Supplement of Hydrol. Earth Syst. Sci., 30, 2651–2666, 2026
https://doi.org/10.5194/hess-30-2651-2026-supplement

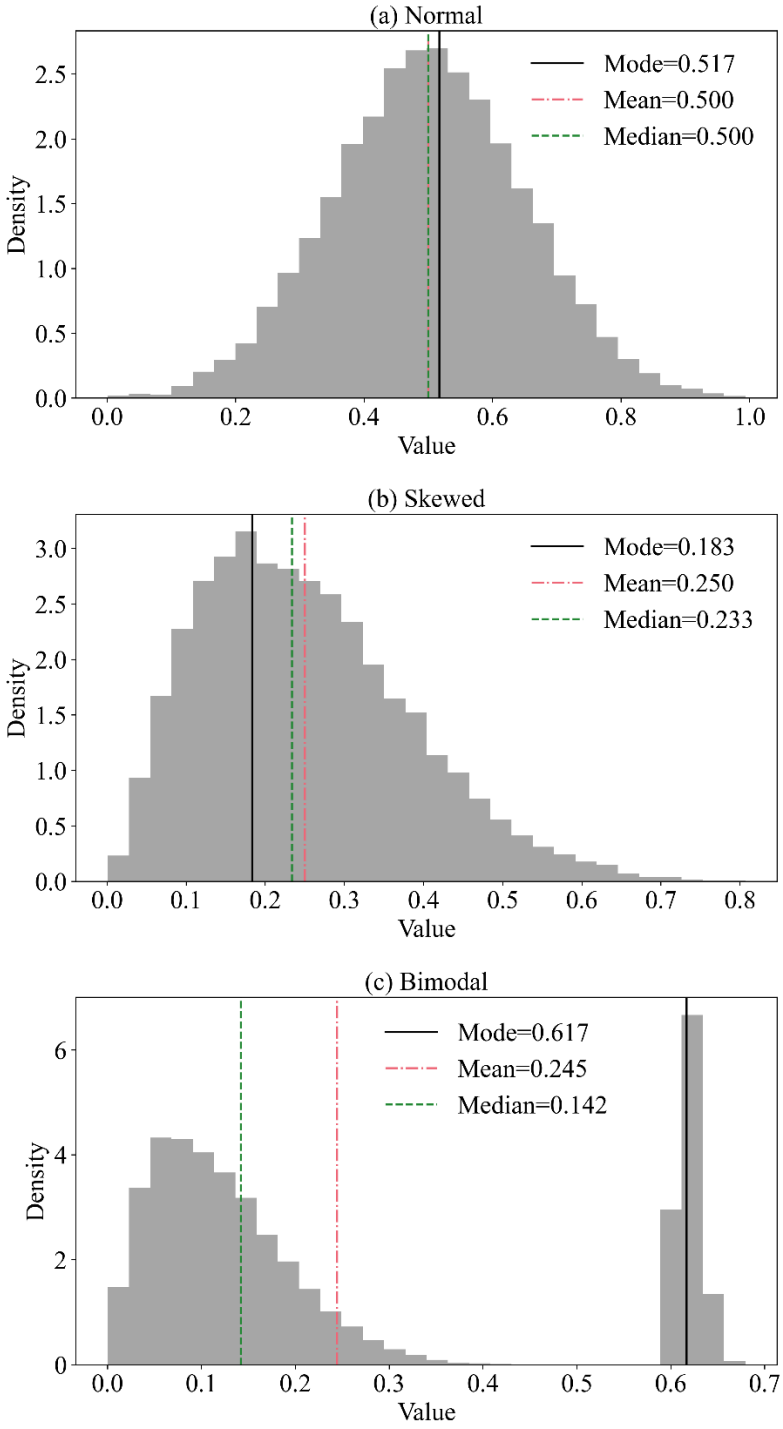
Hydrology and Earth System Sciences
Open Access
EGU

Supplement of

Introducing the Model Fidelity Metric (MFM) for robust and diagnostic land surface model evaluation

Zezhen Wu et al.

Correspondence to: Zhongwang Wei (weizhw6@mail.sysu.edu.cn) and Yongjiu Dai (daiyj6@mail.sysu.edu.cn)



15 **Figure S1.** Artifacts of moment-based metrics. **(a)** Normal distribution. Mode, mean, and median converge. (b) Skewed distribution. Mean and median diverge from the mode. **(c)** Bimodal distribution. All three statistics become unrepresentative artifacts.

