Dear Mr. Berigan,  
  
Thank you for submitting your manuscript to Biological Conservation. We regret to inform you that your paper is not acceptable for publication. We have completed the review of your manuscript and a summary is appended below. The reviewers have advised against publication of your manuscript and we must therefore reject it at this time. For your information and guidance, any specific comments explaining why we have reached this decision and those received from reviewers, if available, are listed at the end of this letter.

You have the option of resubmitting a substantially revised version of your paper, which would be considered as a new submission. If you decide to do this, you should refer to the reference number of the current paper and include a cover letter which explains in detail how the paper has been changed or not, in reply to the Editor and Reviewer comments.  
  
Thank you for giving us the opportunity to consider your work.

Kind regards,  
Biological Conservation  
  
  
Comments from the editors and reviewers:

Editor: We apologize for the time it has taken you to get a recommendation on this submission, and thank you for submitting your manuscript to Biological Conservation. As you will find, both reviewers appreciate the topics of your study, while the authors should extensively revise the manuscript by improving the clarity of the writing and providing the details of the analyses for reproducibility. Note that reviewer #2 provided the comments in a separate PDF file.

Reviewer #1: The manuscript "Multi-seasonal species distribution models better facilitate habitat conservation for a migratory bird" used a spatial decision support system (SDSS) to highlight areas of conservation importance for the American woodcock in Pennsylvania state relying on two separated seasonal models of the species (breeding and migratory). The general finding was that habitat preference differs between seasons, a result that should be considered for effective conservation decisions. The framework used here could inform management plans for the species and can be applied to other migratory species, as was nicely discussed in the manuscript. However, some concerns should be addressed before publication, mainly regarding terminology and methods.  
  
1. Using a "multi-season model" to describe this framework could be misleading. The first time I read the term, I thought about building a unique species distribution model incorporating several seasons instead of building separated single-season distribution models to be used as inputs for the SDSS. For example, multi-season occupancy models, also known as dynamic occupancy models, can incorporate in the model building per se occupancy changes over time. Your framework does not include any temporal variable in the model that allows you to estimate trends in species presence. Additionally, there are already examples of SDMs incorporating time in the modeling framework (e.g., Dynamic or time-specific SDMs), which could be an alternative to estimate the species' realized niche throughout the year. This does not mean that the methodology proposed here is invalid, but it should focus more on using SDSS with several single-season models, which is what you are doing here.  
  
2. Along the manuscript, you framed spatiotemporal transferability between seasons. Still, you never performed analyses that compared your model performance between single-season models (e.g., use the breeding season model to estimate the migratory season and vice versa). To check that your method solves this issue, compare the single-season model predictions. The fact that you find a low correlation between the geographic model predictions (as affirmed in Lines 295-296) does not mean poor transferability between them.  
  
3. There is evidence in the literature of the inconvenience of using just AUC values as the only evaluation metric for model performance, especially when there is no real absence to calculate commission error. Please complement the evaluation of your models with other metrics such as omission rate, Continuous Boyce index, partial ROC, etc. Incorporating more metrics will also let you know if the higher AUC values of Random Forest are not due to model overfitting.  
  
  
4. The normalization of the model outputs (Lines 195-196) should be accompanied by their interpretation with the occurrence records. The fact that you have 50% of your normalization does not inform you about where to find the species. For example, the minimum training presence could show how low a suitability value can be to be considered part of the species' potential habitat. Additionally, the fact that your seasonal models use different variable sets makes comparing them harder. For example, the minimum suitability value for a presence could differ between seasons.  
  
5. An important outcome of your study is a Shiny app. Unfortunately, it was unavailable at the moment of this review (your site was down).  
  
  
Minor comments:  
  
Line 156 & 159: Space between units.  
  
Figure 2. Please clarify what you mean by prioritization metrics.  
  
Figure 3. This figure is hard to understand; please consider using a response curve.  
  
  
  
Reviewer #2: Your paper has a lot of promise and is certainly on an important topic! I think this paper would benefit from restructuring, more background and more references that directly explain the key underpinnings of this paper, far more methodological details (particularly related to SDMs), and improvements in writing quality. I believe the study has a great deal of potential and the topic is very interesting and important, but at its present state, it is simply not ready for publication. With some work and restructuring, I am hopeful that it can be published in a reputable journal!

**Introduction**

Line 28: Transferability is a bit vague. Transferability in space? Time? Both? Just add a bit of

clarity here.

Lines 28-31: By “developed” do you mean trained on data in one area? I am pretty sure this is

what you mean, but the word developed is also a bit vague here. You could also say something

like “developed using data taken from one area.

”

Line 37: I am not sure what “carry-over effects” means

Lines 37-39: I understand what you’re trying to say, but it could be written better. Saying “By

building SDMs which focus solely on occurrence data…

” You are trying to say something like

“When SDMs are built only using occurrence data collected during the breeding season, we

disregard…

”

Lines 42-42: I was following until “during the decision-making process” – I am not sure what

this exactly means in this context. During modeling? During data collection? Prior to beginning

modeling?

Lines 43-44: Maybe give an example of a “spatial prioritization decision” because I do not fully

grasp what that means. I am guessing.

Lines 51-52: “with multiple spatial data layers, such as species distribution models…” This is a

bit confusing to me because the output of an SDM is typically a single layer of binary presence-

absence or habitat suitability (sometimes referred to as probability of presence). Do you mean

the layers of the predictor variables used in SDMs?

Lines 52-52: What are these circumstances? Such as…?

Line 51: SDSS is an interface? Or is it a framework? It sounds like it uses a GIS interface.

