Dear Pr. Hastings,

We have edited our manuscript according to the constructive feedback of the reviewers, for which we are very grateful. You will find our answers to comments and the corresponding changes to the manuscript in bold after each reviewer comment.

Reviewers raised some issues regarding the focus that we chose for the manuscript, as well as the terminology employed. We edited the manuscript in response to both.

Regarding the focus, we have changed the title to "How self-regulation, the storage effect and their interaction contribute to coexistence in stochastic and seasonal environments". This new title is, in our opinion, more representative of the paper: it explicitly refers to how seasonality influences coexistence, which is a central element of our work, and how coexistence mechanisms interact, which is also a key feature here. We have then sought to hold this focus throughout the manuscript.

Both reviewers noted that we were incorrect to equate "stabilizing niche differences" (SNDs) with cases in which intraspecific interaction coefficients are much larger than interspecific coefficients. Indeed, one could say as well that a storage effect induces stabilizing niche differences. We thank the reviewers for this remark, which will no doubt improve greatly the clarity of the manuscript, and we have corrected our wording to "strong self-regulation" (wherever we previously wrote SNDs).

In the same spirit of improving the readability of the manuscript, we changed the name of one of our variable, ρ, the ratio of intra-to-intergroup coefficient, to κ, to avoid any confusion with niche overlap in niche theory, which is usually denoted ρ as well.

We hope that you will find the updated manuscript improved, and that it may now be suitable for publication in Theoretical Ecology,

Sincerely,

Coralie Picoche & Frédéric Barraquand

**** Responses to reviewer comments are provided in bold below ***

COMMENTS FOR THE AUTHOR:

Reviewer #1: This ms appears to be about multiple things. On the one hand, it discusses the relative contributions of the storage effect and differences in within- and between-species competition coefficients. (I will avoid calling this stabilizing niche differences, for reasons explained below.) It further notes that these effects combine superadditively in this model. On the other hand, it's also a paper about how the abundances of the coexisting species are

distributed in trait space: in particular, whether or not we have a multi-modal distribution ("clumpy coexistence"). Or perhaps, like Scranton and Vasseur, it is less about the abundance distribution and more about how many species can be supported, the idea being that a bunch of very similar species can transiently (but lingeringly) coexist in each species clump.

What I find most useful about this ms is that it extends Scranton and Vasseur 2016 to consider noisy seasonal variation instead of white noise.

>> Extending the work of Scranton and Vasseur (2016) to seasonal environments was indeed our original motivation. The new title now explicitly mentions seasonality which is a main theme of the paper, and avoid to make any reference to stabilizing niche differences (see below for a disambiguation of the wording employed).

It is useful to know that a significant seasonal component can substantially reduce the number of coexisting species, at least if we're considering just a storage effect or just differences in between-/within-species competition. I'm not sure that uni- vs multi-modal distributions have a significance beyond that. The paper seems to focus on the shape of the distribution beyond the effect of the shape on the number of coexisting species, and it's not clear to me why this is important.

>> We agree with the referee that biomass-trait distributions can at times be misleading and should not be overinterpreted. We now make this clear at the end of the abstract (l. 21-23). But we believe that theoretical work should demonstrate if, and when, trait distributions are or are not indicative of coexistence mechanisms. Trait patterns are *de facto* being used to infer niche differentiation and environmental filtering (see refs in d'Andrea et al. 2016, Loranger et al. 2018); we therefore believe that it is useful to help accumulate knowledge on biomass-trait distributions.

Scranton and Vasseur (2016) reported multimodality in biomass-trait distributions. Since the models that we consider have been derived from theirs, we think that it is crucial to check if multimodality also arises with/without strong self-regulation and with/without seasonality. We certainly agree that such shapes are not sufficient on their own to form any conclusion as to the coexistence mechanism(s) - we think of these as "clues". We now make their limitations clearer in the discussion (l. 424 - 426).

