Supplement of Nat. Hazards Earth Syst. Sci., 23, 2111–2131, 2023
https://doi.org/10.5194/nhess-23-2111-2023-supplement
© Author(s) 2023. CC BY 4.0 License.

Natural Hazards and Earth System Sciences
Open Access
EGU

Supplement of

Analyzing the informative value of alternative hazard indicators for monitoring drought hazard for human water supply and river ecosystems at the global scale

Claudia Herbert and Petra Döll
Correspondence to: Claudia Herbert (c.herbert@em.uni-frankfurt.de)

The copyright of individual parts of the supplement might differ from the article licence.

S1 Suitability of WaterGAP to quantify streamflow drought hazard indicators

In several model intercomparison studies, WaterGAP was often among the best performing global hydrological models (GHMs). Kumar et al. (2022) assessed the ability of nine catchment-scale models and eight GHMs to simulate hydrological droughts in eight large catchments around the world. Comparing simulated and observed streamflow deficits and SSI1 (SRI) (their Tables 2 and 3), WaterGAP is among the two to three best performing GHMs with performance indicators (R^2 and Nash-Sutcliffe efficiency) comparable to those of the catchment-scale models. In another assessment of streamflow drought based on observed or simulated streamflow at 293 locations in Europe (Tallaksen and Stahl, 2014), WaterGAP performed well as compared to six other GHMs. In their Fig. 3, WaterGAP results are better than the multi-model median for all four performance measures. Moreover, WaterGAP performed best regarding the simulation of drought persistence (their Fig. 4). Prudhomme et al. (2011) analyzed the ability of three GHMs to reproduce historical streamflow drought events in European basins using the regional deficiency index (RDI). While all three models are found to broadly capture the spatiotemporal drought development, the authors conclude that WaterGAP "is arguably best suited to reproduce most regional characteristics of large-scale high and low flow events in Europe" (Prudhomme et al., 2011: 1181). However, WaterGAP tends to overestimate the variability in RDI, which is explained by insufficient soil storage capacity. In an intercomparison study among six GHMs (Zaherpour et al., 2018), WaterGAP showed the best results in simulating monthly streamflow in 27 out of 40 river basins worldwide and in each of the eight hydrobelts (their Fig. 2 and Table 3). In five out of eight hydrobelts, the mean weighted absolute error of Q95 was lowest for WaterGAP. Nevertheless, the study revealed that WaterGAP tends to overestimate low flows, and that discrepancies between simulated and observed seasonality and interannual variability can be significant. In a different multi-model validation study based on five global hydrological and land surface models (Veldkamp et al., 2018), WaterGAP was the only model that slightly underestimated variability in monthly streamflow while the others overestimated variability. Correlation with observed monthly streamflow was highest for WaterGAP in both managed and near-natural basins across the globe (their Fig. 3h). Döll et al. (2016) compared monthly low-flow Q90 as computed by the GHMs WaterGAP and PCRGLOB-WB to observations at 821 WaterGAP calibration stations across the globe. Overall, low flows could be simulated with reasonable accuracy by both GHMs and were overestimated at most stations. WaterGAP results showed a better fit to observations since it is calibrated against mean annual streamflow at the considered stations (their Fig. 3). Despite calibration, WaterGAP simulations show a lower fit to small observed Q90 values below 1 $km^3\ month^{-1}$.

S2 Validation of the applied WaterGAP model version 2.2d

As a limited validation exercise in the present study, percent deviations of simulated Q80 from observed Q80 values were computed per calendar month (Fig. S1). The latter were based on monthly streamflow observations from the GRDC database (GRDC, 2019). Out of 1319 WaterGAP calibration stations, 220 stations with continuous monthly observations between 1986 and 2015 were assessed, which are mainly distributed over the Northern Hemisphere (U.S., Canada, Europe, Russia). The

analysis reveals that Q80 is overestimated by WaterGAP in 63% of all months and stations and in 53% if only relevant deviations > 10% are considered. The median percent deviation ranges between 35% in February and -7% in July. Figure 3 indicates a tendency of WaterGAP 2.2d to overestimate observed Q80 between October and April, while Q80 during the low-flow period in the Northern Hemisphere (May to September) is better captured.

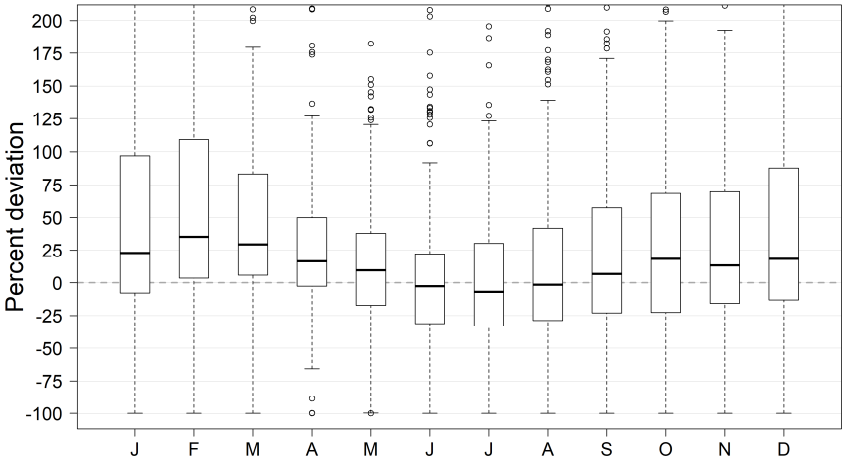

Figure S1: Percent deviations of simulated Q80 per calendar month from Q80 based on GRDC observations using the reference period 1986-2015.

