Authors' response to the reviews of

"Suitability of 17 rainfall and temperature gridded datasets for largescale hydrological modelling in West Africa" by Dembélé et al. (HESS-2020-68)

Reply to the Editor

Dear Authors,

Two good reviewers reviewed the manuscript (I was hoping three but one review didn't come true, this is partly why it took a bit longer than planned, aplogies for that). They rate the scientific significance as good/fair, the scientific quality as fair/fair and the presentation quality as good/poor. So there is room for improvement and both reviewers provide valuable feedback for improving your manuscript, some comments overlap/address the same issue.

Please adjsut your manuscript accordingly and provide an new manuscript including a list with changes made/answers to the reviewer.

Sincerely,

Albrecht Weerts

Dear Prof. Albrecht Weerts,

We thank you for handling our manuscript. Below we provide point-by-point responses to the referees' comments and indicate modifications done to our manuscript.

Best regards,

Moctar Dembélé
on behalf of all co-authors

Reply to Anonymous Referee #1

1) This is an interesting paper that is reasonably well written. Although the assessment includes a large number of datasets, the study area is relatively small, and the model is not recalibrated for each variable, which has led to some questionable conclusions.

Response: *We thank the referee#1 for the positive overall appreciation of our work. As it can be read from the title, our study focuses on the poorly studied region of West Africa, when it comes to hydrological modelling in general and hydrological evaluation of meteorological datasets in particular. Accordingly, our contribution clearly represents added-value, in terms of regional hydrology as well as in terms of hydrological modelling of semi-arid areas. In fact, independent regional evaluation of globally and regionally available datasets are of key importance for hydrology as they can provide new insights that might not be fully highlighted in global studies.*

Regarding the general critics on the model calibration, we would like to refer the reader to our detailed answer to comment 4. Below we provide more details and answers to the detailed comments of this reviewer.

2) You state that "rainfall datasets have contrasting performances across the four climatic zones present in the VRB, suggesting that, in general, basin-wide hydrological model performance might be misleading and invalid for a smaller spatial domain." What makes you think that your results, which also represent a relatively small spatial domain, are not "misleading and invalid" as well?

Response: *We thank the referee#1 for pointing out this potentially misleading statement. We agree with this referee that our statement has some strong wordings, which have led to a different interpretation by the referee#1 than what we intended to say. Our original idea is that the overall model performance in the entire, relatively large modelling domain (from a catchment hydrology perspective, the Volta River Basin, VRB, is indeed a large domain) might not be representative for all subdomains. This is especially the case if the modelling domain extends over multiple climatic zones as in the VRB case. For instance, in global studies, the overall global performance of a rainfall dataset is likely different from its performance in sub-regions such as West Africa (e.g. see Figure 3 in Beck et al. 2017b). Therefore, by "smaller spatial domain" we meant a portion of a large domain under evaluation. To avoid ambiguities, we have now modified this statement in the Abstract as follows: "…rainfall datasets have contrasting performances across the four climatic zones present in the VRB".*

3) It is stated that "the results can be considered valid for West Africa and regions with similar hydroclimatic and physical features" which is highly speculative and likely not true given the variation in precipitation dataset performance and gauge network density. To improve the generalizability of the results, the assessment should be expanded to other regions across Africa or the globe. Alternatively, the abstract and discussion should clearly state that the conclusions and the performance ranking of the datasets are not representative of other regions.

Response: *We agree with the referee#1 that this isolated statement can be interpreted as speculative. However, we did not want to imply any certainty but the possibility that the results might be transferable to other places. This is expressed in the sentence following the aforementioned sentence: "A wider generalization of the findings should be done with caution and after repeating similar evaluation studies in other places". Also, we did not intend to generalize our findings to other regions, which is very clear from the mentions "West Africa" in the title, and "Volta River Basin (VRB) in West Africa" in the very first sentence of our manuscript.*

We agree that we should not have mentioned transferability to other similar climates outside Africa since the performance of any remote sensing-based meteorological data set for hydrological modelling varies across the globe due to many other factors not only related to regional aridity.

To avoid ambiguities, the statements have been reformulated as follows in the discussion: "The results are primarily valid for the study region in West Africa, while a wider generalization of the findings should be done with caution and after repeating similar evaluation studies at other places".

4) The soil moisture, terrestrial water storage, and actual evaporation assessments were carried out without recalibrating the model and therefore the results for these variables are subject to substantial uncertainty. This is supported by the fact that MSWEP, which was used to force GLEAM, does not exhibit good actual evaporation scores. The model should be recalibrated for each variable.

Response:

Performance of MSWEP: In our opinion, MSWEP has very good scores for modelled actual evaporation compared to GLEAM as it always exceeds a Pearson correlation coefficient (r) of 0.9 (Figure 8), with an average r=0.94 for the entire VRB (Appendix A3), and it has the highest spatial pattern score (Esp=0.26) among all the rainfall datasets (Appendix A3).

Model recalibration: We would like to emphasize here that the model is indeed recalibrated for each meteorological input product combination (i.e. rainfall and temperature), but it is recalibrated with streamflow (Q) only and not with soil moisture (Su), terrestrial water storage (St), and actual evaporation (Ea). The logic behind this approach is twofold:
i) We would like to know how well the model performs in combination with the different input variables; we therefore use Su, St and Ea as evaluation variables.
ii) Further calibrating our model with Su, St, and Ea would lead to additional model improvement due to the information content of these variables as demonstrated by Dembélé et al. (2020). In this case, it becomes difficult to disentangle the contribution of the rainfall datasets and the contribution of the calibration variables (Su, St, and Ea) to the overall model performance. Calibrating the model on one reference output variable (in-situ streamflow) and evaluating it against other output variables remains in our view a powerful method to assess the usefulness of a meteorological input dataset for hydrological modelling. By calibrating on streamflow, we give each meteorological data set "a chance" to perform as well as possible for streamflow; we then further discriminate between the usefulness of the input variables for hydrological modelling by assessing whether they can do a good job for streamflow and Su, St and Ea simultaneously. The "dream" input variable should indeed perform well for all variables if only calibrated on one.

We would like to emphasize here, that the evaluation of Su, St, and Ea is not done with the absolute values (i.e. raw data) of the satellite products, but rather we evaluate their temporal dynamics and spatial patterns using bias-insensitive metrics. Therefore, we substantially mitigate uncertainties that might arise from the assessment of these variables when using their absolute values (Dembélé et al. 2020; Nijzink et al., 2018; Mendiguren et al., 2017; Wambura et al., 2018). We had already discussed our choice for the Q-only calibration and its limitations at lines 493-497 and the potential uncertainties related to the satellite datasets used for evaluation at lines 477-479.

However, we have now made the choice of the Q-only calibration clearer by adding the following in the discussion: "The model is calibrated only on Q data despite the known limitations of the Q-only calibration (Demirel et al., 2018). However, calibrating the model on additional variables would result in additional model performance improvement that would not be separable from the contribution of the input datasets to the model performance. Therefore, regarding the goal of this study, the Q-only calibration was the best option to obtain the impact of various meteorological forcing datasets on the plausibility of hydrological processes."

5) The word "gauge" is not used in the abstract and the datasets are only classified as either satellite or reanalysis. However, the amount of gauge data incorporated in the datasets may well be the overriding factor in determining the performance, given the good performance of TAMSAT and CHIRPS in terms of streamflow.

Response: *In the abstract, we have now modified the statement "Seventeen precipitation products based on satellite data (…)" into "Seventeen precipitation products based essentially on gauge-corrected satellite data (…)". Moreover, Table 1 provides information on rainfall datasets developed with gauge data.*

6) Figures 7 and 10 are impossible to interpret, way too much information. Should be condensed.

Response: We agree with referee#1 that Figure 7 and 10 contain a lot of information. We have now reduced the contents of Figure 7 and 10. Thereby, only showing the model performance for the entire VRB. The previous figures are moved to the supplementary materials.

Reply to Referee #2 (Nadav Peleg)

In their paper, Dembélé et al. explore the suitability of combining time series of rainfall and temperature from different climate products as inputs into a hydrological model. The manuscript is well structured and written, methods are robust and results are presented adequately. The research question of the possibility of combining gridded climate variables from different sources to simulate various hydrological components is relevant and timely, and I believe will be of interest for the readers of HESS. Nevertheless, I have a few comments and suggestions for the authors to consider before I can recommend the paper for publication.

Sincerely,

Nadav Peleg

Response: *We thank the referee#2 for the positive overall appreciation of our work. Below we provide answers to the referee's comments.*

Major comments

1. I found one-step in the methodology (i.e. as presented in Figure 1) to be missing. I think it will be meaningful to know how the climate variables (rainfall, temperature) from each climate products are ranked in comparison to observed data (i.e. from ground stations) before ranking the 102 input combinations based on various hydrological components. I think this step is critical to understand the presented results. For example, JRA-55 and ERA5 yield poor correlation with Ea (Figure 8), but isn't this because they are poorly reproducing the rainfall statistics over the VRB? GSMaP-std V6 reproduces well the streamflow (Figure 3), St (Figure 4), Su (Figure 5) and Ea (Figure 8) – will this product be ranked #1 when compared to ground stations? I assume there will be a high correlation between the ranks emerging from the comparison to ground stations and hydrological outputs from the model. If this case, wouldn't it be sufficient to evaluate the best products to use in hydrological simulations simply by comparing them to the few climate stations that are available in the catchment of interest or a nearby area? This is a point for discussion.

Response:

Comparison with ground observations

We agree with the referee that knowing the performance of the meteorological datasets in comparison with ground measurement could be an interesting starting point. However, it is noteworthy that the Volta River basin (VRB) in West Africa is a data scarce region, not like other places in Europe and USA (e.g. Beck et al., 2019a) where a large amount of ground measurements is widely and freely accessible. The few datasets collected by local organizations in the VRB are not easily accessible due to the transboundary nature of the basin that is shared among six countries. It took us one year to obtain *streamflow data, which was further subject to a thorough gap-filling and quality control of time series (Dembélé et al., 2019).*

The VRB region has a low density of meteorological stations (see Figure 1 of Dembélé and Zwart 2016; and Figure 1 of Satgé et al., 2020). A thorough evaluation of satellite/reanalysis datasets with ground measurements in the VRB cannot be limited to a few stations because the basin is about 415,600 km^2 (ten times the size of Switzerland), with a unique and complex climate (see Section 2.4 Study Area), and a strong spatial variability of rainfall.

Even in case of ground measurement availability, the validity of point-to-pixel comparison is questionable (e.g. JRA-55 is 1.25°, and ERA5 is 0.25°) because the gauge measurement will hardly represent the spatial variability of rainfall in a pixel. Moreover, the rainfall datasets used in our study are essentially gauge-corrected data. Therefore, a robust ground evaluation would require independent ground measurements that are not used in the development of the rainfall datasets (Beck et al., 2019a), which is a luxury in West Africa.

Validity of ground evaluation for hydrological modelling

The skill of a product in reproducing well ground measurement under a point-to-pixel evaluation does not necessarily guarantee its high performance for hydrological modelling, mainly in complex hydroclimatic environments such as the VRB. The performance of isolated pixels might not be representative of all pixels. Usually, hydrological modelling is undertaken at daily or higher temporal resolution. However, mismatches between gauge and satellite reporting times are a major issue in ground evaluation (Beck et al., 2019a). This is confirmed by the substantial increase in the evaluation performance of rainfall datasets from daily to monthly time scale (Dembélé and Zwart (2016); see Figure 3 vs. Figure 8 of Satgé et al. (2020)).

We have now added the following to our discussions: "Moreover, when comparing the results of this study to the findings of Satgé et al. (2020) based on a point-to-pixel evaluation of gridded rainfall datasets in West Africa, it is noticeable that the ground evaluation might lead to different results as compared to the hydrological evaluation adopted in the current study. The skill of a rainfall product in well reproducing ground measurements under a point-to-pixel evaluation does not necessarily correlate with its performance for hydrological modelling, particularly in large and complex hydroclimatic environments such as the VRB."

2. The modelling experiment includes 6 years for model calibration and 4 years for model evaluation. These are very short periods, not necessarily representing well the natural climatic and hydrological variability and not necessarily guarantying a successful calibration of the hydrological model parameters. First, I suggest demonstrating with a simple graph (can be presented as SI) that the natural variability is somehow represented in your 10-year data. Second, consider adding a short discussion regarding the sensitivity (quantified) of the hydrological model parameters to the short period that is used for the model training.

Response:

Length of the calibration and simulation period

We agree that the modelling period of 10 years, which includes 6 years for calibration and 4 years for evaluation, might not seem very long, but is long enough to obtain a well calibrated model in our case, *as previously demonstrated by Dembélé et al., (2020). Moreover, a 3-year model warm up period (2000-2002) precedes the calibration period. The choice of the modelling period is constrained by the availability and the quality of the in-situ streamflow measurements in the data-scarce VRB (Dembélé et al., 2019).*

Moreover, it is important to stress that we adopt a daily streamflow calibration, which means a time series of 2192 time steps to simulate and match for each of the 11 gauging stations during the 6-year model calibration period (2003-2008), or 3653 time steps for the 10-year simulation period (2003-2012). In our opinion, this is a robust model calibration approach, additionally supported by the fact that we adopt a multi-site calibration simultaneously at 11 streamflow gauging points located in very distinct hydroclimatic zones within the basin (see Figure 2). It is worth mentioning that the computational cost for each of the 102 input data combinations is about 6 days for 4000 parameter iterations during the model calibration on a computer Intel Xeon Processor E5-2697 v3 with 64 GB of RAM.

Natural variability of streamflow

We thank the referee#2 for this important comment on natural variability that was not appropriately discussed. Natural variability of daily streamflow can be observed at each of the 11 streamflow gauging sites used in this study, and inter-site variability of streamflow can be observed as well for the 10-year period (2003-2012). As it can be seen in Figure R1 below, the modelling period covers years with considerably different streamflow volumes during the wet season and with considerably different peak discharges, ranging e.g. for station 4 from 250 m3/s (year 2011) to 900 m3/s (year 2003). In general, years 2004, 2005 and 2009 can be considered as dry while 2003, 2006, and 2010 are wet for station 2 and 4, which have low flows as compared to the station 11.

Now, we have added Figure S16 showing the hydrographs of all the eleven stations in the supplementary materials, and it is indexed in the text of the revised manuscript (Section 2.5).

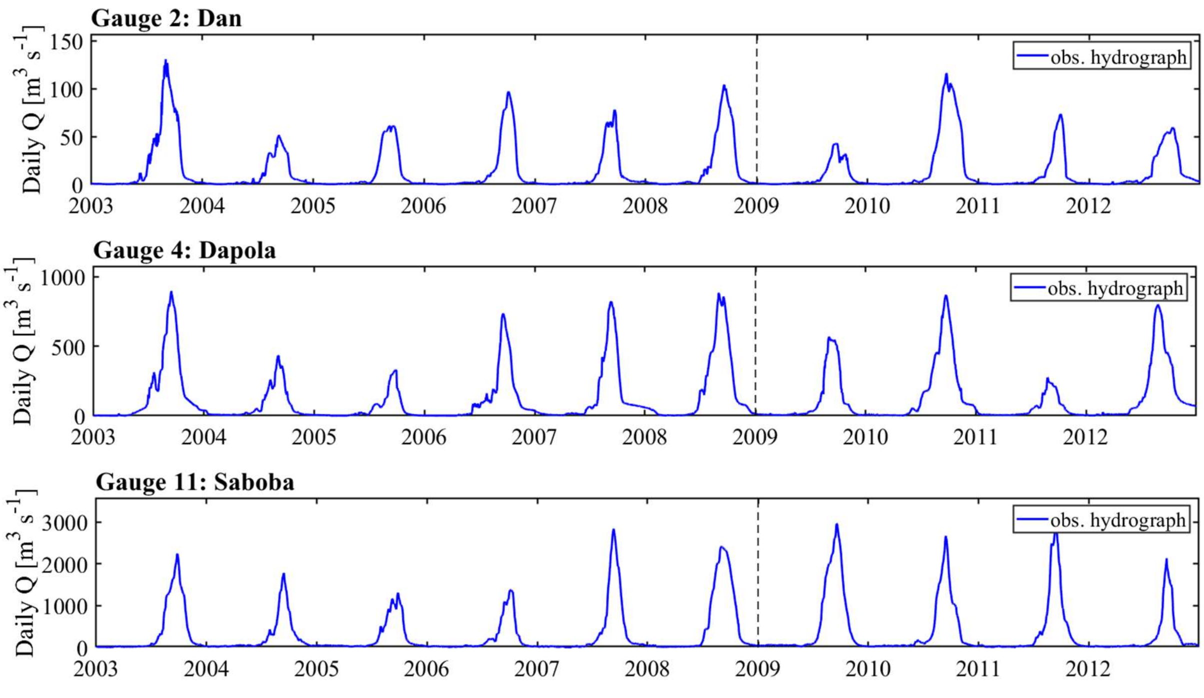


Figure R1: Hydrographs at three different gauging stations in the VRB during the modelling period comprised of the calibration period (2003-2008) and the evaluation period (2009-2012)

Natural variability of meteorological datasets

Natural variability can also be observed in the rainfall datasets as shown below in Figures R2-R3. It can be seen that rainfall varies both in time and space across different climatic zones in the VRB, which makes it an interesting case study for rainfall evaluation. Moreover, the inter-product variability is very high, often higher than the natural variability, which further justifies this study, especially in the absence of in-situ data.

These figures have been added to the supplementary materials (Figures S1 and S5; Figures S3-S4 and S6-8) and indexed in the text of the revised manuscript (Section 2.5).

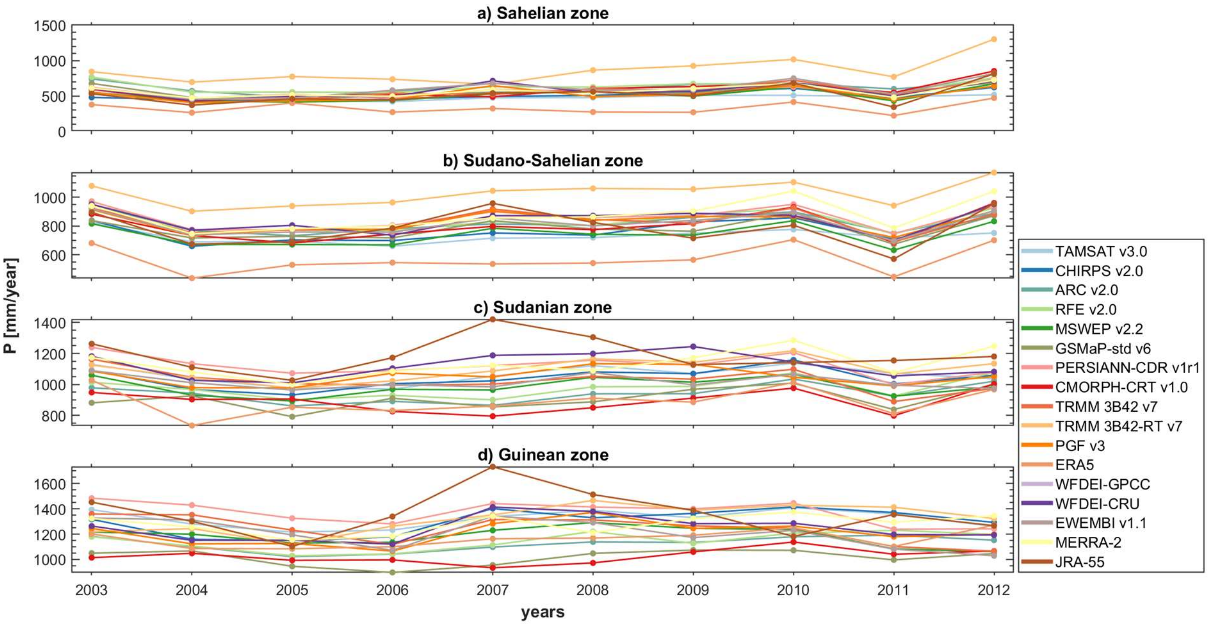


Figure R2: Annual total rainfall for 17 datasets for different climatic zones in the VRB

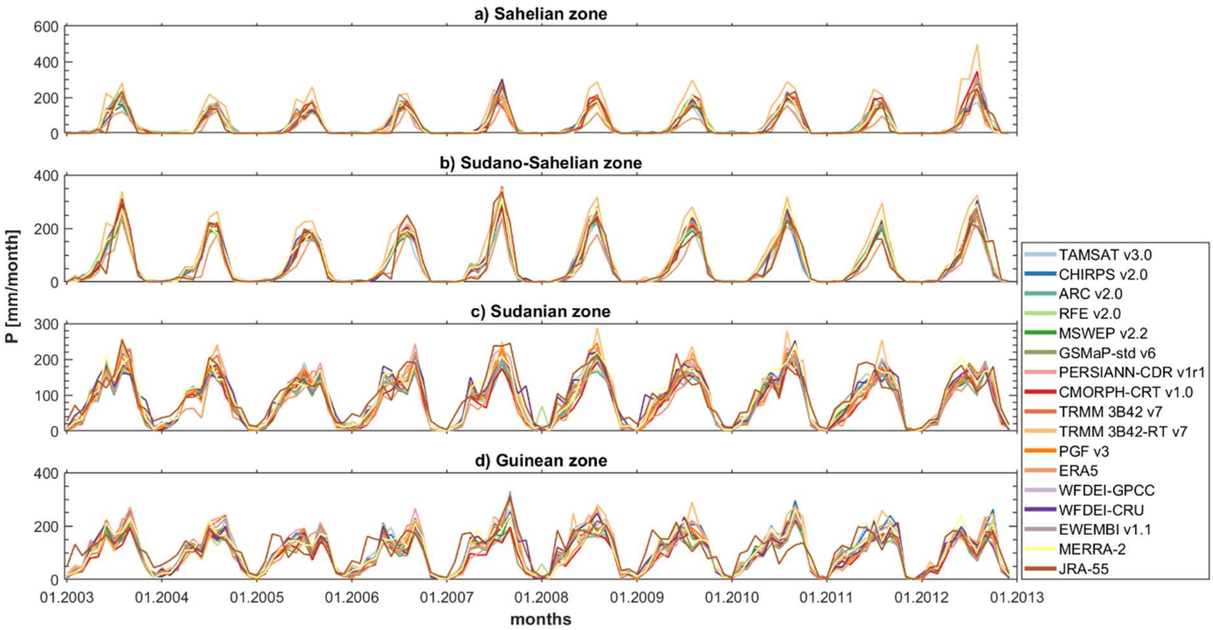


Figure R3: Monthly total rainfall for 17 datasets for different climatic zones in the VRB

Sensitivity of model parameters

In the supplementary materials (Figure S79 at Section 12.2), we provide a figure that shows the distribution of each of the 36 global model parameters and their sensitivity (i.e. second-order coefficient of variation) to different input meteorological data. It can be seen that most of the model parameters vary considerably as a response to the change of rainfall and temperature data.

The following has been added to the discussion: "(…). Moreover, it can be noticed that most of the model parameters are sensitive to the change in meteorological input datasets (Figure S79)."

Minor comments

1. Usually, when considering using gridded climate variables from climate re-analysis/other products as inputs into hydrological models the following steps are taken: (i) computing the skills (i.e. temporal dynamics, magnitude, and occurrence) of the climate variables in comparison to observed data; (ii) choosing the (individual) climate product with the best skill to use; and (iii) performing a bias correction to the climate variables to improve the fit to the observed data. I am missing a paragraph in the introduction/discussion explaining why not simply following this practice which should improve the hydrological outputs from the model.

Response: *The approach proposed by the referee#2 is usually applied for hydrological climate change impact studies, where climate projection data known to be biased are first evaluated and corrected with observations. In our Introduction, we have described the usually adopted approaches for the evaluation of gridded (satellite and reanalysis) datasets as follows:*
"The errors quantification of SRPs and reanalysis products is usually done by comparing them with in-situ measurements (e.g. Dembélé and Zwart, 2016;Thiemig et al., 2012;Beck et al., 2019a;Caroletti et al., 2019;Satgé et al., 2020), or by assessing their reliability as forcing for hydrological models (e.g.Duethmann et al., 2013;Pan et al., 2010;Nkiaka et al., 2017). Other evaluation approaches include triple collocation, which is a technique that estimates the variance of unknown errors of three independent variables without a reference or observed variable (e.g. Massari et al., 2017;Alemohammad et al., 2015;McColl et al., 2014;Roebeling et al., 2012). Compared to the ground-truthing approach, the hydrological evaluation approach has received limited attention (Camici et al., 2018;Poméon et al., 2017)."

Among those approaches, we adopted the hydrological evaluation, which consists in assessing the reliability of the gridded datasets in reproducing plausible spatiotemporal patterns of hydrological processes when used as input to a model, knowing that they might still present some discrepancies with ground measurements. This approach is particularly interesting in data scarce regions where ground evaluation is challenging or impossible. It is important to mention here that the gridded datasets that we are evaluating in our study are essentially gauge-corrected datasets as mentioned in Section 2.2, also see Table 1. In this case, the datasets are already bias-corrected.

We will make this clearer in the abstract by mentioning the use of gauge-corrected datasets in our study. Therefore, as also requested by the referee#1, the statement "Seventeen precipitation products based on satellite data (…)" will be replaced by "Seventeen precipitation products based essentially on gauge-corrected satellite data (…)". Moreover, Table 1 provides information on rainfall datasets developed with gauge data.

2. Results (Figure 3, for example). 22 values are used to represent the combined performance for the calibration and evaluation periods. This is not clear to me. Why not using a single Ekg value for the entire simulation period (merging the calibration and validation periods to a single period) for each gauge, i.e. 11 values in total per combination of temperature and precipitation product? What is the logic in separating the Ekg values to calibration and validation periods?

Response: *The decision for using 22 values of EKG (11 for calibration + 11 for evaluation) was based on the necessity to have enough elements for plotting the boxplots. For simplicity in reporting, a new Figure 3 is now provided only showing the median EKG of the entire simulation period, similarly to Figure 4, 5 and 8.*

3. Table 1. I suggest adding in the table additional column indicating if the product refers to rainfall, temperature or both. Also, please double-check the space-time resolutions reporter in the table. I think that the CMORPH-CRT product, for example, has a resolution of 8-km and 30-min.

Response: *We thank the referee for the suggestion. We have now added a new column indicating if the product refers to rainfall, temperature or both in Table 1. We are aware that different versions of the datasets exist, so that we have carefully mentioned in the caption of Table 1 that the information provided refer to the version of the datasets we have used. The provided information for CMORPH-CRT is correct, and the data was accessed from this web link:* ftp://ftp.cpc.ncep.noaa.gov/precip/CMORPH_V1.0/CRT/0.25deg-DLY_00Z/

4. The use of second-order CV is interesting, I do not recall seeing it in the context of hydrological statistics. Why use it and not simply using Pearson's CV skill? A sentence explaining the motivation is needed.

