Supplement of Hydrol. Earth Syst. Sci., 27, 1809–1825, 2023
https://doi.org/10.5194/hess-27-1809-2023-supplement

Hydrology and Earth System Sciences
Open Access

Supplement of

Benchmarking high-resolution hydrologic model performance of long-term retrospective streamflow simulations in the contiguous United States

Erin Towler et al.

Correspondence to: Erin Towler (towler@ucar.edu)

Supplement

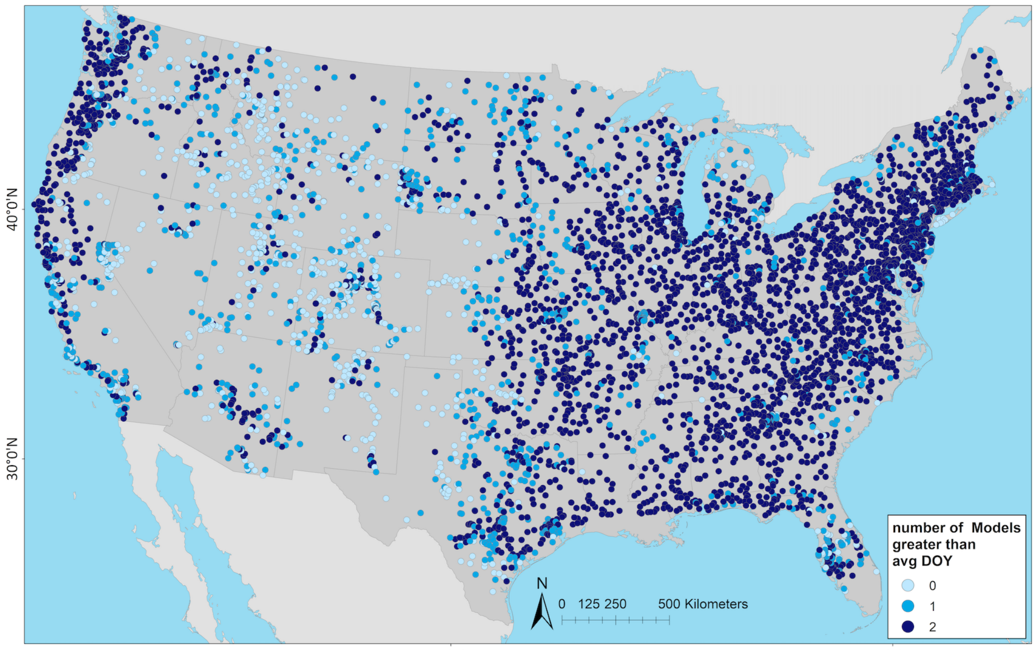


Figure S1. For the National Water Model v2.1 (NWMv2.1) and the National Hydrologic Model v1.0 (NHMv1.0), the number of models where the KGE value is greater than the AvgDOY; both models are better (n=3396), one model is better (n = 1083), or neither model is better (n=911). Map Source: (Grannemann, 2010; Natural Earth Data, 2009; ESRI, 2022a; ESRI, 2022b).

