Author response to reviewer comments

We thank both anonymous reviewers for their detailed and constructive comments. We are able to respond to all of the two reviewer's comments, and in doing produce a much improved manuscript. We are pleased that Reviewer 1 considers that our manuscript "addresses the important and interesting topic of bifurcations in natural systems, and the possibility to understand and predict system stability using statistical indicators." and that the "the manuscript generally shows skilled and clear writing and a relatively thorough literature background." and that our "paper identifies a very promising research direction."

That said, we acknowledge that some significant revisions are required. We agree with Reviewer 1 in "that a potential revision of the paper should consist in a fundamental change and offer more specific and novel results". We are able to address these issues by clearly indicating the strength of the model and the novelty of the manuscript results. In addition to this, we will include new results in response to interesting avenues highlighted by reviewer comments. The major proposed changes are:

- The analysis of the core model in section 3 i) and ii) will be extended to build upon results from Dyke & Weaver (2013). In doing so we will clearly justify the scope of the model. We will also explain how the behaviour of this model is importantly independent from the particulars of the implementation. We have a potentially very general model that explores the dynamics between populations and environmental conditions.
- We will clearly explain why we do not explore non-stationary behaviour of the model in this manuscript. Previously, we have looked at this issue in a related model via a discussion on the separation of timescales (Weaver & Dyke 2013). While clearly important to the dynamics of the model, including results and analysis on different timescales here would detract from the central messages of this manuscript.
- The timeseries analysis of section 5 offers little in terms of new results. In large part, the presence of EWS is neatly explained by linear stability analysis although the variable skewness signal has a less apparent explanation. The analysis will be extended to discuss the existence of the skewness signal in detail, explaining correlations between perturbation strength, signal skewness and transitions.
- New results will exploit the multidimensional model by examining variations in EWS as a system is driven towards the edge of a 2+ dimensional attractor. The precise choice of multidimensional forcing is important; different forcing directions result in transitions to distinct new attractors while also displaying characteristic EWS. This better highlights the domain of multidimensional forcing terms over which an attractor exists, and the occurrence of EWS in multiple forcing dimensions.
- We will include a more detailed treatment and discussion on transition networks of multiple dimensions. This will explain why multitude environmental states prove to be inaccessible without direct control of multiple environmental variables. Consequently, we will show how the model exhibits not just large hysteresis loops, but potentially irreversible critical transitions – if perturbations acting on the system are limited to the initial drivers of the critical transition.

Below we respond directly to general comments from both reviewers, and then more detailed, specific comments.

Response to reviewer 1 general comments

We take the main thrust of Reviewer 1 comments to be that the manuscript does not sufficiently identify a novel conclusion or contribution to either general conceptual or specific empirical studies. We understand this conclusion. We are able to revise the manuscript such that it clearly identifies important contributions that would be of interest to the Earth Systems Dynamics audience.

In the following, reviewer comments are numbered and identified with indented text. Our replies our similarly numbered.

> 1.1 Whereas it is often stated that early warning signals are not necessarily applicable in complex systems, the paper does not seem to take this statement seriously in practice. Instead, it seems to be assumed that the model the authors apply would be representative of a huge class of ecological systems. In fact, the purpose to understand the relation between biota and its environment is rather a description of the whole field of ecology, rather than a specific research question. I therefore wonder what the analysis of such an unspecific model approach can tell us about reality.

1.1 We understand and agree with this comment. We failed to sufficiently stress the strengths of the model which are its simplicity and applicability to higher environmental dimensions. Our motivation was to generalize our understanding of EWS to multidimensional signals and more complicated networks of potential transitions than those that emerge from simple 1D examples. The analysis of this general model is intended to explore and better understand transitions in multidimensional dynamical systems and the relation to EWS.

> 1.2 It is of course well possible that there is something in the model that can be learned about certain systems that I am overlooking. But in this case I suggest that the authors work out this aspect in much more detail, and fundamentally revise their manuscript. I basically see two possibilities: either focus on certain properties of real ecosystems, or focus on a certain mathematical aspect of the model that is relevant for the application of statistical indicators in complex systems in general.

