, Maxent is among the algorithms with best predictive power across sample sizes (as low as ten records) (see Hernandez *et al.* 2006, *Ecography* 29:773-85; and Wisz *et al.* 2008, *Diversity Distrib.* 14:763-73), having moderate sample size sensitivity combined with excellent predictive ability (Wisz *et al.* 2008, *Diversity Distrib.* 14:763-73). Maxent's strong and consistent performance across sample size manipulations may be explained by the way it uses regularization to avoid over-fitting (Wisz *et al.* 2008, *Diversity Distrib.* 14:763-73). The amount of regularization varies flexibly with sample size to ensure consistent performance (Warren & Seifert 2011, *Ecol. Appl.* 21:335-42; Scheglovitova & Anderson 2013, *Ecol. Model.* 269:9-17), which we have considered when evaluating a number a models across a range of regularization values and features classes (see lines 214-232, as suggested in Morales *et al.* 2017, *PeerJ* 5:e3093). Some of this is now discussed in the text (see lines 333-347).

• <u>Reviewer #2</u>: L201-202: Please, also provide a negative argument to small dataset. What level of precision is obtained whit few data vs large dataset?

<u>Authors' answer</u>: this issue was briefly discussed between lines 333-336. We recognize that model quality is clearly influenced by the number of records used in model building. Yet, as the number of occurrence records of any *Philornis* species is limited, we cannot investigate the precision of model accuracy across artificial manipulations of sample size. Otherwise, we risk speculation.

• *Reviewer #2*: L206: Some reference?

Authors' answer: the accessible area is properly defined and referenced in the following lines 204-213.

• <u>Reviewer #2</u>: L 209-211: Labud *et al.* 2003 show data about the movement? I don't think so. Contrarily, Showler & Osbrink 2015 efectively show movement >13 km in some cases. Please provide information about *Philornis* species that you use for modelling.

<u>Authors' answer</u>: The reviewer is right about Labud *et al.* 2003 and Showler & Osbrinck 2015. We replaced those references. Unfortunately, there is no information available on the dispersal capacity of the *Philornis torquans* complex, reason why we used information on *Philornis downsi* and other Muscidae flies.

• <u>Reviewer #2</u>: L250-251: Careful! The comparison that you mention has been studied in *Philornis* species? Do you have information about physiological curve of thermal tolerance? Metabolic exchange? Temperature stress o resistance? Thermal limits? Hypoxia? Provide any evidence about this comparison!

<u>Authors' answer</u>: the phrase was a general comment regarding the value of determining the most limiting variables, and was not intended as an affirmation concerning *Philornis* species. All the same, the sentence was rephrased to avoid confusion (lines 252-254).

• <u>Reviewer #2</u>: L253-256: Remove this and incorporate in the legend of the figure! L301. Figure "3a" change capital letter and number

Authors' answer: Both suggestions were considered.

• <u>*Reviewer #2*</u>: L336-342: Please, provide a brief sentences mention that could happen with a high number of records not autocorrelated? The model should be the same?

<u>Authors' answer</u>: as previously acknowledged elsewhere, we are fully aware that data paucity is of concern because model quality is clearly influenced by the number of records used in model building, which we have briefly discussed between lines 333-336. However, we would prefer not to speculate about the results of modeling with a higher number of non-autocorrelated records. In turn, we added a brief sentence indicating that this modelling approach and its results could benefit from the use of a larger database of occurrence records (lines 345-347).

• <u>Reviewer #2</u>: L343-344: You not mentioned how obtain the primary dataset. This is completely necessary for any modelling niche! Please provide all information in the Methodology section and subsections.

<u>Authors' answer</u>: despite it is not completely clear what the reviewer considers as the "primary dataset", the description of how the occurrence data was obtained had been expanded (see lines 113-123).

• <u>Reviewer #2</u>: L349: Migrate is the same of movement? Please clarify this because the torquans complex move of some way. How move by day? By year? There is literature?

<u>Authors' answer</u>: The term "migrate" seems to cause some confusion to readers, so as this information was not highly relevant for the MS, we removed it from text for a clearer reading. There is no literature supporting the idea that *Philornis* flies migrate in any ways. On the contrary, Causton *et al.* 2019 (*PLoS One*, 14(10): e0224125) stated that active *P. downsi* adults were found year-round in Puerto Ayora (Santa Cruz, Galapagos), meaning that flies inhabit the same area. Here, we need to make clear that in Puerto Ayora there are two different areas from which *Philornis* may locally migrate: 1) the lowland region (elevation 15–41 m), where vegetation was predominantly *Opuntia* and *Jasminocerus* cacti, *Cordia lutea, Acacia* sp, and *Parkinsonia aculeata* trees; and 2) the highland region (elevation 589–616 m) that is vegetated primarily by endemic *Scalesia pedunculata* forest. The lowlands and highlands on Santa Cruz are distinct vegetation zones and rainfall is typically much lower in the lowlands. However, adult *Philornis* (specially females) were found in both areas year-round.

