Decision for round #1 : ***Revision needed***

Moderate changes requested

Dear Authors,

Thank you for sharing this manuscript with us.

Based on the two reviews and my own read through, I believe this study asks a very interesting question, and I want to recommend it as a proof of concept that this question can matter. I also believe, like the reviewers, that the manuscript should be more precise about what it demonstrates exactly, and I would therefore request some effort at clarification.

I think all the issues that we collectively raise can be addressed with a moderate amount of work, essentially changes in writing, but that this work is important for the manuscript.

I would not consider it to require major revisions, but I ask that you seriously consider the suggested changes, and do not hesitate to contact us for further discussion if you disagree or have questions.

I would boil down the issues to two main aspects (but please also address reviewer comments that are not necessarily emphasized here)

1) one that is more a question of presentation (which would mainly require changes in the abstract, intro and bits of discussion),

2) another that is about better understanding the results themselves (and in particular which structure matters in structured IV).

R: We thank the recommender for his very thoughtful and helpful comments that we all addressed as specified below.

------

1) Presentation and scope:

The study makes some restrictive assumptions, at different levels, that may be intuitive for forest dynamics but may not be very representative of other systems.

As a proof of concept, I think it deserves to be out there and it certainly does not have to account for every notion of coexistence or IV in the literature; but it will certainly get a warmer reception if it explicits its choices (see reviewer comments).

Here is a suggested set of clarifications:

- A certain type of structure: since there are no interactions between neighbors nor limitations on dispersal range, spatial structure -- in the classical sense of who is next to whom, and spatial autocorrelation in the environment -- is irrelevant here.

You could have unequal distributions of environments without any autocorrelation and get the same results.

This is a perfectly acceptable starting point, but a very particular choice when talking about coexistence in space and about structured IV. It rules out many potential structural features and coexistence mechanisms.

R1: We fully agree: the fact that environmental variables are spatially auto-correlated does not play a role in our results, as now better emphasised l.176-179. We implemented spatially-autocorrelated variables for the sake of realism, for example if one wants to observe the species distribution in space (at the end of the simulations, species are grouped spatially in sites with similar environments because of environmental filtering), but the dynamic model is spatially-implicit in the end. We specify below why we however refer to spatial structure (see answer R7).

- A certain type of coexistence: since there are no density-dependent interactions nor any role of spatial structure, coexistence rests entirely on a) environment-driven variation in local competitive hierarchy, and b) a little bit of chance (see point 2 below). Again this automatically rules out many coexistence mechanisms (including others that are *also* known or interpreted as niche partitioning, e.g. stabilizing/density-dependent effects, even though they act in entirely different ways).

R2: Our community dynamics model takes its root in a well known model developed by Hurtt and Pacala (1995, Journal of Theoretical Biology). We acknowledge it lacks many ingredients of what could lead to coexistence, as highlighted by several previous studies and now better acknowledged in the text (e.g. l.180, l.384-387). As emphasised above by the Recommender, our objective was to provide a simple proof of concept to raise awareness on this almost undiscussed aspect of the literature dealing with the role of IV in models of community dynamics, rather than to explore the role of the type of IV on all coexistence mechanisms possible, which is out of the scope of our study.

- A certain type of IV: I am interested in the idea that the only driver of individual performance could be a fine-grained high-dimensional external environment, but it is a very strong assumption, practically as strong as e.g. Hubbell's premises in neutral theory. Thus, just like that other work, I think a clear framing of "let us explore this extreme idea and see how far it goes" would be an easier sell.

R3: We agree that the “Perfect knowledge model” corresponds to an extreme case, where performance is entirely determined by known environmental variables (be they biotic or abiotic), and our study represents a virtual experiment in that respect. This is now better emphasised l.40-46, l.81-84 and l.384-387. The different intermediate “Imperfect knowledge models” could also represent the continuum between this extreme case and the completely random IV, another extreme case. Our choice was motivated by the numerous evidence that environmental variability at fine scales in space and time strongly influences individual variation in plant communities (see e.g. l.27-29), but the consequences of such control have been largely overlooked in community ecology studies dealing with the effect of IV. We however fully acknowledge that IV in nature is also under genetic control, and that heritability and dispersal contribute to the observed patterns of individual variation, as now better discussed l.20-27, l.384-387.

A more specific comment on that last point: different readers will have different opinions on whether we don't know many important abiotic factors (depending on the system), but we could imagine a high-dimensional "environment" also due to biotic factors (e.g. local presence of pathogens). However, that environment would then not only be high-dimensional, but changing in time, and performance would be more context dependent. If my reasoning in point 2) below is correct, then you are not testing a general effect of structure in IV so much as the effect of local performance being constant in time (once determined on the basis of species identity and unchanging external factors), which would rule out many other aspects of a microenvironment.