In a recent study, WaterGAP 2.2d model output was validated against GRDC data by comparing SSI3 based on simulated and observed monthly streamflow (SSI3(sim) and SSI3(obs)) during 1971-2000 at 183 globally distributed GRDC stations (Wan et al., 2021). Applying drought hazard classes for SSI according to Agnew (2000), the agreement between simulated and observed hazard classes in each month was analyzed. Among all stations, the agreement ranged between 29 to 88% of all 360 months (their Fig. S4 and Table S3). At 68% of all stations (covering 83% of the assessed basin area), SSI3(sim) and SSI3(obs) resulted in the same drought hazard class in 70 to 88% of the time. Moreover, the goodness-of-fit was evaluated based on the Nash-Sutcliffe efficiency (NSE) for monthly streamflow and SSI3 (their Fig. S3). With a median NSE of 0.5 and an interquartile range of 0.2-0.7 for SSI3 and 0.14-0.7 for streamflow, WaterGAP 2.2d model output showed a moderate agreement with the observations. Both NSEs exceeded 0.7 at 25 out of the 183 stations, which are located in central and eastern Europe (twelve stations), the United States (ten stations), and South Africa (one station).

S3 Observation-based streamflow drought hazard indicators at four selected gauging stations

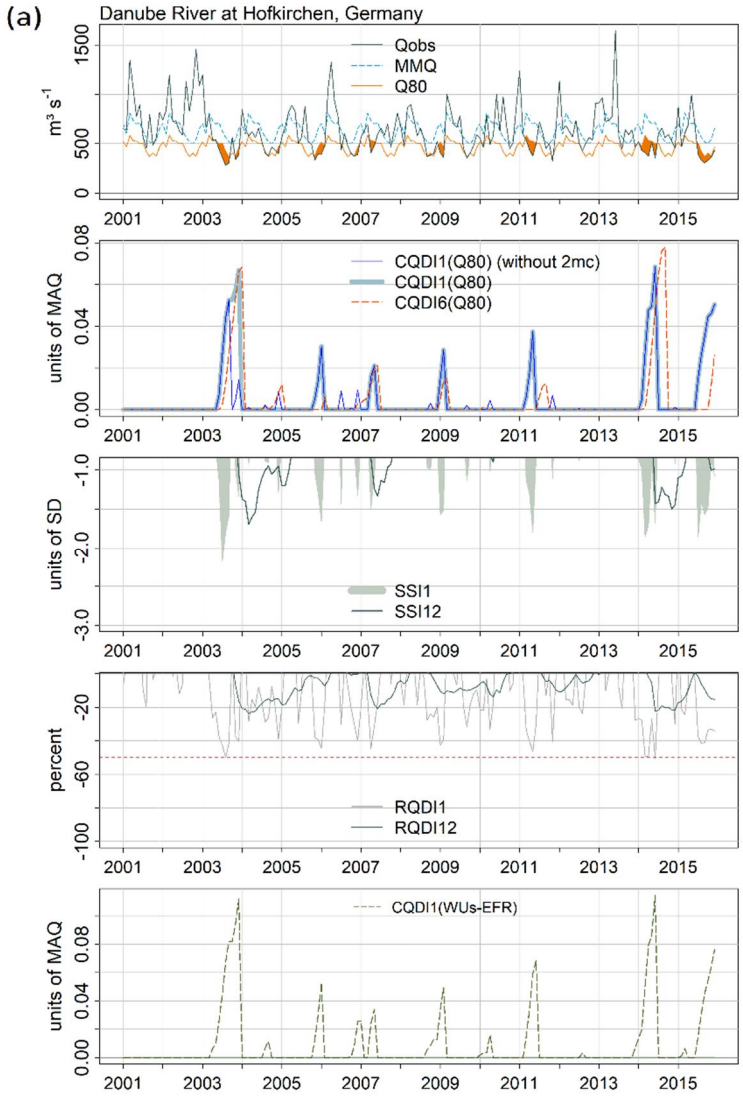

Figure S2: Streamflow drought hazard during 2001-2015 based on observed streamflow during 1986-2015 at four gauging stations in Germany (Danube River, Hofkirchen), Russia (Angara River, Boguchany), U.S. (White River, Oacoma), and South Africa (Orange River, Vioolsdrif): Monthly observed streamflow Qobs, mean monthly streamflow MMQ and Q80 (a), CQDI variants (b), SSI1 and SSI12 (c) RQDI1, RQDI12 (d), and CQDI1-WUs-EFR (e). The cumulative indicators in (b) and (e) indicate drought severity; the non-accumulated indicators (c) and (d) indicate drought magnitude. In (a), periods where Qobs < Q80 are highlighted in orange. In (c), only the value range below -0.84 is shown. In (d), only negative relative deviations are depicted. In (e), EFR is derived using observed instead of naturalized streamflow. 2mc: 2 months criterion. MAQ: mean annual streamflow. SD: standard deviation.

