THE ROYAL SOCIETY
PUBLISHING

PROCEEDINGS B

Increasing temporal variance leads to stable species range limits

John W. Benning, Ruth A. Hufbauer and Christopher Weiss-Lehman

Article citation details
Proc. R. Soc. B **289**: 20220202.
http://dx.doi.org/10.1098/rspb.2022.0202

Review timeline

Original submission:	9 August 2021
1st revised submission:	1 February 2022
2nd revised submission:	6 April 2022
3rd revised submission:	13 April 2022
Final acceptance:	14 April 2022

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

Review History

RSPB-2021-1783.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?
Good

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?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
Please see attached file (see Appendix A).

Review form: Reviewer 2

Recommendation
Reject – article is scientifically unsound

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

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

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

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?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
See attached file (see Appendix B).

Review form: Reviewer 3

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?
Good

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?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
In this work, the authors explore how environmental stochasticity affects the interplay of local adaptation and demography underlying species' range limits. They do so by including random temporal fluctuations in the intercept, slope, or both, of a spatial cline in optimum phenotype, representing spatio-temporal variation in the environment, in an individual based simulation model with explicit genetics (using the SLiM software). In doing so, they extend previous simulation and/or analytical work by Bridle, Polechova, Barton, and others, to also include temporal stochasticity in the environment.

Their main findings are that random fluctuations in the cline intercept (in other words: identical fluctuations through space) do not lead to a restricted range, and allow range expansion over steeper gradients when fluctuations are more autocorrelated. Fluctuations in gradient slope, leading to larger fluctuations away from the range core, do lead to stable (albeit fluctuating) range limits, with a smaller range under steeper (mean) gradient, and less autocorrelated fluctuations. Combining fluctuations in slope and intercept further reduces the range size and increases extinction risk.

These results make sense and seem overall well interpreted, I enjoyed how they are presented in the figures, and I found the paper well written and easy to follow. Still I think the paper could be improved by dealing with the points below.

First, I think one should always be careful with priority claims, such as "all evolutionary range limit models to date assume temporally constant environments." (89-90) Perhaps this is true if you think of models with multiple demes along a near-continuous gradient as here, but not for two-patch models that are often used to investigate what happens in marginal populations, which is critical to range limits. I'm thinking for instance of Holt et al (2004 Am Nat Temporal Variation Can Facilitate Niche Evolution in Harsh Sink Environments), but there may be more recent ones by these or other authors.

Second, I appreciate the value of simulations, but it would be better if you also used even very simple analytical predictions to guide their interpretation. More specifically, most of your results certainly arise from a load that random environmental fluctuations in the optimum impose on the population. These fluctuations causes an expected reduction in mean fitness, for which there are quite a few theoretical predictions in the literature (e.g. Burger & Lynch 1995 Evolution, Lande & Shannon 1996 Evolution, Chevin et al 2017 Am Nat). Some of these predictions are more complex than others, depending on how they account for evolution, for environmental autocorrelation, etc. But even the simplest one that neglects evolutionary responses to temporal fluctuations (something like (s/2) tau^2 in your model, derived by taking the expectation of squared deviations from the fluctuating optimum, assuming constant z) would probably do a decent job here. In the random intercept scenario, this expected load is constant over space, which is equivalent to reducing the growth rate by the same amount everywhere, regardless of local adaptation. In the fluctuating slope scenario, the stochastic load changes over space (with tau), such that even a perfectly adapted mean phenotype (on average) may have negative expected growth rate in extreme environments. I think that dealing with this load explicitly would make results much clearer.

Minor points
81-83: perhaps also cite here the broad analysis across birds & mammals by de Villemereuil et al (your ref 39), since it explicitly tested for a fluctuating optimum.
126-134: you could mention that both models imply a spatial autocorrelation = 1 (at least this is true of the variable intercept, the other may need checking)
182 (eq. 1): this fitness function allows negative values, how did you deal with that?
201 (Eq.2): this depends on the value of s, right?
225-228: It is not very clear where this argument on selection gradients is coming from. Perhaps it would help to write that selection gradients depend on s times the deviation from optimum in such a model.

270-272: “the rate of range expansion slowed as spatial gradients steepened and environments became less positively autocorrelated.” But I seem to see slower expansion under higher autocorrelation in Fig 2A
279: what do you mean by “stationary individuals”? I find the term misleading, as stationary has a meaning in a context of stochastic processes, which you also investigate here.
287-288: “Thus, temporal stochasticity introduced an extinction risk due to fluctuating phenotypic optima that had strong effects on mean fitness”. Here you would need to cite the relevant literature recommended above
366-368: “This is because steeper spatial gradients increased genetic variance across the landscape via gene flow, which better equipped populations to withstand temporal fluctuations in optima”. Interesting, but can you show evidence for that?
402: A reference seems required for this statement: “Increasing temporal variability is one predicted (and observed) consequence of contemporary climate change”
422-424: “there is a critical threshold of temporal environmental variance that can stop range expansion and enforce a stable range limit.” Yes, and you could use the argument outlined above to explain this: temporal fluctuations reduce the expected growth rate in a predictable manner, causing an additional load to that due to local maladaptation. For instance, I predict that the fluctuating slope model would produce stable limits even in the absence of any gene swamping (which you could check by setting dispersal to 0), simply because environmental fluctuations that are too large cause a negative expected long-term growth rate, and thus the populations decline on average in extreme environments.

Decision letter (RSPB-2021-1783.R0)

01-Nov-2021

Dear Dr Benning:

I am writing to inform you that your manuscript RSPB-2021-1783 entitled "Increasing temporal variance leads to stable species range limits" has, in its current form, been rejected for publication in Proceedings B.

This action has been taken on the advice of referees, who have recommended that substantial revisions are necessary. With this in mind we would be happy to consider a resubmission, provided the comments of the referees are fully addressed. However please note that this is not a provisional acceptance.

The resubmission will be treated as a new manuscript. However, we will approach the same reviewers if they are available and it is deemed appropriate to do so by the Editor. Please note that resubmissions must be submitted within six months of the date of this email. In exceptional circumstances, extensions may be possible if agreed with the Editorial Office. Manuscripts submitted after this date will be automatically rejected.

Please find below the comments made by the referees, not including confidential reports to the Editor, which I hope you will find useful. If you do choose to resubmit your manuscript, please upload the following:

1) A ‘response to referees’ document including details of how you have responded to the comments, and the adjustments you have made.
2) A clean copy of the manuscript and one with 'tracked changes' indicating your 'response to referees' comments document.
3) Line numbers in your main document.

4) Data - please see our policies on data sharing to ensure that you are complying (https://royalsociety.org/journals/authors/author-guidelines/#data).

To upload a resubmitted manuscript, 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 Resubmission." Please be sure to indicate in your cover letter that it is a resubmission, and supply the previous reference number.

Sincerely,
Professor Gary Carvalho
mailto: proceedingsb@royalsociety.org

Associate Editor
Board Member: 1
Comments to Author:
This is an MS that presents important and timely findings in terms of how temporal variance in various forms seems to limit adaptation along ecological gradients. If your findings can be shown to be more widely applicable (see below), I see it as highly suitable for publication in Proceedings.

However, as with all exclusively simulation studies, the question often remains how general its conclusions are and how much they depend on the structure of the model - i.e. how much of parameter space has been explored - remains. In particular, reviewer 2 raises some quite serious concerns around the ecological generality of your assumptions, while reviewer 3 requests that you incorporate some analytical modelling into your treatment, in order to justify that the design of your simulation, and parameter values chosen, are justified in terms of occupying sufficiently general regions of parameter space to be biologically relevant.

While I see your findings as important and (potentially) generalisable across a range of situations, I would like to see more justification for the biological set up of the model (esp bearing in mind Rev 2's comments) and the consequences for your findings of changing at least some of these assumptions, and especially your set up of the selective function, and calculations of effective population size.

I would also like you to include analytical predictions to guide your intepretation (particularly of the way that load operates) - and ideally include these analytical expectations in your figures, as suggested by reviewer 3.

Finally, I agree with rev 3 that it is not entirely true that all existing models have only considered variation in space or time...there have been various attempts to incorporate temporal change within a spatial context (even if not variance) - see also Pease et al 1989; and Polechova et al 2009 Am Nat.in addition to the Holt papers already mentioned.

Reviewer(s)' Comments to Author:
Referee: 1
Comments to the Author(s)
Please see attached file.

Referee: 2
Comments to the Author(s)
See attached file.

Referee: 3
Comments to the Author(s)
In this work, the authors explore how environmental stochasticity affects the interplay of local adaptation and demography underlying species’ range limits. They do so by including random

temporal fluctuations in the intercept, slope, or both, of a spatial cline in optimum phenotype, representing spatio-temporal variation in the environment, in an individual based simulation model with explicit genetics (using the SLiM software). In doing so, they extend previous simulation and/or analytical work by Bridle, Polechova, Barton, and others, to also include temporal stochasticity in the environment.

Their main findings are that random fluctuations in the cline intercept (in other words: identical fluctuations through space) do not lead to a restricted range, and allow range expansion over steeper gradients when fluctuations are more autocorrelated. Fluctuations in gradient slope, leading to larger fluctuations away from the range core, do lead to stable (albeit fluctuating) range limits, with a smaller range under steeper (mean) gradient, and less autocorrelated fluctuations. Combining fluctuations in slope and intercept further reduces the range size and increases extinction risk.

These results make sense and seem overall well interpreted, I enjoyed how they are presented in the figures, and I found the paper well written and easy to follow. Still I think the paper could be improved by dealing with the points below.

First, I think one should always be careful with priority claims, such as “all evolutionary range limit models to date assume temporally constant environments.” (89-90) Perhaps this is true if you think of models with multiple demes along a near-continuous gradient as here, but not for two-patch models that are often used to investigate what happens in marginal populations, which is critical to range limits. I’m thinking for instance of Holt et al (2004 Am Nat Temporal Variation Can Facilitate Niche Evolution in Harsh Sink Environments), but there may be more recent ones by these or other authors.

Second, I appreciate the value of simulations, but it would be better if you also used even very simple analytical predictions to guide their interpretation. More specifically, most of your results certainly arise from a load that random environmental fluctuations in the optimum impose on the population. These fluctuations causes an expected reduction in mean fitness, for which there are quite a few theoretical predictions in the literature (e.g. Burger & Lynch 1995 Evolution, Lande & Shannon 1996 Evolution, Chevin et al 2017 Am Nat). Some of these predictions are more complex than others, depending on how they account for evolution, for environmental autocorrelation, etc. But even the simplest one that neglects evolutionary responses to temporal fluctuations (something like (s/2) tau^2 in your model, derived by taking the expectation of squared deviations from the fluctuating optimum, assuming constant z) would probably do a decent job here. In the random intercept scenario, this expected load is constant over space, which is equivalent to reducing the growth rate by the same amount everywhere, regardless of local adaptation. In the fluctuating slope scenario, the stochastic load changes over space (with tau), such that even a perfectly adapted mean phenotype (on average) may have negative expected growth rate in extreme environments. I think that dealing with this load explicitly would make results much clearer.

Minor points
81-83: perhaps also cite here the broad analysis across birds & mammals by de Villemereuil et al (your ref 39), since it explicitly tested for a fluctuating optimum.
126-134: you could mention that both models imply a spatial autocorrelation = 1 (at least this is true of the variable intercept, the other may need checking)
182 (eq. 1): this fitness function allows negative values, how did you deal with that?
201 (Eq.2): this depends on the value of s, right?
225-228: It is not very clear where this argument on selection gradients is coming from. Perhaps it would help to write that selection gradients depend on s times the deviation from optimum in such a model.
270-272: “the rate of range expansion slowed as spatial gradients steepened and environments became less positively autocorrelated.” But I seem to see slower expansion under higher autocorrelation in Fig 2A

279: what do you mean by “stationary individuals”? I find the term misleading, as stationary has a meaning in a context of stochastic processes, which you also investigate here.
287-288: “Thus, temporal stochasticity introduced an extinction risk due to fluctuating phenotypic optima that had strong effects on mean fitness”. Here you would need to cite the relevant literature recommended above
366-368: “This is because steeper spatial gradients increased genetic variance across the landscape via gene flow, which better equipped populations to withstand temporal fluctuations in optima”. Interesting, but can you show evidence for that?
402: A reference seems required for this statement: “Increasing temporal variability is one predicted (and observed) consequence of contemporary climate change”
422-424: “there is a critical threshold of temporal environmental variance that can stop range expansion and enforce a stable range limit.” Yes, and you could use the argument outlined above to explain this: temporal fluctuations reduce the expected growth rate in a predictable manner, causing an additional load to that due to local maladaptation. For instance, I predict that the fluctuating slope model would produce stable limits even in the absence of any gene swamping (which you could check by setting dispersal to 0), simply because environmental fluctuations that are too large cause a negative expected long-term growth rate, and thus the populations decline on average in extreme environments.

Author's Response to Decision Letter for (RSPB-2021-1783.R0)

See Appendix C.

RSPB-2022-0202.R0

Review form: Reviewer 2

Recommendation
Reject – article is scientifically unsound

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

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

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

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
Dear Editor, February 14th 2022

Thank you for the opportunity to review "Increasing temporal variance leads to stable species range limits". In this article the authors use simulations in the population genomics simulator SLiM to explore the impact of temporal environmental fluctuations on the expansion of ranges into novel habitats. The authors examine expansion across a linear environmental gradient contrasting temporal fluctuations in the environmental intercept (impacting the spatial mean) and environmental slope (impacting the spatial variance) on outcomes such as range width.

Temporal variation is increasingly appreciated for its impact on eco-evolutionary dynamics as the authors highlight in their introduction. The basis of this paper is therefore well-founded and relevant to understanding adaptation in a changing world. However, there are several crucial errors in the development and design of the theoretical models that jeopardize the validity of their results and the conclusions drawn. Given these errors I will focus my review solely on the motivation and model formulation rather than the results.

Major concern 1: Simulations are not hypothesis-driven. The authors do a great job of drawing on the existing literature to emphasize the potential impacts of temporal fluctuations on demography and adaptation and hence the possibility of its impact on range expansion. These existing results are never synthesized however into specific hypotheses that clearly motivate the model. Rather, the model and the results are largely observational. While there is deep value to observational science there are distinct risks to employing an observational approach in the context of simulation-based science. In particular, as illustrated by this paper, taking an observational approach greatly reduces the ability to check the intuition of an outcome to identify flaws in model design. It also limits the ability of the reader to understand and justify the design of the simulations and results. For this work to achieve what it could, I highly recommend the authors explicitly formulate the hypothesis underlying the contrast between varying slope and varying intercept designs.

