The manuscript highlights the important correlation between the satellite-derived fAPAR product and observed cereal yield in Europe, clearly positioning the study and its results in drought impact monitoring efforts. Therefore, this work fits the journal’s scope and can be seen as a timely contribution to the growing early warning and climate services research. In particular, the unique scope of the paper, looking at longer, homogeneous periods with robust, significant correlations makes for an interesting, novel article that is written well.

While I enjoyed reading it, some part could benefit for extra clarifications. Besides, I would like to propose a few comments/ideas which could maybe be considered as additional discussion points to strengthen and broaden the manuscript.

- I wonder if the Corine land cover map arable land area is similar to the Eurostat cereal production area data for each NUTS2. It would strengthen the method if this is the case.
- Could you please better justify why you would take insignificant (but “at least different than zero”) values into account (doesn’t this reduce precision of the analysis?) and why the particular threshold of 0.15 was chosen?
- I do not fully understand the following part in the method (L182): “Starting with a minimum length of 2 dekads, up to 990 periods of various length (L, from 2 up to 45 dekads) can be analyzed for each region, and for each of these periods four main metrics are computed”. In the rest of the manuscript, it seems the analysis is done on a 1 dekad level (then only combining dekads based on their correlation, not for calculating the correlation).
- It would strengthen the full analysis to check for sensitivity regarding the chosen -1 thresholds for the drought conditions (L202 and onwards). While not a necessary addition, the paper mainly focusses on its use in drought monitoring systems hence it would be interesting if a similar result would be obtained with other standard deviations as thresholds.
Another sensitivity that could be considered to be evaluated is the detrending method.

- Besides, in this regard (L205), I wonder why drought years are defined as yield anomaly years (these are agricultural impact years, maybe caused by droughts but potentially by other shocks). The authors could consider reversing the analysis, looking at the average yield anomaly during years with a low average fAPAR during the optimal period. This would not guarantee excluding other shocks (that might impact fAPAR too) but would be more interesting in terms of its capacity to be used as an impact monitoring of prediction product (also give insight on the FA rate for example). (Similar remark could be made for Figure 2: it shows that indeed, during large drought episodes in Europe, the fAPAR is low, but does not show anything about potentially low fAPAR values during years not considered droughts)
- I wonder what happened if two periods with Fp+ = 1 are equally long? Could you please explain how this is handled in the analysis?
- In the method, multiple exclusion criteria (related to fAPAR and EUROSTAT data) are stated, however the results show a full map without data gaps. Does that mean no NUTS2 were excluded based on these criteria?
- In figure 4; it is a pity no spatial signal could be visualized. It would be a great addition to show which region of Europe contributes to what here.
- Technical remark: “It is possible to observed two “flexing points” …” à OBSERVE
- I feel the method behind figure 7 could be explained better. Could the authors please reflect on this choice for the bounded length of 6 to 8 months (rather than only the optimal, correlated period)?
- I wonder if the overall limited correlation in central Europe might be caused by changing crop types over the years / fAPAR is calculated based on the full period thus assuming homogeneity in land cover over this period.
- I would like for the authors to better explain figure 8: what is meant with performance? Here, how are Fp+/ F+ calculated? Based on the average fAPAR over the optimised period? I think this is missing in the method.
- In the discussion (eg L363), it seem the authors equate growing season with the season where fAPAR relates to the yield; but can this be supported by agronomic observations? Is this correlation not a sign of a crop growing period vulnerable to droughts rather than a representation of the full season?
- In the discussion (L396-400); the interpretation is not extremely clear. So the inverse relationship between fAPAR and yield is a result of hot-dry summer months (dry spell months or all months?) and their lagged effect: how? Is this correct(ly interpreted?)

With kind regards