Supporting Information for

Quantification of Data and Model Uncertainty for Deep Learning-based Streamflow Prediction

Kaidi Peng[1], Daniel B. Wright[1], Lei Yan[1], G. Aaron Alexander[1], Yagmur Derin[2]

[1]Department of Civil and Environmental Engineering, University of Wisconsin-Madison, Madison, WI, USA.

[2]Department of Civil and Environmental Engineering, University of Iowa, Iowa City, Iowa, USA.

Text S1. Evaluation Metrics

$$\mathrm{CRPS} = \frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{N}\sum_{n=1}^{N}\left|\Delta_{n,t} - y_t\right| - \frac{1}{2N^2}\sum_{n=1}^{N}\sum_{m=1}^{N}\left|\Delta_{n,t} - \Delta_{m,t}\right|\right],$$

where $\Delta_{n,t}$ and $\Delta_{m,t}$ are the individual ensemble members and y_t is the observation at time t. T is the number of time steps in the time series. N is the ensemble size.

$$\mathrm{CR} = \frac{1}{T}\sum_{t=1}^{T}\mathrm{I}_t\left(\Delta_{\min,t} \le y_t \le \Delta_{\max,t}\right),$$

where $\Delta_{\min,t}$ and $\Delta_{\max,t}$ are the smallest and largest values of the predicted ensemble at time t, respectively. I_t is an indicator that equals one when the expression enclosed in parentheses is fulfilled at time t and zero otherwise.

Table S1. Characteristics of 17 study watersheds

Gauge ID	Area [km^2]	Aridity [-]	Annual PET [mm]	Annual Precip. [mm]	Snow fraction [%]	Class
01144000	1787	0.88	1000	1143	18	Natural
01532000	557	0.97	1108	1148	12	Natural
02443500	1852	1.09	1640	1511	0	Disturbed
02425000	4574	1.13	1665	1471	0	Disturbed
03083500	4442	1.02	1205	1179	8	Disturbed
06799100	1738	1.79	1292	721	11	Disturbed
06814000	715	1.56	1383	888	3	Natural
06820500	4558	1.31	1293	986	6	Disturbed
06934000	8236	1.18	1384	1174	4	Natural
07184000	510	1.28	1436	1125	3	Natural
07227100	2036	5.10	2075	407	14	Disturbed
08150000	4802	3.15	1916	608	2	Disturbed
08150800	557	2.58	1844	713	1	Natural
09504420	603	6.30	2046	325	21	Disturbed
14137000	684	0.77	1092	1423	20	Natural
14189000	4636	0.96	1091	1141	5	Disturbed
14301000	1728	0.50	967	1949	3	Natural

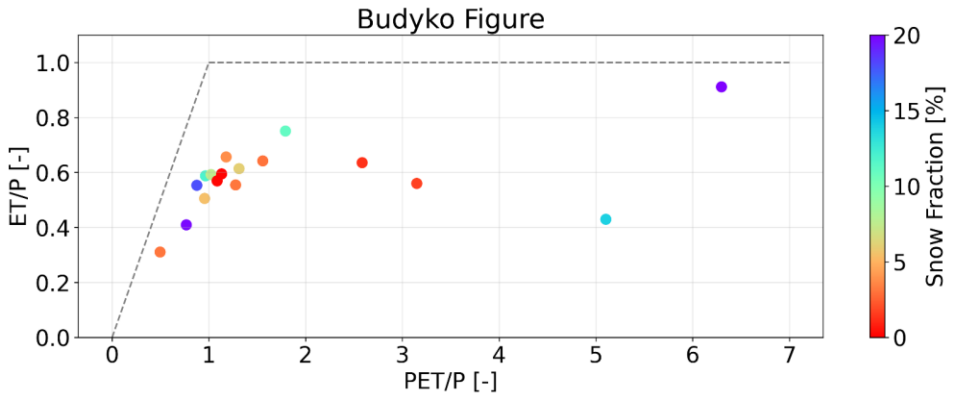


Figure S1. The placement of 17 watersheds on the Budyko curve.

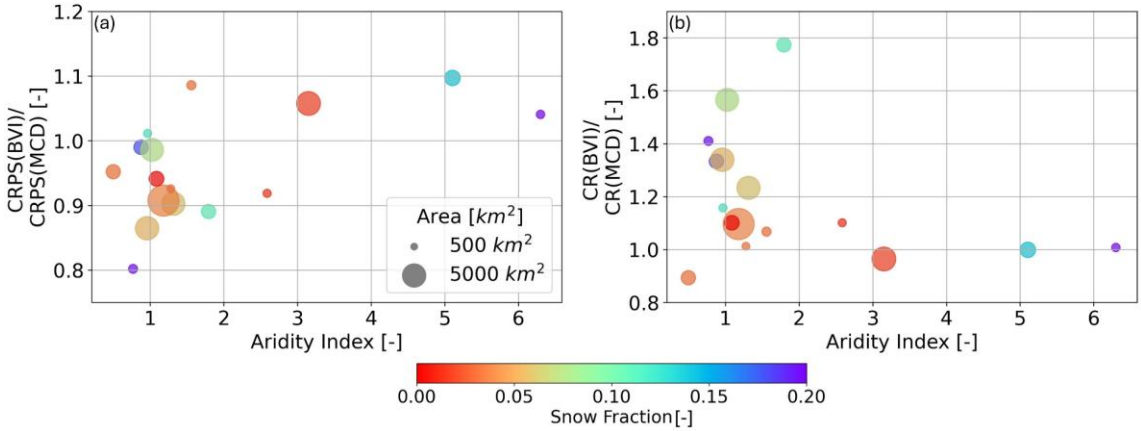


Figure S2. The comparison between BVI and MCD in terms of (a) CRPS and (b) CR. Lower CRPS and higher CR mean a better performance.