Major concern 2: Load.

There are a number of interconnected errors and omissions in the section on "expected genetic load due to temporal variation. First, load is defined relative to a specific model of fitness. This model of fitness should be clearly defined prior to equation 1. Also equation 1 is a function of both z_i and theta_{x,t}. Second, it should be clarified that this analytical model is not a dynamical model. It does not model phenotypic evolution. In lines 174-175 the focus on a population obscures this important distinction. More importantly equation 3 is wrong. Under the model of environmental variation as described, load is distributed like a chi-squared distribution with one degree of freedom (the distribution of a square of a normally distributed random variable) not normally distributed. This can be visualized intuitively. If the load were normally distributed it could technically be negative which wouldn't make sense. This same issue applies to equation 4. Finally, the title of the section (ln 169) as well as on line 201, the authors mention "expected genetic load". However, given the non-dynamical nature of the analytical model the only expectation presented is the expected load of an individual that is perfectly adapted to the average environment of its patch. It is unclear under what conditions such an expectation is meaningful. For example, it may have little meaning within a model of evolution where phenotypes vary both within a population and over time.

Major concern 3: Fitness definition. I didn't dive into Bridle et al.'s definition of fitness but equation 5 is inconsistent with the previously assumed model of Gaussian stabilizing selection that was used to derive the expression for load. The presence of negative values of absolute fitness (Ln 237) is indicative of additional issues with the definition of absolute fitness. In its current form demography can influence the strength of selection directly (the selection coefficient will be a function of r Nx and K) and indirectly through drift. One way to avoid this is to define fitness by multiplying the logistic growth model by the expression for relative fitness (Gaussian curve or 1-load if following equation 5) rather than by subtracting it.

Additional suggestions:

Ln 214: While modelling explicit sexes may seem like a standard and innocuous choice, there are important implications to modeling a gonochoric population. Namely, this introduces the potential for allee effects, where populations fail to establish due to lack of mates. This certainly is a potential realistic element limiting range expansion, however in a simulation alone it will be difficult to separate the impacts of allee effects from other demographic features. Therefore, it may be helpful to first consider a hermaphroditic population. Or contrast the existing results to the hermaphrodidic case to isolate the potential consequences of allee effects.

Decision letter (RSPB-2022-0202.R0)

07-Mar-2022

Dear Dr Benning:

Your manuscript has now been peer reviewed and the reviews have been assessed by an Associate Editor. The reviewers' comments (not including confidential comments to the Editor) and the comments from the Associate Editor are included at the end of this email for your reference. As you will see, the reviewers and the Editors have raised some concerns with your manuscript and we would like to invite you to revise your manuscript to address them.

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 by the Associate Editor, 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), 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,
Professor Gary Carvalho
mailto: proceedingsb@royalsociety.org

Associate Editor
Comments to Author:
Thank you for sending a resubmitted version of your MS, as well as a detailed response to referees' comments, including your additional efforts to establish the generality of your model.

In particular, I appreciate the addition of some analytical approximations for load to illustrate a connection to existing theory (rather than as a full treatment of the load generated in the main simulations). This I think is appropriate given Rev. 3's comments made it clear that s/he thought a simple model would be a good enough approximation to make the key point about the impact of fluctuations in intercept or slope on the load and its spatial pattern.

Your resubmission was sent to one of the original reviewers to consider, who continues to highlight several concerns with the MS - these are detailed in the attached document. Although I do not share all of these concerns, I would like to give you the opportunity to respond to them, and to modify your MS accordingly. In particular, - you might consider changes to ensure that the function of the load model, and its limitations are made more clear.

Many thanks again for submitting your work to Proc Roy Soc, and for your willingness to make extensive revisions on resubmission.

Reviewer(s)' Comments to Author:
Referee: 2
Comments to the Author(s).
Dear Editor, February 14th 2022

Thank you for the opportunity to review “Increasing temporal variance leads to stable species range limits”. In this article the authors use simulations in the population genomics simulator SLiM to explore the impact of temporal environmental fluctuations on the expansion of ranges into novel habitats. The authors examine expansion across a linear environmental gradient contrasting temporal fluctuations in the environmental intercept (impacting the spatial mean) and environmental slope (impacting the spatial variance) on outcomes such as range width.

Temporal variation is increasingly appreciated for its impact on eco-evolutionary dynamics as the authors highlight in their introduction. The basis of this paper is therefore well-founded and relevant to understanding adaptation in a changing world. However, there are several crucial errors in the development and design of the theoretical models that jeopardize the validity of their results and the conclusions drawn. Given these errors I will focus my review solely on the motivation and model formulation rather than the results.

Major concern 1: Simulations are not hypothesis-driven. The authors do a great job of drawing on the existing literature to emphasize the potential impacts of temporal fluctuations on demography and adaptation and hence the possibility of its impact on range expansion. These existing results are never synthesized however into specific hypotheses that clearly motivate the model. Rather, the model and the results are largely observational. While there is deep value to observational science there are distinct risks to employing an observational approach in the context of simulation-based science. In particular, as illustrated by this paper, taking an observational approach greatly reduces the ability to check the intuition of an outcome to identify flaws in model design. It also limits the ability of the reader to understand and justify the design of the simulations and results. For this work to achieve what it could, I highly recommend the authors explicitly formulate the hypothesis underlying the contrast between varying slope and varying intercept designs.

Major concern 2: Load.

There are a number of interconnected errors and omissions in the section on "expected genetic load due to temporal variation. First, load is defined relative to a specific model of fitness. This model of fitness should be clearly defined prior to equation 1. Also equation 1 is a function of both z_i and theta_{x,t}. Second, it should be clarified that this analytical model is not a dynamical model. It does not model phenotypic evolution. In lines 174-175 the focus on a population obscures this important distinction. More importantly equation 3 is wrong. Under the model of environmental variation as described, load is distributed like a chi-squared distribution with one degree of freedom (the distribution of a square of a normally distributed random variable) not normally distributed. This can be visualized intuitively. If the load were normally distributed it could technically be negative which wouldn't make sense. This same issue applies to equation 4. Finally, the title of the section (ln 169) as well as on line 201, the authors mention "expected genetic load". However, given the non-dynamical nature of the analytical model the only expectation presented is the expected load of an individual that is perfectly adapted to the average environment of its patch. It is unclear under what conditions such an expectation is meaningful. For example, it may have little meaning within a model of evolution where phenotypes vary both within a population and over time.

Major concern 3: Fitness definition. I didn't dive into Bridle et al.'s definition of fitness but equation 5 is inconsistent with the previously assumed model of Gaussian stabilizing selection that was used to derive the expression for load. The presence of negative values of absolute fitness (Ln 237) is indicative of additional issues with the definition of absolute fitness. In its current form demography can influence the strength of selection directly (the selection coefficient will be a function of r Nx and K) and indirectly through drift. One way to avoid this is to define fitness by multiplying the logistic growth model by the expression for relative fitness (Gaussian curve or 1-load if following equation 5) rather than by subtracting it.

Additional suggestions:

Ln 214: While modelling explicit sexes may seem like a standard and innocuous choice, there are important implications to modeling a gonochoric population. Namely, this introduces the potential for allee effects, where populations fail to establish due to lack of mates. This certainly is a potential realistic element limiting range expansion, however in a simulation alone it will be difficult to separate the impacts of allee effects from other demographic features. Therefore, it may be helpful to first consider a hermaphroditic population. Or contrast the existing results to the hermaphrodidic case to isolate the potential consequences of allee effects.

Author's Response to Decision Letter for (RSPB-2022-0202.R0)

See Appendix D.

Decision letter (RSPB-2022-0202.R1)

11-Apr-2022

Dear Dr Benning

I am pleased to inform you that your Review manuscript RSPB-2022-0202.R1 entitled "Increasing temporal variance leads to stable species range limits" has been accepted for publication in Proceedings B.

The referee(s) do not recommend any further changes. Therefore, please proof-read your manuscript carefully and upload your final files for publication. 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 me know immediately.

To upload your manuscript, 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.

You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, upload a new version through your Author Centre.

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. Please note that PowerPoint files are not accepted.

3) Electronic supplementary material: this should be contained in a separate file from the main text and the file name should contain the author’s name and journal name, e.g authorname_procb_ESM_figures.pdf
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 see: https://royalsociety.org/journals/authors/author-guidelines/

4) Data-Sharing and data citation
It is a condition of publication that data supporting your paper are made available. Data should be made available either in the electronic supplementary material or through an appropriate repository. Details of how to access data should be included in your paper. Please see https://royalsociety.org/journals/ethics-policies/data-sharing-mining/ for more details.

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=RSPB-2022-0202.R1 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.

5) 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 final version. If you have any questions at all, please do not hesitate to get in touch.

Sincerely,
Professor Gary Carvalho
Editor, Proceedings B
mailto:proceedingsb@royalsociety.org

Associate Editor Board Member
Comments to Author:
Many thanks for these corrections and clarifications in response to the outstanding concerns of the reviewer. I feel the MS has improved as a result.

In terms of your overall conclusions, the recent review of Drosophila field experiments by O'Brien et al (2022) Phil Trans Roy Soc B 377 may be a useful empirical example of how temporal environmental variation limits local adaptation - could be useful to cite in your discussion?

I think it's a shame that you have had to move your paragraph on future work to the SI - I would prefer to see it reinstated in the main discussion, if at all possible.

Decision letter (RSPB-2022-0202.R2)

14-Apr-2022

Dear Dr Benning

I am pleased to inform you that your manuscript entitled "Increasing temporal variance leads to stable species range limits" 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 please 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

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

Data Accessibility section
Please remember to make any data sets live prior to publication, and update any links as needed when you receive a proof to check. It is good practice to also add data sets to your reference list.

Open Access
You are invited to opt for Open Access, making your freely available to all as soon as it is ready for publication under a CCBY licence. Our article processing charge for Open Access is £1700. Corresponding authors from member institutions (http://royalsocietypublishing.org/site/librarians/allmembers.xhtml) receive a 25% discount to these charges. For more information please visit http://royalsocietypublishing.org/open-access.

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,
Editor, Proceedings B
mailto: proceedingsb@royalsociety.org

Appendix A

This interesting MS advances modelling of the evolution of range margins by adding temporal variation in environmental optima. Since temporal variation is ubiquitous in nature and there are good reasons to expect it to influence margins, based on its known effects on persistence of individual populations, this is an important step to take. There is much current interest in range margins because of the impact of anthropogenic environmental change, which makes the paper topical. Since anthropogenic changes in mean climatic variables are likely to be accompanied by changes in temporal fluctuations, the contribution is all the more timely.

The paper is very clearly written and illustrated so I have very few comments related to presentation. I give some thoughts and suggestions below but they are relatively minor in nature. Perhaps the most important point is that the MS considers only a small part of the large range of possible patterns of temporal variation, and in the context of a specific set of other modelling choices. Therefore, inferences about the general effects of temporal fluctuation must always be accompanied by caveats.

Specific comments by line number:

72-3 – biotic interactions of this type are missing from many of the theoretical treatments described up to this point. I would suggest noting this fact here and citing one or two studies that do include biotic interactions in range-margin models (starting with the classic work of Case & Taper).

98-100 – this is an important point. There are many different ways in which temporal fluctuations might be manifest and their consequences might be quite different. The simulations presented here use only two forms of temporal fluctuations – this is fine as a starting point but it does mean that extrapolations to the impact of temporal fluctuations in general need to be made with great care. One issue not mentioned here is the extent of spatial correlation of temporal fluctuations (not the same as variance changing across space). The way the simulations are set up here results in very strong spatial correlation, which might be a rather special case. For example, the impact of gene flow will be different if optima for adjacent demes change always in the same direction rather than sometimes changing in opposite directions.

170-5 – There is a significant departure here from previous models that assumed unlinked loci. Here, linkage is rather unrealistically strong because the single chromosome has length roughly 10^{-3}cM (10^5bp x 10^{-8} recombination events per bp per gamete per generation).

192-4 – how was migration handled for edge populations?

201 – this is for s=0.125, which has not been specified is but is presumably used throughout

234 – was this increase linear, or more than linear?

262 – tau was held constant – so the exploration is across varying slope and autocorrelation, but not across varying extent of temporal fluctuations?

264 – before looking at temporal variation, if would be good to see that the simulation recovered the effects of changing slope predicted from previous work (with a = 1, is the critical slope at the predicted value?). This comparison is currently given at the end of this section. It would be helpful to explain here how the Polechova and Barton prediction was obtained for the conditions of this simulation. A precise match is not expected.

328-33 – can this be understood as an effective steepening of the selective gradient created by adding fluctuations to the linear gradient? If so, it would be compatible with previous work. This is the reason for the question on line 234 about the way the selection gradient changes.

376-393 – these conclusions are specific to the simulation conditions – they might be more generally true but that is not yet demonstrated. I think an explicit recognition of this is important.

422-4 – here it might be appropriate to discuss, as mentioned above, whether the variable slope model generates something equivalent to the stable steepening slope models, or whether the margin is created in a fundamentally different way.

Appendix B

Summary:

By adding environmental noise to the slope and intercept of an environmental gradient, the authors examine the role of environmental stochasticity in determining population extinction (varying intercept) and range limits (varying slope). They find that environmental stochasticity can have important implications for range dynamics and the formation of stable range limits which are not present in the absence of environmental noise. This work has important implications for our understanding of the eco-evolutionary range dynamics and the formation of range limits.

While I think the results of this work have potentially important implications, I have several major technical concerns about the model used that limit my mechanistic understanding of the results and my confidence in their generality to natural ranges. I elaborate on these major concerns in detail below, followed by some relatively minor comments. Overall, the paper is well written and was exciting to read.