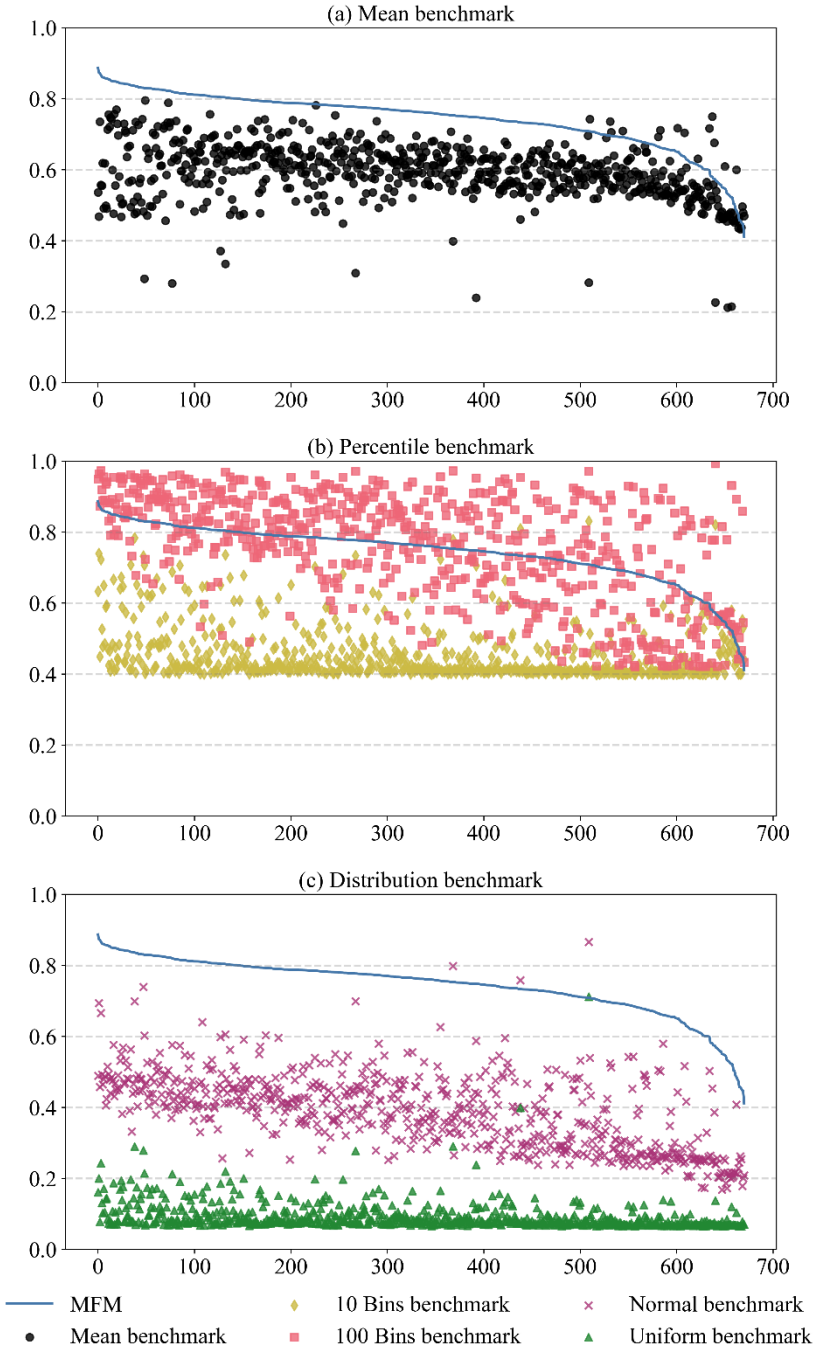


Figure S2. Benchmark evaluation of MFM (default parameters: $c = 4.0, p = 1.0, n_{\mathrm{SUSE}} = n_{\mathrm{PHI}} = 10$). Grey lines are 0.2, 0.4, 0.6, and 0.8. CAMELS observations are benchmarked against synthetic simulations paired in rank order. Each point represents a site. **(a)** Mean benchmark. Synthetic data equal the observed mean. **(b)** Percentile benchmark. Red squares denote 100-bin discretization (percentile) and yellow diamonds represent 10-bin discretization. **(c)** Distributional benchmark. Purple crosses indicate normal sampling (matching observed mean and standard deviation). Green triangles denote uniform sampling across the observed range.

