https://doi.org/10.5194/nhess-2022-174-RC1, 2022
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

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: *In this paper the authors present an approach for the selection and calculation of streamflow drought hazard indicators for monitoring human surface water supply and for river ecosystems. In doing so, the authors discuss and propose to consider the habituation of people and ecosystems to the streamflow regime. For this purpose, eight existing drought indicators are compared and quantified and three new ones are proposed based on the global hydrological model WaterGAP2.2d.*

This article has positive aspects, in particular the effort in modelling that the authors have made is remarkable, however it needs to be reviewed thoroughly as there are several points that need to be clarified and addressed by the authors. In short, the article has potential, but the authors need to make an effort to focus the analysis and make a readable manuscript. I hope that the comments below will help in this direction.

AC: Thank you for the positive feedback.

Structure of the article

RC: *First, the structure of the paper could be improved. It is very long, repetitive in some aspects, lacking clarity on the objectives and results, that should be better highlighted both in the description and conclusions sections.*

Similarly, the order of the sections does not follow a story, some sections could be shortened, removed, or moved to increase the readability. This structure surely makes it very difficult to read. For example, the introduction should be significantly shortened to focus on the description relevant to the objectives of this work. Section 3.4 should be moved into the methodologies section. In addition, there is an annex that includes only one figure that is quite relevant. I suggest including it in the main text, if the number of figures is an editorial requirement I suggest removing some other figures (e.g. Figure 9).

AC: We deleted and shortened repetitive or too detailed paragraphs throughout the text and used more concise language. We restructured and renamed several sections to improve the storyline of the manuscript. "Results and Discussion" (old Sect. 3) are now divided into Sect. 3 and 4, with Sect. 3 focused on the proposed systematic approach for selecting streamflow drought hazard indicators (SDHIs) and Sect. 4 comprising the global-scale assessment.

3 Proposed systematic approach for selecting and computing SDHIs (old Sect. 3.1)

3.1 Assumptions about habituation inherent in drought hazard indicators (old Sect. 3.1.1)

3.2 Levels of drought characterization (old Sect. 3.1.2)

3.3 Illustration of habituation-based classification approach (old Sect. 3.2, new title)

4 Quantification of global streamflow drought hazard by a global hydrological model (old Sect. 3.3, new title)

4.1 Model validation (old Sect. 3.4)

4.2 Discrepancies in drought hazard as quantified by different SDHIs (old Sect. 3.3, new title)

4.2.1 Drought magnitude (level 1) (old Sect. 3.3.1)

4.2.2 Drought severity (level 2) (old Sect. 3.3.2)

4.2.3 Drought severity expressed as frequency of non-exceedance (level 2) (old Sect. 3.3.3)

4.2.4 Relation between the various SDHIs (old Sect. 3.3.4)

4.2.5 Suitability of SPIn to quantifiy streamflow drought hazard (old Sect. 3.5, new title)

AC: We added the following paragraph at the beginning of the new Sect. 3 to clarify the objectives and the structure:

Wilhite und Glantz (1985) suggested distinguishing between a conceptual and an operational drought definition, with the former referring to the general qualitative concept of drought and the latter allowing for a quantitative drought characterization including onset, severity, termination, and spatial extent. In the following Sect. 3.1, aspects that relate to the conceptual drought definition are discussed comprising the description of the targeted drought risk and the system at risk. In particular, assumptions about the habituation of the system at risk to the streamflow regime are discussed, an aspect that is currently not taken into account or not made explicit in drought hazard studies. In order to translate these conceptual definitions into operational drought hazard indicators, a new classification system for hazard indicators is proposed in Sect. 3.2. The new systematic approach is illustrated in Sect. 3.3 for selected SDHIs using streamflow observations at two gauging stations with different streamflow regimes.