Major Concerns:

First is the functional form of fitness in equation 1. Defined as such fitness is density-dependent, the relative fitness, and hence the selection coefficient, is a function of the local population size. This can have important and unintended consequences on range dynamics as range edges and range cores can differ markedly in their population sizes. I believe this may be the case here. Written as in equation 1, the strength of selection will be greatest when population size is large (the second term drops out and hence the third term dominates) and is relatively weak when population sizes are small (the second term is large and may overwhelm selection). The nature of this density-dependence has some worrisome implications for the results. Are stable range limits formed because range edge populations are too small to experience significant selection? I understand the authors are adapting equation 1 from Bridel et al., but the biological motivation for the equation remains unclear. Unless the (-) in equation 1 has a very clear biological basis (as opposed to using a (*) which would be density-independent selection) I would not choose to use this functional form of fitness. Lastly, there form of equation 1 has the unsettling property that it can become negative when populations deviate significantly from the optima. It was unclear to me what measures were taken to prevent this from happening or to even assess if it was happening.

Second, I have some concerns about the effective population size in the model and the role this has in preventing adaptation. While the exact computation of the effective population size is difficult in ranges there are several reasons to believe the effective population size may be quite small in this model. First, K=50 is quite small, especially given two sexes and hence a female census population size of at most 25. Second, while I am not sure I understand the mating system (see below), my understanding of it seems to imply that there will be a lot of variation in reproductive success. This will further reduce the effective population size. In Figure 2D there is a sudden crash in the genetic variance but relatively little correlation between range-wide population size and genetic variance which may be an indication of some issues with effective population size leading to a sudden threshold value. To truly understand the role of Ne, additional simulations are needed that vary K and assesses the sensitivity of the model to this parameter.

Third, maybe I am understanding this wrong but the migration rate seems very high. For Poisson distribution with mean of 0.8 more than half of the individuals (~55%) should move at least one deme. A general rule of thumb (which comes from Fst) is that populations will not be able to spatially differentiate if there is more than 1 migrant per generation. This high of a level of migration may mean that there is a lot of genetic swamping and no ability to adapt to local environmental conditions. To address this one would need to greatly reduce m and measure Fst across the range to the measure whether failure to adapt to local environmental conditions is a result of swamping versus lack of local genetic variation. Note that this level of migration may be bolstering Ne despite the small population size, leading to high levels of neutral genetic variation but very little potential to adapt.

Finally, the envisioned life-cycle of the model is unclear. Why are there separate sexes? Are offspring created by random sampling of gametes or by random sampling of parents? If parents and not gametes are sampled at random then this could lead to large variation in reproductive success and dramatic decreases in effective population size. What is the motivation for using relative fitness in males?

Minor Comments:

Ln 72: One part of the introduction that really confused me was what was meant by “fluctuations” vs “stochasticity” and if these two terms were being used interchangeably. The mention of predators and pathogens in this sentence made me think “fluctuations” was referring to an sinusoidal phenotypic optimum dynamics which are approximately coevolution (e.g., Takahata and Ni 1990), however, as far as I can tell this is not what is meant by “fluctuations”.

Ln 75: I was confused by what references are included here. Why are 25,26,31, and 32 not included? I understand these are discussed later but it made me think that I was misinterpreting this sentence.

Ln 78: Once again I am a confused by the connection between stochasticity and fluctuating selection. The two words are used interchangeably in the introduction but one refers to the (a) in equation 3 and one refers to the (d) in equation 3, right? But then d is related to a so I am having a hard time understanding what is meant by “fluctuation” and “stochasticity” in the context of this model.

Ln 104: "experience more temporal variation" or "exhibit higher temporal variance". Variance is a numerical value whereas variation is a property.

Ln 108: forward-in-time

Ln 109: I would really appreciate more information on the connection to Polechova and Barton and Bridele et al. When you say `building on’, it what way?

Ln 119-125: I had to read this paragraph many times to processes it, its hard to figure out what each "two" refers to and what the second is to the first in line 120.

Ln 123-124: I think one thing that confused me is you never mention the leg of the autocorrelation. It is implied in line 123-124 that the lag is also kept at 1 generation.

Figure 1: Add panel labels.

Figure 1: Rather than or in addition to the different colours could you show example dynamics over time for positive, negative and no autocorrolation?

Ln 158: SLiM, a forward-in-time population genetic simulator,

Ln 170: Why model separate sexes? This will only decrease the effective population size. The role of separate sexes never seems to come up again.

Ln 173: Units on the mutation rate, is this per locus or per allele? On that note why model diploids? What is dominance?

Ln 175: I am guessing this is between any two consecutive loci right?

Ln 180: This isn't actually a "carrying capacity". The carrying capacity of the population is the equilibrium population size which is not what will happen in this model.

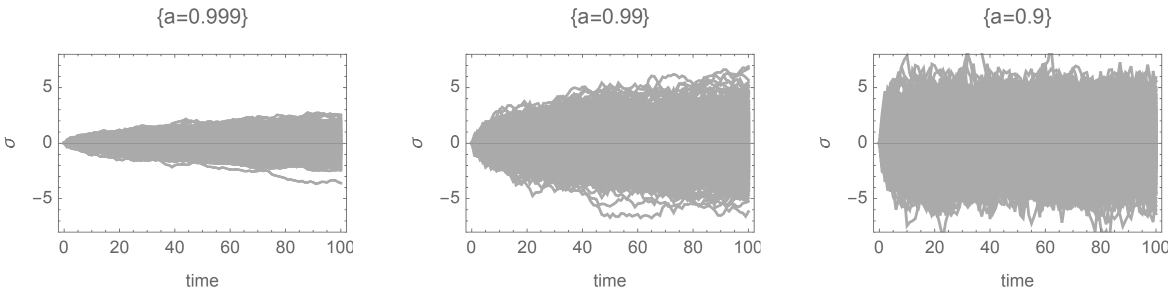
Ln 180: Once again K=50 is pretty small. Motivation for this?

Equation 1: Can you include the intercept of the phenotypic optima explicitly in this equation? E.g., b0+b1 x. Given your genome structure what is the range of z relative to the quadratic? You try to capture this via the linear selection gradient later on but a figure of this would be good? How do you know that W won’t go negative?

Ln 188: I'm not sure I 100% understand. Is each female guaranteed to mate? If so, are all of the offspring of a given female full-sibs or are gametes drawn at random? If male-female mated pairs are drawn at random and not gametes being drawn at random this could result in a significant random variation in reproductive success which could decrease population size even further.

Ln 199-200: This illustrates some of the challenges of density-dependent selection in modelling range dynamics. If you used a density-independent definition of fitness then the selective disadvantage of a particular phenotypic mis-match would depend only on the slope of the gradient and not on the variation in population size among patches.

Ln 210-211: This does not make sense to me. Consider for example \sigma_1=0+(1-a^2)^{0.5}N(0,\tau). The variance in \sigma then is: var(\sigma)=(1-a^2)\tau^2. Is this some equilibrium property? How long does it take to reach equilibrium? Should sigma_0 not be set to 0 then but rather a random number? See example figures below for different values of a given \tau=2.



Ln 214 and 218: Capitalize B and C to match figure caption

Ln 216: I think equation 1 should explicitly include the intercept.

Ln 228: Just to make sure I understand, for every point in space you include two points in the dataset, one which is the current optima and one which is the long-term average optima. Then you run a linear regression to calculate the slope. If my understanding above is correct, I don't understand why the long-term average optima is included. This makes the temporal variation in the gradient hard to interpret. Note: relative fitness is a function of population size, so there could be a range edge effect.

Ln 229: Not that it will matter too much with a \sigma_P of 1.7 but why are you multiplying? I would think you would want to divide by the phenotypic variance (not the SD) to get the the variance in fitness as a function of the phenotypic optimum in units of phenotypic standard deviations.

Ln 237: If you are only measuring 40 patches away from the centre then is there 4 points (2 from the left and 2 from the right) of your linear regression? Or do you run multiple replicate simulations etc.

Ln 246-249: What do you mean by "minimizing the impact of initial conditions". To me this means that I want the simulations to reach an eco-evolutionary equilibrium such that the initial patch in the range remains near this equilibrium throughout the main simulations. Is this true? I would think not as the reduction in population size would result in a loss of genetic variation. Does the genetic variation change significantly at the central patch during the simualtions?

What about the population size? Do you have graphs that show that you do indeed reach an equilibrium during the burnin? Why use a different carrying capacity and stochastic distribution? Why these values?

Figure 3D: How is his measured? Is this for selected, neutral, or genome wide? From line 251 I am guessing this is the neutral pi. Can you plot the selected pi as well? What does the genetic variance in each deme look like? What does Fst look like? I am asking some of these questions because I'm a little surprised by the dynamics of genetic variance especially to how C and D relate. Why do the earlier bottlenecks at say gen 3770 not have similar effects on genetic variance as the one at 3790?

Figure 3D: Why does the orange line end before extinction?

Appendix C

Associate Editor

This is an MS that presents important and timely findings in terms of how temporal variance in various forms seems to limit adaptation along ecological gradients. If your findings can be shown to be more widely applicable (see below), I see it as highly suitable for publication in Proceedings.

Thank you for these encouraging words. We have addressed all AE and reviewer comments, which greatly improved the manuscript.

1. However, as with all exclusively simulation studies, the question often remains how general its conclusions are and how much they depend on the structure of the model - i.e. how much of parameter space has been explored - remains. In particular, reviewer 2 raises some quite serious concerns around the ecological generality of your assumptions, while reviewer 3 requests that you incorporate some analytical modelling into your treatment, in order to justify that the design of your simulation, and parameter values chosen, are justified in terms of occupying sufficiently general regions of parameter space to be biologically relevant.

   Indeed; we feel that incorporating the suggestions of all three reviewers (especially regarding sensitivity analyses and the analytical model) has greatly increased the generality of our findings. Relevant comment responses include comments 16-18 and 49.

2. While I see your findings as important and (potentially) generalisable across a range of situations, I would like to see more justification for the biological set up of the model (esp bearing in mind Rev 2's comments) and the consequences for your findings of changing at least some of these assumptions, and especially your set up of the selective function, and calculations of effective population size.

   We have now addressed the reviewers' comments about justifying model parameters and setup (comments 16-19), and include sensitivity analyses across varying dispersal and carrying capacity values in Supplementary Material. Our main results are largely robust to changes in these parameters.

3. I would also like you to include analytical predictions to guide your interpretation (particularly of the way that load operates) - and ideally include these analytical expectations in your figures, as suggested by reviewer 3.

   Yes, this suggestion by Reviewer 3 was quite insightful and we have now incorporated it into the paper (comment 49; Material and methods: *Expected genetic load due to temporal variation*).

4. Finally, I agree with rev 3 that it is not entirely true that all existing models have only considered variation in space or time...there have been various attempts to incorporate temporal change within a spatial context (even if not variance) - see also Pease et al 1989; and Polechova et al 2009 Am Nat.in addition to the Holt papers already mentioned.

   There were indeed important references and history missing. They have been added where appropriate throughout the manuscript (e.g. lines 81 and 95 in the Introduction).

Referee 1

This interesting MS advances modelling of the evolution of range margins by adding temporal variation in environmental optima. Since temporal variation is ubiquitous in nature and there are good reasons to expect it to influence margins, based on its known effects on persistence of individual populations, this is an important step to take. There is much current interest in range margins because of the impact of anthropogenic environmental change, which makes the paper topical. Since anthropogenic changes in mean climatic variables are likely to be accompanied by changes in temporal fluctuations, the contribution is all the more timely.

The paper is very clearly written and illustrated so I have very few comments related to presentation. I give some thoughts and suggestions below but they are relatively minor in nature. Perhaps the most important point is that the MS considers only a small part of the large range of possible patterns of temporal variation, and in the context of a specific set of other modelling choices. Therefore, inferences about the general effects of temporal fluctuation must always be accompanied by caveats.

Specific comments by line number:

5. **72-3** – biotic interactions of this type are missing from many of the theoretical treatments described up to this point. I would suggest noting this fact here and citing one or two studies that do include biotic interactions in range-margin models (starting with the classic work of Case & Taper).

   Indeed, models examining biotic interactions are important to include. We now include this text at line 57: *Metapopulation dynamics (Holt and Keitt 2000) and biotic interactions (Case and Taper 2000; Case et al. 2005) also can enforce range limits.*

6. **98-100** – this is an important point. There are many different ways in which temporal fluctuations might be manifest and their consequences might be quite different. The simulations presented here use only two forms of temporal fluctuations – this is fine as a starting point but it does mean that extrapolations to the impact of temporal fluctuations in general need to be made with great care. One issue not mentioned here is the extent of spatial correlation of temporal fluctuations (not the same as variance changing across space). The way the simulations are set up here results in very strong spatial correlation, which might be a rather special case. For example, the impact of gene flow will be different if optima for adjacent demes change always in the same direction rather than sometimes changing in opposite directions.

   This is a very good point, which we now include in the Discussion at line 521: *Lastly, temporal stochasticity as modeled here is strongly spatially autocorrelated — i.e., adjacent patches always experience temporal variation in optima in the same direction, positive or negative. It would be fruitful to relax this assumption to allow less spatial autocorrelation in temporal variation, as the effects of gene flow on adaptation would likely differ.*

7. **170-5** – There is a significant departure here from previous models that assumed unlinked loci. Here, linkage is rather unrealistically strong because the single chromosome has length roughly 10^{-3}cM (10^{5}bp x 10^{-8} recombination events per bp per gamete per generation).

   This is a very good point. We have kept the chromosome the same length (10^{5} bp) but have now increased the recombination rate to 10^{-5}, thus making the chromosome one Morgan long. Our results did not show qualitative differences with this change.

8. **192-4** – how was migration handled for edge populations?

   We have now clarified at line 239: *Individuals set to disperse beyond the edges of the landscape were instead routed to the most distal patch in that direction. To avoid any potential edge effects, one end condition for our simulations (see below) was if the population of either of the most distal patches reached half the carrying capacity or more.*

   Also noted at line 299: *Simulations ended after 20,000 generations, or if all populations went extinct, or if at least one of the most peripheral landscape patches reached a population size at least half the carrying capacity (i.e., the species had filled the entire landscape).*

9. **201** – this is for s=0.125, which has not been specified is but is presumably used throughout

   Indeed, which we have now clarified at line 252.

10. **234** – was this increase linear, or more than linear?

    The increase in $|\beta|$ and σ_β with distance from the landscape center in the varying slope scenario is linear, but $Var(\beta)$ increases non-linearly. This is a very salient point that is now discussed at lines 313 and 478 (see also comment 49).

