Response to Reviewer 1 Comments

1) The idea behind this work, of organizing a gridded meteorological drought index into clusters describing spatially widespread dry periods and correlating these with reports of drought impacts, is meritorious, but details of the execution need more justification and improvement:

   We thank the reviewer for the recognition of the value of the overall idea of the study. In the revised manuscript, we will address these concerns and revise the relevant sections accordingly.

2) How PET is calculated for determining SPEI doesn't seem to be explained.

   PET used for the SPEI calculation is derived using the Thornthwaite equation (Thornthwaite, 1948), based on mean monthly temperature and latitude. We agree that this was not explained clearly enough in the original manuscript. In the revised manuscript, we will clarify this explicitly. We will also note that the framework and code that we make publicly available are flexible and allow users to substitute alternative PET formulations or drought indices, provided that the input data are available on the same grid system as ERA5. The downstream clustering and impact-linkage framework is therefore not tied to a specific PET method or drought index formulation.

   Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. Geographical review, 38(1), 55-94.

3) It appears that only point data from GDIS are used to localize drought reports. However, according to its documentation, GDIS provides polygons of the affected provinces in the "geometry" field. Comparing this areal information to the ERA5 meteorological drought extent should show much clearer correspondences compared to only centroids or other individual points from GDIS.

   We thank the reviewer for this important comment. We agree that using the polygon information provided by GDIS can offer a meaningful spatial comparison with the ERA5 drought extent. In the revised workflow, we will include an additional option that uses the GDIS polygons to derive the corresponding statistics. We will also provide, in the Supplementary Material, a comparison between results based on polygons and those based on centroids in order to assess the added value of the areal information.

4) At l. 267, "Frequency" seems like the wrong word for the number of months of drought; "Duration" would be more appropriate. Also, the definition of "Severity" is not clear.

   We agree that the term “frequency” may be ambiguous. However, the metric used here is not the continuous “duration” of a drought event, because it counts the number of drought-affected months within the target year, and these months do not need to be consecutive. We would like to keep the term “frequency,” and to avoid confusion we will explicitly state the units and interpretation: the metric is reported as the number of drought-affected months per year (months/year). We will also clarify the definition of severity by stating that the severity weight is based on the mean SPI/SPEI value across drought-affected months at each pixel, linearly scaled between 0 and −2 and capped for more extreme values

5) Figure 3: The caption fails to state what the blue areas in the maps are.

   The intention of Figure 3 is to show different drought clusters, with different colors used only for visual distinction. As already noted in the caption, different clusters are shown in different colors; therefore, the blue areas do not represent a separate category, but simply one of the identified clusters. To avoid any possible ambiguity, in the revised manuscript we will make this wording more explicit in the caption.

   ‘Different drought clusters are shown in different colors for visual distinction only’

6) Since "detection percentages are consistently higher for SPEI than for SPI", I recommend for the SPEI based drought definition to be used as the primary one for reporting the results, and SPI-based results to be given as secondary, whereas now it's mostly the opposite.

   We thank the reviewer for this thoughtful comment. We agree that SPEI shows higher detection percentages in several cases. However, the purpose of this study is not to recommend one index as the primary definition of drought based solely on detection percentages. Indeed, if this were the logic then it would lead to favouring very lenient drought definitions to have a very high detection percentage. Rather, SPI and SPEI are both included to demonstrate how the proposed framework can be applied to different drought indices. In addition, the relative performance of the two indices is not uniform and can vary across continents and cases, as shown in Tables S7-S9. Nevertheless, following the reviewer’s suggestion, we will revise the text to give greater emphasis to the SPEI based results where appropriate, while still presenting SPI-based results to illustrate the broader applicability of the framework.

7) There are no clear conclusions drawn as to what SP(E)I timescale is considered to define drought. Most of the figures arbitrarily only show the 1 month timescale, which admittedly can be a "flash drought" but seems too short to correspond to impactful drought in most cases. I suggest to first analyze which SP(E)I timescale matches the drought disaster dataset the best, and then report findings primarily for that timescale.

   The reviewer noted that most figures in the main text show only the 1-month timescale. However, the analysis in this study includes four different SPEI timescales, with the figures for the 3-, 6-, and 12-month timescales presented in the Supplementary Material. We agree that the relationship between drought impacts and SPEI timescale is important. However, the purpose of this manuscript is not to determine a single “best” timescale for matching the drought disaster dataset, as this will be highly dependent on the user-case and no universal "best" timescale for drought-related impacts exists, Instead, the study aims to present a flexible framework that users can be apply across multiple drought indices and accumulation timescales. Accordingly, we considered four SP(E)I timescales in the analysis.

8) As well as population and GDP, considering measures of agriculture intensity may be helpful in predicting the impacts of drought, since agriculture is by far the most water intensive major economic sector. Oddly, agriculture is not mentioned at all except in the literature review.

   We agree that agriculture is a major sector affected by drought and that measures of agricultural intensity could be highly relevant for understanding drought impacts. However, the present manuscript focuses on the impact data available in EM-DAT. EM-DAT includes no information on agriculture, and the impact categories that provide the most data entries for drought relate to human and economic losses. This is why we chose to focus on population and GDP exposure. If we were to include agricultural intensity, we would have no corresponding impact data to relate the exposure to in the analysis. We are further not aware of any global drought impact dataset for agriculture. In the revised manuscript, we will acknowledge this limitation more clearly and note that the inclusion of agricultural indicators is an important direction for future work.

9) Seasonality is also never mentioned. It might be hypothesized that droughts occurring during the growing season have much bigger impacts than those at other times of year.

   We agree that seasonality is likely to influence drought impacts, and that droughts during the growing season may be particularly consequential. However, the growing season is location and crop-specific. As noted above, no agricultural impacts are included in EM-DAT (nor in other global impact datasets such as DesInventar or Wikimpacts). This is an important knowledge gap, but would require a project of its own to: (i) maps local growing seasons across the globe; and (ii) attribute agricultural impacts to droughts. We therefore see it as well beyond the scope of the present study (or indeed of any one single study). In the revised manuscript, we will acknowledge this limitation and highlight seasonality as an important avenue for future work.

10) The mostly 3-D figures (e.g. 5-12) are not interpretable. I strongly recommend to find a different way to show results, and adjust the discussion in the text accordingly.

    We thank the reviewer for this important comment. In the revised manuscript, we will replace these figures with clearer alternative visualizations (we are currently experimenting with 2-d heat-maps) and will revise the corresponding discussion in the text accordingly to improve readability and interpretation.

11) 2×10^7 USD seems like a small amount of average damage for a large-scale drought in North America, since the USA has experienced quite a number of droughts that caused multiple billions in damage. It would be helpful to show more statistical information about the set of EM-DAT droughts included in the analysis, including their

minimum, maximum, mean, and median damages, and to compare this to information from other databases.

In the revised manuscript, we will include additional summary statistics, such as the minimum, maximum, mean, and median reported damages, and provide these in the Supplementary Material.

12) Additionally, since this is for publication in a data description journal, the paper should say more about the format of the generated dataset, including what fields it contains and what are some anticipated use cases.

We thank the reviewer for this helpful comment. We agree that the manuscript should more clearly describe the format and contents of the generated dataset, particularly given the scope of a data description journal. In the revised manuscript, we will add a dedicated section outlining the dataset structure, the variables and fields included.