REVIEW #2

The author conducted attribution analysis of a soil moisture drought event in West-Central Europe in 2022 based on a large number of soil moisture estimates. It is found that the return period of the drought event is decreased substantially in current climate as compared with that in pre-industrial climate, although there are uncertainties in soil moisture estimates. The manuscript is well written and easy to follow, and there are only a few minor comments below.

We appreciate the positive assessment of our work and the questions listed below, to which we reply individually in blue font.

1. GLDAS_ CLSM data includes v2.0 (1948-2014) and v2.1 (2000-2022). Is v2.0 and v2.1 used together in the article? How are they merged? Have corrections been made before merging? Only the June data was used in 2022? How about July and August?

   Thank you for this comment; in short, we have repeated much of our analysis for an updated GLDAS-CLSM dataset, and explain the merging procedure in the following. We use GLDAS-CLSM data forced with Princeton v2.0 meteorological forcing and without data assimilation (i.e., open loop, available for 1948–2012), and then ECMWF-analysis forced CLSM output obtained with GRACE data assimilation from 2003 onwards. The latter is scaled to Princeton-forced data based on 2003–2012 (that is, mean and standard deviation are matched), employing a 7-day moving window. Since GRACE data was not available for July and August 2022 when we performed the initial analysis, CLSM was originally run in open-loop mode for those months instead (but still driven by the ECMWF analysis fields).

   We have since updated our CLSM dataset, such that the entire summer of 2022 is now based on CLSM with GRACE data assimilation. For both WCE and NHET, the updated mean June–August 2022 soil moisture is slightly higher than the initial version without GRACE data assimilation for July and August 2022, but the differences are so small that they are not visible in most figures. A noteworthy exception concerns the event return periods, which remain nearly unchanged for NHET (root-zone: 7 years→ 6 years, surface: 17 years → 14 years), but decrease a bit for WCE (root-zone: 33 years → 23 years, surface: 47 years → 33 years), consistent with the slightly higher summer 2022 soil moisture (i.e., a bit less dry and hence less extreme). Since this only results in minor changes of the average 2022 soil moisture drought return periods (always based on GLDAS-CLSM and ERA5-Land, and additionally ERA5 for surface soil moisture), we keep a return period of 20 years for the models in all our soil moisture analyses.

Even though the probability ratios of GLDAS-CLSM decreased, they are still (more than) high enough to be limited to 10'000 (as described in the Methods). Consequently, the PRs in the present-vs-past synthesis figures remain unchanged. The more robust intensity changes of GLDAS-CLSM differ only slightly from our initial analysis, such that the synthesised values barely change compared to our first analysis.

We updated all the text, figures and tables that relate to or make use of GLDAS-CLSM data, and edited the description of GLDAS-CLSM to hopefully make the scaling procedure more clear:

The 2003 to present data is scaled ~~climatology is sealed~~ to the open loop, using scaling factors determined for each grid cell and with a 7-day moving window such that the mean and standard deviation of soil moisture obtained with GRACE data assimilation across 2003–2012 matches the Princeton-forced (1948-2012) climatology in the same period ~~prior to computing drought/wetness indicators~~ (Houborg et al., 2012). For some grid cells in high latitudes, this results in negative and hence not physically meaningful values, which we remove for our analysis of the northern extratropics. ~~Note that the GLDAS-CLSM dataset employed here was produced with GRACE/GRACE-FO data up to and including June 2022. We do not expect this to affect our conclusions, since (i) the attribution statements ultimately depend on the linear trend in soil moisture as a function of global warming for the entire period of 1950-2022, and (ii) the return periods of the 2022 event are similar in ERA5-Land and GLDAS-CLSM.~~

2. Why does CMIP6 use SSP585 and historical scene splicing (usually SSP245), while CMIP5 uses RCP4.5 and historical splicing?

We appreciate the reviewer's question and remark that the choice of future scenario is not as important as it may seem in our analysis framework, although this was not clearly described in the initial manuscript. We employ global warming levels, -1.2 and +0.8 °C with respect to 2022 in our analysis. Therefore, it is not problematic if we fit our statistical model for a GMST range slightly larger than what we ultimately need for our evaluation, which is in fact even desirable (unlike the opposite case, i.e. fitting the statistical model for say -1.2 to only +0.1 °C, and then querying the fitted model at +0.8 °C).

