

I think this is a promising and timely work. The idea of linking SPI/SPEI-based drought information with population and GDP exposure, and reported impacts from EM-DAT/GDIS is valuable. I also like the aim of moving beyond meteorological drought indicators and towards a dataset that is more relevant for exposure and impact studies. However, I do not believe the current version is ready for publication. The main issue is not about the idea, but about the way the dataset is presented and documented. For an ESSD paper, the data product per se needs to be front and centre. Users should be able to understand what the dataset contains, how it was produced, what its limitations are, and how they can reproduce or reuse it. The manuscript now still reads partly like a research paper, while the actual dataset description, uncertainty discussion, validation, and user guidance are not yet strong enough. I would therefore recommend a major revision. My detailed comments are as follows:

We sincerely thank Reviewer 3 for the detailed and constructive evaluation of our manuscript. We are grateful that the reviewer finds the work promising and timely. To address the Reviewer's concerns regarding our presentation of the dataset and user guidance, we plan to significantly restructure the manuscript. In the revised manuscript, we will add a clearer data description section, including the dataset structure, variable definitions, units, spatial and temporal coverage, quality flags, and recommended use cases. In addition, we will reduce some of the broad interpretation of continental exposure and impact patterns and use that space to better document the data product, its limitations, and its reproducibility.

My first major concern is that the current title and abstract give the impression of a general global drought dataset. However, after reviewing the manuscript, IMHO, the workflow is anchored to drought disaster records in EM-DAT/GDIS and searches for drought clusters around those events. This means that the product is not a global catalogue of meteorological drought events. It is more like a dataset of drought clusters associated with reported drought disasters. This should be made very clear in the title, abstract, introduction, and conclusion.

We agree with the reviewer. In the revised manuscript, we will make this distinction clearer throughout the title, abstract, introduction, and conclusion.

We will describe the dataset as an event based drought disaster data product. The dataset does not aim to identify every meteorological drought globally. Instead, it identifies SPI/SPEI based drought characteristics within the spatial and temporal context of globally reported drought disaster records and links these characteristics to population exposure, GDP exposure, and reported socioeconomic impacts.

**Abstract:** You state that events are tracked as spatially and temporally evolving objects. From the current manuscript, if I understand correctly, it looks more like monthly clusters are identified within a target year, but it is not clear whether there is a formal object-tracking algorithm across months. If real temporal tracking is done, please describe it clearly. If not, please rephrase this expression.

We agree that the original wording may be misleading. In the current workflow, monthly drought clusters are identified from SPI/SPEI fields within the relevant event window, and event level characteristics are then derived from these monthly drought affected areas. The workflow does not implement a formal temporal object tracking algorithm that links individual clusters dynamically from one month to the next.

In the revised manuscript, we will therefore remove or rephrase the statement that drought events are “tracked as spatially and temporally evolving objects.” We will instead state that monthly drought clusters are identified and summarized over the event window to derive event level characteristics such as detected duration, mean detected area, severity, and exposure.

**Section 2: ERA5 is at 0.25°, the population data are at 1 km, and the GDP data are at 5 arcmin. Please clearly elaborate on the spatiotemporal integration of these datasets.**

In the revised manuscript, we will add a dedicated explanation of how the ERA5 based SPI/SPEI grid, population data, and GDP data are spatially harmonized before exposure is calculated.

In the exposure workflow, the drought affected pixels are first identified on the SPI/SPEI grid for each event and month. The population and GDP rasters are reprojected and resampled to the same grid as the SPI/SPEI drought index data. After this harmonization, each drought grid cell has corresponding population and GDP values. For each event and month, the drought cluster mask is overlaid with the aligned socioeconomic rasters, so that only the population and GDP values within the detected drought affected pixels are included in the exposure calculation.

**Similarly, please clarify the actual study period and address the temporal mismatch between datasets. While the meteorological data and GDIS events cover 1960–2018, the population and GDP datasets only begin in 1975 and 1990, respectively.**

In the revised manuscript, we will explicitly distinguish between the temporal coverage of the meteorological data, GDIS/EM-DAT records, population data, and GDP data. We will clarify that not all variables are available for the same period, and we will state how this affects the availability of exposure and impact variables for individual events.