1.2 Again, we agree with this conclusion; in attempting to tread the line between these two possibilities, the message is unclear. We will revise the manuscript to clearly pursue the second possibility. We will begin by presenting the results of the original (Dyke & Weaver 2013) model analysis, emphasising the generality of analysed behaviour and those model choices which can be shown as arbitrary; the form of the “biotic effect function”, the nonlinear model component, is independent from the precise choice of individual response functions, their number and diversity. In this sense, further analysis can be seen as applicable to a broad class of (simplified) systems. Along with this, the direction of analysis should exploit the strengths of the model by rooting the analysis in multi-dimensional EWS, and avoiding the (introductory level) time series analysis altogether.

> 1.3 As far as I can see, the main message that the current results offer is the notion that the signal strength of early warning signals is not related to the degree of irreversibility or the distance from the lost equilibrium to the new equilibrium. This is of course a true statement, because statistical indicators can only yield information about the local flow, but not on the global bifurcation structure of a system. However,

this is not a new discovery that would merit a publication, and it does not require the application of such a complex model.

1.3 There is potential for some simple analysis to formalize these observations. EWS in the model are related to the gradient of the "biotic effect function". The rate of change of EWS as the system approaches a bifurcation is therefore associated with the second derivative of the function which may be used to predict the occurrence of the next stable point. Due to the underlying Gaussian process analysis however, uncertainty increases exponentially with the characteristic scale of the underlying function, providing very limited information over the expected transition scale.

1.4 The fact that the vector field F(E) is not conservative is another interesting aspect that could be chosen as a focus. If the authors should decide to explore this in more detail, they should connect to other literature on early warning signals in highdimensional systems such as ecology and climate. Examples are (besides the cited Held and Kleinen, 2004), Hastings and Wysham (2010), Dakos et al. (2010, 2011), Bathiany et al. (2013), Mheen et al. (2013), and Williamson and Lenton (2015).

1.4 The fact that the biotic effect function is not conservative only appears surprising in light of common EWS illustrations where a potential well analogy naturally leads to assumptions of conservation. Indeed, conservation would exclude a range of phenomena from occurring. Perhaps the fact that more phenomena do not emerge in spite of this implies a "missing" component, although as discussed in more detail in 1.9, such phenomena can emerge through variation of the timescales of change within model components. In either case, this certainly warrants a thorough representation within the discussion and analysis of the model, and we thank the reviewer for providing suitable literature as a starting point for this discussion.

Response to reviewer 1 specific comments

1.5 Title Very vague and general

1.5 We believe that a more appropriate title for the revised manuscript would be

"On the relationship between early warning signals and critical transitions in higher dimensional ecosystems"

1.6 Introduction p. 2509, line 10-15. Drawing general and far-reaching conclusions from such an idealised model is a very big promise that is not possible to keep, in my opinion. The main problem is, that one cannot have both, complexity and generality. If there is such a relation it is specific to a certain aspect of reality, and it normally emerges from many studies, not just one. This shows that the scope of the paper is too broad.

1.6 Our claim of generality for the model is based on previous work that we did not sufficiently explain (Dyke & Weaver 2013). We are motivated to explore interactions between populations and environmental conditions and how such interactions can produce critical transitions. A classic real world example is a positive feedack loop operating in freshwater shallow lakes that drive them into eutrophic states. The dynamics are potentially complex; while some simple dynamical systems treatments empirically parameterise certain functions in order to capture key dynamics (Scheffer & Carpenter 2003), our approach is to produce a process-based model that produces critical transitions. We do so by making a relatively small number of assumptions, such as biological organisms have a narrow fundamental niche of environmental conditions. What we have found, is that a model formulated with such simple assumptions displays both stability, that is the emergence of attractive fixed points, and critical transitions, where rapid changes in state occur which flip the system from one fixed point to another.

While simple, the model allows analysis of multiple environmental variables and consequently multiple potential EWS. We believe this to be a novel and important contribution to the EWS literature, and the results, discussion and conclusion will be improved to reflect the strengths of the model in providing insights of critical transitions and EWS beyond a 1D projection.