Response: *We agree with the referee that the use of the second-order CV is uncommon. We have now added the reasons of the use of the second-order CV and the limitations of the classical Pearson's CV in Section 2.7.*

5. Figures 7 and 10. Too many box-plots are presented. Perhaps present only the median (avoid using box-plots) to compare between products and climatic zones. This will considerably reduce the size and information plotted.

Response: *We agree with the referee#2 that Figure 7 and 10 contain a lot of information. As also responded to the referee#1 (comment #6), we have now reduced the contents of Figure 7 and 10, thereby, only showing the model performance for the entire VRB. The previous figures are now moved to the the supplementary materials.*

6. Generalization of the results. In lines 437-438 you mentioned that: "The results can be considered valid for West Africa and regions with similar hydroclimatic and physical features. A wider generalization of the findings should be done with caution and after repeating similar evaluation studies in other places". I do not think that you can generalize the results - they are likely to differ between locations as the quality of climate variables from different climate products differ between locations. In my view, the key message of your paper is that for each large catchment you should consider multiple sources of climate data to find the climate variables combination that is suitable for your region. The VRB is simply a case study used to demonstrate this point.

Response: *We agree with the referee#2 and we would like to stress that we did not intend to generalize our results as we carefully draw the reader's attention on the necessity to repeat the same experiment in other regions. As also responded to the referee#1 (comment #3), and to avoid ambiguities, the statement has been modified as follows in the discussions: "The results are primarily valid for the study region in West Africa, while a wider generalization of the findings should be done with caution and after repeating similar evaluation studies at other places".*

Corrigendum

We have realized an erroneous reporting of the objective function used for model calibration with streamflow data. Instead of using E_{KG} as currently reported in the manuscript, we used a combination of the Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970) of streamflow (E_{NS}) and the Nash-Sutcliffe efficiency of the logarithm of streamflow (E_{NSlog}), similarly to Dembélé et al. (2020). This setting has the advantage of identifying a parameter set that better predicts both high and low flows because E_{NS} is known to be very sensitive to high flows, while E_{NSlog} is a metric for low flows (Krause et al., 2005; Oudin et al., 2006; Pushpalatha et al., 2012). The objective function for streamflow was therefore modified consequently (equation 3). These modifications do not affect the current results, rather they reinforce the analysis as we now report on the model performance for streamflow with multiple skill scores (i.e. E_{NS}, E_{NSlog} and E_{KG}). Moreover, Appendix A3 has been modified accordingly.

References

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. J. M., Weedon, G. P., Brocca, L., Pappenberger, F., Huffman, G. J., and Wood, E. F.: Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling, Hydrol Earth Syst Sc, 21, 6201-6217, https://doi.org/10.5194/hess-21-6201-2017, 2017b.

Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., Van Dijk, A. I., Huffman, G. J., Adler, R. F., and Wood, E. F.: Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS, Hydrol Earth Syst Sc, 23, 207-224, https://doi.org/10.5194/hess-23-207-2019, 2019a.

Dembélé, M., and Zwart, S. J.: Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa, Int J Remote Sens, 37, 3995-4014, https://doi.org/10.1080/01431161.2016.1207258, 2016.

Dembélé, M., Oriani, F., Tumbulto, J., Mariéthoz, G., and Schaefli, B.: Gap-filling of daily streamflow time series using Direct Sampling in various hydroclimatic settings, Journal of Hydrology, 569, 573-586, https://doi.org/10.1016/j.jhydrol.2018.11.076, 2019.

Dembélé, M., Hrachowitz, M., Savenije, H. H., Mariéthoz, G., and Schaefli, B.: Improving the predictive skill of a distributed hydrological model by calibration on spatial patterns with multiple satellite datasets, Water Resources Research, e2019WR026085, https://doi.org/10.1029/2019WR026085, 2020a.

Demirel, M. C., J. Mai, G. Mendiguren, J. Koch, L. Samaniego, and S. Stisen (2018), Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model, Hydrology and Earth System Sciences, 22(2), 1299-1315, https://doi.org/10.5194/hess-22-1299-2018.

Krause, P., D. Boyle, and F. Bäse (2005), Comparison of different efficiency criteria for hydrological model assessment, Advances in geosciences, 5, 89-97, https://doi.org/10.5194/adgeo-5-89-2005.

Kvålseth, T. O.: Coefficient of variation: the second-order alternative, Journal of Applied Statistics, 44, 402-415, https://doi.org/10.1080/02664763.2016.1174195, 2017.

Mendiguren, G., J. Koch, and S. Stisen (2017), Spatial pattern evaluation of a calibrated national hydrological model - a remote-sensing-based diagnostic approach, Hydrology and Earth System Sciences, 21(12), 5987-6005, https://doi.org/10.5194/hess-21-5987-2017.

Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models part I—A discussion of principles, Journal of hydrology, 10(3), 282-290, https://doi.org/10.1016/0022-1694(70)90255-6.

Nijzink, R. C., et al. (2018), Constraining Conceptual Hydrological Models With Multiple Information Sources, Water Resources Research, 54(10), 8332-8362, https://doi.org/10.1029/2017wr021895.

Oudin, L., V. Andréassian, T. Mathevet, C. Perrin, and C. Michel (2006), Dynamic averaging of rainfall-runoff model simulations from complementary model parameterizations, Water Resources Research, 42(7), https://doi.org/10.1029/2005WR004636.

Pushpalatha, R., C. Perrin, N. Le Moine, and V. Andreassian (2012), A review of efficiency criteria suitable for evaluating low-flow simulations, Journal of Hydrology, 420, 171-182, https://doi.org/10.1016/j.jhydrol.2011.11.055.

Satgé, F., Defrance, D., Sultan, B., Bonnet, M.-P., Seyler, F., Rouché, N., Pierron, F., and Paturel, J.-E.: Evaluation of 23 gridded precipitation datasets across West Africa, Journal of Hydrology, 581, 124412, https://doi.org/10.1016/j.jhydrol.2019.124412, 2020.

Wambura, F. J., O. Dietrich, and G. Lischeid (2018), Improving a distributed hydrological model using evapotranspiration-related boundary conditions as additional constraints in a data-scarce river basin, Hydrological processes, 32(6), 759-775, https://doi.org/10.1002/hyp.11453.

Suitability of 17 rainfall and temperature gridded datasets for large-scale hydrological modelling in West Africa

Moctar Dembélé[1], Bettina Schaefli[1,3], Nick van de Giesen[2] & Grégoire Mariéthoz[1]

[1]Institute of Earth Surface Dynamics, Faculty of Geosciences and Environment, University of Lausanne, CH-1015 Lausanne, Switzerland
[2]Water Resources Section, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands
[3]Now at: Institute of Geography, Faculty of Science, University of Bern, CH-3012, Switzerland

Correspondence to: Moctar Dembélé (moctar.dembele@unil.ch)

Abstract. This study evaluates the ability of different gridded rainfall datasets to plausibly represent the spatiotemporal patterns of multiple hydrological processes (i.e. streamflow, actual evaporation, soil moisture and terrestrial water storage) for large-scale hydrological modelling in the predominantly semi-arid Volta River Basin (VRB) in West Africa. Seventeen precipitation products based essentially on gauge-corrected ~~on~~ satellite data (TAMSAT, CHIRPS, ARC, RFE, MSWEP, GSMaP, PERSIANN-CDR, CMORPH-CRT, TRMM 3B42, TRMM 3B42RT) and on reanalysis (ERA5, PGF, ~~JRA-55,~~ EWEMBI, WFDEI-GPCC, WFDEI-CRU, MERRA-2 and~~, PGF and ERA5~~ JRA-55) are compared as input for the fully distributed mesoscale Hydrologic Model (mHM). To assess the model sensitivity to meteorological forcing during rainfall partitioning into evaporation and runoff, six different temperature reanalysis datasets are used in combination with the precipitation datasets, which results in evaluating 102 combinations of rainfall-temperature input data. The model is recalibrated for each of the 102 input combinations, and the model responses are evaluated by using in-situ streamflow data and satellite remote sensing datasets from GLEAM evaporation, ESA CCI soil moisture, and GRACE terrestrial water storage. A bias-insensitive metric is used to assess the impact of meteorological forcing on the simulation of the spatial patterns of hydrological processes. The results of the process-based evaluation show that the rainfall datasets have contrasting performances across the four climatic zones present in the VRB~~suggesting that, in general, basin-wide hydrological model performance might be misleading and invalid for a smaller spatial domain~~. The top three best performing rainfall datasets are: TAMSAT, CHIRPS and PERSIANN-CDR for streamflow; ARC, RFE and CMORPH-CRT for terrestrial water storage; MERRA-2, EWEMBI/WFDEI-GPCC, and PGF for the temporal dynamics of soil moisture; MSWEP, TAMSAT and ARC for the spatial patterns of soil moisture; ARC, RFE and GSMaP-std for the temporal dynamics of actual evaporation; MSWEP, TAMSAT and MERRA-2 for the spatial patterns of actual evaporation. No single rainfall or temperature dataset consistently ranks first in reproducing the spatiotemporal variability of all hydrological processes. A dataset that is best in reproducing the temporal dynamics is not necessarily the best for the spatial patterns. In addition, the results suggest that there is more uncertainty in representing the spatial patterns of hydrological processes than their temporal dynamics. Finally, some region-

tailored datasets outperform the global datasets, thereby stressing the necessity and importance of regional evaluation studies for satellite and reanalysis meteorological datasets.

Keywords: Precipitation; Atmospheric forcing; Hydrological consistency; Process-based evaluation; Data uncertainty propagation; Ungauged basins; Data scarce regions

1 Introduction

Our understanding of environmental systems is underpinned by observational data whose unavailability and uncertainties hinder research and operational applications. Among other factors, atmospheric data quality is of prime importance for the reliability of hydro-meteorological and climatological studies (Ledesma and Futter, 2017;Zandler et al., 2019). Precipitation is one of the major components of the water cycle, which has led to numerous initiatives on understanding its generation, and estimating its amount and variability on Earth (Maidment et al., 2015;Cui et al., 2019). In hydrological modelling (Singh, 2018;Beven, 2019), precipitation is the most important driver variable that determines the spatiotemporal variability of other hydrological fluxes and state variables (Thiemig et al., 2013;Bárdossy and Das, 2008).
With the development of distributed hydrological models that facilitate large-scale predictions (Clark et al., 2017;Fatichi et al., 2016;Ocio et al., 2019), there is a growing need to inform and evaluate those models with distributed observational datasets to improve spatiotemporal process representation (Baroni et al., 2019;Paniconi and Putti, 2015;Hrachowitz and Clark, 2017). A key challenge is the spatiotemporal intermittency of precipitation, which is a major challenge for its measurement and its spatial interpolation (Tauro et al., 2018;Acharya et al., 2019;Bárdossy and Pegram, 2013;Wagner et al., 2012a), especially in regions with particular features such as complex topography, convection-driven precipitation or snowfall occurrence. A comprehensive description of precipitation measurement techniques can be found in previous studies (e.g. Tapiador et al., 2012;Stephens and Kummerow, 2007;Kidd and Huffman, 2011;Levizzani et al., 2020). The drawbacks of in-situ measurements of precipitation include limited and uneven areal coverage, deficiencies in instruments and costly maintenance (Kidd et al., 2017;Awange et al., 2019;Harrison et al., 2019), and have led to the advent of precipitation estimation from space (Barrett and Martin, 1981). Precipitation estimates from space are spatially homogeneous and cover inaccessible regions with uninterrupted records over time (Beck et al., 2019b;Funk et al., 2015).
The advent of satellite-based rainfall products (SRPs) has opened up new avenues for water resources monitoring and prediction, especially in data sparse regions (Serrat-Capdevila et al., 2014;Sheffield et al., 2018;Hrachowitz et al., 2013). Although, the use of SRPs in hydrology is increasing (Xu et al., 2014;Chen and Wang, 2018), they have not been fully adopted for operational purposes yet (Ciabatta et al., 2016;Kidd and Levizzani, 2011). The limited uptake of SRPs in hydrology is due to measurement bias, inadequate spatiotemporal resolutions (e.g. for extreme event simulation) and shortness of the records for some applications (e.g., climate change impact assessments), and the skepticism of some potential users with regard to the data quality (Marra et al., 2019). In the past decades, a large number of SRPs have been developed with different objectives,

spatial and temporal resolutions, input sources, algorithms and acquisition methods (Ciabatta et al., 2018;Ashouri et al., 2015;Brocca et al., 2019). Several studies provide a review of SRPs (e.g. Maidment et al., 2014;Sun et al., 2018;Maggioni et al., 2016;Le Coz and van de Giesen, 2019).

In addition to SRPs, there are also atmospheric retrospective analysis (or reanalysis) datasets of precipitation. A reanalysis system is composed of a forecast model and a data assimilation scheme that integrates spatiotemporal observations of meteorological variables (i.e. temperature, humidity, wind and pressure) to generate gridded atmospheric data (Lorenz and Kunstmann, 2012;Schröder et al., 2018). Precipitation is one of the reanalysis model-generated fields that generally has more uncertainties than the meteorological state fields (Roca et al., 2019). Reanalysis datasets are often used in hydrological modelling (Tang et al., 2019;Duan et al., 2019;Gründemann et al., 2018), and sometimes they are preferred over SRPs because of their usually long-term records suitable for climate change studies, and because of their higher performance in predictable large-scale stratiform systems (Seyyedi et al., 2015;Potter et al., 2018).

Despite the progress in satellite instruments, which has led to substantial advances in improving precipitation estimates (Sorooshian et al., 2011;Tang et al., 2019), there are known inconsistencies among the available SRPs (Sun et al., 2018;Tapiador et al., 2017). SRPs are subject to inherent errors originating mainly from precipitation retrieval instruments and algorithms, sampling frequency, and inadequate representation of cloud physics in some regions (Laiti et al., 2018;Alazzy et al., 2017;Romilly and Gebremichael, 2011). While on the one hand SRPs are subject to systematic biases, reanalysis products on the other hand have uncertainties resulting from their model forcing parameters, low spatial resolution with poor representation of sub-grid processes, and the model physics (Bosilovich et al., 2008;Laiti et al., 2018). Uncertainty quantification both in SRPs and reanalysis data is subject to intense research (e.g. Maggioni et al., 2016;Gebremichael, 2010;Awange et al., 2016;Westerberg and Birkel, 2015). The errors quantification of SRPs and reanalysis products is usually done by comparing them with in-situ measurements (e.g. Dembélé and Zwart, 2016;Thiemig et al., 2012;Beck et al., 2019a;Caroletti et al., 2019;Satgé et al., 2020), or by assessing their reliability as forcing for hydrological models (e.g.Duethmann et al., 2013;Pan et al., 2010;Nkiaka et al., 2017). Other evaluation approaches include triple collocation, which is a technique that estimates the variance of unknown errors of three independent variables without a reference or observed variable (e.g. Massari et al., 2017;Alemohammad et al., 2015;McColl et al., 2014;Roebeling et al., 2012). Compared to the ground-truthing approach, the hydrological evaluation approach has received limited attention (Camici et al., 2018;Poméon et al., 2017).

In rainfall-runoff modelling (Peel and McMahon, 2020), the non-linearity of hydrological processes (Blöschl and Zehe, 2005;Clark et al., 2009) can reduce or amplify the errors in the used input rainfall data and result in a satisfactory or poor representation of the hydrological responses (Maggioni and Massari, 2018;Nijssen, 2004). Consequently, the hydrological model can give a good representation of a hydrological state or flux variable for the wrong reasons (cf. Kirchner, 2006), thereby potentially leading to unfortunate consequences for water resources management (Zambrano-Bigiarini et al., 2017). When testing models as hypotheses (Beven, 2018;Pfister and Kirchner, 2017), type I errors (i.e. false positive model acceptability; Beven, 2010) should be avoided to ensure a high predictive skill of the model and its correctness for good decision-making.

This sheds light on the importance of assessing the reliability of hydrological predictions generated with the use of SRPs and reanalysis products (Behrangi et al., 2011;Kuczera et al., 2010). In this context, knowing the adequacy and coherence of meteorological data in reproducing hydrological processes is a prerequisite to data selection for water resources management (Casse et al., 2015;Laiti et al., 2018).

In the context of hydrological evaluation of precipitation datasets, some limitations can be identified in previous studies. Some studies only evaluate a small number of precipitation datasets or do not consider reanalysis products (e.g. Bitew and Gebremichael, 2011;Ma et al., 2018;Liu et al., 2017;Bhattacharya et al., 2019). Usually, the influence of temperature datasets in combination with rainfall datasets is not tested (e.g. Satgé et al., 2019;Camici et al., 2018;Casse et al., 2015;Qi et al., 2016;Zhang et al., 2019), with the exception of a few studies (e.g. Laiti et al., 2018;Lauri et al., 2014), despite the importance of this interaction for evaporation simulation. Most studies evaluate a single hydrological state or flux variable, generally streamflow (e.g. Poméon et al., 2017;Seyyedi et al., 2015;Shayeghi et al., 2020;Li et al., 2012b), or soil moisture (e.g. Brocca et al., 2013). Some studies use lumped or semi-distributed models, therefore averaging the rainfall amount on large areas (e.g. Duan et al., 2019;Tang et al., 2019;Tobin and Bennett, 2014;Gosset et al., 2013;Shawul and Chakma, 2020), which reduces the bias effect that could occur at the pixel level with a fully distributed model. Often, the model is not recalibrated for each precipitation dataset (e.g. Voisin et al., 2008;Su et al., 2008;Li et al., 2012a;Tramblay et al., 2016), which is, however, a prerequisite for reliable input field assessment (Stisen et al., 2012). Moreover, some studies perform a global-scale analysis and ignore regionally tailored products (e.g. Beck et al., 2017b;Mazzoleni et al., 2019;Fekete et al., 2004), which can outperform global products (e.g. Thiemig et al., 2013). Finally, to the best of our knowledge, no study evaluated the simultaneous impact of various precipitation and temperature datasets on the spatial patterns of several hydrological processes (i.e. soil moisture and evaporation).

In light of the above, we propose to study the adequacy of different combinations of 17 precipitation datasets (10 SRPs and 7 reanalysis products) and 6 temperature datasets from reanalysis, when used as forcing data for a fully distributed hydrological model, in reproducing the spatiotemporal variability of multiple hydrological processes (i.e. streamflow, actual evaporation, soil moisture, and terrestrial water storage). In total, 102 rainfall-temperature input data combinations are tested with the mesoscale Hydrologic Model (mHM) by recalibrating the model for each of the input data combinations. The experiment is carried out in the poorly gauged and predominantly semi-arid Volta River Basin (VRB) located in West Africa, over the period 2003-2012. It is noteworthy that the goal of this study is not to estimate the intrinsic quality of the meteorological forcing (i.e. precipitation and temperature) but rather to understand the impact of the propagation of associated uncertainties on the simulation of hydrological processes (Bhuiyan et al., 2019;Falck et al., 2015;Marthews et al., 2020).

The VRB case study is particularly interesting from both scientific and societal perspectives. On the one hand, precipitation modelling in tropical monsoon climates is a challenging task due to strong seasonality and diurnal variations of rainfall (Turner et al., 2011;Pfeifroth et al., 2016;Cook and Vizy, 2019), and due to isolated convection systems in semi-arid regions (Taylor et al., 2017;Mathon et al., 2002;Parker and Diop-Kane, 2017). On the other hand, open access and good quality datasets are needed for water resources management in West Africa (Roudier et al., 2014;Serdeczny et al., 2017;Di Baldassarre et al., 2010;Dinku, 2019). The following research questions are addressed:

1) What is the impact of different gridded rainfall and temperature datasets on the simulation of hydrological fluxes and state variables?
2) How important is the choice of meteorological datasets for the representation of spatial patterns versus temporal dynamics?

Overall, the objective of this work aligns with the efforts in solving the current scientific challenges in hydrology (i.e. uncertainty in large-scale measurements and data, spatial heterogeneity and modelling methods; Blöschl et al., 2019;Wilby, 2019). Moreover, a growing interest in using satellite remote sensing data in hydrological modelling is expected (McCabe et al., 2017;Peters-Lidard et al., 2017;Wilkinson et al., 2016). Therefore, knowing the suitability of the input data for hydrological modelling is a prerequisite for reliable spatiotemporal predictions, as the goal is to increase model performance with minimum uncertainty (Beven, 2016;McMillan et al., 2018;Savenije, 2009).

2 Methodology

2.1 Overview of the modelling experiment

The adequacy of the rainfall and temperature datasets to plausibly reproduce various hydrological processes is tested with all the 102 possible combinations of 17 rainfall and 6 temperature datasets used as meteorological forcing (see section 2.2). Different temperature datasets are used to allow flexibility in rainfall partitioning into evaporation and runoff because temperature is a key variable for the calculation of potential evaporation (Kirchner and Allen, 2020;Zheng et al., 2019;Van Stan et al., 2020). The hydrological model is recalibrated for each of the 102 combinations of rainfall-temperature datasets (Figure 1~~Figure 1~~).

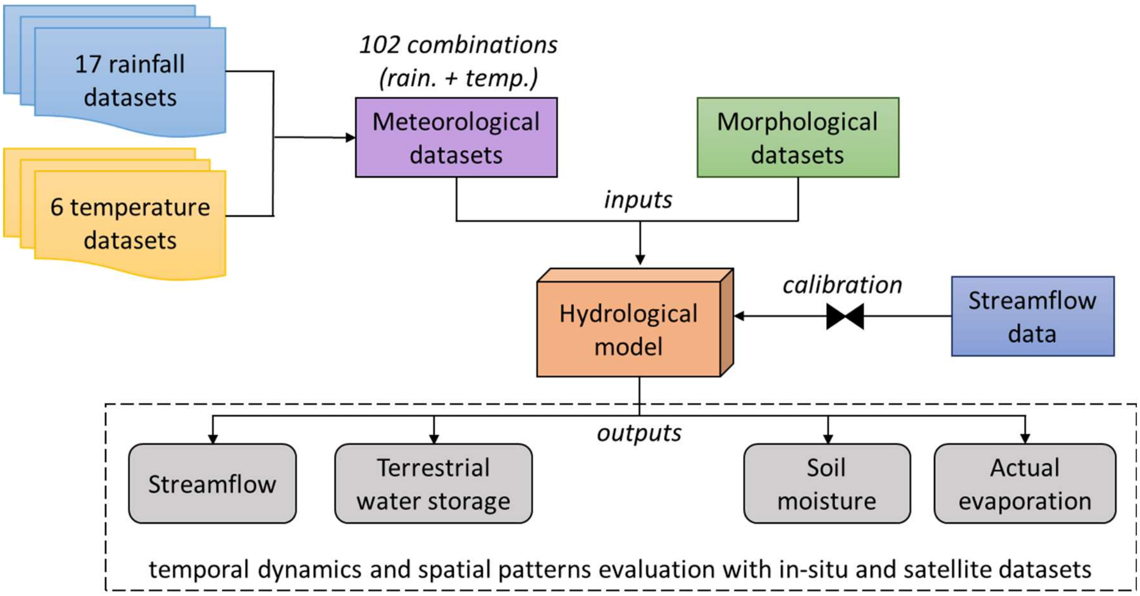


Figure 1. Flowchart of the methodology used to evaluate the suitability of meteorological datasets in reproducing plausible hydrological processes.

The differences in the performance of model outputs are assumed to result from the propagation of the input data uncertainty through the model simulations (Nikolopoulos et al., 2010;Fallah et al., 2020). In case of uncertainties resulting from the hydrological model structure, these uncertainties can be assumed to remain consistent for all the input datasets and therefore it should not hinder the interpretation of the results, because only the parameters change during model calibration and not the model structure (Raimonet et al., 2017).

2.2 Meteorological datasets

This study evaluates 17 rainfall products composed of 10 satellite-based products: TAMSAT, CHIRPS, ARC, RFE, MSWEP, GSMaP, PERSIANN-CDR, CMORPH-CRT, TRMM 3B42 and TRMM 3B42RT; and 7 reanalysis products: JRA-55, EWEMBI, WFDEI-GPCC, WFDEI-CRU, MERRA-2, PGF and ERA5 (Table 1~~Table 1~~). Widely used global and Africa-tailored datasets were selected based on their availability in the period for which streamflow data is available for the hydrological modelling (2000-2012). For SRPs having multiple versions, the gauge-corrected version was selected to avoid the known systematic biases found in the SRPs as compared to ground measurements (Jiang and Wang, 2019;Pellarin et al., 2020). The selected rainfall datasets include single and multi-sensor, with various merged and gauge-corrected products obtained from rain gauges, microwave sensors on low Earth orbits and infrared sensors on geostationary satellites (Maggioni and Massari, 2018;Thiemig et al., 2013;Golian et al., 2019). Moreover, six different datasets of air temperature (at 2 m above ground) are used for the calculation of potential evaporation and they are obtained from the reanalysis products: JRA-55, EWEMBI, WFDEI, MERRA-2, PGF and ERA5.

