https://doi.org/10.5194/nhess-2022-174
Nat. Hazards Earth Syst. Sci. Discuss.
"Analyzing the informative value of alternative hazard indicators for monitoring drought risk for human water supply and river ecosystems at the global scale" by Claudia Herbert and Petra Döll

Response to Editor

Dear Ms. Van Loon,

We are very grateful for the many helpful, informative comments and the constructive and very detailed suggestions that you have provided. We thank you very much for your time and effort to help us improve the manuscript.
Below, each of your comments (in italics, indicated by "**EC**") is followed by our answer (normal font, indicated by "**AC**"). Changes in the manuscript are written in bold.

EC: *Thanks for the revised version of the paper. I appreciate the effort and also the reviewer who looked at the paper again sees improvement. However, some issues indicated by the reviewers were not addressed sufficiently and the new reviewer suggests to reject the paper because they find it unreadable. I still want to give you another opportunity to improve the paper, because I see value in the classification of indicators that you are proposing and I think there are some interesting findings in the paper. But I do need you to more thoroughly re-work the manuscript. Otherwise I will not be able to accept the paper for publication in NHESS.*

> **AC**: Thank you for the positive feedback. We shortened the manuscript by ten pages, rephrased many sections, and restructured the results section following your first suggestion (see below).

EC: *One option is to move a lot of material to supplementary material, for example Section 3.3, 4.1 and 4.2.5, and to focus the paper much more on the core message and its illustration. A suggestion then could be to restructure the paper so that the illustrations better fit the message, in a way that you make a few subsection that each have a different message based on a comparison of different indicators and in each you use a different visualization (a global map at a specific point in time or a time series in a selected grid cell of a selection of indicators) to show the point you are making in that subsection.*

> **AC**: We moved part of Sect. 2.1 (suitability of WaterGAP to model drought hazard indicators) and Section 4.1 (model validation) to the supplement. We deleted large parts of Sect. 3.3 and we provide results for the (new) gauging stations in the supplement. We excluded the indicators SPI and SPEI from our analysis and deleted Sect. 4.2.5 (Suitability of SPIn to quantifiy streamflow drought hazard).
>
> Sect. 3 now only describes the theoretical basis for the suggested classification system, which is then exemplified in Sect. 4 using a) modeled output from WaterGAP and b) observed monthly streamflow at four selected gauging stations. As suggested, Sect. 4 is now divided into subsections with different core messages illustrated using different figures.

EC: *Another option is to divide the paper in two separate ones. One opinion paper on habituation and introducing the classification (Section 3.1 and 3.2). And one research paper on a global comparison of hazard indicators (referring to the suggested classification). In the latter paper it would then be good to show more maps of other droughts years, for example selecting months that had drought hotspots in each of the continents (so one selected month that was a drought hotspot in South America, one in North America, one in Africa, etc.), because processes and therefore indicators can be very different.*

> **AC**: We now illustrate the core messages using four gauging stations as well as modeling results of two selected months of the reference period. We describe the approach at the beginning of Sect. 4 as follows:
>
> **The objective of this chapter is to identify which of the SDHIs presented in Table 1 can be meaningfully quantified at the global scale using WaterGAP 2.2d and which SDHIs are appropriate for monitoring**

different drought risks in large-scale DEWS. We emphasize that the objective is not a drought impact assessment, which is beyond the scope of this study. We want to show how the conceptual discrepancies and similarities between SDHIs (Sect. 3), which are of a general nature and apply to any month of the reference period, are translated into global-scale hazard indicators and how these indicators should be interpreted by end-users of a large-scale DEWS. The indicators are illustrated in global maps for two example months, July 2003 and July 2015, with known drought events in Europe. Following the classification of Table 1, SDHIs indicate drought magnitude (Figs. 2 and S3) or drought severity, the latter either expressed as volume-based anomaly or deficit (Figs. 3 and S4) or as frequency of non-exceedance (denoted with the suffix "_f") (Figs. 5 and S5). In addition, CQDI1(Q80) and CQDI1(Q80-HS) are compared at the global scale with respect to drought occurrence during the whole reference period (Fig. 4). SDHIs are further illustrated for four selected gauging stations with different streamflow regimes and assumed vulnerabilities of the risk system to streamflow anomalies (Figs. 6, S2 and S6). These include two stations with low interannual streamflow variability (Danube River at Hofkirchen, Germany (probably low vulnerability), and Angara River at Boguchany, Russia (possibly higher vulnerability)) and two stations with high interannual variability (White River near Oacoma, U.S. (probably low vulnerability), and Orange River at Vioolsdrif, South Africa (possibly higher vulnerability)).

