

## What processes must we understand to forecast regional scale population dynamics?

Jesse R. Lasky, Mevin B. Hooten and Peter B. Adler

### Article citation details

*Proc. R. Soc. B* **287**: 20202219.

<http://dx.doi.org/10.1098/rspb.2020.2219>

### Review timeline

Original submission: 8 September 2020  
1st revised submission: 30 October 2020  
2nd revised submission: 11 November 2020  
Final acceptance: 12 November 2020

Note: Reports are unedited and appear as submitted by the referee. The review history appears in chronological order.

## Review History

### RSPB-2020-2219.R0 (Original submission)

#### Review form: Reviewer 1

##### Recommendation

Accept with minor revision (please list in comments)

##### Scientific importance: Is the manuscript an original and important contribution to its field?

Excellent

##### General interest: Is the paper of sufficient general interest?

Good

##### Quality of the paper: Is the overall quality of the paper suitable?

Excellent

##### Is the length of the paper justified?

Yes

##### Should the paper be seen by a specialist statistical reviewer?

Yes

**Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.**

No

**It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.**

**Is it accessible?**

No

**Is it clear?**

N/A

**Is it adequate?**

N/A

**Do you have any ethical concerns with this paper?**

No

### **Comments to the Author**

The authors address the question about which biological processes should be incorporated in prediction models in order to increase forecasting accuracy. This is a very timely and important discussion and the title of the manuscript is raising high expectations.

The manuscript is divided in two major sections, a first section describing processes and different levels of biological complexity (genetics, phenotypic complexity, biotic interactions and population demography) and a second section in which they mainly introduce a novel modelling framework making use of an hierarchical Bayesian approach that allows for the integration of genetic, phenotypic and demographic components via reaction norms.

Overall, the manuscript is clearly written and a nice and easy read. I will not comment on the details of the introduced modelling framework since this is not my focal expertise. From the perspective of an empirical biologist, I do however work closely with modelers and therefore want to raise some issues that might be worth considering for a minor revision:

1) The first section gives a nice overview of different biological processes that might be considered in modeling frameworks. The authors did a good job in combining many different perspectives, it feels like an open dialogue between experts of different disciplines. The only thing that I miss, is a statement of the recent advancements to include genomic data in eco-evo models (see e.g. Rudman et al 2017, NEE or Gienapp et al. 2017, TREE). Ideas also exist of how to use genomic data to model selection processes through indirect fitness proxies (Gienapp et al. 2019, EvolLet, Waldvogel et al. 2020, EvolLet).

2) As mentioned above, the manuscript is opened with a strong title and readers are expecting a clear-cut answer. Not surprisingly, though somehow a bit disappointing, the answer finally is that one should continue to assess all the different processes and then use a comparative approach to find out which processes in an individual case give strongest support to the forecast. This brings me to my second issue, something more general. When referring to prediction models to assess species'/population's responses to environmental change, these forecasts are mostly required for non-model organisms and wild populations that are mostly difficult to experimentation. It is therefore questionable if a modeling framework that still requires the full set of components, including phenotypic trait data, will finally provide an approach applicable to a wide range of taxa. The question is, how much accuracy we loose without phenotype data or better how we can circumvent such data by approximating traits via e.g. genetic relatedness. This being said, I don't want to criticize your approach and hope that you can make it accessible

through some kind of implementation (e.g. R package).

## Review form: Reviewer 2

### Recommendation

Reject – article is not of sufficient interest (we will consider a transfer to another journal)

**Scientific importance: Is the manuscript an original and important contribution to its field?**

Marginal

**General interest: Is the paper of sufficient general interest?**

Good

**Quality of the paper: Is the overall quality of the paper suitable?**

Marginal

**Is the length of the paper justified?**

No

**Should the paper be seen by a specialist statistical reviewer?**

No

**Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.**

No

**It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.**

**Is it accessible?**

N/A

**Is it clear?**

N/A

**Is it adequate?**

N/A

**Do you have any ethical concerns with this paper?**

No

### Comments to the Author

There is, and indeed has been since ecology began, great interest in predictive ecology. That is, in understanding the underlying processes in ecology so as to be able to predict how population dynamics might respond across environmental gradients. The authors review the literature and conclude that such predictions are only possible by linking several different disciplines, and provide an integrative, mathematical framework for doing so. However, they are unable to make any actual predictions or to test their model with actual data because no such comprehensive dataset exists. In that sense, I'm not sure how helpful this paper will be. As someone who would very much like to forecast regional scale population dynamics, I struggled to understand how this paper or their model could be applied to systems that I work with. I think few people would argue against the idea that combining genetics, demography, biotic interactions, etc. are all

important (often the statements in the first half of the paper seemed rather trite), but that the subsequent dimensionality is difficult to deal with. I'm not sure, as written, the paper will be very helpful in identifying what factors need to be measured to determine such population dynamics.

A few minor comments below.

Line numbers would be very helpful.

Abstract. "Biologists suggest that we must..." Specifically, three biologists (the authors) make that suggestion. I suggest changing to "We suggest..." It just sounds a little trite otherwise.

Likewise, I learned nothing at all from the Abstract. Sure, we need several different fields to forecast population dynamics, and it depends on the situation. Please incorporate some actual results into the Abstract. For instance, you emphasize the importance of reaction norms--that could go in here. Or tell us a little about your "Integrative, hierarchical model framework".

Overview.

Predictability isn't just an 'acid test' for biology. Ability to predict other situations is really a hallmark of science. This 'movement' has been ongoing for quite some time in that sense.

I also don't find it very surprising that population ecologists, community ecologists, physiologists and evolutionary biologists view their own fields as important. Of course they are.

2.2. Is this individual variation necessary genetic? There is growing awareness of epigenetic and related phenomena. Or what if a particular trait is developmentally programmed?

2.3 "Direct observation of individual responses to changing environments, especially under controlled conditions," But if variation is being affected by so many different factors, how can we control for them all?

5. Many managers are accurately able to predict future population dynamics, and so or many systems we already know "what processes must we understand to forecast regional scale population dynamics" Nonetheless, I found the Conclusion more helpful than the Abstract and should perhaps replace that.