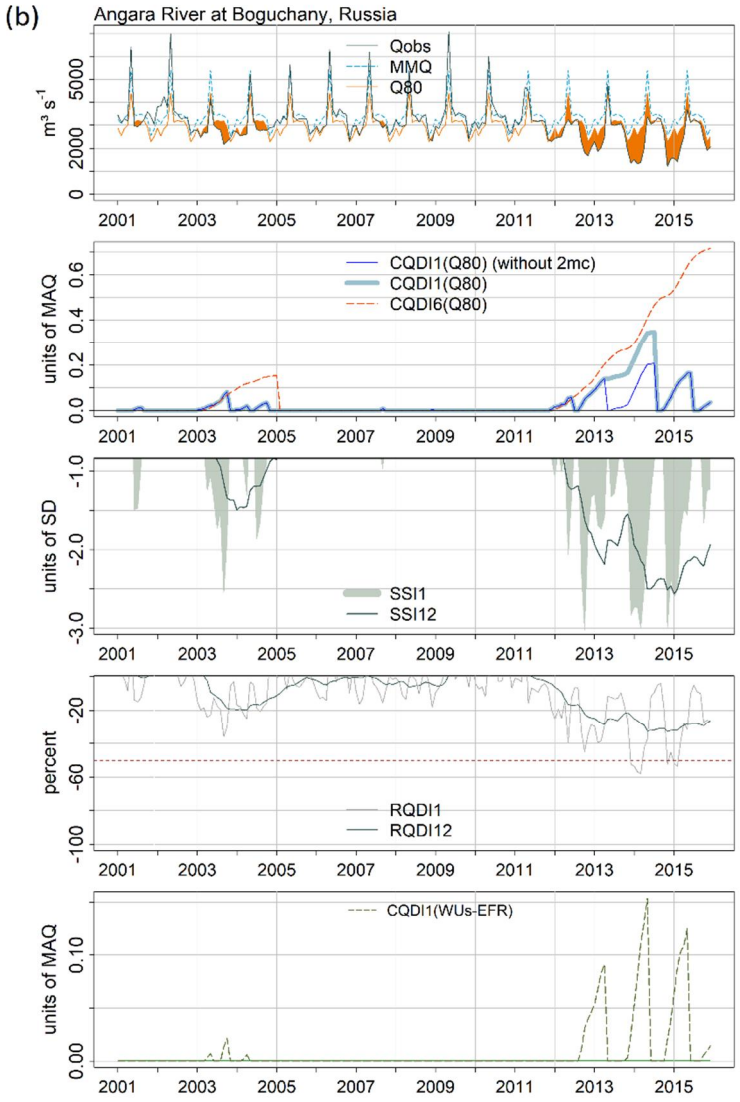

Figure S2: continued.