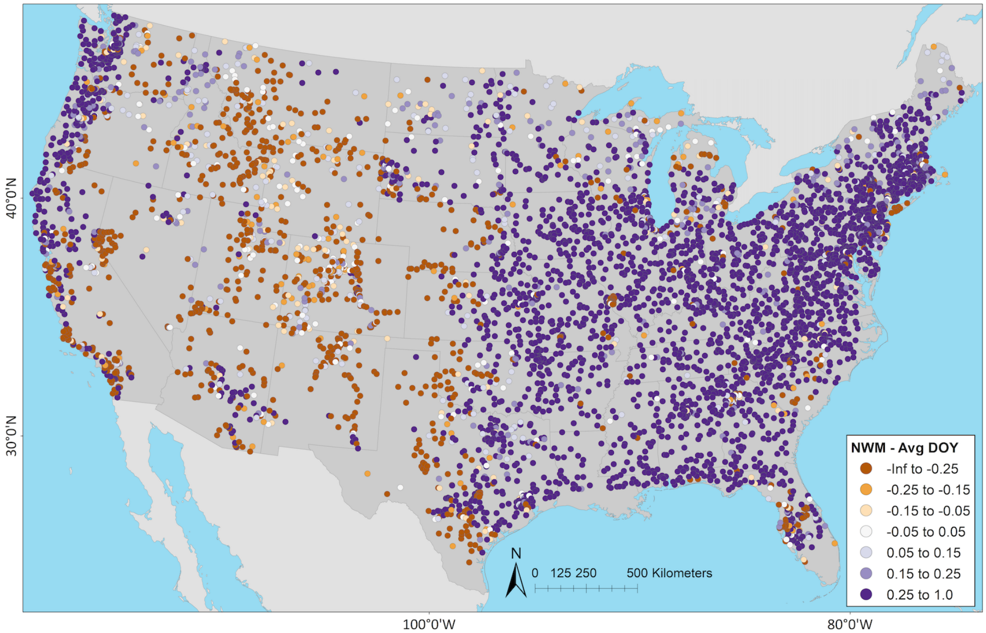


Figure S2. Difference between the Kling–Gupta efficiency (KGE) from the National Water Model v2.1 (NWMv2.1) and the seasonal benchmark based on the average day-of-year flows (AvgDOY); negative (orange) indicates where AvgDOY has a higher (better) KGE, positive (purple) indicates that the NWMv2.1 has a higher (better) KGE. Map Source: (Grannemann, 2010; Natural Earth Data, 2009; ESRI, 2022a; ESRI, 2022b).

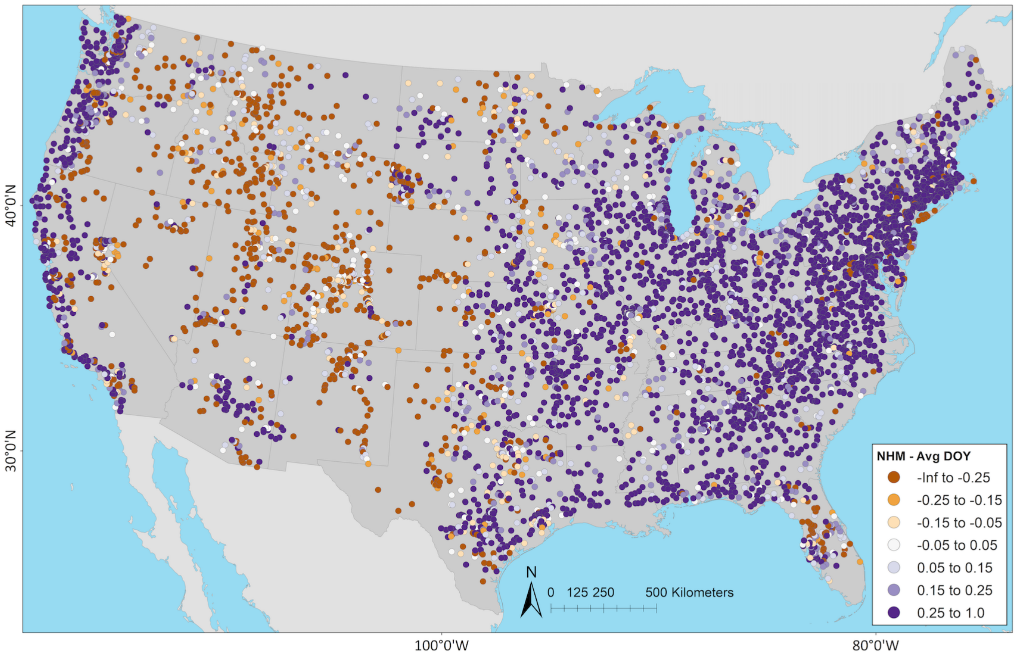


Figure S3. Difference between the Kling–Gupta efficiency (KGE) from the National Hydrologic Model v1.0 (NHMv1.0) and the seasonal benchmark based on the average day-of-year flows (AvgDOY); negative (orange) indicates where AvgDOY has a higher (better) KGE, positive (purple) indicates that the NHMv1.0 has a higher (better) KGE. Map Source: (Grannemann, 2010; Natural Earth Data, 2009; ESRI, 2022a; ESRI, 2022b).