## Review form: Reviewer 3

### **Recommendation**

Major revision is needed (please make suggestions in comments)

### **Scientific importance: Is the manuscript an original and important contribution to its field?**

Good

### **General interest: Is the paper of sufficient general interest?**

Good

### **Quality of the paper: Is the overall quality of the paper suitable?**

Acceptable

### **Is the length of the paper justified?**

Yes

**Should the paper be seen by a specialist statistical reviewer?**

No

**Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.**

No

**It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.**

**Is it accessible?**

N/A

**Is it clear?**

N/A

**Is it adequate?**

N/A

**Do you have any ethical concerns with this paper?**

No

#### **Comments to the Author**

The manuscript is a review paper with a perspective component, where the authors propose to review the arguments about which processes to include in mechanistic models for regional dynamics. They use these arguments to further propose a conceptual framework and a statistical framework for the modelling cycle (iterations of model evaluation with data followed by model improvement). The review on mechanistic models and the modelling cycle perspective are not entirely novel, as there are other reviews on the mechanistic models necessary to describe biodiversity dynamics and perspectives stressing the necessity of the data-model iteration cycles and assimilation of model uncertainty. However, I do like the argument that different fields like to stress the role of their pet processes and the fact that the authors then focus on regional dynamics. In this regard, I was all invested already from the overview. The authors then provide a review on how different mechanistic components have been modelled, but raised the point that these efforts have been mostly isolated. As a consequence, the authors propose an integrative model in a Bayesian framework. How the authors propose the integration of the different mechanistic components is the most novel and exciting bit of this manuscript. The text has a really nice flow and the proposed framework is very impressive. I do see this as an important way forward for more integrative models with a careful complexity appraisal. Still, I have some major comments, mostly at the review part and with the conclusion, as the proposed framework is quite nice already.

Major points:

- Section 2.1, second paragraph: this is all good, but high dimensionality also limits mechanistic approaches (including data-driven ones). Models depicting different processes can generate the same patterns. The identifiability or equifinality mentioned in the last sentence of the previous paragraph relates to this. So, the shortcoming presented in this subsection is to both approaches, not only phenomenological. This must be better presented, because right now this is not the best selling point for the sections that follow, if the authors already stress in the end of the first paragraph of section 2.1 that mechanistic models may not be identifiable due to inherent mechanistic complexity of nature.

- It is somewhat unclear by the end of Section 2.1 how the authors came up with the four aspects they decided to devote the following sections (also presented in Figure 1).

- Section 3.1: the advocacy for simpler models is well careful here. Indeed, if one wants to forecast population dynamics of a given species, a simpler model may serve best. However, the authors arguments read as a critic to the movement towards more complex models (although they end up proposing a fairly complex model anyway). Whereas complexification of models must be done careful as detailed in following sections, the motivation for more complex models normally includes going beyond population dynamics of a single species. Increases in model complexity should be accompanied by an increase in the number of emergent patterns. Say, adding interactions would mean we are now able to predict multiple species, entire communities and the interactions per se. Thus, the strive for more complex model goes together with the hope to increase the number of patterns to be explained \*simultaneously\*. This must be clearer in the section; right now it reads as researchers complexify models just to improve one particular pattern.

- I just want to make a general comment here to express how well designed the framework is. I like it very much, including the potential to integrate loci linkage and spatially-structured gene flow. I also like the resulting inference section. Well done.

- Conclusion: I suggest toning down the text because 1) it is fair to advocate for careful complexification based on data-driven model comparisons, but for this you actually propose a fairly complex model. Hence, you start with a complex model and if necessary, you simplify it. 2) where exactly are the computational methods, in terms of codes? You just \*mention\* potential ways how to the implement the model within a Bayesian framework. I would not see this as 'providing' the computational methods. In this regard, do the authors have some code with this framework that can be provided as supporting information? 3) I would rather argue that you pretty much answered the question posed by proposing an integration of i) allelic model with gene flow (i.e. with spatial structure); ii) phenotypic model; iii) demographic model. Whether one sticks to the full proposed complexity or to a simpler version only the data and model comparison will tell (in this I totally agree). So, you have a clear proposal of maximum complexity, so be honest about this in the conclusion! 4) The interspecific biotic interactions (section 2.4) constitute the least developed component of the proposed framework and this must be mentioned in the conclusion as well.

Minor points:

- Section 2.1, first paragraph: The authors stress the role of temporal transferability, but they could also stress spatial transferability problems, which are particularly problematic for invasive species.

- Section 2.2., last paragraph: mutations might be of lesser importance, but if regional models intend to model range dynamics under changing environments, mutations happening in the expanding front may be fixed. There are plenty of models showing how mutations 'surf' on the expanding front.

- Section 2.3, second paragraph: True that machine learning algorithms must be taken with a grain of salt in terms of interpretability. This is true for ML-based optimization of either phenomenological or mechanistic models. Key here would be that ML-based techniques are only as good as their training, so to be able to predict a system in the future, the future conditions must be somehow integrated in the training data (e.g. by including particular geographic regions, or knowledge from past conditions). The usefulness of ML-based predictions (or any optimization method, including Bayesian frameworks) should be, in any case, tested via virtual ecologist frameworks.

- Section 2.4, second paragraph: note that abiotic change, variability or disturbances should work as coexistence mechanism (as an equalizing mechanism) by decreasing fitness

advantages of stronger competitors or specialists. This is particularly true for NOVEL disturbances/changes, to which an assemblage has not yet had time to evolve a response. So, I liked to see something in these lines in the text, but I missed a bit an explicit mention of coexistence theories here.

- Section 2.4: this section seems to focus a lot on the role of how changing abiotic conditions affect biotic interactions. I would further suggest the authors to complement that with how changing biotic conditions affect biotic interactions as well. This is particularly important considering exploitation by humans or human-induced invasions. In this aspect, missing species (that go locally extinct) or novel species should pose enough change to influence local/regional biotic interactions. There should be plenty of interaction network studies on this topic. Also the complexity of high-order interactions and multiplexes are current expanding fields of research that deserve some attention, or at least, some mention here. The authors mention 'a problem or high dimensionality' in the context of multiple community members, but not explicitly about pairwise vs. high-order interactions, which is obviously a bigger challenge.
- Figure 3: please indicate in the figure caption where to find parameter explanations (e.g, refer to the text or tables). Consider making a figure version with the equations explicitly.

## Decision letter (RSPB-2020-2219.R0)

14-Oct-2020

Dear Dr Lasky:

Your manuscript has now been peer reviewed and the reviewers' comments (not including confidential comments to the Editor) are included at the end of this email for your reference. As you will see, the reviewers have rather divided opinions on the value of the ms, but one thing all agree on is that it is very well written and accessible. Referee 1, as an empiricist who works with modellers, praises the clarity of the explanation and I, as someone in a quite different field, echo that sentiment. Referee 2's view, however, could be summarised as: we know about the value of such an integrative approach but the bottom line is that no-one has yet collected all the necessary data for even one system, so where does this review get us? (The bluntness is all mine -- I am reading between the referee's lines.) Referee 3 has similar misgivings, but thinks that the review would still be valuable if their major concerns could be addressed. So, there are no technical issues that could not be addressed, and it comes down to my judgement call on whether this framework has a realistic chance of implementation or is pie-in-the-sky. Well, I like aspirational reviews and someone may have the ambition (and time and money) to collect all the necessary data to implement the framework, so I would like to invite you to revise your manuscript to address the referees' concerns.