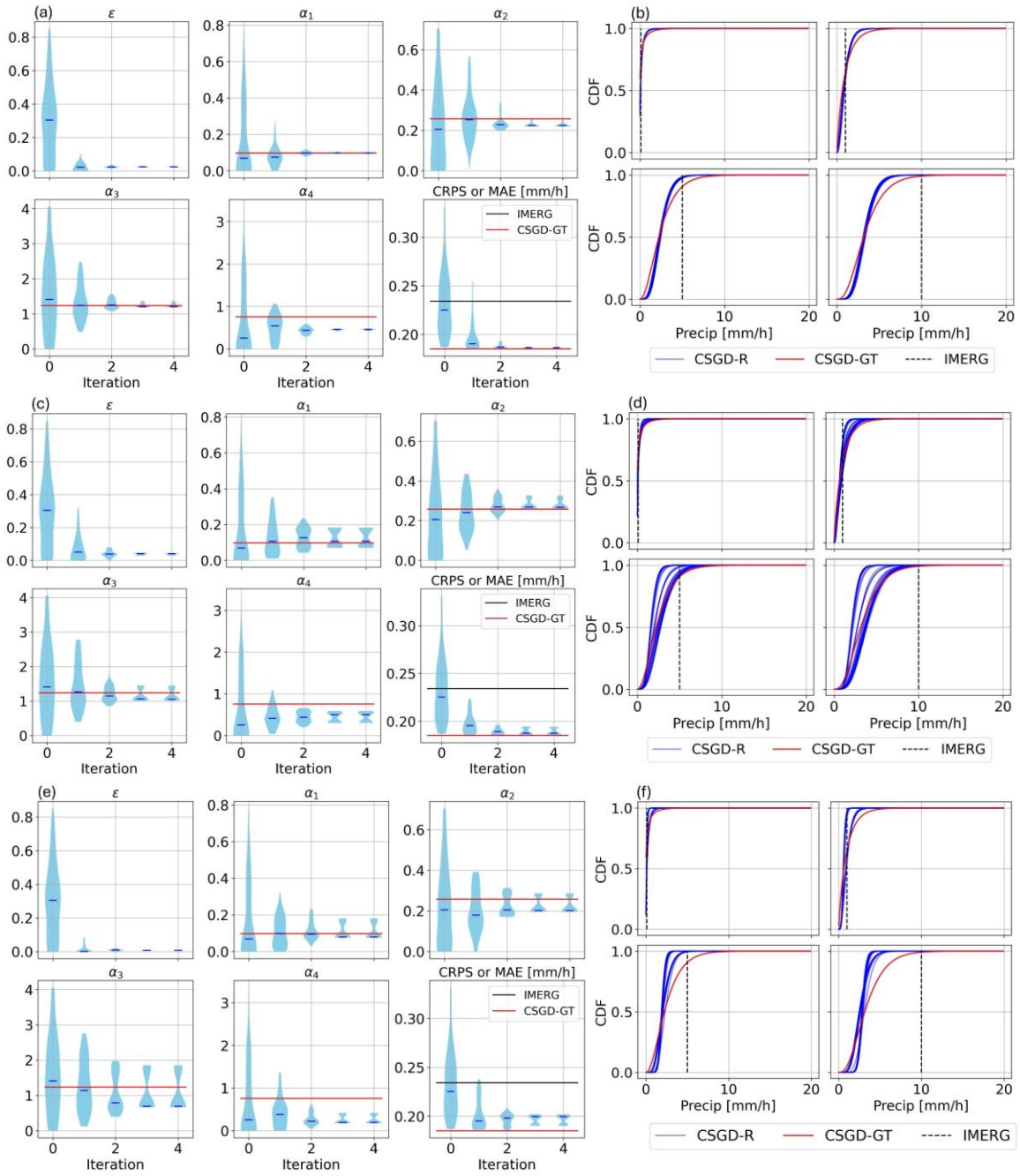


Figure S3. The same as Figure 3, but for the synthetic study of watershed 03083500, with added streamflow uncertainty (a&b) ε =0.01, (c&d) ε =0.03, (e&f) ε =0.05. The retrieved ε in (a) is 0.02, (b) is 0.04, and (c) is 0.008.

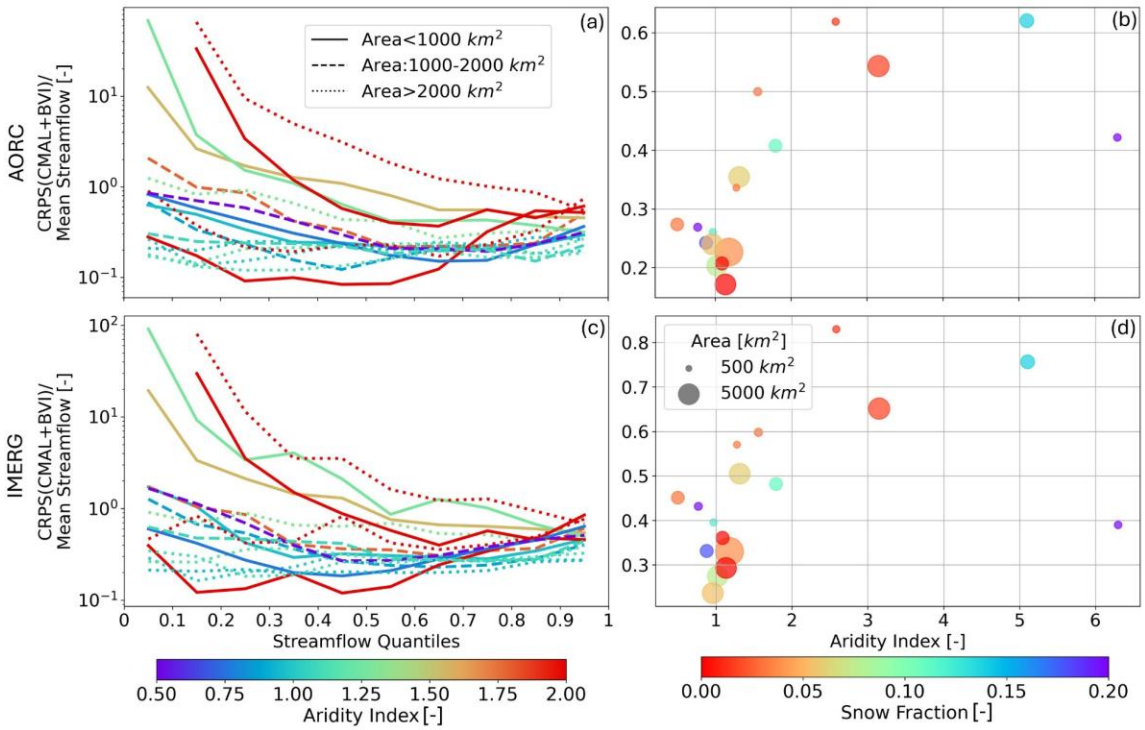


Figure S4. The same as Figure 8, but for AORC and IMERG.

