Added value of seasonal hindcasts for UK hydrological drought outlook
Response to reviewers

Reviewer #2

Review of "Added value of seasonal hindcasts for UK hydrological drought" by Chan et al. This paper describes the use of storylines to assess worst-case scenarios of continuation of drought conditions into the next year, using 2022/2023 as an example. Meteorological hindcasts are from the SEAS5 system together with the GR6J hydrological model in the Greater Anglian region. The storylines uses clustering of the meteorological forcings using large-scale indices. The study highlights a useful tool to assess the risk of persisting drought over the winter period, and I recommend that it is published following a minor revision. However, I think the paper would eb strengthened by a more rigorous discussion on the skill and reliability of seasonal forecasts in Europe.

Response: We thank the reviewer for their positive comments about the manuscript. Please find below our responses to each of the comments.

Major comments

Uncertainty in seasonal forecasts. Seasonal forecasts are inherently uncertain, especially across Europe. The skill of the forecasts are not great on lead times longer than a month, and beyond that it is very questionable how much use the forecasts have. I think this needs to be discussed more in the paper

Verification. The authors could have used for example ERA5 reanalysis for verification of the seasonal forecasts. This could also have shed some light on the reliability of the forecasts, and through that have assessed the reliability and usefulness of the methodology

Response: Thank you for raising this point. We agree it is important to comment on the reliability of the SEAS5 hindcasts. This will be made much clearer in the revised version, also in response to comments from reviewer #3, as it appears that the objectives of our study were not well conveyed. However, we would like to point out that this paper is not about forecast verification and we are not attempting to predict the 2022 drought or assess the skill of the forecasts. Instead, we are using the SEAS5 hindcasts over all years, not just 2022, to explore "what-if" questions – i.e. how bad the 2022 drought could have been if followed by a dry winter characterized by a specific combination of atmospheric circulation indices. The reviewer is entirely correct in pointing out the limited skill of sub-seasonal and seasonal forecasts of drought, especially across Europe and particularly for the UK as shown by previous studies, such as Richardson et al. (2020). This underscores the very rationale behind our proposed methodology, which serves to complement traditional hydrological forecasting approaches. In light of the unreliability of forecasts, it becomes imperative for water managers to grasp the potential implications of a plausible worst-case scenario to make informed decisions.

We propose to add the following figure to the revised manuscript. This follows the inclusion of similar figures in Thompson et al. (2017) and Kelder et al. (2020) to illustrate the reliability of large ensemble hindcast datasets (treated not as forecasts, but as plausible outcomes). Figure 1a shows observed total winter rainfall over the 1983-2016 period from the CEH-GEAR dataset (red dots) and the spread of modelled rainfall from the hindcasts across all ensemble members and lead times for all years over the same period (boxes). The observational data points fall within the box plots for roughly one-half of the years, and fall within the maximum and minimum in all years. Figure 1b shows the distribution of observed winter rainfall for each year standardised against the modelled distribution of winter rainfall for the same year. This shows that for 22 out of the 38 years, the observed value for a particular year is within 1 standard deviation of the modelled rainfall of that year. We hope the reviewer would agree that this figure helps show that the SEAS5 hindcasts are reliable and useful over the selected period. We will expand the methods section to outline the reliability of the SEAS5 hindcasts in the revised manuscript.

