REVIEWER #2

Request: *2018 is the target year for the study, but during the parameter calibration, 2018 was also used as the validation period. Usually, it might be better to set the validation period and the target year separately. This way, the validated model parameters can be better applied to the simulation of the target period; this is especially true if the model is used for forecasting.*

Response:

Dear Reviewer,

Thank you for your comments and recommendations. We were not entirely sure what the comment was here. If our assumption is correct, that the concern is that we used 2018 as both the target and calibration year. Our reply to this point is that 2018 is only consider as the validation year, and it is not used for calibration.

Request: *The year 2018 is the target year of the study in the paper, and an extreme drought happened in this year. In the article, the years 2016 - 2017 were used as the time for parameter calibration, but there was no extreme drought during these two years, which will inevitably lead to the calibrated parameters may not be applicable. This may also be one of the reasons why the results are not ideal. Additionally, the periods of model calibration and validation may be longer for a better and stable result.*

Response: It is acceptable to point out that a dry year could have been used during the calibration process. However, the primary objective of our study was to calibrate the model and assess its reliability to extrapolate to extreme drought conditions not seen during calibration and produce similar values of low water levels. The previous extreme hydrological drought was in 2003; hence, the computational requirements of the model for the Rhine basin were our main restriction. It is worth mentioning that despite not incorporating a dry year, the model successfully captured the extreme event.

We suggest that will add the explanation above to the manuscript.

Request: *The final simulation results do not seem ideal. Usually, an NSE result above 0.7 would be better, but in the paper, many of the results are below 0.5, and even there are some negative values. Although the hydrological situation in a larger basin is indeed more complex and difficult to simulate perfectly, is it possible to make further attempts to adjust the results better?*

Response: We appreciate your comment. Indeed, the results are not ideal, but they are acceptable. It was chosen to focus not only on one statistical score but on the combination of them (NSE, KGE, CC). Additionally, we had to consider all the hydrological stations, therefore using the median value. The lower values are located downstream of the basin, and this result is because there is a positive BIAS from the up-to-middle-stream of the Rhine, which contributes to the overestimation downstream, as stated in L324. Therefore, we do not see a necessity to further improve due to computational time.

We suggest that the explanation above be added to the manuscript.

Request: *The ERA5 data was used to drive the model operation, but it was also used as observational data to verify the model results, which is usually not recommended. It is advisable to see if there are other data, especially observational data, for verification, so that the results would be more reliable and convincing.*

Response: Thank you for your comment. A similar point was made by Reviewer #1. WRF-Hydro requires eight variables in a gridded setting as input forcing data. Therefore, ERA5 was chosen due to the open access availability of the necessary variables in a grid, and the transboundary conditions would create a complex task to obtain information of all the nine countries contained in the basin.

ERA5-Land was not used as input data for the model, and its generation is diverse from ERA5. Balsamo et al. (2009) stated that ERA5-Land uses the Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land to obtain the soil moisture values and ERA5 atmospheric variables as forcing. Therefore, it is not a direct comparison with the input data to the model. We found the necessity to search for a gridded data set that will allow us to compare with the soil moisture results obtained from the model, which is the case of ERA5-Land. The objective was to make a sanity check with an uncalibrated variable.

We suggest that the distinction between the datasets, as mentioned above, will be added to the manuscript.

Request: *The title of this paper is "Model calibration and streamflow simulations for the extreme drought event of 2018 on the Rhine River Basin using WRF-Hydro 5.2.0", but there is not much analysis in the article on the simulation results of the extreme drought event in 2018. More space is devoted to the calibration of model parameters and the discussion of parameters related to the lake process. It is suggested to analyze in detail this extreme drought event and the model's simulation of this event.*

Response: Thank you for your comment. A similar point was made by Reviewer #1. We can add more detail of the drought period (July-November 2018) with the plot below in the Annex and the statistical scores on the tables in the manuscript. We will also emphasize the use of NSE(log) for performance with low streamflow values.