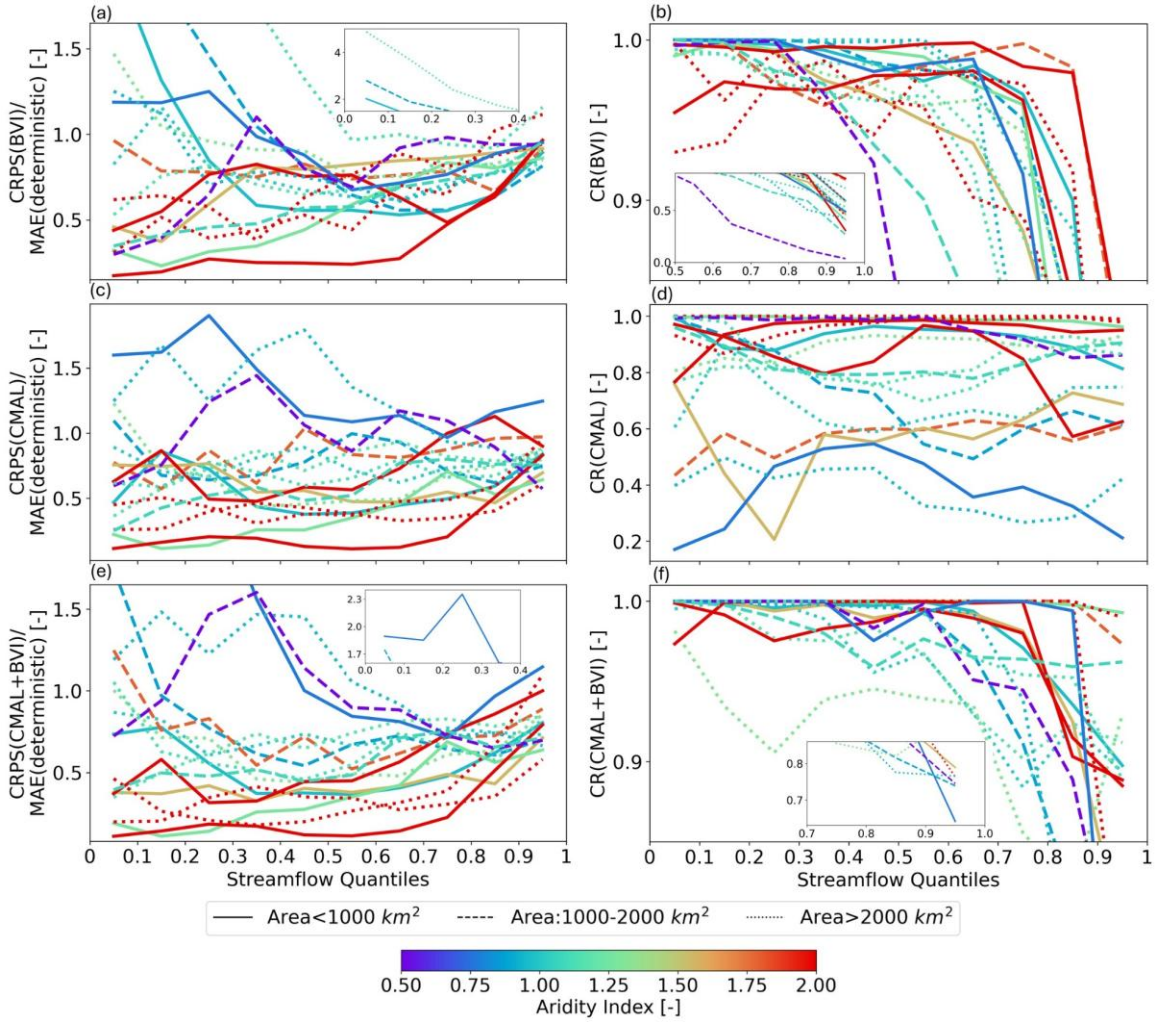


Figure S5. The same as Figure 10, but for AORC.

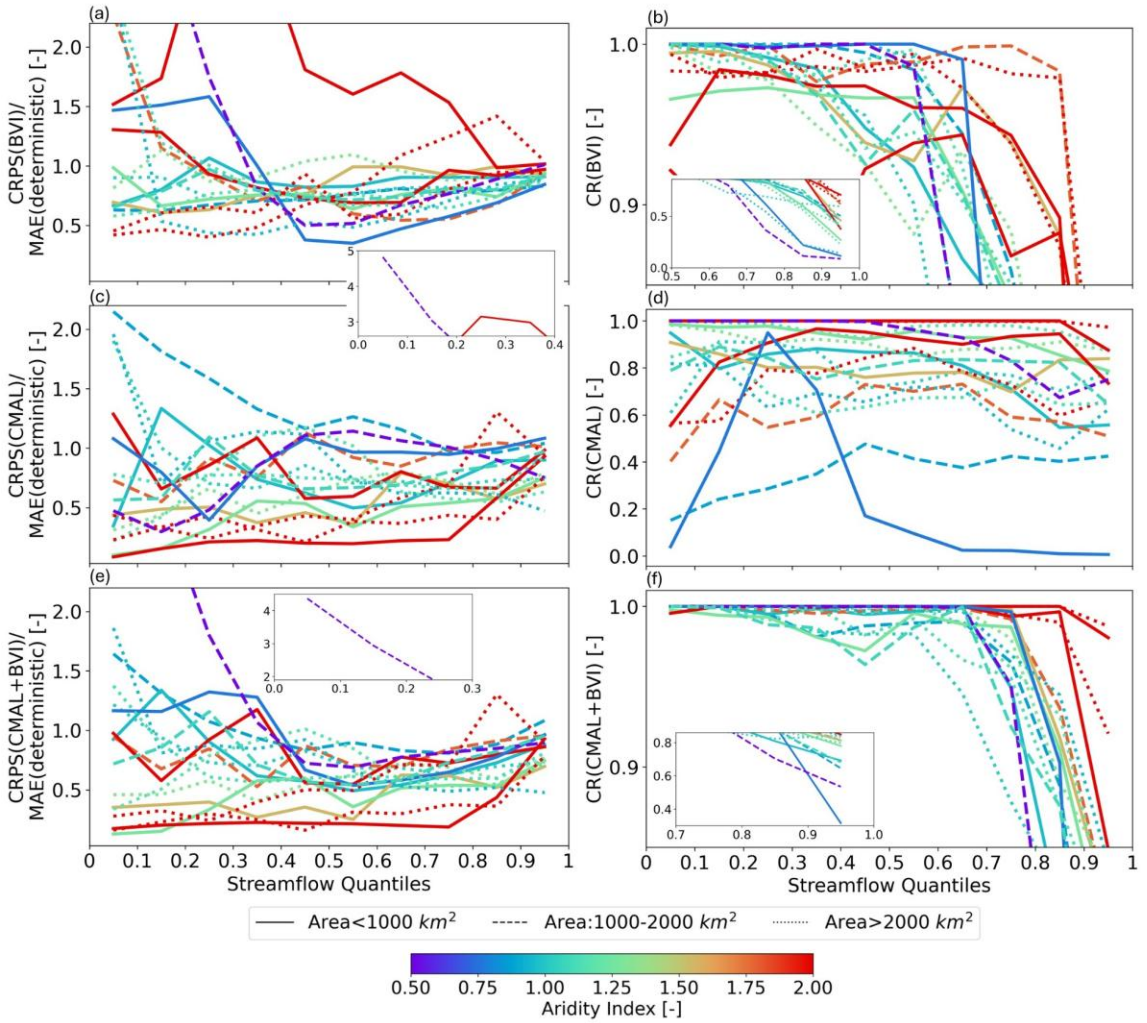


Figure S6. The same as Figure 10, but for IMERG.

