Referee report 1:

Assessment

I appreciate the authors' careful consideration of my main concerns in their revised manuscript and improved the clarity of the paper. I accept the manuscript with only very minor revision, mainly textual.

Response: We would like to thank the reviewer for his valuable and thorough assessment of our manuscript.

Line by line comments

L refers to line and P refers to page.

P2L38-39: In this sentence, suddenly the authors mention compound droughts and heatwaves. I am thinking maybe just write heatwaves as an example? "In this context, the compound effects of droughts with e.g., heatwaves have……"

Response: Accepted and corrected.

P3L89-90: I would like the authors to check carefully again about the ET method. The reference seems for Hargreaves method while the authors stated PM method. I think you should remove reference Hargreaves.

Response: Reviewer is correct, and we apologize for the confusion. In this work, the mHM model uses the Hargreaves-Samani method; therefore, we removed the second part of the sentence: "~~for which the Penman-Monteith method as described by Allen et al. (1998) was used~~"

P6L159: Fig. 1(a)-(c) maybe write it Fig. 1A(a)-(c) like you did for the rest of the figure references.

Response: Accepted and corrected.

P8L178: You cite Figure 3Aa twice.

Response: Accepted and corrected.

P9L198: It is not eight because N. America drought occurred in July 1999 so it was before 2000.

Response: Accepted and corrected.

P9L200-201: Explanation in the rebuttal should be written here so the readers understand that the long droughts were not occurred only in one location.

Response: Accepted, we added an explanation as follows:

"The most extreme identified drought, which began in November 2004 in eastern Eurasia, was still ongoing until December 2022 and achieved a maximum extent of 6.7 million km2 in 2021 and was moving through particularly Eastern Europe and Sibiria during this period, changing its location and extent."

P10: Table 1: could you check all the alphabets for the scores in all tables? Some use small letter some use capital letter. For example, Table 2 for C, table 3 for B and A, etc.

Response: Accepted and corrected.

P21L365: Maybe explain again what is c.

Response: Accepted, “c.” changed to “approximately”.

Referee report 2:

The article brings a specific approach to global evaluation of the drought in a certain statistical categories at global level and it differs from the theoretical framework of generalized drought of smaller areal dimension. The positive value of the calculated drought characteristics in this work is the concept of ‘‘consecutive’’ occurrences of water deficiency. This approach brings a valuable view in theory on drought categorization at global scale and as stated in the abstract it gives a good opportunity to analyze the evolving features of spatiotemporal linked drought events. On the other hand it does not have concrete practical application and it would be difficult to investigate the impacts of such defined drought events as the impacts are of rather smaller dimension with strong spatial variations depending on the local conditions. An innovative approach is presented in using two model systems as well as a specific approach in analyzing the results. However, I have several questions that should be answered in the paper and several formal comments that should be addressed.

Response: We would like to thank the reviewer for his/her valuable and thorough assessment of our manuscript.

We would just like to note that while this study, (conducted on a global scale) could not study drought impacts on a regional scale, we believe our catalog has an application potential as the delimited and classified GLDEs on individual continents can be investigated in terms of impacts in following studies.

General notices

The severity of the drought is defined in a special way depending on the area affected by the drought. Drought as understood in this study results from a deficiency of water in surface or subsurface components of the hydrologic system expressed in AWR and SM terms. Its start is defined by a statistical percentile limit while the severity is attached to the areal extent of the drought within its duration. This approach expresses a certain level of the water deficiency in the area but it does not effectively take into account the aggravating effect behind the defined statistical threshold. This calculation allows to express a certain intensity threshold and its duration but not the actual intensity of the drought and its variation within this time interval I terms of water deficiency level. Intensity drought balance is an important component of drought mainly with the respect to long lasting (even years) drought which is limited by applying above described approach. The areal extent does not express the real deficit of water. This fact should be discussed in text.

Response: The selected definition of drought allows for objective and repeatable determination of drought event start, duration and also intensity since within the analyzed parameters both maximum extent and total time x area affected has been considered in clustering the events. It should be stressed that the approach is primarily used to catalogue drought events and categorize them to allow for testing hypothesis with regards to impact of individual events using e.g. status

of vegetation data, wildfire extent or reported drought impacts on stream flow, reservoir water levels or yields. Moreover, we would like to mention that our severity classification uses the intensity parameter defined by lower percentile to account for occurrence of very severe water deficit withing the defined "drought area".

To clarify that our approach does not study the impacts of sustained drought in one location, we extended the discussion as follows:

"… We also acknowledge that the fixed/prescribed model functions representing hydrological processes might not be equally plausible across different parts of the world and our approach also does not aim to study impacts of sustained drought in one location. …"

Introduction of the dynamic factor in terms of drought geographical spread and motion gives very valuable dimension to drought evaluation. Nevertheless, the dynamic characteristic expressed by the centroid positions and their movement is in a certain way interconnected with the spatiotemporal characteristics of drought severity and these two characteristics are not fully independent one from another. All the centroids, thought their occurrence is time dependent, must necessarily be located within the area defined in the severity categories.