We do not allow multiple rounds of revision so we urge you to make every effort to fully address all of the comments at this stage. If deemed necessary, your manuscript will be sent back to one or more of the original reviewers for assessment. If the original reviewers are not available we may invite new reviewers. Please note that we cannot guarantee eventual acceptance of your manuscript at this stage.

To submit your revision please log into <http://mc.manuscriptcentral.com/prsb> and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions", click on "Create a Revision". Your manuscript number has been appended to denote a revision.

When submitting your revision please upload a file under "Response to Referees" in the "File Upload" section. This should document, point by point, how you have responded to the reviewers' and Editors' comments, and the adjustments you have made to the manuscript. We require a copy of the manuscript with revisions made since the previous version marked as 'tracked changes' to be included in the 'response to referees' document.

Your main manuscript should be submitted as a text file (doc, txt, rtf or tex), not a PDF. Your figures should be submitted as separate files and not included within the main manuscript file.

When revising your manuscript you should also ensure that it adheres to our editorial policies (<https://royalsociety.org/journals/ethics-policies/>). You should pay particular attention to the following:

#### Research ethics:

If your study contains research on humans please ensure that you detail in the methods section whether you obtained ethical approval from your local research ethics committee and gained informed consent to participate from each of the participants.

#### Use of animals and field studies:

If your study uses animals please include details in the methods section of any approval and licences given to carry out the study and include full details of how animal welfare standards were ensured. Field studies should be conducted in accordance with local legislation; please include details of the appropriate permission and licences that you obtained to carry out the field work.

#### Data accessibility and data citation:

It is a condition of publication that you make available the data and research materials supporting the results in the article (<https://royalsociety.org/journals/authors/author-guidelines/#data>). Datasets should be deposited in an appropriate publicly available repository and details of the associated accession number, link or DOI to the datasets must be included in the Data Accessibility section of the article (<https://royalsociety.org/journals/ethics-policies/data-sharing-mining/>). Reference(s) to datasets should also be included in the reference list of the article with DOIs (where available).

In order to ensure effective and robust dissemination and appropriate credit to authors the dataset(s) used should also be fully cited and listed in the references.

If you wish to submit your data to Dryad (<http://datadryad.org/>) and have not already done so you can submit your data via this link

[http://datadryad.org/submit?journalID=RSPB&manu=\(Document not available\)](http://datadryad.org/submit?journalID=RSPB&manu=(Document not available)), which will take you to your unique entry in the Dryad repository.

If you have already submitted your data to dryad you can make any necessary revisions to your dataset by following the above link.

For more information please see our open data policy <http://royalsocietypublishing.org/data-sharing>.

#### Electronic supplementary material:

All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI. Please try to submit all supplementary material as a single file.



Online supplementary material will also carry the title and description provided during submission, so please ensure these are accurate and informative. Note that the Royal Society will not edit or typeset supplementary material and it will be hosted as provided. Please ensure that the supplementary material includes the paper details (authors, title, journal name, article DOI). Your article DOI will be 10.1098/rspb.[paper ID in form xxxx.xxxx e.g. 10.1098/rspb.2016.0049].

Please submit a copy of your revised paper within three weeks. If we do not hear from you within this time your manuscript will be rejected. If you are unable to meet this deadline please let us know as soon as possible, as we may be able to grant a short extension.

Thank you for submitting your manuscript to Proceedings B; we look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Best wishes,  
Innes Cuthill

Prof. Innes Cuthill  
Reviews Editor, Proceedings B  
mailto:proceedingsb@royalsociety.org

Reviewer(s)' Comments to Author:  
Referee: 1

#### Comments to the Author(s)

The authors address the question about which biological processes should be incorporated in prediction models in order to increase forecasting accuracy. This is a very timely and important discussion and the title of the manuscript is raising high expectations.

The manuscript is divided in two major sections, a first section describing processes and different levels of biological complexity (genetics, phenotypic complexity, biotic interactions and population demography) and a second section in which they mainly introduce a novel modelling framework making use of an hierarchical Bayesian approach that allows for the integration of genetic, phenotypic and demographic components via reaction norms.

Overall, the manuscript is clearly written and a nice and easy read. I will not comment on the details of the introduced modelling framework since this is not my focal expertise. From the perspective of an empirical biologist, I do however work closely with modelers and therefore want to raise some issues that might be worth considering for a minor revision:

1) The first section gives a nice overview of different biological processes that might be considered in modeling frameworks. The authors did a good job in combining many different perspectives, it feels like an open dialogue between experts of different disciplines. The only thing that I miss, is a statement of the recent advancements to include genomic data in eco-evo models (see e.g. Rudman et al 2017, NEE or Gienapp et al. 2017, TREE). Ideas also exist of how to use genomic data to model selection processes through indirect fitness proxies (Gienapp et al. 2019, EvolLet, Waldvogel et al. 2020, EvolLet).

2) As mentioned above, the manuscript is opened with a strong title and readers are expecting a clear-cut answer. Not surprisingly, though somehow a bit disappointing, the answer finally is that one should continue to assess all the different processes and then use a comparative approach to find out which processes in an individual case give strongest support to the forecast. This brings me to my second issue, something more general. When referring to prediction models to assess species'/population's responses to environmental change, these forecasts are mostly required for non-model organisms and wild populations that are mostly difficult to experimentation. It is therefore questionable if a modeling framework that still requires the full set of components, including phenotypic trait data, will finally provide an approach applicable to a wide range of taxa. The question is, how much accuracy we lose without phenotype data or better how we can circumvent such data by approximating traits via e.g. genetic relatedness.

This being said, I don't want to criticize your approach and hope that you can make it accessible through some kind of implementation (e.g. R package).

Referee: 2

Comments to the Author(s)

There is, and indeed has been since ecology began, great interest in predictive ecology. That is, in understanding the underlying processes in ecology so as to be able to predict how population dynamics might respond across environmental gradients. The authors review the literature and conclude that such predictions are only possible by linking several different disciplines, and provide an integrative, mathematical framework for doing so. However, they are unable to make any actual predictions or to test their model with actual data because no such comprehensive dataset exists. In that sense, I'm not sure how helpful this paper will be. As someone who would very much like to forecast regional scale population dynamics, I struggled to understand how this paper or their model could be applied to systems that I work with. I think few people would argue against the idea that combining genetics, demography, biotic interactions, etc. are all important (often the statements in the first half of the paper seemed rather trite), but that the subsequent dimensionality is difficult to deal with. I'm not sure, as written, the paper will be very helpful in identifying what factors need to be measured to determine such population dynamics.