R4: Both abiotic and biotic factors can be considered as environmental variables in our model. Variables x in our model should be considered as any niche axis and the term “environment” should be considered in its broad sense (including both biotic and abiotic factors). It is true that we have considered fixed environmental variables in our simulations for simplicity, which, in particular, prevents from realistically representing some biotic factors. In our simulations, we illustrate the environmental filtering in space: the fact that everything is fixed in time means that it is associated with the site. Each species outperforms the others in a particular environment which is varying across sites. But our model and results could be easily transposed to structured (i.e. not random, see R7) environmental changes in time, e.g. seasonal variation or periodic inter-annual variation. In the same way, observed IV would result from the variation of unknown environmental dimensions in space and time (cf. theoretical model in Girard-Tercieux et al. 2023 Ecology and Evolution) and including

unstructured IV in community models would not correctly represent species specific response to this changing environment.
The important point to consider in the “Perfect Knowledge” model is that individuals of the same species performs similarly in the same environment (whether it is changing in space, time, or both) while it is definitely not the case with unstructured IV, which has strong consequences on community dynamics and species coexistence as shown by our results.

2) Clarification of the results:

To me, the results' interpretation suffers from an ambiguity between three contributions:

a) possible differences between the Perfect Knowledge Model (eq 1) and even the best fit using eq 2. Having used different equations there seems like a counterproductive and risky decision, since you might be conflating effects of structure with effects of model mis-specification.

However, I don't think this choice invalidates your results (n_obs=15 gives different results from perfect knowledge, but not dramatically different), so I am fine with letting that go (rather than ask you redo your simulations entirely with the n_obs=15 eq2 as the ground truth!)

R5: We fully agree that the comparison between results from Eq. 2 with n_obs<15 and from the Perfect Knowledge Model mixes the effects of both the environment mis-characterization and of the model mis-specification. Our choice to keep a different model structure for Eqs. 1 and 2 was motivated by the fact that this is often what happens when analysing field data: ecologists might not have sufficient knowledge on the fundamental processes controlling the observed data and thus to what corresponds to Eq. 1 in our manuscript, instead they use statistical fits, often polynomial, like Eq. 2. We introduced the model with Eq. 2 and n_obs=15 precisely to allow the reader to distinguish the effects of model mis-specification from environment mis-characterization. Our results (Fig. 2) clearly shows that the error coming from model mis-specification is much smaller than the difference coming from environment mis-characterization. This is now explicitly specified l.125-126.

b) the specific deterministic structure in IV

c) "frozenness" in time, i.e. the fact that in models without uIV, all future individuals of one species that may come to occupy a given site will have the same performance.

My current intuition is that this third ingredient, rather than anything else, is crucial to explain your results.

I now try to explain this intuition:

First, notice that without any dynamical simulation, you could make predictions about diversity in a purely deterministic case: you draw a performance value for each species in each site, and the fittest species lives there forever. Lack of coexistence would only happen due to some unlucky species being the fittest nowhere, which becomes a simple question of probabilities.

You could already have asked at that level whether the way you create a deterministic structure (distance to some niche centre) guarantees more or less coexistence than drawing each species' fitness at random for every site. The case of totally random performance draws is easy: given S=20 species, each of which has probability 1/S to be the best in a given site, and M=625 sites, then the probability of a species going extinct would be

(1-1/S)^M ~ 10^-14

You can of course make things more unequal between species, so that some have a lower probability of being best anywhere, but then the key is the level of inequality, not any particular assumption about structure, and it could only reduce coexistence. Note also that this totally random situation should be at least roughly equivalent to having infinitely-many environmental dimensions in a perfectly deterministic model.

Unless I am missing something, any result of your model that goes against this basic result must therefore have to do with three ingredients:

- any bias introduced by mortality and fecundity mechanisms (the only one I could see really matter is the rich-get-richer effect of abundance-based fecundity)

- the random distribution of propagules among empty sites, which also impacts the Perfect Knowledge model

- temporal fluctuations in the model with uIV, since each individual is now another chance for a species to perform well or badly in a given site. I strongly suspect that this is the main source of stochastic extinctions.

These three are very "neutral-like" ingredients in your model. Since your results are not terribly different in the case of fixed fecundity, we can likely ignore the rich-get-richer part (though I am not surprised that it seems to make things even worse), and random effects are probably the key player here.

In other words, rather than actual structure or not in the IV, what matters might simply be that species performance per site are drawn once and for all in the Perfect Knowledge Model, and changing over time in models with uIV, creating random drift which allows for extinctions.

An easy test could be to run a simulation with independently drawn random performance values for each species x site pair, held constant in time (with some arbitrary level of inequality between species averages if you want). I would expect at least as much coexistence as the Perfect Knowledge model. If this falls within your definition of structure -- and that could make sense, as I believe a reviewer is pointing -- then this would benefit from being clarified.