Response: Looking for connections between drought severity (particularly concerning its extent) and dynamic characteristics is one of the reasons why we created the presented catalog. It is logical, that the extent of a drought event is interconnected with its movement and evolution in time generally, and our approach enables basically quantify and compare these drought event aspects. So, we do not see this interconnection as a negative thing.

The soil water supply in presented work is basically based on the precipitation only. Nevertheless, the continental dimension of the drought evaluation includes the vast regions supported by water from surrounding mountains both by surface and underground flows. The earth surface in this study is taken rather as a flat area from this point of view (what is in a contradiction to the statement regarding the drought propagation in Australia). This should be mentioned or discussed in the text, as well.

Response: We appreciate the reviewer's comment on the potential role of lateral subsurface flows into soil water supply in hydrological processes, especially in mountainous regions. We fully recognize the importance of subsurface flow contributions at local to regional scales, we would like to emphasize that our study considers soil moisture at a spatial resolution of 0.5° (50 km at the equator). Sub-grid topographic features and local-scale heterogeneities (such as hillslopes and small catchments that typically drive lateral subsurface flow) are not explicitly resolved at this coarser resolution. Therefore, the dominant hydrological processes captured at this scale are those governed by large-scale atmospheric inputs, like precipitation, as the reviewer correctly pointed out, as well as the surface processes (e.g., evapotranspiration and runoff), with soil moisture dynamics primarily influenced by vertical fluxes. Many previous large-scale hydrological and land surface modelling studies (Telteu et al 2021 or Müller Schmied et al 2024) have similarly assumed that lateral subsurface flow can be neglected at resolutions of 0.25°–1°, as its integrated effect tends to be minimal when averaged over larger grid cells.

Telteu, C.-E., et al. (2021): Understanding each other's models: an introduction and a standard representation of 16 global water models to support intercomparison, improvement, and communication, Geosci. Model Dev., 14, 3843–3878, https://doi.org/10.5194/gmd-14-3843-2021

Müller Schmied, H., et al. (2024): Graphical representation of global water models, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2024-1303

Specific notices

Data – why are the databases for SoilClim calculation and mHM calculation different even in the precipitation and temperature parameters ERA5-Land (Muñoz-Sabater et al., 2021) vs ERA5 reanalysis (Hersbach et al., 2020)?

Response: We acknowledge the reviewer's point regarding the different meteorological inputs (ERA5-Land vs. ERA5) used in the SoilClim and mHM model setups. These setups were developed independently by separate teams prior to this study. While we recognise the potential for differences arising from the use of distinct datasets, it's important to note that our primary objective is not a direct comparison of SoilClim and mHM, but rather the creation of a robust soil drought catalogue. As ERA5-Land is essentially a land-surface enhanced downscaling of ERA5, the differences in temperature and precipitation fields are relatively small, particularly when aggregated and averaged out to the 0.5 ° spatial resolution at the end, as we used in our analysis. This aggregation process effectively minimizes the influence of fine-scale variations. Therefore, we think that the minor discrepancies between ERA5-Land and ERA5 do not significantly affect the large-scale drought patterns identified in our catalogue.

-S1 and S2 reflect the differences in methodology; but how far the differences in databases can also contribute to these differences?

Response: As mentioned in the previous response, ERA5/ERA5-Land forcings differ only marginally at the coarser (spatially aggregated) resolution. However, there is no reason why this difference should make significant changes on a continental scale. The differences between GLDEs in S1 and S2 are primarily the result of the difference between SoilClim and mHM.

-R86 “This study considers soil moisture (SM) simulations averaged over the entire 2 m soil column depth (aggregating values over six soil layers) to quantify shallow water availability”. How far is SM comparable with AWR?

Response: AWR is basically relative soil moisture but considers only values between the wilting point (0%) and the field capacity (100%). In the case of soil moisture below/above these thresholds, AWR just stays 0/100%. SM is “standard” relative soil moisture defined as the ratio between volumetric soil moisture and maximum available soil water capacity, so it has a wider scale than AWR, but otherwise both variables do not differ. Nevertheless, we would like to highlight that for the purposes of our analysis, both AWR and SM were subsequently transformed into percentile-based indicators. This transformation effectively normalized the two variables, which allowed for a direct comparison of their relative dryness or wetness across different spatial locations and periods, regardless of their original scale. Therefore, while AWR

and SM differ in their raw representation of soil moisture, their percentile-based transformations allow for meaningful comparisons within our study's context.

