Climate-dependent propagation of precipitation uncertainty into the water cycle

Ali Fallah[1,2], Sungmin O[2], and Rene Orth[2]

[1]Department of Civil and Environmental Engineering, Shiraz University, Shiraz, Iran

5 [2]Department of Biogeochemical Integration, Max Planck Institute for Biogeochemistry, D-07745, Jena, Germany

Correspondence to: Ali Fallah (alifallah@shirazu.ac.ir; afallah@bgc-jena.mpg.de)

Abstract. Precipitation is a crucial variable for hydro-meteorological applications. Unfortunately, rain gauge measurements are sparse and unevenly distributed, which substantially hampers the use of in-situ precipitation data in many regions of the world.

10 The increasing availability of high-resolution gridded precipitation products presents a valuable alternative, especially over gauge-sparse regions. Nevertheless, uncertainties and corresponding differences across products can limit the applicability of these data. This study examines the usefulness of current state-of-the-art precipitation datasets in hydrological modelling. For this purpose, we force a conceptual hydrological model with multiple precipitation datasets in >200 European catchments. We consider a wide range of precipitation products, which are generated via (1) interpolation of gauge measurements (E-OBS and GPCC V.2018), (2)

15 combination of multiple sources (MSWEP V2) and (3) data assimilation into reanalysis models (ERA-Interim, ERA5, and CFSR). For each catchment, runoff and evapotranspiration simulations are obtained by forcing the model with the various precipitation products. Evaluation is done at the monthly time scale during the period of 1984-2007. We find that simulated runoff values are highly dependent on the accuracy of precipitation inputs, and thus show significant differences between the simulations. By contrast, simulated evapotranspiration is generally much less influenced. The results are further analysed with respect to different

20 hydro-climatic regimes. We find that the impact of precipitation uncertainty on simulated runoff increases towards wetter regions, while the opposite is observed in the case of evapotranspiration. Finally, we perform an indirect performance evaluation of the precipitation datasets by comparing the runoff simulations with streamflow observations. Thereby, E-OBS yields the best agreement, while furthermore ERA5, GPCC V.2018 and MSWEP V2 show good performance. In summary, our findings highlight a climate-dependent propagation of precipitation uncertainty through the water cycle; while runoff is strongly impacted in

25 comparatively wet regions such as Central Europe, there are increasing implications on evapotranspiration towards drier regions.

1. Introduction

30 Precipitation is a key quantity in the water cycle since it controls water availability including both blue and green water resources (Falkenmark and Rockström, 2006; Orth and Destouni, 2018). This way, changes in precipitation translate into changes in water resources which could have severe impacts on ecosystems, and consequently economy and society (Oki and Kanae, 2006; Kirtman et al., 2013; Abbott et al., 2019). Changes in precipitation can be induced or intensified by climate change and consequently lead to amplified impacts (Blöschl et al., 2017; Blöschl et al., 2019b). Thus, accurate precipitation information is essential for

35 monitoring water resources and managing related impacts.

Despite the necessity of accurate precipitation datasets, reliable gauge measurements are not widely available. Further, these measurements need to be corrected for potential errors such as wind-induced inaccuracies or precipitation undercatch, especially in higher altitudes (Mekonnen et al., 2015). Next to gauge measurements, precipitation information can be inferred from satellite observations and/or model simulations. Based on these sources, a variety of global gridded precipitation datasets have emerged.

40 While some of these datasets make direct use of gauge measurements to interpolate them in time and space, others make indirect use of the gauge information to calibrate satellite retrieval algorithms or models, enabling them to estimate gridded large-scale precipitation.

Across these datasets, there are ample discrepancies in space and time, highlighting the need for comparative assessments (e.g. Koutsouris et al., 2016; Alijanian et al., 2017, 2019; Balsamo et al., 2018; Sun et al., 2018; Massari et al., 2019; Levizzani and

45 Cattani, 2019; Fallah et al., 2019; Satgé et al., 2020). In particular, indirect evaluation of the datasets through application in hydrological modelling is a valuable alternative in this context, as precipitation is translated into variables with more (reliable) large-scale observations such as runoff (Thiemig et al., 2013; Nerini et al., 2015; Beck et al., 2017a,b,2019a,b; Arheimer et al., 2019; Fereidoon et al., 2019; Bhuiyan et al., 2019; Mazzoleni et al., 2019). However, while this approach relies on the propagation of precipitation uncertainty into runoff it is largely underexplored when and where this propagation pathway is active. Vice versa,

50 it is unclear in which regions or conditions, gridded datasets of runoff (Gudmundsson and Seneviratne, 2016) or evapotranspiration (e.g. Martens et al., 2017; Jung et al., 2019) are impacted by the existing precipitation uncertainties.

In this study, we investigate the uncertainty across six widely used gridded precipitation datasets, including its propagation into the hydrological cycle, i.e. runoff and evapotranspiration (ET). Thereby, we consider gauge-interpolated (E-OBS, GPCC V.2018), multi-source (MSWEP V2), and reanalyses (ERA-Interim, ERA5, CFSR) datasets. With each of them, we force a conceptual land

55 surface model and compare the respectively simulated runoff and ET. This is done separately for different hydro-climatological regimes. In addition, validating the runoff simulations against respective observations we can indirectly infer the performance of the precipitation datasets. This further allows us to obtain guidelines with respect to the usefulness of the different types of precipitation products in the considered regimes.

Section 2 introduces the reference, forcing datasets and model calibration used in the study, and Section 3 illustrates results and

60 discussion. Finally, in Section 4 the conclusions of this study are presented.

2. Data and methodology

2.1. Forcing data

Runoff and ET are modelled with a conceptual hydrological model, the Simple Water Balance Model (SWBM). The underlying framework was initially presented by Koster and Mahanama 2012 where runoff (normalised by precipitation) and ET (normalised

65 by net radiation) are assumed to be polynomial functions of soil moisture (Whan et al., 2015). We use here the model version introduced by Orth and Seneviratne 2013. As inputs, the model uses temperature, net radiation, and precipitation. For each catchment, temperature and net radiation are used from the respective grid cells from the E-OBS (Haylock et al., 2008) and ERA-Interim (Dee et al., 2011) datasets, respectively. Corresponding grid cell-based precipitation data is used from various datasets derived from different sources: gauge-based (E-OBS, GPCC V.2018), multi-source (MSWEP V2) and reanalysis datasets (ERA-

70 Interim, ERA5, CFSR). A summary of all precipitation datasets and their respective characteristics is shown in Table 1.

Before using the precipitation datasets to force the SWBM, they are re-gridded to a common 0.5° spatial resolution, if necessary. This was done through conservative remapping which preserves the water mass (Jones, 1999) using climate data operators

(Schulzweida, 2019). While the SWBM simulations are performed with a daily time step, we focus on monthly averaged data throughout the analyses in this study to mitigate the influence of synoptic weather variability.