11. **262** – tau was held constant – so the exploration is across varying slope and autocorrelation, but not across varying extent of temporal fluctuations?

    This is correct for $\tau_{intercept}$ and τ_{slope}, but in the varying slope scenario, τ_{patch} (the temporal variance in optima for a given patch) increases with distance from the landscape center (Fig. 2).

12. **264** – before looking at temporal variation, if would be good to see that the simulation recovered the effects of changing slope predicted from previous work (with a = 1, is the critical slope at the predicted value?). This comparison is currently given at the end of this section. It would be helpful to explain here how the Polechova and Barton prediction was obtained for the conditions of this simulation. A precise match is not expected.

    The details of how this comparison was done are now in supplementary material S3, and the Results text moved to the beginning of the section as recommended (line 324).

13. **328-33** – can this be understood as an effective steepening of the selective gradient created by adding fluctuations to the linear gradient? If so, it would be compatible with previous work. This is the reason for the question on line 234 about the way the selection gradient changes.

    In essence, yes — the expected genetic load increases non-linearly with distance from the landscape center (please see Fig. 2 and line 478).

14. **376-393** – these conclusions are specific to the simulation conditions – they might be more generally true but that is not yet demonstrated. I think an explicit recognition of this is important.

We have added this qualification to the first sentence of the Discussion.

15. **422-4** – here it might be appropriate to discuss, as mentioned above, whether the variable slope model generates something equivalent to the stable steepening slope models, or whether the margin is created in a fundamentally different way.

    Yes, this insight in regards to steepening gradients is now a central part of the paper (e.g., Fig. 2), thanks to you and Reviewer 3!

Referee 2

By adding environmental noise to the slope and intercept of an environmental gradient, the authors examine the role of environmental stochasticity in determining population extinction (varying intercept) and range limits (varying slope). They find that environmental stochasticity can have important implications for range dynamics and the formation of stable range limits which are not present in the absence of environmental noise. This work has important implications for our understanding of the eco-evolutionary range dynamics and the formation of range limits.

While I think the results of this work have potentially important implications, I have several major technical concerns about the model used that limit my mechanistic understanding of the results and my confidence in their generality to natural ranges. I elaborate on these major concerns in detail below, followed by some relatively minor comments. Overall, the paper is well written and was exciting to read.

Major Concerns:

16. First is the functional form of fitness in equation 1. Defined as such fitness is density dependent, the relative fitness, and hence the selection coefficient, is a function of the local population size. This can have important and unintended consequences on range dynamics as range edges and range cores can differ markedly in their population sizes. I believe this may be the case here. Written as in equation 1, the strength of selection will be greatest when population size is large (the second term drops out and hence the third term dominates) and is relatively weak when population sizes are small (the second term is large and may overwhelm selection). The nature of this density-dependence has some worrisome implications for the results. Are stable range limits formed because range edge populations are too small to experience significant selection? I understand the authors are adapting equation 1 from Bridel et al., but the biological motivation for the equation remains unclear. Unless the (-) in equation 1 has a very clear biological basis (as opposed

to using a (*) which would be density independent selection) I would not choose to use this functional form of fitness. Lastly, there form of equation 1 has the unsettling property that it can become negative when populations deviate significantly from the optima. It was unclear to me what measures were taken to prevent this from happening or to even assess if it was happening.

We understand this concern, but believe we can allay it. One impetus for using this particular fitness function is precedent — this is the (general) fitness function used in Polechová and Barton 2015, Polechová 2018, and Bridle et al. 2019, the models most closely aligned with ours, and the modeling frameworks that our manuscript seeks to build on. Comparisons to prior models, as we do at line 324 for Polechová and Barton 2015, would be difficult if our model used a different fitness function that did not include density dependent selection. Furthermore, for fecundity selection as modeled here, there is good biological reason to think that when populations are at carrying capacity, deviation from optima is more important than at low population sizes when resources are abundant and intraspecific competition is low. Also, one relevant point in regard to small population sizes at the range edge: the small size of range edge populations, which decreases the efficacy of selection relative to drift, is one of the key mechanisms and subsequent insights from Polechová and Barton 2015.

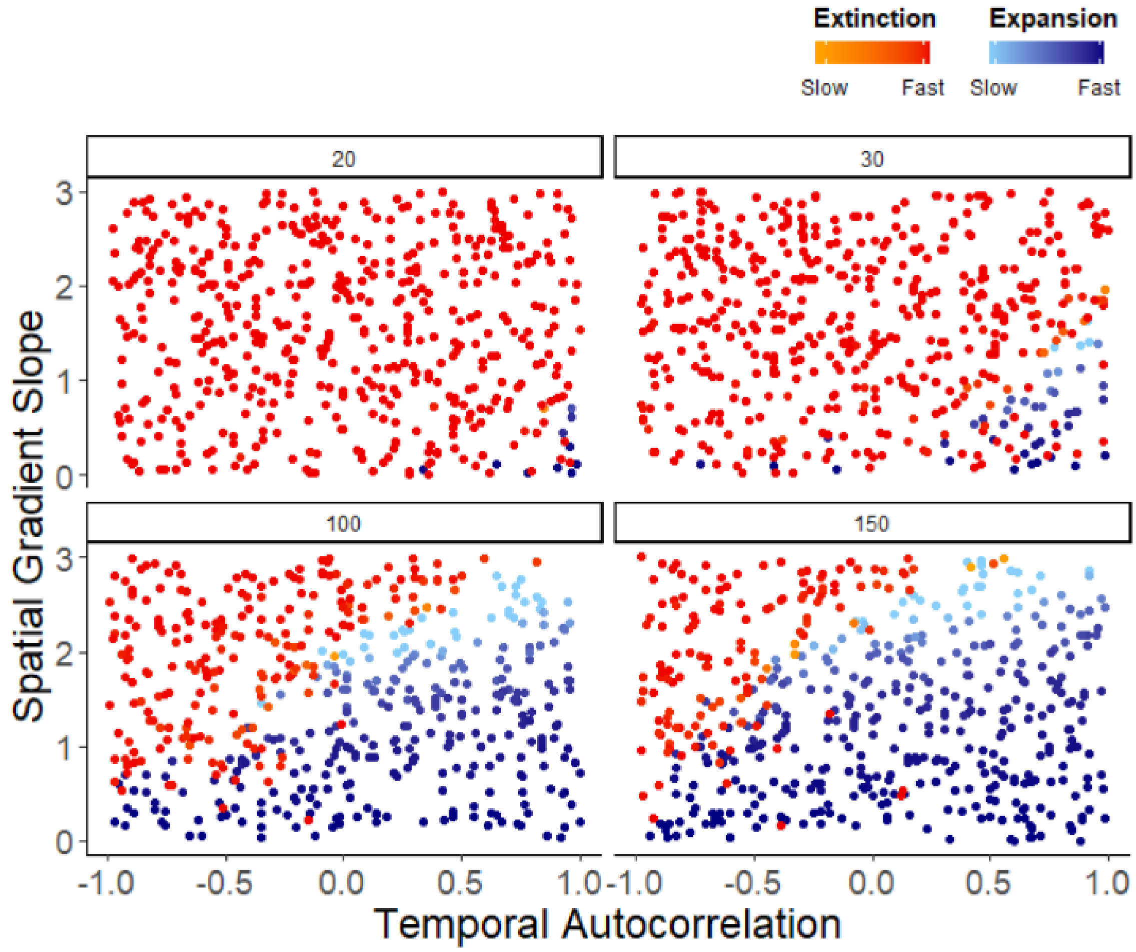
Also, it is true that the function can become negative. But in the simulation model, any $W_i < 0$ was set to 0.00001 (as the mean for a Poisson distribution in SLiM must be > 0), which we now clarify at line 237.

17. Second, I have some concerns about the effective population size in the model and the role this has in preventing adaptation. While the exact computation of the effective population size is difficult in ranges there are several reasons to believe the effective population size may be quite small in this model. First, K=50 is quite small, especially given two sexes and hence a female census population size of at most 25. Second, while I am not sure I understand the mating system (see below), my understanding of it seems to imply that there will be a lot of variation in reproductive success. This will further reduce the effective population size. In Figure 2D there is a sudden crash in the genetic variance but relatively little correlation between range-wide population size and genetic variance which may be an indication of some issues with effective population size leading to a sudden threshold value. To truly understand the role of Ne, additional simulations are needed that vary K and assesses the sensitivity of the model to this parameter.

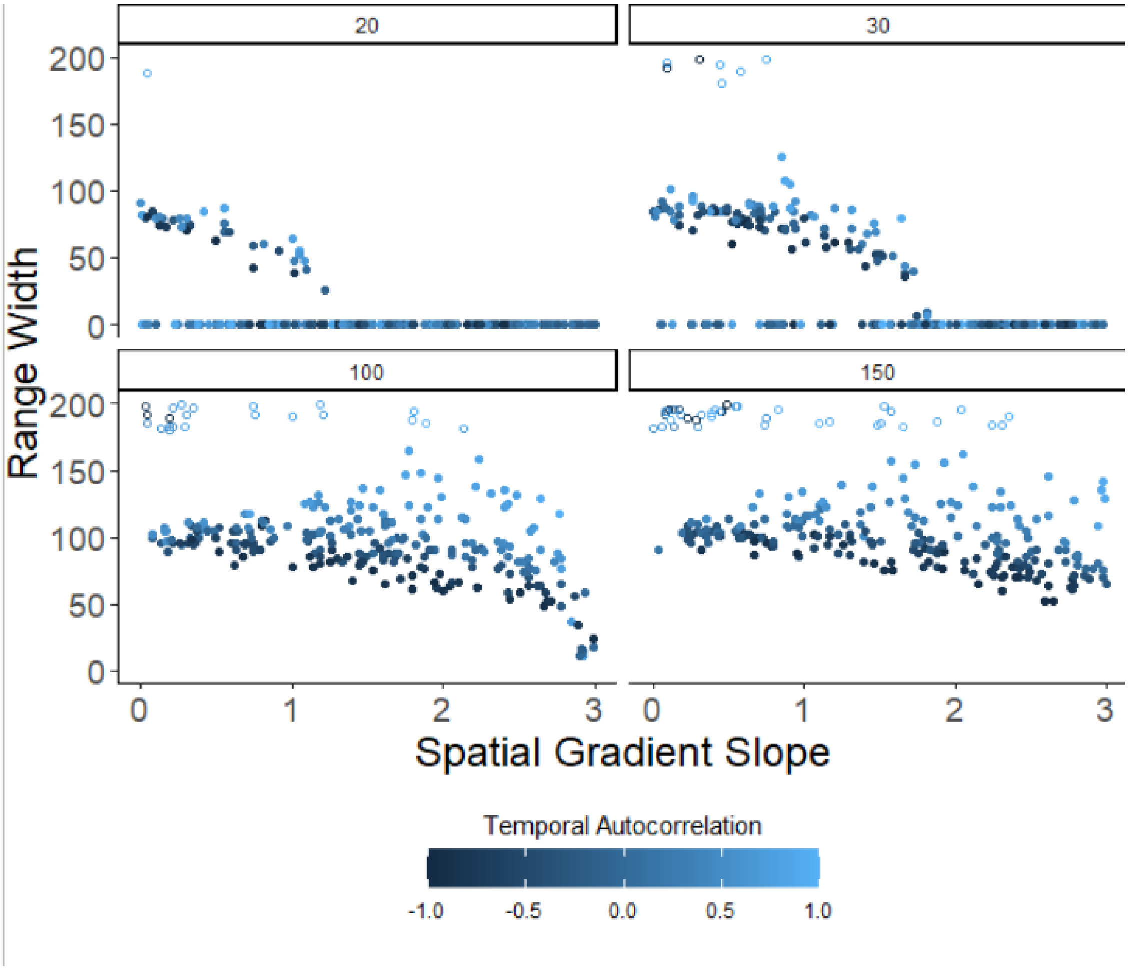
    This is a good idea, and we now have added sensitivity analyses varying K to the Supplementary Material (Figs. S3, S4, and below). In short, our results are largely robust to changes in K (and dispersal; see next comment), with little change in the qualitative relationship between gradient slope and temporal stochasticity for either the varying intercept or varying slope scenarios. The changes we did observe are in line with what we would expect from earlier work on the relationship between population size, dispersal, and adaptation (e.g., Lande 1993; Samani and Bell 2010; Polechová and Barton 2015). From line 403:

The results presented above for the varying intercept and varying slope scenarios were qualitatively robust to changes in carrying capacity (K) and dispersal (m) (Supp. mat. S2: Sensitivity Analyses). In general, larger carrying capacities decrease the parameter space where populations go extinct, likely due to lower demographic stochasticity [e.g., 49] and more efficient selection [e.g., 50]. In contrast, higher dispersal lowered the maximum spatial gradient slope where expansion was possible, likely because this introduced a higher mean fitness cost of dispersal [e.g., 14].

With a varying intercept, the intercept of the "separation line" between extinction and expansion increases with larger K (facet labels below are K values — K = 20, 30, 100, or 150; in the main model, K = 50).