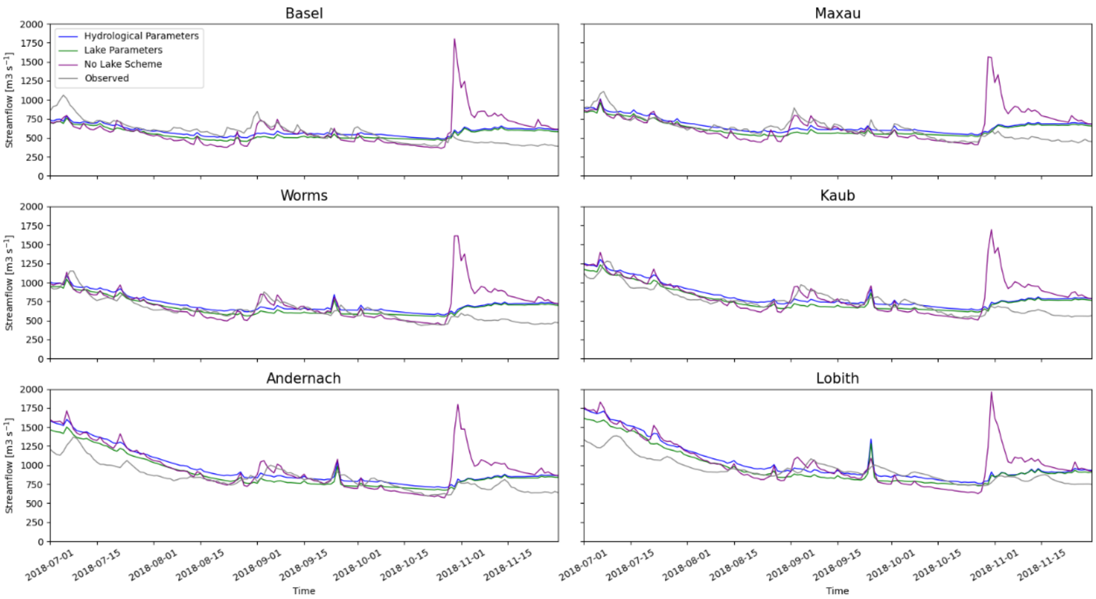


Fig S2. Daily streamflow hydrographs of the hydrological drought period of 2018 (July-November). Model simulation is the blue line and the observed data is the gray line from the stations (see Fig. 1 for their location).

Table S1. Statistical analysis results of the WRF-Hydro model performance regarding the model's hydrological parameters during the hydrological drought in 2018.

Station	Hydrological Drought (01 July – 27 October 2018)				
	NSE	NSE(log)	KGE	CC	Bias
Basel	0.52	0.54	0.50	0.83	-5.87
Maxau	0.69	0.66	0.71	0.83	0.36
Worms	0.64	0.58	0.78	0.85	6.29
Kaub	0.65	0.67	0.82	0.87	4.64
Andernach	-0.09	0.22	0.44	0.88	12.90
Lobith	-0.43	0.03	0.17	0.90	9.84
Median	0.58	0.56	0.60	0.86	5.46

Table S2. Statistical analysis results of the WRF-Hydro model performance regarding the model's lake scheme parameters during the hydrological drought in 2018.

Station	Hydrological Drought (01 July – 27 October 2018)				
	NSE	NSE(log)	KGE	CC	Bias
Basel	0.25	0.24	0.49	0.79	-12.21
Maxau	0.54	0.51	0.69	0.81	-6.53
Worms	0.69	0.65	0.77	0.83	-0.63
Kaub	0.68	0.65	0.84	0.86	-2.34
Andernach	0.38	0.49	0.54	0.88	5.14
Lobith	0.06	0.21	0.26	0.89	2.76
Median	0.46	0.50	0.62	0.85	-1.49

Table S3. Statistical analysis results of the WRF-Hydro model performance without the model's lake scheme during the hydrological drought in 2018.