75 ## 2.2. Reference data

Modelled runoff is evaluated against monthly mean streamflow observations obtained from 426 catchments distributed across Europe (Stahl et al., 2010). These data are available for the period 1984-2007. There is no or little human influence on the streamflow in these catchments, which are mostly between 10-100 km^2 in size.

2.3. Model calibration

80 In a first step, the best possible model performance was determined in each catchment to test the respective applicability of the model. For this purpose, the model is calibrated against streamflow observations in each catchment. The >400 catchments are distributed across Europe, and across different hydro-climatological regions (Fig. 1). The agreement between modelled and observed runoff is determined by computing the Nash-Sutcliffe efficiency (NSE, Nash and Sutcliffe, 1970) using monthly data during 1984-2007. Only catchments where NSE>0.36 (Motovilov et al., 1999; Moriasi, 2007) are retained for the further analyses, 85 which leaves 243 catchments that are well distributed across the continent.

As shown in Fig. 1, the hydro-climatological regime is characterised through long-term average temperature and aridity (Budyko, 1974). Thereby, for each catchment, the temperature derived from the E-OBS dataset, and aridity is computed as the ratio of mean annual net radiation to mean annual precipitation calculated from ERA-Interim and E-OBS respectively.

In each of the 243 catchments, the SWBM is forced with temperature, net radiation and the different precipitation datasets, 90 respectively (Fig. 2). This way, six simulations with the six different precipitation datasets are performed for each catchment, leaving the temperature and net radiation data unchanged. The model parameters are thereby obtained from the above-mentioned calibration using E-OBS precipitation.

All analyses are performed during the warm season (May-September) to exclude the impact of snow and ice, and because ET is of minor relevance during cold months.

95

3. Results and discussion

3.1. Impact of precipitation uncertainty on runoff and ET

Figure 3 illustrates the propagation of precipitation uncertainty into simulated runoff and ET. Each point denotes the standard deviation across the six simulations obtained with the different precipitation datasets and represents a particular month in a specific 100 catchment. Runoff simulations are strongly influenced by precipitation uncertainty while the ET simulations are much less influenced by precipitation uncertainty, as indicated by the regression slope. The strong relationship between runoff and precipitation is in line with previous studies (e.g. Beck et al., 2017a,b; Sun et al., 2018, Blöschl et al., 2019b). It is related to the fact that most of the considered catchments are located in relatively wet climate (aridity<1) such that soils are often saturated, triggering a rather direct runoff response to precipitation. Also, in these climate regimes ET is typically energy-controlled rather 105 than water-controlled (Pan et al., 2019), leading to the observed low sensitivity of ET to precipitation (uncertainty).

3.2. Climate-dependent propagation of precipitation uncertainty

In addition to examining the role of precipitation uncertainty for runoff and ET across all considered catchments, we analyse this uncertainty propagation within individual hydro-climatological regimes (Fig. 4). For this purpose, we compute the median of the standard deviations from catchments within each regime, considering all respective warm season months. As shown in Fig. 4a, the

110 precipitation uncertainty is higher in comparatively cold and wet regions. This could be related to especially sparse gauge networks and more intense rainfall in these regions which are known to increase precipitation uncertainty (Dinku et al., 2008; Hu et al., 2016; Beck et al., 2017b; O and Kristetter, 2018).

Similarly, Figs. 4b and 4c illustrate the fraction of precipitation uncertainty propagating into runoff and ET, respectively. Interestingly, we find systematic variations in this uncertainty propagation with respect to climate. In wet and cold regions,

115 precipitation uncertainty almost exclusively affects runoff whereas ET remains unchanged; Towards drier and warmer climate the uncertainty propagation shifts, affecting runoff less and increasingly influencing ET.

Figure S1 shows the number of catchments located within each hydro-climatological regime. Only boxes with >5 catchments are considered in the analysis. The uneven distribution of catchments across the regimes induces higher uncertainties in the results obtained for the wettest and driest regimes.

120 As the calibration of the SWBM using E-OBS precipitation data (see Section 2.3) can induce biases in our analyses, we re-compute Figure 4 using model parameters obtained from calibration with GPCC V.2018 precipitation forcing, the results are shown in Figure S2. The clear similarity between Figures 4 and S2 suggests minor relevance of the precipitation dataset used in the SWBM calibration. Further, we repeat the uncertainty propagation analysis using all months instead of focusing on the warm season, also showing similar results (Figure S3).

125 **3.3. Indirect validation of precipitation dataset qualities**

Given the preferential propagation of precipitation uncertainty to runoff in the considered European catchments, we focus in the following on runoff only. In this context, we use streamflow measurements from the catchments to validate the modelled runoff, which allows us to draw conclusions also on the usefulness of the employed precipitation forcing datasets. For the runoff validation, we consider the correlation of monthly anomalies in each catchment and the absolute bias. To obtain anomalies, we remove the

130 mean seasonal cycle from the observed and modelled runoff time series of each catchment. The six runoff simulations in each catchment are then ranked with respect to (i) correlation and (ii) bias, and the sum of these 2 ranks is used to obtain an overall ranking of runoff simulations and corresponding precipitation forcing datasets in each catchment.

Figure 5 shows the number of catchments in which each precipitation product yields the best-ranked runoff simulation. Our findings show that overall the performance of modelled runoff is clearly dependent on the employed precipitation product. This is expected

135 given the considerable disagreement between precipitation products, and the preferential propagation of this uncertainty to runoff (Fig. 4). Generally, runoff computed with E-OBS precipitation agrees best with observations. Also, ERA5, MSWEP V2, and GPCC V.2018 yield comparatively good runoff estimates. In contrast, runoff simulations obtained with ERA-Interim and CFSR agree less well with observations. Repeating this evaluation with all months (Fig. S4) and GPCC-derived SWBM parameters (Fig. S5) largely confirms the described results.

140 Furthermore, we compute runoff performance assessments separately for anomaly correlation and absolute bias (Fig. S6). This reveals that the performance of the precipitation datasets is rather similar in terms of resulting runoff biases. Only ERA5 seems to lead to reduced biases compared with the other products, probably as it does not suffer from gauge-based precipitation undercatch. In contrast, there are considerable differences in terms of the runoff anomaly correlation performance across the products. This

reveals that the differences across products shown in Fig. 5 are mostly resulting from contrasting performance with respect to
145   runoff anomaly correlation.