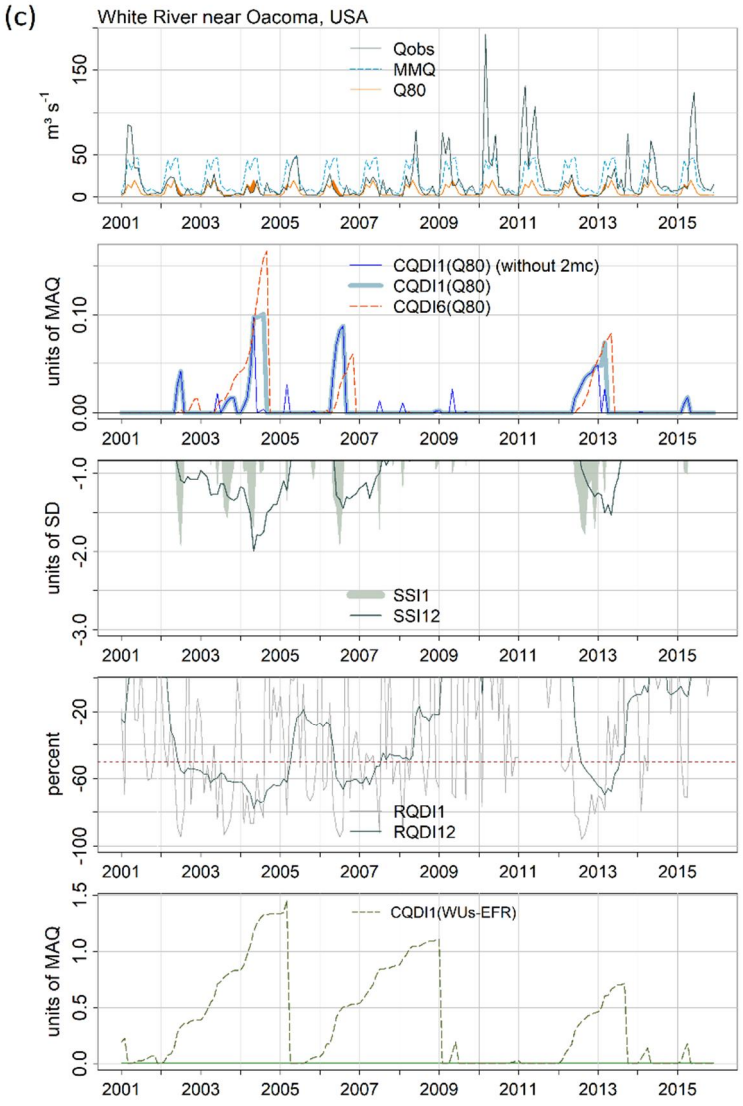

Figure S2: continued.

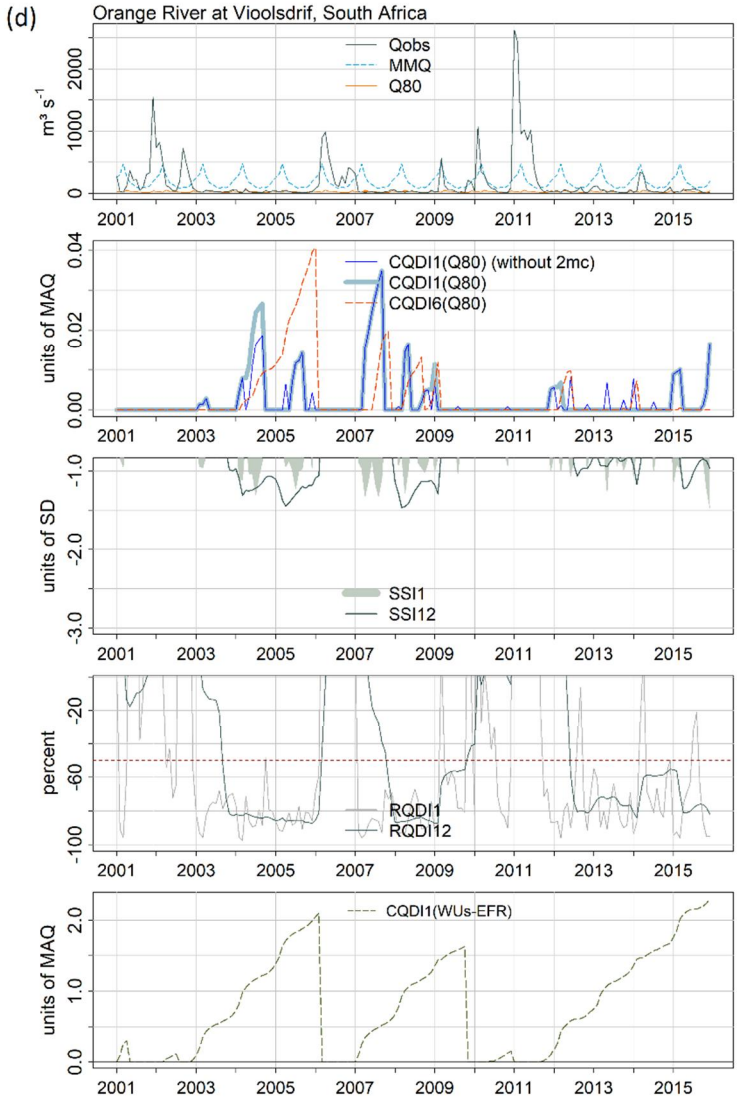

Figure S2: continued.

