

**Manuscript:** Why scaling up uncertain predictions to higher levels of organisation will underestimate change

**Authors:** James Orr, Jeremy Piggott, Andrew Jackson, and Jean-François Arnoldi

Dear Recommender and Reviewers,

Thank you for your thorough reviews, which have helped us to improve the quality of the manuscript. Please find attached our revised manuscript, a version of the manuscript with changes indicated and a point by point response to the Recommender's and Reviewers' comments. [Changes to the manuscript are highlighted in blue.](#) We hope that our revised manuscript is worthy of recommendation from PCI Ecology.

Kind Regards,

James Orr, Jeremy Piggott, Andrew Jackson, and Jean-François Arnoldi

– 22nd of September, 2020

## 1 Response to Recommender

*I have now received two reviews of your manuscript. Both reviewers and I are in agreement that this is an interesting study considering how scaling up uncertain predictions of individual properties in complex systems affects the estimation of system-level properties. The results have important implications in ecology as well as in other research disciplines. However, several issues have been identified which, in my views, require revision before recommendation. Such revised contribution would need to address all of the reviewer comments. In particular, reviewer #1 raises an issue regarding the assumptions on the specific distribution of the “error” used in the mathematical derivation. In addition, reviewer #2 highlights several points that would deserve to be further clarified and discussed (e.g. further discussion of the implications of the results for other research areas, including consequences of intraspecific variations). In addition to the*

*comments of the reviewers, I have a few additional suggestions to help improve the clarity of the manuscript:*

We thank the Recommender for their work and for their helpful comments. As well as changes inspired by the Recommender's and Reviewers' comments we have also slightly modified Figure 1 for aesthetics and we have included a url to the GitHub repository containing all code used for simulations and figures (Line 604).

*Figure 2: When reading first the manuscript, I didn't understand the meaning of the blue and red circles in this figure, and globally this figure is rather difficult to understand. This part only becomes clear when reading the next section with Figure 3. I would suggest either removing this figure, or simplifying it by summarizing more the main steps and goals of the approach taken in the manuscript (as an illustration for the end of the introduction).*

We agree that this figure is confusing when it comes before the explanation of our geometric approach in the text. We have removed it from the introduction and inserted it at the beginning of the discussion to provide an overview of our main results.

*Box 1 is very useful but it is cited only rarely in the text. I think further reference to this box would be very helpful to remind readers of critical steps and definitions of the approach (e.g. how change is measured at the system level in the geometrical approach).*

We have added references to Box 1 in some important places throughout the text, at lines 62, 122 and 195.

*Legend of Figure 3: in (c), please explain what corresponds to  $x$  and  $y$  in the equation and what it means (i.e. expected relationship between error and underestimation as derived from equation 4). From what I understood, the dashed red lines and the black points correspond to  $(\text{mean} - \text{sd})$  and  $(\text{mean} + \text{sd})$  and not to the values of the variances. This needs to be clarified. In addition, I would also explain that "underestimation" refers to the relative magnitude of underestimation as defined in equation (2).*

In the revised figure legend we have clarified what  $x$  and  $y$  correspond to and we refer to equation

4. Based on comments from Reviewer #1 we have also changed the statistics calculated in this figure (and in the following figure). Instead of mean and standard deviation we now consider median and quartiles. We have therefore included the following text in the figure legend:

Blue points are simulated results, red points are the actual median values and dashed lines show the quantiles for vertical subsets of the simulated data. As dimensionality increases the width of the distribution decreases and converges towards its median, which effectively increases the probability of underestimation (b).

*Legend of Figure 4, “The variance around the mean expectation was accurately predicted using the IPR instead of species richness”: I would explain why more clearly in the text. Indeed, if the variance around the mean expectation was well predicted by species richness, we would have the same variance in the two studied cases of biomass distribution as they have the same number of species.*

We have clarified this point in the main text by adding the following sentences (Lines 188-191):

If species richness accurately predicted the width of the distribution of underestimation and therefore the probability of underestimation, the two simulated communities in Fig. 3 would behave in the same way. However, the probability of underestimation is lower than expected based on richness, particularly for the community with a more uneven biomass distribution.

*Line 442 page 21: “we still see below”*

*Line 470 page 23: “probability of underestimation” instead of “probability of synergism”*

Both corrected, thanks.

*Examples page 25: it is not fully clear how these examples are related to what is presented in the main text, this would need to be clarified. More globally, I think the appendices could be linked a little more clearly to the main text.*

We have added some explanations and references to the main text. At the beginning of the section on aggregate properties we now write:

We measure  $S_F$  linear functions of species biomass of the form

$$F_\alpha(B) = \sum_{i=1}^S F_{\alpha,i} B_i \quad \alpha = 1, \dots, S_F$$

And then, in example 1:

This example is the one treated in the main text, where the functions are statistically independent of one another.

followed by:

At first order, the effective dimensionality  $S_{eff}$  is the harmonic mean

$$S_{eff} \approx \frac{1}{\frac{1}{S_F} + \frac{1}{IPR}}$$

as presented in the main text.