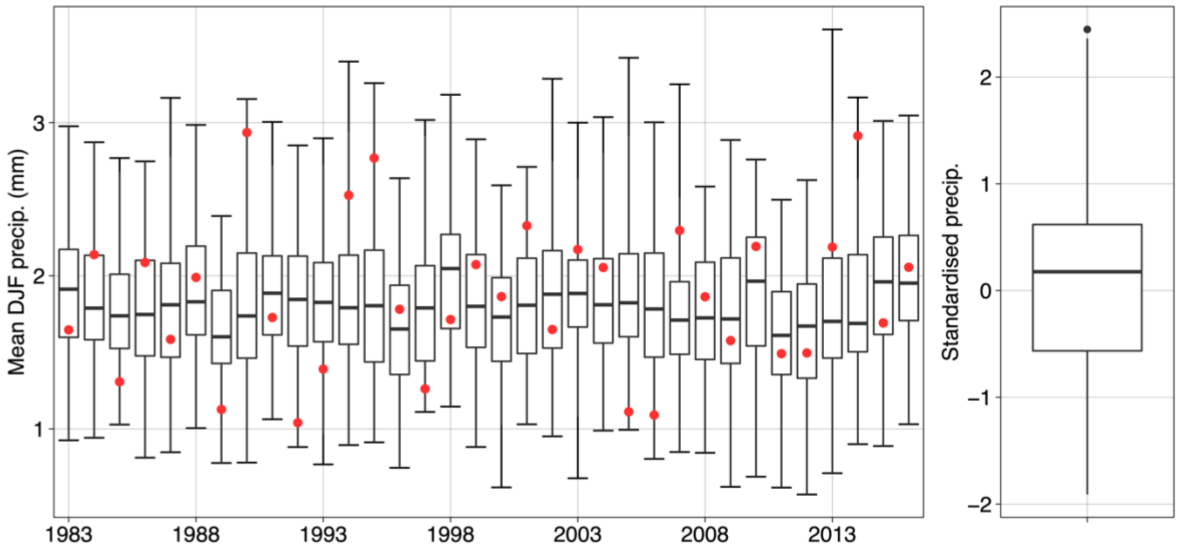

Figure 1 (left) Observed and simulated mean winter (DJF) rainfall over 1983-2016 from the SEAS5 hindcasts across 25 ensemble members and three lead times (Sept, Oct and Nov) over the Anglian region. The whiskers of the box plot extend to the maximum and minimum modelled value. (right) Distribution of observed winter rainfall for each year over 1983-2016 standardised against the distribution of simulated rainfall of the same year.

Minor comments

1. L31. You mention UK as the area here, but the study area is rather eastern England.

Response: We will make it clearer that although the study area is eastern England, the combination of NAO and EA influences winter rainfall for all regions of the UK and conditional outlooks can be applied to other regions.

2. The abstract is a bit too long and the methods is described quite convoluted. I would suggest to shorten this a bit.

Response: We will shorten the abstract. We believe the methods section is rightly longer for a paper like this given we are trying to set out a new framework to create drought storylines. We will try to make the methods section shorter but this might be challenging and might not be possible, especially in light of the later comments from the reviewer (i.e. comment 4 below) and related comments from reviewer #3 which would require elaborating on the use of the UNSEEN fidelity test and explaining more clearly the aims and objectives of our paper.

3. L42. SEAS5 provides forecasts up to 215 days, which is just over 7 months.

Response: We will correct this.

4. L76-78. The argument that there are 2850 winters in the hindcast dataset is a bit exaggerated since the ensemble members from each initialisation are not independent runs. Also, the model runs tend to approach the model cliamte over time. I think this can be further elaborated.

Response: Regarding the reviewer's point on drift towards model climatology, this is why we have tested for the independence and stability of the SEAS5 hindcasts across the three lead times in our paper (Lines 87-91 in the original manuscript). We will aim to make this clearer in the revised methods section. The use of initialised ensembles to explore plausible worst cases is well established, either by pooling climate model hindcasts as in Thompson et al. (2017) or by pooling weather forecasts as in Kelder et al. (2021) and Brunner and Slater (2022) among others. As detailed in Kelder et al. (2022), techniques have been developed to assess the model fidelity of seasonal hindcasts. First, the stability test recognizes that forecasts may drift towards the model climatology over time. This is tested by comparing the distribution of simulated winter rainfall between the three lead times. As seen from Fig 1b, the distribution of simulated winter rainfall is similar across the three lead times, indicating that there is no discernible drift. Second, the independence test assesses whether each ensemble member for each lead time is correlated with each other. This is done by considering the correlation between each pair of ensemble members over the 1982-2015 period for each lead time. As seen from Figure 1c in the manuscript, the median correlation for all three lead times is close to zero, thus showing that the ensemble members can be considered independent from each other over the considered timescale.

L99. The indices were calculated from MSLP. I would put that in the first sentence for clarity.

Response: We will correct.

5. Figure 2. From the figure it looks like the EA has a clear correlation with rainfall anomaly, whereas NINO3.4 and NAO has very weak correlation with rainfall.

Response: We will make the suggested change in the text to describe that 1) there is no clear relationship between ENSO phase and rainfall anomalies, 2) there is a positive relationship between the EA index and rainfall anomalies, and 3) there is a weak negative relationship between the NAO index and rainfall anomalies.