It would be a pity, in our opinion, if we did not report these results that can enlighten debates on trait distributions while our models produce them as a by-product. But it is not the main focus of the paper - we only devote one paragraph to trait patterns in the discussion, for instance. The main focus is on coexistence mechanisms that allow species-rich species communities to maintain in noisy and seasonal environments.

I am not particularly excited by the result that multiple mechanisms may be contributing to coexistence in real systems --- we've been talking about that at least since Chesson started writing in the 1990s --- or that these mechanisms may combine nonadditively. It is, perhaps, a message that cannot be overemphasized, but I don't think it is

especially new.

>> We understand the reviewer's concern. However, it is one thing to know intuitively that a combination of mechanisms *could* sustain biodiversity and another to show that such combination *is* indeed important *in a species-rich context*.

We are aware of only a handful of references showing decisively non-additive effects of mechanisms (Kuang & Chesson 2009, Comita et al. 2014, Mordecai 2015) but we have not found any in species-rich contexts (dozens of species or more), that are the focus of the study. Reviewer #2 was interested by the non-additive effect of mechanisms and we tend to concur.

As we understand it, one's background reading has always a large impact on what one finds novel or not. If reviewer #1 has additional references to the ones we cite (especially older ones), that would be worth discussing in the context of non-additive effects, we will of course be glad to include and discuss those.

Comments:

1. ll. 110--114: I found this a little confusing and ended up calculating the storage effect for both models so that I could understand --- not an option available to most readers. My confusion stemmed from the ambiguity of "competitive strength." I took this to mean the competition coefficients, α_{ij}, which do not in fact covary with r. What's going on is that if we define competition to be $\sum_{j=1}^{S} \alpha_{ij} N_j$ in model 1, then per capita growth depends on the product of the environment ($r_j(\tau)$) and competition, and this interaction produces a storage effect. In the second model, the environment and competition enter additively, so that there is no covariation. Maybe you need to say that the net effect of competition is given by that sum, and that in model 1, that sum is multiplied by the environmental response, $r_j(\tau)$, producing covariation between environment and competition.

>> This is a correct interpretation, which we now explain more clearly l. 140-146. To make sure there can be no confusion, we now use two terms: "competitive strength" refers to the coefficient alpha_{ij} and "net effect of competition" refers to the sum of products of the intrinsic growth rate, alpha_{ij} and the density of other species N_j.

2. I am grumpy about defining stabilizing niche differences as equivalent to the within-species competition coefficients being greater than the between-species competition coefficients. This is certainly one way to get stabilizing niche differences. But normally SND refers to any differences that increase the per capita growth rate of all species as invaders. Thus, the different temperature optima, which give rise to a storage effect in model 1, are also normally considered stabilizing niche differences.

>> Upon closer inspection of the literature, we found this remark to be perfectly accurate and we apologize for the confusion that we created.

We have been unfortunately influenced by the vocabulary used in different subfields: studies on the storage effect seem to use more often "stabilizing mechanisms" and studies using Lotka-Volterra models more often SNDs (e.g., Levine et al. 2017). We erroneously thought that SNDs were a subset of stabilizing mechanisms (while they are in fact one and the same thing). We have now corrected this mistake throughout the manuscript and have changed the terminology to "strong self-regulation", or SSR, whenever intraspecific >> interspecific competition (e.g., l.154, l.213, etc., and Fig.2-4).

3. Results section: For this model, the storage effect alone was able to enable the coexistence of fewer species than competitive differences alone. That's worth knowing, but I would be hesitant to imply that this is a general result. I don't think the paper ever declares this to be universally true, but I would feel most comfortable if it were made clear that this result is contingent on this model.

>> We completely agree that the result is contingent upon the model. Now mentioned explicitly in the discussion l. 361-368. We did, however, our best to inform the model by estimates coming from empirical work (see l. 355-359).

4. ll. 181--183: "On the contrary, seasonality tended to favor..." All of this is still without the storage effect, correct? I'm pretty sure that's the case, but I had to read the paragraph multiple times.

>> Indeed, the paragraph (l. 225 to 240) describes the effect of the random and seasonal noise without any of the coexistence mechanisms. We now repeat that we are considering a case with no coexistence mechanism (l.225-226 and l.232-233).