175 **Table 1. Meteorological datasets with used spatial resolution; the table presents the characteristics of the datasets used in this study, although different spatial and temporal resolutions can be available from the data providers. G: gauge, S: satellite, R: reanalysis, P: precipitation, T: temperature, NP: near-present.**

Datasets	Name/ website	Data sources	Used variables	Spatial coverage	Spatial resolution	Temporal coverage	Temporal resolution	References
TAMSAT v3.0	Tropical Applications of Meteorology using SATellite (TAMSAT), African Rainfall Climatology and Time-series (TARCAT) https://www.tamsat.org.uk/data/archive	S, G	P	Africa 38°N – 36°S, 19°W – 52°E	0.0375°	1983-NP	daily	Maidment et al. (2017), Tarnavsky et al. (2014), Maidment et al. (2014), Maidment et al. (2020)
CHIRPS v2.0	Climate Hazards group InfraRed Precipitation with Stations (CHIRPS) V2.0 http://chg.ucsb.edu/data/chirps/	S, G, R	P	Land 50° N/S, 180° E/W	0.05°	1981-NP	daily	Funk et al. (2015)
ARC v2.0	Africa Rainfall Estimate Climatology (ARC 2.0) https://www.cpc.ncep.noaa.gov/products/international/data.shtml	S, G	P	Africa 40°N – 40°S, 20°W – 55°E	0.1°	1983-NP	daily	Novella and Thiaw (2013)
RFE v2.0	Climate Prediction Center (CPC) African Rainfall Estimate (RFE) https://www.cpc.ncep.noaa.gov/products/international/data.shtml	S, G	P	Africa 40°N – 40°S, 20°W – 55°E	0.1°	2001-NP	daily	Xie and Arkin (1996), Herman et al. (1997)
MSWEP v2.2	Multi-Source Weighted-Ensemble Precipitation (MSWEP) V2.2 http://www.gloh2o.org/	S, G, R	P	Global	0.1°	1979-NP	3-hourly	Beck et al. (2017a)
GSMaP-std v6	Global Satellite Mapping of Precipitation (GSMaP) Moving Vector with Kalman (MVK) Standard V6 https://sharaku.eorc.jaxa.jp/GSMaP/	R, G	P	60◦ N/S, 180° E/W	0.1°	2001-2013	daily	Ushio et al. (2009), Ushio et al. (2019), Kubota et al. (2020)
PERSIANN-CDR v1r1	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) Climate Data Record (CDR) V1R1 http://chrsdata.eng.uci.edu/	S, G	P	60◦ N/S, 180° E/W	0.25°	1983-2016	6-hourly (daily)	Ashouri et al. (2015)
CMORPH-CRT v1.0	Climate Prediction Center (CPC) MORPHing technique (CMORPH) bias corrected (CRT) V1.0 www.cpc.ncep.noaa.gov	S, G	P	60◦ N/S, 180° E/W	0.25°	1998-2015	daily	Joyce et al. (2004), Xie et al. (2017)

TRMM 3B42 v7	TRMM Multi-satellite Precipitation Analysis (TMPA) 3B42 V7 https://mirador.gsfc.nasa.gov/	S, G	P	50◦ N/S, 180° E/W	0.25°	2000-2017	3-hourly	Huffman et al. (2007)
TRMM 3B42 RT v7	TRMM Multi-satellite Precipitation Analysis (TMPA) 3B42 Real Time V7 https://mirador.gsfc.nasa.gov/	S	P	50◦ N/S, 180° E/W	0.25°	2000-NP	3-hourly	Huffman et al. (2007)
WFDEI-CRU	WATCH Forcing Data ERA-Interim (WFDEI) corrected using Climatic Research Unit (CRU) dataset www.eu-watch.org	R, G	P, T	Global	0.5°	1979-2018	3-hourly	Weedon et al. (2014)
WFDEI-GPCC	WATCH Forcing Data ERA-Interim (WFDEI) corrected using Global Precipitation Climatology Centre (GPCC) dataset ftp://rfdata:forceDATA@ftp.iiasa.ac.at/	R, G	P, T	Global	0.5°	1979-2016	3-hourly	Weedon et al. (2014)
PGF v3	Princeton University global meteorological forcing (PGF) http://hydrology.princeton.edu/data/pgf/	R, G	P, T	Global	0.25°	1948-2012	3-hourly	Sheffield et al. (2006)
ERA5	European Centre for Medium-range Weather Forecasts ReAnalysis 5 (ERA5) hourly data on single levels https://cds.climate.copernicus.eu/	R	P, T	Global	0.25°	1979-NP	hourly	Hersbach et al. (2018), Hersbach et al. (2020)
MERRA-2	Modern-Era Retrospective Analysis for Research and Applications 2 (rainfall: M2T1NXFLX_V5.12.4; temperature: M2SDNXSLV_V5.12.4) https://disc.gsfc.nasa.gov/datasets/	S, G, R	P, T	Global	0.625° x 0.5°	1980-NP	hourly	Gelaro et al. (2017), Reichle et al. (2017)
EWEMBI v1.1	EartH2Observe, WFDEI and ERA-Interim data Merged and Bias-corrected for ISIMIP (EWEMBI) http://doi.org/10.5880/pik.2016.004	R, G	P, T	Global	0.5°	1976-2013	daily	Lange (2016)
JRA-55	Japanese 55 year ReAnalysis (JRA-55); rainfall: fcst_phy2m125; temperature: anl_surf125 https://jra.kishou.go.jp/JRA-55/index_en.html	R	P, T	Global	1.25°	1959-NP	3-hourly	Kobayashi et al. (2015)

2.3 Modelling datasets

In addition to the meteorological datasets (Table 1~~Table 1~~), an ensemble of datasets is required for the set-up and the calibration and evaluation of the hydrological model (Table 2~~Table 2~~). The streamflow datasets obtained from different organizations (see acknowledgements) were pre-processed (i.e. gap-filling and quality control) in the work of Dembélé et al. (2019).

Table 2. Modelling datasets. ESA CCI SM: European Space Agency Climate Change Initiative Soil Moisture; GIMMS: Global Inventory Modelling and Mapping Studies; GLEAM: Global Land Evaporation Amsterdam Model; GLiM: Global Lithological Map; GMTED: Global Multi-resolution Terrain Elevation Data; GRACE: Gravity Recovery and Climate Experiment; WFDEI: WATCH Forcing Data methodology applied to ERA-Interim data.

Variables	Products	Spatial resolution	Temporal resolution	References
Morphological data				
Terrain characteristics (elevation, slope, aspect, flow direction and flow accumulation)	GMTED 2010	225 m (0.0021°)	static	Danielson and Gesch (2011) https://topotools.cr.usgs.gov/
Soil properties (horizon depth, bulk density, sand and clay content,)	SoilGrids	250 m (0.0023°)	static	Hengl et al. (2017) https://www.isric.org/explore/soilgrids
Geology	GLiM v1.0	0.5°	static	Hartmann and Moosdorf (2012) https://doi.pangaea.de/10.1594/PANGAEA.788537
Land use land cover	Globcover 2009	300 m (0.0028°)	static	Bontemps et al. (2011) http://due.esrin.esa.int/page_globcover.php
Phenology (leaf area index)	GIMMS	8 km (0.0833°)	bimonthly	Tucker et al. (2005), Zhu et al. (2013) http://cliveg.bu.edu/modismisr/lai3g-fpar3g.html
Model calibration/evaluation				
Streamflow	-	point	daily	Multiple organizations (see acknowledgements)
Terrestrial water storage anomaly (S_t)	GRACE TellUS v5.0	1°	monthly	Tapley et al. (2004), Landerer and Swenson (2012) https://grace.jpl.nasa.gov/
Surface soil moisture (S_u)	ESA CCI SM v4.2	0.25°	daily	Dorigo et al. (2017) https://www.esa-soilmoisture-cci.org/
Actual evaporation (E_a)	GLEAM v3.2a	0.25°	daily	Martens et al. (2017), Miralles et al. (2011) https://www.gleam.eu/

Multiple satellite datasets are used to evaluate the modelled hydrological fluxes and state variables. For the evaluation of the modelled water storages, the GRACE-derived terrestrial water storage (S_t) anomaly data release RL05 (Landerer and Swenson, 2012;Swenson, 2012) is used. The ensemble mean of different products from three processing centers (i.e. Jet Propulsion Laboratory, Center for Space Research at University of Texas, and Geoforschungs Zentrum Potsdam) is preferred because it is more effective in reducing noise in the Earth's gravity signal as compared to the individual products (Sakumura et al., 2014). The surface soil moisture (S_u) data representing the first soil layer (i.e. 2-5 cm depth) is obtained from ESA CCI (Dorigo et al., 2017) using the combination of both active and passive microwave products (Gruber et al., 2017;Wagner et al., 2012b). Actual evaporation (E_a) data is obtained from the GLEAM land surface model that aggregates components of terrestrial evaporation based on the fraction of land cover types per grid cell (Martens et al., 2017). A full description of the datasets is accessible through the references and web links provided in Table 1~~Table 1~~ and Table 2~~Table 2~~.

2.4 Study Area

The transboundary Volta River Basin (VRB) covers approximately 415,600 km^2 (Figure 2~~Figure 2~~) shared among six countries of West Africa (i.e. Burkina Faso, Ghana, Togo, Mali, Benin and Côte d'Ivoire). The relief is predominantly flat with 95% of the basin below 400 m a.s.l (De Condappa and Lemoalle, 2009). The Volta River flows over 1,850 km with a drainage system composed of four sub-basins known as Black Volta (152,800 km^2), White Volta (113,400 km^2), Oti (74,500 km^2), and Lower Volta (74,900 km^2). Before reaching the Atlantic Ocean at the Gulf of Guinea, the Volta River transits in the Lake Volta (area: 8,502 km^2; volume: 148 km^3) formed by the Akosombo dam (7.94 10^6 m^3) (Williams et al., 2016;Dembélé et al., 2020b). The dominant land cover is savannah composed of grassland interspersed with shrubs and trees over 75% of the basin area, followed by cropland (13%), forest (9%), water bodies (2%) and bare land and settlements (1%). Climate in West Africa is unique and complex (Berthou et al., 2019;Bichet and Diedhiou, 2018;Nicholson et al., 2018a). The seasonal and latitudinal oscillation of the Inter-Tropical Convergence Zone (ITCZ) is the predominant rainfall generation mechanism in West Africa (Biasutti, 2019), thereby depicting a south-north gradient of increasing aridity in the VRB. The ITCZ is a narrow belt of clouds associated with intense convective activity resulting from the near-surface convergence of warm and moist trade winds (Schneider et al., 2014;Dezfuli, 2017). The warm northeasterly Harmattan winds emanate from the Sahara and the moist southwest monsoon winds originate in the Atlantic ocean (Nicholson, 2013;Vizy and Cook, 2018). Rainfall in West Africa is characterized by its interannual and multidecadal variability (Biasutti et al., 2018;Thorncroft et al., 2011;Nicholson et al., 2018b). Four eco-climatic zones (i.e. Sahelian, Sudano-Sahelian, Sudanian and Guinean; Table 3~~Table 3~~) are commonly identified based on the average annual precipitation and agricultural features (FAO/GIEWS, 1998;Mul et al., 2015). The aridity index in Table 3~~Table 3~~ is derived from the global aridity index database (Trabucco and Zomer, 2018). The maps of spatial patterns of rainfall and temperature in the VRB for different datasets are shown in Appendix A1~~A1~~ and Appendix A2~~A2~~. The climatology of rainfall and temperature per climatic zones are provided in the Supporting Information (SI, Figures ~~S3~~S11-~~S6~~S14).

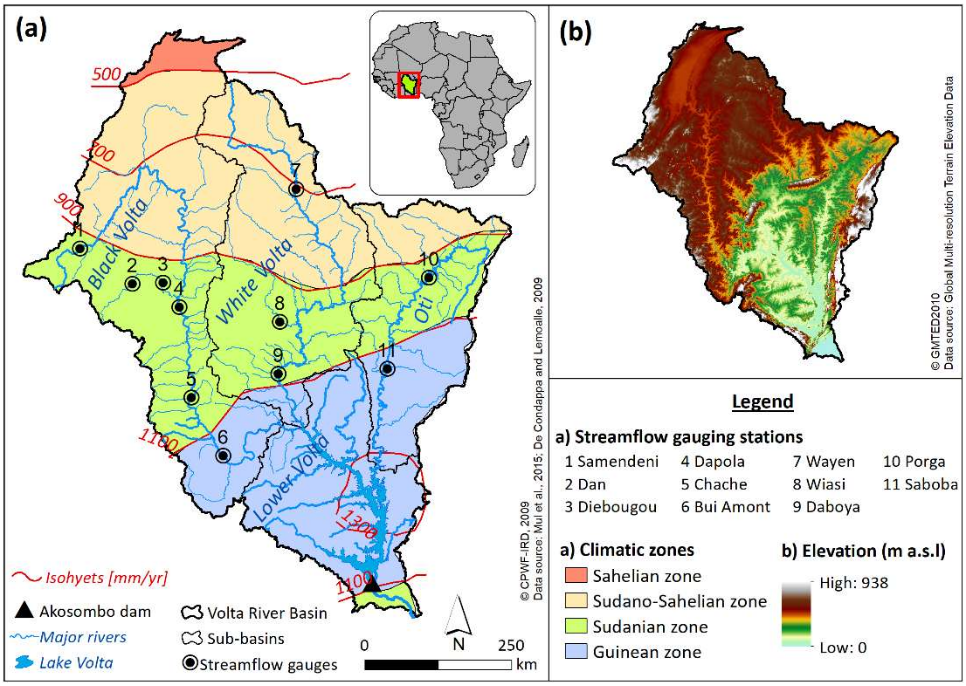


Figure 2. Physical and hydroclimatic characteristics of the Volta River basin.

Table 3. Characteristics of the four eco-climatic zones in the Volta River basin. The mean and range ([min-max]) values are given for the Aridity Index (*AI*).

Eco-climatic zones	Climate class	*AI* (-)
Sahel Savanna	Arid	0.16 [0.12-0.20]
Sudano-Sahelian	Semi-arid	0.29 [0.16-0.43]
Sudanian Savanna	Semi-arid/ Dry sub-humid	0.47 [0.33-0.98]
Guinean Savanna	Dry sub-humid/ Humid	0.70 [0.49-1.22]

2.5 Hydrological Model Setup

The fully distributed mesoscale Hydrologic Model (mHM, version 5.9; Samaniego et al., 2010;Kumar et al., 2013) is used in this study. It is a conceptual model that simulates dominant hydrological processes (e.g. evaporation, soil moisture, subsurface storage, and discharge) per grid cell in the modelling domain. The Muskingum-Cunge method (Cunge, 1969) is used for routing the total grid-generated runoff using a multiscale routing model (Thober et al., 2019). A multiscale parameter regionalization technique (MPR; Samaniego et al., 2017) is used to account for sub-grid variability of the basin physical characteristics (e.g. soil texture, topography and land cover). For this study, 36 global parameters are determined through model calibration (Table S24~~18~~ in the Supporting Information).

In this study, the Hargreaves and Samani method (Hargreaves and Samani, 1985), solely based on air temperature data, is used to calculate the reference evaporation (E_{ref}). Potential evaporation (E_{p}) is calculated by adjusting E_{ref} to vegetation cover (Allen et al., 1998;Birhanu et al., 2019). A dynamical scaling function (F_{DS}) (cf. Demirel et al., 2018) is used to account for vegetation-climate interactions (Bai et al., 2018;Jiao et al., 2017). E_{p} is formulated as follows:

$$E_{\mathrm{p}} = F_{\mathrm{DS}} \cdot E_{\mathrm{ref}}, \text{ with} \tag{1}$$

$$F_{\mathrm{DS}} = a + b\left(1 - e^{(c \cdot I_{\mathrm{LA}})}\right) \tag{2}$$

where I_{LA} represents the leaf area index, a is the intercept term, b represents the vegetation dependent component, and c describes the degree of nonlinearity in the I_{LA} dependency. The coefficients a, b, and c are determined during model calibration. Actual evaporation (i.e. all evaporative fluxes including transpiration, E_{a}) depends on plant water availability, i.e. on root distribution in the subsurface and soil moisture availability (Feddes et al., 1976); this is emulated in mHM by computing E_{a} as a fraction of E_{p} at different soil layers. A multi-layer infiltration capacity approach is used to calculate soil moisture based on a three-layer soil scheme (5 cm, 30 cm and 100 cm depths). As no snow occurs in the VRB, terrestrial water storage is calculated per grid cell by summing up the surface water storage on impervious areas and all subsurface water storage (i.e. reservoirs generating soil moisture, baseflow and interflow). The model is run at a daily time step with a spatial discretization of 0.25° (~28 km at the equator).

The modelling experiment covers the period 2000-2012 with 3-year model warm-up period (2000-2002), 6 years for model calibration (2003-2008) and 4 years for model evaluation (2009-2012). The model is calibrated and evaluated with the available daily in-situ streamflow datasets from 11 locations (Figure 2~~Figure 2~~a), while the evaluation with satellite datasets of evaporation, soil moisture and terrestrial water storage is done at a monthly time step to avoid the impact of mismatches in the daily data retrieval periods among the satellite data sources. An illustration of natural variability of streamflow (Figure S16), precipitation (Figures S1 and S5) and temperature (Figures S3-S4 and S6-8) are provided in the Supporting Information (SI).

2.6 Multisite model calibration on streamflow data

A multisite calibration strategy is adopted by simultaneously constraining the model with the 11 streamflow (Q) gauging stations (Figure 2~~Figure 2~~) to infer a unique parameter set for the whole basin. The objective function Φ_{Q} combines the Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970) of streamflow (E_{NS}) and the Nash-Sutcliffe efficiency of the logarithm of streamflow (E_{NSlog}), and it is formulated such that it has to be minimized:

$$\Phi_{\mathrm{Q}} = \frac{1}{g}\sum_1^g \sqrt{(1 - E_{\mathrm{NS}})^2 + \left(1 - E_{\mathrm{NSlog}}\right)^2}, \text{ with} \tag{3}$$

$$E_{\mathrm{NS}} = 1 - \frac{\sum_1^t \left(Q_{\mathrm{mod}}(t) - Q_{\mathrm{obs}}(t)\right)^2}{\sum_1^t \left(Q_{\mathrm{obs}}(t) - \overline{Q_{\mathrm{obs}}}\right)^2} \text{ and} \tag{4}$$

$$E_{\mathrm{NSlog}} = 1 - \frac{\sum_1^t \left[\log\left(Q_{\mathrm{mod}}(t)\right) - \log\left(Q_{\mathrm{obs}}(t)\right)\right]^2}{\sum_1^t \left[\log\left(Q_{\mathrm{obs}}(t)\right) - \overline{\log\left(Q_{\mathrm{obs}}\right)}\right]^2} \tag{5}$$

where Q_{mod} and Q_{obs} are the modelled and the observed streamflow, t is the number of time steps of the calibration period, and g is the number of streamflow gauging stations present within the modelling domain. Φ_Q is calculated with all the streamflow gauging stations, and it ranges from its ideal value that is 0 to positive infinity.

~~The multi-objective Kling-Gupta efficiency~~ ~~(E_{KG}) (Kling et al., 2012) is used for the formulation of the objective function~~ ~~Φ_Q, which has to be minimized and is formulated as follows:~~

$$\cancel{\Phi_Q = 1 - \left[\frac{1}{g}\sum_{i=1}^{g} E_{KG,i}\left(Q_{mod,i}, Q_{obs,i}\right)\right], \text{ with}} \quad \cancel{(3)}$$

$$\cancel{E_{KG} = 1 - \sqrt{(r_{KG} - 1)^2 + (\beta_{KG} - 1)^2 + (\gamma_{KG} - 1)^2}} \quad \cancel{(4)}$$

~~Where~~ ~~g is the number of gauging stations,~~ ~~r_{KG} is the Pearson correlation coefficient,~~ ~~β_{KG} is the bias term (i.e. the ratio of the means), and~~ ~~γ_{KG} is the variability term (i.e. the ratio of the coefficients of variation) between the observed~~ ~~(Q_{obs}) and modelled~~ ~~(Q_{mod}) streamflow, with~~ ~~μ and σ representing the mean and the standard deviation. The~~ ~~E_{KG} ranges from negative infinity to its optimal value that is unity. As a reference,~~ ~~E_{KG} > -0.41 indicates that the model is better than the mean observed flow (Knoben et al., 2019).~~ ~~Φ_Q ranges from its optimal value that is 0 to positive infinity.~~

The model is calibrated solely with Q data because it is the only available in-situ measurement, and to avoid potential trade-offs of a multivariate calibration that would result in difficulties in identifying the source of variation in the model performance (i.e. input data vs. model parametrization) (Dembélé et al., 2020b). The parameter estimation is done with the dynamically dimensioned search algorithm (Tolson and Shoemaker, 2007) using 4,000 iterations for each of the 102 rainfall-temperature dataset combinations.

2.7 Multivariable model evaluation with streamflow and satellite data

In addition to E_{NS} and E_{NSlog}, the Kling-Gupta efficiency (E_{KG}) (Kling et al., 2012) is used to evaluate the model performance for streamflow.

$$E_{KG} = 1 - \sqrt{(r_{KG} - 1)^2 + (\beta_{KG} - 1)^2 + (\gamma_{KG} - 1)^2} \quad (6)$$

where r_{KG} is the Pearson correlation coefficient, β_{KG} is the bias term (i.e. the ratio of the means), and γ_{KG} is the variability term (i.e. the ratio of the coefficients of variation) between Q_{obs} and Q_{mod}. The E_{KG} ranges from negative infinity to its optimal value that is unity. As a reference, E_{KG} > -0.41 indicates that the model is better than the mean observed flow (Knoben et al., 2019).

In addition to Q, several non-commensurable and satellite-based variables are used for model evaluation (Table 2~~Table 2~~). ~~The model performance for Q is evaluated with E_{KG}.~~ The bias-insensitive Pearson's correlation coefficient (r) is used to assess the temporal dynamics of S_t, S_u and E_a because the model is not calibrated on these variables, and their evaluation datasets are satellite-derived products that encompass uncertainties and can be biased.

The spatial pattern representation of hydrological processes is assessed by using a bias-insensitive and multi-component metric developed by Dembélé et al. (2020b). The proposed spatial pattern efficiency (E_{SP}) metric is formulated similarly to the E_{KG}

(Equation 4) but it focuses only on the spatial pattern of variables rather than on their absolute values (like the SPAEF; Koch et al., 2018). E_{SP} simultaneously assesses the dynamics, the spatial variability, and the locational matching of grid cells between the observed (X_{obs}) and modelled (X_{mod}) variables. Considering two variables X_{obs} and X_{mod} composed of n cells, E_{SP} is defined as follows:

$$E_{\mathrm{SP}} = 1 - \sqrt{(r_{\mathrm{s}} - 1)^2 + (\gamma - 1)^2 + (\alpha - 1)^2}, \text{ with} \tag{7}$$

$$r_{\mathrm{s}} = 1 - \frac{6 \sum_1^n d_{\mathrm{i}}^2}{n(n^2-1)}, \tag{8}$$

$$\gamma = \frac{\frac{\sigma_{\mathrm{mod}}}{\mu_{\mathrm{mod}}}}{\frac{\sigma_{\mathrm{obs}}}{\mu_{\mathrm{obs}}}} \text{ and} \tag{9}$$

$$\alpha = 1 - E_{\mathrm{RMS}}\left(Z_{X_{\mathrm{mod}}}, Z_{X_{\mathrm{obs}}}\right) \tag{10}$$

where r_s is the Spearman rank-order correlation coefficient with d_i the difference between the ranks of the i^{th} cell of X_{mod} and X_{obs}. γ is the variability ratio (i.e. the ratio of the coefficients of variation) that assesses the similarity in the dispersion of the probability distributions of X_{mod} and X_{obs}, with μ and σ representing the mean and the standard deviation, and α the spatial location matching term calculated as the root mean squared error (E_{RMS}) of the standardized values (z-scores, Z_X) of X_{mod} and X_{obs} (Dembélé et al., 2020b). E_{SP} ranges from negative infinity to 1, which is its optimal value. E_{SP} does not have an inherent benchmark, also like E_{KG} (Knoben et al., 2019). For $E_{SP} = 0$, the ranks of the observed and modelled variables are moderately related (i.e. $r_s = 0.55$), while no association among the ranks (i.e. $r_s = 0$) results in $E_{SP} = -0.67$ (cf. Supplementary Material of Dembélé et al., 2020b). However, the main point of using E_{SP} here is not to strictly conclude how well the modelled spatial patterns reproduce the observed patterns, otherwise a benchmark should be used (Schaefli and Gupta, 2007;Seibert et al., 2018), but rather to determine if a modelled spatial pattern is better than another. The spatial pattern evaluation is completed for S_u and E_a, while only the temporal dynamics of S_t are assessed due to the coarse spatial resolution of the GRACE data.

The relative variation in model performance is assessed with the second-order coefficient of variation (V_2) (Kvålseth, 2017). V_2 is an alternative to the classic Pearson's coefficient of variation (~~V~~CV), which has significant limitations that are comprehensively discussed by Kvålseth (2017). The limitations of the CV include its difficult and non-intuitive interpretation because of the lack of an upper bound, its high sensitivity to outliers, its dependence on the sample mean and problems with negative values. For all sample data $x = (x_1, \ldots, x_n) \in R^n$, with $R = (-\infty, \infty)$, V_2 is defined as follows:

$$V_2 = \left(\frac{s^2}{s^2 + \bar{x}^2}\right)^{1/2} \tag{11}$$

where s is the standard deviation and $\bar{x}$ is the mean of x. V_2 varies from 0 to 1 or 0% to 100%, and represents the distance between x and $\bar{x}$ relative to the distance between x and the origin zero.

3 Results

The results are presented and discussed for the entire simulation period (2003-2012, i.e. combined calibration and evaluation periods) because reliable meteorological datasets are expected to produce a plausible representation of hydrological processes independently of the modelling period (Bisselink et al., 2016). Separated results are provided for the calibration and evaluation periods in the SI ~~Supporting Information (SI)~~.

3.1 Model performance for streamflow

Similar model performance patterns are obtained with E_{KG}, E_{NS} and E_{NSlog} of daily streamflow (Q) (Figure 3~~Figure 3~~). Therefore, only E_{KG} is retained for the description of the results. ~~For daily streamflow (Q), a~~All input dataset combinations show a median $E_{KG} > 0.5$, except those having JRA-55 as rainfall input (Figure 3~~Figure 3~~), which~~;, this~~ can be justified by the coarse spatial resolution of that product. The ranking of the rainfall and temperature datasets based on the model performance for Q is provided in Appendix A3~~A3~~. The analysis of model performance for Q is done for the entire VRB and not per climatic zone due to the limited number of stations. As expected, the discrepancies in median E_{KG} are more pronounced across rainfall datasets than across temperature datasets, as visible in the color-coded ranking of the products in Figure 3~~Figure 3~~. For a given rainfall product, the ranking among all rainfall products hardly varies with different temperature products. The ranking of all the datasets for the model performance for Q is also summarized in Appendix A3~~A3~~. The overall stronger impact of the choice of the rainfall dataset on E_{KG} of Q becomes also clear from the second-order coefficient of variations (V_2) of the median E_{KG} (Table S3~~1~~ in SI). For rainfall datasets, the V_2 across temperature datasets varies between 0.5% for GSMaP-std and 4% for JRA-55, with an average V_2 of 2%. For temperature datasets, the V_2 of median E_{KG} of Q across rainfall datasets varies between 10% for MERRA-2 and 12% for ERA5, with an average V_2 of 11%. This result suggests that the choice of a rainfall dataset has a stronger impact on the E_{KG} of Q than the choice of a temperature dataset.

The analysis of the components of E_{KG} (i.e. the Pearson correlation r_{KG}, the bias β_{KG} and the variation γ_{KG}) reveals that, when choosing a rainfall dataset, there is more uncertainty in the bias of Q (V_2 = 14%) than in its variability (V_2 = 6%) and in its dynamics (V_2 = 3%), which is in agreement with the work of Thiemig et al. (2013). Detailed results on the performance for Q (i.e. E_{NS}, E_{NSlog}, E_{KG}, r_{KG}, β_{KG} and γ_{KG}) and the ranking of the datasets with separate results for the calibration and evaluation periods are provided in the SI (Tables S1-S18~~2~~, Figures S17-S26~~11~~).