A few minor comments below.

Line numbers would be very helpful.

Abstract. "Biologists suggest that we must..." Specifically, three biologists (the authors) make that suggestion. I suggest changing to "We suggest..." It just sounds a little trite otherwise.

Likewise, I learned nothing at all from the Abstract. Sure, we need several different fields to forecast population dynamics, and it depends on the situation. Please incorporate some actual results into the Abstract. For instance, you emphasize the importance of reaction norms--that could go in here. Or tell us a little about your "Integrative, hierarchical model framework".

Overview.

Predictability isn't just an 'acid test' for biology. Ability to predict other situations is really a hallmark of science. This 'movement' has been ongoing for quite some time in that sense.

I also don't find it very surprising that population ecologists, community ecologists, physiologists and evolutionary biologists view their own fields as important. Of course they are.

2.2. Is this individual variation necessary genetic? There is growing awareness of epigenetic and related phenomena. Or what if a particular trait is developmentally programmed?

2.3 "Direct observation of individual responses to changing environments, especially under controlled conditions," But if variation is being affected by so many different factors, how can we control for them all?

5. Many managers are accurately able to predict future population dynamics, and so or many systems we already know "what processes must we understand to forecast regional scale population dynamics" Nonetheless, I found the Conclusion more helpful than the Abstract and should perhaps replace that.

Referee: 3

Comments to the Author(s)

The manuscript is a review paper with a perspective component, where the authors propose to review the arguments about which processes to include in mechanistic models for regional

dynamics. They use these arguments to further propose a conceptual framework and a statistical framework for the modelling cycle (iterations of model evaluation with data followed by model improvement). The review on mechanistic models and the modelling cycle perspective are not entirely novel, as there are other reviews on the mechanistic models necessary to describe biodiversity dynamics and perspectives stressing the necessity of the data-model iteration cycles and assimilation of model uncertainty. However, I do like the argument that different fields like to stress the role of their pet processes and the fact that the authors then focus on regional dynamics. In this regard, I was all invested already from the overview. The authors then provide a review on how different mechanistic components have been modelled, but raised the point that these efforts have been mostly isolated. As a consequence, the authors propose an integrative model in a Bayesian framework. How the authors propose the integration of the different mechanistic components is the most novel and exciting bit of this manuscript. The text has a really nice flow and the proposed framework is very impressive. I do see this as an important way forward for more integrative models with a careful complexity appraisal. Still, I have some major comments, mostly at the review part and with the conclusion, as the proposed framework is quite nice already.

Major points:

- Section 2.1, second paragraph: this is all good, but high dimensionality also limits mechanistic approaches (including data-driven ones). Models depicting different processes can generate the same patterns. The identifiability or equifinality mentioned in the last sentence of the previous paragraph relates to this. So, the shortcoming presented in this subsection is to both approaches, not only phenomenological. This must be better presented, because right now this is not the best selling point for the sections that follow, if the authors already stress in the end of the first paragraph of section 2.1 that mechanistic models may not be identifiable due to inherent mechanistic complexity of nature.
- It is somewhat unclear by the end of Section 2.1 how the authors came up with the four aspects they decided to devote the following sections (also presented in Figure 1).
- Section 3.1: the advocacy for simpler models is well careful here. Indeed, if one wants to forecast population dynamics of a given species, a simpler model may serve best. However, the authors arguments read as a critic to the movement towards more complex models (although they end up proposing a fairly complex model anyway). Whereas complexification of models must be done careful as detailed in following sections, the motivation for more complex models normally includes going beyond population dynamics of a single species. Increases in model complexity should be accompanied by an increase in the number of emergent patterns. Say, adding interactions would mean we are now able to predict multiple species, entire communities and the interactions per se. Thus, the strive for more complex model goes together with the hope to increase the number of patterns to be explained \*simultaneously\*. This must be clearer in the section; right now it reads as researchers complexify models just to improve one particular pattern.
- I just want to make a general comment here to express how well designed the framework is. I like it very much, including the potential to integrate loci linkage and spatially-structured gene flow. I also like the resulting inference section. Well done.
- Conclusion: I suggest toning down the text because 1) it is fair to advocate for careful complexification based on data-driven model comparisons, but for this you actually propose a fairly complex model. Hence, you start with a complex model and if necessary, you simplify it. 2) where exactly are the computational methods, in terms of codes? You just \*mention\* potential ways how to implement the model within a Bayesian framework. I would not see this as 'providing' the computational methods. In this regard, do the authors have some code with this framework that can be provided as supporting information? 3) I would rather argue that you pretty much answered the question posed by proposing an integration of i) allelic model with

gene flow (i.e. with spatial structure); ii) phenotypic model; iii) demographic model. Whether one sticks to the full proposed complexity or to a simpler version only the data and model comparison will tell (in this I totally agree). So, you have a clear proposal of maximum complexity, so be honest about this in the conclusion! 4) The interspecific biotic interactions (section 2.4) constitute the least developed component of the proposed framework and this must be mentioned in the conclusion as well.

Minor points:

- Section 2.1, first paragraph: The authors stress the role of temporal transferability, but they could also stress spatial transferability problems, which are particularly problematic for invasive species.

- Section 2.2., last paragraph: mutations might be of lesser importance, but if regional models intend to model range dynamics under changing environments, mutations happening in the expanding front may be fixed. There are plenty of models showing how mutations 'surf' on the expanding front.

- Section 2.3, second paragraph: True that machine learning algorithms must be taken with a grain of salt in terms of interpretability. This is true for ML-based optimization of either phenomenological or mechanistic models. Key here would be that ML-based techniques are only as good as their training, so to be able to predict a system in the future, the future conditions must be somehow integrated in the training data (e.g. by including particular geographic regions, or knowledge from past conditions). The usefulness of ML-based predictions (or any optimization method, including Bayesian frameworks) should be, in any case, tested via virtual ecologist frameworks.

- Section 2.4, second paragraph: note that abiotic change, variability or disturbances should work as coexistence mechanism (as an equalizing mechanism) by decreasing fitness advantages of stronger competitors or specialists. This is particularly true for NOVEL disturbances/changes, to which an assemblage has not yet had time to evolve a response. So, I liked to see something in these lines in the text, but I missed a bit an explicit mention of coexistence theories here.