> 1.7 Sect. 2 It is indicated that "critical transitions" naturally emerge in complex systems and that they would occur "in almost all real systems". I do not believe this. For example, the climate is a very complex system, and there is indeed evidence of abrupt changes (as the authors remind us). But such transitions are the exception rather than the norm, and multiple states are not so common in comprehensive models. I think that the notion that such phenomena can occur would be more appropriate than saying that they are everywhere.

1.7 We appreciate this comment. We did not sufficiently defend such statements. In part, this is an issue of timescales. The Earth's climate is currently in an essentially bi-stable state of glacial-interglacial, a state it has resided in for some 2.2 milllion years. Much of the contemporary concern over anthropogenic climate change, comes from potential tipping elements in the climate system (Lenton et al 2008) which can be understood as critical transitions. At risk of complicating matters, we again refer to our motivation of producing a very general model - a model that in principle could capture dynamics over a very wide range of temporal and spatial scales. We will revise the text to remove "almost all real systems" which is too strong, and more carefully refine the scope of systems we are primarily interested in.

> 1.8 Sect. 3 It should be much better described what the purpose of the model is. The popular and adequate notion that "all models are wrong, but some are useful" and that "there is no model OF something, but a model is made FOR something" nicely summarise what I perceive as the problem of this paper. It could also be that the structure of the model already poses an interesting setup (e.g. due to the many dimensions, the niche approach and the possibility of non-autonomous forcing. But then there should be experiments that systematically explore what this structure specifically means for the applicability of early warnings.

1.8 We have addressed this comment with replies 1.1 to 1.7 above.

> 1.9 By the way, the model seems to be complex enough to show all kinds of complex behaviour such as chaos. Yet, only stable states seem to occur. I wonder why the solutions of the model are so simple despite its complexity.

1.9 The initial novel result found in this model (Dyke & Weaver 2013) is the emergence of stable states, given potential assumptions that a randomly connected system would produce a much richer and less stable behaviours (May 1973). The model has the potential to exhibit chaos and cyclic behaviour. The key factor here are the timescales of interaction, which we have explored previously in a similar model (Weaver & Dyke 2012); separation of timescales between the changes in the biota and environment suppress more complex dynamics. It would be possible to conduct such analysis on this model, but to do so would run the risk of further complicating the key message – something that we acknowledge we did not successfully do with our initial submission!

> 1.10 Sect. 4 p. 2519, line 19-20. The possibility of hidden equilibria that can only be realised by changing another parameter or by perturbing the state directly, is a very interesting idea. This fact alone is not anything new yet, but might be another aspect to explore in more detail in some way. At least I am not aware of publications that focus on this aspect.

1.10 This feature can certainly be highlighted and illustrated in detail, receiving a similar treatment to cusp bifurcations which are discussed in the manuscript. While we feel this is captured by existing figures, we will extend these figures and the surrounding text to discuss this feature more clearly. In particular, we have previously relied on 1D projections of multidimensional systems to convey our results while this feature will be better demonstrated in a 2+D representation.

> 1.11 The sections where actual results are shown are very short in comparison with the rest of the paper. This should be better balanced.

1.11 We agree. This was a consequence of insufficient explanation of some of the important elements and results of the model. The inclusion of mathematical analysis and results in the ways suggested will result in a more balanced manuscript.

> 1.12 Sect. 5 The discussion about skewness and why it behaves that way in the model is unclear to me, and I see no experiments or calculations that would reveal the reason.

1.12 Two of the EWS measures (CSD and variance) are relatively easily explained by linear-stability analysis of the model fixed points, while the skewness signal, and how this can be shown to coincide with critical transitions, requires some additional explanation. Other EWS are related to the derivative of the biotic effect function, skewness is a measure of the second derivative; relatively large skewness measures indicate the gradient is changing relatively rapidly, and therefore that the second derivative of the function is large compared to the first derivative. We will show analytically that a large perturbation magnitude is correlated both with transitions and a large second derivative.

Response to reviewer 2 general comments

Reviewer 2 raises some similar issues to Reviewer 1, in particular proposed novelty. Again, we understand this assessment. In addressing Reviewer 1 comments, we believe that we also address the most substantive comments of Reviewer 2. Below we reply to directly to all Reviewer 2 comments.