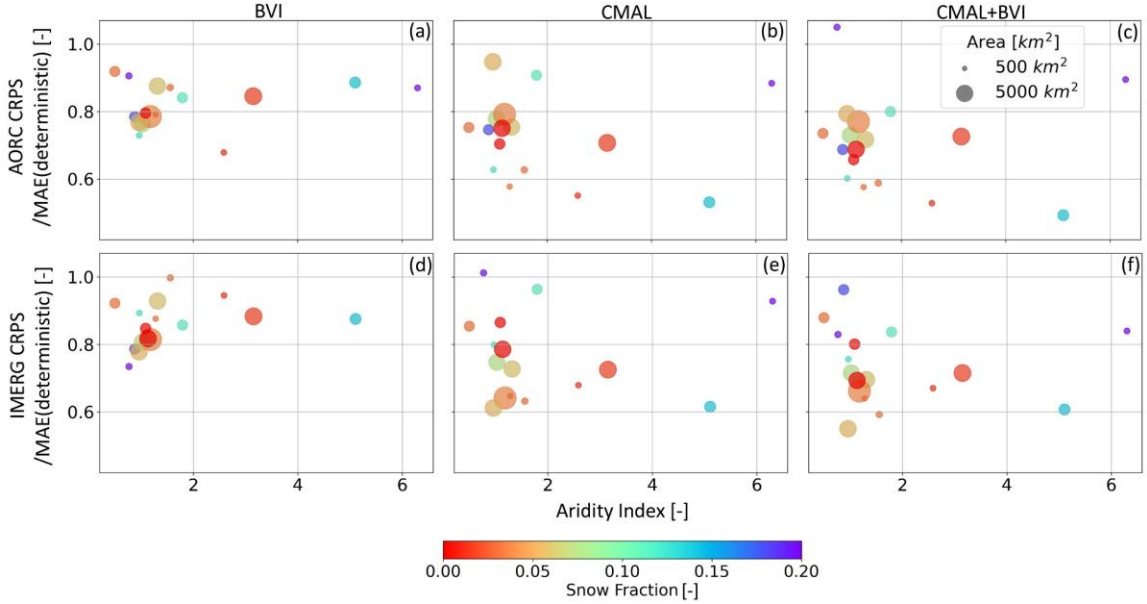


Figure S7. The same as Figure 11, but for AORC and IMERG.

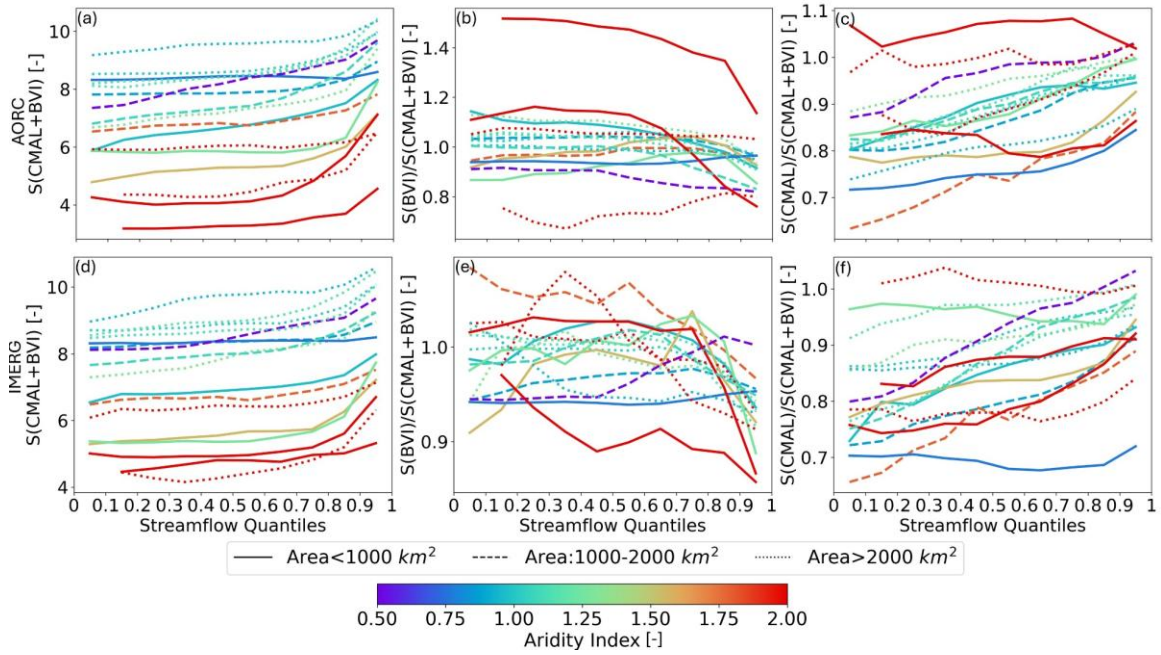


Figure S8. The same as Figure 12a-c, but for AORC and IMERG.