Line 54-55: Decision making process of what? Building the model? In conservation planning? I

apologize, but I am confused.

Lines 40-55: I do not fully grasp what a SDSS is and clearly that is key to understanding portions

of this manuscript. Including examples could help. Is it just a tool that lets you combine species

distribution models? In this case, from different seasons? It may be worth contrasting that with

ensemble SDMs because they do something similar, but with different SDM algorithms.

Line 56: Migratory birds are not sensitive to “cross-seasonal transferability.” Models of

migratory birds are sensitive to issues of cross-seasonal transferability.

Line 67-70: This is helpful for understanding why an SDSS approach is useful (combining

different SDMs into a single layer), but how this sentence is written is somewhat funny because

you’re suggesting that you *possibly* did this, but not directly saying. Be direct. Something like

“To account for these spatio-temporal differences in seasonal occurrence patterns, we used an

SDSS approach… “ or something like that.

Line 71: If this is going to be a paper aimed at informing conservation management, then there

should be more context provided on that earlier on in the Introduction. Right now, you only have

a single sentence that mentions how SDMs can be useful for conservation (Lines 27-28). There is

a lot more literature on how SDMs can be useful for conservation and this really should be much

more comprehensively addressed somewhere earlier in your Introduction.

Line 77: Did you aim to do it or did you do it? “We developed…

”

Line 81: “providing” should just be “provide” to be consistent with the tense of “evaluate”

**Methods**

Lines 85-89: Break this into two sentences

Lines 115-117: “distilled records to presence or likely absence…” is ambiguous. Needs more

detail. What was specifically considered presence vs. assumed absence. Assuming absence is

often contentious so this needs to be well-supported.

Lines 117-119: Were presences therefore broken down by sex? This seems like a potentially

problematic method. If this is considered acceptable to do based on the literature, please state

that more explicitly.

Lines 107-146: Data are clearly obtained from a variety of sources, which is great, but also

makes this paragraph particularly difficult to follow. I would suggest making a table that states

the data source and the sample size of each data source (and maybe other relevant information

too, if applicable). Alternatively, you could write this paragraph more clearly and that might

solve the problem.

Line 149: Use a word other than “accommodate” I am not sure what it really implies in this

context.

Line 156: Were these radii selected for a specific reason? I also think the choice of using multiple

spatial scales should be explained/justified. All of your results are with respect to a specific

spatial scale, but the use of different spatial scales is only mentioned very briefly here.

I think you should make a whole portion of your Methods focused on analyses at different spatial

scales because as I am reading the Discussion, I am confused about the use of spatial scales and

if/how these models were combined?

Lines 160-167: You need to include more detail on this “pilot evaluation” – were all variables

used in all three algorithms? Did you use different numbers of presence and either background or

pseudo-absence points? These different parameterizations affect model performance (i.e., AUC)

and therefore should be discussed since it affected your final choice of algorithm. Additionally, I

would reference the literature here to acknowledge the potential pitfalls of only using AUC to

assess model performance. Nowadays, using metrics like TSS, sensitivity, specificity, Kappa, etc.

is often considered a good idea. Ultimately, you need to include far more detail on how these

models were implemented.

Lines 168-174: You need to provide more detail on random forest classifiers designed for

clustered data vs. a traditional random forest classifier. I use random forest regularly and am not

familiar with the clustered data method, so assume people that do not use RF are even less

familiar.

How many trees did you use? How many replicates? What was your mtry? Did you use this as a

presence-absence model? Presence-pseudoabsence? Presence-background? There should also be

a description of how RF works with respect to Classification and Regression Trees (CARTs). I

am not confident that the algorithm is well-understood by the authors and that is a red flag to me,

but this can easily be addressed by adding details.

Line 183-185: Was this done prior to selecting a model algorithm? It feels like it should have

been and if so, it should be mentioned sooner.

Lines 183-194: Generally, a bit confused by this paragraph… I don’t think I know what you

mean by “To avoid overwhelming final predictive models with highly correlated or

uninformative variables…” Do you mean, to remove correlated variables? But if so, this should

have been done much earlier on and so I am a little confused.

Lines 205-206: Why was variable importance not assessed here using the mean decrease in the

Gini index? Or, is that considered not easily interpretable? The mean decrease in the Gini index

is considered the best way to assess how strongly a given predictor variable influenced model

performance and therefore I do not understand why it was not included. I recognize that

interpreting the very jagged curves is not straightforward, but there is a lot of material out there

on how to interpret them.

Line 255: The link provided does not work.

**Results**

Lines 257-262: I think I would put all of this in a table.

Overall: I think your Results need to be broken into sections so that they are easier to follow. I

felt somewhat lost reading the Results. Sub-headings for each results section would make this

easier to follow.

Realistically, SDMs should be evaluated using more than one performance metric. That is, not

just AUC alone. There should be a statement in the methods somewhere acknowledging the

limitations of AUC. I would recommend calculating TSS at the very least, to also include in your

results.

Are there any results pertinent to the SDSS? The SDSS seems like such an important part of this

paper and it is in the Introduction, Methods, and Discussion, but no mention in the Results.

**Discussion**

Line 323: Does “select habitat at a finer scale during migration” mean that birds are more

particular about their habitat during migration? I am not sure how to interpret this phrase.

The structure of the Discussion is difficult for me to follow.

There is no discussion on the benefits/limitations of RF

SDSS is a large part of your Introduction and Methods, but it is only mentioned in passing in the

Discussion. This should be elaborated on further if it is important.

As of right now, it is difficult for me to properly assess the Discussion because the methods and

results are somewhat unclear to me.