Replies to reviews

"Technical note: What does the Standardized Streamflow Index actually reflect? Insights and implications for hydrological drought analysis

Fabián Lema, Pablo A. Mendoza, Nicolás Vásquez, Naoki Mizukami, Mauricio Zambrano-Bigiarini, and Ximena Vargas

We provide responses to each individual point below. For clarity, comments are given in italics, and our responses are given in plain blue text.

Anonymous Referee #1

The paper demonstrates strong scientific significance, high quality, and effective presentation. The authors investigate hydrological droughts, focusing on Standardized Drought Indices (SDI), notably the Standardized Streamflow Index (SSI), as tools for understanding drought dynamics, including frequency, intensity, duration, and propagation. The SUMMA hydrological model and the mizuRoute routing model were calibrated and used to analyze six case study basins in the western extratropical Andes. This analysis explores the relationship between SSI and various explanatory basin-aggregated variables, such as precipitation and catchment storage, across different time scales.

The authors address the common use of SSI to quantify hydrological drought without clearly understanding its effectiveness in capturing the dynamics of drought propagation in basins with diverse hydrological regimes. This key issue is woven throughout the paper, making it enjoyable to read.

Although the authors work with a limited number of basins, their analysis has implications beyond the extratropical Andean region, particularly in operational contexts where regulators and decision-makers rely on simplified and easily accessible drought indices without further analysis or differentiation between hydrological regimes. The discussion section provides a solid overview of the approach's limitations and potential avenues for future research. The figures convey important information, making it easy to grasp the main findings.

We thank the referee for meticulously reviewing our manuscript and providing several constructive suggestions. We are especially grateful for the referee's positive feedback.

Minor comments:

1. *While the justification for using a hydrological model as a benchmark for evaluating SDIs over observational data is clear, I believe the paper would benefit from a more detailed discussion of the inherent limitations of using a model to represent the long-term behavior of drought in the chosen basins. Analyzing how the model's process representation might lead to inaccuracies in streamflow reproduction—especially in basins where the modeled minimum streamflow values exceed the observed values—could provide additional insight into the analysis.*

We have added the following text in section 4.1 regarding inaccuracies in low flow simulations (L257-L261):

"… there is an overestimation of low flow volumes with exceedance probabilities larger than 90% in the Choapa and Claro catchments (< 2 m^3/s), which could be explained by the inadequate model physics representation including, but not limited to the lack of a common aquifer enabling water exchange among grid cells in our SUMMA configuration, and/or biases in the forcing dataset that impact the accumulation and melting of snow."

In response to this comment and another reviewer's observation, we have added the following text in section 5.1 (L377-L382):

"Although the model's overestimation of low flow volumes in Choapa and Claro (Figure 3) affects the accuracy (i.e., closeness to reality) of the number and duration of detected events (Figure 5), this artifact does not alter our conclusions, as all analyses focus on the impact of methodological choices related to index calculations using simulated variables, regardless of the fidelity of model representations. Even more, all the correlation and drought propagation analyses were performed in the model's world and, therefore, streamflow biases should not impact the extent to which variables or drought indices computed with different time scales relate to each other".

2. *I suggest the authors include more information regarding the model's warm-up period or restrict the time results of their full simulation to account for this warm-up period.*

We did consider a spin-up period before computing all the performance metrics for hydrological model calibration and evaluation. We have clarified this point in L191-L193:

"The observed daily streamflow data is split into a warm-up period (April/2004 – March/2006), a calibration period (April/2010 – March/2017), and two non-consecutive evaluation periods (April/2006 – March/2010 and April/2017 – March/2020)."

We also consider a two-year warm up period before computing the standardized drought indices. We clarify this point in L217-L219:

"We use the calibrated parameters (see section 3.2) to perform hydrologic simulations for the historical period April/1981 – March/2020. All SDI computations consider a spin-up period of two years (April/1981 – March/1983) and the same reference period of 30 years (April/1983 – March/2013)".