S4 Simulated SDHIs at the global scale

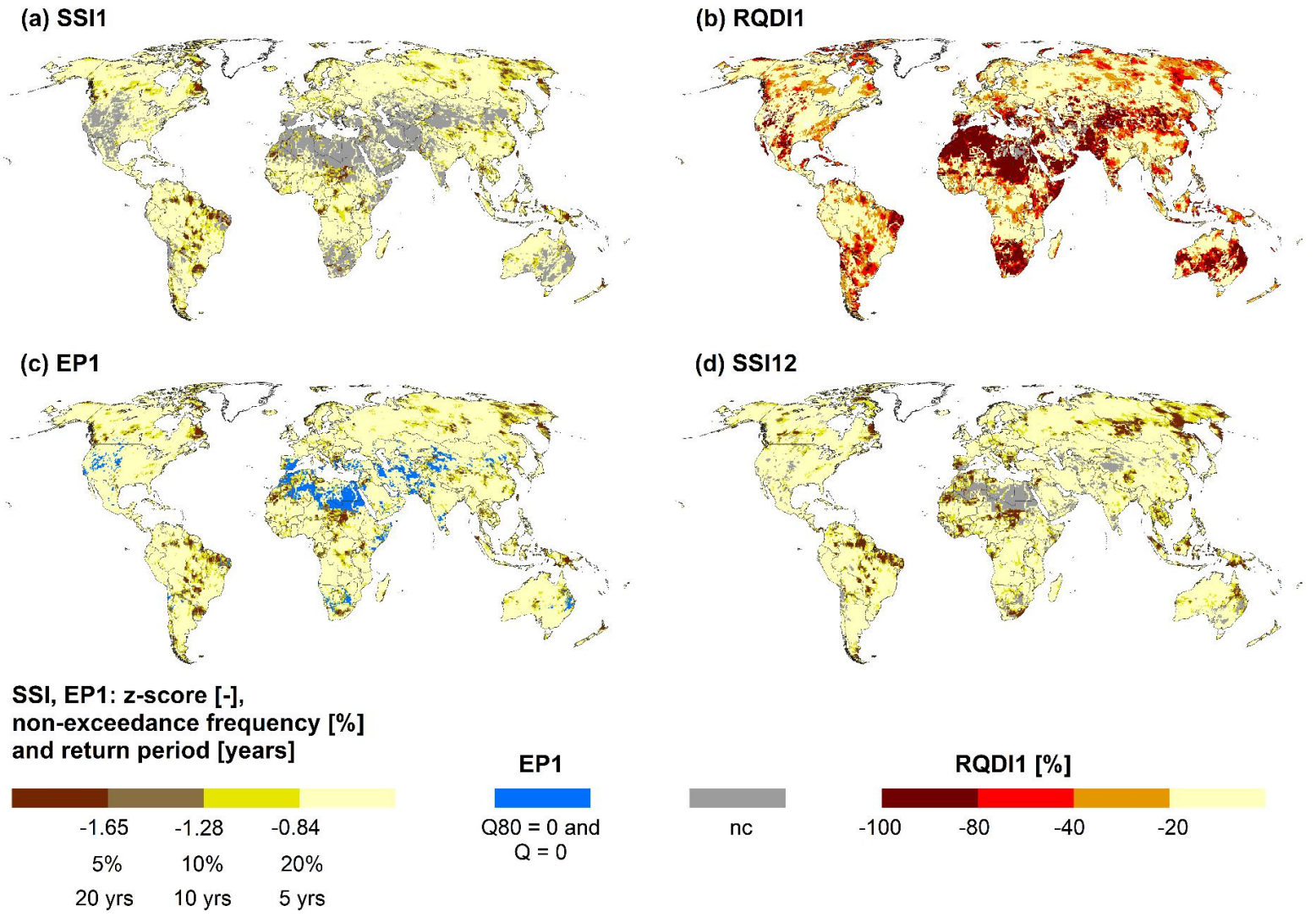

Figure S3: Magnitude of drought hazard (level 1 in Fig. 1): Non-cumulative anomaly in September 1993 as indicated by SSI1 (a), RQDI1 (b), EP1 (c), and SSI12 (d) for the reference period 1986-2015. For the standardized indicators and EP1, the z scores and the corresponding frequencies of non-exceedance and return periods are shown. In the blue grid cells in (c), drought identification is not possible with EP1, since Q80 and Q are zero. "nc": not computable.

