Dear Prof. Luyssaert,

Thank you very much for considering our manuscript for publication in Biogeosciences. We have now carefully revised our manuscript in light of the reviews. In our revision we have:

1) Added additional analyses to the supplementary which only compare regions between 2003 and 2018 that were affected by extreme or moderate drought in both years (Figs. S11-13 as well as additional corresponding statements in methods, results, discussion and conclusions)
2) Included several statements as requested by the reviewers.

Please find a detailed point by point reply to the reviewers' comments below.

Altogether, we incorporated all of the amendments requested by the reviewers and hope that our manuscript now meets the requirements for publication in Biogeosciences.

With my best regard on behalf of all authors

Allan Buras

Report #2

This study compares climatological data and vegetation indices between the 2003 and 2018 drought. It concludes that the 2018 drought response was more extreme i.a. due to a difference in the location and timing of the drought. The manuscript is well written and structured. I think the concerns of the reviewers in particular about the temporal and spatial differences between the two events have been adequately addressed.

Reply: Thank you very much for the positive evaluation of our study. We highly appreciate the time you have spent to review our paper.

I only have a few very minor comments:

Line 159: Shapiro-Wilk?

Reply: Yes, we have rephrased to Shapiro-Wilk.

Line 225: 0.61% should be 61% I guess?

Reply: Thank you for spotting this error. We have corrected to 61 %.

Line 495: '...we observed a different sensitivity of ecosystems to CWB between the two events...' In light of the discussion in 4.2.2, I think this should be rephrased. It is not necessarily the sensitivity of the ecosystems that is different but rather the location and timing of the events subjecting different ecosystems to drought, i.e. the 2018 event hit more sensitive ecosystems earlier in the growing season.

Reply: Here, we use the term sensitivity as in Anderegg et al. (2018) "Hydraulic diversity of forests regulates ecosystem resilience during drought" (Nature), where regression slopes are interpreted as sensitivity (and r^2 as coupling). To account for your comment, we have added the following statement to the corresponding sentence:

More specifically, we observed a different sensitivity of ecosystems to CWB between the two events and a differing sensitivity of land cover classes to drought, with pastures and agricultural fields expressing a higher sensitivity in comparison to forests, which probably was caused by the differing spatial extent of the two events thereby affecting presumably less drought-resistant ecosystems in 2018.

Report #3

The authors have done a good job revising the manuscript. Below I list some (minor) comments and suggestions to further improve the manuscript and clarify a few points.

Reply: Thank you very much for the positive evaluation of our manuscript. We greatly appreciate the time you have spent to review our study.

Abstract:
Abbreviations not explained (NDVI, EVI, VI). Especially VI may need clarification for readers who are not familiar with remote sensing terminology.

Reply: We now also explain the abbreviations in the abstract (L 15).

Introduction:
l. 53: this sentence may need rephrasing. I suggest writing "… heatwave of 2003 was long considered the most extreme compound event… " because it doesn't come as a big surprise that 2018 was worse (and perhaps so was 2019?)

Reply: We have amended the sentence accordingly.

Materials and Methods:

l. 84-86: not totally clear. Did you use April-July for 2003 and May-August for 2018? Please clarify and motivate this choice better. Why not the same period (e.g. April-August) for both years?

Reply: The selection of these two periods was chosen on the basis of climate data, which indicated the peak of the 2003 drought to occur in August, while in 2018 it occurred in July. In the first round of reviews we were asked to use those two periods to have a better comparison between the corresponding peaks of the two drought events. We have now added the explanation for this selection in this section and also refer to section 2.2 where this motivation is explained in detail.

These particular, differing periods were chosen, since they each represent the peak of drought for the corresponding year (see also section 2.2).

l. 198: I suppose this should be Fig. S6 instead of Fig. S5 (which does not show land cover types)? Please also clarify the abbreviations used in Fig. S6 (and other supplementary figures).

Reply: Thank you for spotting this error. We now refer to S6 and have clarified the abbreviations in the supplementary figures.

Results:

Fig. 1: Legend mentions April-July. Should this be August (2003) and July (2018)? Or May-August and April-July? Similar remark for Fig. 3 (while in Fig. 2, the legend has been revised).

Reply: Thank you for spotting this error. We have corrected the figure captions.

Fig. 5: these results are of course strongly determined by the spatial differences in the heatwave of 2018 vs 2003. I would be interested to see the NDVI analysis for Central Europe only, where the heatwave was similarly extreme in both years. In other words, excluding N and S Europe, where the weather was very different for both years. This could give some extra substance to section 4.2.2.