5. l. 182: "Species with a higher thermal optimum are more likely to persist and to reach a higher biomass..." Any intuition about why this is true?

>> The shift to higher values of the average thermal optimum when the community is forced by a seasonal noise can be found for all four models that we considered. The autocorrelation in the seasonal noise allows species with a higher thermal optimum, and, thus, a higher maximum growth rate, to maintain a high productivity for longer contiguous periods of time than a random noise would, as is now explained l. 269-272.

6. 284: "Of course, other mechanisms that we did not include in our models may produce similar patterns." This is important to note. I think it's also important to note that a different model with the same mechanisms might produce different results. How confident are you that a storage effect will always produce clustering on the trait axis in the presence of seasonality? Can you explicitly state your intuitions for why this should be so?

>> We stated in the earlier sentence that "Clumps in the trait axis when averaged across model runs/locations are therefore a signature of the storage effect for the cases that we considered in this article".

We therefore already note implicitly that a different model could produce different results. We do not know for sure that the finding is general, which is why we stay prudent in the latter sentence (modified from the first submission version):

"We therefore view clustering on the trait axis (when averaged over several samples) as an interesting clue suggesting to look for a storage effect, rather than any definite proof that the storage effect is at work." (l.424-426)

Here we only say that clustering could be seen as a clue of the storage effect, not that a storage effect would always lead to clusters.

As to why trait clusters exist in the first place: the storage effect essentially models niches in time, with some species being better for some of the environmental variable values. When this mechanism dominates, we therefore expect that the traits of the remaining species will exhibit contrasted values.

Minor comments:

1. Table 1: b_i and $f_i(\tau)$ appear to be missing values.
>> These values change with the considered species and timestep. This is now clarified in the caption of the Table 1.

2. Fig. 3: It would be really helpful to add row labels to this figure: "White noise" for the top row and "Seasonal" for the bottom row.

>> We have added row labels. Following a remark of reviewer #2, "white noise" has been modified to "random noise" (it is not exactly white in the sense of a stochastic differential equation).

3. Fig. 4: Again, row labels would make this clearer: "Storage effect only" for the top and "Competitive differences only" for the bottom.

>> We have changed the labels to "storage effect" and "strong self-regulation", thank you for this suggestion.

Reviewer #2: The manuscript looks at coexistence in a forced Lotka-Volterra model with two potential coexistence mechanisms: the storage effect and classical frequency dependence operating on shorter timescales than the temporal fluctuations. The Authors find that both mechanisms can individually support some coexistence, but their combination can support more than the sum of the two. Additionally, the forcing signal was made to be both random noise and a noisy sinusoidal forcing, with random noise able to support more species when only one coexistence mechanism was acting.

Overall, I enjoyed reading the manuscript. Below I have some technical suggestions, as well as a few comments on interpreting the results.

First, if I understand correctly, the Authors' white noise isn't really white, and it shouldn't be. I was trying to reproduce their results, and could only do so if the value of tau was kept constant for the duration of one day at a time. Is that what was done?

>> Indeed, that is exactly what was done, to keep the comparison to Scranton and Vasseur (2016) possible.

In that case, since the random jumps in temperature are interspersed by 1-day-long periods of constant temperature, the noise must be somewhat autocorrelated and therefore red.

This is all well, since a stochastic differential equation can only be defined consistently if the white noise term appears linearly on its right hand side - which it doesn't in the manuscript. A "nonlinear SDE" is a non-entity; see, e.g., Simo Särkkä's book "Applied Stochastic Differential Equations" (p. 34) for a discussion. Therefore, having the slightly autocorrelated noise is perfectly fine and even necessary. I would emphasize that temperature was kept constant throughout a day, and would drop the terminology of white noise in the manuscript. Instead, I would just call it "random noise".

>> We completely agree with this remark and thank you for pointing it out.

