Associate Editor Decision:
Reconsider after major revisions (22 Dec 2014) by Dr José M. Moreno

Comments to the Author:

Ecosystem functioning is sensitive to climate extremes, but extremes of a similar magnitude not always elicit the same functional response. Thus, understanding vulnerability to climate extremes is utmost important. Rolinski et al. developed an innovative approach to this question by framing vulnerability, focusing on a given ecosystem function, into a probabilistic framework in relation to the relevant climate extreme (drought in this case, among other). While both reviewers have acknowledged progress from the previous version, and recognized the attractiveness of this approaches to the question of concern, which I agree, some major critical issues remain to be solved before this approach can be accepted. A revised version will be considered once the following issues have been adequately dealt with: 1) greater explanations about mechanisms; 2) Calculations of the SPEI index; 3) the time window chosen and the likely consequences on observed results, including autocorrelation issues. The authors may wish to consider a greater focus on processes within the model that are more robust. The text is much improved but homogenization of the concept of meterological extreme/climate extreme/weather extreme would be helpful to the readers.

We appreciate this thoughtful and comprehensive summary of our effort to improve the previous version and the reviewer's efforts again to provide us with specific recommendations and items to be clarified in the revised manuscript. We respond to the raised issues in the following text with the editor/reviewer comments marked in italic. .

1. Given the paucity of long-term data for the problem object of study, models are used. A first major issue raised during the review is that the paper runs short of in-depth explanations about mechanisms of observed fluxes that could explain the observed patterns. This is important because not all functions in the model are equally robust (eg., photosynthesis vs. are burned).

Although we have confidence in the results of our model, we agree that the application of our probabilistic framework to model results can only elicit the ability of the respective model to represent ecosystem responses to weather conditions and extremes. Therefore, we give more details on papers in which the impact of (extreme) drought on vegetation is quantified and which show that LPJmL can be applied to drought impact studies. Please see our detailed reply to reviewer #1.

2. A second major issue refers to the actual calculations of the index of drought chosen.

We have to apologize for the misleading description (lines 206-208) that led reviewer #1 to conclude we would use long-term (110 years) SPEI

values in the vulnerability calculation, which has never been the case. We have now corrected the misleading sentences and better explain what is done here. Please see the detailed response to reviewer #1.

3. A third important issue refers to the time-window chosen and how this may affect the results, particularly when doing a continental comparison in which sensitivity to previous periods of drought may vary substantially. Actually, droughts are of very different duration and intensities, and long-periods of below-average rainfall, each of which is not much different than the long-term mean, could result in drought. Patterns across the continent for this could vary largely, a fact that needs reflection. Also, problems of autocorrelation related to the temporal period of observation (months) and in relation to the window of reference (12 months) need also reflection.

We have added more information on the reason to choose a 12-month period of negative NBP to define the hazard (see detailed answer to reviewer #1). Furthermore, we would like to note that the different aspects of drought events (intensity, duration and recurrence) are covered in the climate data set that we use to drive the LPJmL model and to calculate the SPEI index. To integrate all these aspects of drought, including their variability, we have specifically chosen a 30-year period which should enable us to quantify vulnerability from immediate and lagged responses to drought.

In your response, please consider each of the items mentioned by reviewer #2. Focusing on those functions that are more robust within the models, and of which a better causal understanding is available, would help clarify the results. For example, a recently burned forest will be differentially vulnerable to the same drought (i.e., duration and intensity) than a mature forest, and issue that is not mentioned. Vulnerability is considered simply as a response variable, but there are many processes behind that, as rightly pointed during the review, need consideration.

We thank the editor and the review to help improving the manuscript text in that respect. We now have added information on the process contributions to the hazardous NBP in the additional Figure 3 for the same sites that are also illustrated in Fig. 2. NBP is decomposed as NPP minus heterotrophic respiration and losses due to fire, and this new analysis shows that the major impact of water shortages on NBP is via the reduction of NPP. We have updated our statements in the results section accordingly and they are now backed up with a description of the dominant process.