Reply: Thank you for this suggestion. We have now added supplementary Figs. S11-S13 which show similar histograms as in Fig. 5 but now only for the intersection of areas featuring extreme drought (CWB anomaly < -2) in both years as well as the intersection of areas featuring moderately dry conditions (-2 < CWB < 0). Since there was hardly any overlap between pixels featuring positive CWB anomalies, we refrained from depicting these comparisons, too. The resulting histograms underline the previously acquired findings that the drought of 2018 superseded the one of 2003. In particular, forest ecosystems featured a higher share of lower NDVI (and EVI for moderate drought areas) quantiles, which we interpret as a consequence of legacies from drought events before 2015. We have added corresponding statements in the methods, results, discussion, and conclusions:

Methods, section 2.2.:

Due to the different spatial patterns in drought severity between the two years, we determined the intersection area where the same CWB anomaly classes were observed in both 2003 and 2018. For this intersection area, we repeated the comparison of the 19 different quantiles between 2003 and 2018 for the five different land-cover by comparing the same pixels for each combination of land-cover class and CWB anomaly class between the two years. This was done, to avoid artefacts related to the fact that CWB anomaly classes were represented by different ecosystems in 2003 compared to 2018. Since the overlap for positive CWB anomalies was very low (altogether only 455 MODIS pixels, and thus no observation for some of the quantiles in some of the land cover classes), we refrained from computing those given their low representativity. Consequently, we only depict the comparison for extreme (CWB < -2) and moderate (-2 < CWB < 0) water deficit. Because of similar areas in both years, we refrained from depicting proportional areas as for the full comparison.

Results:

This observation was confirmed when considering only pixels with extreme (CWB anomaly < -2) or moderate (-2 < CWB < 0) water deficit (Figs. S11-S13) in 2003 and 2018. Although the differences of absolute areas decreased in this comparison, 2018 generally displayed larger areas with lowest quantiles compared to 2003. However, for EVI, subtle differences of opposite sign were observed for regions featuring extreme water deficit, while regions with moderate water deficit expressed similar patterns as for NDVI (Fig. S13 compared to Fig. S12).

Discussion, section 4.2.2

As a first step into this direction, we only compared VI-quantiles of regions that featured extreme or moderate water deficit in both years, thus only considering regions representative of the exactly same ecosystems. On average, we again observed a higher share of low quantiles in 2018 compared to 2003 (Figs. S12 and S13). Only for the EVI in regions featuring extreme drought, 2003 featured a slightly higher share of lower quantiles compared to 2018. Taken together, it

nevertheless seems that even when only considering the same regions for the comparison between both years, the impact of the 2018-drought supersedes the one of 2003. Interestingly, clear differences were only observed for forest ecosystems. This might indicate so-called drought legacy effects (see also section 4.2.4) as a consequence of preceding extreme droughts, such as the one of 2015 after which an increased forest mortality and growth decline was observed in southern Germany and other parts of Central Europe (Buras et al., 2018).

Conclusion:

Finally, in addition to quantifying impacts of the drought 2018 on European ecosystems our results possibly mirror forest drought legacies from preceding drought events. Moreover, additional legacy effects of forest ecosystems are likely to occur in course of the next years.

Fig. 7: include statistics to indicate whether the difference between both years differed between the vegetation types.

Reply: Since we only have one difference for each vegetation type, we struggle to provide a statistic that allows for testing whether the difference between the two years differ among the vegetation types. We have therefore just added a qualitative statement in the results section.

In comparison, the differences between 2003 and 2018 were highest for pastures, followed by arable land, broadleaved forests, mixed forests, and coniferous forests.

Discussion:
l. 400: should this be "5 (±1) percent"?

Reply: This statement refers to the percentages of pixels that featured a positive CWB while the CWB anomaly was below 0 (5 percent) or below -2 (1 percent). These explanations are provided in the caption of Fig. S13 which is referred to in the same sentence. To clarify, we have added:

for details see Fig. S13.

l.407ff: it could also be mentioned in this paragraph that choice of VI can influence differences in responses between vegetation types (e.g. when some indices like NDVI detect drought better in grasslands than in forests, and vice versa for other VIs).

Reply: We agree and have added the following sentence to this paragraph:

Consequently, the choice of VI may also affect the differences in the responses of different vegetation types.

l.450 ff: choice of VI may also matter here (see previous remark). You find a similar response in EVI, supporting your statements. This is worth mentioning here.

Reply: We have added the following statement to the corresponding paragraph:

As mentioned in section 4.2.1 the choice of VI may alter the differences in the responses of different vegetation types to drought. However, the fact that NDVI and EVI indicated differences of similar sign and magnitude among the considered ecosystems, supports our interpretation. Nevertheless, further remote sensing products such as the solar-induced fluorescence may provide additional information on ecosystem-specific drought responses.