Review of: "Dhhed: and ERA5 based global database of dry and hot extreme events from 1950 to 2022"

December 18, 2024

The paper presents a new dataset of compound dry and hot extreme events for the historical period 1950-2022, based on the ERA5 reanalysis data. The methodology introduced by the authors represents a valuable effort towards a more unified approach for the characterization of dry and hot days over different land regions. This could be valuable for both climatological and impact studies. Unfortunately, before the paper can be considered for publication in the journal, there are a serious of major issues that need to be carefully addressed by the authors:

1 Major Concerns

- The analysis is conducted on data with a seasonal cycle. In order to conduct an analysis of extremes at different locations and time of the year, it is usual to remove the seasonal trend from the data when determining their extremeness.

- Very often in literature, as also mentioned by the authors, heatwaves and droughts are calculated with respect to a reference period. The length of the reference period is normally 30 years. This is usually a compromise between having a sample size large enough for the robust computation of high and low percentiles, as well as having a period where trends related to climate change can be neglected. The authors, on the contrary, calculate their threshold for determining extreme events considering, for each grid box, all the days of the years from 1950 to 2022. This could lead to undesirable biases due to strong trends in the considered variables. Possibly acknowledging this, the authors suggest to calculate their threshold for extreme events with respect to a reference period in future work. I think that this is a very important point that could compromise the reliability of the provided dataset and I would suggest the authors to reperform their analysis, properly defining their threshold with respect to a commonly used 30-year period.

- Also, I think that the authors should better motivate the selection of the 1% threshold for defining extremes. Higher percentiles are more sensitive to the size of the sample used for estimating the underlying distribution of the considered data (See Brunner and Voigt 2024 for further clarifications). For this reason, it is usual to characterize heatwaves using thresholds based on the 90th or 95th percentile (Perkins et al. 2012, Russo et al. 2014, Russo et al. 2015). How would your results change for lower thresholds (i.e. 95th or 90th percentile)?

- In general the text should be improved for clarity. In particular, the method section can be sensibly improved: at the moment the employed methodologies do no always result clear.

- There is large confidence in the literature, often based on the same dataset used by the authors in their study (i.e. ERA5), that heatwaves have increased in extent, length, intensity and number during the historical period. I find it alarming that the authors do not find any significant trend in heat days in their dataset. Hence, I would suggest the authors to carefully review their analysis, also taking into consideration my previous comments.

- Also connected to the previous point, the paper proposes a qualitative evaluation of the dataset based on events commonly known from the literature. I think that such an evaluation, at least qualitatively, should be also conducted for the results of the analysis of the trends in the different considered variables.

- In their method for characterizing compound extreme events, the authors use together 3 variables for precipitation and 1 for temperature? Would this not lead to results that are biased towards droughts rather than heatwaves?

- The Labels of all figures should be improved, possibly including more and better details

2 Specific Comments

- The Abstract is very generic. It could be extended and should be reconsidered

- In the abstract, please make it clearer that in your method you consider 1 heat extreme indicator and 3 droughts indicators at the moment.

- Also in the abstract, you mention details of the conducted analysis (such as peak over threshold) that are not reported in the methods

- The structure of the introduction is a bit too general and can be improved

- INtro: you can describe more examples on how extremne events affect the ecosystem and society
- l 12-13: "With Earth climate currently changing": revision needed. The Earth climate has been always changing, over million (if not billions) of years. Here, I think that the authors refer to anthropogenic driven climate change. Please reformulate this period.
- l 16-19: Can you provide more examples on the impact of compound heat-hydrological extremes on vegetation as compared to single events?
- l 22: "The cascading process.. also impact society as a whole": could you provide more details why and in which way?
- l 24-25: I think that there are many studies in the literature working towards a more unified definition of droughts and heatwaves that would be worth to mention here, such as for example:

  Perkins, Sarah E. "A review on the scientific understanding of heatwaves - Their measurement, driving mechanisms, and changes at the global scale." Atmospheric Research 164 (2015): 242-267.
- Also, please mention that the lack of a unified way of defining heatwaves or drought is also due to the fact that the definition of these events often depends on the purpose of the study, the considered region and time of a year.
- l33: "compartment": use different wording
- l39-40: reformulation needed
- l43: "can propagate into impacts": reformulate
- l46-47: What was the goal of the studies you mention? which events they have considered? Please provide more details
- l 48-50: please reformulate
- l 52-54: "For example, it can serve ... to train models predicting ... ": can you better describe here what you mean by minicubes of high resolution satellite imagery and how Dheed can be useful for the purpose you mention?
- In the introduction, I think that the advantages of a reliable identification of Dheed can be better detailed.
- For your analysis, which period of the year you consider? You consider all seasons together? Please specify in section 2.1.
- Most of the data for the period 1950-2022 are expected to be characterized by a trend, especially for T2M. How do you take this into account for your analyses?