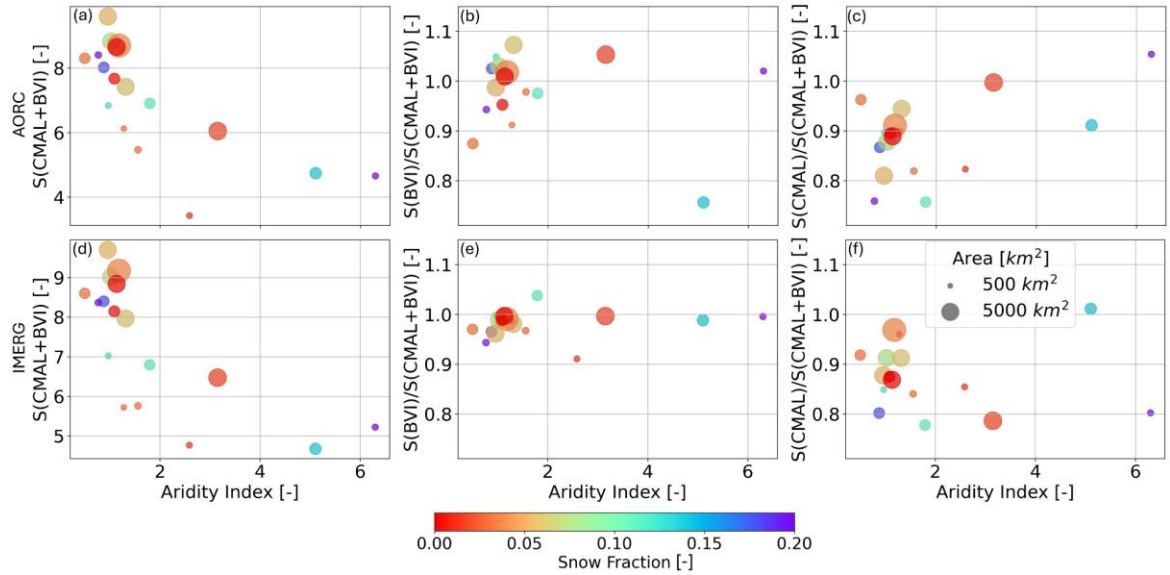
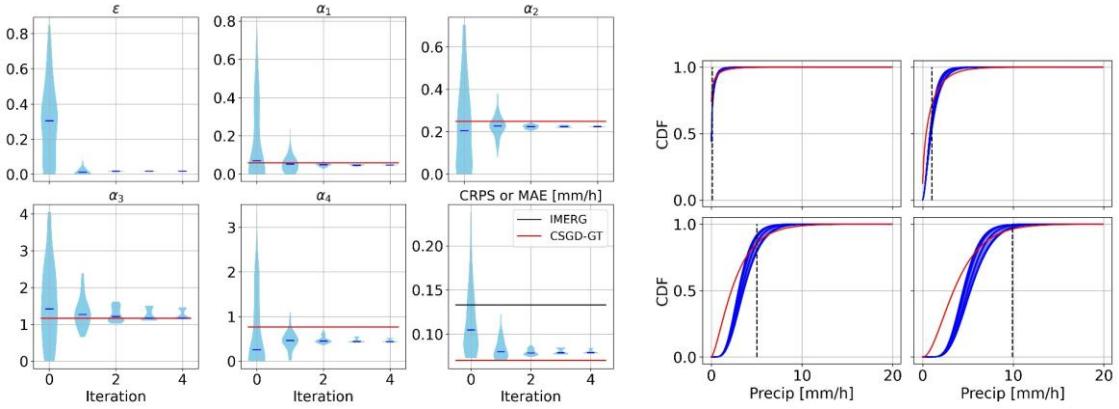
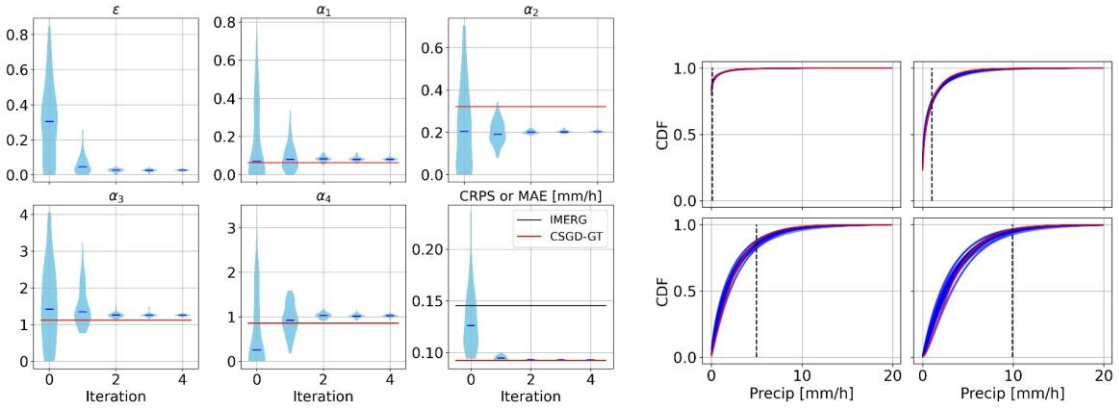


Figure S9. The same as Figure 12d-f, but for AORC and IMERG.