A probabilist colleague once brought to our attention that precisely for that reason, SDEs are in fact a very small part of all random dynamical systems (sensu e.g. Arnold, L. (2013). *Random dynamical systems*. Springer).

We have therefore chosen to use the 'random noise' wording throughout the manuscript. We also define this noise more thoroughly l. 123-129 - explaining that it is not entirely white, but can be considered so at the monthly or annual scale.

Second, I am not convinced the model was run long enough for species richness to stabilize. At least, when I was running it for 15000 instead of 5000 years, the richness against time curve still didn't quite settle down in the pure storage + random noise scenario. This means that Figure 2 probably overestimates species richnesses.

>> This is true that convergence is not fully reached at *t* = 5000 and we now include these details in the Discussion (l. 295-314). However, we do not believe that waiting for convergence would necessarily be meaningful here.

This work originally started as a replication of Scranton and Vasseur (2016)'s results. It was for us a pre-requisite to adding seasonality (we firmly believe that repeatability is paramount for progress in quantitative ecology). Replicating Scranton and Vasseur's model, we realised that even there convergence was not always reached (see in the Figure below the temporal course of the species richness, for a randomly forced community with a storage effect - that is the model of Scranton and Vasseur).
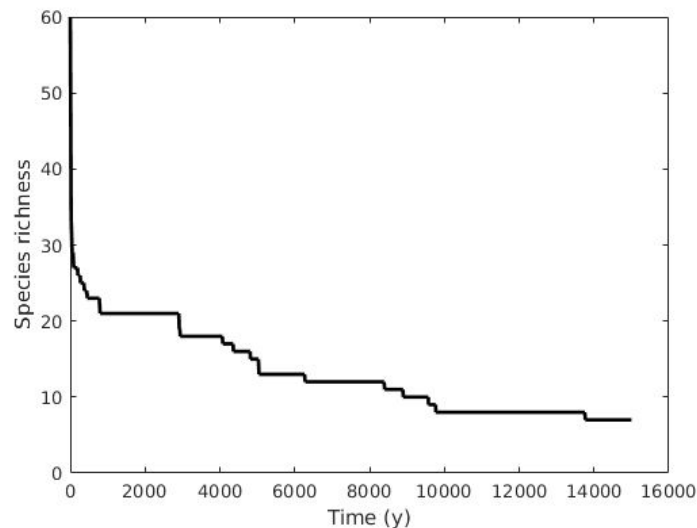
While we considered waiting for convergence, the last transients can be really long, in the sense that exclusion can take an added >10 000 years.

While mathematically speaking the community is not stabilized, we then thought about the biological meaning of waiting an additional 10 000 years and it seemed like a very artificial thing to do. A phytoplankton-like community (the biological target) would necessarily evolve or change in such an amount of time, which would require to add variability to all traits, not only the growth rates. This would lead to another model and we would lose the connection of our models to the one of Scranton and Vasseur (2016). There were also computational difficulties associated to such an increase in simulation length, both in terms of computation duration and memory use.

Thus we decided to keep the 5000 years used by Scranton and Vasseur, which at least allows for a clear-cut comparison. We now explain plainly in the Discussion (l.303-314) that this methodological choice was done to allow for (1) comparison between modelling studies and (2) ecological consistency -- after all, if a model produces robust transient dynamical behaviour between 5000 and 15000 years (i.e., not a short-term transient), perhaps this transient is as important as the asymptotic behaviour (see for similar arguments Hastings et al. Science 2018, now cited as well).

There is a simple way of speeding up convergence, and as an added benefit also making sure that the results are not an artifact of parameter fine-tuning: I would make the m values slightly different across species. What I have done is set m_i equal to 15 plus a uniform random variable, drawn either from [-0.1, 0.1] or [-1, 1]. In both cases, at least for the parameterizations and scenarios I looked at, the species richnesses converged at the same final values. This is useful, since the change in richness is then presumably not due to altering the dynamical behavior through the m_i, but to more rapid convergence stemming from resolving the fine-tuning. I encourage the Authors to do the same, and to check for

convergence (e.g., by checking whether species richness has been constant for 100 years, say) instead of integrating for a fixed time interval.