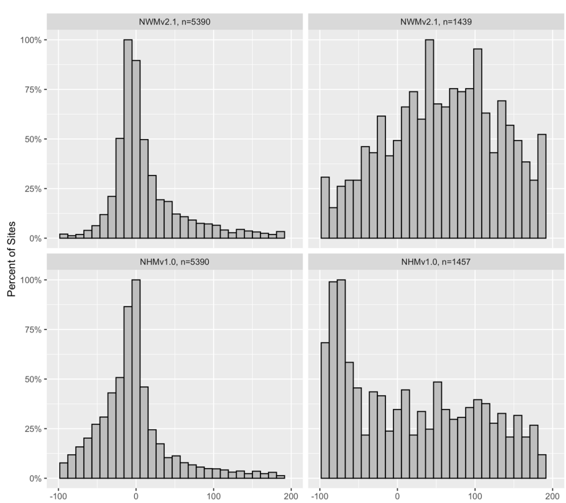


Figure S4: Normalized histograms of PBIAS for National Water Model v2.1 (NWMv2.1, top) and National Hydrologic Model v1.0 (NHMv1.0, bottom), for all sites (left) and for sites where the model's KGE score is less than the average day-of-year flow benchmark (right).

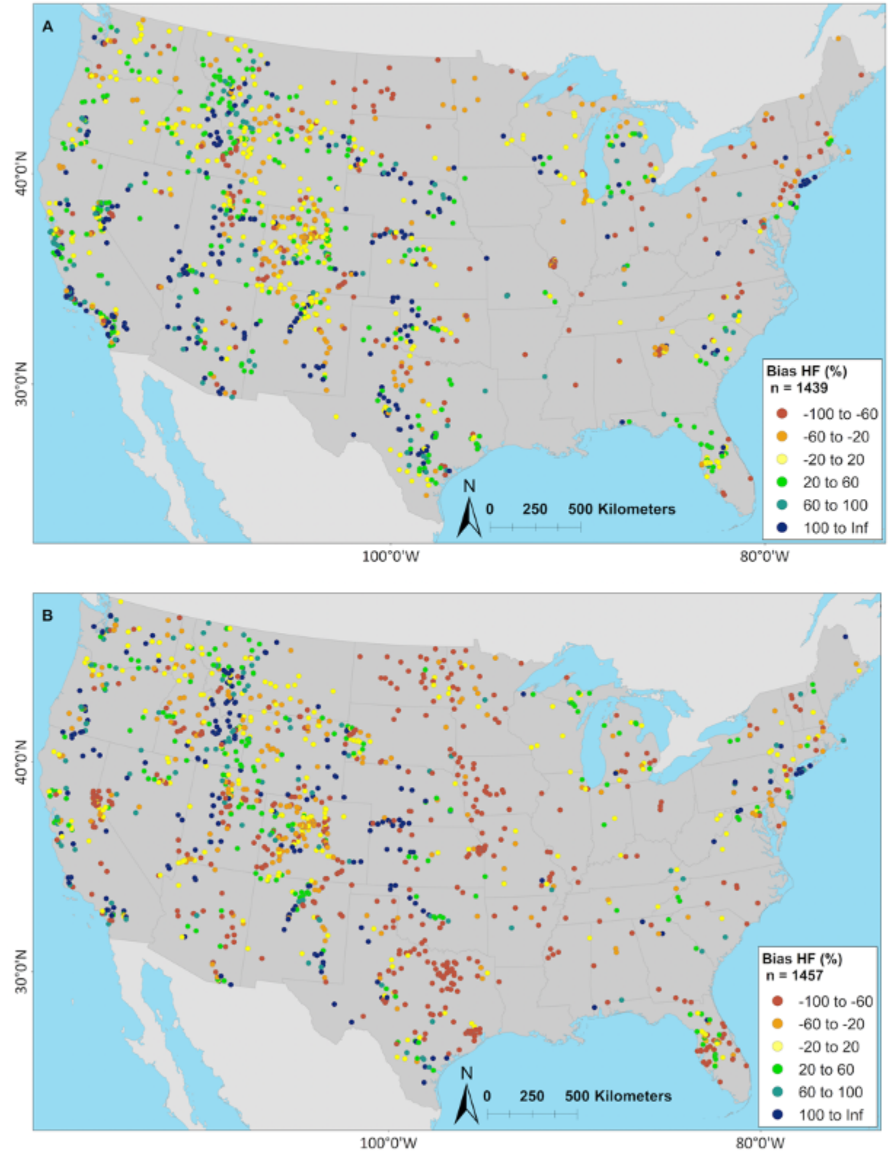


Figure S5: Percent bias of high flow (PBIAS_HF; i.e., exceeding top 2%) maps for National Water Model v2.1 (NWMv2.1) (A) and National Hydrologic Model v1.0 (NHMv1.0) (B), for sites where the KGE score is less than the average day-of-year flow (AvgDOY) benchmark. Cooler colors are where model application is overestimating high flow bias and warmer colors are where model is underestimating high flow bias. Map Source: (Grannemann, 2010; Natural Earth Data, 2009; ESRI, 2022a; ESRI, 2022b).

