Response to Anonymous Referee #2

Received and published: 2 May 2018

Review for bg-2018-130, Flach et al., "Contrasting biosphere responses to hydrometeorological extremes: revisiting the 2010 western Russian Heatwave."

Flach and colleagues, using a multivariate spatiotemporal anomaly detection algorithm on both climate and ecosystem variables, assess the response of productivity to the Russian heat wave of 2010. Motivated by the potential for inconsistencies in the climate event and the biospheric impact (which they suggest is a function of disciplinary divides) they find that an anomalous spring warming event in both the biosphere and climate increased GPP prior to the actual heat wave itself, which occurred later in summer, thus offsetting the negative productivity effects. They note that the compensation occurs in different ecosystems–losses dominated in lower latitude managed ecosytems, such as crop land, while spring gains dominated in higher latitude forested regions. During the heat event itself, they attributed the differential response of forests and crops to different water management strategies of the vegetation classes. Overall the paper is a nice contribution and appears methodologically sound (if not a bit overcomplicated in places). I have a few comments and suggestions for the authors to consider that I hope will help improve the clarity and argument of the paper.

Response: We would like to thank the reviewer for the positive evaluation.

Main comments:

1. Stated motivation: While I am sympathetic to the larger issue that climate extremes and climate impacts are distinct domains and that extremes may not necessarily map to impacts, I find parts of the introduction to be somewhat of a 'straw man.' The hydro and bio perspectives generally do agree on the Russian heat wave–warm temperatures, along with dry soils leads to carbon loss. Consider the fact, for example, that the authors' very own agnostic algorithm finds the same two events in both the met and bio fields; it suggests that the RHW at least, this disconnect does not lead to inconsistent interpretations or conclusions among different disciplines. The notion that there isn't a one-to-one mapping between the geophysical event and the biophysical impact is certainly important for accurately representing the total effects as a

function of the differential vulnerabilities of ecosystems. The authors rightfully emphasize this. However, the notion that this issue is emblematic of some kind of disciplinary divide is over-reach, or at the very least, is not supported by the literature the authors cite here. I heartedly agree that a call for an integrative perspective is a good one, as it can provide both a richer treatment of an extreme event and a basis for better impacts prediction, but the way the introduction is cast at present overstates the extent to which disciplinary perspectives are or were an issue in some kind of misdiagnosis of the RHW. This can be seen, for example, at 3.10, where the authors state that because the GPP declines were not as large as the temperature anomalies in Fig. 1, that this is somehow reflective of "different disciplinary perspectives" rather than of the complexity of the Earth system itself. . .leading the authors to "suspect [. . .it] might become an issue in studies of this kind." If the authors provided a stronger basis in the literature of inconsistent conclusions of the impacts of the RHW or similar events based on disciplinary divides, then sure, the way the intro is written can stand, but I think as is, it overstates it as a problem and diminishes the scientific conclusions of the paper, which are interesting in and of themselves. The point is, those interesting results and the science itself, gets a bit lost in the straw man discourse. Edits to the text can fix this.

Response: We would like to thank the reviewer for the positive view on the scientific conclusions of our manuscript. We agree with the reviewer, that we somehow overstated the disciplinary differences on the existing literature at the basis of the Russian Heatwave. We will carefully revise the abstract and the introduction to fix issues of this kind. In particular, we will rephrase the motivation at 3.10. along the lines the reviewer suggested (more focused on our own results, highlighting the call for an integrated perspective):

"The objective of this paper is therefore to revisit the RHW and to investigate the GPP response during the spring event and the summer heatwave when adopting a hydrometeological driver vs. a biospheric perpective."

Furthermore, we will reformulate the sentence on 3.2-3 to: "However, an integrated assessment including the hydrometeorological and the biospheric domain may facilitate our knowledge about the RHW. In particular, we highlight one aspect of the RHW which can easily be seen, i.e., if we look at the zonal evolution of the RHW in both domains" and remove two sentences in the abstract (1.5 and 1.16-17).