>> This is an interesting suggestion and we are very grateful for your thoroughness. For the reasons mentioned above (comparison to SV16 and ecological interpretation), we would prefer to keep a fixed time interval.

Moreover, changing mortality would also affect the 'emergent neutrality' (sensu Holt) as species would get more dissimilar, i.e., they would then differ in more ways than just their thermal niche and the niches implied by strong self-regulation. We would then have no reason to get several species coexisting in the same trait clump, for instance - changing the mortalities immediately induces new niche differences.

Third, a small comment on terminology. The Authors distinguish between "stabilizing niche differences" (SND) and the storage effect. Within the Chessonian terminology however, the storage effect also counts as an SND (anything does which can boost the long-term invasion growth rate of all species simultaneously). Instead, I would call the Authors' SNDs "classical frequency-dependent stabilization" or something along those lines.

>> As you can see in our response to reviewer 1, we realize that we were wrong to call high intraspecific competition SNDs (we've been mislead by the fact that many articles with Lotka-Volterra models use SNDs to designate high intraspecific competition while storage effects paper usually refer to stabilizing "mechanisms"). We have changed to "strong self-regulation", and define precisely the wording l. 159-162.

Fourth, I wanted to make sure that the interpretation of competition coefficients was correct. Usually, what can be directly measured in an experiment is the reduction in per capita growth rates. But the alpha and rho terms here measure the reduction in intrinsic rates, as per Eq. (1). Is this correct, and does this indeed map onto the values obtained by Barraquand et al. (2018)? I'm just making sure.

>> This is correct. But these two concepts map: if alpha_ii = 10*alpha_ij then r_i(tau) alpha_ii = 10 * r_i(tau) alpha_ij. We therefore think there is no problem here.

Moreover, rho = 10 is taken as an order of magnitude rather than an exact value. We do not pretend that in real data the value is always 10. In fact, the model in Barraquand et al. (2018) is a discrete-time model on the logarithmic scale (Gompertz), because this model fitted best the plankton data. This it is not exactly the same model, but we think that the order of magnitude (10 rather than 1 or 100) will hold. Based on plant data, Adler et al. (2018) identified rho to be closer to 4 or 5. We have done a few simulations with rho = 4 and it did not change the results qualitatively, but we believe that rho = 10 is closer to plankton communities. We now mention all this in the Discussion (l. 362-368). [1] There are other caveats regarding the comparison

[1] The notation rho has been changed to kappa to avoid confusion with the usual niche overlap measure.

between our models and empirical data. For example, we assume in the theoretical models that all intra- or inter-specific interaction coefficients are alike with each category, i.e., we discard the intra- or inter-specific coefficient variances that exist in the empirical data (now fully acknowledged and discussed l. 317-328). The choices that we have made therefore reflect a compromise between being able to compare our results to the theoretical work of SV16 (assuming equal coefficients) and connecting to our empirical work.

Fifth, the presence of frequency-dependence (the rho = 10 case) means there must be differences between species providing larger intra- than interspecific competition. We do not know what these differences are, because it is all encoded phenomenologically, through the competition coefficients. But they must be there, and must be unrelated to temperature optima (indeed, even if two species by accident had the exact same temperature optimum, intraspecific competition would still be ten times greater than interspecific between them). This should be taken into account when interpreting the coexistence of "similar" species with rho = 10: these species are not really similar, it's just that the differences between them aren't being modeled explicitly. See our article (Barabás et al. 2013 Oikos) for more details.

>> Overall we agree and have corrected the wording accordingly: "the diversity within clumps of similar values of thermal optima was considerably decreased" (l.371-372). The use of the word "similar" was influenced by our knowledge of phytoplankton - many diatoms are similar in terms of resource requirements and abiotic drivers. Phytoplankton ecologists usually dub them "similar species" but we completely agree that they have different shapes and life-histories (which, we suspect, is instrumental to their coexistence and producing such weak interspecific competition).