R6: We agree with the Recommender on these points. The fact that the IV term in the Imperfect knowledge model with uIV (epsilon in Eq. 4) is not fixed in a given site for a given species, and can change over time as individuals die and recruit, is indeed precisely what we call "unstructured", as is now clarified l.22-24. We show in our study that adding random IV makes species less different (or conspecific individuals more different), which leads to a reduced niche partitioning. As noted above, on the contrary, having a random but fixed IV term for a given species in a given site would be roughly equivalent to having many or an infinite number of environmental dimensions in a perfectly deterministic model. Additionally, the fact that random draws are independent among conspecific in the Imperfect knowledge model with uIV (and hence could vary for a given site) mimics previous studies that explored the role of IV in coexistence by modelling approach, which was one of our goals. This is now better introduced (Legend of Figure 1, l.63-65).

To conclude, I would be careful to state explicitly what you mean by "structured IV" -- you focus on one aspect of IV (performance), its spatial structure is of no import here, and the most central factor in your results might be constancy in time (of how good a site is for all individuals of a given species) rather than being explicitly structured by a certain number of environmental factors.

Of course, it would be legitimate to interpret this temporal constancy as what it means for IV to be structured by a fixed external microenvironment, rather than by genetics and development, fluctuating abiotic and biotic factors, density-dependent interactions, chance dispersal events, etc. which would all introduce individual or temporal variations.(I would however hesitate to talk about "spatiotemporal structure" in explaining what is important here)

R7: By structured IV, we mean that IV is mainly determined by the structure of the environment, which is variable in space and time in nature, and the specific species response to these environmental variations. We agree that our term "structure", or “spatiotemporal structure”, may have been misleading. As said above, the fact that environmental variables are spatially-autocorrelated does not matter in our approach (the model is spatially-implicit). We use the term “spatial structure” to say that the IV term is associated with a site and its local features, e.g. being located in the bottomland, or fully exposed to sun. The term “structured” IV has been chosen in opposition to “unstructured” IV corresponding to random draws of individual performance. Our work highlights that with environmentally-driven structured IV, conspecific individuals do not differ in their response to the environment (they have the same performance in a given environment), which is not the case when considering random IV. We did not consider any temporal variation here for the sake of simplicity: as previously underlined, our objective was more to provide a proof-of-concept, rather than fully realistic simulations. We kept the term “spatiotemporal structure” where we discuss how to better take into account the real nature of observed IV. We better discuss these points as explained in the previous answers.

If you can address or correct the points made by all three of us, and the text is amended to avoid confusion regarding what you are showing precisely, I see no reason not to recommend what I think is an interesting starting point for asking a worthwhile question.

Sincerely,

Matthieu Barbier

======================

MINOR COMMENTS ON WRITING:

ABSTRACT

"Also, comparing communities simulated with the same level of knowledge of the environment, but adding unstructured IV or not, we found that the effects of incorporating unstructured IV depended on the relative importance of structured versus unstructured IV. In particular, increasing the proportion of unstructured IV into the model moved from a positive to a negative effect on community diversity and similarity in composition with the full knowledge model. "

This whole part is a bit long-winded and hard to read, especially the end of the last sentence.

R: The two sentences have been modified to ease the reading.

MATERIALS AND METHODS

l72 "additive inverse": a bit cumbersome

R: This is now rephrased as “opposite”.

in various places: "thrive" I'd rather say "reside" (thrive implies success)

R: Correction made.

l90 I was also a bit confused by the use of "triangular" here, though I see what you mean

R: This has been clarified.

l141: "Hence" does not really agree with the previous sentence (but rather refers to the step of randomly distributing the propagules), so it is a bit confusing.

R: We kept this order since we first explain how an individual wins in general, and then we precise that an individual can win by forfeit. We however removed “hence” since it was indeed confusing.

DISCUSSION

I also think, like one of the reviewers, that the Discussion is too long and a bit redundant (within itself and with other parts of the manuscript); being more to the point won't be a deciding factor in a recommendation but it would make for an easier read.

R: The discussion has been revised and shortened in several places to improve clarity and avoid redundancy (see R18).

*by **Matthieu Barbier**, 10 Oct 2022 16:49*
Manuscript: **https://doi.org/10.1101/2022.08.06.503032**

Review by Simon Blanchet, 06 Oct 2022 10:10

Dear Authors,

Thanks a lot for giving me the opportunity to read this interesting comment. I'm not a theoretician, so I'm unable to judge the technicam (methodological) part of the MS, and I'll therefore provide more conceptual/general comments.

I think this MS is timely and that it tackles an important question related to the way intraspecific variability (IV) is incorporated into theoretical models. I think this is an important question and that your MS is providing novel and relevant insights. Nonetheless, I have some criticisms regarding the way you envisage.

R: We thank the reviewer for the time he dedicated to the review of our manuscript and his encouraging and constructive comments.