- Section 2.4: this section seems to focus a lot on the role of how changing abiotic conditions affect biotic interactions. I would further suggest the authors to complement that with how changing biotic conditions affect biotic interactions as well. This is particularly important considering exploitation by humans or human-induced invasions. In this aspect, missing species (that go locally extinct) or novel species should pose enough change to influence local/regional biotic interactions. There should be plenty of interaction network studies on this topic. Also the complexity of high-order interactions and multiplexes are current expanding fields of research that deserve some attention, or at least, some mention here. The authors mention 'a problem or high dimensionality' in the context of multiple community members, but not explicitly about pairwise vs. high-order interactions, which is obviously a bigger challenge.

- Figure 3: please indicate in the figure caption where to find parameter explanations (e.g, refer to the text or tables). Consider making a figure version with the equations explicitly.

## Author's Response to Decision Letter for (RSPB-2020-2219.R0)

See Appendix A.

# RSPB-2020-2219.R1 (Revision)

## Review form: Reviewer 2

### Recommendation

Accept as is

### Scientific importance: Is the manuscript an original and important contribution to its field?

Marginal

### General interest: Is the paper of sufficient general interest?

Excellent

### Quality of the paper: Is the overall quality of the paper suitable?

Acceptable

### Is the length of the paper justified?

Yes

### Should the paper be seen by a specialist statistical reviewer?

Yes

### Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.

No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

#### Is it accessible?

N/A

#### Is it clear?

N/A

#### Is it adequate?

N/A

### Do you have any ethical concerns with this paper?

No

### Comments to the Author

I have read through the response to the reviewers and the edited version of the manuscript, and appreciate that the authors have responded to the various concerns. I especially appreciated the new 'feasibility' section. I wonder if rather than focusing on a species that we have a relatively short history of managing (an invasive grass), it would be worth considering managed species (both 'wildlife' such as ungulates or waterfowl or fisheries). While fisheries are often stochastic, many wildlife management regimes do an excellent job of predicting future harvest of their stocks. At the other extreme, indigenous peoples in the Northwest harvested berries and camas lilies through many large environmental fluctuations, and were able to forecast upcoming responses without resorting to complex math.

## Decision letter (RSPB-2020-2219.R1)

11-Nov-2020

Dear Dr Lasky

I am pleased to inform you that your revised manuscript RSPB-2020-2219.R1 entitled "What processes must we understand to forecast regional scale population dynamics?" has been accepted for publication in Proceedings B.

The referee has recommended publication, but also suggest some minor revisions to your manuscript. I leave it at your discretion whether you want to change any of your manuscript in the light of these. Because the schedule for publication is very tight, it is a condition of publication that you submit the revised version of your manuscript within 7 days. If you do not think you will be able to meet this date please let us know.

To upload the final version of your manuscript, log into <https://mc.manuscriptcentral.com/prsb> and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision. You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript and upload a new version through your Author Centre.

When submitting your revised manuscript, you will be able to respond to the comments made by the referee and upload a file "Response to Referees". You can use this to document any changes you make to the original manuscript. We require a copy of the manuscript with revisions made since the previous version marked as 'tracked changes' to be included in the 'response to referees' document.

Before uploading your revised files please make sure that you have:

- 1) A text file of the manuscript (doc, txt, rtf or tex), including the references, tables (including captions) and figure captions. Please remove any tracked changes from the text before submission. PDF files are not an accepted format for the "Main Document".
- 2) A separate electronic file of each figure (tiff, EPS or print-quality PDF preferred). The format should be produced directly from original creation package, or original software format. PowerPoint files are not accepted.
- 3) Electronic supplementary material: this should be contained in a separate file and where possible, all ESM should be combined into a single file. All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

Online supplementary material will also carry the title and description provided during submission, so please ensure these are accurate and informative. Note that the Royal Society will not edit or typeset supplementary material and it will be hosted as provided. Please ensure that the supplementary material includes the paper details (authors, title, journal name, article DOI). Your article DOI will be 10.1098/rspb.[paper ID in form xxxx.xxxx e.g. 10.1098/rspb.2016.0049].

- 4) A media summary: a short non-technical summary (up to 100 words) of the key findings/importance of your manuscript.

#### 5) Data accessibility section and data citation

It is a condition of publication that data supporting your paper are made available either in the electronic supplementary material or through an appropriate repository (<https://royalsociety.org/journals/authors/author-guidelines/#data>).

In order to ensure effective and robust dissemination and appropriate credit to authors the dataset(s) used should be fully cited. To ensure archived data are available to readers, authors should include a 'data accessibility' section immediately after the acknowledgements section. This should list the database and accession number for all data from the article that has been made publicly available, for instance:

- DNA sequences: Genbank accessions F234391-F234402
- Phylogenetic data: TreeBASE accession number S9123
- Final DNA sequence assembly uploaded as online supplemental material
- Climate data and MaxEnt input files: Dryad doi:10.5521/dryad.12311

NB. From April 1 2013, peer reviewed articles based on research funded wholly or partly by RCUK must include, if applicable, a statement on how the underlying research materials – such as data, samples or models – can be accessed. This statement should be included in the data accessibility section.

If you wish to submit your data to Dryad (<http://datadryad.org/>) and have not already done so you can submit your data via this link

[http://datadryad.org/submit?journalID=RSPB&manu=\(Document not available\)](http://datadryad.org/submit?journalID=RSPB&manu=(Document not available)) which will take you to your unique entry in the Dryad repository. If you have already submitted your data to dryad you can make any necessary revisions to your dataset by following the above link. Please see <https://royalsociety.org/journals/ethics-policies/data-sharing-mining/> for more details.

6) For more information on our Licence to Publish, Open Access, Cover images and Media summaries, please visit <https://royalsociety.org/journals/authors/author-guidelines/>.

Once again, thank you for submitting your manuscript to Proceedings B and I look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Best wishes,

Innes

Prof. Innes Cuthill

Reviews Editor, Proceedings B

mailto: [proceedingsb@royalsociety.org](mailto:proceedingsb@royalsociety.org)

Reviewer(s)' Comments to Author:

Referee: 2

#### Comments to the Author(s)

I have read through the response to the reviewers and the edited version of the manuscript, and appreciate that the authors have responded to the various concerns. I especially appreciated the new 'feasibility' section. I wonder if rather than focusing on a species that we have a relatively short history of managing (an invasive grass), it would be worth considering managed species (both 'wildlife' such as ungulates or waterfowl or fisheries). While fisheries are often stochastic, many wildlife management regimes do an excellent job of predicting future harvest of their stocks. At the other extreme, indigenous peoples in the Northwest harvested berries and camas lilies through many large environmental fluctuations, and were able to forecast upcoming responses without resorting to complex math.

## Decision letter (RSPB-2020-2219.R2)

12-Nov-2020

Dear Dr Lasky

I am pleased to inform you that your manuscript entitled "What processes must we understand to forecast regional scale population dynamics?" has been accepted for publication in Proceedings B.