**Section 3.2: You state that ERA5 precipitation and temperature were used to calculate SPI and SPEI at diverse timescales, but several key details in determining the indices are still missing.**

In the revised manuscript, we will provide a clearer description of the SPI and SPEI calculation, including the accumulation timescales, reference period, fitting or standardization procedure, and treatment of zero precipitation where relevant. For SPEI, we will also clarify how potential evapotranspiration is calculated from ERA5 temperature data. These details will be added to the Methods section.

**Section 3.3: The description of maximum distance threshold, spatial buffer distance, merging distance, bounding boxes, target-point grouping, and cluster merging is hard to follow. Please define each parameter clearly and explain how they were chosen. In addition, you treat the GDIS locations as target points, but it is not clear whether these are centroids of affected areas, administrative locations, reported place names, or something else? Please describe the nature/limitations of the event coordinates recorded in GDIS more explicitly.**

In the revised manuscript, we will define each parameter separately, including buffer distance, merging distance, bounding boxes, and cluster merging. We will also explain the purpose of each parameter and whether it is used as a fixed processing choice, an optional configuration in our toolbox, or part of the sensitivity analysis.

In addition, we will clarify the nature and limitations of the GDIS location information. In response to this and other reviewer comments, we will use GDIS affected area polygons as the primary spatial reference wherever they are available. The point based or rectangular search option will be retained only as an optional configuration in the toolbox. This can be useful should users for example wish to extend the analysis to other impact datasets not providing full polygons.

**Section 3.4: You count all population or GDP pixels that fall in a drought cluster for at least one month during the target year. This is understandable, but it can include short-lived like flash droughts or relatively weak drought conditions in the total exposure. The weighted exposure partly addresses this, but the weighting scheme is still a methodological choice rather than a standard measure of impact. Please make the distinction between exposure, weighted exposure, and actual impact clearer, and avoid implying that exposure is the same as impact or risk.**

In the revised manuscript, we will distinguish more clearly between raw exposure, weighted exposure, detected duration, and reported impact. Raw exposure will be described as the population or GDP located within the detected drought affected area. We will also report detected duration as an additional event level metric, defined as the number of months within the event window for which drought clusters are identified. Weighted exposure will be described as a methodological metric that incorporates both the duration and severity of drought conditions within the event window. We do not intend to equate exposure to risk (or impacts), but we do argue that a joint analysis of exposure and impacts can inform on vulnerability or possible discrepancies between exposure and impact data.

**Section 3: as another major issue, EM-DAT reports impacts mostly at national scale, while GDIS provides geocoded locations for reported events. Linking national-level impacts to subnational drought clusters can introduce uncertainty. Please clearly explain how duplicate allocation is avoided and how users should interpret the matched impact values.**

We thank the reviewer for raising this point, and we will clarify it in the revised manuscript. EM-DAT provides national level impact information, meaning that the impact value is associated with the disaster event record for a given country. However, this does not necessarily mean that the whole country was affected by drought stress. EM-DAT/GDIS can also provide subnational information, such as Admin 1 or Admin 2 level locations, which helps identify the reported affected areas. We will clarify that the comparison between drought characteristics, exposure, and reported impacts is performed at the event level, and that reported impact values are not spatially disaggregated across individual drought clusters, grid cells, districts, or polygons. We see no risk of duplicate allocation since if two distinct drought events affect a country simultaneously, EM-DAT reports these as distinct event numbers and they thus have separate polygon lists in GDIS.

**Sections 4: you spend a lot of space interpreting continental exposure and impact patterns. I would suggest condensing some of the interpretation and using that space to explain the data.**

We agree. In the revised manuscript, we will shorten this section and use that space to strengthen the data description. The results section will be reorganized to support the data product rather than read mainly as a research paper interpreting data.

Section 4: various terms have been used throughout the results, but their exact definitions are not always clear. For a multi-month event, for example, is area calculated as a monthly mean, a maximum, or a union footprint over the year? Is severity averaged over drought pixels only, over months, or over the full bounding box? Please further elaborate on this.

In the revised manuscript, we will add explicit definitions for all derived event variables included in the dataset, particularly detected duration, mean detected area, and mean detected severity. Detected duration will be defined as the number of months within the event window for which at least one drought cluster is identified. Mean detected area will be calculated as the average drought affected area over the detected months. For severity, we will first calculate the spatially averaged drought severity over the detected drought pixels for each detected month, and then average these monthly severity values over the detected months. We will also define any additional variables included in the released dataset, such as detected fraction or minimum detected severity, in the data description table.