6. L184. The standard nowadays is to use KGE rather than NSE

Response: We thank the reviewer for this suggestion. In our paper, we have followed the multi-objective calibration strategy to determine the top parameter set for the GR6J hydrological model simulations. As introduced by Smith et al. (2019), the strategy was used successfully to calibrate hydrological models to reconstruct UK river flows and historic droughts. We have therefore not relied entirely on one statistical measure for model performance but instead have used a variety of model performance statistics (Nash-Sutcliffe Efficiency (NSE), NSE of logarithmic flows, absolute percentage bias, percentage error in Q95 (low flows) and percentage error in mean annual minimum (30-day averaged) flow) to take into account the full range of river flow characteristics. We chose to report only the logNSE values in Table 1 for brevity but we propose to add the model performance statistics for all statistical measures in a table in the supplementary materials in the revised manuscript.

7. L199. Why do you state here that "Although winter 2022/23 has now passed.."? The approach is valid regardless if applied to previous cases or in real-time?

Response: The reviewer is correct that this approach is valid whether applied to a prior drought event or to an on-going event in real time. We will add a sentence to clarity this point. The retrospective discussion of winter 2022/23 mainly serves to show that there were still some concerns about water resources after winter 2022/23, and that the rainfall response similar to the winter that actually transpired is included within the larger sample of hindcast winters employed in this study.

8. L214. What is slightly lower? How much of normal rainfall expressed in percentage?

Response: We will add the percentage of normal rainfall.

9. L216. Yes, soil moisture would be depleted during a dry summer, but lower reservoirs would not be as much affected, which drought conditions are you discussing here?

Response: We will clarify the impacts of lower soil moisture on agricultural droughts.

10. Figure 5. The colour scale of the percentiles are not great, since normal conditions have a similar colour scheme as the below normal (yellow-orange-red). I would suggest to use a greenish colour to denote normal conditions.

Response: This particular colour scale is employed by the UK Hydrological Outlook (https://hydoutuk.net/), the UK Hydrological Summaries (https://nrfa.ceh.ac.uk/monthly-hydrological-summary-uk) and more recently the WMO HydroSOS (Status and Outlooks) portal (https://eip.ceh.ac.uk/hydrology/HydroSOS/portal/). The color scale has been tested in multiple iterations with a range of key stakeholders in workshops and focus groups in past projects aiming at better communicating hydro-meteorological information (e.g. AboutDrought, RADAR and EdGE). The colour scale takes into consideration factors such as colour-blind friendliness, user-friendliness and intuitive interpretation. Normal conditions are set to a tan colour in order to be relatively "unnoticable" so that outcomes that focus on extremes are more noticeable. We would prefer to keep the color scale as it is but propose to add thin borders to the bars to make them visually clearer.

References

Brunner, M. I. and Slater, L. J.: Extreme floods in Europe: going beyond observations using reforecast ensemble pooling, Hydrology and Earth System Sciences, 26, 469–482, https://doi.org/10.5194/hess-26-469-2022, 2022.

Kelder, T., Müller, M., Slater, L. J., Marjoribanks, T. I., Wilby, R. L., Prudhomme, C., Bohlinger, P., Ferranti, L., and Nipen, T.: Using UNSEEN trends to detect decadal changes in 100-year precipitation extremes, npj Clim Atmos Sci, 3, 1–13, https://doi.org/10.1038/s41612-020-00149-4, 2020.

Kelder, T., Marjoribanks, T. I., Slater, L. J., Prudhomme, C., Wilby, R. L., Wagemann, J., and Dunstone, N.: An open workflow to gain insights about low-likelihood high-impact weather events from initialized predictions, Meteorological Applications, 29, https://doi.org/10.1002/met.2065, 2022.

Richardson, D., Fowler, H. J., Kilsby, C. G., Neal, R., and Dankers, R.: Improving sub-seasonal forecast skill of meteorological drought: a weather pattern approach, Natural Hazards and Earth System Sciences, 20, 107–124, https://doi.org/10.5194/nhess-20-107-2020, 2020.

Smith, K. A., Barker, L. J., Tanguy, M., Parry, S., Harrigan, S., Legg, T. P., Prudhomme, C., and Hannaford, J.: A multi-objective ensemble approach to hydrological modelling in the UK: an application to historic drought reconstruction, Hydrology and Earth System Sciences, 23, 3247–3268, https://doi.org/10.5194/hess-23-3247-2019, 2019.

Thompson, V., Dunstone, N. J., Scaife, A. A., Smith, D. M., Slingo, J. M., Brown, S., and Belcher, S. E.: High risk of unprecedented UK rainfall in the current climate, Nature Communications, 8, 107, https://doi.org/10.1038/s41467-017-00275-3, 2017.