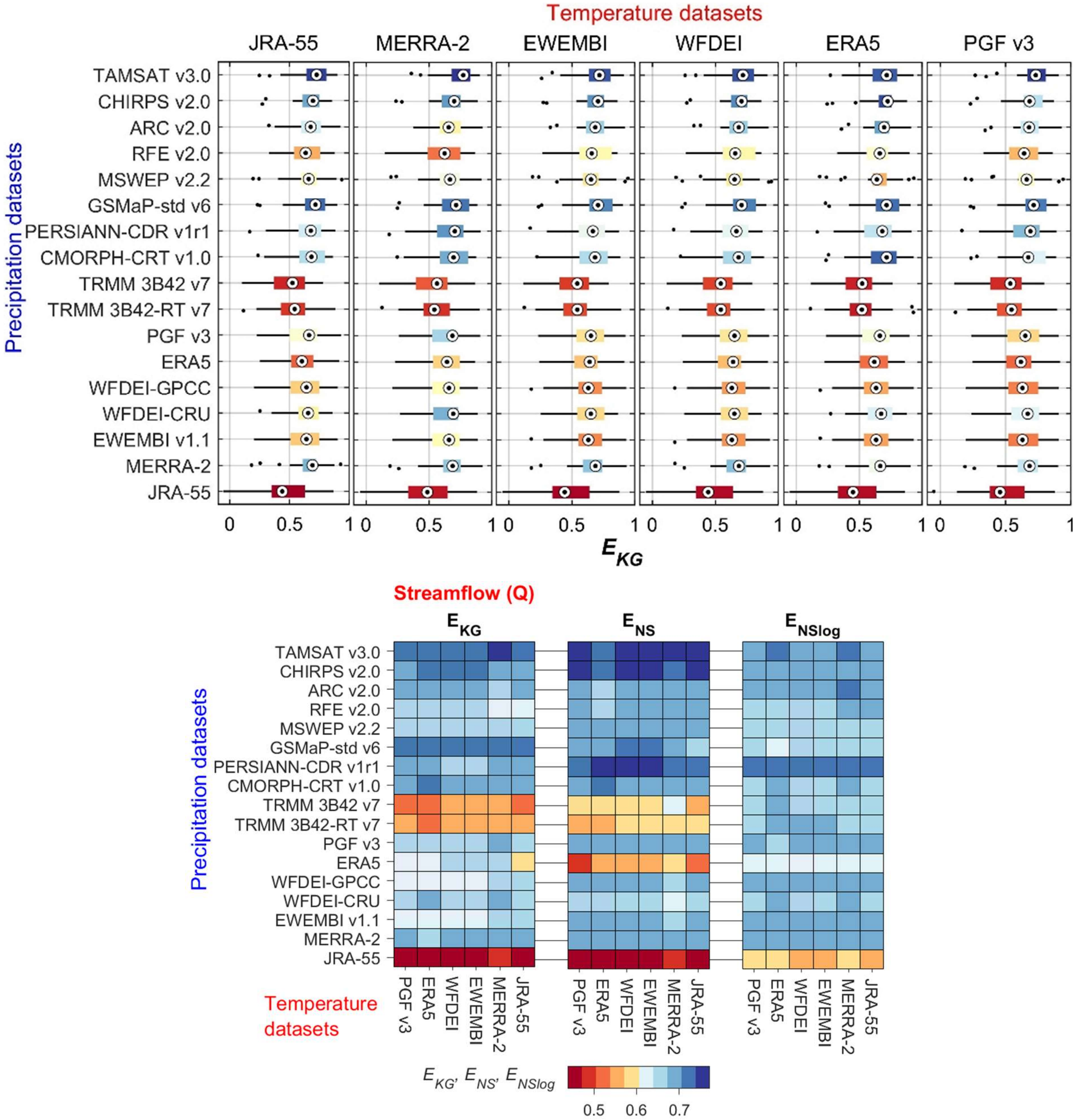


Figure 3. <u>**Kling-Gupta efficiency (E_{KG}), Nash-Sutcliffe efficiency (E_{NS}) and Nash-Sutcliffe efficiency of the logarithm (E_{NSlog}) of daily streamflow (Q) over the simulation period (2003-2012) for 102 combinations of 17 rainfall datasets (y-axis) and 6 temperature datasets (x-axis) used as forcing for the hydrological model.**</u> ~~**Kling-Gupta efficiency (E_{KG}) of daily streamflow (Q) over the simulation period (2003-2012) for 102 combinations of 17 rainfall datasets (y-axis) and 6 temperature datasets (subplots on x-axis) used as inputs in the mHM model. Each boxplot has 22 values representing the combined performance for the calibration and evaluation periods for 11 streamflow gauging stations. The boxplots are colored from the best (blue) to the worst performance (red) based on the median value.**~~

3.2 Model performance for terrestrial water storage

The model performance for the temporal dynamics of monthly terrestrial water storage (S_t) compared to the GRACE product is shown in Figure 4~~Figure 4~~ (see the SI for monthly time series, Figures ~~S23~~S38-S42~~27~~). The average Pearson correlation coefficient (r) of S_t for all datasets in the entire VRB is 0.80, with discrepancies across climatic zones. The driest and wettest climatic zones show the lowest performances, i.e. Sahelian (r = 0.67) and Guinean (r = 0.60) zones, compared to the intermediate climatic zones, i.e. Sudano-Sahelian (r = 0.72) and Sudanian (r = 0.79) zones. Appendix A3~~A3~~ provides the ranking of all the meteorological datasets for the model performance for S_t.

The rainfall datasets show different performances across climatic zones, with ARC showing the highest score for all the climatic zones except the Guinean zone, where CMORPH-CRT ranks first. The choice of the rainfall dataset leads to an average V_2 of 15% for the r of S_t, while the average V_2 is 5% for the choice of the temperature dataset. Detailed results are provided in the SI (Tables S13~~9~~, Figures ~~S12~~S27-S37~~22~~).

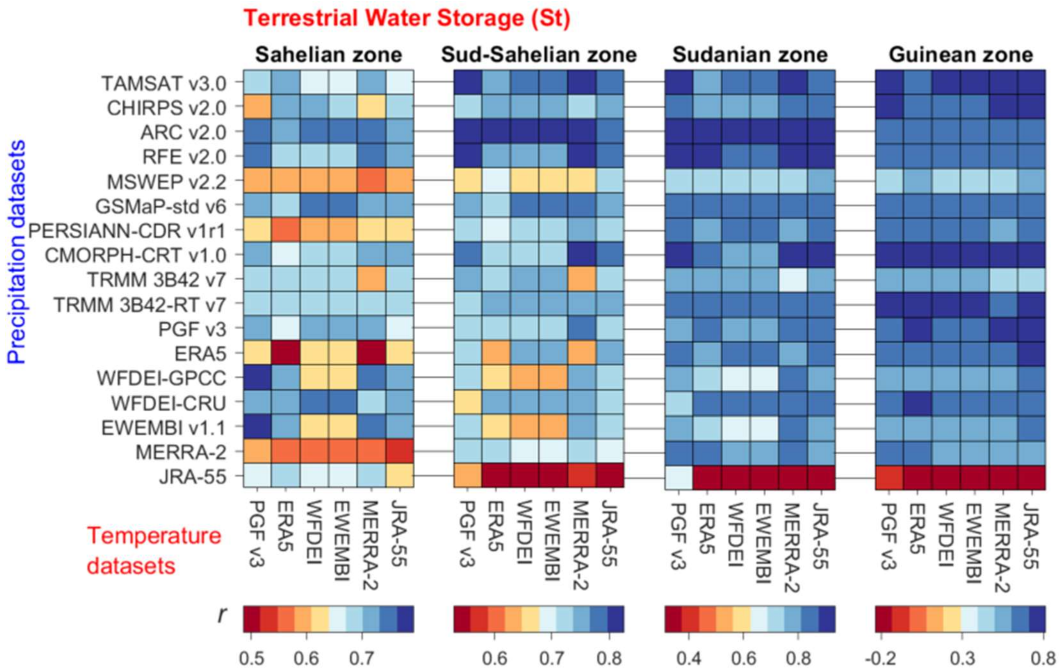


Figure 4. Pearson correlation coefficient (*r*) of modelled terrestrial water storage compared to GRACE data in four climatic zones in the Volta River basin over the simulation period (2003-2012) considering 102 combinations of rainfall (y-axis) and temperature datasets (subplots on x-axis) used as forcing for the hydrological model.

3.3 Model performance for soil moisture

Figure 5~~Figure 5~~ shows the model performance for the temporal dynamics of monthly soil moisture (S_u) compared to the ESA CCI product (see the SI for monthly time series, Figures S54~~39~~-S58~~43~~). The average r of S_u for the entire VRB over all datasets is 0.93. The r of S_u decreases from the drier to the wetter climatic zones: Sahelian (r = 0.94), Sudano-Sahelian (r = 0.94), Sudanian (r = 0.92) and Guinean (r = 0.86). The ranking of the meteorological datasets based on the model performance for S_u is provided in Appendix A3~~A3~~. EWEMBI and WFDEI-GPCC show the highest performance in the Sahelian and Suadano-Sahelian zones respectively, while MERRA-2 shows the highest performance in the Sudanian and Guinean zones. The choice of the rainfall dataset leads to an average V_2 of 4% for the temporal dynamics of S_u, while the average V_2 is 2% for the choice of the temperature dataset.

The spatial patterns of S_u show considerable differences when using different combinations of rainfall and temperature input datasets, as illustrated in Figure 6~~Figure 6~~ (see similar maps for all the meteorological datasets in the SI, Figures ~~S44~~S59-S60~~45~~). The south-north gradient of increasing aridity is not similarly spread among the rainfall-temperature dataset combinations. More interestingly, west-east differences in the spatial patterns of S_u can be observed. These differences in spatial pattern reproduction can also be seen in the spatial pattern efficiency metric (E_{SP}) of S_u for the 102 rainfall-temperature dataset combinations (Figure 7~~Figure 7~~). The average E_{SP} of S_u in the VRB over all datasets is -0.11.

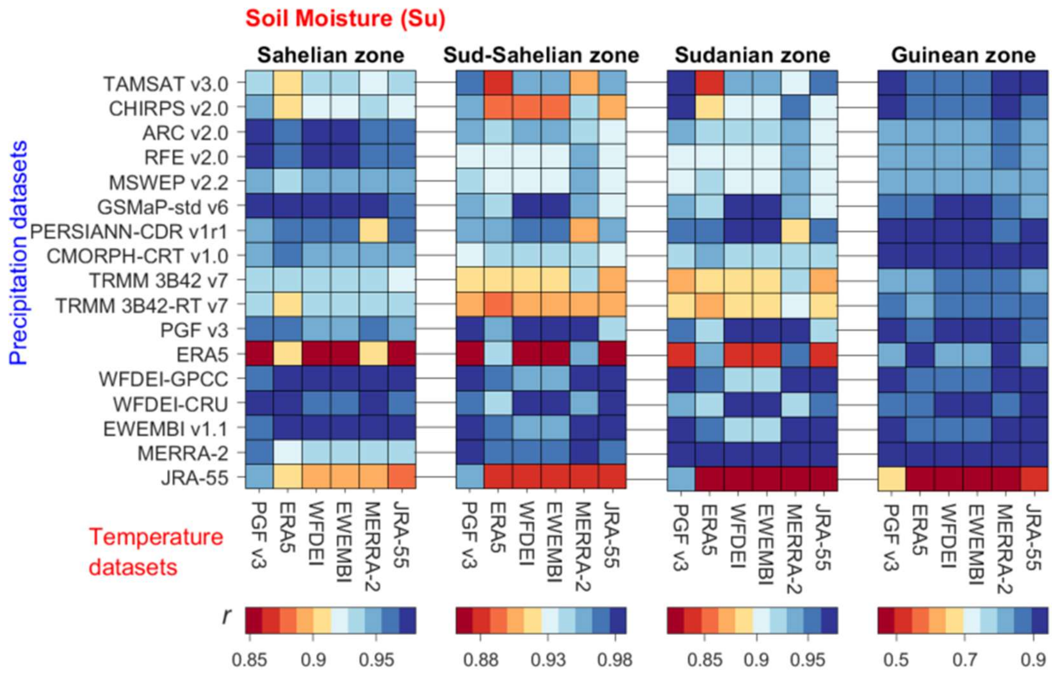


Figure 5. Pearson correlation coefficient (r) of modelled soil moisture (S_u) compared to ESA CCI data over the simulation period (2003-2012) considering 102 combinations of rainfall (y-axis) and temperature datasets (subplots on x-axis) used as forcing for the hydrological model.

For the entire VRB, the choice of the rainfall dataset leads to an average variation of 61% for the E_{SP} of S_u, while the choice of the temperature dataset involves a variation of 45%. Lower impacts of data choices are observed in the climatic zones where the climate is homogeneous as compared to the entire VRB. The choice of a rainfall dataset is more critical for the E_{SP} of S_u in the driest and wettest climatic zones, i.e. Sahelian (E_{SP} = -0.47, V_2 = 25%) and Guinean (E_{SP} = -0.40, V_2 = 26%) zones, than the intermediate zones, i.e. Sudano-Sahelian (E_{SP} = -0.37, V_2 = 11%) and Sudanian (E_{SP} = -0.39, V_2 = 17%) zones. A smaller impact on the E_{SP} of S_u is observed for the choice of the temperature dataset: Sahelian (V_2 = 8%), Guinean (V_2 = 19%), Sudano-Sahelian (V_2 = 5%) and Sudanian (V_2 = 9%) zones. Detailed results on the model performance for S_u and the ranking of the datasets for the calibration and evaluation periods are provided in the SI (Tables S20~~14~~-S21~~15~~, Figures ~~S28~~S43-S53~~38~~).

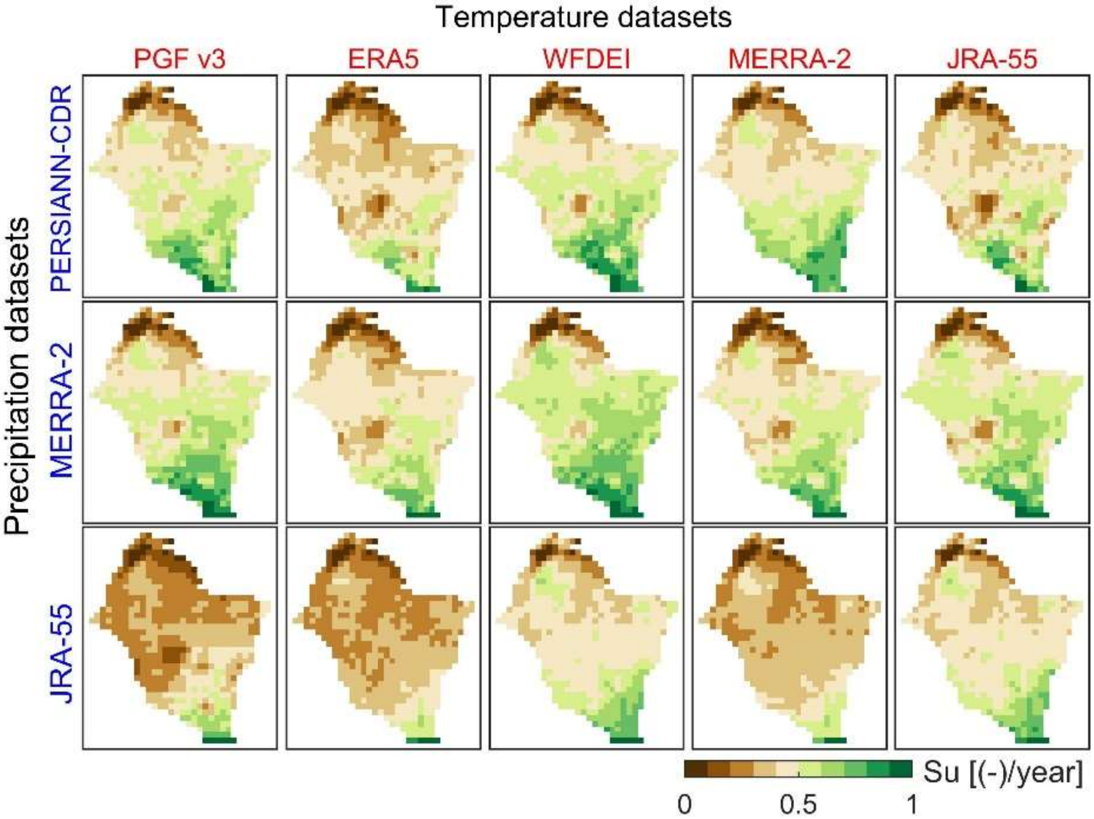
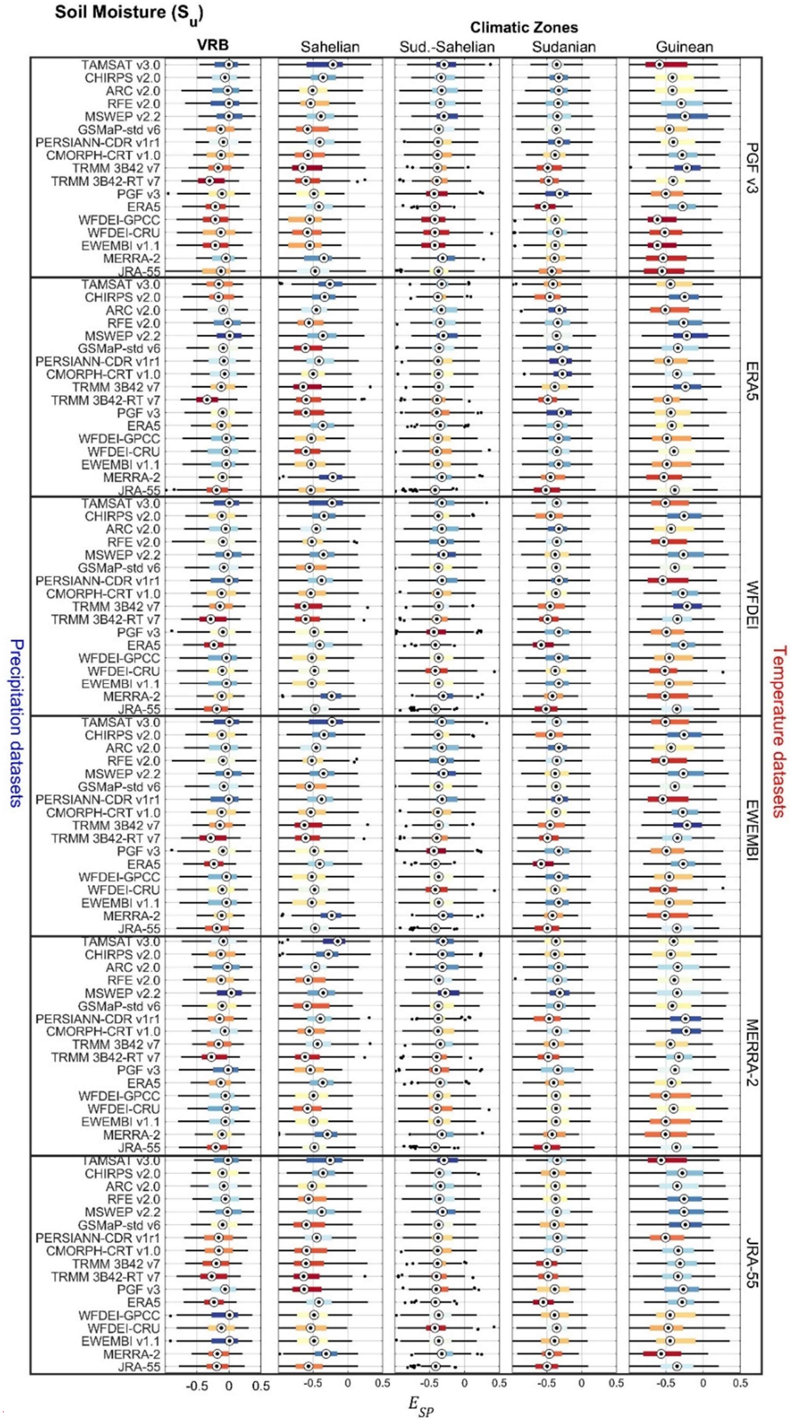


Figure 6. Maps of long-term (2003-2012) average of annual soil moisture (S_u) obtained with different forcing of rainfall (y-axis, blue font) and temperature (x-axis, red font) datasets. The values are normalized between 0 and 1 to emphasize spatial patterns and to use a unique color scale.



Soil Moisture (S_u)
Climatic Zones
VRB
Sahelian
Sud.-Sahelian
Sudanian
Guinean
TAMSAT v3.0
CHIRPS v2.0
ARC v2.0
RFE v2.0
MSWEP v2.2
GSMaP-std v6
PERSIANN-CDR v1r1
CMORPH-CRT v1.0
TRMM 3B42 v7
TRMM 3B42-RT v7
PGF v3
ERA5
WFDEI-GPCC
WFDEI-CRU
EWEMBI v1.1
MERRA-2
JRA-55
PGF v3
ERA5
WFDEI
EWEMBI
MERRA-2
JRA-55
Precipitation datasets
Temperature datasets
-0.5
0
0.5
E_{SP}

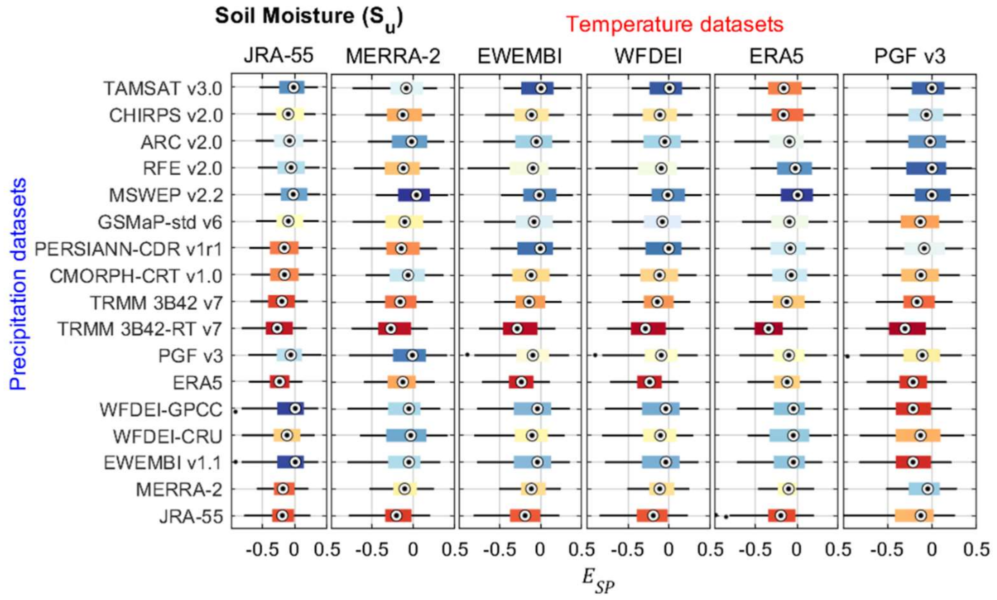


Figure 7. Spatial pattern efficiency (E_{SP}) of soil moisture (S_u) over the entire simulation period (2003-2012) for the Volta River basin (VRB) ~~and its climatic zones,~~ using different combinations of precipitation and temperature datasets used as input for hydrological modelling. Each boxplot has 120 values corresponding to the number of months. The boxplots are colored from the best (blue) to the worst performance (red) based on the median value.

3.4 Model performance for actual evaporation

The model performance for the temporal dynamics of monthly actual evaporation (E_a) compared to the GLEAM product is shown in Figure 8~~Figure 8~~ (see the SI for monthly time series, Figures ~~S57~~S72-S76~~61~~). The average r of E_a for the entire VRB over all datasets is 0.93. Similarly to S_u, the r of E_a is higher is the driest climatic zones: Sahelian (r = 0.94), Sudano-Sahelian (r = 0.94), Sudanian (r = 0.89) and Guinean (r = 0.81). However, the predictive skill of the model for the temporal dynamics of E_a is higher than its predictive skill for E_a in the wetter climatic zones. Appendix A3~~A3~~ shows the ranking of all the meteorological datasets for the model performance for E_a. The rainfall datasets show different performances across climatic zones, with the following best datasets: PERSIANN-CDR in the Sahelian zone, EWEMBI and WFDEI-GPCC in the Soudano-Sahelian zone, ARC in the Sudanian and Guinean zones. The choice of the rainfall dataset leads to an average V_2 of 4% for the temporal dynamics of E_a, while the average V_2 is 1.5% for the choice of the temperature dataset, which aligns with the findings of Jung et al. (2019).

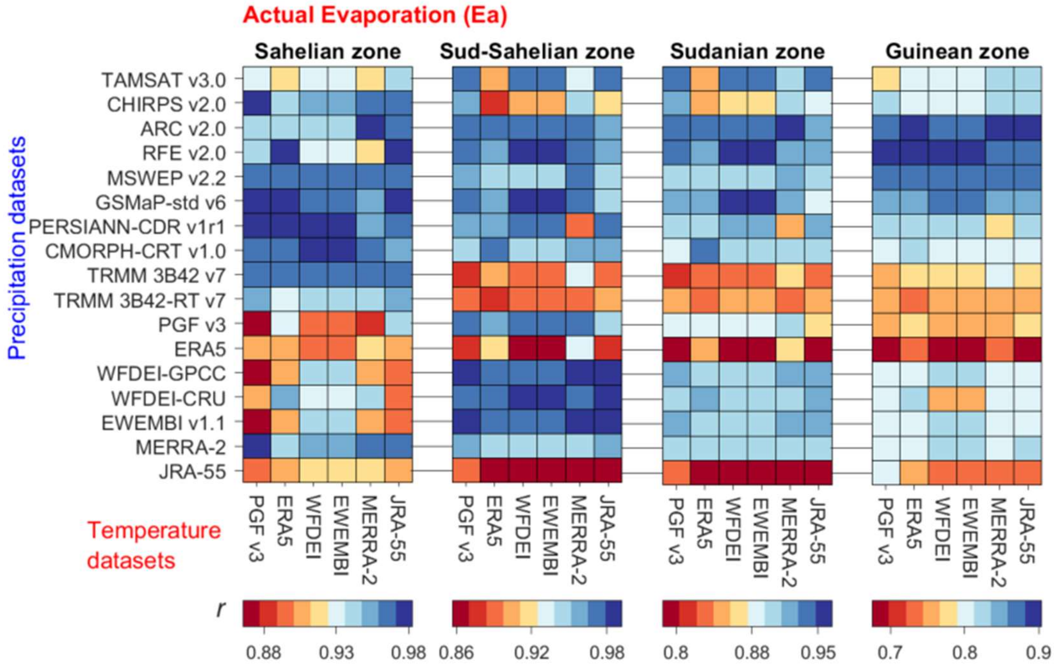


Figure 8. Pearson correlation coefficient (*r*) of modelled actual evaporation (E_a) compared to GLEAM data over the simulation period (2003-2012) considering 102 combinations of rainfall (y-axis) and temperature datasets (subplots on x-axis) used as forcing for the hydrological model.