Figure 6 shows the runoff performance resulting from the various precipitation products for the previously considered hydro-
climatological regimes. We find remarkable performance differences across the regimes, suggesting differential usefulness of
precipitation products for hydrological modelling across different climate zones. Also, we can identify regimes where the
precipitation products perform particularly well or not. For example, MSWEP V2 leads to strong agreement between modelled
150   and observed runoff mostly in comparatively cold and wet climate and less so in warmer and drier regimes. This might be related
to problems of the products incorporated in MSWEP in capturing convective rainfall in warm and dry regions while this is less
problematic in colder regions (Ebert et al., 2007; Beck et al., 2017a,b; Massari et al., 2017; Fallah et al., 2019). The opposite
performance pattern is observed for GPCC V.2018. The lower performance in cold climate, which is also present in the case of E-
OBS, might be related to smaller gauge network density, and more complex topography in colder areas (Ziese et al., 2018). For
155   the other products such as CFSR and ERA-Interim, the performance is less dependent on the hydro-climatological regime.

4. Conclusions

In this study, we investigate how the remarkable discrepancy across state-of-the-art gridded precipitation datasets propagates
through the water cycle. This is essential for hydrological modelling and the applicability of resulting simulations of water balance
160   components such as runoff or ET. Our findings reveal that the uncertainty across precipitation datasets propagates mainly into
runoff rather than ET simulations in Europe. In addition, the partitioning of precipitation uncertainty between runoff and ET is
climate-dependent. In comparatively cold and wet regions such as Europe runoff is more impacted, whereas in drier and warmer
regions the uncertainty partitioning shifts towards ET.

The results in this study are obtained with a single model and are potentially dependent on the choice of that model. Even though
165   this model has been validated thoroughly and applied in previous studies (Orth and Seneviratne, 2014; Orth et al., 2015; Orth and
Seneviratne, 2015, O et al., 2019), future research needs to explore precipitation error propagation with other models (as in Bhuiyan
et al., 2019). This should particularly include distributed models adding to our use of a lumped scheme. However, we do obtain
similar results with different calibrations of this model, while previous research indicated that differences across model calibrations
can be similar to that across models (Tebaldi and Knutti, 2007).

170   The strong link between precipitation and runoff in Europe allowed us to perform an indirect validation of precipitation products
through the performance of the respectively modelled runoff. Overall, the E-OBS precipitation dataset yields the most reliable
streamflow simulations in Europe. Weaker but still comparatively good agreement between modelled and observed streamflow is
obtained with ERA5, GPCC V.2018 and MSWEP V2. Thereby the products differ mostly with respect to the temporal dynamics
rather than the overall amount of precipitation. The interpolated products overall outperform the satellite-derived products in
175   Europe. This is probably due to the high density of gauge observations, as previous research found contrasting conclusions in
regions with low gauge density (e.g. Thiemig et al., 2013 for Africa). Further, we study the precipitation product performance with
respect to climate. We find systematic variations for datasets like MSWEP and GPCC whereas ERA5, ERA-Interim, and CFSR
perform more similarly across climate regimes. Revealing climate-dependent accuracies in some precipitation datasets supports
focused development of these products. This way, innovative hydrological validation of precipitation data, in addition to direct

180   validation against ground truth, can contribute to address the still considerable uncertainty across state-of-the-art gridded products in future efforts.

Further, these findings allow a more targeted combination of products to compensate for individual weaknesses and preserve respective strengths. The climate-dependent (propagation of) precipitation uncertainties illustrates that there is no best overall product but instead a careful regional, climate-based selection can support hydrological applications. Overall, these findings

185   highlight the usefulness of streamflow measurements capturing truly large-scale hydrological dynamics which can even be used to make inference on the accuracy of precipitation datasets (Behrangi et al., 2011; Thiemig et al., 2013; Beck et al., 2017a, 2019a; Arheimer et al., 2019; Bhuiyan et al., 2019; Mazzoleni et al., 2019).

Another important outcome of our analyses is that ET simulations are mostly insensitive to precipitation uncertainty in European climate, confirming previous studies (Bhuiyan et al., 2019). However, in warmer and drier regions such as the Middle East, Central

190   North America or Australia, the link between ET and precipitation should be stronger. Wherever available in these regions, ET measurements can and should be used for indirect evaluation of large-scale precipitation products to complement the results in this study where we focused more on comparatively wet regions.

Moreover, our findings suggest that, across Europe and regions with similar climate, gridded runoff datasets (e.g. Gudmundsson and Seneviratne, 2016) inevitably suffer from the existing uncertainty in state-of-the-art precipitation datasets, although this

195   depends on the extent to which they rely on precipitation data. In contrast, gridded ET products (e.g. Martens et al., 2017, Jung et al., 2019) are not impacted by precipitation uncertainty in these regions. In warmer and drier regions, however, the gridded ET products are more challenged than the runoff products.

Overall, our findings highlight the important role of precipitation accuracy and the understanding of the propagation of existing inaccuracies through the water cycle. Revealing the climate-dependency of this propagation, this study contributes to improved

200   modelling and monitoring of water resources which is of particular relevance in the case of extreme events such as floods and droughts, which might increase in a changing climate.

Competing interests. The authors declare no conflicts of interest.

Acknowledgment. The authors thank Ulrich Weber for preparing the precipitation datasets. Ali Fallah acknowledges financial

205   support from Ministry of Science, Research and technology of the I.R. of Iran, and also the support in the form of hosting and supervision provided by the Max Planck Institute for Biogeochemistry in Jena, Germany. Rene Orth and Sungmin O acknowledge funding support by the German Research Foundation (Emmy Noether grant number 391059971). We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES ([http://ensembles-eu.metoffice.com], accessed 9 December 2019) and the data providers in the ECA&D project ([http://www.ecad.eu], accessed 9 December 2019). Further, we are thankful for streamflow data

210   from a dataset compiled by Stahl et al., 2010, who collected data from the European water archive ([http://www.bafg.de/GRDC/], accessed 9 December 2019), from national ministries and meteorological agencies, as well as from the WATCH project ([http://www.eu-watch.org], accessed 9 December 2019).

References

215 **References**

Abbott, B. W., Bishop, K., Zarnetske, J. P., Minaudo, C., Chapin, F. S., Krause, S., Hannah, D. M., Conner, L., Ellison, D., Godsey, S. E., Plont, S., Marçais, J., Kolbe, T., Huebner, A., Frei, R. J., Hampton, T., Gu, S., Buhman, M., Sara Sayedi, S., Ursache, O., Chapin, M., Henderson, K. D. & Pinay, G.: Human domination of the global water cycle absent from depictions and perceptions. Nature Geoscience, 12(7), 533-540, https://doi.org/10.1038/s41561-019-0374-y, 2019.

220 Alijanian, M., Rakhshandehroo, G. R., Mishra, A. K., & Dehghani, M.: Evaluation of satellite rainfall climatology using CMORPH, PERSIANN-CDR, PERSIANN, TRMM, MSWEP over Iran. International Journal of Climatology, 37(14), 4896-4914, https://doi.org/10.1002/joc.5131, 2017.

