



Supplement of

Improving large-scale river routing models with ESA long-term CCI discharge data assimilation: advancing accuracy for climate studies

Malak Sadki et al.

Correspondence to: Malak Sadki (malak.sadki@magellium.fr)

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

S1 Comparison of CCI discharge products with in-situ observations

This section provides a quantitative evaluation of the CCI discharge products used in this study. Both altimetry-based (Q_{alti}) and multispectral-based ($Q_{\text{multispec}}$) discharge estimates are compared against independent in-situ measurements from hydrometric networks in the Niger and Congo river basins. The objective of this comparison is to assess the reliability of the CCI satellite-derived discharge products prior to their assimilation into hydrological models. Metrics such as the Kling–Gupta Efficiency (KGE), bias, and correlation are presented, along with selected hydrographs for representative stations.

For the Niger Basin, we compared the CCI data with 13 stations of the Niger Basin Authority (ABN) network, covering the period from approximately 2010 to 2017. In the Congo Basin, the comparison is made with the So-Hybam database, which includes two stations that overlap with the CCI data, providing a longer temporal record from 1947 to 2020.

Figure S1 illustrates the performance of the KGE index and its three components — β (bias), r (correlation), and γ (internal variability) — for the CCI discharge products across stations in the Niger and Congo basins, with Q_{alti} represented in blue and $Q_{\text{multispec}}$ in green. Following this, Figure S2 presents the time series of discharge estimates from satellite altimetry (Q_{alti}) and multispectral imagery ($Q_{\text{multispec}}$) compared to in-situ observations for five stations in the Niger basin and two stations in the Congo basin, with an overview of the period from 2010 to 2020.

In the Niger basin, the analysis of the 20-year time series (2000–2020) across stations such as Koulikoro, Ansongo, Niamey and Lokoja reveals distinct patterns in the performance of Q_{alti} and $Q_{\text{multispec}}$ products. Figure S2 shows that Q_{alti} data generally capture the seasonal hydrological cycle well, particularly in upstream stations like Koulikoro and Ansongo, where the correlations with in-situ data are 0.99 and 0.92, respectively. However, Q_{alti} tends to overestimate high flows, as indicated by biases of 1.35 at Koulikoro and 1.06 at Ansongo. In Niamey, the availability of Q_{alti} data from 2017 only limits the analysis. However, the two-peak hydrological cycle is reasonably well-represented, with a strong correlation of 0.96, although the peak flows are slightly underestimated (bias of 0.85). At Lokoja, near the Niger River’s outlet, a similar underestimation is observed with a bias of 0.90, but the seasonality remains well represented, with a correlation of 0.79.

