Response letter of hess-2022-124-RC2

Dear Mariana Madruga de Brito,
Please find the responses to the comments.
Comments made by the reviewer were highly insightful. They allowed us to greatly improve the quality of the manuscript. We described the response to the comments.

Each comment made by the reviewers is written in italic font. We numbered each comment as (n.m) in which n is the reviewer number and m is the comment number.

Sincerely,
Yuya Kageyama and Yohei Sawada

Responses to the comments of Referee #2

In this manuscript, the authors use the new GDIS data and link it to hazards (i.e. soil moisture). The contribution is timely and innovative. I enjoyed reading the paper. Currently, we have little understanding of the relationships between impacts and hazard data. Hence, the proposed approach provides a way to understand these linkages. I particularly appreciate that instead of the traditional approach (e.g. hazard x vulnerability x exposure = risk), the authors go from the disaster itself, and from it try to identify the hazard drivers.

The manuscript is well written, yet, some issues need to be addressed:

→ Many thanks for the comments and suggestions.

(2.1) *My main criticism is that the authors claim to have assessed the "regional vulnerability to drought" (see line 258). Figure 5 shows that the SDI alone cannot explain drought disasters. Indeed, in Sub-Saharan Africa and South Asia, the SDI is not as high as in Europe and North America, yet, many disasters are observed. One explanation for this is indeed, the socio-economic vulnerability in these regions is higher. However, it could also be that it is the population exposure. Since you have not accessed any of these, you cannot make claims regarding the vulnerability.*

→ We appreciate this comment. As the reviewer mentions, we did not directly include exposure, and "regional vulnerability to drought" was overstated. There are several difficulties in treating exposure in our study. To estimate exposure, it is necessary to identify what has been damaged (e.g., people, crops, forests, etc.). However, EM-DAT provides no information on impact types (e.g., water shortage, famine, wildfire, etc.) in many drought events, which inhibits the identification of what has been damaged. In addition, EM-DAT provides no information about the amount of damage in multiple drought events (line 329). Therefore, we will change "regional vulnerability to drought" to "proxy of regional vulnerability to drought". In addition, we will explicitly say that our analysis did not directly include exposure in the discussion (See also our response to the comment 2.7). In the abstract (line 11), we will change "Our comparison between GDIS and ERA5-Land can quantify vulnerability to drought" to "Our comparison between GDIS and ERA5-Land could benefit the quantification of vulnerability to drought."

(2.2) *You mentioned that you disregarded drought events that lasted for less than 2 months. However, I assume most of the EM-DAT data only has a year, and few have a start and end date. In such cases, you mentioned that you considered the whole year as a drought year. That can be very problematic, as probably many of these events could actually be 2 months. It would be great if you could bring this limitation in the discussion section, where you already mentioned the limitation of the impact data.*

→ We fully agree with this comment. There are some GDIS drought events in which the details of drought duration information are not provided. Therefore, drought events shorter than two months may be included in our analysis, although we intended to exclude them. We will add this limitation in the discussion section from line 331. "Although we excluded GDIS drought events shorter than two months from our analysis, some of the analyzed events might be shorter than two months. This is because we applied January for the start and December for the end of the event if the start and/or end months of events shown in EM-DAT were unclear."

(2.3) *Why was landcover data from 2020 used (line 127) if the drought events go from 1964 to 2018? Likewise, why did you consider soil moisture data from 2019-2020 (line 132) when your impact data is from previous years?*

→ The detailed landcover datasets such as MODIS landcover (used in our study), GlobCover, and Global Land Cover Characterization (GLCC) do not provide landcover data over a long period of time. The MODIS land cover dataset starts in 2001. In addition, we only used landcover data to exclude the barren or sparsely vegetated areas from drought-prone areas by the drought clustering method, so that the timing of land cover does not significantly change our results. We confirmed that there was little or no difference in the extent of the barren or sparsely vegetated areas when we compared with the oldest 2001 data and the latest 2020 data. We will not change the manuscript at this point.