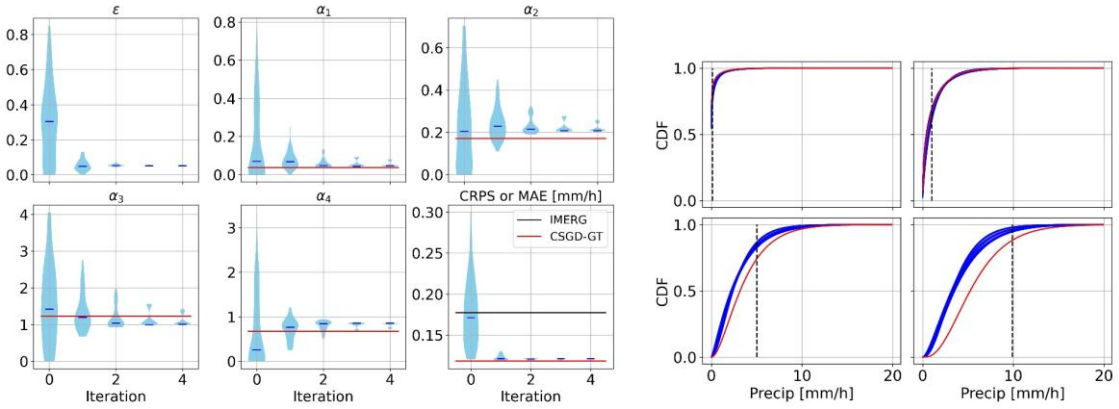
(a) 01144000



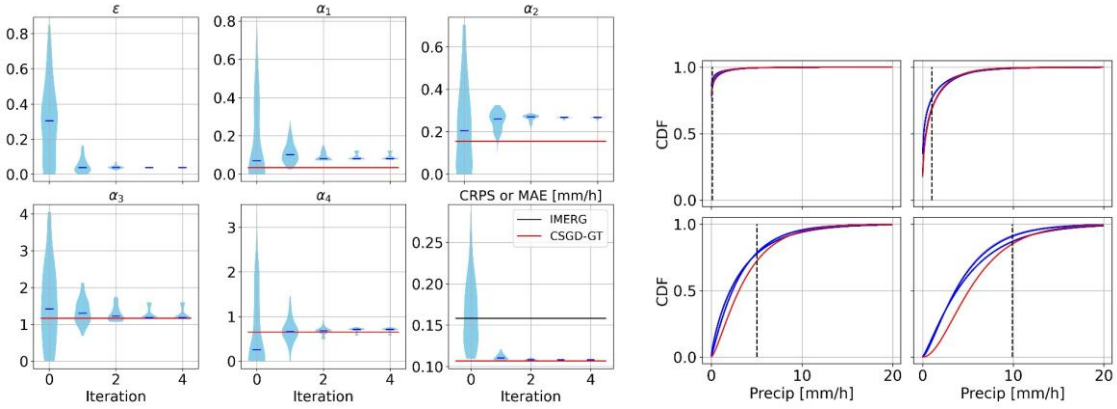
(b) 01532000



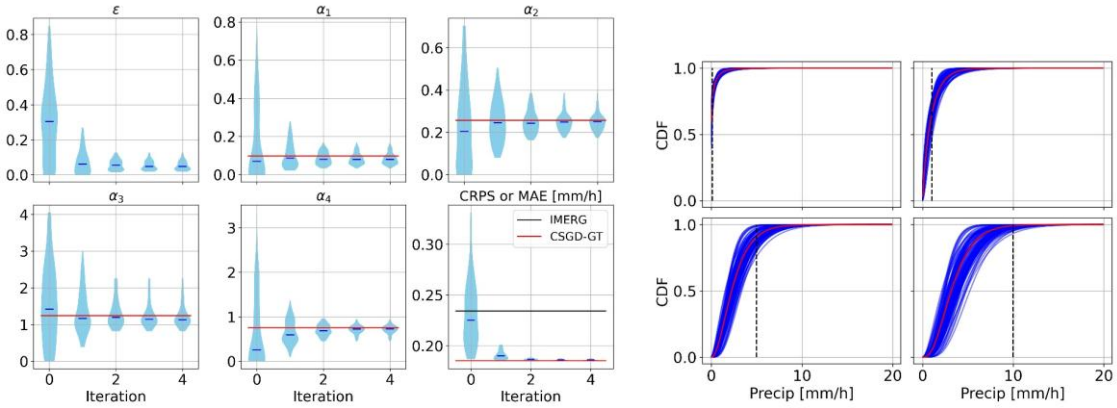
(c) 02425000



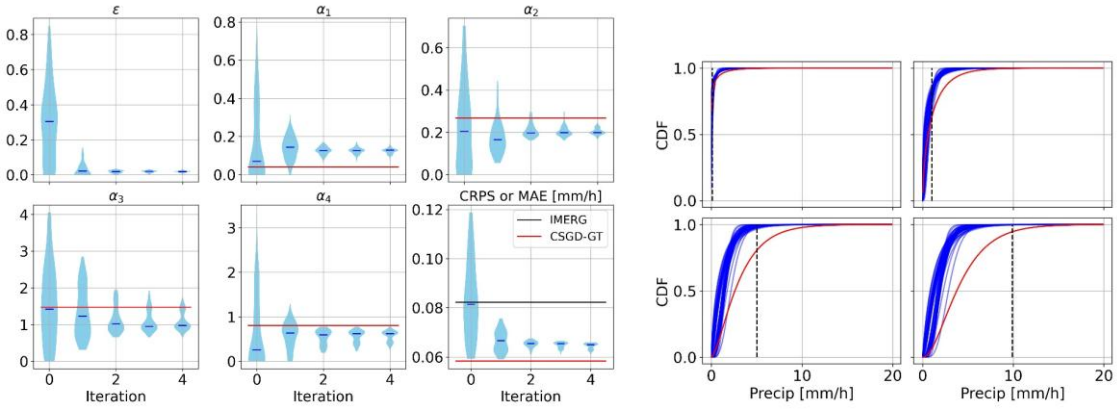
(d) 02443500



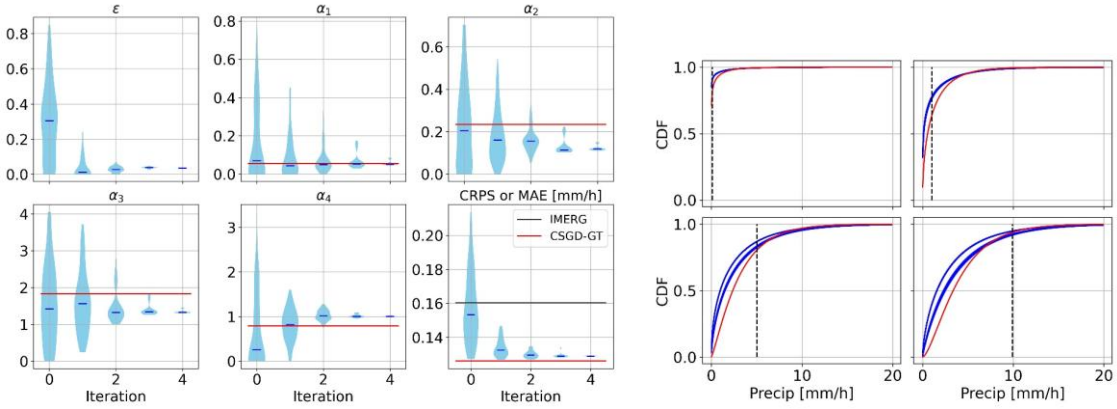
(e) 03083500