As for S_u, the choice of input datasets has a considerable impact on the reproduction of the spatial patterns of E_a (Figure 9~~Figure 9~~). Similar maps for all the meteorological datasets are provided in the SI (Figures S77~~62~~-S78~~63~~). It can be observed that different rainfall-temperature combinations used to force the model result in large discrepancies in the spatial pattern of E_a, especially in the southern region. The south-north gradient of increasing aridity with west-east differences is represented differently among the rainfall-temperature dataset combinations (see e.g., the difference between the first two columns of the first row in Figure 9~~Figure 9~~)

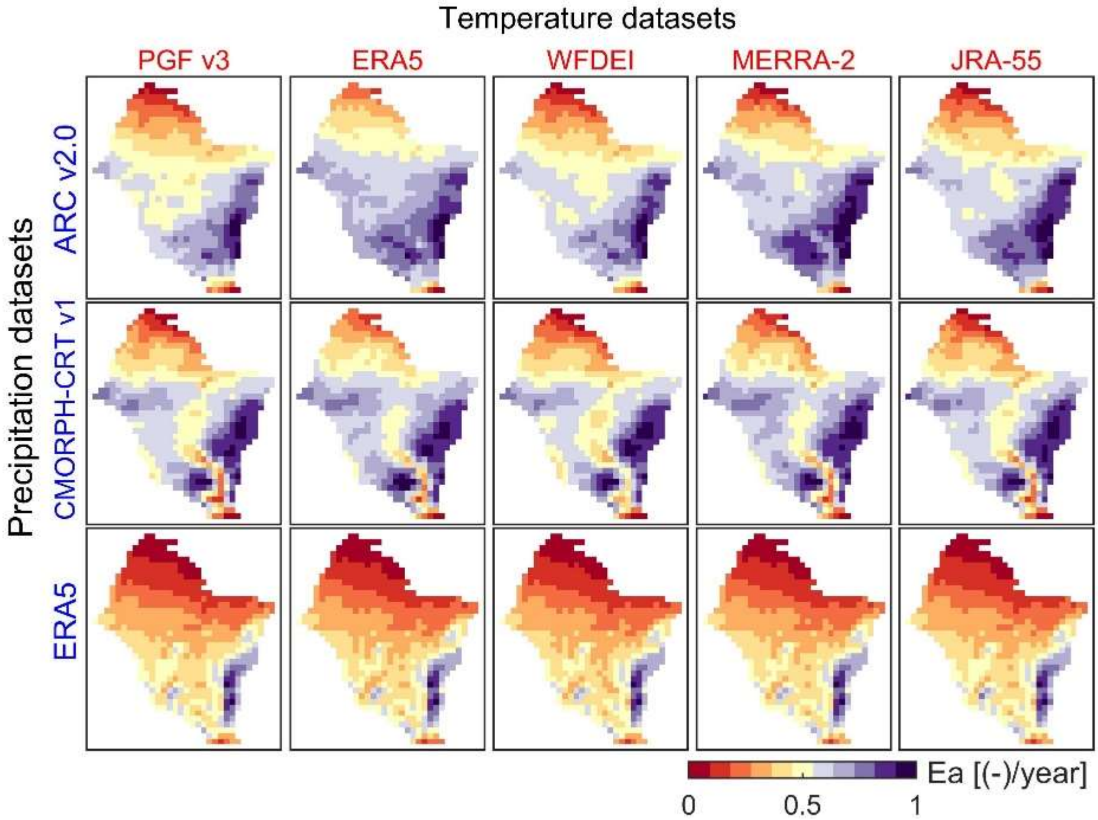
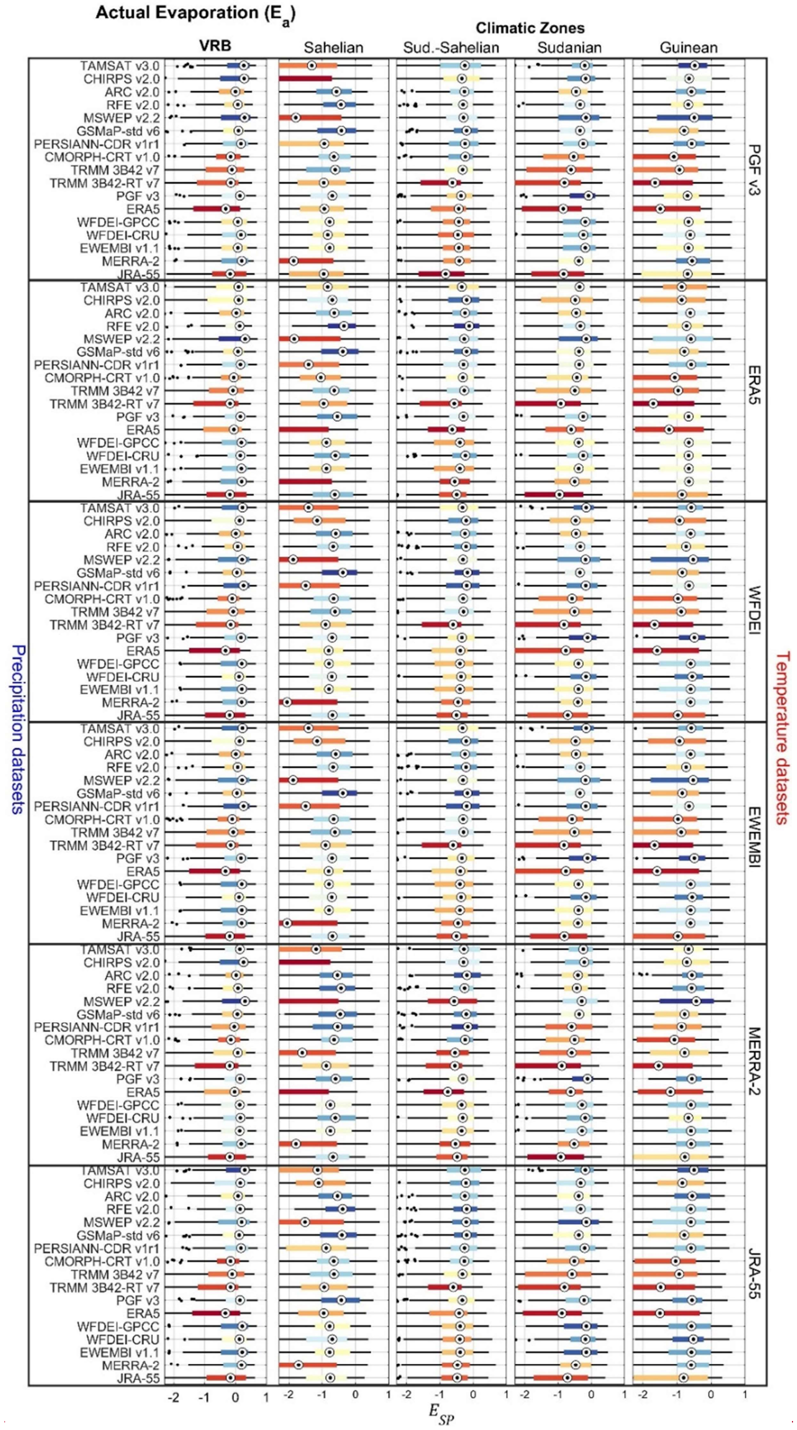


Figure 9. Maps of long-term (2003-2012) average of annual actual evaporation (E_a) obtained with different forcing of rainfall (y-axis, blue font) and temperature (x-axis, red font) datasets. The values are normalized between 0 and 1 to emphasize spatial patterns and to use a unique color scale.

The E_{SP} of E_a for the 102 rainfall-temperature dataset combinations in the VRB ~~and the climatic zones~~ is given in Figure 10~~Figure 10~~. The average E_{SP} of E_a in the VRB over all datasets is 0.07, which is higher than for S_u (E_{SP} = -0.11). The choice of the rainfall dataset for the VRB affects the E_{SP} of E_a on average by 93%, while the choice of the temperature dataset involves a variation 33%. However, lower impacts of data choices are observed in the climatic zones. The choice of rainfall dataset is more critical for the E_{SP} of E_a in the driest and wettest climatic zones, i.e. Sahelian (E_{SP} = -0.99, V_2 = 49%) and Guinean (E_{SP} = -0.79, V_2 = 37%) zones, than the intermediate zones, i.e. Sudano-Sahelian (E_{SP} = -0.35, V_2 = 36%) and Sudanian (E_{SP} = -0.42, V_2 = 49%) zones. A smaller impact on the E_{SP} of E_a is observed for the choice of the temperature dataset: Sahelian (V_2 = 21%), Guinean (V_2 = 10%), Sudano-Sahelian (V_2 = 17%) and Sudanian (V_2 = 21%) zones. Detailed results on the model performance for E_a and the ranking of the datasets for the calibration and evaluation periods are provided in the SI (Tables ~~S16~~S22-S23~~17~~, Figures ~~S46~~S61-S71~~56~~).



Actual Evaporation (E_a)
Climatic Zones
VRB
Sahelian
Sud.-Sahelian
Sudanian
Guinean
TAMSAT v3.0
CHIRPS v2.0
ARC v2.0
RFE v2.0
MSWEP v2.2
GSMaP-std v6
PERSIANN-CDR v1r1
CMORPH-CRT v1.0
TRMM 3B42 v7
TRMM 3B42-RT v7
PGF v3
ERA5
WFDEI-GPCC
WFDEI-CRU
EWEMBI v1.1
MERRA-2
JRA-55
PGF v3
ERA5
WFDEI
EWEMBI
MERRA-2
JRA-55
Precipitation datasets
Temperature datasets
-2 -1 0 1
E_{SP}

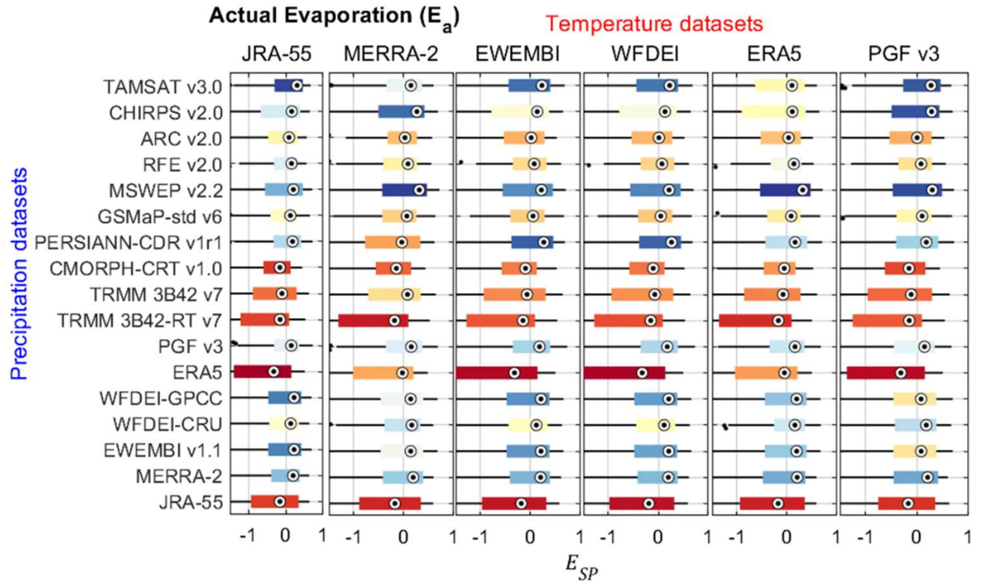


Figure 10. Spatial pattern efficiency (E_{SP}) of actual evaporation (E_a) over the entire simulation period (2003-2012) for the Volta River basin (VRB)~~ and its climatic zones,~~ using different combinations of precipitation and temperature datasets used as input for hydrological modelling. Each boxplot has 120 values corresponding to the number of months. The boxplots are colored from the best (blue) to the worst performance (red) based on the median value.

4 Discussions

This study builds upon and expands existing research studies on the evaluation of meteorological datasets in several ways:

(i) The evaluation of the spatial patterns of multiple hydrological processes (i.e. streamflow, actual evaporation, soil moisture, and terrestrial water storage) in addition to the more classically evaluated temporal dynamic.

(ii) The evaluation of a high number of both satellite-based and reanalysis rainfall datasets considered in combination with different temperature datasets.

(iii) The assessment of the model performance across four considerably different climatic zones from semi-arid to sub-humid.

The overall outcome of this analysis is the ranking of all the meteorological datasets based on their ability to simulate various hydrological processes across different climatic zones in the VRB (Appendix A3~~A3~~). It is worth noting that the overall ranking shows which product is best or worst at simulating a given hydrological flux or state variable. However, the ranking does not systematically tell whether a dataset is good or bad. Only the skill scores can be used to draw a judgement on the adequacy of a given dataset to produce plausible model outputs.

The results show that there is no single rainfall dataset outperforming the others in reproducing all hydrological processes across different climatic zones. These findings align with previous studies in the sense that there is no rainfall dataset that is the best everywhere (Beck et al., 2017b;Sylla et al., 2013). For datasets providing both rainfall and temperature data, the combination of the two variables as model input is not necessarily the best option for obtaining the highest performance in modelling a given hydrological state or flux variable. The best rainfall-temperature combinations for the spatiotemporal representation of each hydrological flux and state variable are provided in the SI (Figure ~~S7~~S15).

The results are primarily valid for the study region in West Africa, while a wider generalization of the findings should be done with caution and after repeating similar evaluation studies at other places~~The results can be considered valid for West Africa and regions with similar hydroclimatic and physical features. A wider generalization of the findings should be done with caution and after repeating similar evaluation studies in other places~~. Nevertheless, the key message is that: ~~“~~*there is no rainfall dataset of all hydrological processes*~~”~~ and ~~“~~*the best rainfall dataset for temporal dynamics might not be the best for spatial patterns*~~”~~. Therefore, different rainfall datasets should be evaluated before choosing the most suitable one for hydrological modelling in large catchments.

Moreover, when comparing the results of this study to the findings of Satgé et al. (2020) based on a point-to-pixel evaluation of gridded rainfall datasets in West Africa, it is noticeable that the ground evaluation might lead to different results as compared to the hydrological evaluation adopted in the current study. The skill of a rainfall product in well reproducing ground measurements under a point-to-pixel evaluation does not necessarily correlate with its performance for hydrological modelling, particularly in large and complex hydroclimatic environments such as the VRB.

Despite the efforts to produce a comprehensive evaluation of the meteorological datasets, the results obtained might be subject to uncertainties related to the potential model structural deficiencies as well as errors in the observational datasets used for the model evaluation (McMillan et al., 2010;Renard et al., 2010;Gupta and Govindaraju, 2019). The distribution of the final model parameters (Figures S79~~65~~-S80~~66~~) highlights the possibility of obtaining equally good model performances for different parameter sets (i.e. equifinality), which can be a justification for model recalibration. Moreover, it can be noticed that most of the model parameters are sensitive to the change in meteorological input datasets (Figure S79). A detailed analysis of parameter variability as a function of input data is beyond the scope of the current study, but could build the basis of future research, namely to identify data errors by analyzing parameter patterns (e.g. rooting depth), and resolve potential structural deficiencies of the mHM model. However, the mHM is chosen because of its adequacy for the experiment of this study (for model selection, see Addor and Melsen, 2019). The structure of mHM allows the representation of seamless spatial patterns of hydrological processes through the MPR scheme (Samaniego et al., 2017). In addition, mHM facilitates parameter regionalization and is therefore convenient for large-scale modelling, and it harnesses the full potential of the forcing datasets as it is a fully distributed model that has performed well in previous studies including those in the VRB (e.g. Poméon et al., 2018;Dembélé et al., 2020b). Regarding the model evaluation, the comparison between the observed and modelled hydrological processes is done only on their temporal dynamics and spatial patterns using bias-insensitive metrics, except for streamflow, which limits the potential impact of satellite data uncertainty.

The model is calibrated only on *Q* data despite the known limitations of the *Q*-only calibration (Demirel et al., 2018). However, calibrating the model on additional variables would result in additional model performance improvement that would not be separable from the contribution of the input datasets to the model performance. Therefore, regarding the goal of this study, the

Q-only calibration ~~that~~ was the best option to obtain the impact of various meteorological forcing datasets on the plausibility of hydrological processes. As no rainfall dataset ranks first in simulating all the hydrological processes, this study confirms that model calibration on multiple variables is a way forward in improving the overall representation of the hydrological system and increasing the predictive skill of hydrological models (Dembélé et al., 2020b;Dembélé et al., 2020a). The domain-wide calibration strategy adopted in this study generates a unique parameter set for the simulation of multiple hydrological processes across several catchments with different hydroclimatic features, which has the consequence of having local differences in model performance. However, domain-wide calibration has proved to perform similarly to domain-split calibration in previous studies (Mizukami et al., 2017), and it was ideal for this study because of the interest in simulating seamless spatial patterns, which might have not been possible with separately simulated portions of the basin. Moreover, the main goal of this study is to assess the adequacy of the meteorological datasets for large-scale hydrological modelling, knowing that these datasets usually have a coarse spatial resolution with pixels often averaged over regions with strong sub-grid variability.

Finally, the importance of regional evaluation is emphasized by this study because some region-tailored datasets (e.g. TAMSAT and ARC) which are not included in global scale studies (e.g. Beck et al., 2017b;Mazzoleni et al., 2019;Essou et al., 2016) outperform global datasets. The decision to use a given dataset is not only motivated by the availability or the accuracy of the data, but also by data accessibility (i.e. storage platforms, openness, format, pre-processing requirement, etc.). The findings of this study provide further awareness for the data users and improvement avenues for data producers in their quest of the most accurate products (e.g. Massari et al., 2020;Contractor et al., 2020;Berg et al., 2018;Brocca et al., 2014;Cucchi et al., 2020;Beck et al., 2017a).

5 Conclusion

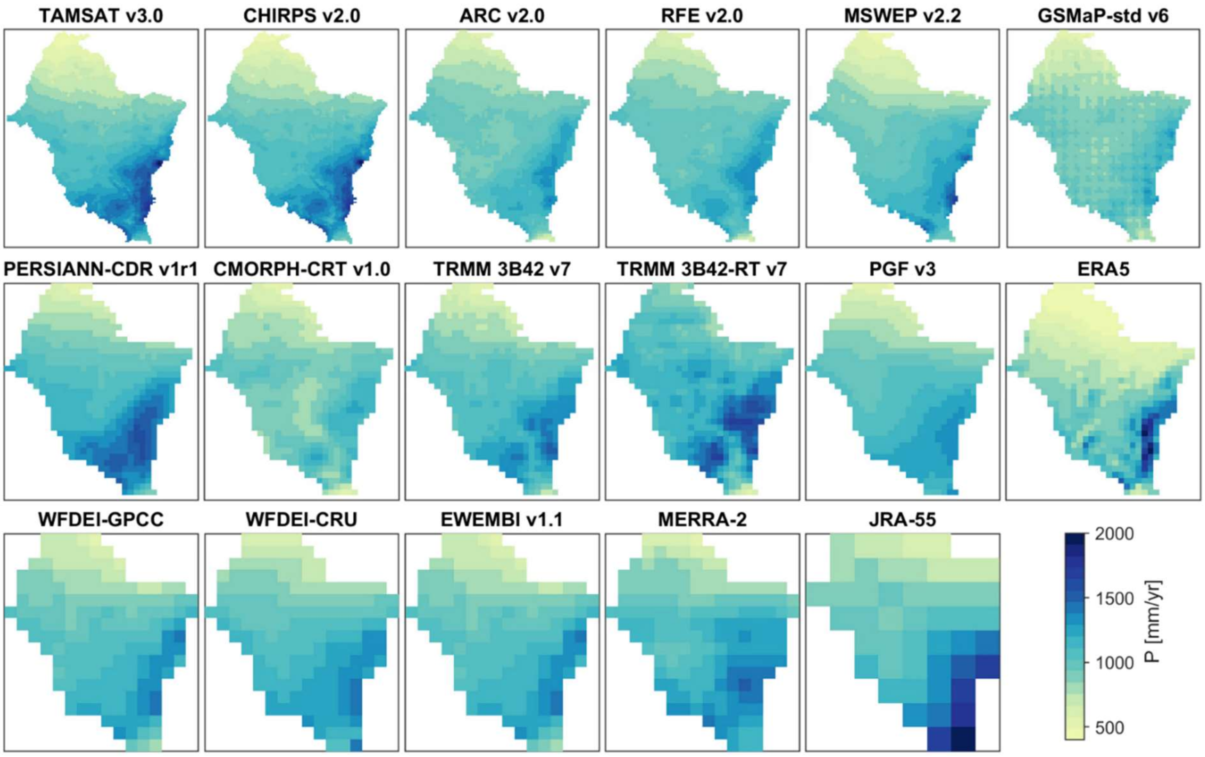
This modelling study evaluates the ability of multiple combinations of rainfall-temperature datasets to reproduce plausible hydrological processes and patterns. The experiment is done in the Volta River basin with the fully distributed mesoscale Hydrologic Model (mHM) over a 10-year period (2003-2012), using 17 rainfall and 6 temperature datasets from satellite and reanalysis sources. The spatial and temporal representation of streamflow, terrestrial water storage, soil moisture and actual evaporation are evaluated using in-situ and satellite remote sensing observational datasets. The key findings are:

- No rainfall dataset consistently outperforms all the others in reproducing the highest model performance for all hydrological processes, and the best dataset for the temporal dynamics is not necessarily the best for the spatial patterns.
- Rainfall datasets have a higher impact on the spatiotemporal representation of hydrological processes than temperature datasets, but the later have a higher influence on the spatial patterns of soil moisture.
- The large-scale performance for the meteorological datasets is not always valid for sub-regions in the same basin.

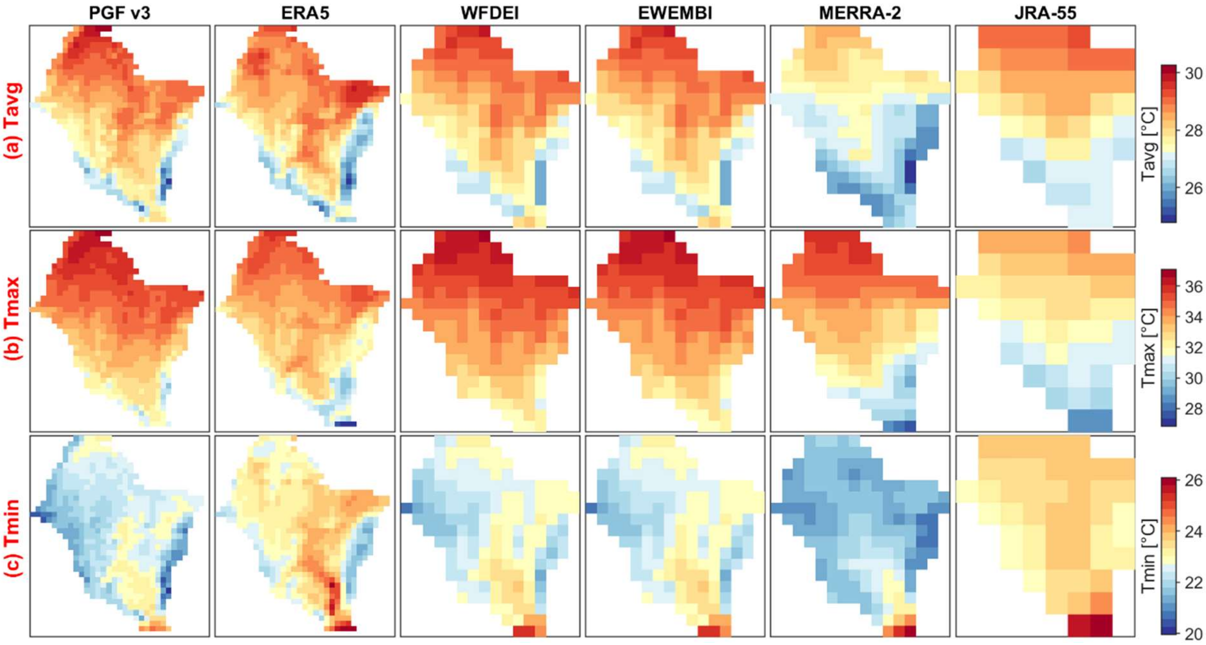
The findings of this study give a critical insight of the performance for several meteorological datasets in the challenging hydroclimatic environment of West Africa. They are expected to foster further research initiatives in improving the gridded meteorological datasets and further draw users' attention on the contrasting performances of these datasets in modelling hydrological fluxes and state variables. Efforts should be devoted in reporting on the impact of data uncertainties on process representation in hydrological modelling, especially when model outputs are used for decision-making.

Future studies can test the transferability of the model's global parameters across different input datasets, i.e. how reliable a parameter set obtained with a given input dataset is for running the same model with a different input dataset. The answer to this research question will shed light on the necessity of model recalibration when using different meteorological forcing. Furthermore, the predictive skill of the model can be improved with a parameter sensitivity analysis to determine parameters that affect the spatiotemporal representation of each hydrological flux and state variable.