Station	Hydrological Drought (01 July – 27 October 2018)				
	NSE	NSE(log)	KGE	CC	Bias
Basel	0.24	0.10	0.70	0.82	-14.20
Maxau	0.65	0.58	0.87	0.90	-7.97
Worms	0.78	0.78	0.89	0.90	-1.27
Kaub	0.66	0.66	0.68	0.91	-0.84
Andernach	-0.01	0.31	0.27	0.92	8.42
Lobith	-0.54	-0.19	-0.05	0.92	5.26
Median	0.45	0.45	0.69	0.91	-1.26

When considering only the low streamflow values period (July – November 2018), Fig. S2 shows that model performs these values until the end of October, as it is shown on the Tables S1, S2 and S3. The median values of the statistical scores show a good agreement with all the set ups. However, there is a better correlation and lower bias when not taking into consideration the lake scheme (Table S3) and Fig. S2 shows a better agreement with the variability of the streamflow even during this low water level period.

It is also visible that for the model there is an overestimation of the discharge from 2018-10-28 and further, with a high peak for the set up without the lake scheme (Fig. S2). This overestimation of observed streamflow is already visible at gauge Diepoldsau at the alpine Rhine, upstream of Lake Constance (Fig. S3). Therefore, the problem clearly originates from the alpine region of the basin.

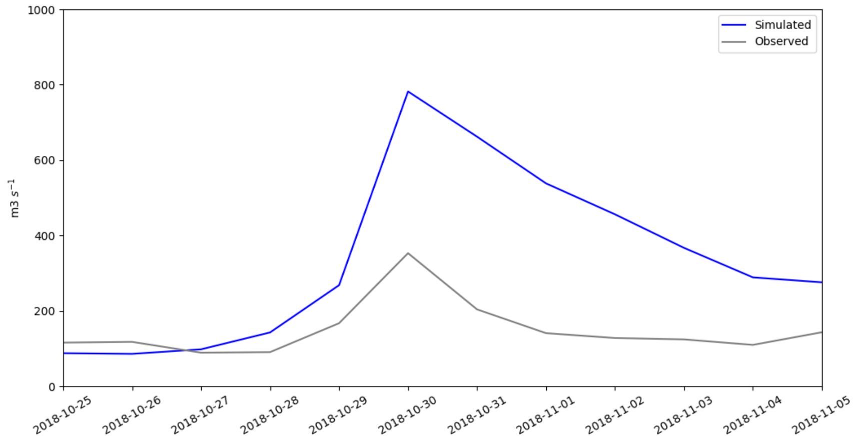


Fig S3. Daily streamflow hydrograph at the Diapoldsau Station

A probable cause is that during that time, ERA5 overestimated observed precipitation in the alpine parts, and more importantly, that it overestimated air temperature, leading to falsely accounting as rain rather than snow in WRF-Hydro.

Regarding precipitation overestimation, it has been stated by other authors (e.g. Rivoire et al. 2021 or Bandhauer et al., 2022) that this is a typical phenomenon of ERA5 in the alpine regions. With respect to falsely assuming rainfall, Fig. S4 shows the spatial mean daily values of precipitation and air temperature of the alpine high-altitude region of the Rhine basin. It is noticeable that the air temperature is greater than zero for the entire period. Therefore, it is possible that WRF-Hydro considers most of the precipitation as rainfall, leading to a direct substantial rainfall-runoff event. In contrast, observations show significant new snow accumulation in the same region in order or magnitude of 0.25 – 2 m during this period (Fig. S5). The snow depth data are taken from EnviDat (Mott, R. 2023). Based on this analysis, we can conclude that the overestimation of the observed streamflow peak at October 30 is due to falsely accounting precipitation mainly as rainfall rather than snow.