Please, also consider polishing the use of concepts (meteorological extremes, climate extremes, weather extremes). Climate extremes are used in the title, but little reference is made thereafter. Most of the time meteorological extremes are used and this is explained by making reference to the terminology used in the IPCC SREX report and other recent (AR5) reports. Yet, there, weather extreme

is the reference concept, not meteorological extremes, which is equated to climate extreme.
Perhaps, to avoid misunderstanding, clarifying these concepts or using the one that is referred to, would help the reader to not believe that, in the end, we are speaking of the same concepts.

The use of the terms weather, meteorological and climate extremes is indeed ambiguous in the literature and was mixed in our manuscript. Using monthly average values, we were unsure which term is most appropriate but you are absolutely right with the reference on the SREX report. Thus, following the definition of the SREX report we changed our wording to 'weather extremes' in the title and throughout the text.

Anonymous Referee #1

I reviewed for the revised manuscript from Rolinski et al. about 'A probabilistic risk assessment for the vulnerability of the European carbon cycle to climate extremes: The ecosystem perspective'. I acknowledge the correction provided and the gain in clarity of the manuscript. However, I am still concerned by a lack of understanding of underlying processes in this nicely presented conceptual risk assessment. In turn, after reading the manuscript, I cannot conclude if the framework captured the actual functioning of the system with ecological meaning or just provides evidences within a bunch of statistical analysis. We definitely lack clarity in the ecological explanations of observed fluxes, particularly how much anomalies are due to NPP, respiration or disturbances, the three components of the NBP. I mentioned few key papers, among other which try to stick a little more to ecological evidences of ecosystem functioning.

We appreciate this detailed and differentiated resume. We have to admit that we were very reluctant to enter the discussion on underlying processes for two reasons: 1) We thought that it would rather dilute the focus of the paper which we saw in the presentation of the method and 2) we thought that it would not contribute to more insight and clarity. Here, we have to reconsider our line of thought and now follow your point of view as we point out in the detailed comments below.

I am also concerned with a last point on SPEI calculations over long periods (century) with climatic trends. In explain below potential biases which would benefit from clarifications. In turn, I am convinced by the potential benefits of this conceptual approach, but frightened about conclusions not supported by an extensive literature review of observations and intermediate informations as NEP, Resp , disturbance that would help to understand the results. I pointed out some potential bias in SPEI that would benefit to be clarified.

We regret this misleading text on the calculation of the SPEI. Please see the detailed comments.

Detailed informations below(document):
L123-L126: references to add

Baudis et al. 2014. Intraspecific differences in responses to rainshelter-induced drought and competition of Fagus sylvatica L. across Germany. Forest Ecology and management 330: 283-292.

Baudis et al. 2014 study the impact of drought in leaf stomatal conductance during an experimental drought. However, the point that we want to make in the sentence (L. 126- L128) is that ecosystems can be resilient to a single extreme event and allow for recovery to pre-conditions. This part has not been analysed by Baudis et al.

Barbeta A. Ogaya R. Penuelas J. 2013. Dampening effects of long-term experimental drought on growth and mortality rates of a Holm oak forest. Global change biology 19(10): 3133-3144
Martin Stpaul et al. 2014 The temporal response to drought in a Mediterranean evergreen tree: comparing a regional precipitation gradient and a throughfall exclusion experiment. Global Change biology 19(8): 2413-2426

We thank the reviewer for pointing us to these valuable papers. We now cite both to give an example of drought impacts and ecosystem recovery or plant acclimation (lines 121-125).

L 300-320: very sad again to have not even a single reference on observations and conclude 'Therefore, the LPJmL model is indeed capable of capturing dynamic responses to, e.g., single or consecutive drought events'.

The paragraph was intended to explain in more detail the processes implemented in the model that are responsible for the resulting responses on water shortages under different temperatures taking long-term (memory) effects of the plants into account. Here, we included some references to paper in which LPJmL results are compared to other models and data and in which water-plant interactions play a dominant role (lines 307-320).