*Appendix page 29: This is not fully clear how the different aggregate functions are defined here. For instance, do they depend on species biomass or on other species properties? This point would deserve to be explained in the main text too.*

We modified the text so that in now reads:

Linear aggregate functions of the form

$$F(B) = \sum_{i=1}^S F_i B_i$$

were defined via the coefficients  $F_i$ , i.e. their sensitivity to the change in the biomass of species  $i$ . The sensitivity of an aggregate function to each species was randomly drawn from a normal distribution of mean 0 and standard deviation 1. This corresponds to the case of statistically independent functions (see example 1 in subsection S3.2)

## 2 Response to Reviewer 1

*In their manuscript, the authors argue that scaling up individual properties of complex systems to a system-level properties will necessarily result in an underestimation. The authors show that this effect is dimension dependent and they argue that in general the dimension should be computed as the inverse participation ratio. It is a well-written manuscript and, especially, I find the geometric approach very intuitive. I think this result deserves a recommendation in PCI.*

We thank the reviewer for their thorough evaluation of our manuscript, and for their positive comments. The reviewer's main point about properties of mathematical expectations, and the lack of explicit assumptions about probability distribution, is an important criticism. We address it in detail in this new version. We hope that it will clarify our results and their domain of validity.

*I have one major comment that the authors should first address. I do not think that the result applies to any type of "error"; they should be an implicate assumption about the "error" that the authors have to make explicit. I arrive to this conclusion as in general in probability theory we cannot switch between taking the expectation of a random variable and an arbitrary function, i.e. in general  $f(E(x))$  is not equal to  $E(f(x))$ . For example, let us assume  $x \sim$  Uniform distribution between  $-1$  and  $+1$  and  $f(x) = x^2$ . Clearly,  $E(x) = 0$ , but  $E(x^2) > 0 = (E(x))^2$ . So equations (4), (5), (7) and the mathematical derivation in the appendix work only for specific assumptions on the distribution of the "error". Stated as it is, they are simply wrong. The authors have to find under which assumption their mathematical derivation works and make it explicit in their manuscript. I guess the assumption is an independence assumptions between the "error" between, i.e., they may have to be *i.i.d.* distributed. I also find the authors should provide more mathematical reference justify their mathematical derivation.*

The Reviewer is right, the mean of a non linear function of a random variable is typically not equal to the function of the mean. The correct statistic to use is the median, and as long as the function of the random variable is monotonous, the median of the function is equal to the function of the median. Furthermore the median is more directly related to probabilities (it is the value such that the probability of a random variable taking a value greater than the median is 0.5). We thus use median and quartiles rather than mean and standard deviation in our

simulations and figures (Figs. 2 and 3) and we have changed our presentation accordingly:

Lines 144-147:  $P(y > 0; x)$  is driven by the distribution of the random term  $\cos \theta$ . If this distribution is narrow, realisations of  $y$  will fall close to the median  $\tilde{y}$ . Because the latter is positive and increases predictably with  $x$ , so will the probability of any realised  $y$  to be positive.

We also agree that our assumptions on the distributions of errors were not sufficiently explicit. We now explain that we only consider unbiased and unskewed distributions, i.e. distributions whose mean and median is zero. We then explain that our starting point is i.i.d. distributions of error, but that the whole point of our IPR construction is to relax this assumption by allowing errors to depend on the system's component (e.g. the species, with the error scaling with its abundance or biomass). We changed the text accordingly:

Lines 137-140: the term  $\cos \theta$  can take any values between  $-1$  and  $+1$ . For the sake of simplicity, in what follows we will suppose that its mean and median are zero. This is the case if the errors associated with individual variables are drawn from independent symmetric distributions centred on zero (unbiased and unskewed predictions at the component level).

Lines 154-159: We use the expression for the variance as a *definition* of *effective dimension*. In doing so, we can free ourselves from the strict Euclidean representation, and generalize the theory beyond i.i.d. normal error distributions. This will be useful when applying our theory to ecological problems, where components are the biomass of species, are their contribution to ecosystem change are not equivalent, thus errors not i.i.d.

### 3 Response to Reviewer 2

*Authors develop a framework to quantify the underestimation of the magnitude of a system level change when scaling up from species-level to ecosystem function (i.e. aggregated biomass). Authors argue that underestimation – and uncertainty – grow with the system dimensionality, with dimensionality not meaning more constituent species, but more diversity (i.e., diversity metrics like the inverse participation ratio or Hill numbers) – The explanation from authors is based on the geometric observation that in high dimensions there are more ways to be more different, than ways to be more similar. Authors provide a linear and nonlinear approx. to proof this statement. They go deeper to explain that nonlinearity controls the sensi-*

*tivity to underestimate upscaled predictions. Authors make a connection to stressed ecosystems: there will be bias towards synergism when multiple stressors predictions are scaled up to higher levels of organization.*