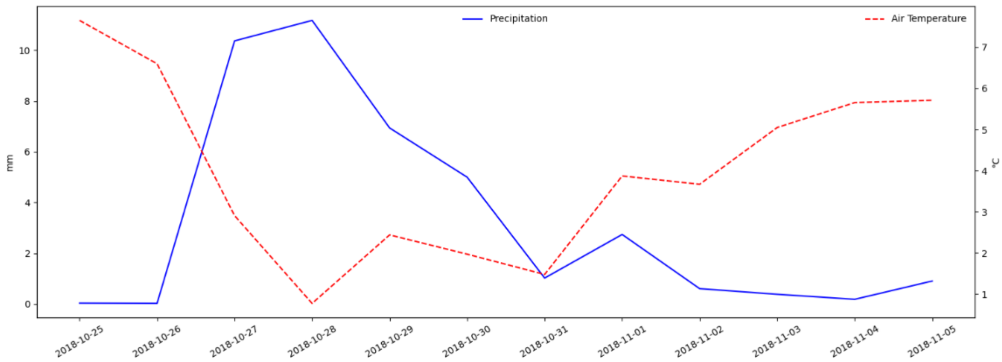
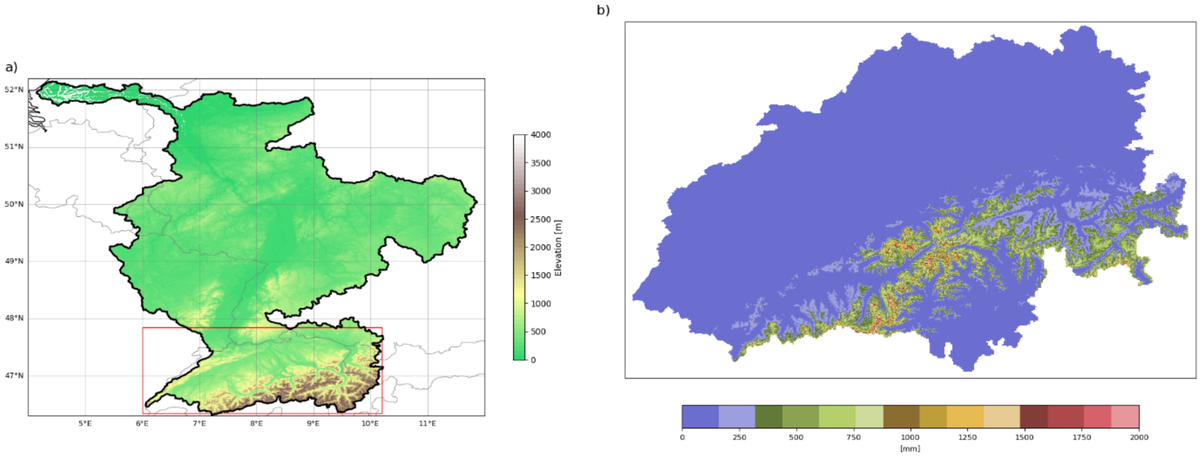


Fig S4. Spatial daily mean precipitation and air temperature from ERA5 for the alpine high-altitude regions of the Rhine basin.



S5. Snow depth accumulation from 2018-10-28 to 2018-11-05 (b). The location of this area is in a red square in (a).

When studying drought conditions, we cannot omit to include the behavior of the basin before and after the event took place. Thus, it is important that our model is calibrated in accordance with a regular year and extrapolate this performance to an extended dry period. Even though the results are not perfect, we were able to find a suitable set up of the model that can achieve the basin's behavior from wet to dry events.

We suggest that the argument above be added to the manuscript as a sub chapter the result section.

References

Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., and Betts A.K.: A Revised Hydrology for the ECMWF Model: Verification from Field Site to Terrestrial Water Storage and Impact in the Integrated Forecast System, J. Hydrometeorol., 10, 623–643, https://doi.org/10.1175/2008JHM1068.1, 2009.

Bandhauer, M., Isotta, F., Lakatos, M., Lussana, C., Båserud, L., Izsák, B., Szentes, O., Tveito, O.E., and Frei, C.: Evaluation of daily precipitation analyses in e-obs (v19.0e) and era5 by comparison to regional high-resolution datasets in European regions. Int. J. Climatol., 42(2), 727–747., https://doi.org/10.1002/joc.7269, 2022

Mott, R.: Seasonal snow data for Switzerland OSHD - FSM2sohd. EnviDat.[data set], https://www.doi.org/10.16904/envidat.404, 2023

Rivoire, P., Martius, O., Naveau, P.: A Comparison of Moderate and Extreme ERA-5 Daily Precipitation With Two Observational Data Sets, Earth and Space Science, 8, 1-16, https://doi.org/10.1029/2020EA001633, 2021