6 Appendix A: Figures



A1. Mean annual rainfall totals over the period 2003-2012 for 17 rainfall datasets the Volta River basin



A2. Mean annual air temperature (average (a), maximum (b) and minimum (c)) over the period 2003-2012 for 6 temperature datasets in the Volta River basin

		Skill scores For temporal dynamics: E_{KG} of *Q*, *r* of *St*, *Su* and *Ea* | For spatial patterns: E_{SP} of *Su* and *Ea*																									
		VRB						Sahelian zone					Sudano-Sahelian zone					Sudanian zone					Guinean zone				
		Temporal dynamics				Spatial patterns		Temporal dynamics			Spatial patterns		Temporal dynamics			Spatial patterns		Temporal dynamics			Spatial patterns		Temporal dynamics			Spatial patterns	
		Q	*St*	*Su*	*Ea*	*Su*	*Ea*	*St*	*Su*	*Ea*	*Su*	*Ea*	*St*	*Su*	*Ea*	*Su*	*Ea*	*St*	*Su*	*Ea*	*Su*	*Ea*	*St*	*Su*	*Ea*	*Su*	*Ea*
Rainfall datasets	TAMSAT v3.0	0.73	0.86	0.94	0.94	-0.04	0.21	0.69	0.93	0.93	-0.22	-1.21	0.78	0.94	0.95	-0.31	-0.29	0.86	0.93	0.92	-0.37	-0.22	0.74	0.90	0.81	-0.50	-0.61
	CHIRPS v2.0	0.70	0.86	0.93	0.92	-0.11	0.18	0.67	0.92	0.96	-0.34	-1.60	0.74	0.92	0.92	-0.36	-0.24	0.82	0.92	0.88	-0.41	-0.36	0.71	0.89	0.82	-0.32	-0.82
	ARC v2.0	0.68	0.91	0.93	0.97	-0.05	0.03	0.73	0.97	0.95	-0.48	-0.58	0.80	0.95	0.97	-0.33	-0.24	0.92	0.93	0.94	-0.34	-0.44	0.67	0.84	0.89	-0.42	-0.59
	RFE v2.0	0.65	0.88	0.92	0.96	-0.07	0.10	0.71	0.97	0.95	-0.55	-0.49	0.78	0.94	0.98	-0.34	-0.22	0.89	0.91	0.94	-0.35	-0.35	0.65	0.83	0.89	-0.38	-0.67
	MSWEP v2.2	0.66	0.72	0.92	0.94	0.00	0.26	0.58	0.94	0.96	-0.36	-1.90	0.68	0.94	0.95	-0.30	-0.32	0.73	0.91	0.90	-0.36	-0.19	0.49	0.82	0.86	-0.27	-0.53
	GSMaP-std v6	0.71	0.84	0.94	0.95	-0.10	0.08	0.72	0.97	0.97	-0.58	-0.40	0.76	0.96	0.97	-0.37	-0.19	0.84	0.94	0.92	-0.36	-0.36	0.66	0.88	0.86	-0.37	-0.81
	PERSIANN-CDR v1r1	0.68	0.80	0.96	0.94	-0.08	0.17	0.59	0.95	0.97	-0.41	-1.13	0.72	0.96	0.95	-0.36	-0.22	0.81	0.95	0.90	-0.35	-0.29	0.64	0.92	0.82	-0.45	-0.65
	CMORPH-CRT v1.0	0.69	0.87	0.94	0.94	-0.11	-0.12	0.69	0.95	0.97	-0.55	-0.72	0.76	0.94	0.96	-0.38	-0.28	0.86	0.93	0.90	-0.35	-0.53	0.78	0.92	0.81	-0.29	-1.04
	TRMM 3B42 v7	0.54	0.77	0.91	0.88	-0.16	-0.06	0.67	0.93	0.97	-0.60	-0.79	0.72	0.92	0.90	-0.37	-0.34	0.78	0.89	0.83	-0.45	-0.55	0.52	0.84	0.78	-0.28	-0.89
	TRMM 3B42-RT v7	0.54	0.86	0.91	0.89	-0.30	-0.16	0.68	0.93	0.94	-0.61	-0.92	0.74	0.90	0.89	-0.40	-0.59	0.83	0.89	0.84	-0.49	-0.85	0.72	0.85	0.75	-0.38	-1.61
	PGF v3	0.66	0.82	0.96	0.93	-0.08	0.16	0.69	0.95	0.90	-0.54	-0.61	0.73	0.97	0.97	-0.42	-0.33	0.82	0.95	0.89	-0.33	-0.15	0.72	0.90	0.76	-0.43	-0.58
	ERA5	0.63	0.82	0.91	0.87	-0.20	-0.23	0.57	0.87	0.90	-0.40	-1.42	0.70	0.89	0.89	-0.40	-0.50	0.81	0.87	0.81	-0.50	-0.75	0.69	0.87	0.70	-0.33	-1.43
	WFDEI-GPCC	0.64	0.75	0.96	0.95	-0.07	0.17	0.71	0.97	0.91	-0.52	-0.79	0.68	0.98	0.98	-0.38	-0.39	0.76	0.96	0.91	-0.36	-0.30	0.55	0.88	0.81	-0.50	-0.62
	WFDEI-CRU	0.67	0.83	0.96	0.94	-0.09	0.14	0.72	0.97	0.93	-0.54	-0.69	0.73	0.97	0.98	-0.41	-0.36	0.83	0.95	0.91	-0.37	-0.20	0.70	0.90	0.79	-0.47	-0.59
	EWEMBI v1.1	0.64	0.75	0.96	0.95	-0.07	0.17	0.71	0.97	0.91	-0.52	-0.79	0.68	0.98	0.98	-0.38	-0.39	0.76	0.96	0.91	-0.36	-0.30	0.55	0.88	0.81	-0.50	-0.62
	MERRA-2	0.68	0.80	0.97	0.93	-0.11	0.20	0.56	0.93	0.96	-0.27	-2.00	0.71	0.97	0.95	-0.32	-0.48	0.81	0.97	0.90	-0.43	-0.45	0.61	0.93	0.81	-0.53	-0.60
	JRA-55	0.45	0.38	0.83	0.84	-0.18	-0.18	0.66	0.90	0.91	-0.50	-0.73	0.56	0.89	0.87	-0.41	-0.55	0.38	0.84	0.80	-0.49	-0.83	-0.19	0.50	0.75	-0.40	-0.85
Temperature datasets	JRA-55	0.64	0.81	0.94	0.93	-0.12	0.07	0.68	0.95	0.93	-0.49	-0.84	0.73	0.95	0.95	-0.37	-0.33	0.82	0.93	0.89	-0.41	-0.39	0.61	0.87	0.80	-0.37	-0.77
	MERRA-2	0.66	0.79	0.92	0.92	-0.10	0.07	0.66	0.94	0.94	-0.45	-1.14	0.71	0.93	0.94	-0.36	-0.37	0.78	0.91	0.89	-0.39	-0.44	0.59	0.85	0.81	-0.39	-0.76
	EWEMBI	0.64	0.78	0.93	0.93	-0.10	0.06	0.66	0.94	0.94	-0.45	-0.96	0.71	0.94	0.94	-0.37	-0.33	0.78	0.92	0.89	-0.40	-0.42	0.59	0.86	0.80	-0.40	-0.81
	WFDEI	0.64	0.78	0.93	0.93	-0.10	0.06	0.66	0.94	0.94	-0.45	-0.96	0.71	0.94	0.94	-0.37	-0.33	0.78	0.92	0.89	-0.40	-0.41	0.59	0.86	0.80	-0.40	-0.81
	ERA5	0.64	0.81	0.94	0.93	-0.10	0.08	0.66	0.94	0.94	-0.48	-1.01	0.74	0.95	0.95	-0.37	-0.34	0.81	0.93	0.89	-0.37	-0.45	0.59	0.86	0.81	-0.40	-0.83
	PGF v3	0.64	0.81	0.93	0.92	-0.12	0.06	0.67	0.94	0.94	-0.48	-1.01	0.73	0.94	0.94	-0.37	-0.38	0.81	0.92	0.89	-0.38	-0.39	0.62	0.85	0.81	-0.43	-0.78

		VRB								Sahelian zone					Sudano-Sahelian zone					Sudanian zone					Guinean zone				
		Temporal dynamics						Spatial patterns		Temporal dynamics			Spatial patterns		Temporal dynamics			Spatial patterns		Temporal dynamics			Spatial patterns		Temporal dynamics			Spatial patterns	
	Variables	*Q*	*Q*	*Q*	*St*	*Su*	*Ea*	*Su*	*Ea*	*St*	*Su*	*Ea*	*Su*	*Ea*	*St*	*Su*	*Ea*	*Su*	*Ea*	*St*	*Su*	*Ea*	*Su*	*Ea*	*St*	*Su*	*Ea*	*Su*	*Ea*
	Performance metrics	E_{NS}	E_{NSlog}	E_{KG}	r	r	r	E_{SP}	E_{SP}	r	r	r	E_{SP}	E_{SP}	r	r	r	E_{SP}	E_{SP}	r	r	r	E_{SP}	E_{SP}	r	r	r	E_{SP}	E_{SP}
Rainfall datasets	TAMSAT v3.0	0.745	0.704	0.731	0.856	0.937	0.945	-0.044	0.210	0.686	0.931	0.929	-0.223	-1.215	0.777	0.936	0.954	-0.313	-0.294	0.857	0.927	0.916	-0.371	-0.216	0.740	0.901	0.808	-0.503	-0.612
	CHIRPS v2.0	0.738	0.698	0.703	0.863	0.933	0.919	-0.114	0.178	0.671	0.923	0.958	-0.337	-1.601	0.740	0.915	0.920	-0.359	-0.244	0.824	0.919	0.880	-0.411	-0.357	0.715	0.891	0.819	-0.318	-0.820
	ARC v2.0	0.682	0.693	0.680	0.910	0.932	0.965	-0.055	0.026	0.732	0.968	0.952	-0.477	-0.580	0.799	0.948	0.969	-0.330	-0.241	0.918	0.926	0.940	-0.337	-0.437	0.667	0.840	0.887	-0.416	-0.586
	RFE v2.0	0.689	0.664	0.646	0.882	0.920	0.964	-0.066	0.098	0.709	0.969	0.948	-0.546	-0.490	0.780	0.935	0.976	-0.343	-0.217	0.890	0.913	0.938	-0.354	-0.350	0.648	0.833	0.886	-0.380	-0.671
	MSWEP v2.2	0.690	0.650	0.655	0.719	0.918	0.937	-0.003	0.258	0.579	0.945	0.965	-0.365	-1.898	0.681	0.937	0.950	-0.299	-0.325	0.728	0.911	0.904	-0.359	-0.186	0.490	0.822	0.863	-0.270	-0.528
	GSMaP-std v6	0.695	0.646	0.713	0.842	0.945	0.949	-0.100	0.076	0.718	0.973	0.971	-0.582	-0.404	0.756	0.956	0.970	-0.372	-0.192	0.845	0.938	0.924	-0.360	-0.359	0.656	0.878	0.856	-0.371	-0.805
	PERSIANN-CDR v1r1	0.735	0.711	0.682	0.804	0.957	0.938	-0.082	0.167	0.594	0.952	0.974	-0.405	-1.125	0.718	0.956	0.954	-0.361	-0.224	0.807	0.950	0.905	-0.347	-0.291	0.637	0.921	0.820	-0.453	-0.651
	CMORPH-CRT v1.0	0.696	0.667	0.690	0.875	0.939	0.937	-0.110	-0.122	0.691	0.953	0.966	-0.549	-0.716	0.764	0.940	0.956	-0.383	-0.275	0.856	0.926	0.905	-0.352	-0.525	0.775	0.915	0.809	-0.293	-1.037
	TRMM 3B42 v7	0.590	0.658	0.539	0.769	0.907	0.880	-0.157	-0.062	0.671	0.934	0.965	-0.599	-0.787	0.717	0.918	0.897	-0.373	-0.338	0.780	0.890	0.833	-0.448	-0.551	0.516	0.842	0.776	-0.276	-0.894
	TRMM 3B42-RT v7	0.574	0.669	0.540	0.860	0.907	0.887	-0.295	-0.163	0.677	0.929	0.943	-0.613	-0.921	0.744	0.905	0.890	-0.399	-0.590	0.834	0.890	0.842	-0.486	-0.846	0.724	0.849	0.752	-0.376	-1.613
	PGF v3	0.695	0.679	0.661	0.823	0.960	0.928	-0.080	0.159	0.688	0.954	0.904	-0.540	-0.611	0.729	0.973	0.968	-0.421	-0.327	0.824	0.951	0.888	-0.334	-0.151	0.715	0.900	0.762	-0.431	-0.577
	ERA5	0.540	0.631	0.627	0.823	0.907	0.867	-0.195	-0.228	0.574	0.868	0.903	-0.396	-1.416	0.704	0.893	0.888	-0.398	-0.503	0.814	0.874	0.813	-0.500	-0.749	0.691	0.870	0.697	-0.327	-1.431
	WFDEI-GPCC	0.677	0.687	0.638	0.748	0.965	0.948	-0.065	0.173	0.707	0.974	0.909	-0.517	-0.795	0.678	0.975	0.980	-0.384	-0.393	0.756	0.955	0.912	-0.356	-0.296	0.546	0.884	0.806	-0.496	-0.615
	WFDEI-CRU	0.646	0.666	0.666	0.829	0.958	0.945	-0.091	0.142	0.720	0.967	0.927	-0.545	-0.693	0.730	0.972	0.977	-0.415	-0.355	0.829	0.946	0.909	-0.367	-0.196	0.697	0.902	0.791	-0.470	-0.595
	EWEMBI v1.1	0.677	0.687	0.638	0.748	0.965	0.948	-0.065	0.173	0.707	0.974	0.909	-0.517	-0.795	0.678	0.975	0.980	-0.384	-0.393	0.756	0.955	0.912	-0.356	-0.296	0.546	0.884	0.806	-0.496	-0.615
	MERRA-2	0.687	0.691	0.684	0.800	0.974	0.932	-0.112	0.198	0.558	0.934	0.959	-0.274	-1.997	0.712	0.973	0.952	-0.318	-0.480	0.807	0.970	0.904	-0.429	-0.446	0.615	0.932	0.809	-0.530	-0.598
	JRA-55	0.460	0.581	0.453	0.377	0.830	0.838	-0.185	-0.178	0.658	0.904	0.909	-0.501	-0.731	0.558	0.894	0.865	-0.413	-0.547	0.385	0.839	0.804	-0.494	-0.826	-0.185	0.505	0.748	-0.395	-0.845
Temperature datasets	JRA-55	0.654	0.670	0.640	0.811	0.943	0.928	-0.115	0.070	0.682	0.950	0.934	-0.494	-0.841	0.733	0.950	0.948	-0.373	-0.332	0.817	0.933	0.892	-0.411	-0.390	0.609	0.869	0.801	-0.371	-0.768
	MERRA-2	0.663	0.672	0.656	0.785	0.924	0.921	-0.101	0.067	0.659	0.943	0.943	-0.452	-1.139	0.706	0.934	0.938	-0.361	-0.375	0.785	0.913	0.886	-0.388	-0.439	0.594	0.848	0.810	-0.393	-0.761
	EWEMBI	0.663	0.666	0.642	0.781	0.931	0.925	-0.102	0.061	0.663	0.944	0.942	-0.454	-0.959	0.713	0.940	0.944	-0.368	-0.335	0.776	0.920	0.891	-0.398	-0.416	0.593	0.855	0.804	-0.403	-0.811
	WFDEI	0.663	0.666	0.642	0.781	0.931	0.925	-0.102	0.061	0.663	0.944	0.942	-0.454	-0.958	0.713	0.940	0.944	-0.368	-0.335	0.776	0.920	0.891	-0.400	-0.410	0.593	0.855	0.804	-0.403	-0.811
	ERA5	0.658	0.671	0.644	0.808	0.940	0.927	-0.105	0.076	0.665	0.944	0.939	-0.478	-1.015	0.736	0.947	0.946	-0.367	-0.342	0.808	0.932	0.892	-0.370	-0.453	0.587	0.861	0.807	-0.396	-0.831
	PGF v3	0.657	0.671	0.640	0.808	0.927	0.925	-0.116	0.057	0.670	0.940	0.944	-0.484	-1.008	0.728	0.936	0.943	-0.374	-0.378	0.807	0.916	0.888	-0.385	-0.390	0.620	0.854	0.807	-0.434	-0.781

A3. Model performance for streamflow (Q), terrestrial water storage (S_t), soil moisture (S_u) and actual evaporation (E_a) using various rainfall-temperature dataset combinations as model inputs. Each score for a given rainfall product represents the average over individual combinations with 6 temperature datasets, while the score is the average over combinations with 17 rainfall datasets for each temperature dataset. The skill scores of the temporal dynamics are obtained with the Kling-Gupta efficiency (E_{KG}), the Nash-Sutcliffe efficiency (E_{NS}) and the Nash-Sutcliffe efficiency of the logarithm (E_{NSlog}) ~~-~~for Q, and the Pearson's correlation coefficient (r) for S_t, S_u and E_a. The spatial pattern efficiency (E_{SP}) is used to assess the spatial representation of S_u and E_a. The skill scores are ranked from the best (blue) to the worst (red). The results are shown for the four climatic zones in the Volta River basin (VRB) over the simulation period (2003-2012).

Supplement. The supplement related to this article is available online at: *to be provided by the journal*

Data availability. The meteorological and modelling datasets used in this study are freely available via the web links provided in Table 1~~Table 1~~ and Table 2~~Table 2~~. More information on satellite-based precipitation datasets can be found at http://ipwg.isac.cnr.it/. The modelling database is available at https://doi.org/10.5281/zenodo.3662308~~repository to be provided upon acceptance~~.

Author contributions. MD performed the analyses and drafted the manuscript. All authors contributed to the writing, review and editing process that lead to the final manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank the providers of the datasets used in this study (see Table 1~~Table 1~~ and Table 2~~Table 2~~). We are grateful to the developers of mHM at CHS/UFZ (Germany) for their open-source model. We thank the providers of the streamflow data obtained from the Volta Basin Authority (VBA), the Direction Générale des Ressources en Eau (DGRE) of Burkina Faso, the Hydrological Services Department (HSD) of Ghana and the Direction Générale de l'Eau et de l'Assainissement (DGEA) of Togo. We thank the reviewers for their useful comments.

Financial support. Moctar Dembélé was supported by the Swiss Government Excellence Scholarship (2016.0533 / Burkina Faso / OP), and the Doc.Mobility fellowship (SNF, P1LAP2_178071) of the Swiss National Science Foundation. Bettina Schaefli was supported by a research grant of the Swiss National Science Foundation (SNF, PP00P2_157611).

Review statement. This paper was edited by Albrecht Weerts~~Editor's Name~~ and reviewed by ~~reviewers' name/anonymous~~Nadav Peleg and one anonymous referee.

7 References

Acharya, S. C., Nathan, R., Wang, Q. J., Su, C.-H., and Eizenberg, N.: An evaluation of daily precipitation from a regional atmospheric reanalysis over Australia, Hydrol Earth Syst Sc, 23, 3387-3403, https://doi.org/10.5194/hess-23-3387-2019, 2019.

Addor, N., and Melsen, L.: Legacy, rather than adequacy, drives the selection of hydrological models, Water Resources Research, 55, 378-390, https://doi.org/10.1029/2018WR022958, 2019.

Alazzy, A. A., Lü, H., Chen, R., Ali, A. B., Zhu, Y., and Su, J.: Evaluation of satellite precipitation products and their potential influence on hydrological modeling over the Ganzi River Basin of the Tibetan Plateau, Advances in Meteorology, 2017, https://doi.org/10.1155/2017/3695285, 2017.

Alemohammad, S., McColl, K., Konings, A., Entekhabi, D., and Stoffelen, A.: Characterization of precipitation product errors across the United States using multiplicative triple collocation, Hydrol Earth Syst Sc, https://doi.org/10.5194/hess-19-3489-2015, 2015.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, 326, 1998.

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., and Prat, O. P.: PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies, Bulletin of the American Meteorological Society, 96, 69-83, https://doi.org/10.1175/BAMS-D-13-00068.1, 2015.

Awange, J., Ferreira, V., Forootan, E., Andam-Akorful, S., Agutu, N., and He, X.: Uncertainties in remotely sensed precipitation data over Africa, International Journal of Climatology, 36, 303-323, https://doi.org/10.1002/joc.4346, 2016.

Awange, J., Hu, K., and Khaki, M.: The newly merged satellite remotely sensed, gauge and reanalysis-based Multi-Source Weighted-Ensemble Precipitation: Evaluation over Australia and Africa (1981–2016), Science of The Total Environment, 670, 448-465, https://doi.org/10.1016/j.scitotenv.2019.03.148, 2019.

Bai, P., Liu, X., Zhang, Y., and Liu, C.: Incorporating vegetation dynamics noticeably improved performance of hydrological model under vegetation greening, Sci Total Environ, 643, 610-622, https://doi.org/10.1016/j.scitotenv.2018.06.233, 2018.

Bárdossy, A., and Das, T.: Influence of rainfall observation network on model calibration and application, Hydrol Earth Syst Sc, 12, 77-89, https://doi.org/10.5194/hess-12-77-2008, 2008.

Bárdossy, A., and Pegram, G.: Interpolation of precipitation under topographic influence at different time scales, Water Resources Research, 49, 4545-4565, https://doi.org/10.1002/wrcr.20307, 2013.

Baroni, G., Schalge, B., Rakovec, O., Kumar, R., Schüler, L., Samaniego, L., Simmer, C., and Attinger, S.: A Comprehensive Distributed Hydrological Modeling Intercomparison to Support Process Representation and Data Collection Strategies, Water Resources Research, 55, 990-1010, https://doi.org/10.1029/2018wr023941, 2019.

Barrett, E. C., and Martin, D. W.: Use of satellite data in rainfall monitoring, Academic press, 1981.

Beck, H. E., Van Dijk, A. I., Levizzani, V., Schellekens, J., Gonzalez Miralles, D., Martens, B., and De Roo, A.: MSWEP: 3-hourly 0.25 global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data, Hydrol Earth Syst Sc, 21, 589-615, https://doi.org/10.5194/hess-21-589-2017, 2017a.

Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. J. M., Weedon, G. P., Brocca, L., Pappenberger, F., Huffman, G. J., and Wood, E. F.: Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling, Hydrol Earth Syst Sc, 21, 6201-6217, https://doi.org/10.5194/hess-21-6201-2017, 2017b.

Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., Van Dijk, A. I., Huffman, G. J., Adler, R. F., and Wood, E. F.: Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS, Hydrol Earth Syst Sc, 23, 207-224, https://doi.org/10.5194/hess-23-207-2019, 2019a.

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Van Dijk, A. I., McVicar, T. R., and Adler, R. F.: MSWEP V2 global 3-hourly 0.1° precipitation: methodology and quantitative assessment, Bulletin of the American Meteorological Society, 100, 473-500, https://doi.org/10.1175/BAMS-D-17-0138.1, 2019b.

Behrangi, A., Khakbaz, B., Jaw, T. C., AghaKouchak, A., Hsu, K., and Sorooshian, S.: Hydrologic evaluation of satellite precipitation products over a mid-size basin, Journal of Hydrology, 397, 225-237, https://doi.org/10.1016/j.jhydrol.2010.11.043, 2011.

Berg, P., Donnelly, C., and Gustafsson, D.: Near-real-time adjusted reanalysis forcing data for hydrology, Hydrol Earth Syst Sc, 22, 989-1000, https://doi.org/10.5194/hess-22-989-2018, 2018.

Berthou, S., Rowell, D. P., Kendon, E. J., Roberts, M. J., Stratton, R. A., Crook, J. A., and Wilcox, C.: Improved climatological precipitation characteristics over West Africa at convection-permitting scales, Climate Dynamics, 53, 1-21, https://doi.org/10.1007/s00382-019-04759-4, 2019.

Beven, K.: Preferential flows and travel time distributions: defining adequate hypothesis tests for hydrological process models, Hydrological Processes, 24, 1537-1547, https://doi.org/10.1002/hyp.7718, 2010.

Beven, K.: Facets of uncertainty: epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication, Hydrological Sciences Journal, 61, 1652-1665, https://doi.org/10.1080/02626667.2015.1031761, 2016.

Beven, K.: On hypothesis testing in hydrology: Why falsification of models is still a really good idea, Wiley Interdisciplinary Reviews: Water, 5, e1278, https://doi.org/10.1002/wat2.1278, 2018.

Beven, K.: How to make advances in hydrological modelling, Hydrology Research, https://doi.org/10.2166/nh.2019.134, 2019.

Bhattacharya, T., Khare, D., and Arora, M.: A case study for the assessment of the suitability of gridded reanalysis weather data for hydrological simulation in Beas river basin of North Western Himalaya, Applied Water Science, 9, 110, https://doi.org/10.1007/s13201-019-0993-x, 2019.

Bhuiyan, E., Abul, M., Nikolopoulos, E. I., Anagnostou, E. N., Polcher, J., Albergel, C., Dutra, E., Fink, G., Martínez-de la Torre, A., and Munier, S.: Assessment of precipitation error propagation in multi-model global water resource reanalysis, Hydrol Earth Syst Sc, 23, 1973-1994, https://doi.org/10.5194/hess-23-1973-2019, 2019.

Biasutti, M., Voigt, A., Boos, W. R., Braconnot, P., Hargreaves, J. C., Harrison, S. P., Kang, S. M., Mapes, B. E., Scheff, J., and Schumacher, C.: Global energetics and local physics as drivers of past, present and future monsoons, Nature Geoscience, 11, 392, https://doi.org/10.1038/s41561-018-0137-1, 2018.

Biasutti, M.: Rainfall trends in the African Sahel: Characteristics, processes, and causes, Wiley Interdiscip Rev Clim Change, 10, e591, https://doi.org/10.1002/wcc.591, 2019.

Bichet, A., and Diedhiou, A.: West African Sahel has become wetter during the last 30 years, but dry spells are shorter and more frequent, Climate Research, 75, 155-162, https://doi.org/10.3354/cr01515, 2018.

Birhanu, D., Kim, H., and Jang, C.: Effectiveness of introducing crop coefficient and leaf area index to enhance evapotranspiration simulations in hydrologic models, Hydrological Processes, https://doi.org/10.1002/hyp.13464, 2019.

Bisselink, B., Zambrano-Bigiarini, M., Burek, P., and De Roo, A.: Assessing the role of uncertain precipitation estimates on the robustness of hydrological model parameters under highly variable climate conditions, Journal of Hydrology: Regional Studies, 8, 112-129, https://doi.org/10.1016/j.ejrh.2016.09.003, 2016.

Bitew, M. M., and Gebremichael, M.: Evaluation of satellite rainfall products through hydrologic simulation in a fully distributed hydrologic model, Water Resources Research, 47, https://doi.org/Artn W06526 10.1029/2010wr009917, 2011.

Blöschl, G., and Zehe, E.: On hydrological predictability, Hydrological Processes: An International Journal, 19, 3923-3929, https://doi.org/10.1002/hyp.6075, 2005.