My main criticism is that I have the feeling that you consider that fitness traits (performance) of an individual can be entirely captured by environmental parameters, which results that IV is highly determined by the environment, as soon as the environment is properly described. This is therefore a deterministic view of IV. You seem to suggest that if IV is not well explained by field ecologists, this is because "most of the environmental variation that actually influences individual's attributes is not properly monitored in ecological studies". This is a strong statement with which -as a field ecologist- I disagree. My personal opinion is that IV (in most species) is actually driven by both deterministic and neutral processes (as most biodiversity facets actually). For instance, in fish, inter-individual differences in body size (for a same cohort, same parents, in the same environment) in juveniles is indeed partly governed by microhabitat use, but also by prior effects (i.e. fish are not hatching at the exact same time, which determine for instance dominance ranking, etc...). Even within the shoal of fish of the same age, inter-individual differences in body size exist whereas they use the exact same micro-habitat. There are plenty of examples like that in the natural world that suggest that not all differences are governed deterministically. Other examples include patterns of "maladaptation"; there has been plenty of studies that recently demonstrated that organisms can actually be maladapted to their local environment (eg because of source-sink dispersal), and I doubt that these patterns emerged only because field biologists failed to properly measure the appropriate parameters. If we were to estimate (across several phyla) the part of IV observed in natural populations that is explained by environmental parameters, we would be probably around 30-70%. I personally think that most of the unexplained variance is due to neutral processes (dispersal, drift, random mortality...) that are hard to capture. I'm actually surprised that you did not cite empirical papers having quantified the link between IV and environmental parameters, and that you don't use this flourishing literature to actually rank your models within this existing data. The only paper you cite to tell that IV is highly structured is one of your paper (Girard-Tercieux et al. 2022); this is fine but I think your arguments should be built on other papers that actually address this issue from empirical data (btw, the example you use in the Introduction about the clones rather suggests that a non-negligible part of IV is not deterministic as inter-clone variation seems high).

R8: In the real world, we acknowledge that genetic IV can indeed impact performance and that neutral mechanisms can also impact community dynamics. We agree that the Perfect knowledge model we present in this paper is not fully realistic. As emphasised in our responses to the Recommender, our aim was to provide a proof-of-concept starting from an extreme case. As you highlighted, the “Perfect knowledge model” is indeed both a reference and an extreme situation where performance is completely determined by the local environment. We aimed to explore the effect of modelling environmentally-driven IV using random IV, as it is often the case, through a virtual experiment. We further clarified these points (see R9, R10, R25). We did not mean that uncertainty arises because ecologists do not know their subject well; we modified the sentence accordingly (l.350-353).

Our study explicitly refers to the plant-ecology literature and our model was developed as a plant community model (l.43, l.81). We furthermore acknowledge that, as any model, it lacks realism for perfectly representing plant communities, and certainly even more for representing communities of other organisms such as fish. However we believe all fields of community ecology would benefit from rethinking the nature of IV and its representation in models. Generally, predictions from ecological models are surrounded by a large uncertainty, which can reflect the fact that all processes are not taken into account in models. While neutral processes can explain some of this uncertainty, individual attributes often vary with the characteristics of the environment (see added references l.28-29). In addition, the ecological niche is known to be multidimensional (Hutchinson 1957, Barbier et al. 2021 https://doi.org/10.1103/PhysRevX.11.011009), but ecological models integrate few of them, be it in population dynamics models (Clark et al. 2003 https://doi.org/10.1890/0012-9658(2003)084[0017:CHTITT]2.0.CO;2) or in niche models (Druon et al. 2015 https://doi.org/10.1016/j.pocean.2014.11.005). Our study aims at raising awareness of the often undiscussed hypothesis of random IV in modelling studies that have dominated research on the role of IV on community dynamics, and we are grateful for your comments that helped clarify our approach and the use of such an extreme case.

Based on this comment/critic, I can make several concrete suggestions:

-Tone down the fact that IV is determined only by the environment and that ecologists are not properly measuring the multiple dimensions of the environment. Perhaps the latter is true but this is unlikely, as we know that most resources are actually not limiting: the niche of most species can be predicted based on a relatively low number of dimensions.
The empirical reality demonstrates that part of IV is determined by the environment, and that another part is determined non-deterministically. Stick on this empirical reality to explore to which extent modelling IV according to this gradient (from purely deterministic to random, or unstructured, as you prefer) affects the outputs of community assembly in theoretical models. And please, avoid the sentences as the one I quote above; in general empirical ecologists know relatively well their biological models and there are plenty of tools that are now used to estimate the environment very precisely according to several dimensions. This can be improved, but I'm not sure this is the core of the problem (and the goal of your study).

R9: Following the feedback of both Reviewers and the Recommander, we modified the text to clarify that we proposed an extreme situation in order to build a proof-of-concept (l.40-46, l.81-84 and l.384-387). As said earlier, we do not claim that *all* observed IV in nature is due to the sole mechanism of unobserved or mischaracterized environmental dimensions. Our main point here is that this produces a non-negligible part of observed IV. For example, as technological progress enables to better monitor both the environment and the response of species, the role of additional environmental dimensions or of environmental variations at finer scales than previously considered has been evidenced in plant communities, e.g. in micro-environment studies (Tymen et al. 2017, Annals of Forest Science; Zellweger et al. 2019, Trends in Ecology and Evolution) and studies showing the importance of many edaphic parameters (Zinger et al. 2019, Molecular Ecology). Importantly, the ecological niche is not only about limiting resources, for instance it also encompasses many aspects of the habitat of the species (*e.g.* temperature, composition of the soil, slope, presence of pathogens etc.). There is not doubt that field ecologists know very well their subject, but unfortunately they typically do not have access for various reasons to all determinants of performance, at least in plant communities, at the scale that matters to individuals: because it has not been brought to light yet, because it is too costly to measure at fine scales, *etc*. We improved discussion of this aspect (l.344-353). Also, it is important to note that the "dimensions" included in our model encompass both resources and non-resource environmental variables as well as their non-linear contribution to individual performance: more than two parameters are often required to properly represent individual response to a given environmental variable known to be important by field ecologists. So even in the case where few variables actually matter to describe a species niche, be they resources or not, and abiotic or biotic factors, much more parameters may be needed to describe the individual responses to these few variables (see Figure 1 and l.344-353).