-The severity scores in Table 1 should be better defined/clarified. Drought occurrence was identified using the 10th-percentile drought. In Table 1 (a) represents maximum areal extent of calculated out of 10th percentile; (d) represents drought intensity expressed as the total sum of areal extents of 2nd-percentile drought; but (d) is equal and even bigger than a) in table 1. This is confusing and should be better explained/clarified in text or in Table 1 legend.

Response: The scores are explained in detail in Section 3.2:

"… Subsequently, for each of these characteristics and both datasets, the identified drought events (775 for SoilClim and 630 for mHM) were placed in order from the lowest to the highest values of the given characteristic, and orders of the values were used as scores. For example, a score of 1 was attributed to the event with the lowest value of the given characteristic, and a score of 775 or 630 was attributed to the event with the highest value of this characteristic for SoilClim and mHM, respectively."

So, the scores represent just an order of the given GLDE in given characteristic, which was calculated separately for each characteristic. Therefore, the fact that values of characteristic (a) are inherently higher than characteristic (d) does not reflect into the scores, which range 1–775 (or 1–630 for mHM) for all four characteristics.

The explanation of scores is too long to be put into Table 1 caption, however we clarified the caption as follows:

"… according to the severity classification. The severity scores calculation is described in Sect. 3.2 from severity characteristics which are specified in points a–d in Sect. 3.2 (* indicates ongoing droughts)."

- P11, R 227-28 The claim „The box plots of these categories are shown in Fig. 5, ... with nearly no overlap in ... categories doesn`t look to be the full truth – This claim is not fully visible in (A) soil clim (b) and (c) as well as in (B)mMH (b) and (c) in Fig. 5.

Response: Accepted, we adjusted the sentence as follows:

"The box plots of these categories are shown in Fig. 5, which reveals that the employed characteristics decrease in a stepwise manner from the category of extremely dynamic droughts (7d) to that of extremely static droughts (1d), with only small overlaps in values among the interquartile ranges of the seven individual categories."

-P20 -there is a number of uncertainties originating from the different databases, methodological performances and clustering described. Nevertheless, there is no estimate of any of these uncertainties. Percentages The level of elimination of the effect of uncertainties that affect the long-term means within individual grids should be documented or at least discussed in some example(s).

Response: Using both SoilClim and mHM independently in the entire study is our main approach to demonstrate the level of uncertainty and how the results change with a different approach. However, we cannot exactly quantify the magnitude of uncertainties for most variables such as meteorological inputs.

-P20, R 344-346 Table S2 does not show the disconnection of GLDE event but the overlap in 2013/06/26 – 2014/07/10.

Response: There is a temporal overlap but these two GLDEs are not spatiotemporally connected. The second GLDE started while the first one was diminishing in a different location, so the algorithm defined them as separate but in the case of SoilClim based GLDEs there was a spatial connection and a very long-lasting GLDE was delimited.

P23, R419-421 "…Australia, including a lack of larger mountain ranges and other elements, which could be expected to constrain drought propagation". This is questionable statement which should be supported by some arguments regarding Australian climatography. The influence of orography o the drought spread is region specific and the influence of atmospheric circulation must not be neglected in such evaluation.

Response: Accepted, we acknowledge that our statement was inaccurate, and we changed the sentence (deleting our speculation regarding the geography of Australia) as follows:

"This finding is due to the fact that it is by far the smallest of the analyzed continents; however, in terms of the 7d category, the relative proportion was much greater~~, which fits the geography of Australia, including a lack of larger mountain ranges and other elements, which could be expected to constrain drought propagation~~."

- As no impacts of any drought event listed in Tab 1 and 2 was mentioned the whole study has to be taken as a purely theoretical approach to evaluate the spread and dynamic of drought in a certain relative categories. There are some unspoken simplifications of the influence of geographical features of the earth's surface and the effect of atmospheric circulation was not considered. Nevertheless, the article provides a comprehensive robust methodology for global drought assessment and can serve as a basis for further investigation of identified the drought events and their propagation at regional and global scale.

Response: This study primarily presents our approach to delimit, catalog and classify drought events. We believe this catalog will be used further for the study of drought impacts and drought drivers such as atmospheric circulation, however we did not see any more space within this study to include these aspects (which are thematically out of the main aim of the study) as the study is rather extensive as it stands.

Formal errors

-Capital D in Table 1 indicating the drought intensity (should be small d). Similarly in Table 2 C->c.

Response: Accepted and corrected.

-No keywords list available

Response: We used the suggested "Copernicus_Word_template" that does not include keywords.

-P3, R90 Allen et al. (1998) used in text is not listed in References

Response: Accepted and corrected.

-P21, R353 Table 4 is not available neither n the manuscript or i the supplements

Response: Accepted and corrected to Table 2B.

-P21, R359 Fig. 11 is not available neither n the manuscript or i the supplements

Response: Accepted and corrected to “Figs. 1 and, 4”.

-Table 3 – there are both capital and small letters used to mark the absolute and relative frequencies in severity classification section of the table

Response: Accepted and corrected.