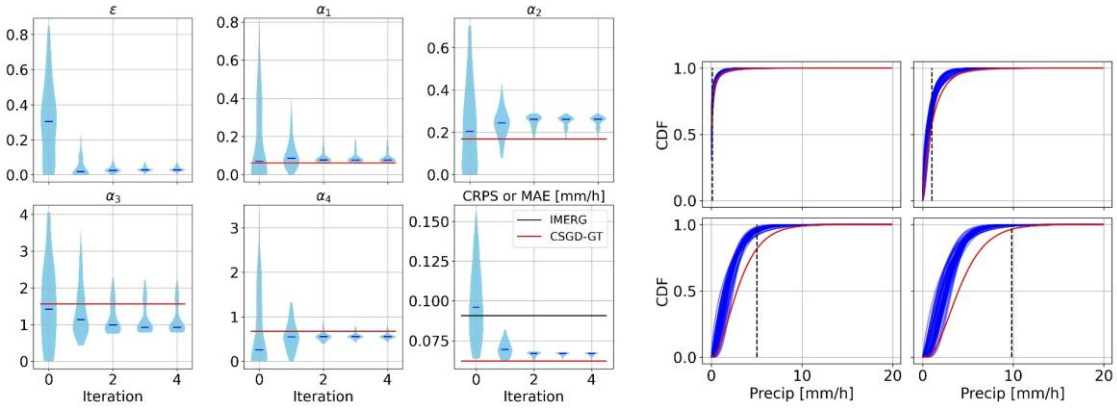
(f) 06799100



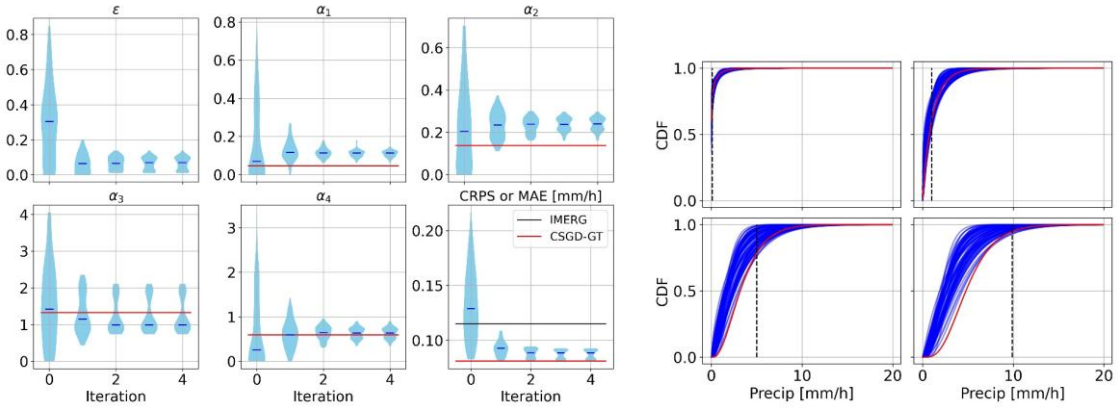
(g) 06814000



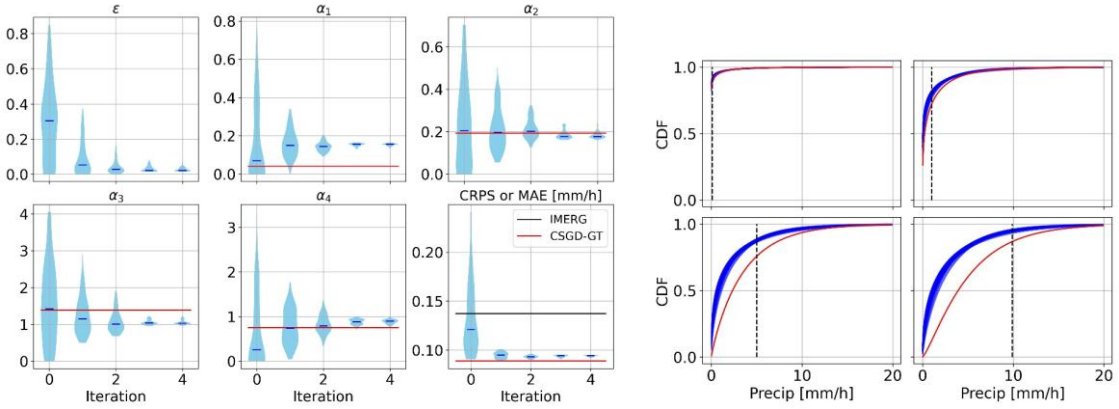
(h) 06820500



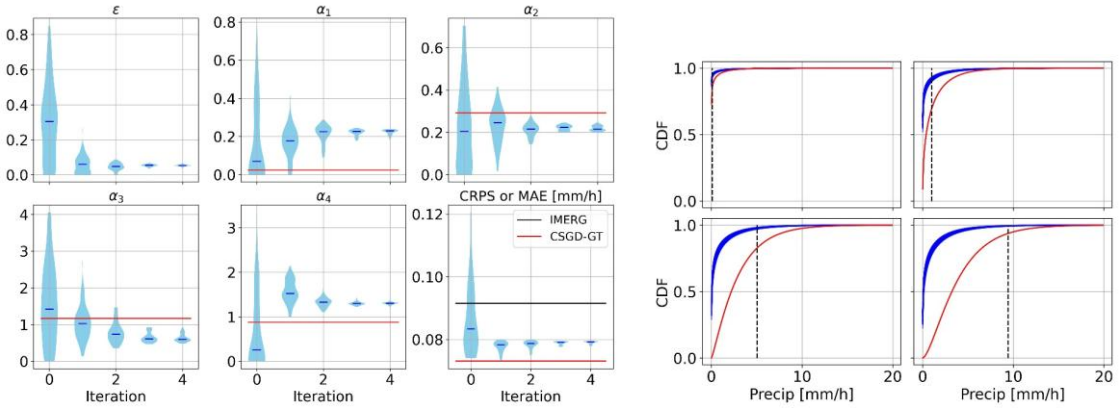
(i) 06934000



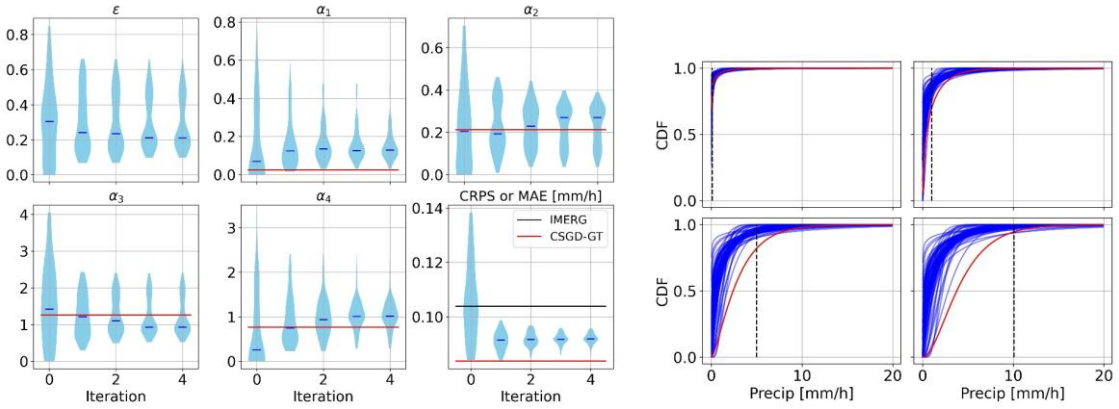
(j) 07184000



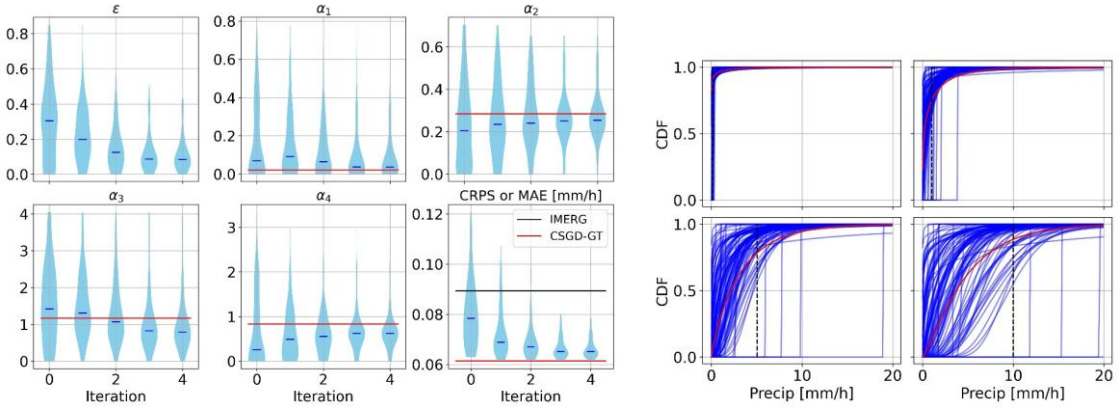