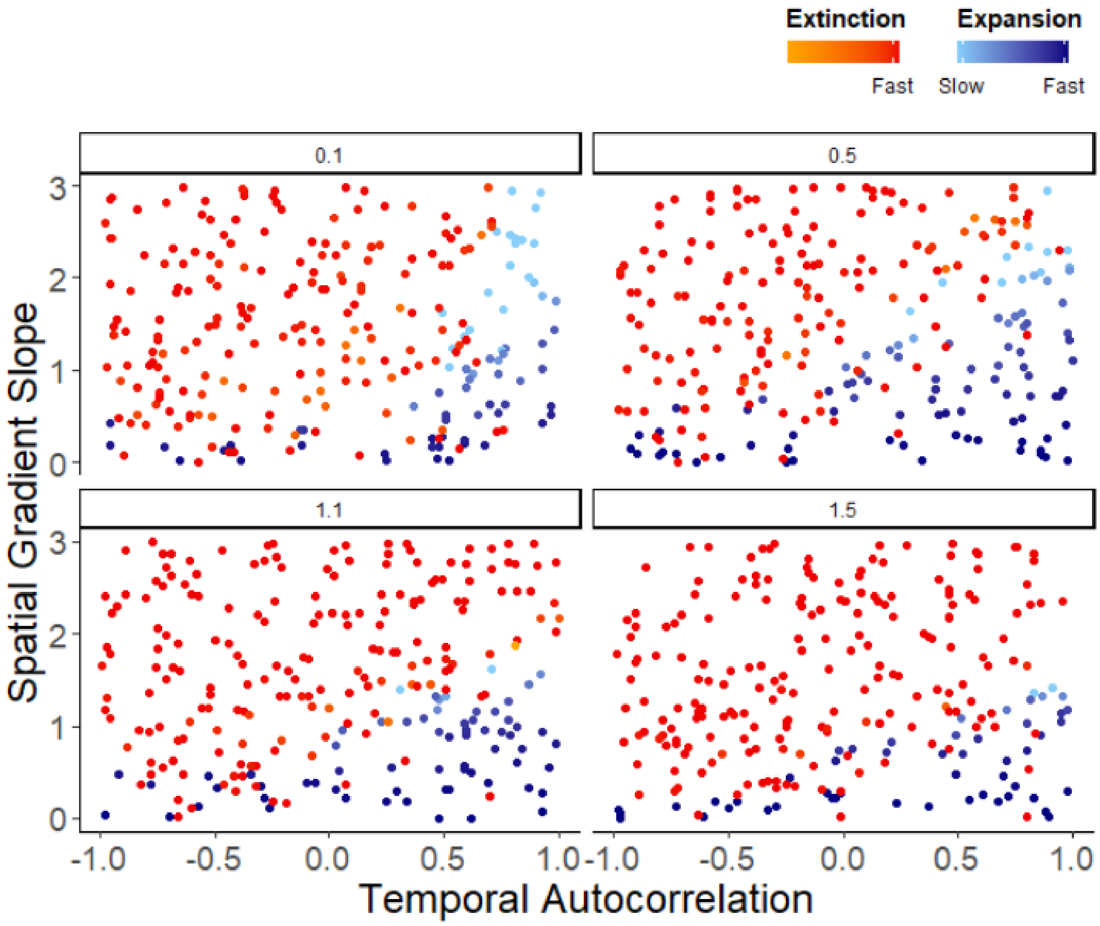
With a varying slope (below), lower K values increase extinction risk (range width = 0 in the plots below implies extinction) which causes fewer simulations to form stable range limits. With higher K values, most simulations form stable range limits, but higher population sizes slightly increase range width and lessen the effects of spatial gradient slope. (Again, facet labels below are K values.) The hollow points below (as in Fig. 4 in the MS) are simulations where populations reached the edge of the landscape — these populations experienced a stretch of "benign" generations with temporal variation low enough to allow them to reach the landscape edge (thus ending the simulation) before there was a large temporal fluctuation.



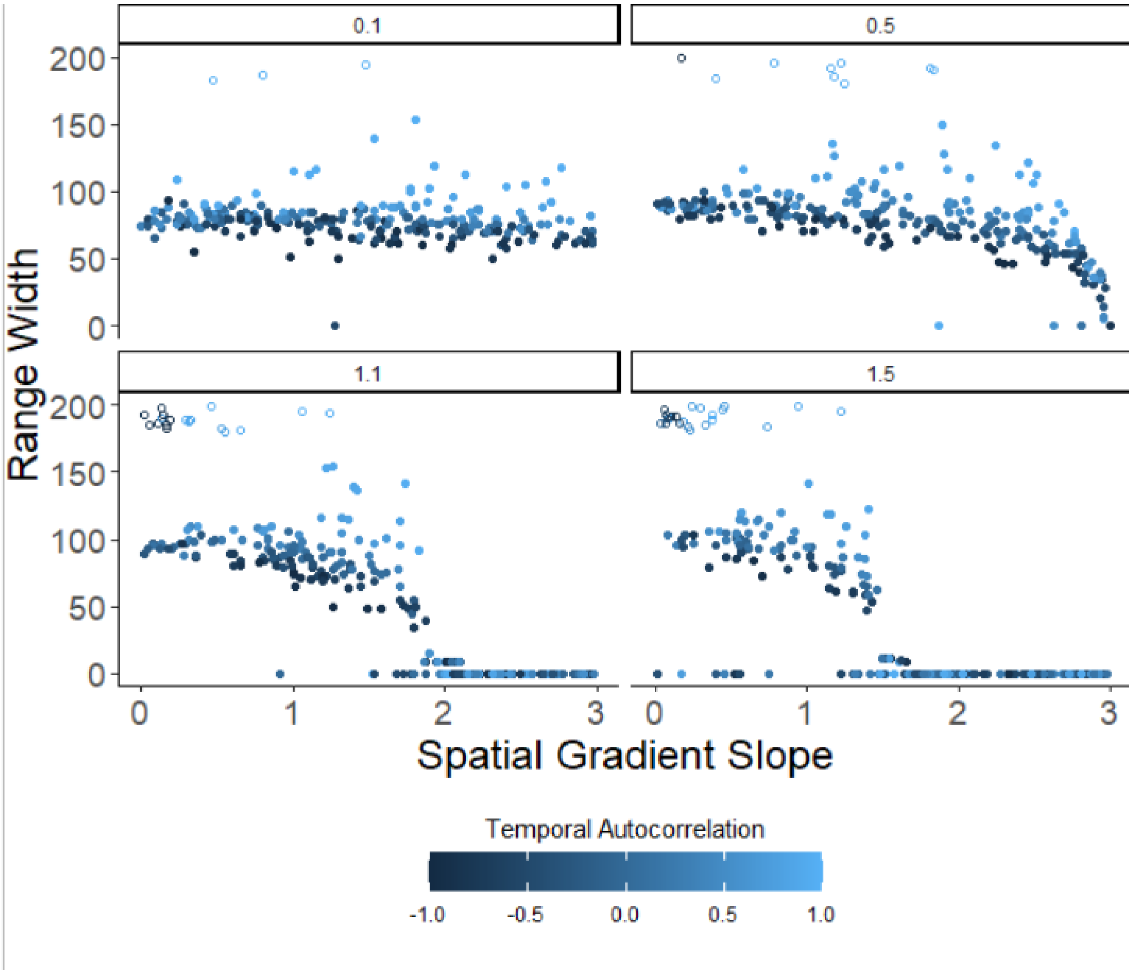
We now note at line 244: *It is important to note that dispersal (m) and carrying capacity (K) are both hard to estimate and highly variable in natural populations; for our main simulations we use values for these parameters near the mode of their distributions as estimated by Polechová and Barton (2015). We explored the sensitivity of our results to* *varying m and K in Supplementary Material; the qualitative results presented below were largely robust to changes in these parameters.*

18. Third, maybe I am understanding this wrong but the migration rate seems very high. For Poisson distribution with mean of 0.8 more than half of the individuals (~55%) should move at least one deme. A general rule of thumb (which comes from Fst) is that populations will not be able to spatially differentiate if there is more than 1 migrant per generation. This high of a level of migration may mean that there is a lot of genetic swamping and no ability to adapt to local environmental conditions. To address this one would need to greatly reduce m and measure Fst across the range to the measure whether failure to adapt to local environmental conditions is a result of swamping versus lack of local genetic variation. Note that this level of migration may be bolstering Ne despite the small population size, leading to high levels of neutral genetic variation but very little potential to adapt.

    Indeed, this is another parameter where sensitivity analyses are warranted, as migration/dispersal rates can differ greatly among taxa. When we vary migration (m), in the varying intercept scenario, we see that the main effect of increasing m is to lower the maximum spatial gradient slope where expansion is possible (Figs. S5, S6). (Facet labels below are m values — $m = 0.1, 0.5, 1.1, 1.5$; in the main model, $m = 0.8$.)



In the varying slope scenario, increasing dispersal (*m*) mainly serves to truncate the range of spatial gradient slope parameter space where persistence is possible (see plot below) — i.e., high dispersal exacerbates the negative effects of maladaptive gene flow and the fitness cost of dispersal and causes extinction with steep spatial gradient slopes. Regardless of dispersal levels, as we see in the main simulation results, most surviving populations end up with stable range limits due to increasing temporal variance.



19. Finally, the envisioned life-cycle of the model is unclear. Why are there separate sexes? Are offspring created by random sampling of gametes or by random sampling of parents? If parents and not gametes are sampled at random then this could lead to large variation in reproductive success and dramatic decreases in effective population size. What is the motivation for using relative fitness in males?

    The main motivation for modeling a diploid, sexual species is the increased relevance to the majority of plant and animal taxa that are of conservation or management concern. For reproduction in the model, each female's expected fecundity (realized fecundity is then drawn from a Poisson distribution) is estimated based on Eq. 5. If the Poisson draw returns > 0, her mate is chosen from the pool of males in her patch, with P(mating) proportional to a male's fitness as calculated by Eq. 1. Weighting males' probability of mating by their matching to environmental optima is based on the assumption that male phenotypes (e.g., size) depend in part on this matching, and that these phenotypes often influence mating success (e.g., Partridge and Farquhar 1983; Oddou-Muratorio et al. 2018). Gametes are generated from parents and subjected to mutation and recombination before fertilization.

References:
Partridge, Linda, and Marion Farquhar. "Lifetime mating success of male fruitflies (Drosophila melanogaster) is related to their size." Animal Behaviour 31, no. 3 (1983): 871-877.
Oddou‐Muratorio, Sylvie, Julie Gauzere, Aurore Bontemps, Jean‐François Rey, and Etienne K. Klein. "Tree, sex and size: Ecological determinants of male vs. female fecundity in three Fagus sylvatica stands." Molecular Ecology 27, no. 15 (2018): 3131-3145.

Minor Comments:

20. **Ln 72**: One part of the introduction that really confused me was what was meant by "fluctuations" vs "stochasticity" and if these two terms were being used interchangeably. The mention of predators and pathogens in this sentence made me think "fluctuations" was referring to an sinusoidal phenotypic optimum dynamics which are approximately coevolution (e.g., Takahata and Ni 1990), however, as far as I can tell this is not what is meant by "fluctuations".

    We mean "fluctuation" in the general sense, akin to "change." We now clarify at line 67: *Biotic environments also fluctuate (stochastically or regularly) through time as populations of predators, mutualists, pathogens, and competitors wax and wane.*

21. **Ln 75**: I was confused by what references are included here. Why are 25,26,31, and 32 not included? I understand these are discussed later but it made me think that I was misinterpreting this sentence.

    Those references were omitted simply because they were expanded upon in the subsequent sentences, but we agree it is useful to cite them here, and have done so at line 70.

22. **Ln 78**: Once again I am a confused by the connection between stochasticity and fluctuating selection. The two words are used interchangeably in the introduction but one refers to the (a) in equation 3 and one refers to the (d) in equation 3, right? But then d is related to a so I am having a hard time understanding what is meant by "fluctuation" and "stochasticity" in the context of this model.

    Here, and throughout, we are assuming that environmental stochasticity leads to fluctuating selection. In essence, for our purposes, environmental stochasticity *manifests* as fluctuations in phenotypic optima, and thus selection (i.e., line 249: *Environmental* *variation manifested as changes in the optimum phenotype of the quantitative trait across space and time.*)

23. **Ln 104**: "experience more temporal variation" or "exhibit higher temporal variance". Variance is a numerical value whereas variation is a property.

    Thank you; we have fixed this at line 116 to read *may sometimes be more temporally variable*

24. **Ln 108**: forward-in-time

    Corrected.

25. **Ln 109**: I would really appreciate more information on the connection to Polechova and Barton and Bridele et al. When you say `building on', it what way?

    We have clarified this connection at line 209: *Our simulation model builds most directly on the modeling frameworks of Polechová and Barton [14] and Bridle et al. [16], which explored the role of spatial environmental gradients on range dynamics. We build on the insights from those models by asking how temporal variation in the environment influences population spread and the formation of range limits .*

26. **Ln 119-125**: I had to read this paragraph many times to processes it, its hard to figure out what each "two" refers to and what the second is to the first in line 120.

    Agreed, this paragraph was clunky; we have made substantial edits to this section of the Introduction and think it is now much clearer.

27. **Ln 123-124**: I think one thing that confused me is you never mention the leg of the autocorrelation. It is implied in line 123-124 that the lag is also kept at 1 generation.

    Yes; we have clarified this at line 267: *Temporal autocorrelation could be positive ($0<a\leq1$), negative ($-0.99\leq a<0$), or uncorrelated ($a=0$) temporal stochasticity, with autocorrelation lag equal to one generation.*

28. **Figure 1**: Add panel labels.

    Labels have been added.

29. **Figure 1**: Rather than or in addition to the different colours could you show example dynamics over time for positive, negative and no autocorrolation?