We used (absolute) soil moisture data during 1950–2020 in calculating the percentiles. After the percentiles were calculated, only (percentile) data during the period of the GDIS drought events (1964–2018) was used (line 131). We expected that a longer data period would contribute to more robust percentile values. For calculating SPI, at least 30 years of data is required (McKee et al., 1993). We also used percentile values which were calculated from absolute values during 1964–2018 (the same period with the GDIS drought events) and the results were slightly better when used percentile values which were calculated from absolute values during 1950–2020; the differences of drought indices during the GDIS drought period and the whole period was bigger and hydro-meteorological drought-prone areas were more consistent with GDIS drought-prone areas (not shown). We will add a comment from line 133. "We used the longer period of original ERA5-Land data (1950–2020) than the study period (1964–2018) to yield more robust percentile values."

(2.4) *Item 3.3 and 4.2. Similar to the first comment, stratifying your SDI data by different regions is not equivalent to "understanding the vulnerability in each region". What you did here was to analyze the SDI according to different regions. This is definitely an important analysis, but it is not a vulnerability analysis. You can hypothesise that the vulnerability causes these differences, but you cannot confirm this with the present data. The statements in line 255 are pretty strong and cannot be made without considering the socio-economic vulnerability. I suggest toning them down and writing that "the vulnerability could explain these differences". However, for this, further analysis is required.*

→ We appreciate this comment. Relating to the first comment (2.1), we will tone down like "vulnerability could explain these differences." We will also change the subtitle "Regional vulnerability to drought" to "Regional levels of drought indices associated with GDIS drought events". The detailed comment on exposure will be described in the discussion section (See also comment 2.7). Below is the revised version of these sections.

3.3 Regional levels of drought indices associated with GDIS drought events

The levels of hydro-meteorological drought indices associated with socio-economic drought events identified in GDIS are different in different regions. Vulnerability could explain these differences (Delbiso et al. 2017; Gasparrini et al. 2015; Tschumi and Zscheischler, 2020) Please note that vulnerability is not the only explanation for these differences; exposure is another factor that influences the linkage between hazards and impact (Visser et al., 2014; see also the discussion section). Since we did not directly include exposure, we recognized these differences as "the proxy of vulnerability". Following Bachmair et al. (2016), the levels of SDI which are associated with socio-economic drought events in GDIS were quantified and analyzed. The levels of SDI were stratified by geographical regions to understand the distribution of the proxy of vulnerability in each region.

4.2 Regional levels of drought indices associated with GDIS drought events

Figure 5 shows ~. A large number of small SDI events indicates that less severe hydro-meteorological droughts have caused serious socio-economic impacts, meaning that the regions are vulnerable to drought. On the other hand, the regions with a large number of large SDI events can be recognized as robust regions to drought. Thus, Sub-Saharan Africa and South Asia are vulnerable to drought, while North America and Europe are less vulnerable to drought. This regional characteristic of vulnerability to drought can be found when SDI is generated by soil moisture in different soil layers (not shown). Note that Middle East & North Africa were excluded from the analysis because the sample size was too small (n = 4).

In the introduction (line 72), we will change "we quantified the vulnerability to drought in different geographical regions" to "we quantified the levels of drought indices associated with GDIS drought events in different geographical regions, whose differences could benefit the quantification of vulnerability to drought".

(2.5) *Item 3.4. It is not clear to me what the advantage of clustering is. Why would the results be different with or without clustering? Could you please elaborate?*

→ The idea of drought clustering is to check whether spatially large hydro-meteorological droughts typically lead to disasters that can be seen in GDIS. For more detail, please see our responses to the comment of the reviewer #1 (comment 1.6).

(2.6) *In line 259, you say that your findings are consistent with previous ones that used GDP, however, you have not used GDP here. Please rephrase the sentence, or do the analysis*

→ We intended to say that high-income countries were associated with lower vulnerability to drought. We did not use GDP, as the reviewer mentions. We will change this sentence like "Previous studies have shown that higher GDP per capita is associated with lower vulnerability to natural hazards (e.g., Kim et al., 2019; Tanoue et al., 2016). North America and Europe are high-income countries, and these previous works support our findings."