We have edited Sect. 3.3 as follows to make our approach more clear:

We note that regardless of the underlying emission scenario, model data from 1850 to 2022 and from 1850 to 2050 are used to conduct the present-vs-past and future-vs-present climate analyses, but these time periods only indicate the amount of data used to fit the statistical model and hence infer the relationship between event indices and GMSTs. We then rely on global warming levels ~~We~~to calculate the return periods, the probability ratio (PR — the factor-change in the event's probability) and change in intensity of the drought event.~~ in order to compare the climate of now and the climate of the past, defined respectively by the GMST values of now and of the preindustrial past (1850–1900, based on the Global Warming Index https://www.globalwarmingindex.org)~~. For our comparison of the present (2022) to the past (1850–1900) climate, the GMST changes with respect to the present amount to -1.2 ° C according to the Global Warming Index https://www.globalwarmingindex.org), and for comparing additional changes in the future to the present, we use +0.8 ° C relative to the 2022 GMSTs (+2.0 ° C with respect to pre-industrial conditions). As such, it does not matter when the future warming is reached in any given model simulation, which allows us to combine models with different emission scenarios and still perform a consistent analysis.

3. The uncertainty range for the surface and rootzone soil moisture anomalies (Figures 1a and 3a) could be shown in the supplement materials.

Only GLDAS-CLSM is a good candidate to supplement Fig. 1a (global 1950–2022 data available), and the anomaly patterns are generally similar to ERA5-Land (used for Fig. 1a). We provide a comparison below for root-zone soil moisture (cf. **Fig. R5 & R6**):

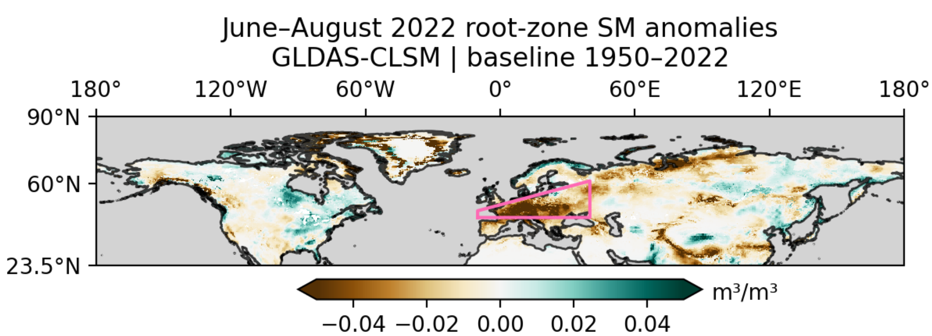

Fig. R5: 2022 summer root-zone soil moisture anomalies based on GLDAS-CLSM and using 1950–2022 as a baseline.

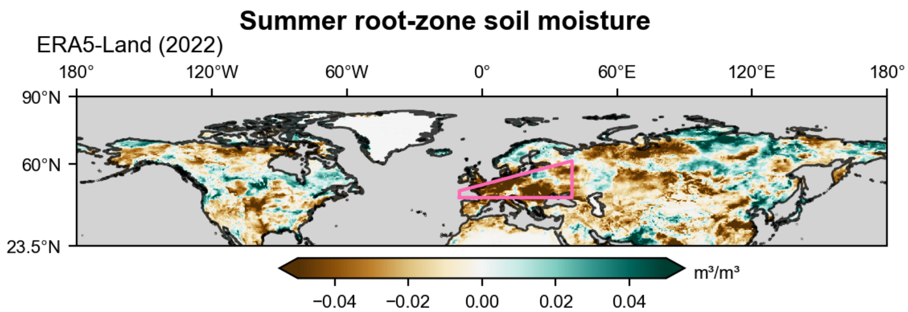

Fig. R6: Fig. 1a from the manuscript; like Fig. R5, but using ERA5-Land.

Besides West-Central Europe, where both datasets point to widespread negative anomalies, drought conditions were also reported by other regions, such as the central and southwestern United States or China. Both ERA5-Land and GLDAS-CLSM indicate soil moisture deficits in these areas.

We also show GLDAS-CLSM based surface soil moisture anomalies below (**Fig. R7**), which also does not agree with ERA5-Land everywhere, but certainly supports the main message of Figs. 1a and 3a: West-Central Europe experienced a strong soil drought in the summer of 2022, and despite positive soil moisture anomalies in regions such as northeastern North America or parts of Siberia, soil moisture deficits also dominated the northern extratropics as a whole.