Blöschl, G., Bierkens, M. F. P., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., Kirchner, J. W., McDonnell, J. J., Savenije, H. H. G., Sivapalan, M., Stumpp, C., Toth, E., Volpi, E., Carr, G., Lupton, C., Salinas, J., Széles, B., Viglione, A., Aksoy, H., Allen, S. T., Amin, A., Andréassian, V., Arheimer, B., Aryal, S. K., Baker, V., Bardsley, E., Barendrecht, M. H., Bartosova, A., Batelaan, O., Berghuijs, W. R., Beven, K., Blume, T., Bogaard, T., Borges de Amorim, P., Böttcher, M. E., Boulet, G., Breinl, K., Brilly, M., Brocca, L., Buytaert, W., Castellarin, A., Castelletti, A., Chen, X., Chen, Y., Chen, Y., Chifflard, P., Claps, P., Clark, M. P., Collins, A. L., Croke, B., Dathe, A., David, P. C., de Barros, F. P. J., de Rooij, G., Di Baldassarre, G., Driscoll, J. M., Duethmann, D., Dwivedi, R., Eris, E., Farmer, W. H., Feiccabrino, J., Ferguson, G., Ferrari, E., Ferraris, S., Fersch, B., Finger, D., Foglia, L., Fowler, K., Gartsman, B., Gascoin, S., Gaume, E., Gelfan, A., Geris, J., Gharari, S., Gleeson, T., Glendell, M., Gonzalez Bevacqua, A., González-Dugo, M. P., Grimaldi, S., Gupta, A. B., Guse, B., Han, D., Hannah, D., Harpold, A., Haun, S., Heal, K., Helfricht, K., Herrnegger, M., Hipsey, M., Hlaváčiková, H., Hohmann, C., Holko, L., Hopkinson, C., Hrachowitz, M., Illangasekare, T. H., Inam, A., Innocente, C., Istanbulluoglu, E., Jarihani, B., Kalantari, Z., Kalvans, A., Khanal, S., Khatami, S., Kiesel, J., Kirkby, M., Knoben, W., Kochanek, K., Kohnová, S., Kolechkina, A., Krause, S., Kreamer, D., Kreibich, H., Kunstmann, H., Lange, H., Liberato, M. L. R., Lindquist, E., Link, T., Liu, J., Loucks, D. P., Luce, C., Mahé, G., Makarieva, O., Malard, J., Mashtayeva, S., Maskey, S., Mas-Pla, J., Mavrova-Guirguinova, M., Mazzoleni, M., Mernild, S., Misstear, B. D., Montanari, A., Müller-Thomy, H., Nabizadeh, A., Nardi, F., Neale, C., Nesterova, N., Nurtaev, B., Odongo, V. O., Panda, S., Pande, S., Pang, Z., Papacharalampous, G., Perrin, C., Pfister, L., Pimentel, R., Polo, M. J., Post, D., Prieto Sierra, C., Ramos, M.-H., Renner, M., Reynolds, J. E., Ridolfi, E., Rigon, R., Riva, M., Robertson, D. E., Rosso, R., Roy, T., Sá, J. H. M., Salvadori, G., Sandells, M., Schaefli, B., Schumann, A., Scolobig, A., Seibert, J., Servat, E., Shafiei, M., Sharma, A., Sidibe, M., Sidle, R. C., Skaugen, T., Smith, H., Spiessl, S. M., Stein, L., Steinsland, I., Strasser, U., Su, B., Szolgay, J., Tarboton, D., Tauro, F., Thirel, G., Tian, F., Tong, R., Tussupova, K., Tyralis, H., Uijlenhoet, R., van Beek, R., van der Ent, R. J., van der Ploeg, M., Van Loon, A. F., van Meerveld, I., van Nooijen, R., van Oel, P. R., Vidal, J.-P., von Freyberg, J., Vorogushyn, S., Wachniew, P., Wade, A. J., Ward, P., Westerberg, I. K., White, C., Wood, E. F., Woods, R., Xu, Z., Yilmaz, K. K., and Zhang, Y.: Twenty-three unsolved problems in hydrology (UPH) – a community perspective, Hydrological Sciences Journal, 64, 1141-1158, https://doi.org/10.1080/02626667.2019.1620507, 2019.

Bontemps, S., Defourny, P., Bogaert, E. V., Arino, O., Kalogirou, V., and Perez, J. R.: GLOBCOVER 2009-Products description and validation report, 53, 2011.

Bosilovich, M. G., Chen, J., Robertson, F. R., and Adler, R. F.: Evaluation of global precipitation in reanalyses, Journal of applied meteorology and climatology, 47, 2279-2299, https://doi.org/10.1175/2008jamc1921.1, 2008.

Brocca, L., Moramarco, T., Melone, F., and Wagner, W.: A new method for rainfall estimation through soil moisture observations, Geophys Res Lett, 40, 853-858, https://doi.org/10.1002/grl.50173, 2013.

Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W., and Levizzani, V.: Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data, Journal of Geophysical Research: Atmospheres, 119, 5128-5141, https://doi.org/10.1002/2014JD021489, 2014.

Brocca, L., Filippucci, P., Hahn, S., Ciabatta, L., Massari, C., Camici, S., Schüller, L., Bojkov, B., and Wagner, W.: SM2RAIN–ASCAT (2007–2018): global daily satellite rainfall data from ASCAT soil moisture observations, Earth System Science Data, 11, 1583-1601, https://doi.org/10.5194/essd-11-1583-2019, 2019.

Camici, S., Ciabatta, L., Massari, C., and Brocca, L.: How reliable are satellite precipitation estimates for driving hydrological models: A verification study over the Mediterranean area, Journal of hydrology, 563, 950-961, https://doi.org/10.1016/j.jhydrol.2018.06.067, 2018.

Caroletti, G. N., Coscarelli, R., and Caloiero, T.: Validation of Satellite, Reanalysis and RCM Data of Monthly Rainfall in Calabria (Southern Italy), Remote Sensing, 11, 1625, https://doi.org/10.3390/rs11131625 2019.

Casse, C., Gosset, M., Peugeot, C., Pedinotti, V., Boone, A., Tanimoun, B., and Decharme, B.: Potential of satellite rainfall products to predict Niger River flood events in Niamey, Atmospheric Research, 163, 162-176, https://doi.org/10.1016/j.atmosres.2015.01.010, 2015.

Chen, L., and Wang, L.: Recent advance in earth observation big data for hydrology, Big Earth Data, 2, 86-107, https://doi.org/10.1080/20964471.2018.1435072, 2018.
Ciabatta, L., Brocca, L., Massari, C., Moramarco, T., Gabellani, S., Puca, S., and Wagner, W.: Rainfall-runoff modelling by using SM2RAIN-derived and state-of-the-art satellite rainfall products over Italy, International journal of applied earth observation and geoinformation, 48, 163-173, https://doi.org/10.1016/j.jag.2015.10.004, 2016.
Ciabatta, L., Massari, C., Brocca, L., Gruber, A., Reimer, C., Hahn, S., Paulik, C., Dorigo, W., Kidd, R., and Wagner, W.: SM2RAIN-CCI: A new global long-term rainfall data set derived from ESA CCI soil moisture, Earth System Science Data, 10, 267, https://doi.org/10.5194/essd-10-267-2018, 2018.
Clark, M. P., Rupp, D. E., Woods, R. A., Tromp-van Meerveld, H., Peters, N., and Freer, J.: Consistency between hydrological models and field observations: linking processes at the hillslope scale to hydrological responses at the watershed scale, Hydrological Processes: An International Journal, 23, 311-319, https://doi.org/10.1002/hyp.7154, 2009.
Clark, M. P., Bierkens, M. F. P., Samaniego, L., Woods, R. A., Uijlenhoet, R., Bennett, K. E., Pauwels, V. R. N., Cai, X., Wood, A. W., and Peters-Lidard, C. D.: The evolution of process-based hydrologic models: historical challenges and the collective quest for physical realism, Hydrol Earth Syst Sc, 21, 3427-3440, https://doi.org/10.5194/hess-21-3427-2017, 2017.
Contractor, S., Donat, M., Alexandre, L. V., Ziese, M., Meyer-Christoffer, A., Schneider, U., Rustemeier, E., Becker, A., Durre, I., and Vose, R. S.: Rainfall Estimates on a Gridded Network (REGEN)–a global land-based gridded dataset of daily precipitation from 1950 to 2016, Hydrology and Earth System Sciences (HESS), 24, 919-943, https://doi.org/10.5194/hess-24-919-2020, 2020.
Cook, K. H., and Vizy, E. K.: Contemporary climate change of the African monsoon systems, Current Climate Change Reports, 5, 145-159, https://doi.org/10.1007/s40641-019-00130-1s, 2019.
Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Schmied, H. M., Hersbach, H., and Buontempo, C.: WFDE5: bias adjusted ERA5 reanalysis data for impact studies, Earth System Science Data Discussions, 1-32, https://doi.org/10.5194/essd-2020-28, 2020.
Cui, X., Guo, X., Wang, Y., Wang, X., Zhu, W., Shi, J., Lin, C., and Gao, X.: Application of remote sensing to water environmental processes under a changing climate, Journal of Hydrology, https://doi.org/10.1016/j.jhydrol.2019.04.078, 2019.
Cunge, J. A.: On The Subject Of A Flood Propagation Computation Method (Musklngum Method), Journal of Hydraulic Research, 7, 205-230, https://doi.org/10.1080/00221686909500264, 1969.
Danielson, J. J., and Gesch, D. B.: Global multi-resolution terrain elevation data 2010 (GMTED2010), US Geological Survey2331-1258, 34, https://doi.org/10.3133/ofr20111073, 2011.
De Condappa, D., and Lemoalle, J.: Water atlas of the Volta Basin-Atlas de l'eau dans le basin de la Volta, 93, 2009.
Dembélé, M., and Zwart, S. J.: Evaluation and comparison of satellite-based rainfall products in Burkina Faso, West Africa, Int J Remote Sens, 37, 3995-4014, https://doi.org/10.1080/01431161.2016.1207258, 2016.
Dembélé, M., Oriani, F., Tumbulto, J., Mariéthoz, G., and Schaefli, B.: Gap-filling of daily streamflow time series using Direct Sampling in various hydroclimatic settings, Journal of Hydrology, 569, 573-586, https://doi.org/10.1016/j.jhydrol.2018.11.076, 2019.
Dembélé, M., Ceperley, N., Zwart, S. J., Mariéthoz, G., and Schaefli, B.: Potential of Satellite and Reanalysis Evaporation Datasets for Hydrological Modelling under Various Model Calibration Strategies, Advances in Water Resources, https://doi.org/10.1016/j.advwatres.2020.103667, 2020a.
Dembélé, M., Hrachowitz, M., Savenije, H. H., Mariéthoz, G., and Schaefli, B.: Improving the predictive skill of a distributed hydrological model by calibration on spatial patterns with multiple satellite datasets, Water Resources Research, e2019WR026085, https://doi.org/10.1029/2019WR026085, 2020b.
Demirel, M. C., Mai, J., Mendiguren, G., Koch, J., Samaniego, L., and Stisen, S.: Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model, Hydrol Earth Syst Sc, 22, 1299-1315, https://doi.org/10.5194/hess-22-1299-2018, 2018.
Dezfuli, A.: Climate of western and central equatorial Africa, in: Oxford Research Encyclopedia of Climate Science, 10.1093/acrefore/9780190228620.013.511 2017.
Di Baldassarre, G., Montanari, A., Lins, H., Koutsoyiannis, D., Brandimarte, L., and Blöschl, G.: Flood fatalities in Africa: from diagnosis to mitigation, Geophys Res Lett, 37, https://doi.org/10.1029/2010GL045467, 2010.

Dinku, T.: Challenges with availability and quality of climate data in Africa, in: Extreme Hydrology and Climate Variability, Elsevier, 71-80, https://doi.org/10.1016/B978-0-12-815998-9.00007-5, 2019.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.: ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions, Remote Sens Environ, 203, 185-215, https://doi.org/10.1016/j.rse.2017.07.001, 2017.

Duan, Z., Tuo, Y., Liu, J., Gao, H., Song, X., Zhang, Z., Yang, L., and Mekonnen, D. F.: Hydrological evaluation of open-access precipitation and air temperature datasets using SWAT in a poorly gauged basin in Ethiopia, Journal of hydrology, 569, 612-626, https://doi.org/10.1016/j.jhydrol.2018.12.026, 2019.

Duethmann, D., Zimmer, J., Gafurov, A., Güntner, A., Kriegel, D., Merz, B., and Vorogushyn, S.: Evaluation of areal precipitation estimates based on downscaled reanalysis and station data by hydrological modelling, Hydrol Earth Syst Sc, 17, 2415-2434, https://doi.org/10.5194/hess-17-2415-2013, 2013.

Essou, G. R., Sabarly, F., Lucas-Picher, P., Brissette, F., and Poulin, A.: Can precipitation and temperature from meteorological reanalyses be used for hydrological modeling?, Journal of Hydrometeorology, 17, 1929-1950, https://doi.org/10.1175/JHM-D-15-0138.1, 2016.

Falck, A. S., Maggioni, V., Tomasella, J., Vila, D. A., and Diniz, F. L.: Propagation of satellite precipitation uncertainties through a distributed hydrologic model: A case study in the Tocantins–Araguaia basin in Brazil, Journal of Hydrology, 527, 943-957, https://doi.org/10.1016/j.jhydrol.2015.05.042, 2015.

Fallah, A., O, S., and Orth, R.: Climate-dependent propagation of precipitation uncertainty into the water cycle, Hydrol. Earth Syst. Sci., 24, 3725-3735, https://doi.org/10.5194/hess-24-3725-2020, 2020.

FAO/GIEWS: Sahel Weather And Crop Situation 1998, Food and Agriculture Organization (FAO) / Global Information and Early Warning System (GIEWS), 8, 1998.

Fatichi, S., Vivoni, E. R., Ogden, F. L., Ivanov, V. Y., Mirus, B., Gochis, D., Downer, C. W., Camporese, M., Davison, J. H., Ebel, B., Jones, N., Kim, J., Mascaro, G., Niswonger, R., Restrepo, P., Rigon, R., Shen, C., Sulis, M., and Tarboton, D.: An overview of current applications, challenges, and future trends in distributed process-based models in hydrology, Journal of Hydrology, 537, 45-60, https://doi.org/10.1016/j.jhydrol.2016.03.026, 2016.

Feddes, R. A., Kowalik, P., Kolinskamalinka, K., and Zaradny, H.: Simulation of Field Water-Uptake by Plants Using a Soil-Water Dependent Root Extraction Function, Journal of Hydrology, 31, 13-26, https://doi.org/10.1016/0022-1694(76)90017-2, 1976.

Fekete, B. M., Vörösmarty, C. J., Roads, J. O., and Willmott, C. J.: Uncertainties in precipitation and their impacts on runoff estimates, Journal of Climate, 17, 294-304, https://doi.org/10.1175/1520-0442(2004)017<0294:Uipati>2.0.Co;2, 2004.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., and Hoell, A.: The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes, Scientific data, 2, 150066, https://doi.org/10.1038/sdata.2015.66, 2015.

Gebremichael, M.: Framework for satellite rainfall product evaluation, Geophys. Monogr. Ser, 191, 265-275, https://doi.org/10.1029/2010gm000974, 2010.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., and Reichle, R.: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), Journal of Climate, 30, 5419-5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Golian, S., Javadian, M., and Behrangi, A.: On the use of satellite, gauge, and reanalysis precipitation products for drought studies, Environmental Research Letters, 14, 075005, https://doi.org/10.1088/1748-9326/ab2203, 2019.

Gosset, M., Viarre, J., Quantin, G., and Alcoba, M.: Evaluation of several rainfall products used for hydrological applications over West Africa using two high-resolution gauge networks, Quarterly Journal of the Royal Meteorological Society, 139, 923-940, https://doi.org/10.1002/qj.2130, 2013.

Gruber, A., Dorigo, W. A., Crow, W., and Wagner, W.: Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals, Ieee T Geosci Remote, 55, 6780-6792, https://doi.org/10.1109/Tgrs.2017.2734070, 2017.

Gründemann, G. J., Werner, M., and Veldkamp, T. I.: The potential of global reanalysis datasets in identifying flood events in Southern Africa, Hydrol Earth Syst Sc, 22, 4667-4683, https://doi.org/10.5194/hess-22-4667-2018, 2018.

Gupta, A., and Govindaraju, R.: Propagation of structural uncertainty in watershed hydrologic models, Journal of Hydrology, 575, 66-81, https://doi.org/10.1016/j.jhydrol.2019.05.026, 2019.

Hargreaves, G. H., and Samani, Z. A.: Reference crop evapotranspiration from temperature, Applied engineering in agriculture, 1, 96-99, https://doi.org/10.13031/2013.26773, 1985.

Harrison, L., Funk, C., and Peterson, P.: Identifying changing precipitation extremes in Sub-Saharan Africa with gauge and satellite products, Environmental Research Letters, 14, 085007, https://doi.org/10.1088/1748-9326/ab2cae, 2019.

Hartmann, J., and Moosdorf, N.: The new global lithological map database GLiM: A representation of rock properties at the Earth surface, Geochem Geophy Geosy, 13, https://doi.org/10.1029/2012gc004370, 2012.

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotic, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G., Ribeiro, E., Wheeler, I., Mantel, S., and Kempen, B.: SoilGrids250m: Global gridded soil information based on machine learning, PLoS One, 12, e0169748, https://doi.org/10.1371/journal.pone.0169748, 2017.

Herman, A., Kumar, V. B., Arkin, P. A., and Kousky, J. V.: Objectively determined 10-day African rainfall estimates created for famine early warning systems, Int J Remote Sens, 18, 2147-2159, https://doi.org/10.1080/014311697217800, 1997.

Hersbach, H., de Rosnay, P., Bell, B., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Alonso-Balmaseda, M., Balsamo, G., Bechtold, P., Berrisford, P., Bidlot, J.-R., de Boisséson, E., Bonavita, M., Browne, P., Buizza, R., Dahlgren, P., Dee, D., Dragani, R., Diamantakis, M., Flemming, J., Forbes, R., Geer, A. J., Haiden, T., Hólm, E., Haimberger, L., Hogan, R., Horányi, A., Janiskova, M., Laloyaux, P., Lopez, P., Muñoz-Sabater, J., Peubey, C., Radu, R., Richardson, D., Thépaut, J.-N., Vitart, F., Yang, X., Zsótér, E., and Zuo, H.: Operational global reanalysis: progress, future directions and synergies with NWP, ERA Report Series 27, ECMWF, Reading, UK, European Centre for Medium Range Weather Forecasts, https://doi.org/10.21957/tkic6g3wm, 2018.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., and Schepers, D.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, https://doi.org/10.1002/qj.3803, 2020.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E., and Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB)—a review, Hydrological Sciences Journal, 58, 1198-1255, https://doi.org/10.1080/02626667.2013.803183, 2013.

Hrachowitz, M., and Clark, M. P.: HESS Opinions: The complementary merits of competing modelling philosophies in hydrology, Hydrol Earth Syst Sc, 21, 3953-3973, https://doi.org/10.5194/hess-21-3953-2017, 2017.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., and Stocker, E. F.: The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, Journal of hydrometeorology, 8, 38-55, https://doi.org/10.1175/JHM560.1, 2007.

Jiang, D., and Wang, K.: The Role of Satellite-Based Remote Sensing in Improving Simulated Streamflow: A Review, Water, 11, 1615, https://doi.org/10.3390/w11081615, 2019.

Jiao, Y., Lei, H. M., Yang, D. W., Huang, M. Y., Liu, D. F., and Yuan, X.: Impact of vegetation dynamics on hydrological processes in a semi-arid basin by using a land surface-hydrology coupled model, Journal of Hydrology, 551, 116-131, https://doi.org/10.1016/j.jhydrol.2017.05.060, 2017.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P.: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution, Journal of Hydrometeorology, 5, 487-503, https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2, 2004.

Jung, H. C., Getirana, A., Arsenault, K. R., Holmes, T. R., and McNally, A.: Uncertainties in Evapotranspiration Estimates over West Africa, Remote Sensing, 11, 892, https://doi.org/10.3390/rs11080892 2019.

Kidd, C., and Huffman, G.: Global precipitation measurement, Meteorological Applications, 18, 334-353, https://doi.org/10.1002/met.284, 2011.

Kidd, C., and Levizzani, V.: Status of satellite precipitation retrievals, Hydrol Earth Syst Sc, 15, 1109-1116, https://doi.org/10.5194/hess-15-1109-2011, 2011.

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., and Kirschbaum, D. B.: So, how much of the Earth's surface is covered by rain gauges?, Bull Am Meteorol Soc, 98, 69-78, https://doi.org/10.1175/BAMS-D-14-00283.1, 2017.

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resources Research, 42, https://doi.org/10.1029/2005wr004362, 2006.

Kirchner, J. W., and Allen, S. T.: Seasonal partitioning of precipitation between streamflow and evapotranspiration, inferred from end-member splitting analysis, Hydrol Earth Syst Sc, 24, 17-39, https://doi.org/10.5194/hess-24-17-2020, 2020.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, Journal of Hydrology, 424, 264-277, https://doi.org/10.1016/j.jhydrol.2012.01.011, 2012.

Knoben, W. J., Freer, J. E., and Woods, R. A.: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores, Hydrol Earth Syst Sc, 23, 4323-4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., and Endo, H.: The JRA-55 reanalysis: General specifications and basic characteristics, Journal of the Meteorological Society of Japan. Ser. II, 93, 5-48, https://doi.org/10.2151/jmsj.2015-001, 2015.

Koch, J., Demirel, M. C., and Stisen, S.: The SPAtial EFficiency metric (SPAEF): multiple-component evaluation of spatial patterns for optimization of hydrological models, Geosci Model Dev, 11, 1873-1886, https://doi.org/10.5194/gmd-11-1873-2018, 2018.

Kubota, T., Aonashi, K., Ushio, T., Shige, S., Takayabu, Y. N., Kachi, M., Arai, Y., Tashima, T., Masaki, T., and Kawamoto, N.: Global Satellite Mapping of Precipitation (GSMaP) products in the GPM era, in: Satellite precipitation measurement, Springer, 355-373, https://doi.org/10.1007/978-3-030-24568-9_20, 2020.

Kuczera, G., Renard, B., Thyer, M., and Kavetski, D.: There are no hydrological monsters, just models and observations with large uncertainties!, Hydrological Sciences Journal, 55, 980-991, https://doi.org/10.1080/02626667.2010.504677, 2010.

Kumar, R., Samaniego, L., and Attinger, S.: Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations, Water Resources Research, 49, 360-379, https://doi.org/10.1029/2012wr012195, 2013.

Kvålseth, T. O.: Coefficient of variation: the second-order alternative, Journal of Applied Statistics, 44, 402-415, https://doi.org/10.1080/02664763.2016.1174195, 2017.

Laiti, L., Mallucci, S., Piccolroaz, S., Bellin, A., Zardi, D., Fiori, A., Nikulin, G., and Majone, B.: Testing the Hydrological Coherence of High-Resolution Gridded Precipitation and Temperature Data Sets, Water Resources Research, 54, 1999-2016, https://doi.org/10.1002/2017WR021633, 2018.

Landerer, F. W., and Swenson, S. C.: Accuracy of scaled GRACE terrestrial water storage estimates, Water Resources Research, 48, https://doi.org/10.1029/2011wr011453, 2012.

Lange, S.: EartH2Observe, WFDEI and ERA-Interim data Merged and Bias-corrected for ISIMIP (EWEMBI), in, edited by: Services, G. D., 2016.

Lauri, H., Räsänen, T., and Kummu, M.: Using reanalysis and remotely sensed temperature and precipitation data for hydrological modeling in monsoon climate: Mekong River case study, Journal of Hydrometeorology, 15, 1532-1545, https://doi.org/10.1175/Jhm-D-13-084.1, 2014.

Le Coz, C., and van de Giesen, N.: Comparison of rainfall products over sub-Sahara Africa, Journal of Hydrometeorology, https://doi.org/10.1175/JHM-D-18-0256.1, 2019.

Ledesma, J. L., and Futter, M. N.: Gridded climate data products are an alternative to instrumental measurements as inputs to rainfall–runoff models, Hydrological Processes, 31, 3283-3293, https://doi.org/10.1002/hyp.11269, 2017.

Levizzani, V., Kidd, C., Kirschbaum, D. B., Kummerow, C. D., Nakamura, K., and Turk, F. J.: Satellite Precipitation Measurement, Springer, https://doi.org/10.1007/978-3-030-24568-9, 2020.

Li, L., Ngongondo, C. S., Xu, C.-Y., and Gong, L.: Comparison of the global TRMM and WFD precipitation datasets in driving a large-scale hydrological model in southern Africa, Hydrology Research, 44, 770-788, https://doi.org/10.2166/nh.2012.175, 2012a.

Li, X.-H., Zhang, Q., and Xu, C.-Y.: Suitability of the TRMM satellite rainfalls in driving a distributed hydrological model for water balance computations in Xinjiang catchment, Poyang lake basin, Journal of Hydrology, 426, 28-38, https://doi.org/10.1016/j.jhydrol.2012.01.013, 2012b.

Liu, X., Yang, T., Hsu, K., Liu, C., and Sorooshian, S.: Evaluating the streamflow simulation capability of PERSIANN-CDR daily rainfall products in two river basins on the Tibetan Plateau, Hydrol Earth Syst Sc, 21, 169-181, https://doi.org/10.5194/hess-21-169-2017, 2017.

Lorenz, C., and Kunstmann, H.: The hydrological cycle in three state-of-the-art reanalyses: Intercomparison and performance analysis, Journal of Hydrometeorology, 13, 1397-1420, https://doi.org/10.1175/Jhm-D-11-088.1, 2012.

Ma, Q., Xiong, L., Liu, D., Xu, C.-Y., and Guo, S.: Evaluating the Temporal Dynamics of Uncertainty Contribution from Satellite Precipitation Input in Rainfall-Runoff Modeling Using the Variance Decomposition Method, Remote Sensing, 10, 1876, https://doi.org/10.3390/rs10121876, 2018.

Maggioni, V., Meyers, P. C., and Robinson, M. D.: A review of merged high-resolution satellite precipitation product accuracy during the Tropical Rainfall Measuring Mission (TRMM) era, Journal of Hydrometeorology, 17, 1101-1117, https://doi.org/10.1175/Jhm-D-15-0190.1, 2016.

Maggioni, V., and Massari, C.: On the performance of satellite precipitation products in riverine flood modeling: A review, Journal of hydrology, 558, 214-224, https://doi.org/10.1016/j.jhydrol.2018.01.039, 2018.

Maidment, R., Black, E., Greatrex, H., and Young, M.: TAMSAT, in: Satellite Precipitation Measurement, Springer, 393-407, https://doi.org/10.1007/978-3-030-24568-9_22, 2020.

Maidment, R. I., Grimes, D., Allan, R. P., Tarnavsky, E., Stringer, M., Hewison, T., Roebeling, R., and Black, E.: The 30 year TAMSAT African rainfall climatology and time series (TARCAT) data set, Journal of Geophysical Research: Atmospheres, 119, 10,619-610,644, https://doi.org/10.1002/2014jd021927, 2014.

Maidment, R. I., Allan, R. P., and Black, E.: Recent observed and simulated changes in precipitation over Africa, Geophys Res Lett, 42, 8155-8164, https://doi.org/10.1002/2015gl065765, 2015.

Maidment, R. I., Grimes, D., Black, E., Tarnavsky, E., Young, M., Greatrex, H., Allan, R. P., Stein, T., Nkonde, E., and Senkunda, S.: A new, long-term daily satellite-based rainfall dataset for operational monitoring in Africa, Scientific data, 4, 170063, https://doi.org/10.1038/sdata.2017.63, 2017.

Marra, F., Nikolopoulos, E. I., Anagnostou, E. N., Bardossy, A., and Morin, E.: Precipitation Frequency Analysis from Remotely Sensed Datasets: A Focused Review, Journal of Hydrology, 574, 699-705, https://doi.org/10.1016/j.jhydrol.2019.04.081, 2019.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci Model Dev, 10, 1903-1925, https://doi.org/10.5194/gmd-10-1903-2017, 2017.

Marthews, T. R., Blyth, E. M., Martínez-de la Torre, A., and Veldkamp, T. I.: A global-scale evaluation of extreme event uncertainty in the eartH2Observe project, Hydrol Earth Syst Sc, 24, 75-92, https://doi.org/10.5194/hess-24-75-2020, 2020.

Massari, C., Crow, W., and Brocca, L.: An assessment of the performance of global rainfall estimates without ground-based observations, Hydrol Earth Syst Sc, 21, 4347-4361, https://doi.org/10.5194/hess-21-4347-2017, 2017.