(2.7) *In line 266, you write that the vulnerability is explained by the lack of structural measures. Sure, these are contributing factors, but socio-economic variables and exposure could be even more important. I understand the author's point of view, being an engineer myself. However, "command and control" approaches in which drought risk is controlled solely with conventional engineering measures is an outdated view. The authors cannot ignore the vast literature on socio-economic vulnerability and risk perception. You do not need to cite these papers, but have a look at the GAR chapters on vulnerability and droughts:*

https://gar.undrr.org/sites/default/files/chapter/2019-05/Chapter_3.pdf

https://gar.undrr.org/chapters/chapter-6-special-section-drought

→ We appreciate this implication. We stated the complex nature of vulnerability in the discussion sections (from line 319) and included non-structural countermeasures. We will strengthen the content including exposure in this section (from line 323), "Exposure is another important factor that influences the linkage between hazards and impact (Visser et al., 2014). The inconsistency between hydro-meteorological drought-prone areas and drought-prone areas found in GDIS in Namibia implies that the quantification of exposure is necessary to strengthen the analysis in our study. It is necessary to identify what has been damaged (e.g., people, crops, forests, etc.) to quantify exposure. However, EM-DAT provides no information on impact types (e.g., water shortage, famine, wildfire, etc.) in many drought events, which inhibits the identification of what has been damaged. Like vulnerability, exposure is complex and dynamic, for example, affected crops change with the seasons (Bodner et al., 2015). To improve our analysis on vulnerability shown in Section 4.2, detailed analyses on the complex and dynamic nature of both vulnerability and exposure are necessary."

(2.8) *Line 282: Spain is not shown in Figure 5. Hence, you cannot make this statement based only in the Figure. It could well be that the other European countries are "robust" and push the mean up, whereas Spain pushes the mean down. I suggest removing the sentence or doing the analysis for Spain only.*

→ Spain is actually less vulnerable to drought. We will add the mean value of SDI for Spain in the manuscript.

(2.9) *Line 301: I find it good that you can find good linear correlations. However, I was surprised as soil moisture and drought impacts usually do not have a linear relationship. Indeed, there are non-linear relationships (there is* *a lag between soil moisture and forestry impacts for instance). Furthermore, you have multi-year drought events, which add complexity to the interpretation of spatial and temporal differences in drought prevalence. Indeed, it could be that a drought is not so severe in biophysical terms. However, because they last so long, the impacts are high. I assume you got good correlations because EMDAT usually do not have long term impacts but rather focuses on immediate ones. Please add a discussion of these issues in your discussion section.*

→ We fully agree that the relationship between soil moisture and drought processes is non-linear. As the reviewer mentions, there are non-linear relationships between the drought severity and drought stress in vegetation (e.g., Chen et al. 2020; Meyer et al., 2014), where damage increases suddenly when the drought severity exceeds a certain critical threshold. Note that we found no "good linear correlations", nor did we explore any linear relationships in this paper. We did not explicitly describe that our analyzed relationship is linear. It is currently unclear for us which part of our analyses the reviewer recognized as a good linear correlation. Maybe we can improve our responses to this comment if the reviewer clarifies this point in the next round.

We agree that there is a lag between hazards and impact. Whether EM-DAT captures the long-lasting drought impacts is an important issue. The reason why drought events shown in GDIS were generally represented by drought hazards quantified from ERA5-Land may be due to the relatively long period of the drought duration in GDIS; the mean duration is approximately 12 months. Anyway, drought phenomenon is complex; the impacts of drought last even after the hydro-meteorological drought ends. Further analysis is needed to focus on the chronological correspondence to drought hazards. We will add the following discussion from line 318 emphasizing the complex nature of the relationship between soil moisture and drought impacts.

"The relationship between hazards and impact is much more complex than addressed in this study. Many studies have revealed the non-linear relationships between the drought severity and the reduction of vegetation growth (e.g., Chen et al. 2020; Meyer et al., 2014), where damage increases suddenly when the drought severity exceeds a certain critical threshold. On the other hand, de Brito et al. (2020) reported that there was a linear relationship between the drought severity and the number of drought articles as a proxy of socio-economic drought impacts. In addition, drought is a long-lasting disaster and there is a time-lag between hazards and impact, so that the period of hydro-meteorological drought is not necessarily consistent with the period considered as a disaster in EM-DAT. Some studies have revealed that the impacts of drought last even after the hydro-meteorological drought ends (e.g., Shahbazbegian and Bagheri, 2010). Yokomatsu et al. (2020) tried to capture the impact of the drought in terms of the economic development after the drought. In any case, further analyses are needed to focus on the chronological correspondence to drought hazards."