EC: *Please consider a thorough restructuring of the paper(s) and also address the reviewers comments. And also clarify these additional point from my side:*
- There still is some confusing about the methodology (see reviewer #1).

AC: The description of SPI and SPEI was deleted, since the indicators are not analyzed any more.

EC: *- There are some explanations of methods or methodological choices in the Results sections, for example, what data was used to calculate the indicators (l.119-120: "Hydrological drought hazard indicators were computed using output from the global water availability and water use model WaterGAP2.2d" and l.470: "We used observations instead of WaterGAP modelling result to exclude model uncertainties."). Clarify this in the Methods section and add "in this subsection" to l.470.*

AC: We added a new chapter at the beginning of the methods section, which describes the streamflow data used in this study (streamflow observations and modeled streamflow). The section with line 470 was deleted.

EC: *- l. 503: "Performance of SPI12 and SPEI12 is equally low" > Avoid phrasing like high or low performance, realistic, good, etc. throughout the manuscript since you are not comparing to objective drought observations or impacts that can tell you which indicator is best. Your point is that they are different and should be picked carefully because they represent different assumptions about habituation, so just discuss the implications of choosing a different indicator.*

AC: We replaced these phrases throughout the manuscript. The comparison of SDHIs with SPI and SPEI was deleted due to the length of the manuscript.

EC: *- Reservoirs and SSI12 / CQDI1-Q80-HS: WaterGAP has reservoirs included, so why accumulate streamflow and why not take into account actual flow in dry areas? If the model would simulate natural flow, then it would make sense to use SSI12 in regions where water is stored in reservoirs, and to calculate CQDI1-Q80-HS ("drought to continue in any month where Q80 is zero also if the current streamflow Q exceeds zero"), but since WaterGAP already has reservoirs, aren't you double counting?*

AC: Thank you for pointing it out. We addressed this issue in the recommendations for reservoir managers in Sect. 5 and the description for Table 2 (previously Table 3):

Line 613: ** **Reservoir managers should be informed to consider SDHIs of the grid cell that represent inflow into the reservoir.**

Line 667: **Importantly, reservoir managers should only consider SDHIs of the grid cell that represent inflow into the reservoir.**

EC: *- The selection of grid cells and month: you now say: "March 2002 was selected as it was among the months with the highest difference between CQDI-Q80 and CQDI-Q80-HS" & "two illustrative grid cells with the same CQDI-Q80 value in March 2002" >> These are not valid justifications. To avoid this problem, I would suggest to add more months and locations (either in the paper if you split the paper in two separate ones, or in Supplementary material), for example the months with the maximum drought in each of the continents.*

AC: We now compare July 2003 and July 2015 and explain the approach at the beginning of Sect. 4 (**The objective of this chapter is to identify […]**) (see paragraph in bold on page 1 and 2 of this document). The results for July 2015 are provided in the supplement.

The comparison of the "two illustrative grid cells" was deleted due to the length of the manuscript.

EC: *- l.784-799: also here there is methods justification in the Results section. Please move.*

AC: The whole section was deleted.

EC: *- Section 4.2.5: You calculated correlation between simulated SSI and different SPIn with observed SSI for the basins in which the model was also calibrated. > You should have used other observed stations for validation. Please change and/or move to Supplementary material.*

AC: The whole section was deleted.

https://doi.org/10.5194/nhess-2022-174
Nat. Hazards Earth Syst. Sci. Discuss.
"Analyzing the informative value of alternative hazard indicators for monitoring drought risk for human water supply and river ecosystems at the global scale" by Claudia Herbert and Petra Döll

Response to Anonymous Referee #1, report 2

We thank you very much for your helpful comments and constructive suggestions for improving the manuscript. Below, each comment (in italics, indicated by "**RC**") is followed by our answer (normal font, indicated by "**AC**"). Changes in the manuscript are written in bold.