Massari, C., Brocca, L., Pellarin, T., Abramowitz, G., Filippucci, P., Ciabatta, L., Maggioni, V., Kerr, Y., and Fernandez Prieto, D.: A daily 25 km short-latency rainfall product for data-scarce regions based on the integration of the Global Precipitation Measurement mission rainfall and multiple-satellite soil moisture products, Hydrol Earth Syst Sc, 24, 2687-2710, https://doi.org/10.5194/hess-24-2687-2020, 2020.

Mathon, V., Laurent, H., and Lebel, T.: Mesoscale convective system rainfall in the Sahel, Journal of applied meteorology, 41, 1081-1092, https://doi.org/10.1175/1520-0450(2002)041<1081:Mcsrit>2.0.Co;2, 2002.

Mazzoleni, M., Brandimarte, L., and Amaranto, A.: Evaluating precipitation datasets for large-scale distributed hydrological modelling, Journal of Hydrology, 124076, https://doi.org/10.1016/j.jhydrol.2019.124076, 2019.

McCabe, M. F., Rodell, M., Alsdorf, D. E., Miralles, D. G., Uijlenhoet, R., Wagner, W., Lucieer, A., Houborg, R., Verhoest, N. E. C., Franz, T. E., Shi, J., Gao, H., and Wood, E. F.: The Future of Earth Observation in Hydrology, Hydrol Earth Syst Sci, 21, 3879-3914, https://doi.org/10.5194/hess-21-3879-2017, 2017.

McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., and Stoffelen, A.: Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target, Geophys Res Lett, 41, 6229-6236, https://doi.org/10.1002/2014GL061322, 2014.

McMillan, H., Freer, J., Pappenberger, F., Krueger, T., and Clark, M.: Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge predictions, Hydrological Processes: An International Journal, 24, 1270-1284, https://doi.org/10.1002/hyp.7587, 2010.

McMillan, H. K., Westerberg, I. K., and Krueger, T.: Hydrological data uncertainty and its implications, Wiley Interdisciplinary Reviews: Water, 5, e1319, https://doi.org/10.1002/wat2.1319, 2018.
Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrol Earth Syst Sc, 15, 453-469, https://doi.org/10.5194/hess-15-453-2011, 2011.
Mizukami, N., Clark, M. P., Newman, A. J., Wood, A. W., Gutmann, E. D., Nijssen, B., Rakovec, O., and Samaniego, L.: Towards seamless large-domain parameter estimation for hydrologic models, Water Resources Research, 53, 8020-8040, https://doi.org/10.1002/2017wr020401, 2017.
Mul, M., Obuobie, E., Appoh, R., Kankam-Yeboah, K., Bekoe-Obeng, E., Amisigo, B., Logah, F. Y., Ghansah, B., and McCartney, M.: Water resources assessment of the Volta River Basin, International Water Management Institute (IWMI)9290908297, 82, 2015.
Nash, J. E., and Sutcliffe, J. V.: River flow forecasting through conceptual models part I—A discussion of principles, Journal of hydrology, 10, 282-290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.
Nicholson, S. E.: The West African Sahel: A review of recent studies on the rainfall regime and its interannual variability, ISRN Meteorology, 2013, https://doi.org/10.1155/2013/453521, 2013.
Nicholson, S. E., Fink, A. H., and Funk, C.: Assessing recovery and change in West Africa's rainfall regime from a 161-year record, International Journal of Climatology, 38, 3770-3786, https://doi.org/10.1002/joc.5530, 2018a.
Nicholson, S. E., Funk, C., and Fink, A. H.: Rainfall over the African continent from the 19th through the 21st century, Global and planetary change, 165, 114-127, https://doi.org/10.1016/j.gloplacha.2017.12.014, 2018b.
Nijssen, B.: Effect of precipitation sampling error on simulated hydrological fluxes and states: Anticipating the Global Precipitation Measurement satellites, Journal of Geophysical Research, 109, https://doi.org/10.1029/2003jd003497, 2004.
Nikolopoulos, E. I., Anagnostou, E. N., Hossain, F., Gebremichael, M., and Borga, M.: Understanding the scale relationships of uncertainty propagation of satellite rainfall through a distributed hydrologic model, Journal of Hydrometeorology, 11, 520-532, https://doi.org/10.1175/2009JHM1169.1, 2010.
Nkiaka, E., Nawaz, N., and Lovett, J. C.: Evaluating global reanalysis datasets as input for hydrological modelling in the Sudano-Sahel region, Hydrology, 4, 13, https://doi.org/10.3390/hydrology4010013, 2017.
Novella, N. S., and Thiaw, W. M.: African rainfall climatology version 2 for famine early warning systems, Journal of Applied Meteorology and Climatology, 52, 588-606, https://doi.org/10.1175/JAMC-D-11-0238.1, 2013.
Ocio, D., Beskeen, T., and Smart, K.: Fully distributed hydrological modelling for catchment-wide hydrological data verification, Hydrology Research, https://doi.org/10.2166/nh.2019.006, 2019.
Pan, M., Li, H., and Wood, E.: Assessing the skill of satellite-based precipitation estimates in hydrologic applications, Water Resources Research, 46, https://doi.org/10.1029/2009WR008290, 2010.
Paniconi, C., and Putti, M.: Physically based modeling in catchment hydrology at 50: Survey and outlook, Water Resources Research, 51, 7090-7129, https://doi.org/10.1002/2015wr017780, 2015.
Parker, D. J., and Diop-Kane, M.: Meteorology of tropical West Africa: The forecasters' handbook, John Wiley & Sons, 468 pp., https://doi.org/10.1002/9781118391297, 2017.
Peel, M. C., and McMahon, T. A.: Historical development of rainfall-runoff modeling, Wiley Interdisciplinary Reviews: Water, e1471, https://doi.org/10.1002/wat2.1471, 2020.
Pellarin, T., Román-Cascón, C., Baron, C., Bindlish, R., Brocca, L., Camberlin, P., Fernández-Prieto, D., Kerr, Y. H., Massari, C., Panthou, G., Perrimond, B., Philippon, N., and Quantin, G.: The Precipitation Inferred from Soil Moisture (PrISM) near Real-Time Rainfall Product: Evaluation and Comparison, Remote Sensing, 12, 481, https://doi.org/10.3390/rs12030481 2020.
Peters-Lidard, C. D., Clark, M., Samaniego, L., Verhoest, N. E. C., van Emmerik, T., Uijlenhoet, R., Achieng, K., Franz, T. E., and Woods, R.: Scaling, Similarity, and the Fourth Paradigm for Hydrology, Hydrol Earth Syst Sci, 21, 3701-3713, https://doi.org/10.5194/hess-2016-695, 2017.
Pfeifroth, U., Trentmann, J., Fink, A. H., and Ahrens, B.: Evaluating satellite-based diurnal cycles of precipitation in the African tropics, Journal of Applied Meteorology and Climatology, 55, 23-39, https://doi.org/10.1175/Jamc-D-15-0065.1, 2016.
Pfister, L., and Kirchner, J. W.: Debates—Hypothesis testing in hydrology: Theory and practice, Water Resources Research, 53, 1792-1798, https://doi.org/10.1002/2016WR020116, 2017.

Poméon, T., Jackisch, D., and Diekkrüger, B.: Evaluating the performance of remotely sensed and reanalysed precipitation data over West Africa using HBV light, Journal of hydrology, 547, 222-235, https://doi.org/10.1016/j.jhydrol.2017.01.055, 2017.

Poméon, T., Diekkrüger, B., and Kumar, R.: Computationally Efficient Multivariate Calibration and Validation of a Grid-Based Hydrologic Model in Sparsely Gauged West African River Basins, Water, 10, 1418, https://doi.org/10.3390/w10101418, 2018.

Potter, G. L., Carriere, L., Hertz, J., Bosilovich, M., Duffy, D., Lee, T., and Williams, D. N.: Enabling reanalysis research using the collaborative reanalysis technical environment (CREATE), Bulletin of the American Meteorological Society, 99, 677-687, https://doi.org/10.1175/Bams-D-17-0174.1, 2018.

Qi, W., Zhang, C., Fu, G., Sweetapple, C., and Zhou, H.: Evaluation of global fine-resolution precipitation products and their uncertainty quantification in ensemble discharge simulations, Hydrol Earth Syst Sc, 20, 903-920, https://doi.org/10.5194/hess-20-903-2016, 2016.

Raimonet, M., Oudin, L., Thieu, V., Silvestre, M., Vautard, R., Rabouille, C., and Le Moigne, P.: Evaluation of gridded meteorological datasets for hydrological modeling, Journal of Hydrometeorology, 18, 3027-3041, https://doi.org/10.1175/JHM-D-17-0018.1, 2017.

Reichle, R. H., Liu, Q., Koster, R. D., Draper, C. S., Mahanama, S. P., and Partyka, G. S.: Land surface precipitation in MERRA-2, Journal of Climate, 30, 1643-1664, https://doi.org/10.1175/JCLI-D-16-0570.1, 2017.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., and Franks, S. W.: Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, Water Resources Research, 46, https://doi.org/10.1029/2009WR008328, 2010.

Roca, R., Alexander, L. V., Potter, G., Bador, M., Jucá, R., Contractor, S., Bosilovich, M. G., and Cloché, S.: FROGS: a daily 1°× 1° gridded precipitation database of rain gauge, satellite and reanalysis products, Earth System Science Data, 11, 1017-1035, https://doi.org/10.5194/essd-11-1017-2019, 2019.

Roebeling, R., Wolters, E., Meirink, J., and Leijnse, H.: Triple collocation of summer precipitation retrievals from SEVIRI over Europe with gridded rain gauge and weather radar data, Journal of hydrometeorology, 13, 1552-1566, https://doi.org/10.1175/JHM-D-11-089.1, 2012.

Romilly, T. G., and Gebremichael, M.: Evaluation of satellite rainfall estimates over Ethiopian river basins, Hydrol Earth Syst Sc, 15, 1505-1514, https://doi.org/10.5194/hess-15-1505-2011, 2011.

Roudier, P., Ducharne, A., and Feyen, L.: Climate change impacts on runoff in West Africa: a review, Hydrol Earth Syst Sc, 18, 2789-2801, https://doi.org/10.5194/hess-18-2789-2014, 2014.

Sakumura, C., Bettadpur, S., and Bruinsma, S.: Ensemble prediction and intercomparison analysis of GRACE time-variable gravity field models, Geophys Res Lett, 41, 1389-1397, https://doi.org/10.1002/2013GL058632, 2014.

Samaniego, L., Kumar, R., and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resources Research, 46, https://doi.org/10.1029/2008wr007327, 2010.

Samaniego, L., Kumar, R., Thober, S., Rakovec, O., Zink, M., Wanders, N., Eisner, S., Müller Schmied, H., Sutanudjaja, E. H., Warrach-Sagi, K., and Attinger, S.: Toward seamless hydrologic predictions across spatial scales, Hydrol Earth Syst Sc, 21, 4323-4346, https://doi.org/10.5194/hess-21-4323-2017, 2017.

Satgé, F., Ruelland, D., Bonnet, M.-P., Molina, J., and Pillco, R.: Consistency of satellite-based precipitation products in space and over time compared with gauge observations and snow-hydrological modelling in the Lake Titicaca region, Hydrol Earth Syst Sc, 23, 595-619, https://doi.org/10.5194/hess-23-595-2019, 2019.

Satgé, F., Defrance, D., Sultan, B., Bonnet, M.-P., Seyler, F., Rouché, N., Pierron, F., and Paturel, J.-E.: Evaluation of 23 gridded precipitation datasets across West Africa, Journal of Hydrology, 581, 124412, https://doi.org/10.1016/j.jhydrol.2019.124412, 2020.

Savenije, H. H.: HESS Opinions" The art of hydrology", Hydrol Earth Syst Sc, 13, 157-161, https://doi.org/10.5194/hess-13-157-2009, 2009.

Schaefli, B., and Gupta, H. V.: Do Nash values have value?, Hydrological Processes, 21, 2075-2080, https://doi.org/10.1002/hyp.6825, 2007.

Schneider, T., Bischoff, T., and Haug, G. H.: Migrations and dynamics of the intertropical convergence zone, Nature, 513, 45, https://doi.org/10.1038/nature13636, 2014.

Schröder, M., Lockhoff, M., Fell, F., Forsythe, J., Trent, T., Bennartz, R., Borbas, E., Bosilovich, M. G., Castelli, E., and Hersbach, H.: The GEWEX Water Vapor Assessment archive of water vapour products from satellite observations and reanalyses, Earth system science data, 10, 1093-1117, https://doi.org/10.5194/essd-10-1093-2018, 2018.

Seibert, J., Vis, M. J., Lewis, E., and Meerveld, H. v.: Upper and lower benchmarks in hydrological modelling, Hydrological Processes, https://doi.org/10.1002/hyp.11476, 2018.

Serdeczny, O., Adams, S., Baarsch, F., Coumou, D., Robinson, A., Hare, W., Schaeffer, M., Perrette, M., and Reinhardt, J.: Climate change impacts in Sub-Saharan Africa: from physical changes to their social repercussions, Regional Environmental Change, 17, 1585-1600, https://doi.org/10.1007/s10113-015-0910-2, 2017.

Serrat-Capdevila, A., Valdes, J. B., and Stakhiv, E. Z.: Water management applications for satellite precipitation products: Synthesis and recommendations, JAWRA Journal of the American Water Resources Association, 50, 509-525, https://doi.org/10.1111/jawr.12140, 2014.

Seyyedi, H., Anagnostou, E. N., Beighley, E., and McCollum, J.: Hydrologic evaluation of satellite and reanalysis precipitation datasets over a mid-latitude basin, Atmospheric Research, 164, 37-48, https://doi.org/10.1016/j.atmosres.2015.03.019, 2015.

Shawul, A. A., and Chakma, S.: Suitability of global precipitation estimates for hydrologic prediction in the main watersheds of Upper Awash basin, Environmental Earth Sciences, 79, 53, https://doi.org/10.1007%2Fs12665-019-8801-3, 2020.

Shayeghi, A., Azizian, A., and Brocca, L.: Reliability of reanalysis and remotely sensed precipitation products for hydrological simulation over the Sefidrood River Basin, Iran, Hydrological Sciences Journal, 65, 296-310, https://doi.org/10.1080/02626667.2019.1691217, 2020.

Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, Journal of climate, 19, 3088-3111, https://doi.org/10.1175/JCLI3790.1, 2006.

Sheffield, J., Wood, E. F., Pan, M., Beck, H., Coccia, G., Serrat-Capdevila, A., and Verbist, K.: Satellite Remote Sensing for Water Resources Management: Potential for Supporting Sustainable Development in Data-Poor Regions, Water Resources Research, 54, 9724-9758, https://doi.org/10.1029/2017wr022437, 2018.

Singh, V. P.: Hydrologic modeling: progress and future directions, Geoscience Letters, 5, 15, https://doi.org/10.1186/s40562-018-0113-z, 2018.

Sorooshian, S., AghaKouchak, A., Arkin, P., Eylander, J., Foufoula-Georgiou, E., Harmon, R., Hendrickx, J. M., Imam, B., Kuligowski, R., and Skahill, B.: Advanced concepts on remote sensing of precipitation at multiple scales, Bulletin of the American Meteorological Society, 92, 1353-1357, https://doi.org/10.1175/2011bams3158.1, 2011.

Stephens, G. L., and Kummerow, C. D.: The remote sensing of clouds and precipitation from space: A review, Journal of the Atmospheric Sciences, 64, 3742-3765, https://doi.org/10.1175/2006jas2375.1, 2007.

Stisen, S., Højberg, A., Troldborg, L., Refsgaard, J., Christensen, B., Olsen, M., and Henriksen, H.: On the importance of appropriate precipitation gauge catch correction for hydrological modelling at mid to high latitudes, Hydrol Earth Syst Sc, 16, https://doi.org/10.5194/hess-16-4157-2012, 2012.

Su, F., Hong, Y., and Lettenmaier, D. P.: Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin, Journal of Hydrometeorology, 9, 622-640, https://doi.org/10.1175/2007jhm944.1, 2008.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K. L.: A review of global precipitation data sets: Data sources, estimation, and intercomparisons, Reviews of Geophysics, 56, 79-107, https://doi.org/10.1002/2017rg000574, 2018.

Swenson, S. C.: GRACE monthly land water mass grids NETCDF release 5.0. Ver. 5.0. PO.DAAC, CA, USA, [Dataset accessed 2018.11.01], https://doi.org/10.5067/TELND-NC005, 2012.

Sylla, M., Giorgi, F., Coppola, E., and Mariotti, L.: Uncertainties in daily rainfall over Africa: assessment of gridded observation products and evaluation of a regional climate model simulation, International Journal of Climatology, 33, 1805-1817, https://doi.org/10.1002/joc.3551, 2013.

Tang, X., Zhang, J., Gao, C., Ruben, G. B., and Wang, G.: Assessing the Uncertainties of Four Precipitation Products for Swat Modeling in Mekong River Basin, Remote Sensing, 11, 304, https://doi.org/10.3390/rs11030304, 2019.

Tapiador, F., Turk, F. J., Petersen, W., Hou, A. Y., García-Ortega, E., Machado, L. A., Angelis, C. F., Salio, P., Kidd, C., and Huffman, G. J.: Global precipitation measurement: Methods, datasets and applications, Atmospheric Research, 104, 70-97, https://doi.org/10.1016/j.atmosres.2011.10.021, 2012.

Tapiador, F., Navarro, A., Levizzani, V., García-Ortega, E., Huffman, G., Kidd, C., Kucera, P., Kummerow, C., Masunaga, H., and Petersen, W.: Global precipitation measurements for validating climate models, Atmospheric Research, 197, 1-20, https://doi.org/10.1016/j.atmosres.2017.06.021, 2017.

Tapley, B. D., Bettadpur, S., Watkins, M., and Reigber, C.: The gravity recovery and climate experiment: Mission overview and early results, Geophys Res Lett, 31, https://doi.org/10.1029/2004gl019920, 2004.

Tarnavsky, E., Grimes, D., Maidment, R., Black, E., Allan, R. P., Stringer, M., Chadwick, R., and Kayitakire, F.: Extension of the TAMSAT satellite-based rainfall monitoring over Africa and from 1983 to present, Journal of Applied Meteorology and Climatology, 53, 2805-2822, https://doi.org/10.1175/JAMC-D-14-0016.1, 2014.

Tauro, F., Selker, J., Van De Giesen, N., Abrate, T., Uijlenhoet, R., Porfiri, M., Manfreda, S., Caylor, K., Moramarco, T., and Benveniste, J.: Measurements and Observations in the XXI century (MOXXI): innovation and multi-disciplinarity to sense the hydrological cycle, Hydrological sciences journal, 63, 169-196, https://doi.org/10.1080/02626667.2017.1420191, 2018.

Taylor, C. M., Belušić, D., Guichard, F., Parker, D. J., Vischel, T., Bock, O., Harris, P. P., Janicot, S., Klein, C., and Panthou, G.: Frequency of extreme Sahelian storms tripled since 1982 in satellite observations, Nature, 544, 475, https://doi.org/10.1038/nature22069, 2017.

Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., Levizzani, V., and De Roo, A.: Validation of satellite-based precipitation products over sparsely gauged African river basins, Journal of Hydrometeorology, 13, 1760-1783, https://doi.org/10.1175/Jhm-D-12-032.1, 2012.

Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., and De Roo, A.: Hydrological evaluation of satellite-based rainfall estimates over the Volta and Baro-Akobo Basin, Journal of Hydrology, 499, 324-338, https://doi.org/10.1016/j.jhydrol.2013.07.012, 2013.

Thober, S., Cuntz, M., Kelbling, M., Kumar, R., Mai, J., and Samaniego, L. J. G. M. D. D.: The multiscale Routing Model mRM v1. 0: simple river routing at resolutions from 1 to 50 km, 2019, 1-26, https://doi.org/10.5194/gmd-12-2501-2019, 2019.

Thorncroft, C. D., Nguyen, H., Zhang, C., and Peyrillé, P.: Annual cycle of the West African monsoon: regional circulations and associated water vapour transport, Quarterly Journal of the Royal Meteorological Society, 137, 129-147, https://doi.org/10.1002/qj.728, 2011.

Tobin, K. J., and Bennett, M. E.: Satellite precipitation products and hydrologic applications, Water International, 39, 360-380, https://doi.org/10.1080/02508060.2013.870423, 2014.

Tolson, B. A., and Shoemaker, C. A.: Dynamically dimensioned search algorithm for computationally efficient watershed model calibration, Water Resources Research, 43, https://doi.org/10.1029/2005wr004723, 2007.

Trabucco, A., and Zomer, R.: Global Aridity Index and Potential Evapo-Transpiration (ET0) Climate Database v2. CGIAR Consortium for Spatial Information (CGIAR-CSI). Published online, available from the CGIAR-CSI GeoPortal at https://cgiarcsi.community, accessed 11.07.2019], 2018.

Tramblay, Y., Thiemig, V., Dezetter, A., and Hanich, L.: Evaluation of satellite-based rainfall products for hydrological modelling in Morocco, Hydrological Sciences Journal, 61, 2509-2519, https://doi.org/10.1080/02626667.2016.1154149, 2016.

Tucker, C. J., Pinzon, J. E., Brown, M. E., Slayback, D. A., Pak, E. W., Mahoney, R., Vermote, E. F., and El Saleous, N.: An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data, Int J Remote Sens, 26, 4485-4498, https://doi.org/10.1080/01431160500168686, 2005.

Turner, A., SPERBER, K. R., Slingo, J., Meehl, G., MECHOSO, C. R., Kimoto, M., and Giannini, A.: Modelling monsoons: Understanding and predicting current and future behaviour, in: The Global Monsoon System: Research and Forecast, World Scientific, 421-454, https://doi.org/10.1142/9789814343411_0025, 2011.

Ushio, T., Sasashige, K., Kubota, T., Shige, S., Okamoto, K. i., Aonashi, K., Inoue, T., Takahashi, N., Iguchi, T., and Kachi, M.: A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data, Journal of the Meteorological Society of Japan. Ser. II, 87, 137-151, https://doi.org/10.2151/jmsj.87A.137, 2009.

Ushio, T., Mega, T., and Kubota, T.: Multi-satellite Global Satellite Mapping of Precipitation (GSMaP)-Design and Products, 2019 URSI Asia-Pacific Radio Science Conference (AP-RASC), 2019, 1-1, Van Stan, J. T., Gutmann, E., and Friesen, J.: Precipitation Partitioning by Vegetation: A Global Synthesis, Springer Nature, Switzerland, 295 pp., https://doi.org/10.1007/978-3-030-29702-2, 2020.

Vizy, E. K., and Cook, K. H.: Mesoscale convective systems and nocturnal rainfall over the West African Sahel: role of the Inter-tropical front, Climate dynamics, 50, 587-614, https://doi.org/10.1007/s00382-017-3628-7, 2018.

Voisin, N., Wood, A. W., and Lettenmaier, D. P.: Evaluation of precipitation products for global hydrological prediction, Journal of Hydrometeorology, 9, 388-407, https://doi.org/10.1175/2007jhm938.1, 2008.

Wagner, P. D., Fiener, P., Wilken, F., Kumar, S., and Schneider, K.: Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions, Journal of Hydrology, 464, 388-400, https://doi.org/10.1016/j.jhydrol.2012.07.026, 2012a.

Wagner, W., Dorigo, W., de Jeu, R., Fernandez, D., Benveniste, J., Haas, E., and Ertl, M.: Fusion of active and passive microwave observations to create an essential climate variable data record on soil moisture, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS Annals), 2012b, 315-321,

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, Water Resources Research, 50, 7505-7514, https://doi.org/10.1002/2014WR015638, 2014.

Westerberg, I. K., and Birkel, C.: Observational uncertainties in hypothesis testing: investigating the hydrological functioning of a tropical catchment, Hydrological Processes, 29, 4863-4879, https://doi.org/10.1002/hyp.10533, 2015.

Wilby, R. L.: A global hydrology research agenda fit for the 2030s, Hydrology Research, https://doi.org/10.2166/nh.2019.100, 2019.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J., Groth, P., Goble, C., Grethe, J. S., Heringa, J., t Hoen, P. A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S. A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Mons, B.: The FAIR Guiding Principles for scientific data management and stewardship, Sci Data, 3, 160018, https://doi.org/10.1038/sdata.2016.18, 2016.

Williams, T. O., Mul, M. L., Biney, C. A., and Smakhtin, V.: The Volta River Basin: Water for food, economic growth and environment, Routledge, 282 pp., 2016.

Xie, P., and Arkin, P. A.: Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions, Journal of climate, 9, 840-858, https://doi.org/10.1175/1520-0442(1996)009<0840:AOGMPU>2.0.CO;2, 1996.

Xie, P., Joyce, R., Wu, S., Yoo, S.-H., Yarosh, Y., Sun, F., and Lin, R.: Reprocessed, bias-corrected CMORPH global high-resolution precipitation estimates from 1998, Journal of Hydrometeorology, 18, 1617-1641, https://doi.org/10.1175/JHM-D-16-0168.1, 2017.

Xu, X., Li, J., and Tolson, B. A.: Progress in integrating remote sensing data and hydrologic modeling, Progress in Physical Geography: Earth and Environment, 38, 464-498, https://doi.org/10.1177/0309133314536583, 2014.

Zambrano-Bigiarini, M., Nauditt, A., Birkel, C., Verbist, K., and Ribbe, L.: Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile, Hydrol Earth Syst Sc, 21, 1295-1320, https://doi.org/10.5194/hess-21-1295-2017, 2017.

Zandler, H., Haag, I., and Samimi, C.: Evaluation needs and temporal performance differences of gridded precipitation products in peripheral mountain regions, Scientific reports, 9, 1-15, https://doi.org/10.1038/s41598-019-51666-z, 2019.

Zhang, D., Liu, X., Bai, P., and Li, X.-H.: Suitability of satellite-based precipitation products for water balance simulations using multiple observations in a humid catchment, Remote Sensing, 11, 151, https://doi.org/10.3390/rs11020151, 2019.

Zheng, H., Yang, Z. L., Lin, P., Wei, J., Wu, W. Y., Li, L., Zhao, L., and Wang, S.: On the sensitivity of the precipitation partitioning into evapotranspiration and runoff in land surface parameterizations, Water Resources Research, 55, 95-111, https://doi.org/10.1029/2017WR022236, 2019.

Zhu, Z. C., Bi, J., Pan, Y. Z., Ganguly, S., Anav, A., Xu, L., Samanta, A., Piao, S. L., Nemani, R. R., and Myneni, R. B.: Global Data Sets of Vegetation Leaf Area Index (LAI)3g and Fraction of Photosynthetically Active Radiation (FPAR)3g

Derived from Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) for the Period 1981 to 2011, Remote Sensing, 5, 927-948, https://doi.org/10.3390/rs5020927, 2013.