AC: We added the following paragraph at the beginning of the new Sect. 4 to clarify the objectives and the structure:

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. After a limited validation of modeled streamflow (Sect. 4.1), SDHIs of drought magnitude and severity are compared separately (Sect. 4.2.1-4.2.3) following the classification system presented in Fig. 1. The SDHIs are shown in global maps for a selected month (March 2002), as it is important to understand the relation between indicators at a certain point in time, especially for the application in DEWS, which are focussed on the current situation or the near future. As patterns of indicators depend on characteristic of the streamflow regime and water use that are temporally constant over the reference period, the reasons for similarities and differences between indicators can be deducted in any month of the reference period. March 2002 was selected as it was among the months with the highest difference between CQDI-Q80 and CQDI-Q80-HS. In addition to the analysis for the selected time step, the latter two indicators are compared at the global scale with respect to drought occurrence during the whole reference period. Discrepancies and similarities of the indicators are discussed in more detail for two illustrative grid cells with the same CQDI-Q80 value in March 2002 (Sect. 4.2.4). Finally, the suitability of SPI with different averaging periods to estimate streamflow drought hazard is assessed using streamflow observations from 218 GRDC gauging stations (Sect. 4.2.5). Based on this global-scale analysis and the proposed habituation-based classification approach, selected SDHIs are recommended for implementation in large-scale DEWS (Sect. 5).

AC: In the introduction, we deleted paragraphs and sentences that anticipated aspects of the proposed new classification system (e.g., lines 92-118) as well as repetitive and too detailed sentences (e.g., lines 141-147). Overall, we shortened the introduction by one page.

AC: In the introduction, we added the following sentence (in bold) to highlight the research gap with respect to drought hazard concepts:

[Clearly, the conception or selection of hazard indicators needs to take into account the habituation and thus vulnerability of the system at risk.] **However, investigations and guidance on how to select the optimal SDHI, considering both the targeted risk and the habituation of the system at risk to the streamflow regime, are missing.**

AC: In the introduction, we added the following sentence to highlight the research gap with respect to a hazard indicator classification:

Certainly, an improved classification of drought hazard indicators would facilitate a better understanding of drought characteristics and provide guidance in selecting appropriate drought hazard indicators.

AC: The first paragraph of the model validation (lines 823-843) describing the performance of WaterGAP in model intercomparison studies was moved to the end of the model description (Sect. 2.1). Here, we added an introductory sentence and a recently published study:

In several model intercomparison studies, WaterGAP was often among the best performing global hydrological models (GHMs). Kumar et al. (2022) assessed the ability of nine catchment-scale models and eight GHMs to simulate hydrological droughts in eight large catchments around the world. Comparing simulated and observed streamflow deficits and SSI1 (SRI) (their Tables 2 and 3), WaterGAP is among the two to three best performing GHMs with performance indicators (R^2 and Nash-Sutcliffe efficiency) comparable to those of the catchment-scale models.

AC: The second paragraph of the model validation (old Sect. 3.4, lines 844-850) was moved to the beginning of the new Sect. 4 (see new structure above) as suggested by referee #3. We added the new Fig. 3 below for a more detailed analysis of Q80. Furthermore, we added the paragraph below describing the results of another recent WaterGAP model validation. Hence, the model validation is now focused on Q80 and a SDHI (SSI3), which are both relevant for the subsequent global-scale analysis of SDHIs.

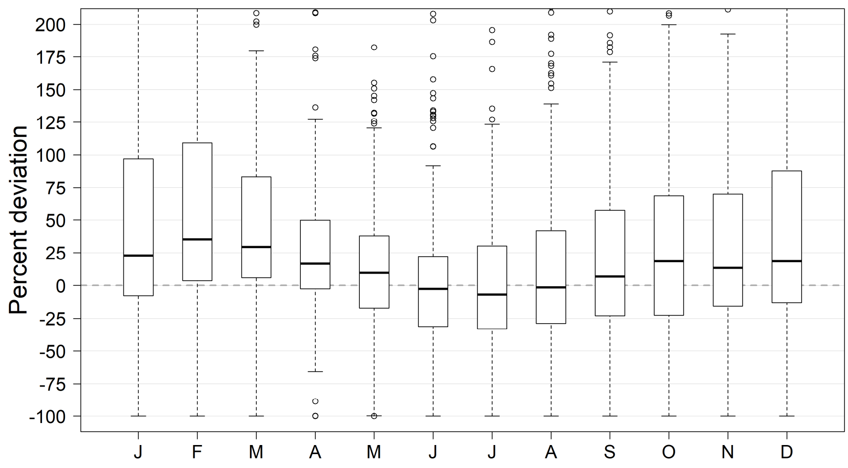

Figure 3: Percent deviations of simulated Q80 per calendar month from Q80 based on GRDC observations using the reference period 1986-2015.