RC: *I thank the authors for their efforts made to improve the manuscript. However, I think that for the sake of clarity and to increase the impact of the article, it could still be reduced and simplified a little. Below some general points that need to be further clarified.*

> **AC**: We shortened the manuscript by ten pages and restructured and rephrased many sections to improve the readability and to clearly communicate the core messages and recommendations. Many sections were moved to a new supplementary material. To exemplify the suggested indicator classification system, we now use four gauging stations as well as two months of the reference period (July 2003 and July 2015) with known drought events in Europe. Many figures were moved to the supplement to keep the manuscript as short as possible.

RC: *Section 2.2.1 still needs to be reviewed. Please, clarify whether the standardised variables have been calculated by adjusting the distribution or by using the z-score as a simplification.*
The approximation of precipitation anomalies as a deviation from the mean is not the best approximation given that precipitation in general is non-normal, for shorter aggregation periods as well as in the context of semi-arid regions or regions with marked dry seasons where the authors focus part of the analysis.

> **AC**: Due to the length of the manuscript, the two indicators SPI and SPEI were removed from the analysis.

RC: *The rationale given for the selection of sites and month of analysis shows that the analysis focuses on a drought indicator modelling study disconnected from a risk analysis. This approach is correct, but what I had stressed in my comment is that in order to make recommendations on how to properly introduce this information into risk analysis, more information and analysis is needed.*

For instance, the authors mention in their reply "…that this kind of information is not available at the global scale and we recommend providing different hazard indicators covering different habituations to the streamflow regime. Then, people with local knowledge on this type of information can decide, which hazard indicator fits best to the targeted risk".
Here I understand that local knowledge includes the introduction of other variables that are defining drought risk, more analyses, etc. This position seems to be a simplification which historically has not helped the evolution of the concept of drought risk.
I encourage the authors to leave all necessary caveats open and to quantify appropriately here the nature of their results.

> **AC**: In different paragraphs throughout the manuscript, we try to highlight that the manuscript is focused on drought hazard, and that an impact assessment is beyond the scope of this study. For example, to underline that the figures are not used for an impact assessment, we added the following sentence in the introduction (line 83 in the new version):

This new methodology is exemplified at the global scale for eight existing and three newly developed SDHIs using a) modeled output from the global water resources and use model WaterGAP2.2d and b) observed monthly streamflow at four selected gauging stations.

AC: At the beginning of Sect. 4 (Similarities and discrepancies in SDHIs as quantified by a global hydrological model), we address this issue as follows:

The objective of this chapter is to identify which of the SDHIs presented in Table 1 can be meaningfully quantified at the global scale using WaterGAP 2.2d and which SDHIs are appropriate for monitoring different drought risks in large-scale DEWS. We emphasize that the objective is not a drought impact assessment, which is beyond the scope of this study. We want to show how the conceptual discrepancies and similarities between SDHIs (Sect. 3), which are of a general nature and apply to any month of the reference period, are translated into global-scale hazard indicators and how these indicators should be interpreted by end-users of a large-scale DEWS.

AC: We conclude the manuscript with the following suggestion for future research:

We suggest that future studies analyze how well these hazard indicators, in combination with suitable vulnerability and exposure indicators, can estimate drought impacts in the targeted risk systems at regional or national scales.

AC: In this study, we give recommendations on how to select and interpret drought hazard indicators from a global model, a topic that has not been covered so far in such detail. From our point of view, this focus on the hazard aspect is not a simplification but a necessary delimitation of the topic.

RC: *The explanation for the choice of March 2002 is not entirely clear. It is also unclear why it is relevant that the two chosen points present the same value for the CQDI-Q80 in a given month and year. Why not using the points already presented in Figure 2. Please further elaborate for the sake of clarity.*

AC: To exemplify the suggested indicator classification system, we now use four gauging stations as well as two months of the reference period (July 2003 and July 2015) with known drought events in Europe. We emphasize in Sect. 4, however, the generality of the results:

We want to show how the conceptual discrepancies and similarities between SDHIs (Sect. 3), which are of a general nature and apply to any month of the reference period, are translated into global-scale hazard indicators and how these indicators should be interpreted by end-users of a large-scale DEWS.