3. *Lastly, I would appreciate it if the authors could elaborate on how they envision the design of regional analysis frameworks that consider more than just hydrological regimes, including similarities in physical features such as slope, elevation, soil properties, and land cover, among others.*

This is an interesting point. The analyses presented here could be expanded to a larger number of basins that consider a greater diversity of features (e.g., Vásquez *et al.*, 2021, Muñoz-Castro *et al.*, 2023), in order to examine whether the time scales of hydrological variables (e.g., precipitacion, soil moisture, SWE) that maximize the correlation (or 'optimal' time scales) with the SSI are related to physiographic attributes such as contributing area, slope, elevation, geology, land cover and soil type, among others. A simple stratification of attribute values by optimal time scale, or any other hydrological descriptor of interest (e.g., Sawicz *et al.*, 2011, Almagro *et al.*, 2024) could provide valuable insights, complementing previous drought investigations using large samples of catchments. For example, Van Loon and Laaha (2015) found that geology and land use were relevant controls for hydrological drought duration. Peña-Gallardo et al. (2019) concluded that elevation and vegetation coverage are the main factors controlling the diverse response of SSI to SPEI time scales. More recently, Brunner and Stahl (2023) confirmed that land surface processes are required to explain the temporal clustering of hydrological droughts. More generally, additional large-sample hydrology analyses could help to improve our understanding of the main drivers affecting drought occurrence and propagation across different hydroclimates. We will incorporate these ideas in the Discussion section (L447-L457).

L235-L245: Can the authors discuss the implications of using a model that overestimates the low flow volumes (exceedance probabilities over 90%) for analyzing hydrological droughts for the Choapa and Claro cases?

We thank the reviewer for making this point. Although the model' s overestimation of low flow volumes in Choapa and Claro (Figure 3) affects the accuracy (i.e., closeness to reality) of the number and duration of detected events (Figure 5), this artifact does not alter our conclusions, as all analyses focus on the impact of methodological choices related to index calculations using simulated variables, regardless of the fidelity of model representations. Even more, all the correlation and drought propagation analyses were performed in the model's world and, therefore, streamflow biases should not impact the extent to which variables or drought indices computed with different time scales relate to each other. We will clarify these points in section 5.1 ("Drought detection and characteristics") of the revised manuscript (L377-L382).

Additionally, what are the most likely causes of the model's misrepresentation of low flow volumes in those basins? Can some of the selected process parameterizations (i.e., snowmelt) negatively affect the obtained results in this respect?

We speculate that both model structural deficiencies and forcing errors may be the main causes of the misrepresentation of low flow volumes. The SUMMA configuration used in this study considers that precipitation is the only water input for each grid cell, and lateral water exchanges among modeling units – in particular, groundwater fluxes – are not allowed, which could explain the relatively lower performance in the Cochiguaz and Claro River basins. On the other hand, CR2MET v2.0 – which is the baseline dataset for model simulations – can be considered a bias correction of daily ERA5 precipitation and extreme temperature outputs. Because such correction relies on meteorological gauges, which are very sparse in the Andes Cordillera, the resulting forcing dataset may contain biases that affect the accumulation and melting of snow, both extremely relevant in the Cochiguaz and Claro River basins since they are snowmelt-driven catchments. To reflect these points, we have modified the text as follows (L257-L261):

"… there is an overestimation of low flow volumes with exceedance probabilities larger than 90% in the Choapa and Claro catchments (< 2 m^3/s), which could be explained by the lack of a common aquifer enabling water exchange among grid cells in our SUMMA configuration, and/or biases in the forcing dataset that impact the accumulation and melting of snow."

Also, did the authors evaluate the influence of multiple model parameterizations on the obtained results?

We did not explore the effects of using multiple parameterizations on the results and conclusions and, therefore, we will add the following text to section 5.4 (L439-L441):

"We did not explore the effects of using alternative model parameterizations (e.g., stomatal resistance, lateral fluxes) or spatial configurations (e.g., spatially varying soil layer depths) on the results and conclusions obtained."

L260-L276: Did the authors consider a spin-up period for the model before starting the analysis in 04/1983? If so, please include this information. If not, I'd recommend neglecting the first two simulation years (1983-1984) in the subsequent analysis to minimize the influence of initial conditions in the obtained results.