- In section 2.1 you say that you calculate daily mean, min and max temperatures from hourly T2M values. But which of these 3 variables you use in your analysis is not clear.
- I think it would be better to introduce ETref before mentioning it in line 76
- l77-82: Please reformulate, since the information you provide are not clear. Why You have 60x60 points in longitude and latitudwe for each Zarr cube? Is this a personal choice of the author? Please clarify. Please, also be aware that the original horizontal resolution of ERA5 data is not 0.25 degrees lon. Specify that you use a gridded product, provided on a regular grid with a horizontal resolution of 0.25 degrees.
- l 89: it should be $\Theta_2 m$
- In equation 1, how do you calculate the parameter C_d? I see that you provide more details at lines 101-102. I think that it would help a better readibility of this part of the text if details on parameter C_d are reported when the parameter is first introduced.
- In Equation 1, from where you derived the value of G? please specify.
- Please provide reference for equation 2
- l 102: what is 10^{-6}? the value of changes in ETref when using constant values of C_d? please specify
- When aggregatinhg the 3 PEIs with Tmax in your dataset, you will be giving more weights to drought events. Why this choice? Wouldn't it be better to consider only one drought indicator together with the heat indicator? for example, for a drought event lasting more than 90 days, this will be counted 3 times according to your definition.
- l 114-115: "Fitting a parametric distribution... proved difficult": why? please provide more details
- l117: Not clear, please reformulate.
- Why you choose the 1 % as a threshold? Higher percentiles are more sensitive to the sample size used for estimating the underlying data distribution. For this reason, it is usual to characterize heatwaves with thresholds based on the 90th or 95th percentile. How would your results change in this case?
- l 122: how do you define a data cube?
- l 125-126: "have uneven values greater than 1": what is the unit you are considering in this case? Which metric are you considering?

- l. 124: you say that you extract DHEEs as labelled groups of dry and hot DEOs. Following your analysis, I would rather reformulate this sentence. In fact I think it would be more correct to say that you extract DHEEs as labelled groups of dry and/or hot DEOs.
- Also check the text for consistency between DHEEs and dhees
- l.124-133: This part of the methods results not clear and needs reformulation. When do you apply the 3-day long condition? You do this at each grid point, before the temporal and spatial aggregation? What do you mean by "group DEOs connecting across the globe along the longitude dimension"? What about latitudes? What do you mean by "possibility to merge labels from contiguous data cubes"? Do you consider the fact that at different latitudes you have different numbers of land points in your final ranking of events?
- Fig. 2: It would be nice to also see some examples of events occurring over Asia and North America. In the caption, please be consistent in the style of the enumeration of the rows (row 1 and 2 vs Third). Also in the caption, what do you show in row 4? please specify
- l139: what is the percentage of events affected by a single indicator? Please clarify
- l. 153: You go from introducing Fig. 2 to Fig. 10. Please reconsider the order of the figures so that they can be referenced in the same order they appear in the text
- l157-159: not clear. Please reformulate
- l. 162-163: Why not excluding then the years from 1950 to 1970 already before-hand, since you know that for these years ERA5 is less reliable?
- l 165: "but there seems to be a positive trend": in which feature?
- l. 165: It is not clear how you define the different categories of Fig. 3
- Fig. 3: some colors miss a label. What do they indicate?
- l. 160: add some references on which your considerations are based
- how the 4th row of Fig. 3 differs from the plot of Fig. 5? pleae clarify
- l. 174: can you better clarify what you mean by volume?
- Caption of Fig 7: please specify that in this case you are only considering hot and dry events
- Table 2: Maybe I would join Fig. 8 and Table 2. This would help readability.

- section 3.3: A similar validation based on available literature should be conducted also for the evinced trends.
- l 213: which database?
- l.214: Provide more context on the specific events, as well as references
- Fig. 8: Maybe you can label the events by year? what are the two events occurring in Russia?
- l 238: "the final threshold": What is this final threshold?
- l. 235-240: Can you better report about this in the methods section?