In a recent study, WaterGAP 2.2d model output was validated against GRDC data by comparing SSI3 based on simulated and observed monthly streamflow (SSI3(sim) and SSI3(obs)) during 1971-2000 at 183 globally distributed GRDC stations (Wan et al. 2021). Applying drought hazard classes for SSI according to Agnew (2000), the agreement between simulated and observed hazard classes in each month was analyzed. Among all stations, the agreement ranged between 29 to 88% of all 360 months (their Fig. S4 and Table S3). At 68% of all stations (covering 83% of the assessed basin area), SSI3(sim) and SSI3(obs) resulted in the

same drought hazard class in 70 to 88% of the time. Moreover, the goodness-of-fit was evaluated based on the Nash-Sutcliffe efficiency (NSE) for monthly streamflow and SSI3 (their Fig. S3). With a median NSE of 0.5 and an interquartile range of 0.2-0.7 for SSI3 and 0.14-0.7 for streamflow, WaterGAP 2.2d model output showed a moderate agreement with the observations. Both NSEs exceeded 0.7 at 25 out of the 183 stations, which are located in Central and Eastern Europe (twelve stations), the United States (ten stations), and South Africa (one station).

AC: We think that Figure A1 should remain in the appendix, since it does not show indicators, and since we refer to this figure in almost all chapters.

AC: We deleted the old Fig. 9 as suggested.

AC: We shortened the conclusion section by almost one page. For instance, we deleted the last paragraph (lines 1033-1045) and summarized the recommendations as follows:

Out of the twelve analyzed SDHIs, we recommend a set of magnitude and severity indicators for large-scale DEWS specific to the risk systems 1) human water supply from surface water and 2) river ecosystems, distinguishing intermittent and perennial streamflow regimes as well as low and large water storage capacities.

RC*: The methodology can also be reduced, focusing on the description of the proposed new indicators and model validation. In addition, some aspects should be clarified, e.g., eq(1) for SPI, X is noted as a generic variable, if this variable is precipitation this representation could be misleading.*

AC: Given the inconsistent indicator descriptions identified in the literature (or the absence of any description), we think that the detailed description of the conventional indicators is as important as the description of the new indicators. We would like to keep equation 1 in this generic form as it can be used for any variable. However, we adjusted the description below the equation as follows:

with X = variable **(e.g. precipitation)**, μ = mean, and σ = standard deviation.

Focus of the manuscript and sectoral risk representation

RC: *The authors indicate that the focus of the article is on analysing indicators for monitoring drought risk in very specific sectors (human water supply and river ecosystems). I agree that the selection of the hazard indicator is key to determine this dimension of drought risk and the discussion on that direction is more than welcome and needed. However, apart from referring to these sectors as the focus of the article and several speculative and unsupported assertions, there is no information in this article on how drought specific indicators affect these sectors. The dynamics of how these sectors or systems are affected is surely complex, depending on various factors that determine their exposure and vulnerability beyond whether they rely on upstream reservoirs or the systems are seasonally dependent. However, this is not enough to characterize the vulnerability of these sectors. Further discussion and analysis in this regard is needed.*

On the one hand, considering the way the article is structured, orienting the analysis in the description and comparison between indicators, derived metrics, etc. in a concise and targeted manner can improve the focus and structure of the article. On the other hand, to strictly evaluate whether the proposed indicators are valid to represent risk, a more detailed analysis of the proposed sectors is needed, with a description of their vulnerabilities and how their impacts are produced as a consequence of the combination of the different dimensions of risk.

AC: WaterGAP 2.2d output was used in two drought risk studies for South Africa and at the global scale (Meza et al. 2021; Meza et al. 2020). We think that such and impact assessment is beyond the scope of this study. Especially the South Africa study showed that vulnerability and exposure indicators can only be meaningfully estimated at the regional to local scale. Such information is not available at the global scale. We do not make “speculative and unsupported assertions” about sectors. On the contrary, we write (e.g., in the conclusions, lines 1023-1029) 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. To clarify that hazard indicators are only one out of three components to estimate drought risk, we added the following sentence in the new Sect. 3.1 (Assumption about habituation inherent in drought hazard indicators):

In conclusion, percentile-based hazard indicators and relative deviations 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.