Alijanian, M., Reza Rakhshandehroo, G., Mishra, A. & Dehghani, M.: Evaluation of Remotely Sensed Precipitation Estimates using PERSIANN-CDR and MSWEP for Spatio-Temporal Drought Assessment over Iran. Journal of Hydrology, 579,
225 124189 , https://doi.org/10.1016/j.jhydrol.2019.124189, 2019.

Arheimer, B., Pimentel, R., Isberg, K., Crochemore, L., Andersson, J. C. M., Hasan, A., and Pineda, L.: Global catchment modelling using World-Wide HYPE (WWH), open data and stepwise parameter estimation, Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-111, in review, 2019.

Balsamo, G., Agusti-Panareda, A., Albergel, C., Arduini, G., Beljaars, A., Bidlot, J., Blyth, E., Bousserez, N., Boussetta, S., Brown,
230 A., Buizza, R., Buontempo, C., Chevallier, F., Choulga, M., Cloke, H., Cronin, M. F., Dahoui, M., De Rosnay, P., Dirmeyer, P. A., Drusch, M., Dutra, E., Ek, M. B., Gentine, P., Hewitt, H., Keeley, S. P. E., Kerr, Y., Kumar, S., Lupu, C., Mahfouf, J.-F., McNorton, J., Mecklenburg, S., Mogensen, K., Muñoz-Sabater, J., Orth, R., Rabier, F., Reichle, R., Ruston, B., Pappenberger, F., Sandu, I., Seneviratne, S. I., Tietsche, S., Trigo, I. F., Uijlenhoet, R., Wedi, N., Woolway, R. I. & Zeng, X.: Satellite and In Situ Observations for Advancing Global Earth Surface Modelling: A Review. Remote
235 Sensing, 10(12), 2038, https://doi.org/10.3390/rs10122038, 2018.

Beck, H. E., van Dijk, A. I., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & de Roo, A.: MSWEP: 3-hourly 0.25 global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. Hydrology and Earth System Sciences, 21(1), 589, 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,
240 G. J. & Wood, E. F.: Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. Hydrol. Earth Syst. Sci., 21(12), 6201-6217, https://doi.org/10.5194/hess-21-6201-2017, 2017b.

Beck, H. E., Wood, E. F., McVicar, T. R., Zambrano-Bigiarini, M., Alvarez-Garreton, C., Baez-Villanueva, O. M., Sheffield, J. & Karger, D. N.: Bias correction of global high-resolution precipitation climatologies using streamflow observations from 9372 catchments. Journal of Climate, https://doi.org/10.1175/JCLI-D-19-0332.1, 2019a.

245 Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Dijk, A. I. J. M. v., McVicar, T. R. & Adler, R. F.: MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. Bulletin of the American Meteorological Society, 100(3), 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,
250 https://doi.org/10.1016/j.jhydrol.2010.11.043, 2011.

Ehsan Bhuiyan, M. A., 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. Sci., 23, 1973–1994, https://doi.org/10.5194/hess-23-1973-2019, 2019.

255 Blöschl, G., Hall, J., Parajka, J., Perdigão, R. A. P., Merz, B., Arheimer, B., Aronica, G. T., Bilibashi, A., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G. B., Claps, P., Fiala, K., Frolova, N., Gorbachova, L., Gül, A., Hannaford, J., Harrigan, S., Kireeva, M., Kiss, A., Kjeldsen, T. R., Kohnová, S., Koskela, J. J., Ledvinka, O., Macdonald, N., Mavrova-Guirguinova, M., Mediero, L., Merz, R., Molnar, P., Montanari, A., Murphy, C., Osuch, M., Ovcharuk, V., Radevski, I., Rogger, M., Salinas, J. L., Sauquet, E., Šraj, M., Szolgay, J., Viglione, A., Volpi, E., Wilson, D., Zaimi, K. & Živković, N.: Changing climate shifts timing of European floods. Science, 357(6351), 588-590, https://10.1126/science.aan2506, 260 2017.

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. 265 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, 270 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., et al.: Twenty-three unsolved problems in hydrology (UPH) – a community perspective. Hydrological Sciences Journal, 64(10), 1141-1158, https://doi.org/10.1080/02626667.2019.1620507, 2019a.

Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A. P., Parajka, J., Merz, B., Lun, D., Arheimer, B., Aronica, G. T., Bilibashi, A., 275 Boháč, M., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G. B., Claps, P., Frolova, N., Ganora, D., Gorbachova, L., Gül, A., Hannaford, J., Harrigan, S., Kireeva, M., Kiss, A., Kjeldsen, T. R., Kohnová, S., Koskela, J. J., Ledvinka, O., Macdonald, N., Mavrova-Guirguinova, M., Mediero, L., Merz, R., Molnar, P., Montanari, A., Murphy, C., Osuch, M., Ovcharuk, V., Radevski, I., Salinas, J. L., Sauquet, E., Šraj, M., Szolgay, J., Volpi, E., Wilson, D., Zaimi, K. & Živković, N.: . (2019a) Changing climate both increases and decreases European river floods. Nature, 573(7772), 108-280 111, https://doi.org/10.1038/s41586-019-1495-6, 2019b.

Budyko, M. I. (Mikhail Ivanovich) & Miller, David H.: Climate and life (English ed. / edited by David H. Miller). Academic Press, New York, https://trove.nla.gov.au/version/26211710, 1974.

Copernicus Climate change Service (C3S), ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), date of 285 access. https://cds.climate.copernicus.eu/cdsapp#!/home, 2017.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M.,

Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. & Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137(656), 553-597, https://doi.org/10.1002/qj.828, 2011.

Dinku, T., Chidzambwa, S., Ceccato, P., Connor, S. J., & Ropelewski, C. F.: Validation of high-resolution satellite rainfall products over complex terrain. International Journal of Remote Sensing, 29(14), 4097-4110, https://doi.org/10.1080/01431160701772526, 2008.

Ebert, E. E., Janowiak, J. E. & Kidd, C.: Comparison of Near-Real-Time Precipitation Estimates from Satellite Observations and Numerical Models. Bulletin of the American Meteorological Society, 88(1), 47-64, https://doi.org/10.1175/BAMS-88-1-47, 2007.

Fallah, A., Rakhshandehroo, G.R., Berg, P., O, S. and Orth, R.: Evaluation of precipitation datasets against local observations in Southwestern Iran. Int J Climatol., in press, https://doi.org/10.1002/joc.6445, 2019.

Falkenmark, M. & Rockström, J. The new blue and green water paradigm: breaking new ground for water resources planning and management. J. Water Resour. Plan. Manag. 132, 129–132, https://doi.org/10.1061/(ASCE)0733-9496(2006)132:3(129), 2006.