(k) 07227100



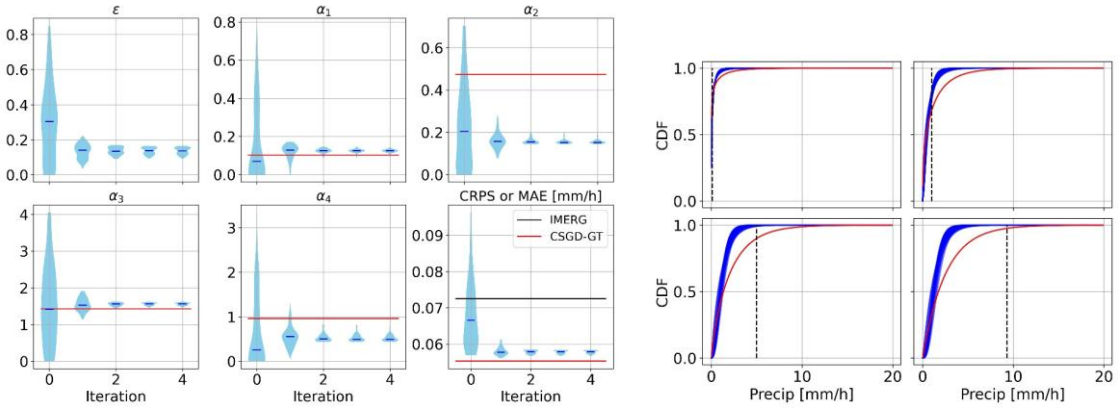
(l) 08150000



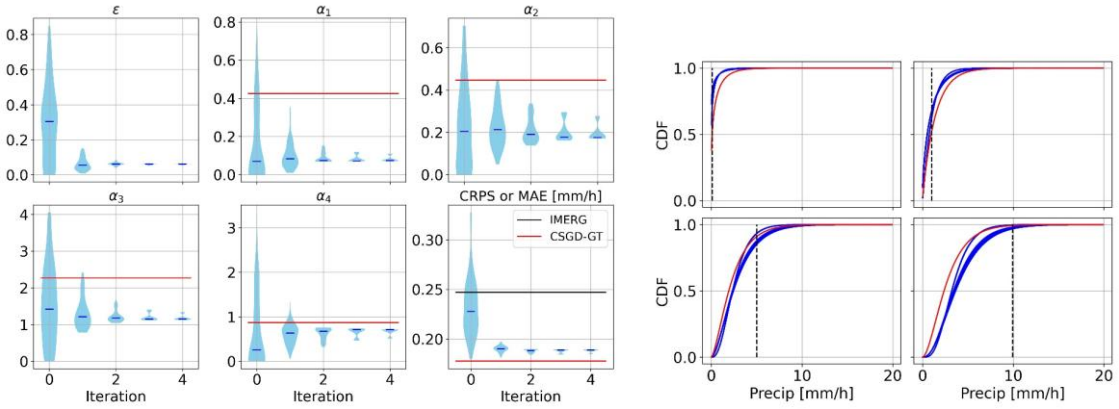
(m) 08150800



(n) 09504420



(o) 14137000



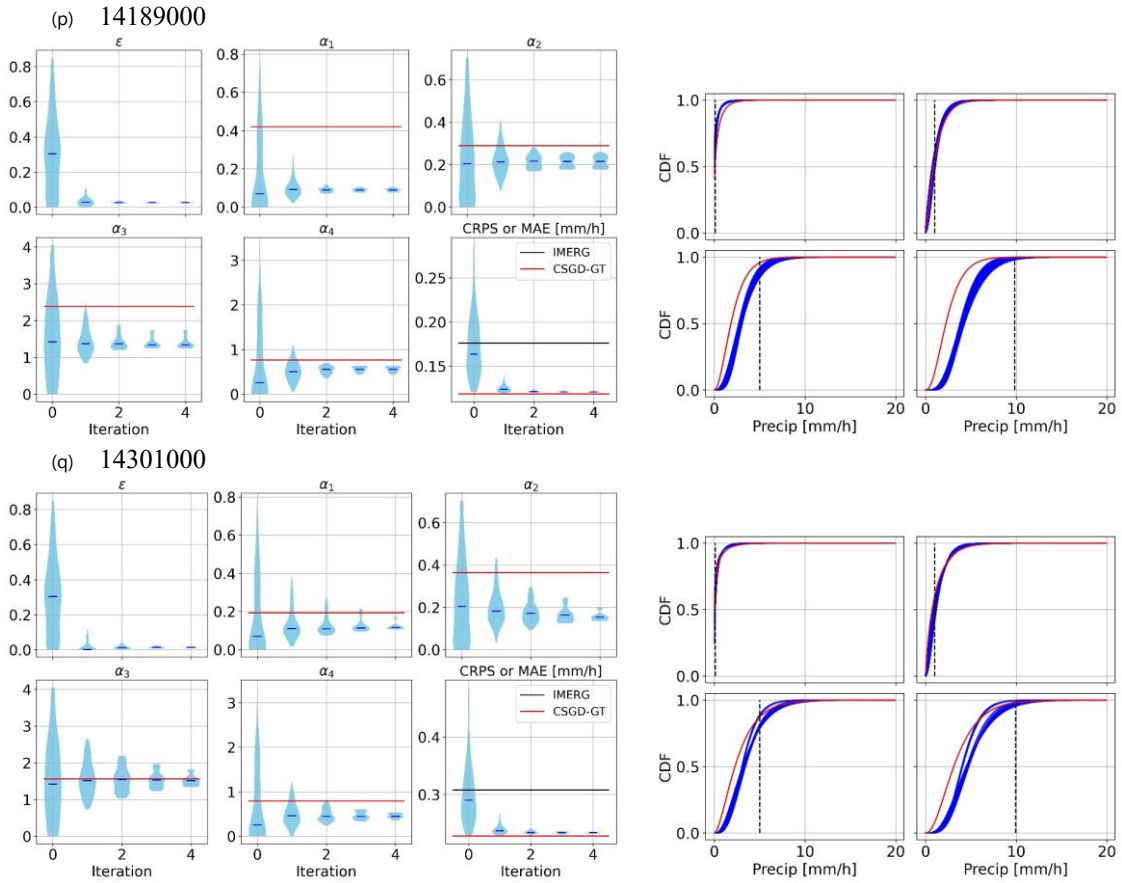


Figure S10. The same as Figure 3, but for all the watersheds in standard retrieval.