-the terms used to name the models might be changed. As an example, the "perfect knowledge model" is a strange naming. It seems that it is the perfect model that mimick the reality. It is always taken as the "reference". No, this is not a reference, this an extreme model from a gradient going from deterministic to random IV assembly. It represents a situation that probably not exist in the wild (how can we believe that the performance of an individual is purely deterministically determined?, nothing in the wild is purely deterministically determined, otherwise variation would not exist anymore). For instance, you mention in the Discussion that the perfect knowledge model represents the "reality" (l. 276): no, it represents one extreme case along a gradient. You actually don't know what is the reality (and me neither). So I would simply named it "Purely environmentally-driven IV", or "Deterministic model", or any other term more "neutral" than perfect model. The same for the other terms used to name the other two models.

R10: Following the remarks of both reviewers, we removed all phrasing containing terms relating to "reality" which led to misleading interpretation of our approach - what we meant was "reference" (l.45, l.83, l.247, l.259, l.264, l.267, l.269, l.296 and l.311). We decided to keep the names of our models, first because it helps with our storytelling, and second because we used the same terminology in our previous work (see https://doi.org/10.1002/ece3.9860). We substantially reformulated the way to introduce and consider them, in agreement with both reviewers comments (l.44-51, l.81-83, l.103-110, l.116-121, l.310-312)

Additionally, rather than proposing that field ecologists should better monitor the environment (which is already what we are trying to do), I think it is more important to quantify -based on existing data- what is the part of IV that is explained by the environment, and what is the part that is unexplained. I think a meta-analysis could be performed based on existing data, which would be highly valuable for theoreticians to know the extent of IV that should be modeled as structured or not.

R11: We fully agree both research efforts should be simultaneously sustained. We clarified that improving environmental monitoring is an ongoing effort that participates in better structuring IV l.351-353. We further underlined that partitioning IV into structured vs unstructured IV or into environmentally-explained vs. unexplained, for example by means of meta-analyses, is an important way forward (l.387-391).

Apart from this main comments, I have a few specific ones:

-It was hard to me what you mean by IV. In my mind, IV is the amount of variation observed for one species within a site. It can be quantified using the CV for a trait or the allelic richness for genetic diversity. If I understood well, this is not really what you mean in your models. For instance, under the perfect knowledge model, I would expect that the CV (in terms of performance) for a given species at a given site would be extremely low, whereas it would be higher for other models; am I right? If yes it means that what you are manipulating is rather inter-individual variation rather than IV per se. I think you can keep the term IV as there are not clear consensus about definition (but see https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.7884 and https://onlinelibrary.wiley.com/doi/abs/10.1111/brv.12472 for some attempts), but just try to be extremely clear about what you mean, clearly define the terms.

R12: We agree that a clear definition of "IV" was missing in our study. Intraspecific variability is not necessarily defined within a "site", and the spatial definition of a "site" remains unclear or context-dependent. Violle *et al.* (2012) defined intraspecific variation without any spatial consideration. We here considered IV as the observed difference in an attribute (here performance) between individuals of the same species, whatever the origin of this difference. We now clearly define "IV" (l.4-5) and cite the paper by Raffard *et al.* l.4 (which was more relevant for our study than Clemens *et al.* since this latter specifically focuses on fishes).

Moreover, we used a constant total variation between individuals across models, as illustrated in Figure 1 and spelled out l.56-59 ("These models share the same amount of total variation, but partitioned differently between sIV and uIV, depending on the amount of knowledge of the environment we consider"): for a given number of environmental dimensions, random IV acts as a surrogate of the amount of environmental variation that is not observed. Eq. 2 allows us to infer the amount of this unexplained variation. This results in similar coefficients of variation (CVs) of individual performance across models. We further clarified this l.121-125.

-in the same way, I was expecting you to manipulate traits, not performance (because as a field ecologists we rarely have access to performance!). Again, be extremely clear about that. In the discussion you have a section about niche vs hierarchical traits; perhaps this should be extended into the Introduction. Or more generally try to introduce how IV is generally modeled in other models (for non theoreticians).