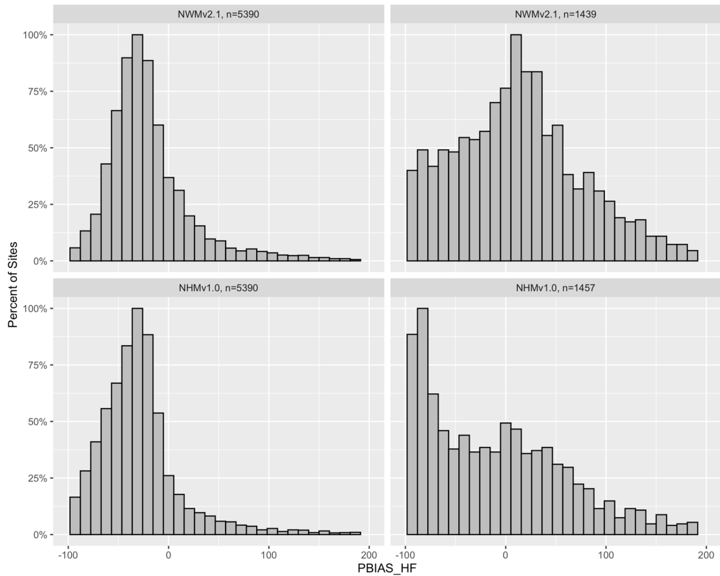


Figure S6: Normalized histograms of Percent bias of high flow (PBIAS_HF; i.e., exceeding top 2%) for National Water Model v2.1 (NWMv2.1, top) and National Hydrologic Model v1.0 (NHMv1.0, bottom), for all sites (left) and for sites where the model's KGE score is less than the average day-of-year flow benchmark (right).