*Authors apply underestimation at two levels in ecology – biomass to diversity. Could the message be that the higher the dimensionality, for example from species to intraspecific or even to intraorganismal level for large populations or communities, the higher the underestimation of system change at these high dimensionality levels? What do authors think about generalizing (or discussing) their method to any number of levels? As authors notice in Box 5, this topic is important in many disciplines. It is relevant to any field of science that contains two or more levels each containing variance, and variance at intraorganismal and intraspecific levels might contain additional dimensions, like the number of traits or trait architecture of individuals.*

This is an interesting point, which certainly deserve clarification. When quantifying dimensional effects there are only two levels of organisation that are important: the level where predictions are made and the level where predictions are scaled up to. Knowledge of the effective dimensionality and the non-linearity of the aggregate property used to measure change at the higher level of organisation is all that is required to quantify the bias towards underestimation due to dimensional effects. The dimensionality of each individual component in the level where predictions are made (e.g. the number of individuals within a population) is not important. However if predictions are scaled up across multiple levels of organisation the dimensionality of multiple levels will be important. For example, if predictions are based at the level of individuals and are first scaled up to the species level and then scaled up to the ecosystem level, there will be two opportunities for dimensional effects to be introduced. Based on your comments we have included the following paragraph in the discussion (Lines 284 to 291):

We have only considered two levels of organisation: the level where predictions are made and the level where predictions are scaled up to. However, intermediate levels could, in principle, be considered. For instance, given the increasing resolution of ecological data, predictions of change may originally be based at the level of individual organisms and could first be scaled up to species-level predictions and subsequently scaled up to ecosystem-level predictions. Here, if non-linear aggregate properties are used, dimensional effects will bias species-level predictions towards underestimation and will further increase this bias for ecosystem-level predictions.

*Overall, with the increasing resolution of data in ecology, usually containing individual level data, accounting for uncertainty to quantify the bias in the mean field approaches is key, especially in the context of management of large ecosystems. Authors should emphasize more how the nature of ecological and other's disciplines data is challenging our understanding of uncertainty when accounting not only for 2 but for many levels. This relates to Box 2 – All these disciplines contain individuals varying in phenotypes, strategies and so on. Yet these heterogeneities within species are just ignored across disciplines. Are authors assuming all diversity metrics are based in mean field phenotypic distributions containing low variance? Why is this so?*

In line with our response to your above comment we agree that the increasing resolution of data means that predictions could be scaled up across multiple levels of organisation making it important to consider the dimensionality of multiple levels of organisation. We have added the following text to address this point (Lines 291-293):

With an ever-increasing resolution of data, scaling predictions across multiple levels of organisation, and potentially introducing dimensional effects at multiple levels, may become more common in the study of complex systems.

*Authors refer to "multidimensional system" to a system containing a species-rich ecosystem – do authors implicitly assume that each species increases ecosystem dimension in one? Why? Does this imply that all species living in a species-rich ecosystem make a perfect partition of one dimension per species? Please clarify.*

In section 3.1 *Effective Dimensionality* we address this point that each species does not necessarily increase the effective dimensionality of the system by one. Indeed, if there is an uneven biomass distribution and if a species' contribution to change is statistically proportional to its biomass, then taking the IPR (a specific diversity index) instead of species richness will give a much better approximation for the effective dimensionality of the system as shown in Figure 3.

*Authors emphasize their method predicts the generation towards non-additive synergism – They use a geometric method to prove this statement yet the processes underlying diversity metrics can be different while*



*the diversity metrics per se might remain similar – For example – rapid negative frequency dependent resource selection increases fitness of rare types, increasing the number of coexisting types within a species. This mechanism also balance species abundances and increases diversity. The mechanism of positive frequency dependent resource selection has the opposite impact, reducing intraspecific diversity while the mean types and the diversity metrics can be the same than in the previous case. How do authors think two opposite processes at intraspecific level change dimensionality (and uncertainty) at the diversity metrics level? Do these two processes provide alternative bias towards synergism and antagonism? For example, can this be tested exploring two different selection regimes using the simulated communities of Fig 5? Please clarify.*

We thank the Reviewer for these comments. To clarify, we use diversity in two different contexts in our paper. We first use a specific measure of diversity (the IPR) to quantify the effective dimensionality of the system, and therefore the expected bias towards underestimation when predictions are scaled up from individual components to the system level. Secondly, we use diversity (Shannon’s diversity index) as a non-linear aggregate property, commonly used by ecologists, to quantify system-level change.

Furthermore, in terms of multiple stressors we are not making predictions about what stressors, perhaps with opposing mechanisms as you suggest, will do. Instead, we are saying that if we have a model for how stressors combine (e.g. an additive model) at one level of organization and we want to deduce a model for their combined effect at higher levels, the process of scaling up will generate an underestimation of the combined effects, interpreted as a synergy between stressors, even if no systematic synergy was observed at the lower level of organization. To clarify this point we have added in the following sentence in our discussion of multiple stressors (Lines 320 to 323):

In fact, we are not making predictions about how stressors will behave at higher levels of organization. What we claim instead is that, if we have a model for the combined effect of stressors at one level of organization and use that model to deduce their combined effect at higher levels, the process of scaling up the model will introduce a bias towards an observed synergy between stressors, even if no systematic synergy was observed at the lower level