AC: We deleted the analysis of the grid cells in Italy and Paraguay (old Sect. 4.2.4 and old Table 2) with the same CQDI1-Q80 value.

https://doi.org/10.5194/nhess-2022-174
Nat. Hazards Earth Syst. Sci. Discuss.
"Analyzing the informative value of alternative hazard indicators for monitoring drought risk for human water supply and river ecosystems at the global scale" by Claudia Herbert and Petra Döll

Response to Anonymous Referee #4, report 1

We thank you very much for your helpful comments and constructive suggestions for improving the manuscript. Below, each comment (in italics, indicated by "**RC**") is followed by our answer (normal font, indicated by "**AC**"). Changes in the manuscript are written in bold.

General comments

RC: *The authors collected 12 drought indicators and discussed which of them are useful to display in large-scale drought early warning systems. The authors concluded "drought magnitude is best quantified by return period or relative deviation from mean, and severity by return period or water volume below a threshold relative to mean annual streamflow (from abstract)".*

Witnessing the recent frequent occurrence of severe drought events in many parts of the world, large-scale drought early warning is undoubtedly important. Reviewing and comparing drought indicators are also important, because it is widely recognized that drought is difficult to define or quantify. Although I fully understand the importance of the topic, I am unable to recommend publication of this work in the current from.

In short, the manuscript is unreadable. I have tried to go through this manuscript twice, but I couldn't complete it. Below I raise concrete examples why I am saying it is unreadable. Also I have to say that this paper is too long. In a nutshell, the authors' conclusions seem "drought magnitude is best quantified by return period or relative deviation from mean, and severity by return period or water volume below a threshold relative to mean annual streamflow (from abstract)". In my view, these conclusions are already well-perceived by hydrologists. I don't see any valid reasons why this long paper is needed to convey these unsurprising conclusions. Actually, I observe many paragraphs can be omitted. Again, I am not deprecating the authors' work. I just want to say that the manuscript is not ready to communicate with potential readers.

> **AC**: We shortened the manuscript by ten pages and restructured and rephrased many sections to improve the readability and to clearly communicate the core messages and recommendations. Many sections were moved to a new supplementary material.

Specific comments

RC: *Line 29-33 "Drought poses…": Likely this paragraph can be omitted. The concepts of "hazard", "exposure", and "vulnerability" seem to appear only in the next paragraph and play marginal role in this work.*

> **AC**: We deleted this paragraph.

RC: *Line 34-52 "Drought risk indicators…": This paragraph contains a lot of information, but in my view, it is undirected. It really puzzled me what is the point of the authors. Finally, I noticed that I can better understand just ignore this paragraph and proceed the next paragraph.*

> **AC**: We deleted lines 34-44. We think that the second part of this paragraph focusing on current regional to global-scale DEWS is relevant for the manuscript.

RC: *Line 62-74 "Streamflow drought hazard can be estimated…": Again, the paragraph includes a lot of information,*

but it is too specific. Readers wants to understand the background and objective of the study in Introduction, not the details. I have to say, this paragraph should be also omitted.

AC: We deleted this paragraph.

Line 75-90 "SDHIs are commonly classified into ": the former part (general classification of drought indicators) is informative, but the latter part (explanation of the standardized streamflow indicator SSI) looks too specific.

AC: We deleted the sentence specifying averaging periods of SSI (lines 82-83). However, we think that the latter part of this paragraph is relevant as it describes why an improved classification system for drought hazard indicators is required from our point of view.

RC: *Line 91-105 "A further consideration in designing SDHI is how to conceptualize drouth in intermittent or highly seasonal streamflow regimes…": Again, the first sentence is okay, but the following part looks simply too specific and unorganized.*

AC: We deleted lines 94-97 and shortened the remaining paragraph. We now describe only one important paper addressing drought in intermittent streamflow regimes in two sentences.