R13: Our main goal here was to test the effect of structuring IV (or not) when simulating communities. For parcimony's sake, we chose to perform this test on a trait, or attribute, that has a direct link with performance, and provide with this a first proof of concept. Testing the effect of the nature of the considered trait and its relationship with performance (niche vs hierarchical trait with a linear or non-linear link to performance) on the outcome of adding unstructured IV has been explored by Stump *et al.* (2022, Ecological Monographs), and was out of the scope of our study. Overall both features, the nature of the traits and its link with performance on the one hand, and its structure or source of variation on the other hand, can explain differences in previous tests of the role of IV on community dynamics. This is now better acknowledged and discussed l.77-79 and l.318-328.

-In figure 1, arrow 1 is not clear; are you comparing the two extremes situation (dots)?

R14: We indeed compare the simulations where some environmental information is replaced by random variation (uIV > 0%) with the reference situation (uIV = 0%). We modified the arrow to avoid confusion.

-In equations 2 & 3 the Pîj are not defined. I'm not sure I understood this part. If I understood well, you first retrieve the parameters of the linear model linking Pij to environmental dimensions. Then you used these parameters to infer Pîj (adding or not a random error to account for uIV); am I right ? Please try to improve this section a bit.

R15: We agree this needed to be clarified. We used different notations (with or without hats) to make clear that these were not the same numbers. Indeed, in Eq. 1, we compute a first performance. Then in Eq. 2, we use this same performance (along with environmental information) to infer parameters. Finally, we use these same parameters to compute a second performance in Eq. 3 and 4. We defined Pîjm l.114 and rephrased the paragraph accordingly.

-initialization is done with 10 individuals per species. This is not a lot...As a population geneticist I expect drift to be the major process affecting these populations. Can you tell us if the initial N as an importance for final results?

R16: We agree that the initial number of individuals per species would be a critical parameter in a model with evolution. However, we did not model any genetic processes, as clarified now l.84-85. We did test our simulation set-up with a unique individual per species at initialization and the results were qualitatively similar (not shown). Overall we expect the initial number of individuals per species to have little effect on our results, except when there cannot be any “win-by-forfeit” anymore (when the initial number of individuals per species is very high).

-l. 132; if mortality is proportional to the performance of an individual, this is not stochastic isn't it? I'm not sure you really modeled a stochastic process of mortality. This is correct but this should be told.

R17: With this “stochastic mortality”, the probability of death is inv_logit(0.5 × performance). Then, the one-percent individuals that die at each timestep are determined through a random sampling without replacement using this probability distribution determined by individual performances. Therefore, this mortality does include some stochasticity as opposed to the completely deterministic mortality. This is now clarified l.155-157.

-The Discussion is too long, there are some repetitions, and some sentences are out of scope. I would work further on this part of the MS by being more succinct and straight to the point.

R18: We made the discussion clearer by removing some redundancies, rephrasing some sentences or adding further explanations (e.g. l.249-250, l.267-269, l.275, l.275-277, l.294-296, l.324-326, l.340-343, l.381, l.387-391, l.413), in agreement with both the Recommender and the reviewer’s suggestions. Please note that the first paragraph of each part of the Discussion is redundant on purpose in order to better guide the reader. We would like to keep this structure for the article.

Good luck

Simon Blanchet

Review by Bart Haegeman, 07 Oct 2022 12:23

This paper deals with the way intraspecific variability (IV) is generated in modelling studies on the role of IV on community diversity. The authors distinguish two types of IV, which they call structured (sIV) and unstructured IV (uIV), and argue that these two types of IV can lead to very different model outcomes.

I think the paper possibly makes an important point, but I'm confused about what exactly it establishes. My confusion starts with the authors' focus on IV generated by the variation in different local environmental conditions that conspecifics experience. Is this type of IV considered to be dominant in nature? While in the introduction and in the discussion other types of IV are briefly mentioned (in particular, IV generated by genetic differences between conspecifics), I am missing justification for the focus on environment-driven IV.

R19: Thank you for your interesting comments on our study. IV can be indeed either genetic or environmental (Violle et al. 2012). In our study, our starting point was that, despite these potentially different natures of IV, most modelling studies have so far introduced IV in models using a purely random IV, a representation that does not allow to capture the spatio-temporal pattern of environmentally-driven IV. We here provided a test of the effect of this rarely discussed modelling choice and showed that it can lead to dramatically different results compared to when IV is considered as emerging from the environment. We have modified the introduction to clarify this point.

Secondly, the modelling approach is largely based on starting with a model that only contains sIV (i.e. no uIV), replacing part of this sIV by uIV, and then exploring the effects of this replacement. As a consequence, the precise implementation of this replacement seems to be key, but the authors do not give much justification for it. For example, when in the model part of the sIV is replaced by uIV, this latter IV is also structured in space and time, in the sense that the variability in performance hat{p}_ij, Eq (4), within species i is still generated by conspecifics being located in different sites j. Hence, applying the authors' definitions of sIV and uIV, I would say that this replacement of part of the initial sIV is still sIV, and not uIV as claimed in the paper. This seems an important issue to me. More generally, even if the replaced sIV would clearly classify as uIV, it should be justified that this particular replacement is appropriate for the purpose of the paper, i.e. able to capture the effects of a specific type of IV independently of the particular implementation.

