We thank the reviewer for their comments and suggestions. Please find hereafter our replies in italics shaded in blue. For the sake of clarity, a clear distinction is made between compound dry and hot (CDH) events and the dry and hot extreme event database (Dheed).

RC1: 'Comment on essd-2024-396', Anonymous Referee #1, 16 Nov 2024

Citation: https://doi.org/10.5194/essd-2024-396-RC1

The manuscript essd-2024-396 presents a dataset identifying historical compound dry and hot extreme events derived from ERA5 global meteorological data. The study's focus on compound events, which are more damaging than univariate extremes, is timely and valuable for the community. However, several issues require attention to enhance the dataset's reliability and utility.

Major Concerns:

1. **Global Trend of Heat Extremes**:

   The data analysis indicates a non-significant trend in heat extreme day numbers globally from 1970 to 2022. This finding contradicts numerous studies showing that heatwaves have become more frequent and severe over time. This discrepancy should be thoroughly investigated, as it undermines the reliability of the dataset.

   We thank the reviewer for bringing our attention to this result, which also puzzled us. After a careful revision of our workflow, we identified an error in our code, which introduced a monthly bias in the temperature and precipitation data. All the analyses have been run again and the new results show a significant trend in the annual number of hot days.

2. **Data & Methodology**:

   a. The dataset relies exclusively on ERA5 data. The authors should include a literature review demonstrating that ERA5 is widely accepted for historical drought and heat stress analyses.

   ERA5 has been widely used in scientific publications related to drought and/or heat. A Google Scholar search on citations of Hersbach et al. 2020 reveals that out of the 19,587 citations, 5,510 contain the keyword drought and 10,500 contain the keywords heat or heatwave. Searching for drought AND (heat OR heatwave) reveals 3,600 citations.
   https://scholar.google.com/scholar?hl=en&as_sdt=2005&sciodt=0%2C5&cites=18403910731188548420&scipsc=1&q=drought+AND+%28heat+OR+heatwave%29&btnG=

   Nevertheless, various studies report limitations regarding precipitations, especially in the tropics.

   b. The use of the 1% threshold for defining "extreme" events requires justification through references to relevant literature.

Most studies on extreme events use a threshold of 5% or 10% (resp. 95% or 90%) on monthly anomalies based on a 30 year reference period, corresponding to a return period for a specific location and month of about 20 or 10 years. In this study, we use a longer reference period (74 years) and daily data, which, for a grid cell, leads to an annual average of 3 to 4 days flagged as extremes for a single indicator.

c. The parameterization G=0.5Rn for ground heat flux is uncommon as this ratio is typically associated with vegetation cover. Now the impact from the surface cover is missing. This choice should be explained or supported with references.

For the reference evapotranspiration, we adopted the parametrization G=0.5Rn for the nighttime soil heat flux density following Allen et al. 1998 (Equations 45 and 46) and Singer et al., 2021. This value is not the actual nor the potential evapotranspiration but a reference for a well watered grass cover.

3. **Seasonal Detrending**:

The proposed methodology lacks seasonal detrending, a standard preprocessing step in drought and heat analyses. Without removing seasonal cycles, anomalies are compared to absolute values rather than seasonal baselines. For example, warmer days in spring might qualify as heat stress even if their absolute temperatures are lower than those in summer. Similarly, seasonal cycles in PEI could influence water stress results. This methodological issue weakens the robustness of the findings and should be addressed.

The rationale behind the choice of not removing the mean seasonal cycle (MSC) is that the primary aim of our database is to analyse the impacts on the vegetation, which are more sensitive to absolute thresholds than to relative ones (see, e.g., Marchin et al. 2022). Especially, in a changing climate with a shifting reference, the MSC should be adapted dynamically. A shift in the start of the growing season or the rainy season would be considered as extreme in terms of anomalies, while the physiological consequences wouldn't necessarily be such.

4. **Omission and Commission Errors**:

The authors conducted a qualitative literature survey of extreme events captured or missed by the dataset. This evaluation is commendable but could be improved by quantitatively summarizing omission and commission errors. These statistics should also be briefly mentioned in the abstract for clarity.