L342-345 Biomass burnt (BB) result from dead and live fuel consumption in surface fires and from crown scorching (Thonicke et al., 2010) and is included in the carbon balance NBP (Eq.4) => do we have necromass assessment and tree mortality as reviewed in Mcdowell et al. 2013.
McDowell N.G. et al. 2013. News Phytol 200(2): 304-321

The paper by McDowell et al. presents a specific data-model comparison study in which the model performance to simulate plant mortality from carbon starvation and hydraulic failure for two temperate tree species in Western US is evaluated. The physiology studied in this paper is of high importance for all dynamic vegetation models and provides suggestions for refinement of mortality functions which are not only empirically based on mortality rates but give detailed insights on how physiological processes can be incorporated that then lead to drought-induced mortality, thus reducing model uncertainty. However, we think the McDowell et al. paper is not ideal in the context of our study, partly because the investigated species are not representative for our study region. The data presented in this paper are thus not directly usable to

evaluate necromass and mortality rates in the LPJmL version that we use in our study. However, the reviewer has made an important point and so we added the Evangeliou et al. 2015 publication where the amount of litter or necromass simulated by LPJmL was evaluated to constrain emissions from biomass burning. We have now added a sentence including the citation in lines 372-374.

L402-404:"The border region of Ukraine, Belarus and Russia is most pronounced with positive VE values of more than 0.4 where fire has a large impact on the carbon balance." => no information/result/graph on simulated fire activity illustrates this point. All along the manuscript, we never know if changes occur in NPP, heterotrophic respiration or fire modify NBP… it's kind of a black box

In order to evaluate the contribution of the different components of NBP for the hazard and non-hazard groups, we derived new figures (Fig. 3 and appendix B). We acknowledge that this additional analysis is a major advancement for our study as it also helps to further explain our vulnerability approach. We are grateful for your perseverance in this respect. We hope to contribute to a better understanding of the model processes and of the resulting vulnerability measure by the additional analysis and additions to the text (Fig. 3, lines 419-441 and 447-458).

L473-474: why only a 12 months period? Leaf life span in evergreen forests can last up to 3 years, while in grasslands it's only 6 months. We might expect different impact delays of a 3 months drought (SPEI3) depending on leaf life span of the ecosystem. Please discuss this point and potential weaknesses of this single 12 months value in your approach.

The selected time window to identify a hazard in the bosphere is not defined in relation to the life span of leafs or plants, but based on weather conditions. We have chosen the 12-month period to distinguish the responses of the biosphere from typical seasonal effects. In climates with dry summers, plants adapt to drought by being not productive during that time period. So, a non-productive summer should not be classified as being vulnerable since it is an adaptation to a regular recurring climate feature. Barbeta et al. (GCB, 2013) also analysed long-term effects of drought. In our study we take the ecosystems perspective when net biome productivity is negative over 1 year. This measure integrates reduced plant productivity beyond average climate conditions, as well as fire disturbance and increased heterotrophic respiration. All of them have the potential to be responsible for carbon loss from the ecosystem, i.e. going beyond the individual plants perspective. Longer-term effects could be included in the analysis by choosing a longer time-window, i.e. a 24-month period. We think that this is beyond the scope of this paper but interesting for further analysis. We have now inserted some additional text explaining this point better (lines 232-241, 531-533, 539-541).

Also, how much LAI adjustment to icreasing drought is actually simulated in LPJmL and fits the Carnicer Results (figure and reference below)?

Carnicer J. et al. 2011. PNAS 108(4): 1474- 1478
This might be crucial for understanding drought effects, highly balanced by an hydroecological adjustment of LAI along increasing drought trend which could mitigate drought impact on NPP. (Figure from Carnicer et al. 2011)

In the LPJmL model, photosynthesis is based on the Farquhar approach (Sitch et al. 2003) where transpiration and CO_2 uptake into the intercellular spaces of the leaf are coupled through canopy conductance. Plant assimilation is then calculated according to the limitation by light and Rubisco activity. The latter includes a temperature inhibition function to define an upper and lower temperature limit where photosynthetic activity is possible in the plants. This assimilation rate is first calculated under non-water-limited conditions, to allow for the calculation of other water balance variables, and then, given the atmospheric demand, recalculated for the water-limited situation. This way, drought or high temperature signals in the climate data set have an impact on photosynthetic capacity in the model.