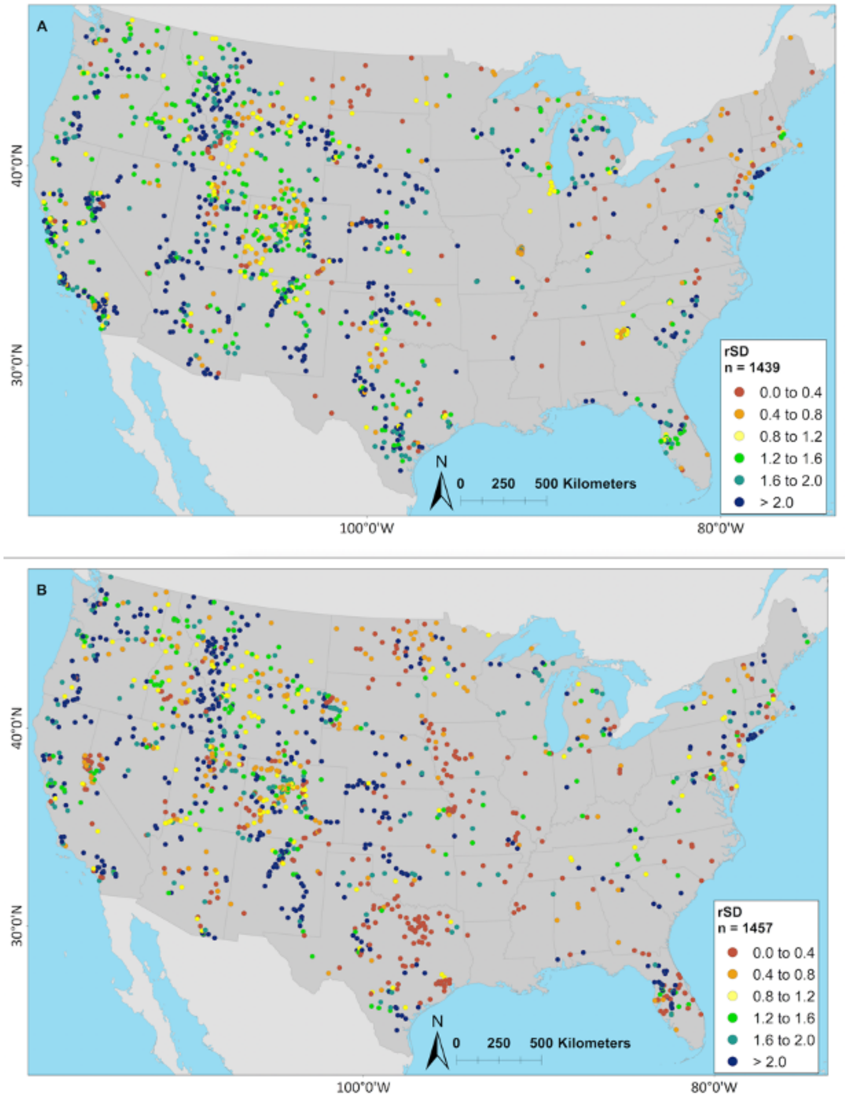


Figure S7: ratio of standard deviation (rSD) maps for National Water Model v2.1 (NWMv2.1) (A) and National Hydrologic Model v1.0 (NHMv1.0) (B), for sites where the KGE score is less than the average day-of-year flow (AvgDOY) benchmark. Cooler colors are where model application is overestimating variability and warmer colors are where model is underestimating variability. Map Source: (Grannemann, 2010; Natural Earth Data, 2009; ESRI, 2022a; ESRI, 2022b).

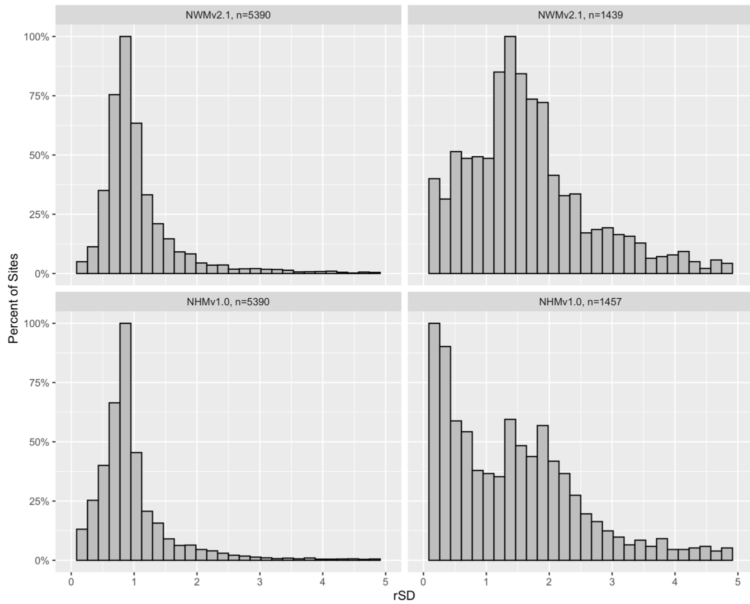


Figure S8: Normalized histograms of standard deviation ratio (rSD) for National Water Model v2.1 (NWMv2.1, top) and National Hydrologic Model v1.0 (NHMv1.0, bottom), for all sites (left) and for sites where the model's KGE score is less than the average day-of-year flow benchmark (right).