• <u>Reviewer #2</u>: L350: *P. downsi* inhabits in Galápagos! That species are limited by the sea! In your case torquans complex is not limited for geographical barriers! if the authors don't suspect to *Philornis* change among states is necessary provide a explain to the potential movement.

<u>Authors' answer</u>: We are aware that *P. downsi* inhabits Galápagos and is limited by the sea. However there could be internal migrations between different environments present at lowlands and highlands, or even between islands. Indeed, Dudaniec *et al.* (*BMC Ecology*, 2008, 8:13) showed strong evidence for high interisland gene flow, evincing that *P. downsi* populations have high connectivity between islands and thus a high dispersal capacity. Made clear in text (lines 210-211).

• <u>Reviewer #2</u>: L362-364: Along the lats is possible that torquans complex present reaction norm of its physiological minimum thermal temperature? Please provide a short sentence with some example or hypothesis please.

<u>Authors' answer</u>: the reviewer's suggestion was considered in the text (lines 365-367).

• <u>*Reviewer #2*</u>: L367-370: Mention species, provide references please. L376: Cursive *Protocalliphora*. L375-378: Some redundant with the previous sentence. Please, shorten the sentence and this paragraph.

Authors' answer: all previous comments were considered.

• *Reviewer #2*: L381-382: How many time live a *Philornis*? There is some reference? Life table?

<u>Authors' answer</u>: the paragraph was modified and this phrase was removed from the text. However, specimens of *Philornis torquans* are known to live for up to 100 days in captivity (Saravia Pietropaolo *et al.* 2018, *Canadian Entomologist*, 150:317-603); while adults of *Philornis downsi* survived almost 200 days under lab conditions (Causton *et al.* 2019, *PLoS ONE* 14(10):e0224125).

• <u>Reviewer #2</u>: L401-406: In global warming scenario, how affect this to your results? Do you thinks tha could increase the infestation? The reproduction increase with the temperature? What other fitness traits increase/decrease with high temperatures?

<u>Authors' answer</u>: We would prefer to avoid this kind of argumentations, and stick to a reasonable explanation of our results. This niche model is based on scenopoetic (environmental) variables and

constrained by accessibility (for details, see Soberon 2010, *Ecography* 33:159-67), but ignoring biotic interactions. To discuss about influence of global warming in a multihost-parasite interaction, without not even a projection of the model to a series of future scenarios, seems a bit too speculative.

• *<u>Reviewer #2</u>*: L409-411: This must be mentioned before in the methodology!

Authors' answer: as suggested, the phrase was moved to the Methods section (lines 172-175).

• <u>Reviewer #2</u>: L443-446: I thinks that this could develop more! Would it be possible for *Philornis torquans* complex to invade Chile through its own mechanisms? certainly, the authors do not have this clear, since they do not know the capacity of movement (or migration) of the complex as well as physiological aspects that could give a better explanation to the invasion in an area of Chile where average temperatures could ensure adequate development of the species.

Authors' answer: a brief discussion about the issue was added to the text (lines 424-432).

Reviewer 3

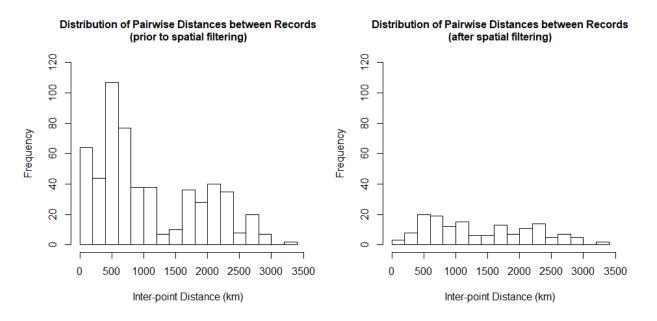
Comments on "Environmental variables determining the presence of an avian parasite: the case of the *Philornis torquans* complex (Diptera: Muscidae) in South America" by Cuervo *et al.*

• <u>Reviewer #3</u>: Line 88-89: this is a great example of how this information could be useful. " to locate habitats with low risk of *Philornis* occurrence for the re-introduction of bird species or reinforcement of their populations"

Authors' answer: Thanks!

• <u>Reviewer #3</u>: Lines 133-136: This filtering method is a bit hand-wavy. It's tough because the presenceonly sampling is from the literature, and sampling bias is therefore a huge issue - but it also means loss of information. Maybe elaborate on how the distribution "results" were evaluated (spatially? histogram of distance distributions?)?

<u>Authors' answer</u>: as pointed out by the reviewer, the loss of information is a necessary setback. The distribution results were visually inspected, in search of homogeneous distribution and lack of evident spatial clustering. In addition, we analyzed the distribution of pairwise distance between records. Prior to spatial filtering, a major cluster of pairwise distances is detected between 0 and 800km, and a 2nd cluster between 1600-2400 km. After spatial filtering and removing sampling bias, no major clusters are evident and frequency of pairwise distances is quite homogeneous.