We thank the reviewer for this valuable suggestion. However, we think that commission and omission errors computed on 40 reported events and 20 events from the Dheed database might underestimate the actual error. Indeed, on the one hand, the chosen events are spatially large and/or last many days, which is not the case for most of the events in the database. On the other hand, searching for references to small events might be unfeasible without dedicated text mining and web crawling tools, which are outside of the scope of this study.

Nevertheless, we will add a sentence in the revised abstract and manuscript stating that: Out of 40 events selected a priori, 38 could be associated with Dheed events. All 10 largest and 10 longest events from Dheed could be linked to events reported in the scientific or grey literature.

5. **Global Trend Mapping**:

   Figure 6 presents trend analyses across continents. A pixelwise global trend map would provide a more detailed spatial representation and is strongly recommended.

   We thank the reviewer for this suggestion. We will add a pixelwise decadal trend map to the revised manuscript. Indeed at the grid cell scale, an annual trend map doesn't make sense because, thankfully, there aren't extremely dry and hot days every year. However, when aggregating the number of extremely dry and hot days over ten years, a significant trend can be observed in many places.

6. **Intercomparison with Other Indices**:

   The dataset should be compared with drought and heat indices from other sources over a long time span to validate its reliability.

   We thank the reviewer for this suggestion. The qualitative validation of the CDH events shows that our indices are suited to detect large events. A more systematic quantitative comparison with indices from other sources could indeed reinforce the confidence in the Dheed database. Under the constraint of the revision deadline, we haven't been able to do this in the time imparted. We would compare the daily PEI with daily SPEI at ICOS sites (EOBS based, Pohls et al 2023 https://doi.org/10.5281/zenodo.7561854; ERA5 based Liu et al 2024 https://doi.org/10.5281/zenodo.8060268).

Additional Comments:

- **Data Citation**: ESSD requires the data URL and citation to be explicitly mentioned in the abstract.

  The data URL and citation will be added to the revised abstract.

- **Figure 1**: The terms "ET0" and "ETref" should be unified for consistency.

  We will use the notation ET0 throughout the manuscript, consistent with Allen et al 1998

- **Hourly ETref Calculation**: Clarify why θ (daily mean temperature) is used instead of hourly temperature.

  There was an error in the manuscript. The hourly temperature is used in the calculation. This will be corrected in the revised manuscript.

- **Missing References**: Two papers by Ima are cited but not included in the reference list.

The two pieces of software have no author and their references were included in the beginning of the list. They have now been moved to follow the alphabetical order, using the software title as reference entry.

- **Figure 2**: Clarify the difference between "no extreme" and "10th–90th percentile."

  All values lying between the 10th and 90th percentiles are flagged as no extreme. The tails of the distributions (values smaller than the 10th percentile or greater than the 90th percentile are depicted in white. There was an error in the description of the figure, stating that the white colour shows the centre of the distribution, while the grey colour does. The caption will be modified accordingly in the revised manuscript.

- **Figure 3**: Add a legend explaining the meaning of the additional colors.

  The following sentence will be added to the revised manuscript:

The shades of blue and purple show the accumulation period of the water balance. The darker the shade the longer the accumulation period: a water balance accumulated over 180 days which is below the 1% threshold is rendered in the darkest shade. The 90 day accumulation period is shown in the medium shade. The 30 day accumulation period has the lightest shade.

- **Table 2**: Specify the units for area and volume.

  The area and volume are not properly area and volume but are proportional to the area and to the area multiplied by the duration of the event in days. The following sentence will be added to the table caption: The area is an adimensional proxy of the spatial area affected by an event obtained by counting the number of voxels in an event multiplied by the cosine of their respective latitude (volume) divided by the number of days between the start and the end of the event (duration). An area of 1 is the size of a 0.1x0.1 degree grid cell at the Equator or about 122 km^2.

- **Data Accessibility**: URLs for the data cubes are not accessible. Please ensure they are active and functional.

  We are sorry that the reviewer was not able to access the data at the provided URL. The zenodo links do work. There was indeed an error in the s3 link, which should have been https://s3.bgc-jena.mpg.de:9000/deepextremes/v4/ReadMe.md. The links to the revised version of the database will be corrected in the revised manuscript.