You can expect to receive a proof of your article from our Production office in due course, please check your spam filter if you do not receive it. PLEASE NOTE: you will be given the exact page length of your paper which may be different from the estimation from Editorial and you may be asked to reduce your paper if it goes over the 10 page limit.

If you are likely to be away from e-mail contact during this period, let us know. Due to rapid publication and an extremely tight schedule, if comments are not received, we may publish the paper as it stands.

If you have any queries regarding the production of your final article or the publication date please contact [procb\\_proofs@royalsociety.org](mailto:procb_proofs@royalsociety.org)

Your article has been estimated as being 10 pages long. Our Production Office will be able to confirm the exact length at proof stage.

### Open access

You are invited to opt for open access via our author pays publishing model. Payment of open access fees will enable your article to be made freely available via the Royal Society website as soon as it is ready for publication. For more information about open access publishing please visit our website at [http://royalsocietypublishing.org/site/authors/open\\_access.xhtml](http://royalsocietypublishing.org/site/authors/open_access.xhtml).

The open access fee is £1,700 per article (plus VAT for authors within the EU). If you wish to opt for open access then please let us know as soon as possible.

### Paper charges

An e-mail request for payment of any related charges will be sent out shortly. The preferred payment method is by credit card; however, other payment options are available.

### Electronic supplementary material:

All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

You are allowed to post any version of your manuscript on a personal website, repository or preprint server. However, the work remains under media embargo and you should not discuss it with the press until the date of publication. Please visit <https://royalsociety.org/journals/ethics-policies/media-embargo> for more information.

Thank you for your fine contribution. On behalf of the Editors of the Proceedings B, we look forward to your continued contributions to the Journal.

Sincerely,  
Proceedings B  
<mailto:proceedingsb@royalsociety.org>



# Appendix A

## Response to reviews

Jesse R. Lasky, Mevin B. Hooten, Peter B. Adler

October 29, 2020

### 1 Editor

Your manuscript has now been peer reviewed and the reviewers' comments (not including confidential comments to the Editor) are included at the end of this email for your reference. As you will see, the reviewers have rather divided opinions on the value of the ms, but one thing all agree on is that it is very well written and accessible. Referee 1, as an empiricist who works with modellers, praises the clarity of the explanation and I, as someone in a quite different field, echo that sentiment. Referee 2's view, however, could be summarised as: we know about the value of such an integrative approach but the bottom line is that no-one has yet collected all the necessary data for even one system, so where does this review get us? (The bluntness is all mine – I am reading between the referee's lines.) Referee 3 has similar misgivings, but thinks that the review would still be valuable if their major concerns could be addressed. So, there are no technical issues that could not be addressed, and it comes down to my judgement call on whether this framework has a realistic chance of implementation or is pie-in-the-sky. Well, I like aspirational reviews and someone may have the ambition (and time and money) to collect all the necessary data to implement the framework, so I would like to invite you to revise your manuscript to address the referees' concerns.

Author response: We thank the editor for their input and for the chance to edit our manuscript in response to these comments. Below we provide point-by-point responses to the referees. Also, we have added text to the manuscript to discuss the prospects for implementing our proposed studies in the real world.

### 2 Referee 1

The authors address the question about which biological processes should be incorporated in prediction models in order to increase forecasting accuracy. This is a very timely and important discussion and the title of the manuscript is raising high expectations. The manuscript is divided in two major sections, a first section describing processes and different levels of biological complexity (genetics, phenotypic complexity, biotic interactions and population demography) and a second section in which they mainly introduce a novel modelling framework making use of an hierarchical Bayesian approach that allows for the integration of genetic, phenotypic and demographic components via reaction norms. Overall, the manuscript is clearly written and a nice and easy read. I will not comment on the details of the introduced modelling framework since this is not my focal expertise. From the perspective of an empirical biologist, I do however work closely with modelers and therefore want to raise some issues that might be worth considering for a minor revision:

Author response: We thank the referee for their support.

1) The first section gives a nice overview of different biological processes that might be considered in modeling frameworks. The authors did a good job in combining many different perspectives, it feels like

an open dialogue between experts of different disciplines. The only thing that I miss, is a statement of the recent advancements to include genomic data in eco-evo models (see e.g. Rudman et al 2017, NEE or Gienapp et al. 2017, TREE). Ideas also exist of how to use genomic data to model selection processes through indirect fitness proxies (Gienapp et al. 2019, EvolLet, Waldvogel et al. 2020, EvolLet).

Author response: We thank the referee for their support. We have added the Rudman et al. 2017 and Gienapp et al. 2019 citations in the appropriate section (2.2). We would like to note that we cited Gienapp et al. 2017, TREE in our original submission.

2) As mentioned above, the manuscript is opened with a strong title and readers are expecting a clear-cut answer. Not surprisingly, though somehow a bit disappointing, the answer finally is that one should continue to assess all the different processes and then use a comparative approach to find out which processes in an individual case give strongest support to the forecast. This brings me to my second issue, something more general. When referring to prediction models to assess species'/population's responses to environmental change, these forecasts are mostly required for non-model organisms and wild populations that are mostly difficult to experimentation. It is therefore questionable if a modeling framework that still requires the full set of components, including phenotypic trait data, will finally provide an approach applicable to a wide range of taxa. The question is, how much accuracy we lose without phenotype data or better how we can circumvent such data by approximating traits via e.g. genetic relatedness. This being said, I don't want to criticize your approach and hope that you can make it accessible through some kind of implementation (e.g. R package).

Author response: We agree that it will be difficult to carry out studies that are able to determine the importance of all of the suggested processes here for regional forecasting. In some systems it will be nearly impossible to dissect and model some of these processes. We have edited the abstract and several other points in the manuscript to clarify our intentions. We have added a paragraph (section 4.5), discussing potential study systems and highlighting how the spirit of our approach can be carried through studies of systems where certain types of data are unavailable.

### 3 Referee 2

There is, and indeed has been since ecology began, great interest in predictive ecology. That is, in understanding the underlying processes in ecology so as to be able to predict how population dynamics might respond across environmental gradients. The authors review the literature and conclude that such predictions are only possible by linking several different disciplines, and provide an integrative, mathematical framework for doing so. However, they are unable to make any actual predictions or to test their model with actual data because no such comprehensive dataset exists. In that sense, I'm not sure how helpful this paper will be.