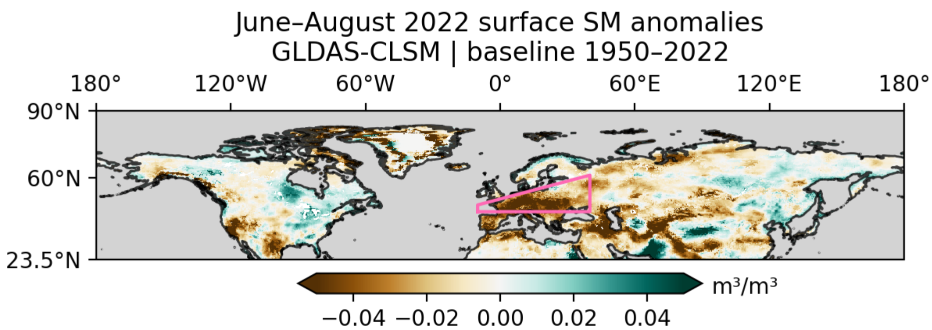

Fig. R7: Fig. 1a from the manuscript; like Fig. R6, but for surface soil moisture.

We believe that the different timeseries in Figs. 1b–c and 3b–c are more efficient to convey the “observational” uncertainties, since they make use of several (and not just two) datasets, and additionally show that the uncertainty decreases in time.

4. What is the definition of Intensity in this study?

The intensity corresponds to the mean event magnitude, which changes as our climate warms. In other words, climate change affects the expected magnitude of an event with any given return period (which is a different but equally viable perspective as considering how the probability of an event with a given magnitude changes). Technically, the intensity is represented by the location parameter of the scaling and shifting Gaussian distributions that we employ, which in turn is modelled as a function of global mean temperature changes, and has the same units as the index of interest (e.g., summer mean root-zone soil moisture).

We use magnitude and intensity synonymously throughout the text, and have edited the first occurrence as follows to hopefully make it more clear:

The essence of the approach we employ here is that event indices — regional summertime averages of soil moisture, precipitation, temperature — are represented with continuous probability distributions conditional on GMST, which enables us to estimate how the intensity (event magnitude) and probability of occurrence have changed under human-induced climate change.

5. Line 479: Does -9% (-13%.. -4%) refer to a change in intensity? Please explain in detail.

   Yes, this does refer to a change in intensity. As stated in our response above, the intensity or event magnitude has the same units as the respective index. We communicate changes in intensity in °C for temperature, and use percentages for precipitation and soil moisture. This is motivated by the fact that the implications of absolute temperature changes are intuitive, whereas relative changes can easily be interpreted even without knowledge of the respective total (or, in this case, past climate state) such as, e.g., the mean volumetric root-zone soil moisture of the northern extratropics.

   We edited the text to make it more clear that we do indeed refer to a change in intensity:

   We also estimate the mean change in WCE summer root-zone soil moisture from the past to the present climate, ~~and obtain~~which yields intensity changes with best estimates (confidence intervals) of -9% (-13% .. -4%) for ERA5-Land and -14% (-16% .. -11%) for GLDAS-CLSM.

6. What is the physical meaning of representation error? Perhaps only when the minimum value of Representation error is significantly greater than 1 (less than 0) can the impact of anthropogenic climate change on Probability ratio (change in intensity) be considered significant?

In the WWA framework, observations — or in the case of this study, observation-derived estimates (of soil moisture) — are considered as equally valid representations of the same climate realisation (i.e., reality), and hence sample the same (true) underlying natural variability. If the observations were perfect, we would expect them to indicate identical best estimates, but this is clearly not the case. Therefore, the mean discrepancy of individual best estimates of observations to the observational mean best estimate serves as a "representation" error.

In the context of the last question (8.), we would like to emphasise here that the observation-derived soil moisture estimates are associated with a large representation error compared to temperature or precipitation (cf. Fig 6 or 7 and Fig. S7, S8). This error is primarily large not because the datasets do not agree on the presence of a clear warming footprint in soil moisture, but rather because they either point to a signal consistent with the models (ERA5-Land) or an even stronger one (GLDAS-CLSM).