(2.10) *Line 9: What is meant by "drought information" here? Do you mean Drought impact information (i.e., the GDIS data?). Please be more specific.*

→ This is drought impact information as shown in GDIS. We will change the language.

(2.11) *Line 55-57: I do not understand the purpose of these 3 last sentences. When you write about vegetation growth, do you mean forest or crops? Regarding the last sentence. Why do we need to treat drought impacts as they are socially perceived? If you think the section is relevant, please expand it to make it clearer to the reader.*

→ Vegetation means plants in general, both forests and crops. We intended to say that not all droughts affect vegetation growth, and not all vegetation decline is caused by drought. "Socially perceived as drought" implies that we should consult a disaster database, which shows events in which the society has actually suffered from drought (we say it as "socially perceived"). We will rephrase like, "It is unclear whether socio-economic drought impacts are associated with declined vegetation growth. It is ideal to treat the socio-economic drought impact based on a disaster database since it directly shows events in which the society has actually suffered from drought."

(2.12) *Line 95: This information is repeated in the introduction. I suggest removing it: "GDIS is the geocoded…."*

The next sentence is about EM-DAT, which made us say "GDIS is the geocoded disaster locations database based on EM-DAT". We will rephrase like "GDIS is generated based on EM-DAT."

(2.13) *Line 122: You do not "show the drought vulnerability". Land cover is an exposure data, not vulnerability.*

→ We will change "show the drought vulnerability" to "show the levels of drought indices associated with GDIS drought events."

We intended to use land cover data as exposure data. We will change it like "As a proxy of exposure data, we used the MODIS land cover~."

(2.14) *Line 207: "significantly higher": could you please provide the test results that made you reach this conclusion.*

→ The test results are below the table. In any case, the values are well below the threshold we set in *K–S* test ($p<0.01$). We will say "significantly higher ($p<0.01$)" in the revised version of the paper.

Table R1: The results of *K–S* test between the GDIS drought period and the whole period.

	DAP	SDI
First (0–7 cm)	2.5e-40	8.4e-33
Second (7–28 cm)	1.0e-21	9.3e-22
Third (28–100 cm)	3.2e-10	4.6e-15
Root zone (0–100 cm)	1.6e-17	2.7e-18

Additional References

Bodner, G., Nakhforoosh, A. and Kaul, HP.: Management of crop water under drought: a review, Agron. Sustain. Dev., 35, 401–442, https://doi.org/10.1007/s13593-015-0283-4, 2015.

Delbiso, T. D., Altare, C., Rodriguez-Llanes, J. M., Doocy, S., and Guha-Sapir, D.: Drought and child mortality: a meta-analysis of small-scale surveys from Ethiopia, Sci. Rep., 7, https://doi.org/10.1038/s41598-017-02271-5, 2017.

Gasparrini, A., Guo, Y., Hashizume, M., Lavigne, E., Zanobetti, A., Schwartz, J., Tobias, A., Tong, S., Rocklov, J., Forsberg, B., Leone, M., De Sario, M., Bell, M. L., Guo, Y.-L. L., Wu, C., Kan, H., Yi, S.-M., Zanotti Stagliorio Coelho, M. de S., Nascimento Saldiva, P. H., Honda, Y., Kim, H., and Armstrong, B.: Mortality risk attributable to high and low ambient temperature: a multicountry observational study, The Lancet, 386, 369–375, https://doi.org/10.1016/S0140-6736(14)62114-0, 2015.

Hayes, M. J., Wilhelmi, O. V., and Knutson, C. L.: Reducing drought risk: bridging theory and practice, Nat. Hazards Rev., 5(2), 106–113, 2004.

Meyer, E., Aspinwall, M. J., Lowry, D. B., Palacio-Mejía, J. D., Logan, T. L., Fay, P. A., and Juenger, T. E.: Integrating transcriptional, metabolomic, and physiological responses to drought stress and recovery in switchgrass (Panicum virgatum L.), BMC genomics, 15(1), 1–15, 2014.

Shahbazbegian, M., and Bagheri, A.: Rethinking assessment of drought impacts: a systemic approach towards sustainability, Sustain. Sci., 5, 223–236, https://doi.org/10.1007/s11625-010-0110-4, 2010.