Fereidoon, M., Koch, M. & Brocca, L.: Predicting Rainfall and Runoff Through Satellite Soil Moisture Data and SWAT Modelling for a Poorly Gauged Basin in Iran. Water, 11(3), 594, https://doi.org/10.3390/w11030594, 2019.

Gudmundsson, L. and Seneviratne, S. I.: Observation-based gridded runoff estimates for Europe (E-RUN version 1.1), Earth Syst. Sci. Data, 8, 279–295, https://doi.org/10.5194/essd-8-279-2016, 2016.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D. & New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. Journal of Geophysical Research: Atmospheres, 113(D20), https://doi.org/10.1029/2008JD010201, 2008.

Hu, Z., Hu, Q., Zhang, C., Chen, X., & Li, Q.: Evaluation of reanalysis, spatially interpolated and satellite remotely sensed precipitation data sets in central Asia. Journal of Geophysical Research: Atmospheres, 121(10), 5648-5663. https://doi.org/10.1002/2016JD024781, 2016.

Jones, P. W.: First- and Second-Order Conservative Remapping Schemes for Grids in Spherical Coordinates. Monthly Weather Review, 127(9), 2204-2210, https://doi.org/10.1175/1520-0493(1999)127<2204:FASOCR>2.0.CO;2 , 1999.

Kirtman, B., S.B. Power, J.A. Adedoyin, G.J. Boer, R. Bojariu, I. Camilloni, F.J. Doblas-Reyes, A.M. Fiore, M. Kimoto, G.A. Meehl, M. Prather, A. Sarr, C. Schär, R. Sutton, G.J. van Oldenborgh, G. Vecchi and H.J. Wang,: Near-term Climate Change: Projections and Predictability. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 953–1028, https://doi.org/10.1017/CBO9781107415324.023, 2013.

Koster, R. D. & Mahanama, S. P. P.: Land Surface Controls on Hydroclimatic Means and Variability. Journal of Hydrometeorology, 13(5), 1604-1620, https://doi.org/10.1175/JHM-D-12-050.1, 2012.

Koutsouris, A. J., Chen, D. & Lyon, S. W. Comparing global precipitation data sets in eastern Africa: a case study of Kilombero Valley, Tanzania. International Journal of Climatology, 36(4), 2000-2014, https://doi.org/10.1002/joc.4476, 2016.

Levizzani, V. & Cattani, E.: Satellite Remote Sensing of Precipitation and the Terrestrial Water Cycle in a Changing Climate. Remote Sensing, 11(19), 2301, https://doi.org/10.3390/rs11192301, 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. & Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture. Geosci. Model Dev., 10(5), 1903-1925, https://doi.org/10.5194/gmd-10-1903-2017, 2017.

Massari, C., Crow, W. & Brocca, L.: An assessment of the performance of global rainfall estimates without ground-based observations. Hydrol. Earth Syst. Sci., 21(9), 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 GPM IMERG Early Run with multiple satellite soil moisture products, Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-387, in review, 2019.

Mazzoleni, M., Brandimarte, L. & Amaranto, A.: Evaluating precipitation datasets for large-scale distributed hydrological modelling. Journal of Hydrology, 578, 124076, https://doi.org/10.1016/j.jhydrol.2019.124076, 2019.

Mekonnen, G. B., Matula, S., Doležal, F. & Fišák, J.: Adjustment to rainfall measurement undercatch with a tipping-bucket rain gauge using ground-level manual gauges. Meteorology and Atmospheric Physics, 127(3), 241-256, https://doi.org/10.1007/s00703-014-0355-z, 2015.

Moriasi, D., Arnold, J., Van Liew, M., Bingner, R., Harmel, R., & Veith, T.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE, 50, 885, https://doi.org/10.13031/2013.23153, 2007.

Motovilov, Y. G., Gottschalk, L., Engeland, K. & Rodhe, A.: Validation of a distributed hydrological model against spatial observations. Agricultural and Forest Meteorology, 98-99, 257-277, https://doi.org/10.1016/S0168-1923(99)00102-1, 1999.

Nash, J. E. & Sutcliffe, J. V.: 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, 1970.

Nerini, D., Zulkafli, Z., Wang, L.-P., Onof, C., Buytaert, W., Lavado-Casimiro, W. & Guyot, J.-L.: A Comparative Analysis of TRMM–Rain Gauge Data Merging Techniques at the Daily Time Scale for Distributed Rainfall–Runoff Modeling Applications. Journal of Hydrometeorology, 16(5), 2153-2168, https://doi.org/10.1175/JHM-D-14-0197.1, 2015.

O, S., & Kirstetter, P.-E.: Evaluation of diurnal variation of GPM IMERG-derived summer precipitation over the contiguous US using MRMS data. Quarterly Journal of the Royal Meteorological Society, 144(S1), 270-281, https://doi.org/10.1002/qj.3218, 2018.

O, S., E. Dutra, and R. Orth, Process-based models show comparatively robust performance in changing climatic conditions, J. Hydrometeorol., revisions pending, manuscript available on request.

Oki, T. & Kanae, S.: Global Hydrological Cycles and World Water Resources. Science, 313(5790), 1068-1072, https://doi.org/10.1126/science.1128845, 2006.

360   Orth, R. & Seneviratne, S. I.: Predictability of soil moisture and streamflow on subseasonal timescales: A case study. Journal of Geophysical Research: Atmospheres, 118(19), 10,963-10,979, https://doi.org/10.1002/jgrd.50846, 2013.

Orth, R. & Seneviratne, S. I.: Using soil moisture forecasts for sub-seasonal summer temperature predictions in Europe. Climate Dynamics, 43(12), 3403-3418, https://doi.org/10.1007/s00382-014-2112-x, 2014.

Orth, R., & Seneviratne, S. I.: Introduction of a simple-model-based land surface dataset for Europe Env. Res. Lett., 10, 044012,
365   https://doi.org/10.1088/1748-9326/10/4/044012, 2015.

Orth, R., Staudinger, M., Seneviratne, S. I., Seibert, J. & Zappa, M.: Does model performance improve with complexity? A case study with three hydrological models. Journal of Hydrology, 523, 147-159, https://doi.org/10.1016/j.jhydrol.2015.01.044, 2015.

Orth, R. & Destouni G.: Drought reduces blue-water fluxes more strongly than green-water fluxes in Europe. Nature
370   Communications, 9:3602, https://doi.org/10.1038/s41467-018-06013-7, 2018.

Pan, S., Pan, N., Tian, H., Friedlingstein, P., Sitch, S., Shi, H., Arora, V. K., Haverd, V., Jain, A. K., Kato, E., Lienert, S., Lombardozzi, D., Ottle, C., Poulter, B., and Zaehle, S.: Evaluation of global terrestrial evapotranspiration by state-of-the-art approaches in remote sensing, machine learning, and land surface models, Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-409, in review, 2019.