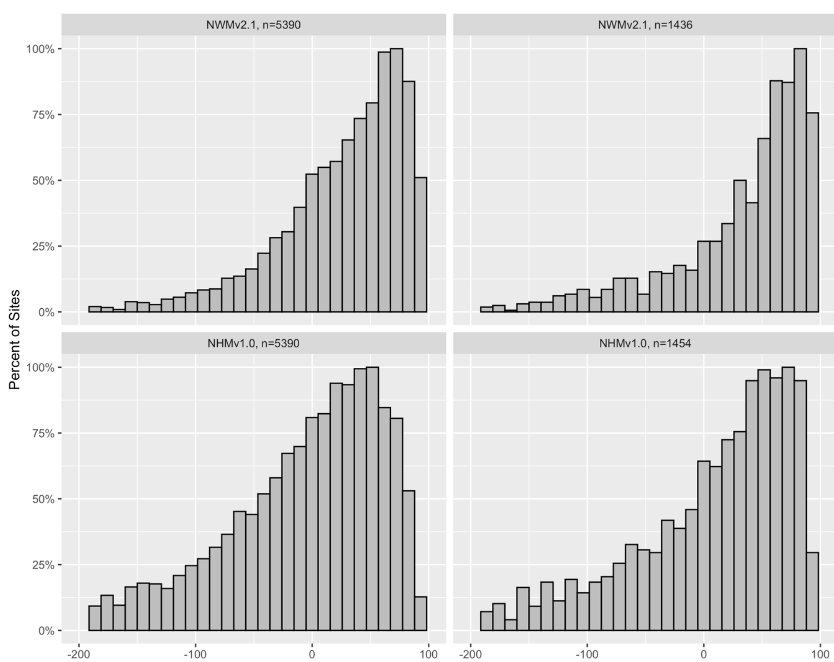
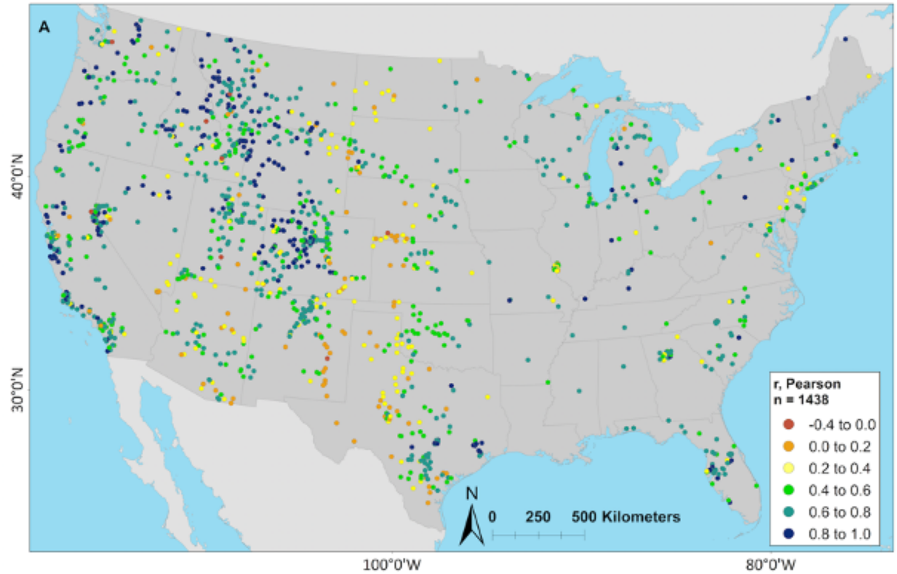


Figure S9: Normalized histograms of percent bias of low flow (PBIAS_LF, flows below 30% percentile) for National Water Model v2.1 (NWMv2.1, top) and National Hydrologic Model v1.0 (NHMv1.0, bottom), for all sites (left) and for sites where the model's KGE score is less than the average day-of-year flow benchmark (right).



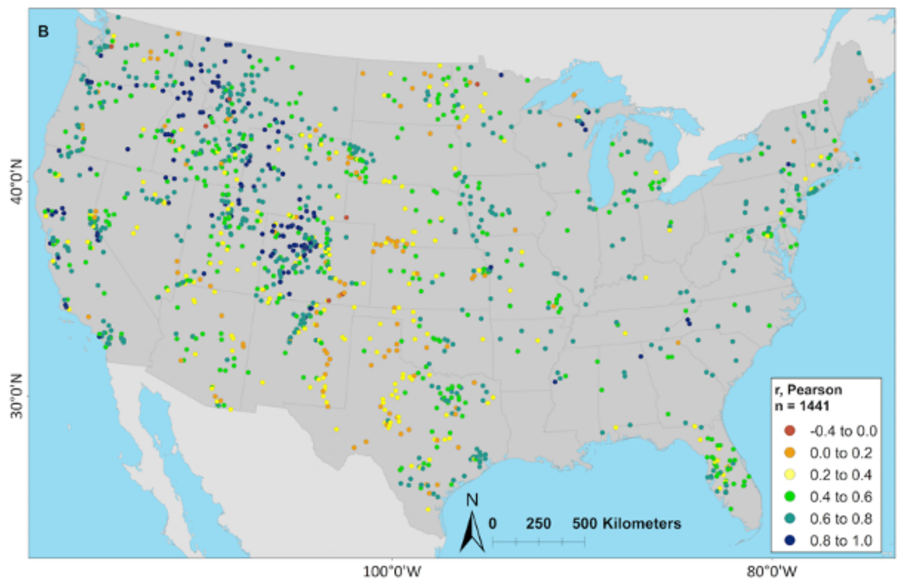


Figure S10: Pearson's correlation coefficient (r) for National Water Model v2.1 (NWMv2.1) (A) and National Hydrologic Model v1.0 (NHMv1.0) (B), for sites where the KGE score is less than the average day-of-year flow (AvgDOY) benchmark. Map Source: (Grannemann, 2010; Natural Earth Data, 2009; ESRI, 2022a; ESRI, 2022b).