Grid cells – case study selection and description

RC: *The selection of the two grid cells seems to be motivated by characteristics derived from some of the modelled variables. However, no description of these sites exists in the manuscript. A quick search turns up that one gridcell is in central eastern Paraguay (perhaps including a portion of the Paraguay river) and the second near Firenze (Arno). Both points with very different realities regarding how they might be exposed and how they are vulnerable to droughts and surely each will have a very specific risk profile. Here I see a missed opportunity, as one of the objectives of this paper is to find out how the exposed systems can be used. Surely, a discussion along these lines would greatly improve the discussion on the usefulness of the various indicators. Indeed, the comparison between indicators is merely informative, which does not enable identification or validation between them (Please refer to my final comment) Similarly, March 2002 has been used to describe Figures 3 and 4. But it is not clear why this period was chosen or how this comparison can be extrapolated to the whole period.*

AC: With respect to the two grid cells, we rephrased the first sentence in the new Sect. 4.2.4 (see below). Discussion of the risk profile in the two grid cells is, from our point of view, beyond the scope of this study as described in our previous answer.

The relation between selected SDHIs from Figs. 4, 5 and 7 is compared for two illustrative grid cells that share the same CQDI-Q80 value of 0.04 in March 2002 (Table 2).

AC: With respect to the selected month March 2002, we added the following sentences at the beginning of the new Sect. 4:

The SDHIs are shown in global maps for a selected month (March 2002), as it is important to understand the relation between indicators at a certain point in time, especially for the application in DEWS, which are focused on the current situation or the near future. As patterns of indicators depend on characteristic of the streamflow regime and water use that are temporally constant over the reference period, the reasons for similarities and differences between indicators can be deducted in any month of the reference period. March 2002 was selected as it was among the months with the highest difference between CQDI-Q80 and CQDI-Q80-HS.

Comparison not validation

RC: *As proposed in this article, the fundamental purpose of any drought indicator is to represent the sectoral impacts in the best possible way. Indeed, the validation of the best indicator should be consistent in how it represent sectoral impacts. In this sense, it is perfectly legitimate to compare indicators, but it is not possible to validate or rank one over another without looking at independent variables that represent potential impacts. Please elaborate further on this issue.*

AC: We think that an impact assessment is beyond the scope of this paper (see above). However, we added this aspect as future research need as last sentence in the conclusions:

Since an impact assessment was beyond the scope of this study, future studies could 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.

References

Kumar, Amit; Gosling, Simon N.; Johnson, Matthew F.; Jones, Matthew D.; Zaherpour, Jamal; Kumar, Rohini et al. (2022): Multi-model evaluation of catchment- and global-scale hydrological model simulations of drought characteristics across eight large river catchments. In: *Advances in Water Resources* 165, S. 104212. DOI: 10.1016/j.advwatres.2022.104212.

Meza, Isabel; Eyshi Rezaei, Ehsan; Siebert, Stefan; Ghazaryan, Gohar; Nouri, Hamideh; Dubovyk, Olena et al. (2021): Drought risk for agricultural systems in South Africa: Drivers, spatial patterns, and implications for drought risk management. In: *The Science of the total environment* 799, S. 149505. DOI: 10.1016/j.scitotenv.2021.149505.

Meza, Isabel; Siebert, Stefan; Döll, Petra; Kusche, Jürgen; Herbert, Claudia; Eyshi Rezaei, Ehsan et al. (2020): Global-scale drought risk assessment for agricultural systems. In: *Nat. Hazards Earth Syst. Sci.* 20 (2), S. 695–712. DOI: 10.5194/nhess-20-695-2020.

Wan, Wenhua; Zhao, Jianshi; Popat, Eklavyya; Herbert, Claudia; Döll, Petra (2021): Analyzing the Impact of Streamflow Drought on Hydroelectricity Production: A Global-Scale Study. In: *Water Res* 57 (4). DOI: 10.1029/2020WR028087.

Wilhite, D.; Glantz, M. (1985): Understanding the drought phenomenon: the role of definitions. In: *Water International* 10 (3), S. 111–120. DOI: 10.1080/02508068508686328.