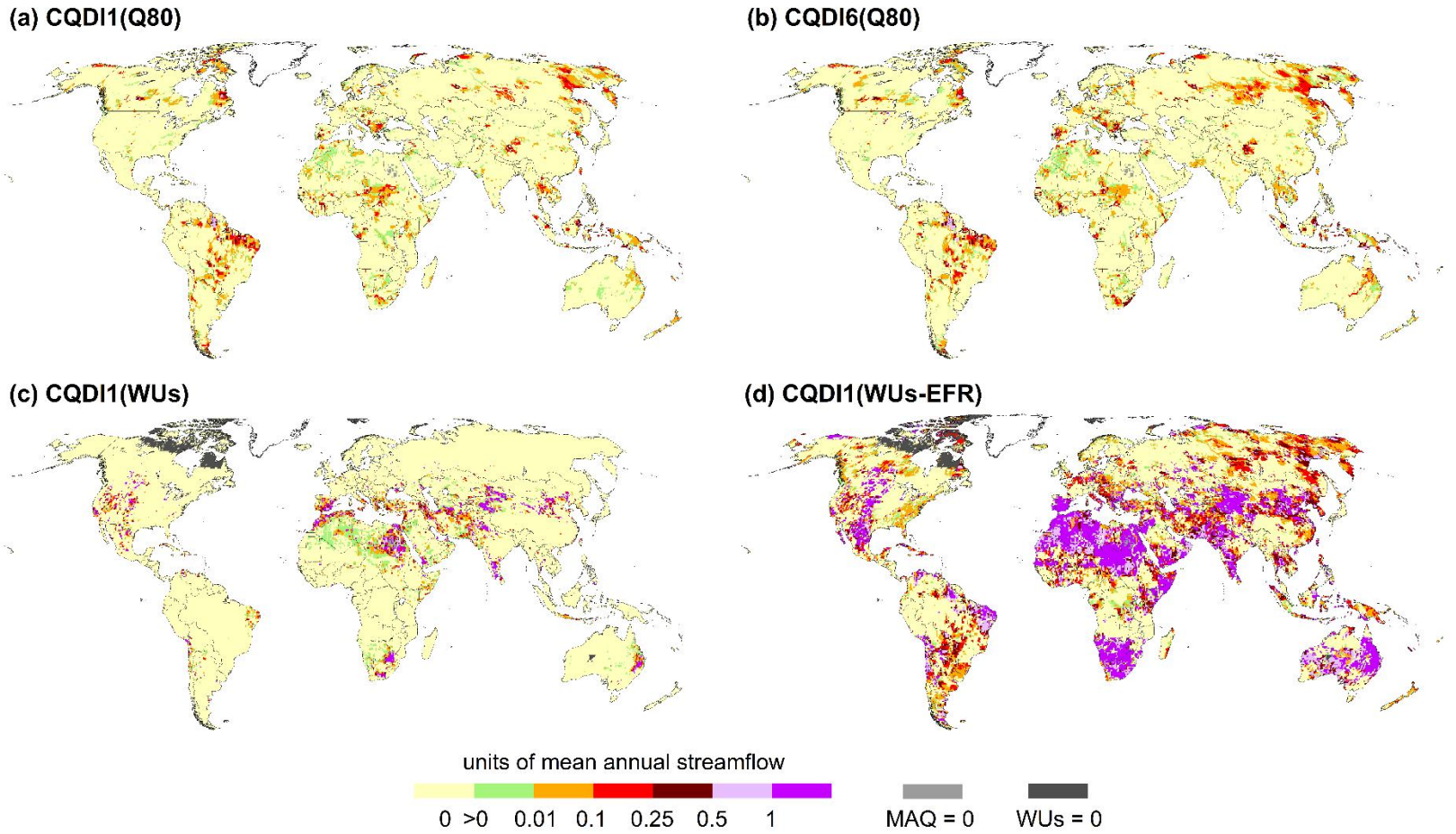

Figure S4: Severity of drought hazard (level 2 in Fig. 1): Cumulative deficit in September 1993 since onset of drought event as indicated by CQDI1(Q80) (a), CQDI6(Q80) (b), CQDI1(WUs) (c), and CQDI1(WUs-EFR) (d) for the reference period 1986-2015. Grid cells with a deficit of zero are shown in beige. Values larger than zero and below 0.1 are shown in green. A value of 0.1, for example, denotes that the current cumulative deficit is equivalent to 10% of mean annual streamflow (MAQ). WUs: mean annual surface water withdrawals.