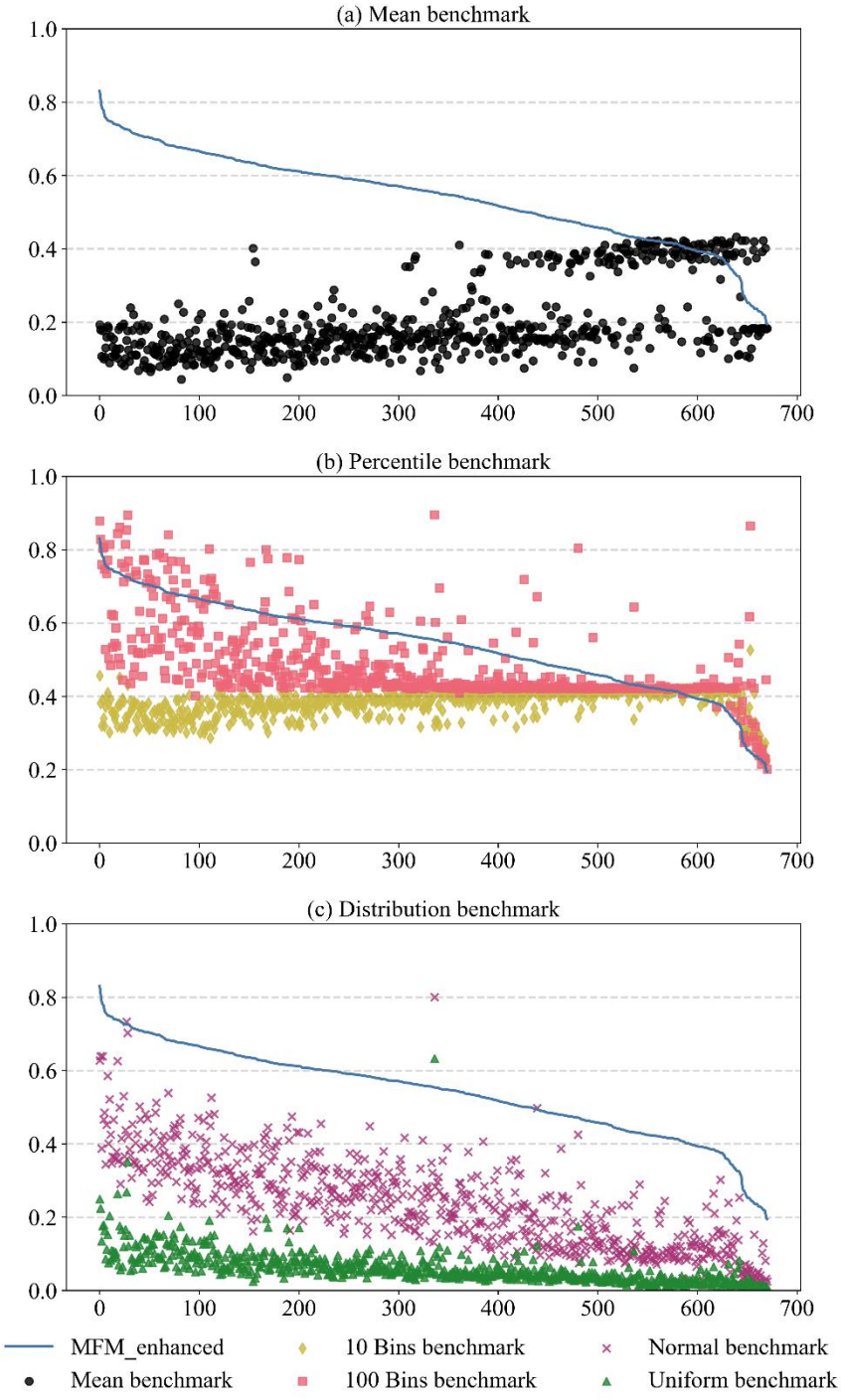


Figure S3. Benchmark evaluation of enhanced MFM (enhanced parameters: $c = p = 2.0, n_{\mathrm{SUSE}} = n_{\mathrm{PHI}} = 100$). Grey lines are 0.2, 0.4, 0.6, and 0.8. The methodology of generating synthetic simulations is the same as in Fig. S2. **(a)** Mean benchmark. **(b)** Percentile benchmark. **(c)** Distributional benchmark.