75

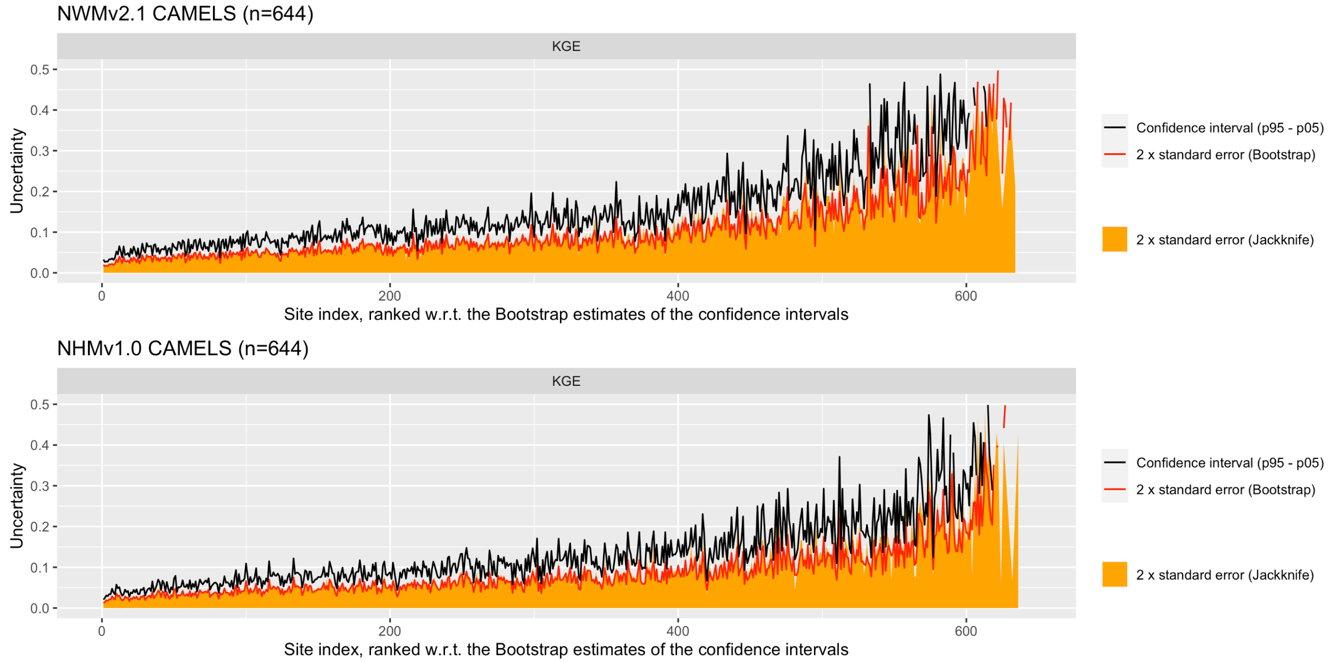


Figure S11. Estimates of uncertainty in the KGE estimates for the CAMELS (Catchment Attributes and Meteorology for Large-sample Studies) basins (Addor et al. 2017) using the gumboot package (Clark and Shook, 2021) in R (R Core Team, 2021) for the National Water Model v2.1 (NWMv2.1; top) and National Hydrologic Model v1.0 (NHMv1.0; bottom). Quantification of the uncertainty is obtained from 2x standard error estimates obtained using jackknife and bootstrap estimates, as well as intervals computed as the difference between the 95th and 5th percentiles of the bootstrap samples (see Clark et al. 2001 for details). The figure shows the uncertainty in the KGE estimates, with the bootstrap and jackknife showing similar estimates for both models. KGE uncertainty estimates for the full set of gages in this study (Foks et al. 2022) are included in Towler et al. (2023a, 2023b).