By substracting maintenance respiration costs from gross primary productivity, net primary productivity is determined. Based on the allometric rules in the model, the amount of carbon which can be allocated to new leaves can be derived so that the leave carbon pool increases. The amount of carbon in leaves, together with SLA and crown area, determine LAI in the model (Sitch et al. 2003, eq. 5). This means that under drought conditions the plant assimilation is down-regulated, reducing the increment of carbon to the leaf carbon pool which then reduces LAI. So, if we see a reduction in plant productivity (NPP) in the model in a specific region, the model also simulates a reduction in LAI, given the model structure.

Unfortunately, simulated LAI values were not stored in the model output, so we regret we cannot provide you with a detailed examination of this model functionality from currently simulated model results. However, given the fact that the hydrology of the model and the productivity has been validated in previous studies, which we cite in our paper, we are confident that the model describes the respective plant physiology reasonably well. Nevertheless, we have learnt that this can be an important point for model evaluation and the cited Carnicer et al. paper as well as follow-up studies can be a valuable test. In that sense, we take your criticism as an encouragement for future model studies.

L 485-495: we appreciate this point.

Thank you very much.

Last concern regarding SPEI: I have a last question about SPEI3 calculations: the authors mention SPEI calculations over the 1901-2010(L206-L208). Then analysis are performed over the 1981-2010 period. I am here concerned about the climate trend over the SPEI3 calculation. SPEI calculations, by construction, fit a normal distribution over the whole dataset to identify deviation for this

distribution. In turn, SPEI values for a given month are dependant on the normal distribution of the whole reference dataset. It means that calculating SPEI3 over the 1981-2010 or the 1901-2010 might lead to different values for the monthly SPEI3 over the 1981-2010 period if the distributions for the two periods are different. They actually might be different considering recent observed climate changes and as is illustrated in the MED SPEI3 trend from the vicente-serrano database. the continental heterogeneity in climate trends and the subsequent bias in SPEI3 calculations. To clarify this point, I suggest the authors to provide continental pixel trend analysis on SPEI3 and the regional values of SPEI3 trend for MED, NEU, CEU for the 1901-2010 period.
I illustrate my concern with some calculations I just performed with the SPEI R package using the default Wichita climate database. I computed SPEI3 for the period provided by the default dataset (1980-2011) and I built a theoretical increasing and decreasing trend by replicating the default time serie 3 times and artificially increasing/decreasing PET and decreasing/increasing Precipitation.

We are very grateful for the reviewer's comment regarding the calculation of the SPEI index and its use in the vulnerability calculation. We apologize to have misguided the reviewer in how the SPEI calculations were performed and used in our analysis. Through the detailed explanation and examples provided we finally understood what the critical point in our manuscript text was that lead to the confusion: The sentence in L206-L208 gave the impression that we used SPEI values from 110 years in the hazard calculation, thus ignoring any trend that has occurred during this long period. This, of course, contradicts with the description of results covering the time period 1981-2010 in the results section. This is not what we are doing: The climate data cover 1901-2010 which we use as forcing for LPJmL. For the analysis, we calculate SPEI values for the investigation period 1981 to 2010 (with exactly the same R package). We have now understood that this was not well described in the text and led to the confusion. We therefore do not run into problems where decadal climatic trends influence our vulnerability analysis as the reviewer rightly pointed out.
We have now corrected the years in the sentence (L207) so that the study period 1981-2010 is clearly distinguished from the input data (section3.1) and the modelling protocol (section 3.3)

Anonymous Referee #2

In my view the authors have significantly improved the ms, including a thoughtful section in the Discussion regarding limitations of their approach. The method is useful for dealing with complez DGVM outputs and for policy makers (in combination with other diagnostic tools).

Thank you very much for this evaluation.