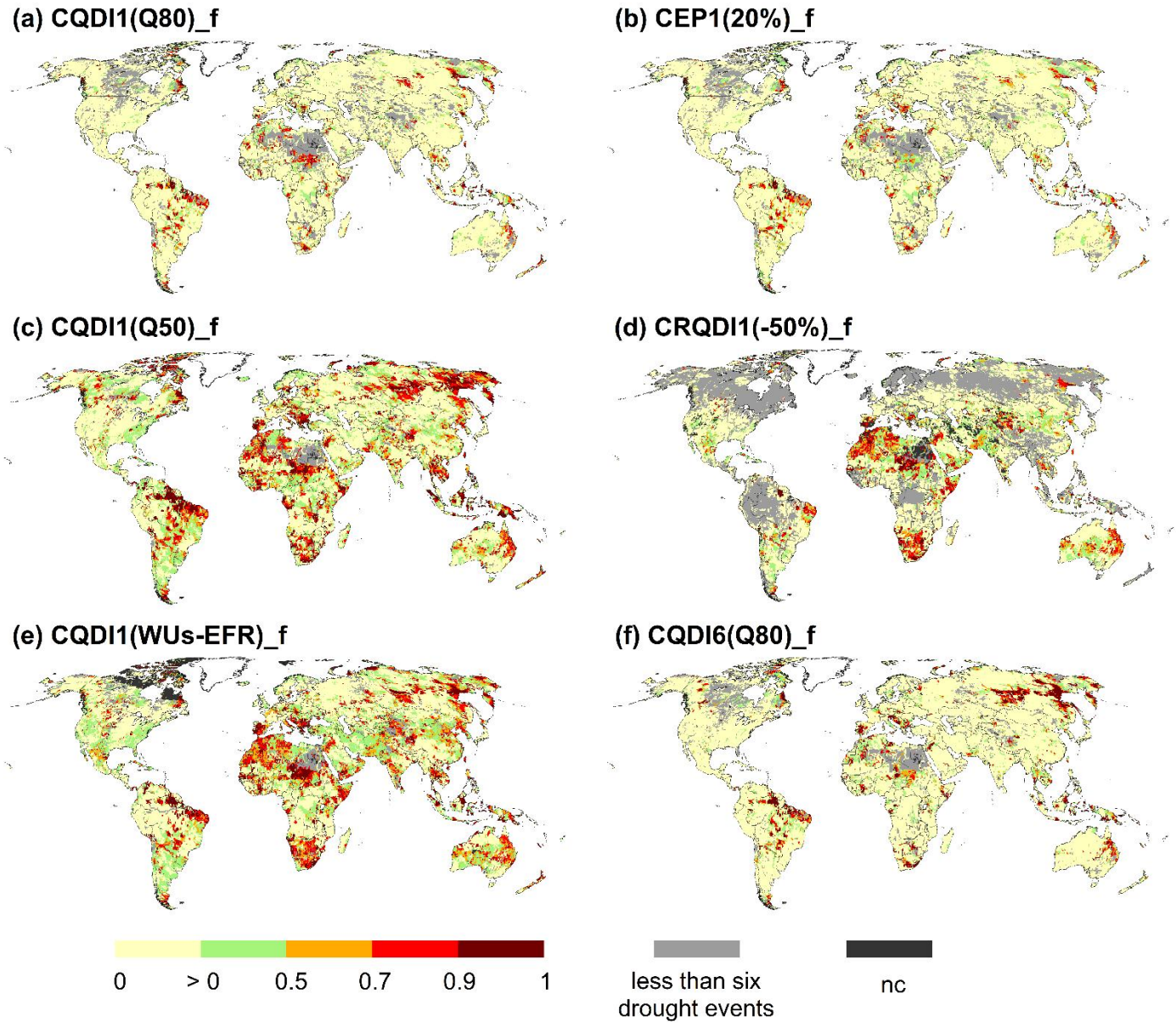

Figure S5: Probability of non-exceedance of drought events (level 2 in Fig. 1) in September 1993 for the cumulative indicators CQDI1(Q80)_f (a), CEP1(20%)_f (b), CQDI1(Q50)_f (c), CRQDI1(-50%)_f (d), CQDI1(WUs-EFR)_f (e), and CQDI6(Q80)_f (f) for the reference period 1986-2015. A value of 0.8, for example, indicates that the cumulative anomaly or deficit, i.e., the severity up to this month, is higher than the severity of 80% of all drought events in the reference period. The probability of non-exceedance was not computed for grid cells shown in light grey, where less than six drought events were computed in the reference period (Sect. 2.3). "nc": not computable.