We appreciate this suggestion and have incorporated plots showing different autocorrelation patterns into Figure S1:

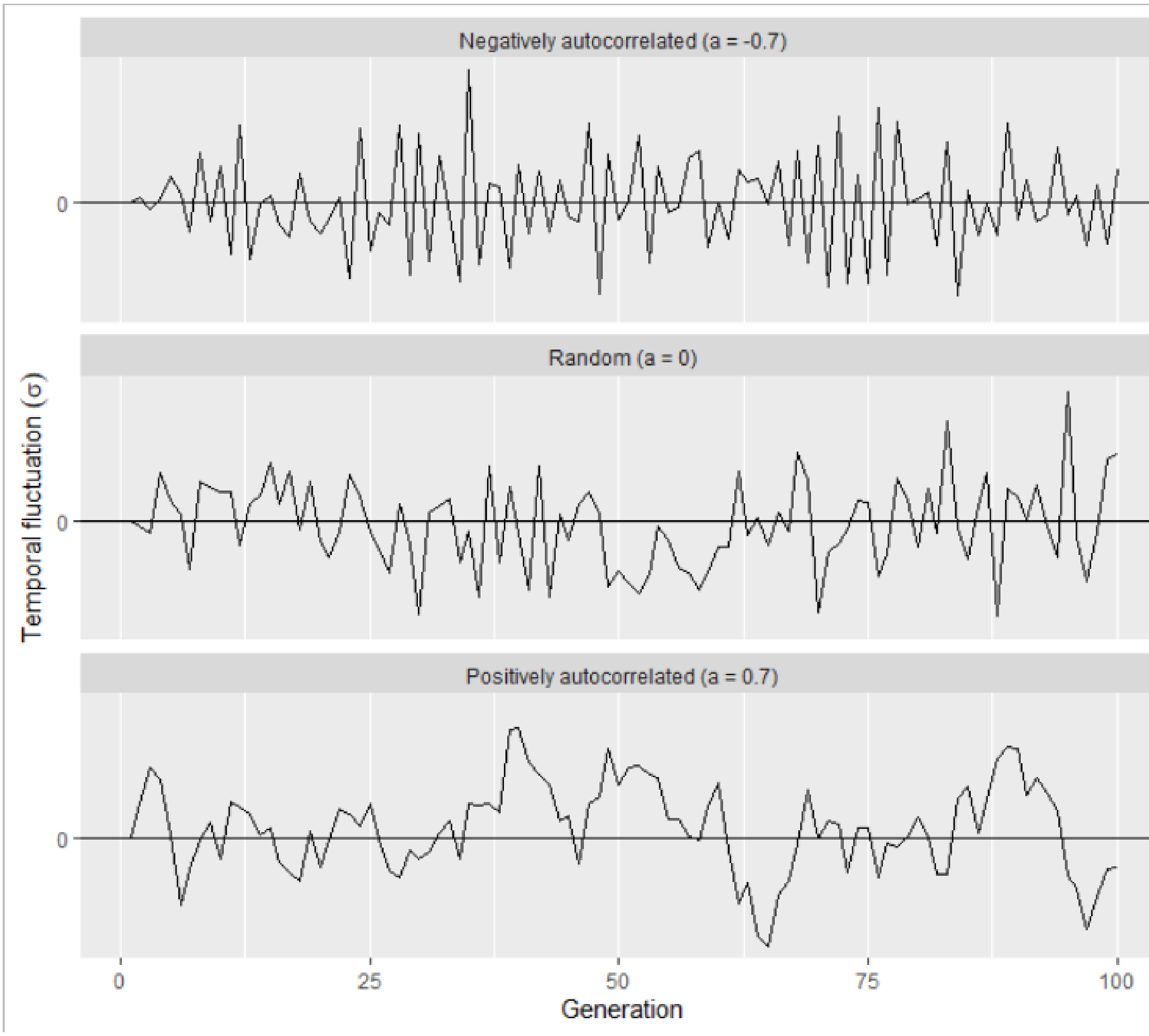


Figure S1. *Examples of temporal fluctuation patterns in the intercept (for varying intercept scenario) or slope (for the varying slope scenario) of the spatial gradient under negative (a = -0.7), random (a = 0), and positive (a = 0.7) temporal autocorrelation.*

30. **Ln 158**: SLiM, a forward-in-time population genetic simulator,

We have changed line 205 to read, *we used SLiM [50], a forward-in-time population* *genetic modeling software, to build a complex, individual-based simulation of species range dynamics in a spatially and temporally varying environment.*

31. **Ln 170**: Why model separate sexes? This will only decrease the effective population size. The role of separate sexes never seems to come up again.

    The main motivation for modeling a sexual, diploid species is the increased relevance to the majority of plant and animal taxa that are of conservation or management concern.

32. **Ln 173**: Units on the mutation rate, is this per locus or per allele? On that note why model diploids? What is dominance?

    Mutation rates in SLiM are per base position per generation (in a gamete, so per allele); we have clarified the units at line 218. Regarding ploidy, SLiM only models diploids, but this also makes the model more relevant to the majority of plant and animal taxa of conservation and management concern. The dominance coefficient for all QTLs in the model is 0.5 (i.e., no dominance, line 217).

33. **Ln 175**: I am guessing this is between any two consecutive loci right?

    Indeed; we clarified this at line 221.

34. **Ln 180**: This isn't actually a "carrying capacity". The carrying capacity of the population is the equilibrium population size which is not what will happen in this model.

    We have revised line 225 to read: *Each patch hosted a local population subject to density-dependent regulation, with carrying capacity (for a perfectly adapted population) constant across the landscape (here, K = 50).*

35. **Ln 180**: Once again K=50 is pretty small. Motivation for this?

    The main motivation was computational efficiency, but sensitivity analyses varying K are now discussed (see comment 17).

36. **Equation 1**: Can you include the intercept of the phenotypic optima explicitly in this equation? E.g., b0+b1 x. Given your genome structure what is the range of z relative to the quadratic? You try to capture this via the linear selection gradient later on but a figure of this would be good? How do you know that W won't go negative?

    The equations in the paper have changed substantially and this equation (5) now includes a load term, $L(\theta)$, which captures the changes in optima as you request here and relates directly to our new analytical model of genetic load (Eq. 1-4).

Regarding z, we place no constraints on z values except selection (i.e. mathematically, phenotypes could theoretically take on any real value, but will be constrained in practice due to selection). Also, it is true that the fitness function can become negative. But in the simulation model, any $W_i < 0$ is forced to 0.00001 (as in SLiM, Poisson λ must be > 0), which we now clarify at line 237.

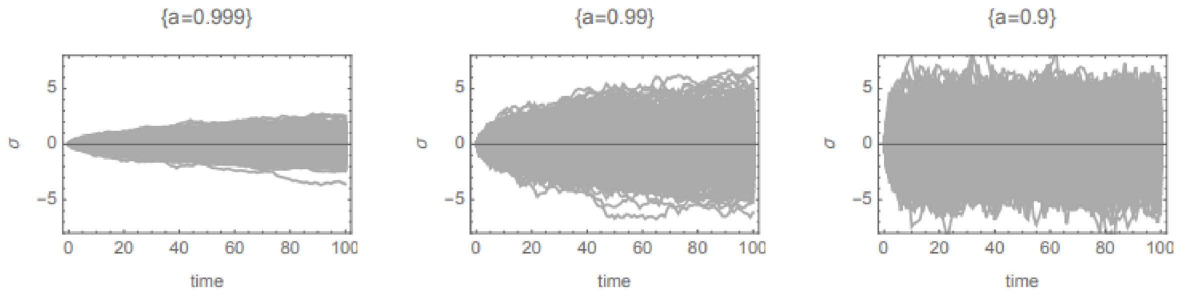
37. **Ln 188**: I'm not sure I 100% understand. Is each female guaranteed to mate? If so, are all of the offspring of a given female full-sibs or are gametes drawn at random? If male-female mated pairs are drawn at random and not gametes being drawn at random this could result in a significant random variation in reproductive success which could decrease population size even further.

    (partly from Comment 19) For reproduction in the model, each female's fecundity (λ for a Poisson draw) is estimated based on Eq. 1. Every female is *not* guaranteed to mate (i.e., some will draw 0 from the Poisson distribution). If the Poisson draw returns > 0, her mate is chosen from the pool of males in her patch, with P(mating) proportional to a male's fitness as calculated by Eq. 1 (thus, all of a female's offspring are full sibs). Gametes are generated from parents and subjected to mutation and recombination before fertilization / karyogamy. This process strikes us as more biologically relevant to most sexual diploids than one where a focal female's gametes are randomly paired with X number of male gametes in a single generation, and allows for a more realistic incorporation of demographic stochasticity into the model.

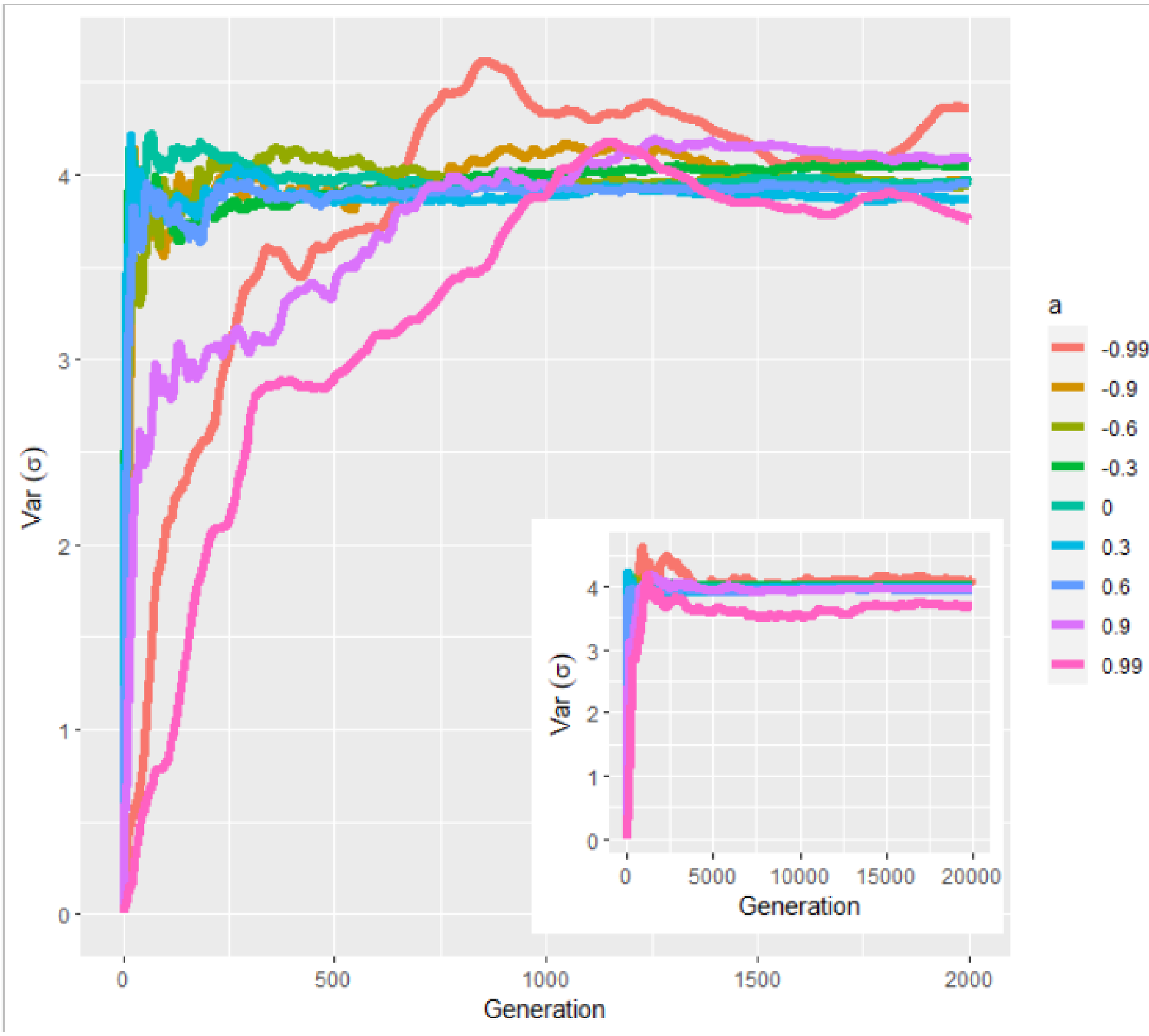
38. **Ln 199-200**: This illustrates some of the challenges of density-dependent selection in modelling range dynamics. If you used a density-independent definition of fitness then the selective disadvantage of a particular phenotypic mis-match would depend only on the slope of the gradient and not on the variation in population size among patches.

    Please see comment 16.

39. **Ln 210-211**: This does not make sense to me. Consider for example \sigma_1=0+(1-a^2)^{0.5}N(0,\tau). The variance in \sigma then is: var(\sigma)=(1-a^2)\tau^2. Is this some equilibrium property? How long does it take to reach equilibrium? Should sigma_0 not be set to 0 then but rather a random number? See example figures below for different values of a given \tau=2.



Setting $d = (1 - a^2)^{0.5}$ (based on Ripa & Lundberg 1996) results in $Var(\sigma)$ within each simulation being (roughly) equal across the timespan of our simulations. It does take some time for $Var(\sigma)$ to equilibrate/converge in our simulations, and at extreme autocorrelation values (e.g., a = -0.99 or 0.99) this equilibrium process is sometimes slower (especially with strongly positive autocorrelation). Below we show $Var(\sigma)$ over time/generations for different values of a (averaged across 10 replicates per a value). Main plot shows generations 1-2000, and the inset shows this pattern across the timescale of our simulations (20,000 generations).



We have clarified at line 269: *We set* $d=(1-a^2)^{0.5}$ *so that Var(ε) over the full time series used in our simulations was approximately equal for all values of a.*

40. **Ln 214 and 218**: Capitalize B and C to match figure caption

    We now use lowercase panel labels in accordance with journal convention.

41. **Ln 216**: I think equation 1 should explicitly include the intercept.

    Please see comment 36.

42. **Ln 228**: Just to make sure I understand, for every point in space you include two points in the dataset, one which is the current optima and one which is the long-term average optima. Then you run a linear regression to calculate the slope. If my understanding above is correct, I don't understand why the long-term average optima is included. This makes the temporal variation in the gradient hard to interpret. Note: relative fitness is a function of population size, so there could be a range edge effect.

    This section is indeed a bit unclear. Essentially we want to give the reader a sense of how the level of temporal stochasticity modeled in our simulation compares to temporal stochasticity, and the attendant fluctuations in selection, that we see in nature, and thought that selection gradients were a straightforward way to do that given that we have many estimates of those in natural populations. The long-term optima phenotype is included because we need two points to calculate the slope, and thus the selection coefficient / gradient. We have reworded this section in hopes of increasing clarity (Supp. mat. S2).

43. **Ln 229**: Not that it will matter too much with a \sigma_P of 1.7 but why are you multiplying? I would think you would want to divide by the phenotypic variance (not the SD) to get the the variance in fitness as a function of the phenotypic optimum in units of phenotypic standard deviations.

    Based on Lande and Arnold 1983, SD-standardized selection gradients are calculated as $\beta_\sigma = \beta\sigma_P$, which results in a gradient in units of the phenotypic SD. Standardized selection differentials, on the other hand, *are* calculated as $\frac{S}{\sigma_P}$, but we do not use differentials here.

44. **Ln 237**: If you are only measuring 40 patches away from the centre then is there 4 points (2 from the left and 2 from the right) of your linear regression? Or do you run multiple replicate simulations etc.