2. Two events v. one event: My comment here is a corollary to the above about how the paper is cast relative to the literature. The authors are taking two separate events in 2010, an anomalous spring and an anomalous summer, and integrating the impacts across those two events and casting it as the net effects of the RHW, rather than simply examining the net consequence of the RHW itself. Certainly the spring event is crucial to providing a picture of GPP over the growing season and this approach makes sense for the effects of the full growing season on GPP: the extent to which the spring anomaly primed, compensated, or otherwise interacted with the RHW is important. But conceptually the authors need to make clear that simply combining them does not constitute the carbon response to the RHW, for as written, the RHW impacts are presented as the net effects of two separate events, rather than just the heat wave. Given the motivation the authors lead with (i.e., that there is an inherent potential for some kind of mismatch from the atmosphere down and the biosphere up), calling the impact of the RHW the integration of two distinct events seems like an issue. Perhaps the results should be recast around the compensation effects of spring growth on total growing season GPP in the year of the Russian heat wave. I think just making this distinction clearer is important. The net impact of the RHW is not growing season GPP, which includes the anomalous spring, it's just the GPP loss during the RHW. These integrations can be seen in Tables 1 and 2, S4.1, etc. Further complicating this is the fact that the actual losses and gains of GPP are domain integrated, and the domain integration is a function of the detection algorithm. Certainly the authors discuss that the compensation occurs in a fundamentally different part of the domain and land cover class than the heat wave impacts, so I find the combination a bit misleading–it occurs in a different location and time than the actual heat wave–1TgC in crops is fundamentally different than that for forests (though from a carbon accounting perspective perhaps not). This again, is just about how the results are presented, particularly the res+/res-, not the results themselves.

Response: We would like to thank the reviewer for this comment. We apologize if the net effects of the Russian Heatwave (RHW) can be misunderstood as integrated spring and summer effect. We will carefully revise the manuscript to address this issue. Furthermore, we did not mention in the manuscript, that there is no event after summer. Thus the annual integration over the events in the growing season in 2010 equals the integration over spring and summer. We add a sentence to clarify this issue on p.9, l.6: "Please note, that we did not find extreme events after summer, which implies a fast recovery of vegetations activity after summer. Integrations over the spring

and summer events thus equals the annual integration.“ Furthermore, we will reformulate “integrated over spring and summer“ to “annually integrated“.
Reviewer #1 expressed concerns along the same lines, particularly with respect to process based connections between the spring and summer event. We will provide a more in-depth discussion abouth how the spring and summer event might be related:
First, we will add a paragraph to the introduction (p.5, l.3): “Temperature anomalies exceeded more than 10~K in both spring and summer, but they lead to distinctive anomalies in gross primary productivity (GPP). Positive GPP anomalies occurred during the spring event, whereas negative GPP anomalies are occurring during the summer heatwave. The positive GPP response in spring might be a reaction to warmer, more optimal spring temperatures (Wang et al, 2017) possibly accompanied by enough water availability. However, negative GPP anomalies in summer occurred only in areas south of 55°N (Fig. 1c) indicating that the GPP response involves much more processes than high temperatures and drought during the unique RHW. As already indicated by Smith, 2011, the connection between biosphere and hydrometeorology is much more complex than just a direct one-to-one mapping. Further complicating this issue is the fact that the summer event cannot be investigated without the previous spring, as both seasons are inherently related via memory effects in water availability. Increased GPP in spring may negatively influence soil moisture and thus GPP during summer (Buermann et al., 2013). In Summary, comparing ... “

Second, we will add a paragraph to the discussion p.12, l. 22 – p. 13, l. 10. as folows: “Another important aspect is that the combination of the anomalous spring and the unique heatwave in summer might be inherently connected via land surface feedbacks. Buermann et al., 2013 showed that warmer springs going in hand with earlier vegetation activity negatively affect soil moisture in summer. It is a general observation that warm and dry springs enhance summer temperatures during droughts, which suggests the presence of soil-moisture temperature feedbacks across seasons (Haslinger et al., 2017). In case of the Russian heatwave 2010, soil moisture was one of the main drivers (Hauser et al., 2016), in hand with persistent atmospheric pressure patterns (Miralles et al., 2014). Thus, we suspect that the spring event is connected to the summer heatwave in 2010, if not setting the preconditions for a heatwave of this unique magnitude. “

3. Merits of the detection approach: Part of the basis of this manuscript is that a much more sophisticated detection approach is needed to accurately represent the biophysical impacts of

climate extremes. If one simply did the detection–as is typical– at the grid point scale on the hydrometeorological fields and then composited on the biophysical fields for the same dates as the meteorological anomaly, would the results and/or conclusions substantially differ? At 5.10 the authors claim that for a short time series a traditional threshold approach would be problematic. Is there evidence for this? The authors still have to perform a sensitivity analysis of their results to the chosen threshold (S4.1). At some places the paper feels needlessly complex– perhaps the authors could better justify their complicated analytical choices?