S5 Comparison of CQDI1(Q80) and CEP1(20%) at three selected gauging stations

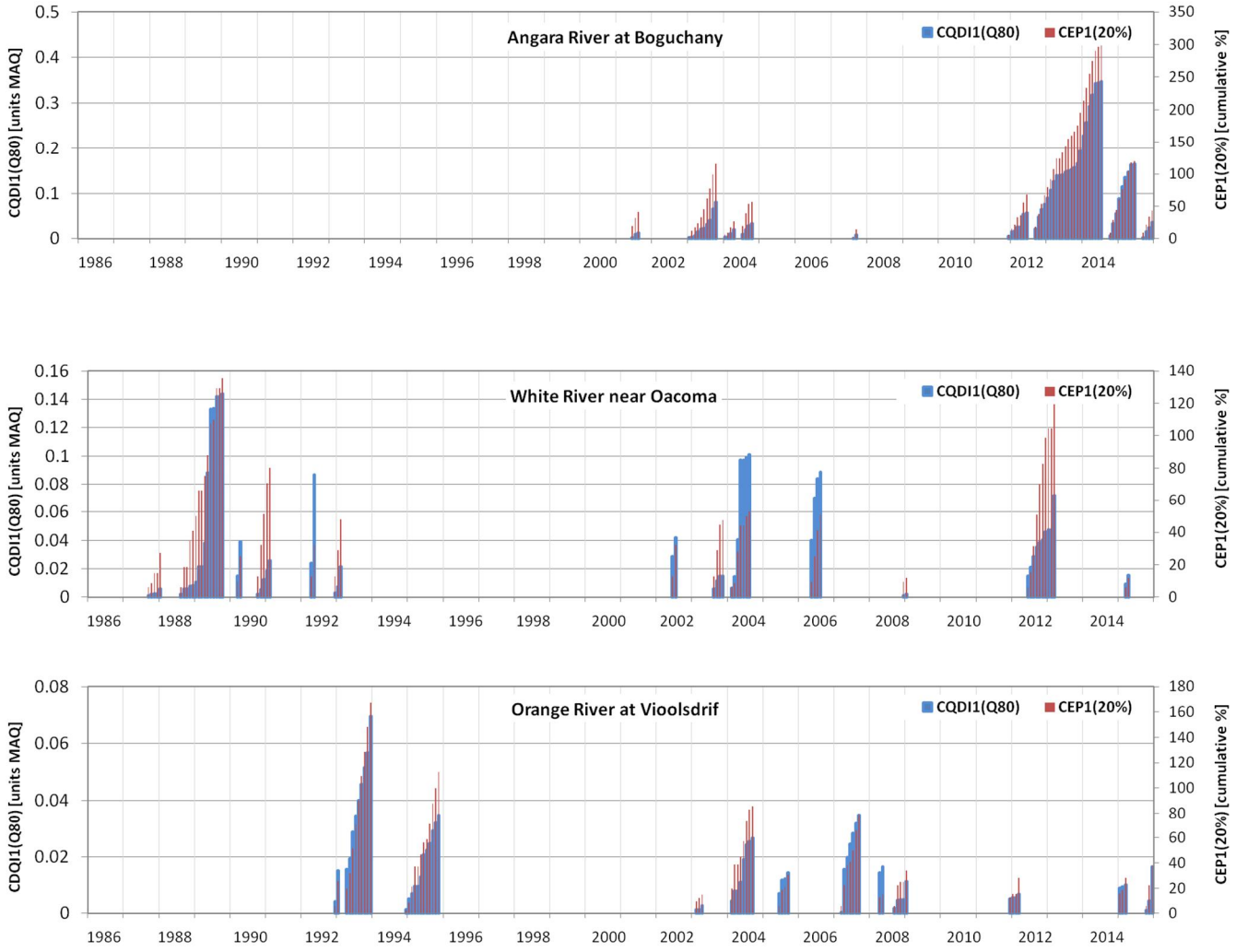

Figure S6: Drought severity per month during the reference period 1986-2015 at three selected gauging stations in Russia (Angara River, Boguchany), U.S. (White River, Oacoma), and South Africa (Orange River, Vioolsdrif), as indicated by CQDI1(Q80) (blue) and CEP1(20%) (red). MAQ: mean annual streamflow.

References

Agnew, C. T.: Using the SPI to Identify Drought, 2000.

Döll, P., Douville, H., Güntner, A., Müller Schmied, H., and Wada, Y.: Modelling Freshwater Resources at the Global Scale: Challenges and Prospects, Surv Geophys, 37, 195–221, https://doi.org/10.1007/s10712-015-9343-1, 2016.

GRDC: Global Runoff Data Centre, Federal Institute of Hydrology, Koblenz, Germany, 2019.

Kumar, A., Gosling, S. N., Johnson, M. F., Jones, M. D., Zaherpour, J., Kumar, R., Leng, G., Schmied, H. M., Kupzig, J., Breuer, L., Hanasaki, N., Tang, Q., Ostberg, S., Stacke, T., Pokhrel, Y., Wada, Y., and Masaki, Y.: Multi-model evaluation of catchment- and global-scale hydrological model simulations of drought characteristics across eight large river catchments, Advances in Water Resources, 165, 104212, https://doi.org/10.1016/j.advwatres.2022.104212, 2022.

Prudhomme, C., Parry, S., Hannaford, J., Clark, D. B., Hagemann, S., and Voss, F.: How Well Do Large-Scale Models Reproduce Regional Hydrological Extremes in Europe?, Journal of Hydrometeorology, 12, 1181–1204, https://doi.org/10.1175/2011JHM1387.1, 2011.

Tallaksen, L. M. and Stahl, K.: Spatial and temporal patterns of large-scale droughts in Europe: Model dispersion and performance, Geophys. Res. Lett., 41, 429–434, https://doi.org/10.1002/2013GL058573, 2014.

Veldkamp, T. I. E., Zhao, F., Ward, P. J., Moel, H. de, Aerts, J. C. J. H., Schmied, H. M., Portmann, F. T., Masaki, Y., Pokhrel, Y., Liu, X., Satoh, Y., Gerten, D., Gosling, S. N., Zaherpour, J., and Wada, Y.: Human impact parameterizations in global hydrological models improve estimates of monthly discharges and hydrological extremes: a multi-model validation study, Environ. Res. Lett., 13, 55008, https://doi.org/10.1088/1748-9326/aab96f, 2018.

Wan, W., Zhao, J., Popat, E., Herbert, C., and Döll, P.: Analyzing the Impact of Streamflow Drought on Hydroelectricity Production: A Global-Scale Study, Water Res, 57, https://doi.org/10.1029/2020WR028087, 2021.

Zaherpour, J., Gosling, S. N., Mount, N., Schmied, H. M., Veldkamp, T. I. E., Dankers, R., Eisner, S., Gerten, D., Gudmundsson, L., Haddeland, I., Hanasaki, N., Kim, H., Leng, G., Liu, J., Masaki, Y., Oki, T., Pokhrel, Y., Satoh, Y., Schewe, J., and Wada, Y.: Worldwide evaluation of mean and extreme runoff from six global-scale hydrological models that account for human impacts, Environ. Res. Lett., 13, 65015, https://doi.org/10.1088/1748-9326/aac547, 2018.