Section 4: The reported economic damages appear quite low for major drought disasters, for example values around  $2.5 \times 10^7$  USD, as the highest total damages, at a 500 km buffer distance. Please double-check this. Also, GDP exposure is reported in constant 2015 USD, so it is not directly comparable with unadjusted historical damage values unless the damage data are also harmonised. This should also be clarified carefully.

We have used the EM-DAT variable Total Damage, Adjusted ('000 US\$), which is expressed in thousands of US dollars and adjusted for inflation using the OECD US Consumer Price Index, following the EM-DAT adjustment procedure. In the revised manuscript, we will clarify that although the OECD CPI series uses 2015 as its base year, EM-DAT rescales the CPI so that the latest available year in the database version is used as the reference year. Therefore, the adjusted damage values should not simply be described as 2015 USD. We will state this explicitly in the data description, figure captions, and relevant results sections. We will also clarify that comparisons between GDP exposure and reported damages should be interpreted cautiously, because GDP exposure represents the economic activity located within the detected drought affected area, whereas reported damage is an EM-DAT impact variable with its own reporting uncertainties.

Section 4: The detection percentage across thresholds, timescales, and buffer distances is a good start. However, it mainly shows parameter sensitivity, not whether the drought clusters are reliable. Please add a more explicit quality assessment, for example, comparison with well-known drought events, independent products, or regional case studies. At least, you should provide clearer guidance on which settings are recommended for which use cases.

The purpose of the sensitivity analysis was to examine how the results change under different combinations of index, timescale, threshold, and spatial configuration, rather than to identify a single best configuration.

In the revised manuscript, we will clarify that the dataset is intended to provide ensemble style information across different methodological settings. The most appropriate configuration may depend on the user's research question, the drought process of interest, the region, the underlying input data, and the design of the study. Therefore, we will avoid presenting any one configuration as universally optimal.

However, to provide a more transparent assessment of the resulting event characteristics, we will also present case studies for which clusters were or were not detected despite reported low or high impacts. This will help users understand both the strengths and limitations of the cluster based event characterization.

You refer to many supplementary figures and tables. Without these materials, it is not possible to assess several key claims in this study. To make this manuscript self-contained, key statistics should be moved into the main text, rather than leaving the main conclusions dependent on supplementary material.

In the revised manuscript, we will reduce some of the extended interpretation of the results and make the manuscript more self contained.

Fig.3: Please add longitude and latitude labels, and other necessary elements.

In the revise manuscript, we will revise Figure 3 to include longitude and latitude labels.

Figs.4-12: These figures are difficult to interpret in the current format. The 3D surface and scatter plots do not work well, and it is hard to read actual values or see the main relationships. I do not think these 3D figures are indeed necessary, and I would suggest replacing most of them with clearer 2D figures.

We agree. In the revised manuscript, we will replace most of the 3D surface and scatter plots with clearer 2D figures. For the sensitivity analysis, we will use 2D heatmaps by index, threshold, and timescale. For exposure and impact relationships, we will use 2D scatter plots with clearer axis labels, units, and correlation annotations.

Please standardise the spelling/abbreviation of the Emergency Events database throughout the manuscript. Currently, it appears inconsistently as EMDAT, EM-DAT, and EM DAT.

In the revised manuscript, we will standardize the spelling as EM-DAT throughout the manuscript.

Lines 102: Please double check the reference (Kageyama and Sawada 2024). You have got it down as 2024 here, but the reference list gives it is 2022 (Line 682).

In the revised manuscript, we will correct the year to 2022.

Lines 146 and 568: A similar issue occurs with (Kummu et al. 2023) The main text says 2023, but the reference list has it pegged at 2025 (Line 682).

In the revised manuscript, we will correct the year to 2025.

Line 147: Please pick either USD or US dollars and stick with it throughout the manuscript.

In the revised manuscript, we will standardize the monetary unit notation throughout the manuscript.

Line 484-486: I am a bit lost on how you pulled that conclusion directly from Figure 1.

We thank the reviewer for pointing this out. This was an incorrect figure reference. The statement was intended to refer to the exposure results shown in Figure 5, not Figure 1. We will correct this in the revised manuscript and check all figure references for consistency.

Line 570-571: Please double-check this link here.

In the revised manuscript, we will double-check the link.

There are a few spelling errors remain, see examples in Lines 386, 430, 469, 515.

In the revised manuscript, we will carefully proofread the revised manuscript and correct the spelling errors noted by the reviewer, including the examples mentioned in lines 386, 430, 469, and 515.