RC: *Line 104-116 "This paper analyzes": I think these are the only fully understandable paragraphs of this section. I just want to see a literature review which is directly relevant to these paragraphs.*

AC: We considerably shortened the introduction to the most relevant aspects.

RC: *Line 136 "In several model intercomparison studies…": Too long. What the authors need to convey here is that WaterGAP is well validated, intensively compared with other models, and the results are usable for this analysis. I think this paragraph can be condensed into a few lines.*

AC: The whole paragraph was moved to the new supplementary material and the core messages are now summarized in this section in a few sentences.

RC: *Line 168-267: Here the authors provide lengthy explanation for 12 indexes. I believe readers first want to know the definition of each index here, but I found it quite difficult, because each part is structured differently and including too many trivial information. I think here definition of index (hopefully with equations), threshold, and the parameters (in particular the time window the authors chose) would far enough to proceed reading. The remaining information can be put into supplemental material or appendix of this paper.*

AC: The sections about SPI and SPEI were deleted, since the indicators are not assessed any more (due to the length of the manuscript). For the severity indicators, equations were added that define the computation of the respective deficit. The time period is included in a new section 2.1 describing the observed and modeled streamflow data used in this study.

We are aware that the detailed indicator description may seem trivial to experts in this field. However, after studying the literature on drought (hazard) indicators, definitions are sometimes imprecise, missing or wrong, and we feel that such a detailed description is valuable especially for people new to the field.

RC: *Line 295 "The choice of drought hazard indicators implies assumptions about the habituation of the system at risk": I was totally puzzled by this part and following discussion. First of all, what is "choice"? Who chooses for what? What does "imply" mean? Whose "assumptions"? The following lines do not answer any of these questions. Actually, hereafter, I felt myself reading a unpolished first draft.*

AC: We rephrased especially the first paragraph of this section to address the issues raised by the reviewer (see paragraph below in bold). The first line now reads **"The selection of drought hazard indicators for a**

DEWS requires a clear definition of […]". We think that it is now clear that it is the end-user and/or developer of a DEWS who chooses, or selects, hazard indicators.

We further write "**Consequently, the selection of an indicator requires a definition, often based on assumptions, about "what is normal or needed", i.e., to what the risk system is habituated to."** to introduce the terms "assumption" and "habituation".

The term "imply" generally means that an idea is involved without saying it directly. So in this case, an indicator implies a certain idea (in this case, an assumption) about what the risk system is habituated to.

Rephrased paragraph:

The selection of drought hazard indicators for a DEWS requires a clear definition of "the risk of what for whom". Drought hazard indicators are risk system-specific (Blauhut et al. 2021), and there is not one that fits all. Drought is usually conceptualized as anomaly ("less water than normal") and/or deficit ("less water than needed"). Consequently, the selection of an indicator requires a definition, often based on assumptions, about "what is normal or needed", i.e., to what the risk system is habituated to. In the case of streamflow, people and ecosystems are assumed to have adapted to certain characteristics of the flow regime. For example, if drought indicators are computed based on the calendar month-specific distribution of streamflow values, it is implicitly assumed that the risk system has adapted to the seasonality of streamflow. But also temporally constant thresholds, which have traditionally been used to define hydrological droughts (Stahl et al. 2020), are suitable for certain systems, e.g., for computing drought risk for electricity generation by thermal power plants, which require a certain minimum streamflow for operation.

RC: *Table 1: Tables in a paper is typically a summary of some specific paragraphs or sections. However, I couldn't figure out how Table 1 and text are related. Furthermore, the table includes a lot of subjective judgement (e.g. "a certain degree of" "suitable for" "better proxy for" without showing rationale. It doesn't help my understanding of discussion.*

AC: Table 1 summarizes for each indicator the assumed habituation, which is the topic of this section. We think that a summary table is valuable here to give a quick overview of the numerous indicators. In the table caption, we added the following sentence to explain the terms "a certain degree", etc.:

The general terms "a certain degree" or "a certain reduction" in the first column are specified in a drought assessment by selected thresholds for drought definition.