• <u>*Reviewer #3*</u>: Figure 1 needs some work with choosing labels that can be distinguished from oneanother. A mix of light and dark colors would help.

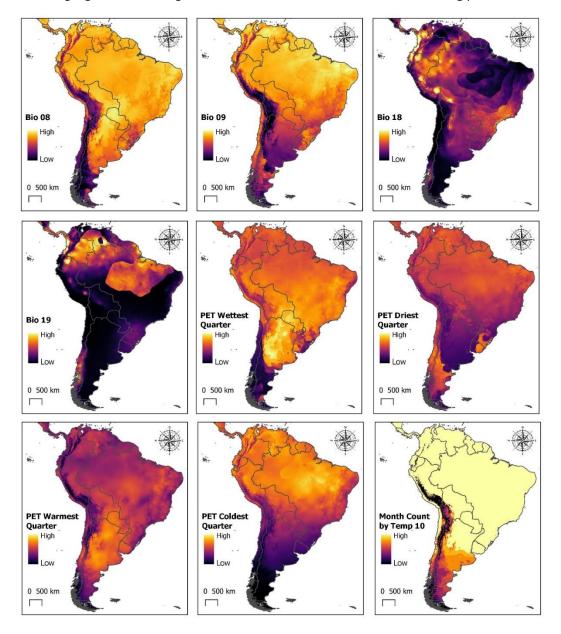
<u>Authors' answer</u>: Figure 1 was improved with light/dark colored labels, and the addition of a detail from the area where occurrences were clustered.

• *Reviewer #3*: Line 152: Need to define PET.

<u>Authors' answer</u>: The excluded variables are now listed in Table S2, and the abbreviation PET is defined there.

• <u>Reviewer #3</u>: Line 150-156: I wasn't aware of this! How interesting! Given the very large spatial scale and very high resolution of the WorldClim data, plus the filtering to remove points clustered close together, I would think this wouldn't be an issue in this case. I would therefor like to see how their inclusion changes the results.

<u>Authors' answer</u>: The issue is not dependent neither on spatial scale nor on sampling bias, but is observed as awkward environmental layers, with evident abrupt differences between neighboring pixels, which in turn should be continuous. Modeling with such data results in equally awkward predictions. For instance, see the following Figure, illustrating the nine variables excluded from our modeling procedure.



• <u>Reviewer #3</u>: Line 158-175: I'm intrigued by the a prior hypotheses, but I'd rather see a bit more information about them (eg., in a table?) - what species were they conducted in. Short of looking up each reference myself, I did look up Sinclair (2015) and it's not specifically on these species, but insects in general. A review of sorts. That is not made clear in the text here: it sounds like it was observed in this species. It makes sense that it's not in the study species as the authors are claiming to be doing the first study on it. But the prior evidence and wherefrom it derives would be helpful to the reader.

<u>Authors' answer</u>: following the reviewer's comment, we clarified the origin of the prior evidence (lines 164-165 and 167-168).

• *Reviewer #3*: Line 204-211: Nice!

Authors' answer: Thanks

• <u>Reviewer #3</u>: Results: My main concern with the methods and results, and thus conclusions drawn, is that only 3 of over 900 models met the criteria for inclusion: is there correction for multiple tests. This is likely just my unfamiliarity with the specific methods here, but I implore the authors to explicitly explain how the methods control for multiple tests to an audience who may look at this with my same concern. I simply fear that the combination of very low initial sample size combined with the multiple model approach puts the study at risk of explaining artifacts. I have not yet made it to the discussion, so perhaps this is discussed, but I'd just like to see that the authors address why these climatic factors, which are enormously correlated, might only have been found significant in the 3 models? Is it simply due to the other variable sets being orthogonal to Set1? If so, a table of the variable sets and the relative variance explained by each one would be helpful.

<u>Authors' answer</u>: the reviewer's concern is understandable. However, this is not a case of multiple tests (or multiple comparisons), where a set of statistical inferences are considered simultaneously. Indeed, it should be considered that models are not tested statistically against each other, but they are ranked in a model selection table in consideration of statistics of performance and complexity (comparison with a null model, omission rate, and AICc) of individual models. The best models are selected from this list based on their ranking, and 3 individual selection criteria (significance as compared with a null model of random prediction; of such significant models, omission criterion of < 5%; of the significant, low-omission models, $\Delta AICc < 2$) (for more details see Cobos *et al.* 2019, *PeerJ* 7:e6281). In certain way, the approach is quite similar to the one of multimodel inference (see Grueber *et al.* 2011, *J. Evol. Biol.* 24:699-711). Furthermore, correlation among variables was accounted during the variable selection process, while individual variables are not statistically tested in each model as in traditional linear models.