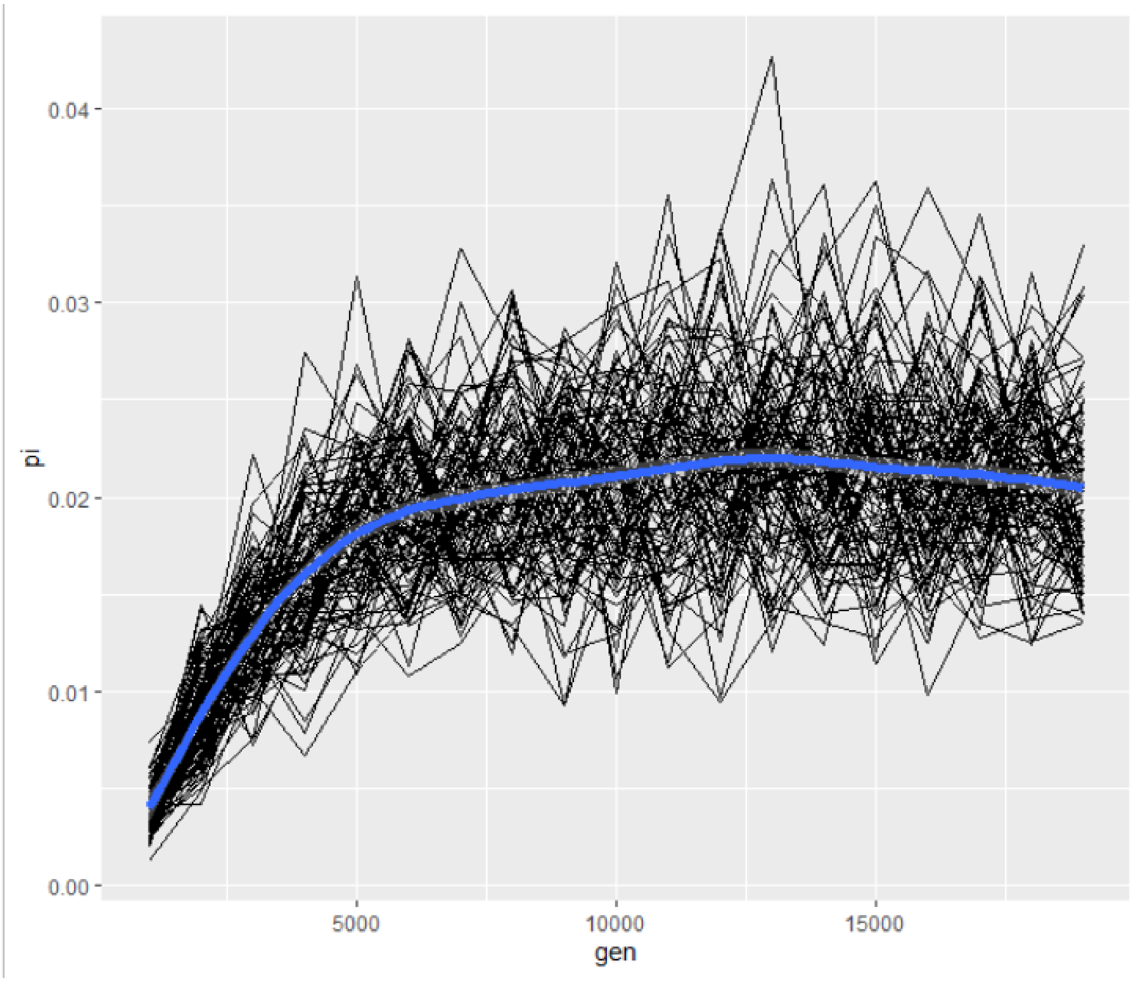
    Here, we simply mean that temporal variance, τ, equals 4 at patches 40 away from the landscape center in the varying slope scenario (whereas $\tau = 4$ everywhere in the varying intercept scenario). We have clarified the text at line 275.

45. **Ln 246-249**: What do you mean by "minimizing the impact of initial conditions". To me this means that I want the simulations to reach an eco-evolutionary equilibrium such that the initial patch in the range remains near this equilibrium throughout the main simulations. Is this true? I would think not as the reduction in population size would result in a loss of genetic variation. Does the genetic variation change significantly at the

central patch during the simualtions? What about the population size? Do you have graphs that show that you do indeed reach an equilibrium during the burnin? Why use a different carrying capacity and stochastic distribution? Why these values?

This phrasing was vague; line 289 now states: *The goal of this burn-in period was to generate independent replicates of genetically variable source populations that had evolved in landscapes of similar spatial and temporal heterogeneity for each simulation run, while allowing enough time for the different simulation burnins to converge on similar levels of genetic diversity. During the burn-in, mean heterozygosity of neutral mutations (π) in the central population usually reached an equilibrium by 10,000 generations (Fig. S2).*

The plot below shows how π equilibrates across the burnin period and is now found as Figure S2 in Supp. mat.



The burnin spatio-temporal landscape was intended to generate independent replicates of genetically variable source populations for each simulation run — essentially we wanted to let these burnin populations evolve across a spatially and temporally heterogeneous landscape to provide the source for the population founding the range expansion simulation. To begin the range expansion simulation, we have to pull founders from somewhere, and pulling randomly from across the source/burnin populations seemed to us a reasonable choice.

46. **Figure 2D**: How is this measured? Is this for selected, neutral, or genome wide? From line 251 I am guessing this is the neutral pi. Can you plot the selected pi as well? What does the genetic variance in each deme look like? What does Fst look like? I am asking some of these questions because I'm a little surprised by the dynamics of genetic variance especially to how C and D relate. Why do the earlier bottlenecks at say gen 3770 not have similar effects on genetic variance as the one at 3790?

    Genetic variance in Fig. 3d (originally 2D) is variance in the focal trait conferring adaptation to the gradient (now specified in the legend), taking the median across all demes. Because this trait is purely controlled by genotype (i.e., no environmental influence), V_G is simply the phenotypic variance ($V_{G} = V_P$). With the updated plots in Fig. 3b-d, we can see how a large environmental deviation around generation 4220 caused a large drop in abundance and V_G, both of which then begin to slowly recover before another “bad year” causes extinction. The relationship between temporal variation, abundance, and V_G is clearer in these updated plots partly because the original plots showed mean, instead of median V_G , and these mean values could be highly skewed by patches with very few individuals and thus inflated estimates of V_G.

47. **Figure 2D**: Why does the orange line end before extinction?

    In Fig. 3d, the line ends one generation before extinction because abundance = 0 at extinction, and thus there is no measure of V_G.

Referee 3

In this work, the authors explore how environmental stochasticity affects the interplay of local adaptation and demography underlying species’ range limits. They do so by including random temporal fluctuations in the intercept, slope, or both, of a spatial cline in optimum phenotype, representing spatio-temporal variation in the environment, in an individual based simulation model with explicit genetics (using the SLiM software). In doing so, they extend previous simulation and/or analytical work by Bridle, Polechova, Barton, and others, to also include temporal stochasticity in the environment.

Their main findings are that random fluctuations in the cline intercept (in other words: identical fluctuations through space) do not lead to a restricted range, and allow range expansion over steeper gradients when fluctuations are more autocorrelated. Fluctuations in gradient slope, leading to larger fluctuations away from the range core, do lead to stable (albeit fluctuating) range limits, with a smaller range under steeper (mean) gradient, and less autocorrelated fluctuations. Combining fluctuations in slope and intercept further reduces the range size and increases extinction risk.

These results make sense and seem overall well interpreted, I enjoyed how they are presented in the figures, and I found the paper well written and easy to follow. Still I think the paper could be improved by dealing with the points below.

48. First, I think one should always be careful with priority claims, such as “all evolutionary range limit models to date assume temporally constant environments.” (89-90) Perhaps this is true if you think of models with multiple demes along a near-continuous gradient as here, but not for two-patch models that are often used to investigate what happens in marginal populations, which is critical to range limits. I’m thinking for instance of Holt et al (2004 Am Nat Temporal Variation Can Facilitate Niche Evolution in Harsh Sink Environments), but there may be more recent ones by these or other authors.

    Yes, this is certainly true and this claim has been tempered, and appropriate citations added in the Introduction. E.g.,

    However, Holt et al. [38] showed how temporal environmental fluctuations can have positive effects and facilitate adaptation in sink habitats if the sink environment becomes more benign long enough to increase population growth and adaptation.

    If the environment is changing directionally through time (e.g., glacial advance, warming temperatures), theory has shown how the rate of environmental change, the amount of genetic variance, and the steepness of an underlying spatial gradient all influence population persistence [43,44].

    Recent work by Holt et al. [43] provided some of the first insights into the influence of non-directional temporal variation on range dynamics, showing how temporal variation in competition can modulate the size of an established species’ range, and that temporal variation in immigration increases the probability of establishment in sink habitats.

49. Second, I appreciate the value of simulations, but it would be better if you also used even very simple analytical predictions to guide their interpretation. More specifically, most of your results certainly arise from a load that random environmental fluctuations in the optimum impose on the population. These fluctuations causes an expected reduction in mean fitness, for which there are quite a few theoretical predictions in the literature (e.g. Burger & Lynch 1995 Evolution, Lande & Shannon 1996 Evolution, Chevin et al 2017 Am Nat). Some of these predictions are more complex than others, depending on how they account for evolution, for environmental autocorrelation, etc. But even the simplest one that neglects evolutionary responses to temporal fluctuations (something like (s/2) tau^2 in your model, derived by taking the expectation of squared deviations from the fluctuating optimum, assuming constant z) would probably do a decent job here. In the random intercept scenario, this expected load is constant over space, which is equivalent to reducing the growth rate by the same amount everywhere, regardless of local adaptation. In the fluctuating slope scenario, the stochastic load changes over space (with tau), such that even a perfectly adapted mean phenotype (on average) may have negative expected growth rate in extreme environments. I think that dealing with this load explicitly would make results much clearer.

    Thank you for this very insightful suggestion! We now include a simple analytical model to help guide our interpretation of how temporal variance affects genetic load and subsequent population persistence (Materials and methods: *Expected genetic load due to temporal variation*; line 169).

Minor points

50. **81-83**: perhaps also cite here the broad analysis across birds & mammals by de Villemereuil et al (your ref 39), since it explicitly tested for a fluctuating optimum.

    This reference has been added.

51. **126-134**: you could mention that both models imply a spatial autocorrelation = 1 (at least this is true of the variable intercept, the other may need checking)

    Indeed; Reviewer 1 also noted this and we have included the following in the Discussion at line 532: *Lastly, temporal stochasticity as modeled here is strongly spatially autocorrelated — i.e., adjacent patches always experience temporal variation in optima in the same direction, positive or negative. It would be fruitful to relax this assumption to allow less spatial autocorrelation in temporal variation, as the effects of gene flow on adaptation would likely differ.*

52. **182 (eq. 1)**: this fitness function allows negative values, how did you deal with that?

    In the simulation model, any $W_i < 0$ is forced to 0.00001 (as in SLiM, Poisson λ must be > 0), which we now clarify at line 237.

53. **201 (Eq.2)**: this depends on the value of s, right?

    Indeed; it is now clarified at line 253 that throughout we are using $s = 0.125$.

54. **225-228**: It is not very clear where this argument on selection gradients is coming from. Perhaps it would help to write that selection gradients depend on s times the deviation from optimum in such a model.

    This section was definitely opaque; we have reworded it in hopes of making our aims more clear (note it has moved to Supp mat. S2 due to space constraints).
    In natural populations, selection on traits is often measured using standardized linear selection gradients, β*, which describe how an individual's fitness varies with its trait value (Lande and Arnold 1983). The magnitude of temporal fluctuations in these gradients can be described with* σ_β *(e.g., de Villemereuil et al. 2020). To compare the fluctuations in selection that emerge in our models to what is observed in natural populations, we estimated* β *and* σ_β *given the parameter values used in our simulations. We first generated 100,000 draws from* $\Phi_t \sim$ *Normal(0, 4); these values represent deviations from the long-term mean optimum across 100,000 generations, assuming no temporal autocorrelation (i.e.,* $a = 0$*). Following (Lande and Arnold 1983), linear selection gradients were estimated for each generation as the slope coefficient of the linear regression of mean-relativized fitness on the phenotypic trait values of two hypothetical phenotypes: (1) a phenotype perfectly adapted to the current trait optimum and (2) a phenotype adapted to the long-term mean optimum. Per (Lande and Arnold 1983), we standardized this gradient by multiplying it by the phenotypic standard deviation,* σ_P*, which we estimate at 1.8 (the mean* σ_P *in the central patch in our "varying intercept" simulations after 20,000 generations). With this simulation of temporal stochasticity across 100,000 generations, the absolute value of the standardized linear selection gradient,* $|\beta|$*, averaged ~0.10 across generations, and* $\sigma_\beta \approx 0.14$*. These selection gradients are well within the range of selection gradients and their variance in nature, [median* $|\beta| = 0.16$ *in Kingsolver et al. (2001);* $\sigma_\beta = 0.099$ *in de Villemeruil et al. (2020), averaging across birds and mammals]. In the varying slope scenario,* $|\beta|$ *and* σ_β *increased with distance from the landscape center as temporal variance increased, matching temporal variance of the varying intercept scenario 40 patches away from the*

landscape center (Fig. 2).

55. **270-272**: "the rate of range expansion slowed as spatial gradients steepened and environments became less positively autocorrelated." But I seem to see slower expansion under higher autocorrelation in Fig 2A

    Yes, this was phrased poorly. Line 334 now states: *The ultimate fate of the species in each simulation was either eventual extinction or continual expansion; stable range limits did not form, though the rate of range expansion slowed as spatial gradients steepened (light blue points in Fig. 3).*

56. **279**: what do you mean by "stationary individuals"? I find the term misleading, as stationary has a meaning in a context of stochastic processes, which you also investigate here.