A few minor comments:

- I would stick to using kilograms, years, electronvolts, and Kelvins consistently. So in Eq. (5) I would not divide the argument of the sine function by 365, and in line 131 I would express the temperature interval in Kelvins.

>> We have indeed changed Eq. (5) and kept the same units, except for the use of the Kelvins. We remain consistent with the already published paper of Scranton and Vasseur for their use of Kelvins and Celsius. Celsius are easier to interpret (for a sizeable fraction of the planet at least).

- Eq. (6): How was the mean r_i obtained? Was it calculated by first generating the full time-series for tau, and then taking the temporal average? Or by some other method? Please explain.
>> This is exactly what we did. This is now explained l. 148-150.

- Lines 247-248: With rho=10, there is no neutral kind of diversity in any sense of the word, since by default intra > 10*inter, precluding neutrality (see also my comment above).

>> Yes, we agree. We have now removed "neutral" (l.371-373) as it could be misinterpreted, even with the quotation marks (we meant within-clump diversity, which consists in species with similar coefficient values related to temperature, but we agree that self-regulation implies other niches differences).

- In general, pitching the article in terms of reconciling niche and neutral perspectives is a bit misleading. The Authors' model does not have any neutrality, and their result of the superadditive influence of various mechanisms is interesting in its own right. I would focus more on this superadditivity as the main thrust.

>> We have changed the title to better represent the focus of the article, hinting at the superadditivity of coexistence mechanisms (which we find indeed quite interesting). Our Introduction mentions studies that have sought to reconcile niche and neutral perspectives, but we do not write anywhere that we try such reconciliation ourselves.

We only contextualize how this idea of "clumpy coexistence" (connected to the biomass-trait distribution) came to be in the Introduction, which requires to talk about "emergent neutrality". This is, in our view, needed to introduce properly Scranton and Vasseur's 2016 model and discuss the meaning of biomass-trait distributions later on. We have therefore re-rewritten a bit the Introduction to introduce the concept of clumpy coexistence and the work of SV16, starting l. 38 and ending l. 65.

References used in this response
1. Adler, P. B. et al. Competition and coexistence in plant communities: intraspecific competition is stronger than interspecific competition. Ecology Letters (2018). doi:10.1111/ele.13098
2. Arnold, L. *Random dynamical systems*. Springer. (2013)
3. Barraquand, F. , Picoche, C. , Maurer, D. , Carassou, L. and Auby, I. Coastal phytoplankton community dynamics and coexistence driven by intragroup density‐dependence, light and hydrodynamics . Oikos (2018). doi:10.1111/oik.05361
4. Comita, L. S. et al. Testing predictions of the Janzen-Connell hypothesis: a meta-analysis of experimental evidence for distance- and density-dependent seed and seedling survival. Journal of Ecology 102, 845–856 (2014).
5. D'Andrea, R. & Ostling, A. Challenges in linking trait patterns to niche differentiation. Oikos 125, 1369–1385 (2016).
6. Hastings, A. et al. Transient phenomena in ecology. Science 361(6406), (2018).
7. Holt, R. Emergent neutrality. Trends in Ecology & Evolution 21, 531–533 (2006).
8. Kuang, J. J. & Chesson, P. Coexistence of annual plants: generalist seed predation weakens the storage effect. Ecology 90, 170–182 (2009).
9. Levine, J. M., Bascompte, J., Adler, P. B. & Allesina, S. Beyond pairwise mechanisms of species coexistence in complex communities. Nature 546, 56–64 (2017).
10. Loranger, J., Munoz, F., Shipley, B. & Violle, C. What makes trait–abundance relationships when both environmental filtering and stochastic neutral dynamics are at play? Oikos (2018). doi:10.1111/oik.05398
11. Mordecai, E. A. Pathogen impacts on plant diversity in variable environments. Oikos 124, 414–420 (2015).
12. Scranton, K. & Vasseur, D. A. Coexistence and emergent neutrality generate synchrony among competitors in fluctuating environments. Theoretical Ecology (2016). doi:10.1007/s12080-016-0294-z