Author response: The referee is correct that we are currently unable to implement our approach because of a lack of appropriate data. However, we are currently leading a large collaborative effort to collect these data for an annual plant of major interest for ecosystem management across western North America, *Bromus tectorum* (cheatgrass). There are a number of other systems amenable to this work and with enough community interest to gather the required data. For example, annual plants, fish (with shorter lifespans), and inter-tidal invertebrates are commonly studied for all of the processes described above. The model plant *Arabidopsis thaliana* has been studied extensively for genetic, physiological, and ecological purposes and is currently the subject of a large distributed experiment <https://grenenet.wordpress.com/>. Even long-lived organisms like trees have species that could

be appropriate for such a study, thanks to a wealth of existing ecological knowledge, existing long-term common garden experiments, and ongoing efforts to characterize range-wide genomic variation (e.g., work by Sally Aitken of UBC on conifers of western North America). We have added a paragraph (section 4.5) that mentions some of these efforts and systems.

As someone who would very much like to forecast regional scale population dynamics, I struggled to understand how this paper or their model could be applied to systems that I work with. I think few people would argue against the idea that combining genetics, demography, biotic interactions, etc. are all important (often the statements in the first half of the paper seemed rather trite), but that the subsequent dimensionality is difficult to deal with. I'm not sure, as written, the paper will be very helpful in identifying what factors need to be measured to determine such population dynamics.

Author response: We emphasize that combining all of these aspects in a model *may not* be the best route to prediction (see section 3.1). Additionally, we have added a paragraph (section 4.5) to highlight some candidate empirical systems to which our approach can be applied (and indeed which we will be applying in an ongoing project on the widespread grass *Bromus tectorum*).

A few minor comments below.  
Line numbers would be very helpful.

Author response: We have added line numbers.

Abstract. "Biologists suggest that we must..." Specifically, three biologists (the authors) make that suggestion. I suggest changing to "We suggest..." It just sounds a little trite otherwise.

Author response: To clarify, the goal of this sentence is to convey the calls by many biologists (authors excluded) to move in the direction of prediction based on process-oriented models, setting the stage for the thesis of our paper. We have edited the abstract to increase clarity on this point. We also emphasize that combining all of these aspects in a model may not be the best route to prediction (see section 3.1).

Likewise, I learned nothing at all from the Abstract. Sure, we need several different fields to forecast population dynamics, and it depends on the situation. Please incorporate some actual results into the Abstract. For instance, you emphasize the importance of reaction norms—that could go in here. Or tell us a little about your "Integrative, hierarchical model framework".

Author response: We have rewritten part of the Abstract, and included additional details about our goals as well as our proposed approach, including reaction norms and the hierarchical Bayes statistical framework.

Overview.

Predictability isn't just an 'acid test' for biology. Ability to predict other situations is really a hallmark of science. This 'movement' has been ongoing for quite some time in that sense.

Author response: We agree with the referee on this point and we have omitted the word ‘movement’ at this location in the text and clarify that we see this as a ‘growing emphasis.’

I also don't find it very surprising that population ecologists, community ecologists, physiologists and evolutionary biologists view their own fields as important. Of course they are.

Author response: We agree that it is unsurprising to claim that all of these fields are important in their own right. However, we emphasize throughout and especially in section 3.1 that including all of these processes in predictive models may not be desirable. This latter point has not been widely made with respect to regional population forecasts.

2.2. Is this individual variation necessary genetic? There is growing awareness of epigenetic and related phenomena. Or what if a particular trait is developmentally programmed?

Author response: Individual variation can be plastic as well as genetic, but in referring to ‘this individual variation’ we presume the reviewer is referring to transgenerational heritable variation. Persistently transgenerationally inherited epigenetic variation is still a phenomenon with little documentation and evidence for a major role in adaptation in any system and thus we believe not appropriate to include at this point. Developmental programming is a form of plasticity and thus can be accommodated in our model as described in sections 3.3 and 4.2, and described by reaction norms as in Fig 2.

2.3 “Direct observation of individual responses to changing environments, especially under controlled conditions,” But if variation is being affected by so many different factors, how can we control for them all?

Author response: We agree with the referee that experimental manipulation of a large number of environmental parameters is highly challenging. Nevertheless, ecophys/developmental models based on such observations have proven to have high utility for prediction across environments for crops. Additionally, we have added a sentence in section 4.5 along the lines of the reviewer’s critique.

5. Many managers are accurately able to predict future population dynamics, and so or many systems we already know “what processes must we understand to forecast regional scale population dynamics” Nonetheless, I found the Conclusion more helpful than the Abstract and should perhaps replace that.

Author response: Here we disagree with the reviewer. We know of few examples where managers, or researchers, can accurately predict population dynamics. Fisheries applications may have the longest track record, but even in that literature forecast skill is often disappointing. More generally, time-series approaches show some promise for predicting short-term trajectories of local populations, often ignoring environmental drivers. Forecasting impacts of environmental change over longer time scales and/or at regional scales is even more challenging, and in most cases we have no measure of skill because long-term projections are not validated. We would be very interested in examples of successful forecasts of this kind if the reviewer has some in mind!

## 4 Referee 3

The manuscript is a review paper with a perspective component, where the authors propose to review the arguments about which processes to include in mechanistic models for regional dynamics. They use these arguments to further propose a conceptual framework and a statistical framework for the modelling cycle (iterations of model evaluation with data followed by model improvement). The review on mechanistic models and the modelling cycle perspective are not entirely novel, as there are other reviews on the mechanistic models necessary to describe biodiversity dynamics and perspectives stressing the necessity of the data-model iteration cycles and assimilation of model uncertainty. However, I do like the argument that different fields like to stress the role of their pet processes and the fact that the authors then focus on regional dynamics. In this regard, I was all invested already from the overview. The authors then provide a review on how different mechanistic components have been modelled, but raised the point that these efforts have been mostly isolated. As a consequence, the authors propose an integrative model in a Bayesian framework. How the authors propose the integration of the different mechanistic components is the most novel and exciting bit of this manuscript. The text has a really nice flow and the proposed framework is very impressive. I do see this as an important way forward for more integrative models with a careful complexity appraisal. Still, I have some major comments, mostly at the review part and with the conclusion, as the proposed framework is quite nice already.

Author response: We thank the referee for their appreciation of our work.

Major points:

- Section 2.1, second paragraph: this is all good, but high dimensionality also limits mechanistic approaches (including data-driven ones). Models depicting different processes can generate the same patterns. The identifiability or equifinality mentioned in the last sentence of the previous paragraph relates to this. So, the shortcoming presented in this subsection is to both approaches, not only phenomenological. This must be better presented, because right now this is not the best selling point for the sections that follow, if the authors already stress in the end of the first paragraph of section 2.1 that mechanistic models may not be identifiable due to inherent mechanistic complexity of nature.