375   Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-y., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Delst, P. V., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., Dool, H. v. d., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G. & Goldberg, M.: The
380   NCEP Climate Forecast System Reanalysis. Bulletin of the American Meteorological Society, 91(8), 1015-1058, https://doi.org/10.1175/2010bams3001.1, 2010.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y., Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., and Becker, E.: NCEP Climate Forecast System Version 2 (CFSv2) Monthly Products. Retrieved from: https://doi.org/10.5065/D69021ZF, 2012.

385   Satgé, F., Defrance, D., Sultan, B., Bonnet, M.-P., Seyler, F., Rouché, N., Pierron, F. & 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.

Schulzweida, Uwe.: CDO User Guide (Version 1.9.6). Zenodo. http://doi.org/10.5281/zenodo.2558193, 2019.

Stahl, K., Hisdal, H., Hannaford, J., Tallaksen, L. M., van Lanen, H. A. J., Sauquet, E., Demuth, S., Fendekova, M. & Jódar, J.:

390   Streamflow trends in Europe: evidence from a dataset of near-natural catchments. Hydrol. Earth Syst. Sci., 14(12), 2367-2382, https://doi.org/10.5194/hess-14-2367-2010, 2010.

Tebaldi, C. & Knutti, R.: The use of the multi-model ensemble in probabilistic climate projections. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365(1857), 2053-2075, https://doi.org/10.1098/rsta.2007.2076, 2007.

395   Thiemig, V., Rojas, R., Zambrano-Bigiarini, M. & 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.

Whan, K., Zscheischler, J., Orth, R., Shongwe, M., Rahimi, M., Asare, E. O. & Seneviratne, S. I.: Impact of soil moisture on extreme maximum temperatures in Europe. Weather and Climate Extremes, 9, 57-67,

400   https://doi.org/10.1016/j.wace.2015.05.001, 2015.

Ziese, M.; Rauthe-Schöch, A.; Becker, A.; Finger, P.; Meyer-Christoffer, A.; Schneider, U.: GPCC Full Data Daily Version.2018 at 1.0°: Daily Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data. https://doi.org/10.5676/DWD_GPCC/FD_D_V2018_100, 2018.

405

Table 1: Summary of the precipitation datasets evaluated in this study

Group	Dataset	Temporal coverage	Spatial coverage	Spatial resolution	Data sources	Reference
Interpolated	E-OBS	1950-2018	Europe	0.25°	Gauge	Haylock et al., 2008
	GPCC V.2018	1901-2016	Global	1°	Gauge	Ziese et al., 2018
Multi-source	MSWEP V2	1979-NRT[1]	Global	0.1°	Satellite + Gauge + Reanalysis	Beck et al., 2019
Modelled	ERA-Interim	1979-2019	Global	0.5°	Reanalysis	Dee et al., 2011
	ERA5	1950-NRT[2]	Global	~0.28°	Reanalysis	Copernicus Climate change Service, 2017
	CFSR	1979-NRT[1]	Global	0.5°	Reanalysis	Saha et al., 2010, 2012

[1] Near Real-Time product available until the present with a delay of several hours.

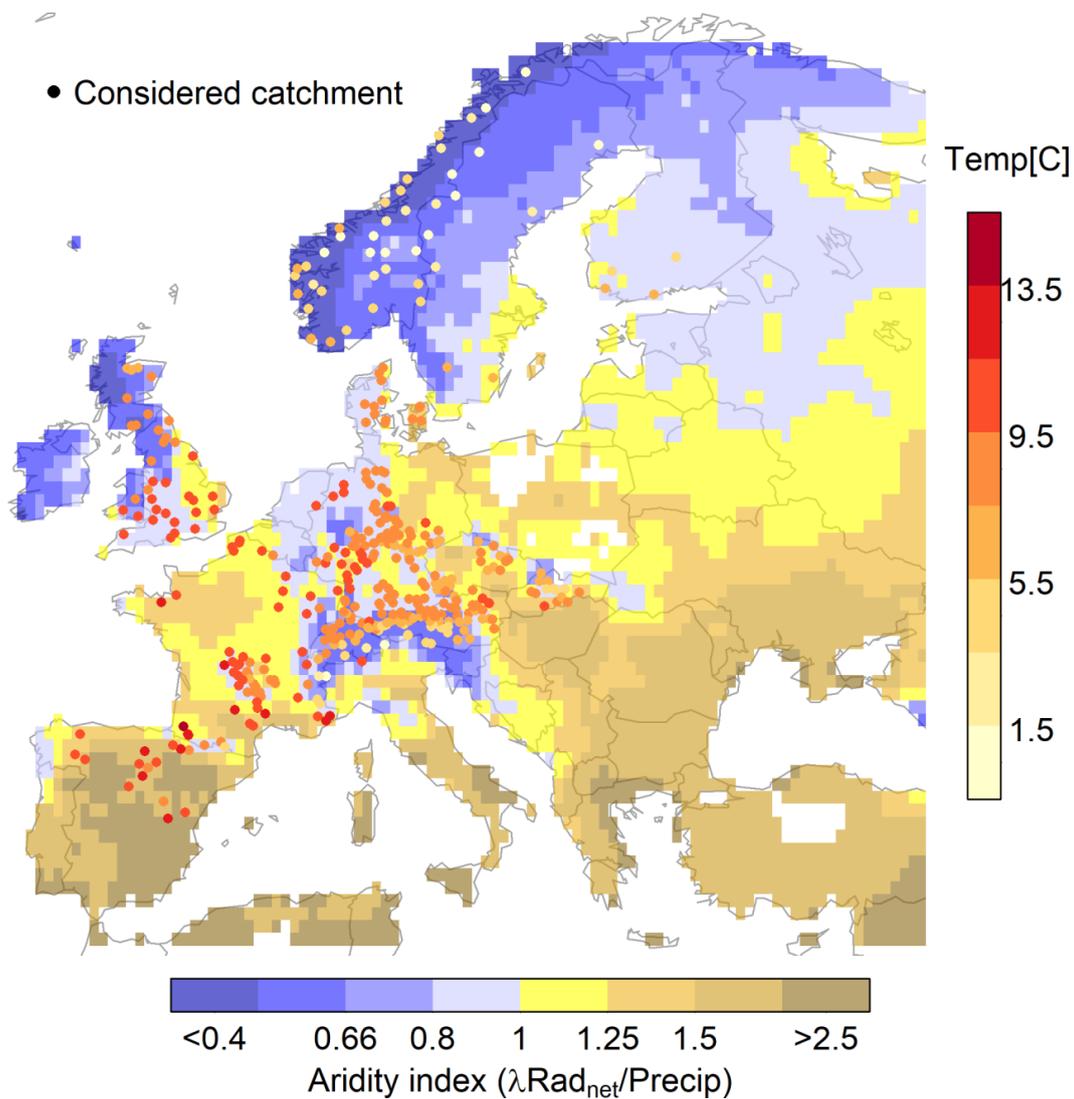
[2] Available until the present with a delay of several months.