Response: We would like to thank the reviewer for the critical analysis of our detection approach. The main advantage of our multivariate detection approach is that we can integrate information about several variables simultaneously and might also detect rare combinations of variables which are not detected as extreme individually (4.9-14). However, the events in Russia 2010 are not an example of this kind. Thus, we agree with the reviewer, that it is possible to get similar results by combining several univariate detection approaches for the Russian heatwave. Combining univariate detection approaches would require to choose a threshold for each variable individually. Performing a full sensitivity analysis of the choosen thresholds would lead to a combination of many possible thresholds which would render high dimensional unfeasable.

We would like to thank the reviewer pointing us to our claim at 5.10. which might be suspect to missunderstanding. Our intention was not to state that the traditional threshold approach itself is problematic. We wanted to state that the underlying assumption (equal distribution of extreme events among all grid cells) is most likely not met for short time series (here: 11 years). A 10% threshold in a 10 year time series would seelct one extreme year in each grid cell (not more, not less). There are regions where extremes events occur more often or are longer than one year, e.g., California (Griffin, D., and K. J. Anchukaitis, 2014 , GRL) or where by chance no extreme event at all is occuring in the given time frame. Our spatiotemporal segmentation is addressing this issue by choosing thresholds over larger areas with compareable climate and phenology. As a request also from Reviewer#1 we will completely revise the section 2.2 including a new schematic figure for the spatiotemporal extraction. In this process, we will rephrase the given part above, which is suspect to missunderstanding. Furthermore, we will justify complicated analytical choices as suggested by the reviewer, remove unneccessary parts (global thresholds, local thresholds), merge it with the information in S1, and avoid jargon whenever possible.

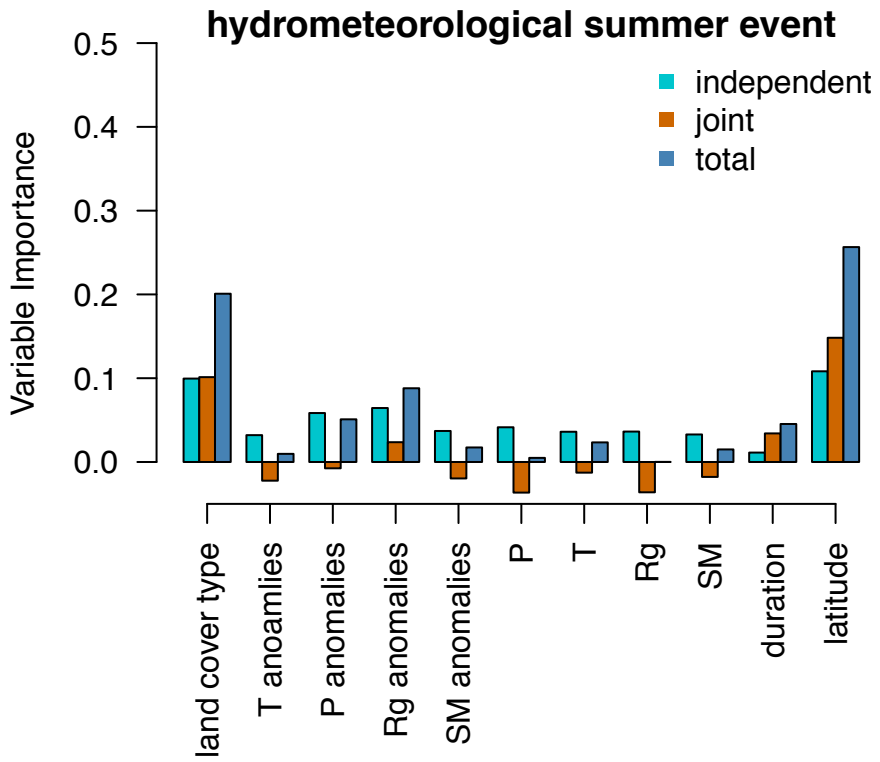
4. Model of factors explaining the GPP response. This section (S3), which is referred to in the main, but relegated to the Supplemental could be better emphasized and explained. For example, the factors in the hierarchical modeling approach are not independent. Are interaction variables used to address this issue? Given the confounding of latitude and temperature and land cover class, why not add latitude to the regression hierarchy to see its explanatory power, given the sentiment at 12.3?