Author response: We have increased the emphasis on this point in Section 2.1: that problems of high dimensionality can plague mechanistic approaches as well. *However*, it is important to note that the approach we describe below also utilizes *data on underlying mechanisms* (Figures 2 and 3). Without these data, there are often multiple, very different, mechanistic models that can reproduce the same high-level patterns (e.g. abundance) but that are poor at predicting under novel conditions. We have reorganized this section and added a sentence emphasizing this point.

- It is somewhat unclear by the end of Section 2.1 how the authors came up with the four aspects they decided to devote the following sections (also presented in Figure 1).

Author response: We have added a sentence at the end of this section clarifying that these are four categories we created to organize the many types of complexity that could cause problems for regional forecasts.

- Section 3.1: the advocacy for simpler models is well careful here. Indeed, if one wants to forecast population dynamics of a given species, a simpler model may serve best. However, the authors arguments read as a critic to the movement towards more complex models (although they end up proposing a fairly complex model

anyway). Whereas complexification of models must be done careful as detailed in following sections, the motivation for more complex models normally includes going beyond population dynamics of a single species. Increases in model complexity should be accompanied by an increase in the number of emergent patterns. Say, adding interactions would mean we are now able to predict multiple species, entire communities and the interactions per se. Thus, the strive for more complex model goes together with the hope to increase the number of patterns to be explained \*simultaneously\*. This must be clearer in the section; right now it reads as researchers complexify models just to improve one particular pattern.

Author response: We agree with the referee's point. In our original submission we discussed an example of this in section 3.3, where ecophysiological/developmental models of crop performance have the benefit of predicting both fitness and ecosystem function (photosynthetic rate and primary productivity). We now have added an additional discussion of this issue in 3.1 with respect to model complexity.

- I just want to make a general comment here to express how well designed the framework is. I like it very much, including the potential to integrate loci linkage and spatially-structured gene flow. I also like the resulting inference section. Well done.

Author response: We thank the referee.

- Conclusion: I suggest toning down the text because 1) it is fair to advocate for careful complexification based on data-driven model comparisons, but for this you actually propose a fairly complex model. Hence, you start with a complex model and if necessary, you simplify it. 2) where exactly are the computational methods, in terms of codes? You just \*mention\* potential ways how to implement the model within a Bayesian framework. I would not see this as 'providing' the computational methods. In this regard, do the authors have some code with this framework that can be provided as supporting information? 3) I would rather argue that you pretty much answered the question posed by proposing an integration of i) allelic model with gene flow (i.e. with spatial structure); ii) phenotypic model; iii) demographic model. Whether one sticks to the full proposed complexity or to a simpler version only the data and model comparison will tell (in this I totally agree). So, you have a clear proposal of maximum complexity, so be honest about this in the conclusion! 4) The interspecific biotic interactions (section 2.4) constitute the least developed component of the proposed framework and this must be mentioned in the conclusion as well.

Author response: With respect to #1 and #3, the referee's critique seems to be that we are proposing a complex model but at the same time hedging as to whether it is a good approach. We must clarify that our complex integrated model is used to evaluate the multiple potentially important processes, and may not be the ideal model. We have added a sentence to the conclusion emphasizing this point.

With respect to #2, we have changed the language to "suggests general computational methods."

With respect to #4, we have included sentences in section 4.5 about limitations on our efforts to build mechanistic models.

Minor points: - Section 2.1, first paragraph: The authors stress the role of temporal transferability, but they could also stress spatial transferability problems, which are particularly problematic for invasive species.

Author response: We have added a sentence highlighting this idea.

- Section 2.2., last paragraph: mutations might be of lesser importance, but if regional models intend to model range dynamics under changing environments, mutations happening in the expanding front may be fixed. There are plenty of models showing how mutations 'surf' on the expanding front.

Author response: The referee is correct, we have included a mention of this and a citation to Peischl et al 2013.

- Section 2.3, second paragraph: True that machine learning algorithms must be taken with a grain of salt in terms of interpretability. This is true for ML-based optimization of either phenomenological or mechanistic models. Key here would be that ML-based techniques are only as good as their training, so to be able to predict a system in the future, the future conditions must be somehow integrated in the training data (e.g. by including particular geographic regions, or knowledge from past conditions). The usefulness of ML-based predictions (or any optimization method, including Bayesian frameworks) should be, in any case, tested via virtual ecologist frameworks.

Author response: Our proposed statistical framework provides a rigorous model-based approach to provide learning that can directly incorporate mechanistic process models (either agent-based or otherwise, including the virtual ecologist concept). We have clarified this in the manuscript and noted that it differs from ML approaches, most of which are not generative models, cannot simulate data based on mechanisms, and lack the ability to formally incorporate virtual ecologist frameworks. Our approach provides a much richer way to investigate processes.

- Section 2.4, second paragraph: note that abiotic change, variability or disturbances should work as coexistence mechanism (as an equalizing mechanism) by decreasing fitness advantages of stronger competitors or specialists. This is particularly true for NOVEL disturbances/changes, to which an assemblage has not yet had time to evolve a response. So, I liked to see something in these lines in the text, but I missed a bit an explicit mention of coexistence theories here.

Author response: We are not clear how explicit consideration of coexistence theory would add to our arguments or change our proposed approach. Our goal here is predicting population dynamics and not predicting community diversity. Our proposed approach already includes interspecific interactions, novel environments, and potential interactions among them. Furthermore, we do not have any strong intuition on how novel environments will impact coexistence, despite our familiarity with the theory.

- Section 2.4: this section seems to focus a lot on the role of how changing abiotic conditions affect biotic interactions. I would further suggest the authors to complement that with how changing biotic conditions affect biotic interactions as well. This is particularly important considering exploitation by humans or human-induced invasions. In this aspect, missing species (that go locally extinct) or novel species should pose enough change to influence local/regional biotic interactions. There should be plenty of interaction network studies on this topic. Also the complexity of high-order interactions and multiplexes are current expanding fields of research that deserve some attention, or at least, some mention here. The authors mention 'a problem or high dimensionality' in the context of multiple community members, but not explicitly about pairwise vs. high-order interactions, which is obviously a bigger challenge.

Author response: We have noted these points in this section, referring now to higher-order interactions as one of the challenges of high dimensionality in biotic interactions.

- Figure 3: please indicate in the figure caption where to find parameter explanations (e.g, refer to the text or tables). Consider making a figure version with the equations explicitly.

Author response: We have provided a detailed description of model parameters in the Supplemental tables, and noted this in the caption. A figure that jointly contains the directed model graph and the equations was so cluttered that it lacks clarity. However, with the descriptions in the main text, the symbols in the model graphic, and the supplemental tables with symbol descriptions, we have the components well covered.