Equations:

The percent bias in the high flows (PBIAS_HF) is defined as (Yilmaz et al. 2008):

$$PBIAS_{HF} = \frac{\sum_{h=1}^{H}(S_h - O_h)}{\sum_{h=1}^{h} O_h}$$

Where h = 1, 2,… H are the low flow indices for flows with exceedance probabilities lower than 0.02.

The percent bias in the low-flow (PBIAS_LF) is defined as (Yilmaz et al. 2008):

$$PBIAS_{LF} = -1 \cdot \frac{\sum_{l=1}^{L}[log(S_l) - log(S_L)] - \sum_{l=1}^{L}[log\ (O_l - log(O_L)]}{\sum_{l=1}^{L}[log(O_l) - log(O_L)]} \times 100$$

where l = 1, 2,…L is the flow value index in the low-flow segment (0.7–1.0 flow exceedance probabilities) of the flow duration curve and L is the minimum flow index.

References:

Addor, N., Newman, A. J., Mizukami, N., and Clark, M. P.: The CAMELS data set: catchment attributes and meteorology for large-sample studies, Hydrol. Earth Syst. Sci., 21, 5293–5313, https://doi.org/10.5194/hess-21-5293-2017, 2017.

Clark M. and Shook, K: Package ‘gumboot: Bootstrap Analyses of Sampling Uncertainty in Goodness-of-Fit Statistics’, available online at: https://cran.r-project.org/web/packages/gumboot/index.html (last access March 6, 2023), 2021.

Clark, M. P., Vogel, R. M., Lamontagne, J. R., Mizukami, N., Knoben, W. J. M., Tang, G., et al.: The abuse of popular performance metrics in hydrologic modeling, Water Resour. Res., 57, e2020WR029001. https://doi.org/10.1029/2020WR029001, 2021.

ESRI: USA States Generalized Boundaries [Data set]. ESRI. https://esri.maps.arcgis.com/home/item.html?id=8c2d6d7df8fa4142b0a1211c8dd66903, 2022a.

ESRI: World Countries (Generalized) [Data set]. ESRI. https://hub.arcgis.com/datasets/esri::world-countries-generalized/about, 2022b.

Foks, S.S., Towler, E., Hodson, T.O., Bock, A.R., Dickinson, J.E., Dugger, A.L., Dunne, K.A., Essaid, H.I., Miles, K.A., Over, T.M., Penn, C.A., Russell, A.M., Saxe, S.W., and Simeone, C.E.: Streamflow benchmark locations for conterminous United States, version 1.0 (cobalt gages): U.S. Geological Survey data release, https://doi.org/10.5066/P972P42Z, 2022.

Grannemann, N. G.: Great Lakes and Watersheds Shapefiles [Data set]. USGS. https://www.sciencebase.gov/catalog/item/530f8a0ee4b0e7e46bd300dd, 2010.

Natural Earth Data.: Ocean (version 5.1.1) [Data set]. Natural Earth Data. https://www.naturalearthdata.com/downloads/10m-physical-vectors/10m-ocean/, 2009.

R Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, available at: https://www.r-project.org (last access: 4 May 2022), 2021.

Towler, E., Foks, S.S., Staub, L.E., Dickinson, J.E., Dugger, A.L., Essaid, H.I., Gochis, D., Hodson, T.O., Viger, R.J., and Zhang, Y.: Daily streamflow performance benchmark defined by the standard statistical suite (v1.0) for the National Water Model Retrospective (v2.1) at benchmark streamflow locations for the conterminous United States (ver 3.0, March 2023): U.S. Geological Survey data release, https://doi.org/10.5066/P9QT1KV7, 2023a.

Towler, E., Foks, S.S., Staub, L.E., Dickinson, J.E., Dugger, A.L., Essaid, H.I., Gochis, D., Hodson, T.O., Viger, R.J., and Zhang, Y.: Daily streamflow performance benchmark defined by the standard statistical suite (v1.0) for the National Hydrologic Model application of the Precipitation-Runoff Modeling System (v1 byObs Muskingum) at benchmark streamflow locations for the conterminous United States (ver 3.0, March 2023): U.S. Geological Survey data release, https://doi.org/10.5066/P9DKA9KQ, 2023b.

Yilmaz, K., Gupta, H., and Wagener, T.: A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model, *Water Resour. Res., 44*, 2008.