We thank the reviewer for their question and think that our revised text pasted below makes the approach more clear:

To combine the two lines of evidence into a synthesised assessment, first, a representation error is added (in quadrature) to the observations~~, to account for the difference between observations-based datasets that cannot be explained by natural variability (light blue bars)~~. ~~This is~~ The rationale behind this is that we consider observations as equally valid representations of a singular climate realisation with the same underlying true natural (internally generated) variability. Therefore, the mean deviation of individual datasets to the overall mean best estimate indicates a representation error (of observations with respect to reality), shown in the synthesis figures as white boxes around the natural — that is, internally generated — variability (light blue bars). The dark blue bar shows the average over the observation-based products (black marker) and the total uncertainty (width of the bar) based on natural variability and representation errors. Instead of representation errors, next, a term to account for intermodel spread is added (in quadrature) to the natural variability of the models. Note that while this term is based on the scatter of model means (analogous to the representation error for observations), we interpret model simulations as independent climate realisations. Consequently, we only add this term granted that the differences between models cannot solely be explained by natural variability, which is the case here. ~~This~~The intermodel spread is shown in the synthesis figures as white boxes around the light red bars.

7. There are a total of 25 models, but only 7 models in Figures 6-7 have passed the test. How about using all models? This may increase the reliability of the results.

Thank you for the suggestion. This study closely follows the WWA approach (Philip et al., 2020) that has been used for numerous rapid attribution and also many peer-reviewed publications. In this detailed protocol, the model validation and subsequent restriction to validated models is pivotal to ensure the main goal of extreme event attribution to human-induced climate change using both observations and climate models: to make the analysis more robust, as climate models have the advantage of providing alternative realisations of climate variability that can supplement observational records (e.g., Perkins & Fischer, 2013). However, since models are necessarily stark simplifications of our complex climate system and their performance depends both on the variable and region of interest, considering models without validation could decrease rather than increase the reliability of an attribution study.

We thus think that our focus should remain on synthesised estimates of probability ratios and changes in intensity. Nevertheless, we already provide a statement in the main text that is based on all (CMIP6) models:

"Finally, we remark that among the 25 available CMIP6 models used here (of which 7 passed the validation), all agree that based on 1950–2022, the best estimate of the probability ratio is at least 1, and oftentimes on the order of 10 or higher."

8. In Figures 6 and 7, the attribution results of the probability ratio are not significantly greater than 1 (the minimum value of the confidence interval is less than 1), which does not seem to suggest that the impact of anthropogenic climate change on drought in 2022 is significant. Please explain.

The probability ratios shown in Figs. 6 and 7 suggest that human influence has most likely increased the probability of the event, but the reviewer is right in that we cannot entirely exclude the possibility that this influence is relatively weak or even renders the event slightly less probable. However, we think that other factors are also relevant here. Firstly, the "observational" uncertainty, which contributes to the synthesised uncertainty together with the model uncertainty (see also response to question 6), is primarily high because the different datasets indicate either a moderate (ERA5-Land, ERA5) or strong (GLDAS-CLSM) warming imprint, and not due to, e.g., opposing signals.

Moreover, the lower bound of the more robust intensity changes is clearly > 0 for Fig. 7 (1950 onwards for models, as for observations), and only very slightly < 0 for Fig. 6.

Furthermore, widespread (summer) soil drying is also clearly expected due to our changing climate: mainly because strong warming increases evaporation which, unless counteracted by changes in precipitation [and/or runoff], causes soil desiccation. Particularly the dependence of soil moisture on precipitation, characterised by large internal variability, greatly increases the uncertainty, whereas the temperature signal is comparatively obvious.

We believe that we already clearly communicate the high level of uncertainty compared to more standard attribution analyses of, e.g., heatwaves. To quote the other reviewer:

> *"The authors are at pains to emphasise the sources of uncertainty in the analysis, not least of which arise from the observations or observations-based products. This uncertainty is explored at length and conclusions have been well phrased in light of this. The authors provide a convincing argument that robust yet conservative conclusions about the change in soil drought can be made and highlight that with lower confidence stronger statements are possible."*

For all the reasons outlined above, we think that our conclusions are not nullified by the fact that the results for WCE point to a low, but non-zero probability that the anthropogenic impact did/does not exacerbate the event's probability of occurrence, and are hence adequate.

References

Perkins, S. E. and Fischer, E. M. (2013): The usefulness of different realizations for the model evaluation of regional trends in heat waves. Geophys. Res. Lett. 40, 5793-5797. doi:10.1002/2013GL057833

Philip, S., Kew, S., van Oldenborgh, G. J., Otto, F., Vautard, R., van der Wiel, K., King, A., Lott, F., Arrighi, J., Singh, R., and van Aalst, M. (2020): A protocol for probabilistic extreme event attribution analyses, *Adv. Stat. Clim. Meteorol. Oceanogr.* 6, 177–203. doi:10.5194/ascmo-6-177-2020