We did consider a two-year warm up period before computing the standardized drought indices. We clarify this point in L217-L219:

"We use the calibrated parameters (see section 3.2) to perform hydrologic simulations for the historical period April/1981 – March/2020. All SDI computations consider a spin-up period of two years (April/1981 – March/1983) and the same reference period of 30 years (April/1983 – March/2013)".

L36: "associated to" replace with "associated with."

We will modify the text following the reviewer's recommendation.

L42: "Despite the drought concept refers" replace with "Despite the drought concept referring."

We will modify the text following the reviewer's recommendation.

L80: "percentile-based thresholds that are commonly" replace with "percentile-based thresholds commonly."

We will modify the text following the reviewer's recommendation.

L92: "What are the effects of different time scales on" replace with "How do different time scales affect"

We will modify the text following the reviewer's recommendation.

L94: "towards" replace with "toward".

We will modify the text following the reviewer's recommendation.

L96: "To seek for answers," replace with "To seek answers,"

We will modify the text following the reviewer's recommendation.

L113: "Hereafter, to" replace with "Hereafter,"

We will modify the text following the reviewer's recommendation.

L117 & L118: "mean annual temperatures between 9 to 16 °C" and "aridity indices between 0.4 to 3" replace with "mean annual temperatures between 9 and 16 °C" & "aridity indices between 0.4 and 3".

We will modify the text following the reviewer's recommendation.

L349: "drought durations ranging 12.3-12.9 months" replace with "drought durations ranging from 12.3-12.9 months"

We will modify the text following the reviewer's recommendation.

L386: "SSI is not as relevant in snowmelt-driven basins, compared to mixed regime and rainfall-dominated catchments." This statement appears weak. I recommend replacing it with: "SSI is less relevant in snowmelt-driven basins than in mixed regimes and rainfall-dominated catchments."

We will modify the text following the reviewer's recommendation.

References

Almagro, A., Meira Neto, A. A., Vergopolan, N., Roy, T., Troch, P. A. & Oliveira, P. T. S. (2024) The Drivers of Hydrologic Behavior in Brazil: Insights From a Catchment Classification. *Water Resour. Res.* **60**(8). doi:10.1029/2024WR037212

Brunner, M. I. & Stahl, K. (2023) Temporal hydrological drought clustering varies with climate and land-surface processes. *Environ. Res. Lett.* **18**(3). doi:10.1088/1748-9326/acb8ca

Loon, A. Van & Laaha, G. (2015) Hydrological drought severity explained by climate and catchment characteristics. *J. Hydrol.* **526**, 3–14. Elsevier B.V. doi:10.1016/j.jhydrol.2014.10.059

Muñoz-Castro, E., Mendoza, P. A., Vásquez, N. & Vargas, X. (2023) Exploring parameter (dis)agreement due to calibration metric selection in conceptual rainfall-runoff models. *Hydrol. Sci. J.* Taylor & Francis. doi:10.1080/02626667.2023.2231434

Peña-Gallardo, M., Vicente-Serrano, S. M., Hannaford, J., Lorenzo-Lacruz, J., Svoboda, M., Domínguez-Castro, F., Maneta, M., et al. (2019) Complex influences of meteorological drought time-scales on hydrological droughts in natural basins of the contiguous Unites States. *J. Hydrol.* **568**(November 2018), 611–625. Elsevier. doi:10.1016/j.jhydrol.2018.11.026

Sawicz, K., Wagener, T., Sivapalan, M., Troch, P. A. & Carrillo, G. (2011) Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. *Hydrol. Earth Syst. Sci.* **15**(9), 2895–2911. doi:10.5194/hess-15-2895-2011

Vásquez, N., Cepeda, J., Gómez, T., Mendoza, P. A., Lagos, M., Boisier, J. P., Álvarez-Garretón, C., et al. (2021) Catchment-Scale Natural Water Balance in Chile. In: *Water Resources of Chile*, 189–208. doi:10.1007/978-3-030-56901-3_9