410

Figure 1: Map of the study area. Signs mark the position of the 426 study catchments, with color indicating their annual average temperature. Map colors show the aridity index of regions as determined by a ratio of long-term average net radiation and precipitation (1984-2007).

415

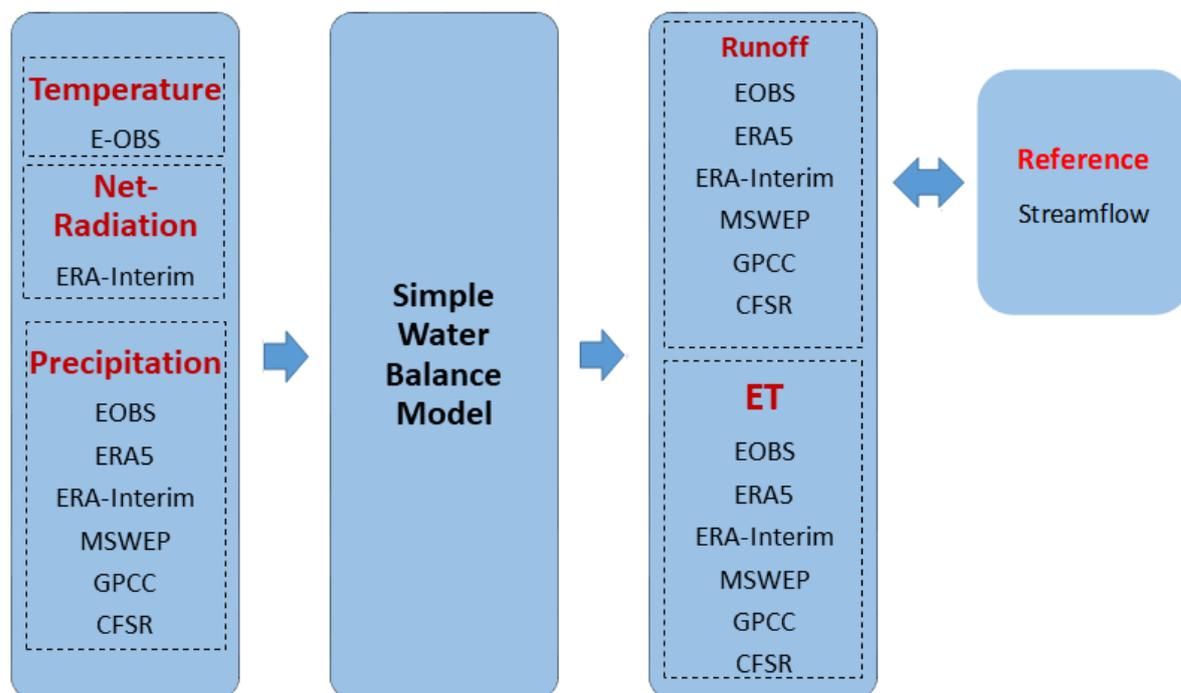



Figure 2: Overview of the modelling approach. The SWBM model is forced with consistent net radiation and temperature data, but six different precipitation datasets. The obtained runoff and evapotranspiration are assessed in terms of the variability between the simulations. The performance of the runoff simulations is determined against streamflow observations.

420

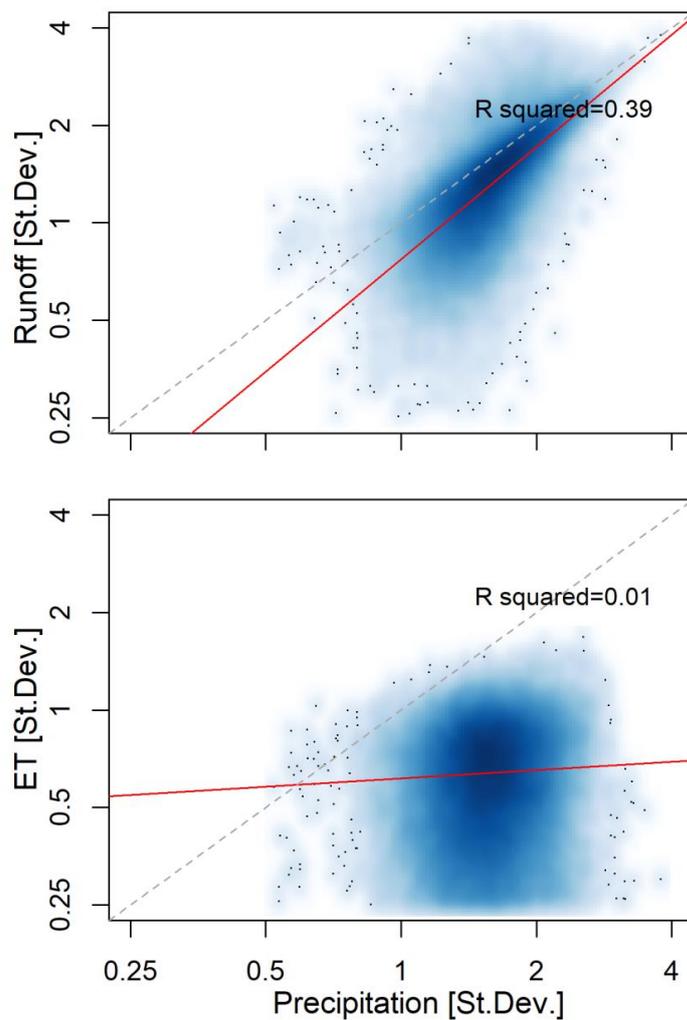



Figure 3: Propagation of precipitation uncertainty into the runoff and ET simulations. Standard deviations are computed across the precipitation estimates and resulting runoff and evapotranspiration values. This is done at every grid cell and every month between May and September. Red lines indicate linear regression lines. Note that a log-log scale is used.