> 2.1 However, with regard to early warning signals (EWS), which seems to be the focus of the manuscript as indicated by the title and the abstract, I do not see any new results.

2.1 As mentioned in the response to the general comments of Reviewer 1, we will emphasise the general and mathematical aspects of the model, particularly its potential to investigate otherwise simple concepts in multidimensional environments including critical transitions, early warning signals and cusp bifurcations. A part of this change is to shift the focus of analysis and discussion away from the (at times introductory level) time series analysis, and towards novel results revolving around the mathematical tractability of the model, and its multidimensional environment.

> 2.2 This is well known (see for example any nonlinear dynamics text book). The EWS autocorrelation and variance are just functions of the eigenvalue of the linearised dynamics and this is also well known (Held Kleinen, 04) so I don't see any novelty here. The calculation of the EWS via numerics therefore seems unnecessary as it confirms what is already demonstrated in the analytics.

2.2 We agree that the manuscript is overly focused on this aspect of the analysis which is essentially well trodden ground, and simply explained by linear-stability analysis. As with our reply 1.2, we have decided that the best way to improve the manuscript is to switch focus onto the strengths of the model; its multidimensionality, simplicity and associated tractability. These features will be leveraged to investigate otherwise simple concepts in multidimensional environments. This will include critical transitions in 2+ dimensions, especially the emergence of cusp-bifurcations, "hidden" attractors which are accessible only by multivariable forcing or environmental displacement (1.10), and multivariate EWS, where driving towards the edge of an attractor basin produces a EWS and transition both sensitive to forcing variables.

> 2.3 Skewness is related to the locally dominant nonlinear term in the dynamics and one may get more information about the bifurcation. This is discussed but not very clearly. See Sieber and Thompson (2012) for work in this direction. Again, this can be investigated directly from the equations presented making the numerics unnecessary. The authors may argue that as this is a multidimensional set of equations this is not necessarily expected which is true.

2.3 This point is already covered by 1.3 and 1.12; a shift towards more mathematical analysis will provide clearer and potentially more succinct explanations for the emergence of the discussed EWS. The ability of each EWS to provide information of the following transition will also be discussed in detail as the defined characteristic length scale of the biotic effect function allows us to formalise notions of locality and compare directly to the characteristic transitions scale.

> 2.4 Stability of an attractor can be lost through global bifurcations (no critical slow down), however their analytics indicate this is a local bifurcation which is defined through its local stability. The authors only consider simple point attractors, rather than more exotic periodic or chaotic attractors that might be present in a multidimensional system.

2.4 There is clear and specific interest in more interesting attractor behaviour, also addressed in our reply 1.8. While certainly possible in environments with the dimensionality we study, the simplified behaviour results directly from explicitly separated timescales between the changes in biota and environment. This is not studied in the manuscript to avoid obfuscating the core message, although the absence of more complex attractors, in spite of the possibility, is noteworthy and will be expressed in a revised manuscript.

> 2.5 The main claims of the manuscript regarding early warnings are not novel, however the ecosystem model, depending on how much of it has been previously published, potentially is, although it is very abstract and difficult to map to specific systems.
>
> This may be interesting to biologists although I am unqualified to comment on this. Most of the manuscript is spent discussing this model. It appears the authors have published variants of this model mainly in biology journals. I wonder whether the paper would be a better fit in these journals?
>
> The manuscript here could fit in ESD under point 3 of the journal scope (Earth system interactions with the biosphere) although at a guess the emphasis should probably need to be more on the properties ecosystem model, its specific relevance to the Earth system and what is new in the manuscript.

2.5 The scope of the systems to which our model analysis is applicable to is addressed largely by our replies 1.1 and 1.2. We agree that our manuscript would be of interest to the ESD audience, in particular:

“**Dynamics of the Earth System**: contributions may include novel concepts and theories and their application to better understand the nature of the Earth as a complex, coupled system; experimental work of dynamic self-regulating systems relevant to Earth system functioning; historical and philosophical perspectives of Earth system theory; models to describe and predict the dynamics of the whole Earth system; and studies on the importance of interactions between component systems.”