R20: Following other comments of the Recommender on what we exactly meant by “structure” (see R4, 6 and 7), we have clarified this point in the introduction (l.25-27). Briefly, we considered what we called uIV as unstructured inasmuch as it is not associated with the site (i.e. the location of the individual) nor with the time of observation. It is instead randomly drawn for each individual at the recruitment step, independently of each other and of the site. This representation typically mimics what most of previous modelling studies did, which was part of our goals.

We however realised thanks to your comments that the notations in our equations were misleading. There is no site random effect in our models, but as there is only one individual per site, we confounded both. This is now corrected, with these notations:

- j = species (unchanged)
- k = environmental variable (unchanged)
- i = individual (changed)

- m = site (changed)

We also simplified the equations in order to make them clearer.

Finally, I would have liked more discussion about the implications of the model results. Would it be possible to illustrate concretely how the findings of this paper change the interpretation of previous modelling studies? Do previous model results change qualitatively when substituting uIV by sIV? Such examples could make the current paper more convincing. It would also be useful to provide guidelines for future modelling studies on IV. How should IV be modelled to get relevant and robust results? And more generally, I would suggest the authors to focus their discussion more on modelling work. For example, how should one adapt the unifying framework of Stump et al in light of the current paper, or how could incorporating genetically-driven IV in the current study modify, or not, the results (beyond simply noticing that genetic differences between individuals would lead to spatial/temporal structure).

R21: These are indeed all very interesting points. Regarding the first question ("Would it be possible to illustrate concretely how the findings of this paper change the interpretation of previous modelling studies?"), we explicitly characterised the different nature and structure of IV present in nature that are however typically all represented through random distributions (e.g. l.331-342). The second one ("Do previous model results change qualitatively when substituting uIV by sIV?") would be an interesting next step but is beyond the scope of the present study. It would indeed require to reproduce and amend previous models, while our study provides a simple and standalone proof-of-concept showing that a random noise is not a good substitute for environmentally-driven observed IV. The third one ("how should one adapt the unifying framework of Stump et al in light of the current paper") is a good point, as both Stump et al's framework and our study are complementary and can together help explain the discrepancies in modelling studies, as explained l.320-326. Indeed, Stump et al. explored the effect of the nature of the trait considered for IV and its link with performance on modelling outcome, but exclusively using random IV, a point raised in their discussion. On the contrary, our study focuses on this latter point, while using a unique trait (or attribute). We believe both raised important points, which should help community ecology move forward, as discussed l.326-328. We believe the last question ("how could incorporating genetically-driven IV in the current study modify, or not, the results") is out of scope of our study and would require complex implementations and further hypotheses (on e.g. heritability) beyond our first simple proof of concept (l.40-42).

In short, to me this paper indicates that modelling work on IV should carefully consider the type of IV and the details of how it's implemented. However, it is unclear to me what the paper establishes beyond this general warning. I think more justification of the chosen approach and more discussion of the implications of the results are needed.

R: We believe the comments of the Recommender and the reviewers helped us improve our manuscript and clarified its significance for community ecology and modelling.

More detailed comments:

* Figure 1: I like the idea of presenting graphically the paper's approach. However, I found it hard to understand certain aspects of this figure before having read the methods section.

Panel A: Not clear to me what the clouds of points are representing. If only one variable is varied while keeping all others fixed, then performance vs this one variable should give a curve, right?

Panel B: Again not clear to me what is represented here. Is this illustrating the fitting procedure of Eq. 2, with only one unobserved environmental variable?

Panel C: Is the %sIV on the x-axis the same as (number of observed environmental variables)/15? If so, it is probably better to just put the number of observed variables. Same for y-axis.

Maybe use the arrow labels to refer to the two questions of the introduction and mentioned several times in the main text, rather than introducing three arrows/labels here.

R22: We thank the reviewer for these suggestion, which help clarify the figure as follows:

- Panel A: The other variables are actually not fixed, but the variation of performance with all environmental variables is projected in 2D for each variable separately. This is now better specified in the legend.
- Panel B: This panel represents the step of the experiment where Eq. 2 is fitted to the observed response with only one explicative variable observed (here X6): observed IV is inferred from data and variation not explained by the observed variable X6 is interpreted as noise around a mean fit. This has been clarified in the legend.
- Panel C: The % of sIV and uIV are computed as you understood it. We added a second axis in order to clarify the link between %uIV and the number of observed dimensions.
- Arrows: We already explicitly refer to the arrows in the introduction.

* Methods, paragraph line 56: Details are missing here, e.g, for the conditional autoregressive model. Please provide everything needed to replicate your work. Why not start with a distribution on the interval [0,1], rather than rescale the normal variables?

R23: The code used to generate the CAR models is provided in the GitHub repository and has been archived on Zenodo.The analysis is thus entirely replicable, as specified in the section "Supplementary information and code access". We first used a CAR process to obtain spatially autocorrelated environmental variables. We then rescaled the values on the interval [0, 1] to ensure that each environmental variable had the same effect on species performance on average. We clarified this point in the method part l.74-75.

