**Review of "Increased vulnerability of European ecosystems to two compound dry and hot summers in 2018 and 2019" by Bastos et al.**

This work focuses on understanding the central European ecosystem response to two consecutive extreme dry and hot summers in 2018 and 2019. The authors combine different observation/reanalysis datasets and both statistical and dynamic vegetation modelling approaches to comprehensively evaluate the impacts of these compound events. Based on different responses to DH2018 and DH2019, 4 categories of ecosystems are identified, of which two show stronger negative effects of the recent DH events on vegetation compared to previous climate-vegetation linear relationships. One of the two classes shows increased ecosystem vulnerability to compound DH events possibly modulated by long-term warming and increased pCO2, and is characterized by grasslands and crops. The second class shows continuing browning from DH2018 to DH2019 even with alleviation of soil moisture deficits and heat stress in DH2019, mainly attributed to legacy effects of DH2018, and is characterized by forests and grasslands. These legacy effects, however, are missing in the current generation land-surface models, suggested by factorial simulations using an ensemble of these models. This may shed doubt on the ability of model skills to predict the resilience of temperature ecosystems to more likely compound events in the future.

The manuscript is clearly written and easy to understand its message, and the results will be of great interest to a wide community of climate and ecosystem researchers. However, I suggest a few points be addressed before it is considered for publication, as follows.

My biggest concern is that the "increased vulnerability" can be exaggerated. First, the authors only focused on the area with negative EVIanom in DH2018 (area_clusters), while the rest area (white color pixels in Fig. 3) was not considered, including core regions such as scattered parts in Germany and Poland, and large assembled parts in southern Sweden and the southeast of the study region. Based on the climate-vegetation relationship estimated from this area_clusters, the 2018 and/or 2019 EVIanom of some ecosystem clusters appears to be deviated from the long-term relationships. However, including the rest area can both affect the departures of 2018/2019 EVIanom as well as the long-term relationship. First, the slopes (EVIanom-SMAnom slope and EVIanom-Tanom slope) can be larger as modulated by vegetation susceptibility in the rest area during 2001-2017, for instance, the southeastern part (southern Sweden) during 2003 (2015) heat wave. Second, including 2018 EVIanom in the rest area can surely offset some of the 2018 departures since it is a positive EVI anomaly. Therefore, in my opinion the approach to just look at DH2018 EVIanom-negative region in this manuscript does not allow to conclude the increased vulnerability of European ecosystems to the compound events (also given the fact that it only covers 20% of the area_clusters).

Soil moisture anomaly is a useful indicator for the climate impacts of DH events. I feel the correlation is not enough to assess the model skills in simulating soil moisture variability (Table 1). First, the record can be too short for a robust correlation (Apr-Sep, sample size=6?), and this can be an issue not only for soil moisture validation, but also for GPP (e.g. L310). Second, even if the correlation is high, the magnitude of the changes in the absolute soil moisture values can be underestimated in the model due to too shallow soil. Therefore, calculating RMSE or normalized-RMSE using the absolute soil moisture values can be useful.

**Other comments**

L16 Should it be dominated by forests and grasslands? It is inconsistent with results shown in L240.

L179&L289 Please add a few lines to justify the use of the RF regression. What's the advantage compared to a linear multivariate regression?

L181 Please explain more on the 3x3 and 17x9 used here.

Sec 3.2.2 & Sec 3.2.3 There are some repetitions in these two sections, such as the definition of residuals.

L245 Is it a correct statement? It seems like a few exceptions can be there (SManom and Tanom of cluster C4).

L260-261 This may be inaccurate. See my major comment.

Fig. 4 Anomalies during EVIanom positive years and EVIanom negative years are supposed to be comparable (add up to 0 eventually). There is a trick that 2018&2019 were not included in this long-term relationship. If they were, "abnormal" positive values could show up.

L266 How can different out of bag scores affect explained variability so much?

L291-293 Please rephrase. The improved predictivity of RF seems to be contradictory to the comparable residuals.

Fig.6 The x-axis labels are hard to read.

Sec 5.2 Is the simulation of 2018 productivity anomalies really so well? At least during the spring season precondition it is not so consistent between model and data.

Fig. A5 I cannot find anywhere in the text that this figure is discussed.

**Technical comments**

L3 though -> through
L72 Modify the citation.
L180 Double-check the variables used here.
L226 excepting -> except
L285 Add . before "In DH2019"
L301 ,since -> , since
L310 GPP should be GPPanom?
L323 EVIanom, subscript anom
L388 related -> be related