Response: We would like to thank the reviewer for his interest in the Section S3 and agree that the section can be much better emphasized and explained. We will carefully revise the section and introduce more information in the main manuscript on 12.1. The factors in the hierarchical modeling approach are indeed not independent. However, the hierarchical partitioning after Chevan and Sutherland (1991), is exactly made for this kind of issues. The method extracts the independent contribution of interacting variables.

We would also like to thank the reviewer for the idea to extent the regression model with the factor latitude. Indeed, latitude has a very high independent explanatory power (Figure below) which is comparable to the importance of land cover type in summertime. The high independent explanatory power indicates that latitude provides additional information, which is not already contained in the other factors (e.g., land cover type or absolute temperatures). In particular access to deeper water (and soil type) might be factors not contained in the model, but also changing with latitude and therefore possibly explaining the importance of latitude. Apart from including more information about the method, we will include more information on S3 in the main manuscript and the new results in section 3.3 with the following paragraph and reformulate other sections if necessary:

"To disentangle the variable importance of the different confounding factors, we run a simple linear regression model which tries to explain GPP as function of the hydrometeorological driver variables (temperature, precipitation, radiation, surface moisture, anomalies and absolute values), as well as vegetation type, duration and latitude (Supplementary S3). We use an algorithm after \citet{Chevan:1991wg} which extracts the independent contribution of the variable importance related to this particular variable regardless of the model complexity or dependencies among variables.

The model reveals from a statistical point of view, that vegetation type and the latitudinal gradient are the most important variables explaining GPP during the summer event, followed by the hydrometeorological drivers. Access to deeper water and soil type as well as non-linear feedbacks are factors which are not represented in the model, but might explain the high variable importance of latitude."



5. Attribution to uWUE differences. The authors attribute the reduced GPP declines during the summer event of forests in part due to the uWUE. Certainly this has a role to play. One could also imagine uWUE being an explanatory variable in the model presented in section S3 as well–could the authors add that? It seems like the authors are positioned to better attribute whether it was the absolute magnitude of the temperature itself (which diminished as a function of latitude) or something innate to the land cover classes (and their underlying WUE), which just so happens to vary as a function of latitude. The model seems like an ideal place to disassociate these factors.

*Response: In general we like the idea to add uWUE as a explanatory variable in the model. However, uWUE is defined as GPP * VPD^0.5 / ET. Thus adding uWUE as factor to the model*

would be somehow circular, as the target variable (GPP) is contained in the possible factor uWUE. Thus, we think adding uWUE would be inappropriate from a statistical point of view.

Regarding the spring event and soil moisture depletion carry-over effects under forcing discussed at 12.22-13.6, Mankin et al., Journal of Climate 2017 and Mankin et al. GRL 2018 note that increased productivity is associated with such carry over effects in some of the models, regionally and globally under forcing.

We would like to thank the reviewer for the two additional references which we found very interesting. We will add the references to the discussion about soil moisture carry over effects.

Minor comments:

11.5: I don't understand the soil moisture in Fig. 7. Is it the normalized measure? Is it the m3/m3? Can the authors add contours if the forests separate by latitude in 7b?

Response: It is m3/m3.

Grammar/spelling throughout could be improved.

Response: We will go through the text once again. We also highly appreciate that Biogeosciences now performs a carful language check previous to publication.

1.16: (e.g., a vegetation index)

Response: The sentence will be removed to address major comment 1)

inconsistency in comma usage after e.g. and i.e.

Response: Will be checked for consistency

2.29: not sure the name is "heat summer"

Response: changed into "European heatwave 2003"

2.32: a, not an hydrometeorological

Response: Will be changed

5.8: grammar (“in high”)

Response: Will be corrected

5.4: Why not leave them as missing data?

Response: We would like to thank the reviewer for this comment. Indeed, it would be an option to leave the data as missing in case all variables are missing at one observation, excluding the observation from the multivariate detection. However, in comparison to univariate event detection, our multivariate algorithm requires that all variables are available. Thus, there are many more missing instances, i.e., cases which have only one of the variables missing, all others are available. We will add the following sentence on that:

“The gap filling is necessary for a multivariate detection approach as there are many more cases in which one variable is missing in the multivariate cube compared to a univariate data stream.“

21.32 “spatiotemporal” not “. . .temporla”

Response: Will be corrected

Author contributions: “wrote” not “ote"

Response: Will be corrected