As explained to the Recommander above, spatial autocorrelation was used to better illustrate the environmental filtering of species. At the end of the simulations with the Perfect Knowledge models, species are grouped spatially in sites with similar environments because of environmental filtering. We added a figure in the Supplementary Materials to illustrate this (Fig. S5.17). In nature, it is also often the case that environmental variables are spatially autocorrelated and that individuals of the same species are grouped spatially (Pélissier et al.). But as underlined by the Recommander, species coexistence in our study is independent of this spatial structure of the environment.

* Methods, Eqs (1) and (2): The shape of Eq (1) is called triangular, and contrasted with the quadratic of Eq (2). I suppose the triangular shape is referring to dependence of d_ij on x when x is close to x^*? And hence the triangular shape would disappear when removing the square root in the definition of d_ij, Eq (1)? This raises a number of questions. Why choose a triangular shape? The empirical performance curves (performance vs environmental variable) I know behave as a quadratic function close to their optimum. Taking instead a triangular-shaped optimum here looks artificial. Moreover, this triangular shape cannot be fitted well by a quadratic function, so this choice will (artificially) increase the difference between perfect and imperfect knowledge models. How would your results change when defining d_ij without the square root in Eq (1)? Maybe use "~" or something similar instead of "=" in Eq (2), because this is a fitting model, and not an equality as in Eqs (3) and (4).

R24: This choice takes its roots in the Hurt & Pacala model we started from to build our model where performance follows the distance between the local conditions and the optimum of the species. While alternative choices could obviously have been made, the differences in the shape of the “Perfect” vs “imperfect” knowledge models only had a negligible effect on the differences in results between the two models, as explained above to the Recommender.. The different shapes were used to mimic what often happens in ecological studies, where ecologists do not necessarily have access to the “true” shape of the actual processes and fit them using statistical models with observed data.

* Methods, line 97: "second question", maybe also briefly recall the procedure for the first question.

R: We removed it for both questions since it seemed unhelpful to the reader.

* Methods, line 101: "estimated from", do you mean "generated as"?

R: Correction made.

* Methods, line 131: "probability to die is proportional to performance", should this be "inversely proportional"?

R: Correction made.

* Methods, line 144: "in the imperfect knowledge model colonization depends on ...", so the performance hat{p}_ij of Eqs (3) and (4) is used here instead of p_ij? I suppose the same is true for the performance-dependent mortality? That is, where p_ij is used in the perfect knowledge model, hat{p}_ij of Eqs (3) and (4) are used in the imperfect knowledge models?

R: You are correct. hat{p}_ij is the computed performance (after statistical inference in Eq.2) while p_ij is the reference performance. We clarified this l.116.

* Methods, line 177: "with the perfect knowledge model", I suppose the same is true when quantifying site sorting for the imperfect knowledge models?

R: We quantified site sorting using the community’s mean performance from the Perfect knowledge model, but for final communities obtained through the dynamics driven by each of the three performance models. Indeed, the performance with the Perfect knowledge model corresponds to how close a species is from its environmental optimum.

* Figure 2: Maybe use pink for the y-axis label on the right, so that it's clear that the right-hand y-axis refers to the pink points/curve/ribbon. In the legend: replace "an configuration" by "a configuration ExS"

R: Done.

* Figure 3: In panel A, why is there a remaining difference between the case of 15 observed environmental variables and the perfect knowledge model?

R: See our response R6 following the Recommender's comment.

* Figure 4: In the legend: replace "one the one hand" by "on the one hand"

R: Done.

* Discussion, line 215: "unrealistic communities", this seems to refer to communities that are different from those of the reference model (the model with only sIV). Are these communities unrealistic in the sense of not corresponding to empirical data? If not, I think it's better to avoid the term "(un)realistic" (used several times in the text).

R25: We indeed used the terms "real", "realistic" and "unrealistic" referring to the Perfect knowledge model, which is used as a reference situation. We did not refer to any empirical data since it is not the goal of this virtual experiment to accurately represent what happens in the field. We understand that these terms were misleading, and changed them, for instance for "unrealistic" to "dissimilar from the reference", "realism" to "similarity with the reference", "more realistic" to "more similar to the reference", in agreement with other comments from the recommender and the first reviewer. The term "real" still appears in the main text but now refers to natural communities/empirical data only.

* Discussion, line 242: "a neutral mechanism", given the variable interpretations of this term among ecologists, maybe describe more explicitly what you mean by this.

R: We clarified this as "affecting all species in the same way".

* Discussion, line 263: "this was however only due to", where is this shown?

R: We did not include any figure to show this, we relied on the previous result (effect of unstructured IV on species richness). We added "probably" to avoid confusion regarding any supplementary analysis.

* Discussion, line 269: "as expected ..." until the end of paragraph: Unclear to me what this is referring to.

R: This has been clarified (l.304-306).