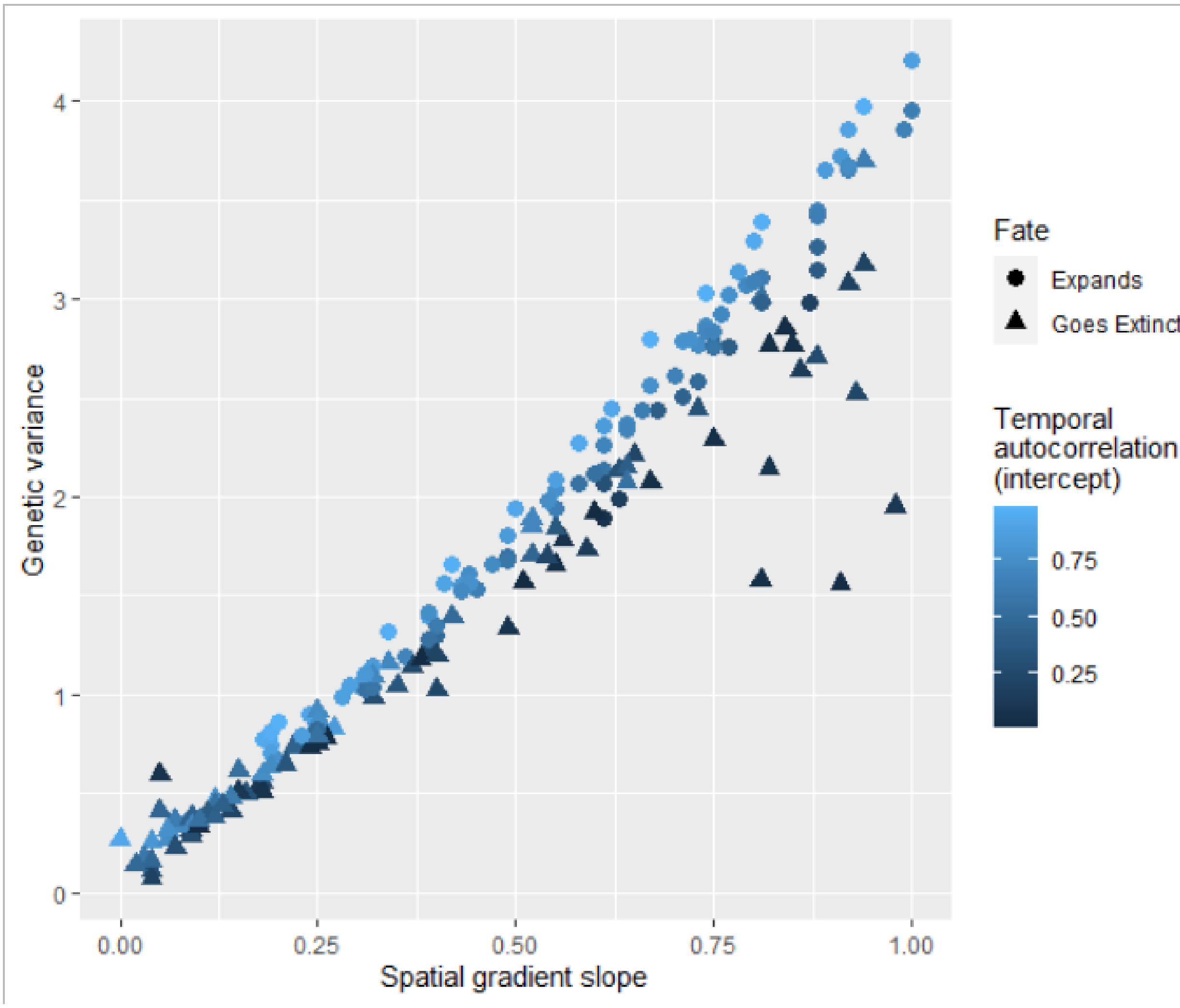
    This was indeed confusing; we have deleted the term "stationary" in this sentence as it was not necessary. We simply meant that individuals who were well adapted to their current patch and did not migrate from that patch.

57. **287-288**: "Thus, temporal stochasticity introduced an extinction risk due to fluctuating phenotypic optima that had strong effects on mean fitness". Here you would need to cite the relevant literature recommended above

    These have been cited (line 348).

58. **366-368**: "This is because steeper spatial gradients increased genetic variance across the landscape via gene flow, which better equipped populations to withstand temporal fluctuations in optima". Interesting, but can you show evidence for that?

    We have now added the plot below as Fig. S11, showing how mean V_G changes with spatial gradient slope in the varying intercept / random slope scenario and influences species' extinction.



59. **402**: A reference seems required for this statement: “Increasing temporal variability is one predicted (and observed) consequence of contemporary climate change”

    Indeed; we have cited the the most recent IPCC report now at line 460.

60. **422-424**: “there is a critical threshold of temporal environmental variance that can stop range expansion and enforce a stable range limit.” Yes, and you could use the argument outlined above to explain this: temporal fluctuations reduce the expected growth rate in a predictable manner, causing an additional load to that due to local maladaptation. For instance, I predict that the fluctuating slope model would produce stable limits even in the absence of any gene swamping (which you could check by setting dispersal to 0), simply because environmental fluctuations that are too large cause a negative expected long-term growth rate, and thus the populations decline on average in extreme

environments.

We have incorporated this idea of increasing genetic load into the paper and it has definitely improved clarity. And you are absolutely correct that temporal stochasticity will produce range limits even without gene swamping, as we see range limits form even with a flat spatial gradient in optima ($b = 0$).

Appendix D

Associate Editor

Thank you for sending a resubmitted version of your MS, as well as a detailed response to referees' comments, including your additional efforts to establish the generality of your model.

In particular, I appreciate the addition of some analytical approximations for load to illustrate a connection to existing theory (rather than as a full treatment of the load generated in the main simulations). This I think is appropriate given Rev. 3's comments made it clear that s/he thought a simple model would be a good enough approximation to make the key point about the impact of fluctuations in intercept or slope on the load and its spatial pattern.

Your resubmission was sent to one of the original reviewers to consider, who continues to highlight several concerns with the MS - these are detailed in the attached document. Although I do not share all of these concerns, I would like to give you the opportunity to respond to them, and to modify your MS accordingly. In particular, - you might consider changes to ensure that the function of the load model, and its limitations are made more clear.

Many thanks again for submitting your work to Proc Roy Soc, and for your willingness to make extensive revisions on resubmission.

Thank you for these encouraging words. We have addressed the reviewer's concerns below, including clarifying certain aspects of the text and providing additional context for certain modeling choices. In particular, we have taken care to more clearly detail the purpose and utility of the simple analytical model in the paper (eg, lines 138, 174).

Please note that we did need to move the "future model extensions" paragraph from the Discussion to Supp. mat. due to space constraints, per instructions from the Proc B office.

Referee: 2

Dear Editor, February 14th 2022

Thank you for the opportunity to review "Increasing temporal variance leads to stable species range limits". In this article the authors use simulations in the population genomics simulator SLiM to explore the impact of temporal environmental fluctuations on the expansion of ranges into novel habitats. The authors examine expansion across a linear environmental gradient contrasting temporal fluctuations in the environmental intercept (impacting the spatial mean) and environmental slope (impacting the spatial variance) on outcomes such as range width.

Temporal variation is increasingly appreciated for its impact on eco-evolutionary dynamics as the authors highlight in their introduction. The basis of this paper is therefore well-founded and relevant to understanding adaptation in a changing world. However, there are several crucial errors in the development and design of the theoretical models that jeopardize the validity of their results and the conclusions drawn. Given these errors I will focus my review solely on the motivation and model formulation rather than the results.

Thank you for these comments; we respond to each below.

Major concern 1: Simulations are not hypothesis-driven. The authors do a great job of drawing on the existing literature to emphasize the potential impacts of temporal fluctuations on demography and adaptation and hence the possibility of its impact on range expansion. These existing results are never synthesized however into specific hypotheses that clearly motivate the model. Rather, the model and the results are largely observational. While there is deep value to observational science there are distinct risks to employing an observational approach in the context of simulation-based science. In particular, as illustrated by this paper, taking an observational approach greatly reduces the ability to check the intuition of an outcome to identify flaws in model design. It also limits the ability of the reader to understand and justify the design of the simulations and results. For this work to achieve what it could, I highly recommend the authors explicitly formulate the hypothesis underlying the contrast between varying slope and varying intercept designs.

While we appreciate this argument, we do not agree with it for two main reasons. First, we would argue that the aim of the model was not to test explicit hypotheses with (simulated) data. Rather, our simulation model is an attempt to answer questions (e.g., how does temporal variation influence range dynamics?) that is motivated by the theory and evidence outlined in the Introduction. We have made this clearer at line 118: *The theoretical and empirical work above indicate that there exists much potential for temporal variation to influence species' range dynamics. In this paper, we use analytical and simulation models to ask how temporal variation in the environment influences population expansion and the formation of range limits across spatial gradients.*

We do note that our models, while motivated by questions, have *generated* testable hypotheses (see Discussion). Much valuable theoretical work has followed a similar tack.

The second point is the more philosophical issue that any hypotheses fit into the manuscript at this point would be necessarily presented as *a priori* hypotheses, but would actually be *post-hoc* hypotheses — we cannot unsee our model results. We are wary of taking such an approach.

Major concern 2: Load. There are a number of interconnected errors and omissions in the section on “expected genetic load due to temporal variation.

First, load is defined relative to a specific model of fitness. This model of fitness should be clearly defined prior to equation 1.

We have now clarified this at line 186: *As in Lande and Shannon [36], under a model of Gaussian stabilizing selection, we define genetic load for an individual i as*

Also equation 1 is a function of both z_i and theta_{x,t}.

Thanks for catching this; amended.

Second, it should be clarified that this analytical model is not a dynamical model. It does not model phenotypic evolution. In lines 174-175 the focus on a population obscures this important distinction.

This is now clarified on line 174: *We first present a simple analytical model to describe how our two geographic modes of temporal environmental variation (varying intercept and varying slope) affect expected individual fitness. This is not a dynamical model, but serves to help us build our intuition regarding how temporal variation in the environment may lead to fitness costs for populations at equilibrium.*

More importantly equation 3 is wrong. Under the model of environmental variation as described, load is distributed like a chi-squared distribution with one degree of freedom (the distribution of a square of a normally distributed random variable) not normally distributed. This can be visualized intuitively. If the load were normally distributed it could technically be negative which wouldn’t make sense. This same issue applies to equation 4.

Thank you for catching this mistake! Indeed, the distribution of L would be chi-squared. We have reworked this section (starting line 173) to focus on the expectation of L (i.e., E[L]). This does not change any of our conclusions, as the expected value of load is what we were concerned with, and plotting, from the beginning.

Finally, the title of the section (ln 169) as well as on line 201, the authors mention “expected genetic load”. However, given the non-dynamical nature of the analytical model the only expectation presented is the expected load of an individual that is perfectly adapted to the average environment of its patch. It is unclear under what conditions such an expectation is meaningful. For example, it may have little meaning within a model of evolution where phenotypes vary both within a population and over time.

We have amended the introduction to this section to make our aims for the model more clear (line 174): *We first present a simple analytical model to describe how our two geographic modes of temporal environmental variation (varying intercept and varying slope) affect expected individual fitness. This is not a dynamical model, but serves to help us build our intuition regarding how temporal variation in the environment may lead to fitness costs for populations at equilibrium.*

Major concern 3: Fitness definition.
I didn't dive into Bridle et al.'s definition of fitness but equation 5 is inconsistent with the previously assumed model of Gaussian stabilizing selection that was used to derive the expression for load. The presence of negative values of absolute fitness (Ln 237) is indicative of additional issues with the definition of absolute fitness. In its current form demography can influence the strength of selection directly (the selection coefficient will be a function of r Nx and K) and indirectly through drift. One way to avoid this is to define fitness by multiplying the logistic growth model by the expression for relative fitness (Gaussian curve or 1-load if following equation 5) rather than by subtracting it.

It is unclear to us why you deem equation 5 inconsistent with a model of Gaussian stabilizing selection — the load term encapsulates the fitness cost due to deviation away from the peak of a Gaussian fitness function. We have clarified this at line 239: *Thus, the first part of equation (5) describes standard logistic growth and the second part introduces the fitness cost proportional to the deviation of an individual's phenotypic value from the local optimum (i.e., genetic load) under a model of Gaussian stabilizing selection.*

There are several reasons we hesitate to change the fitness function. As we argued in our first response letter, one important reason is precedent — this is the (general) fitness function used in Polechová and Barton 2015, Polechová 2018, and Bridle et al. 2019, the models most closely aligned with ours, and the modeling frameworks that our manuscript seeks to build on (line 215). Comparisons to results from prior models, as we do at line 335 for Polechová and Barton 2015, would be difficult if our model used a different fitness function. Furthermore we feel it is an eminently reasonable assumption that maximum fitness decreases with increasing maladaptation, which is what occurs with this particular fitness function (which we now clarify at line 242: *The maximum attainable W_i thus decreases with increasing maladaptation, $L(\theta)$.*).

Regarding the potential for negative values of W_i, there are several points to note. First, the extreme maladaptation needed to generate negative values is infrequent in our simulations [e.g.,

~2% of individuals with $W_i < 0$ in a simulation of expansion across a moderate gradient (b = 0.5) with random temporal fluctuations (a = 0)]. Furthermore, the value W_i defined by Equation 5 is not technically the fitness value for an individual, but rather is the mean of a Poisson distribution from which an individual's fitness value is drawn. We have made this clearer in this section (e.g., line 243). These draws obviously cannot be less than zero, given the properties of the Poisson distribution. The piecewise function that sets any $W_i < 0$ to $W_i = 0$ (or more precisely, $W_i = 0.00001$ due to SLiM programming rules) simply reflects the fact that for a trait determining fitness, there exist upper and lower threshold values beyond which fitness is zero, due to the large magnitude of deviation from the phenotypic optimum, and the peculiar properties of fitness as a "trait" (i.e., there is a hard lower limit of zero). All values beyond those two thresholds will result in the same fitness value (zero), leading to a zero-inflated distribution of fitness as seen in most natural populations.

Additional suggestions:

Ln 214: While modelling explicit sexes may seem like a standard and innocuous choice, there are important implications to modeling a gonochoric population. Namely, this introduces the potential for allee effects, where populations fail to establish due to lack of mates. This certainly is a potential realistic element limiting range expansion, however in a simulation alone it will be difficult to separate the impacts of allee effects from other demographic features. Therefore, it may be helpful to first consider a hermaphroditic population. Or contrast the existing results to the hermaphrodidic case to isolate the potential consequences of allee effects.

This is certainly an interesting and important direction to explore, but we feel it is beyond the scope of this paper. It is also unlikely to change any of our main conclusions because in our model, we see that populations can establish and grow to large abundances beyond the "stable" range limit (i.e., colonization is not limited by Allee effects), but that these populations are ephemeral due to temporal variation in phenotypic optima that drives frequent extinction. However, we now note it in the "model extensions" section of the Supp mat.: *While our model incorporates Allee effects insofar as reproduction depends on the presence of at least one individual of the opposite sex within an individual's patch, modulating the strength of these effects would be fruitful.*