“**Earth system interactions with the biosphere**: contributions may include mechanisms of ecosystem–Earth system interactions; conceptual models of life–Earth co-evolution; simulation of biogeochemical cycles within an Earth system context.”

Response to reviewer 2 general comments

> 2.6 Intro: It seems the authors definition of critical transitions is defined as a loss of stability via a local bifurcation. At least this seems to be the only mechanism explicitly considered in the paper. It would be good to clarify this. Critical transitions can also happen via global bifurcations where CSD is not necessarily expected.

> Abrupt changes in states however can happen in many different ways not associated to bifurcations. This should probably be discussed. See Ashwin et al (2012) for more info.

2.6 We thank the reviewer for this is excellent suggestion which affords us a good opportunity to provide insights into the role of dimensionality in the model by providing a schematic of model fixed points across a transition. Our understanding of transitions in the model is very simple, largely as a result of the absence of anything more interesting than point attractors; transitions occur at the intersection of the independent nullclines in each environmental variable. Stable points are lost when they annihilate with unstable fixed points. This analysis would become much deeper if we relaxed our assumption of (nearly) separated timescales which opens the door for a much more nuanced investigation. While we believe that substantive analysis in this area is probably beyond the scope of this manuscript, we will include initial discussions on some of the consequences of relaxing timescales and exploring more complex attractors in the model.

> 2.7 Section 2, line 6: systems approaching a critical point see increasing variance in state variables if stochastically perturbed. Only true if approaching a local bifurcation because of CSD.

2.7 As with reply 2.6, we will clarify what types of transitions emerge from the model. In this instance it is simple to include in the manuscript that the convergence of stable and unstable fixed points leads to large eigenvalues

> 2.8 Section 2, line 13. Don't understand sentence starting 'This is referred...'

2.8 The sentences in question are *"In addition to large susceptibility, the near-critical dynamics of such systems are slow such that fluctuations which do not cause transitions are reigned back to the steady-state very weakly. This is referred to as an increasing decay time, or as a decreasing recovery rate."*

Terms such as "decay time" of fluctuations and the "recovery rate" from perturbations are common in ecology literature. We hoped to relate them to the concept of CSD which may be interpreted in these ways, although this text is unwieldy and will be made clearer for the audience it intends to serve by the following replacement

"Along with an increasing variance in the state variables, the recovery rate of fluctuations about the mean state is small, or equivalently we may say the decay time of these perturbations is large."

> 2.9 Equation (3) does the term |.| denote vector norm? Seems this is the case but needs to be made clear in text.

2.9 Yes. This will be clarified by the following addition inserted immediately after Equation (3)

"where σ_E is the niche width and $|\boldsymbol{\mu}_i - \boldsymbol{E}|$ denotes the vector norm, yielding the distance of the niche of the ith biotic element, $\boldsymbol{\mu}_i$, from the current state of the environment, $\boldsymbol{E}$."

> 2.10 Lower case omega used in figure 3 not defined.

2.10 It is a part of our convention to use indexed lowercase symbols to denote individual elements of vectors and matrices. This will be clarified by the following addition immediately after Equation (4)

"The effects in $\mathbf{\Omega}$ *are denoted* $\omega_{i,j}$ *which gives the linear influence of the* i*th biotic element on the* j*th environmental variable. These are chosen independently from the standard normal distribution, where the mean value of zero provides no bias towards positive or negative life-environment coupling"*

References

Dyke, J. G. and Weaver, I. S. (2013): The emergence of environmental homeostasis in complex ecosystems, PLoS computational biology, 9, e1003 050, doi:10.1371/journal.pcbi.1003050

Weaver, I. S. and Dyke, J. G. (2012): The importance of timescales for the emergence of environmental self-regulation, Journal of Theoretical Biology, 313, 172–180, doi:10.1016/j.jtbi.2012.07.034

May, R. M (1973): Stability and complexity in model ecosystems, 6, Princeton University Press

Scheffer, M., & Carpenter, S. R. (2003). Catastrophic regime shifts in ecosystems: linking theory to observation. Trends in Ecology & Evolution, 18(12), 648–656. http://doi.org/10.1016/j.tree.2003.09.002