425

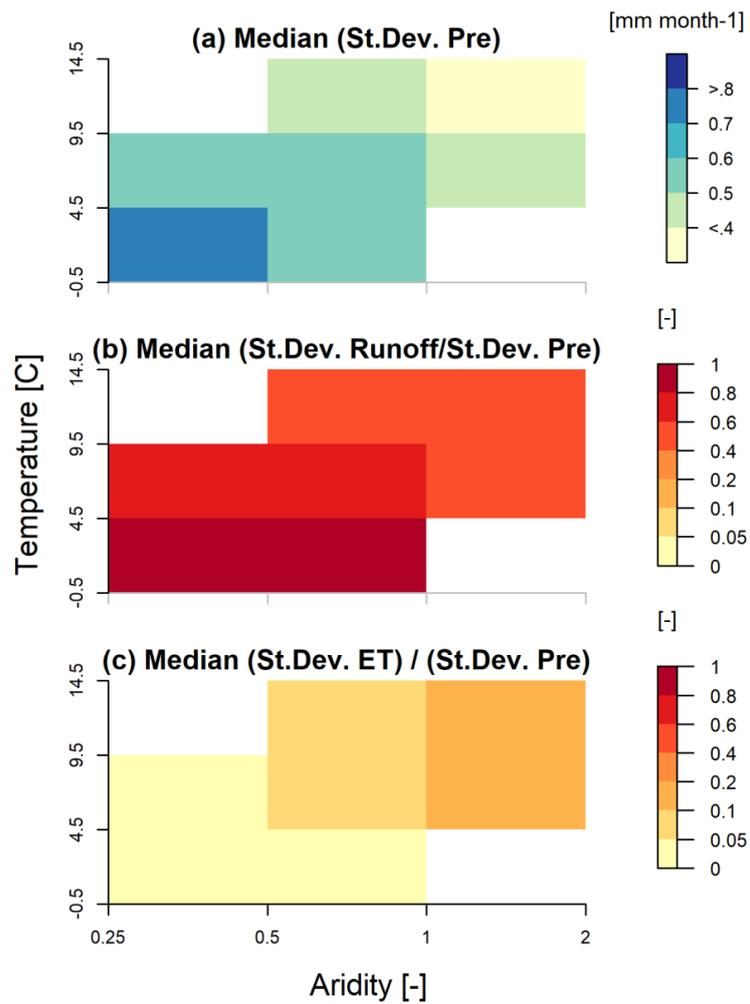



Figure 4: Climate-dependent propagation of precipitation uncertainty into runoff and ET. a) standard deviation across precipitation products, b) and c) relative standard deviation of resulting runoff and ET simulations with respect to that of precipitation, respectively.

430

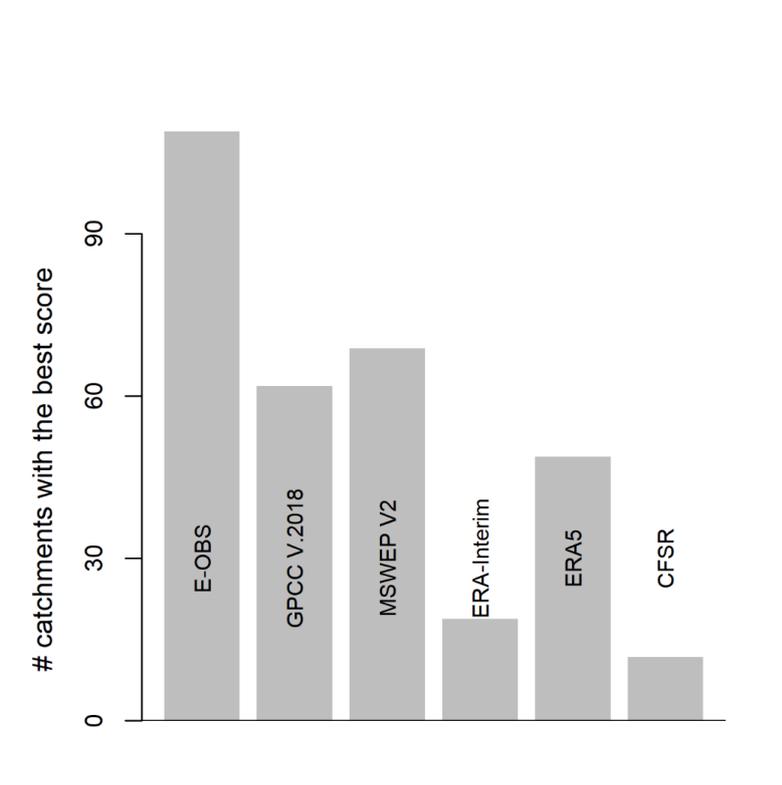



Figure 5: Number of catchments where each precipitation product yields the best agreement with runoff observations (May-September). Multiple data products can be best-performing at a catchment since they are ranked based on a merged score by combining anomaly correlation and absolute error.

435

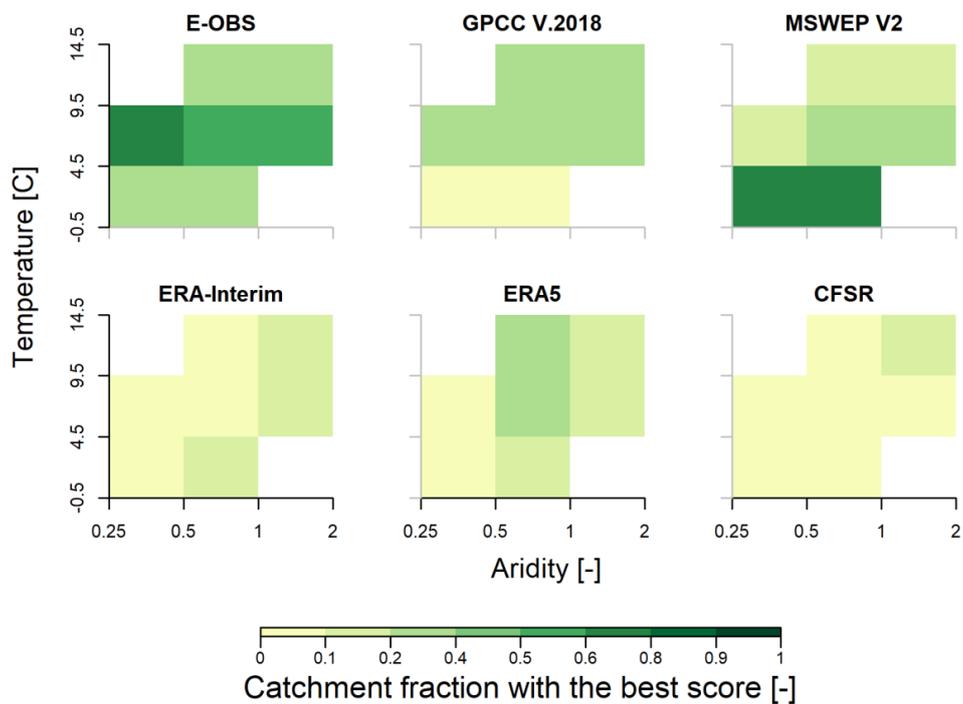



Figure 6: Runoff-based performance of precipitation products across climate regimes. Colors refer to the percentage of catchments within each box recognized as the best performance based on anomaly correlation and absolute bias during May-September.

440