AC: The term "a better proxy for" is not used any more, since the indicators SPI and SPEI were deleted.

RC: *Line 319 "In hydrology, flow duration curves showing the fraction of the time …": I couldn't contextualize the sentence here. I couldn't see any logical flow. Many readers expect the first sentence of paragraph is a summary of the paragraph, but this is not case of this manuscript.*

AC: We deleted this first sentence and moved or deleted most of this paragraph. The first paragraph is now focused on a comparison between percentile-based indicators and relative deviations. Lines 326-334 describing the findings of Kumar et al. (2009) were deleted.

RC: *Line 335 "Similar to the example above from Kumar et al. (2009), the 20th streamflow percentile (or SSI1=-0.84) would correspond to a low relative streamflow deviation (e.g. -20%) in a humid region (low interannual variability) compared to a higher deviation (e.g. -50%) in a semi-arid region (high interannual variability)": Impossible to understand. Is this a specific result of Kumar et al. (2009)? Must the readers go through Kumar et al. (2009)? Does SSI1=-0.84, 20%, 50% have any specific meaning or just some example numbers? Actually, I got a similar impression for numerous lines hereafter. First I blamed myself about my carelessness, but soon I stopped it: simply the text is too unorganized to read.*

AC: We rephrased the paragraph as follows:

Utilization of percentile-based indicators (e.g., SSI12, SSI1, and CQDI1(Q80) in Table 1) implies that people in different climate regions and social systems are equally habituated to a certain interannual variability, which is most likely not the case. The 20th streamflow percentile (or SSI1 = -0.84) would correspond to a low relative streamflow deviation (e.g. -20%) in a humid region (low interannual variability) compared to a higher deviation (e.g. -50%) in a semi-arid region (high interannual variability). Hence, percentile-based indicators might underestimate streamflow drought hazard in semi-arid areas where people (and ecosystems, albeit possibly to a lower degree) are often more vulnerable to reductions in water availability.

RC: *Line 346 "In conclusion, percentile-based hazard indicators and relative deviation from the long-term mean or median should be used complementarily in large scale DEWS in combination with adequate vulnerability and exposure indicators to cover different drought risks": I am happy to see conclusion was given here. It would be even better if this sentence comes at the beginning of this paragraph.*

AC: We shortened and restructured the paragraph comparing percentile-based indicators and relative deviations to improve readability. We think that the sentence in line 346 is a conclusion drawn from the indicator comparison and it should therefore remain at the end.

RC: *Line 382 "This concept is not new…, Nevertheless, only a few …": Hard to read. What is the point?*

AC: We deleted the sentence in line 382.

RC: *Line 409 "The indicator types (columns in Fig1) include the volume-based anomaly, the standardized or percentile-based anomaly, and the relative deviation, all of which are described in the previous section": I couldn't find where exactly these are described. If these categories are important, better to define explicitly in prior.*

AC: We replaced "**all of which are described in the previous section**" by "**(Sect. 2.3)**". Sect. 2.3 (previously Sect. 2.2) comprises the indicator description.

RC: *Figure 1 Caption "Classification system" and diagram: What does "system" mean? Many times the verb "choose" appear in the diagram, but who chooses for what? What the arrows indicate? Again, the Figure 1 doesn't enhance my understanding at all.*

AC: We think that the term "classification system" (i.e., putting items into categories or groups based on their characteristics) is suitable here, since indicators are grouped based on their characteristics.

AC: We replaced the term "choose" by "select" in Figure 1. In the figure caption, we rephrased the following sentence: "**The dark grey boxes indicate decisions that have to be made when computing the indicators, e.g. which averaging period is selected**".

Since Sect. 3.1 now starts with "**The selection of drought hazard indicators for a DEWS**", we think that it is clear at this point that the end-user or developer of a DEWS "chooses" (or selects) indicators for the DEWS. Moreover, the second to last paragraph of the introduction also clarifies this aspect: "**This paper analyzes which SDHIs are suitable for assessing and monitoring drought risk for human water supply from surface water and for river ecosystems in large-scale DEWS. We propose a systematic approach to indicator selection […]**".