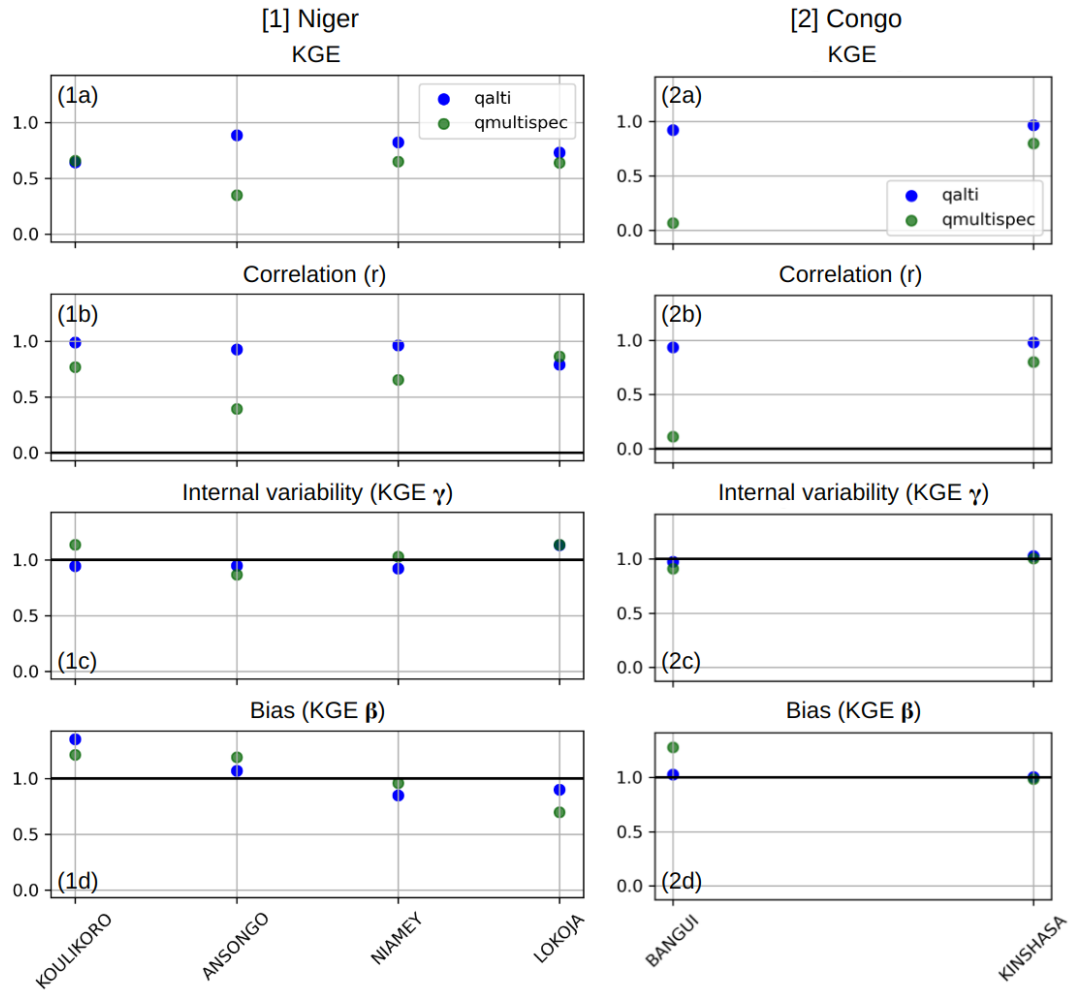


Figure S1. KGE score rate and its 3 components (r , β , γ) for CCI discharge products at stations where in-situ observations are available (Niger Basin left, Congo Basin right): altimetry-based discharge (Q_{alti}) in blue and multispectral imagery discharge ($Q_{\text{multispec}}$) in green.

In contrast, $Q_{\text{multispec}}$ data provide denser coverage, particularly during low-flow periods, as seen in Figure S2. However, this increased data density comes at the cost of increased noise, leading to generally lower performance scores compared to Q_{alti} . For instance, at Ansongo, the correlation with in-situ data drops significantly to 0.40 for $Q_{\text{multispec}}$, compared to 0.92 for Q_{alti} . This decrease in performance is largely due to a temporal shift in the hydrological cycle representation, as well as greater data dispersion, particularly during high-flow periods.

At Koulikoro, $Q_{\text{multispec}}$ data perform better in high-flow periods but exhibit a slight lag during the recession phase, resulting in a correlation of 0.76 compared to 0.92 for Q_{alti} . Despite these issues, $Q_{\text{multispec}}$ data show improved representation of low flows, closer to in-situ observations, compared to Q_{alti} . Downstream at Lokoja, $Q_{\text{multispec}}$ continues to underperform, with a bias of 0.70 compared to 0.90 for Q_{alti} , and lower correlations due to increased noise and data dispersion.

At Niamey, the performance of the $Q_{\text{multispec}}$ product against in-situ observations is subpar, with a KGE score of 0.65 for daily data, and dropping further to 0.57 for monthly data. This underperformance is primarily due to two factors: a temporal lag in the representation of the hydrological cycle and the failure to accurately reproduce the double peak characteristic of the flow at Niamey. The first peak, which results from the contributions of a nearby tributary, is particularly missed. This temporal

lag significantly reduces the correlation, with a monthly correlation of only 0.62, compared to 0.97 for the Q_{alti} product derived from altimetry data.

In the Congo basin, where comparisons were limited to Bangui and Kinshasa due to the lack of independent in-situ data (not available at Chembe-Ferry), Q_{alti} data show strong performance, closely aligning with in-situ observations. At both stations, Q_{alti} achieved KGE scores exceeding 0.92, with correlations greater than 0.93 and minimal bias (1), demonstrating a robust representation of the internal variability of the hydrological cycle, as shown in Figure S1. In contrast, $Q_{\text{multispec}}$ data, while showing similar correlation and internal variability representation at Kinshasa, suffer from greater data dispersion, leading to a lower daily correlation of 0.8 and a KGE score of 0.8.

At Bangui, however, the $Q_{\text{multispec}}$ data are less reliable, showing significant dispersion and failing to accurately represent the hydrological cycle. This is reflected in a very low correlation of 0.11 daily and 0.02 monthly, and a poor KGE score of 0.07, indicating that while $Q_{\text{multispec}}$ can offer denser temporal data, the quality and reliability are significantly compromised in certain locations.

Despite the poor quality of the multispectral imagery discharge products, these data have been retained in the analysis for comparison purposes. Currently, the calibrated multispectral data is available only for Niamey in the Niger basin and for Kinshasa and Bangui in the Congo basin. This limitation constrains the analysis of the added value of data assimilation over the MGB model, considering the size of the basins, and restricts the ability to conduct a coherent comparison with other experiments (dH, Q_{alti}). The calibrated versions are expected to be available in Phase 2 of the CCI-Discharge project, covering all analyzed stations. It will be valuable to revisit and reassess this work with the calibrated datasets once they become available to evaluate their performance.

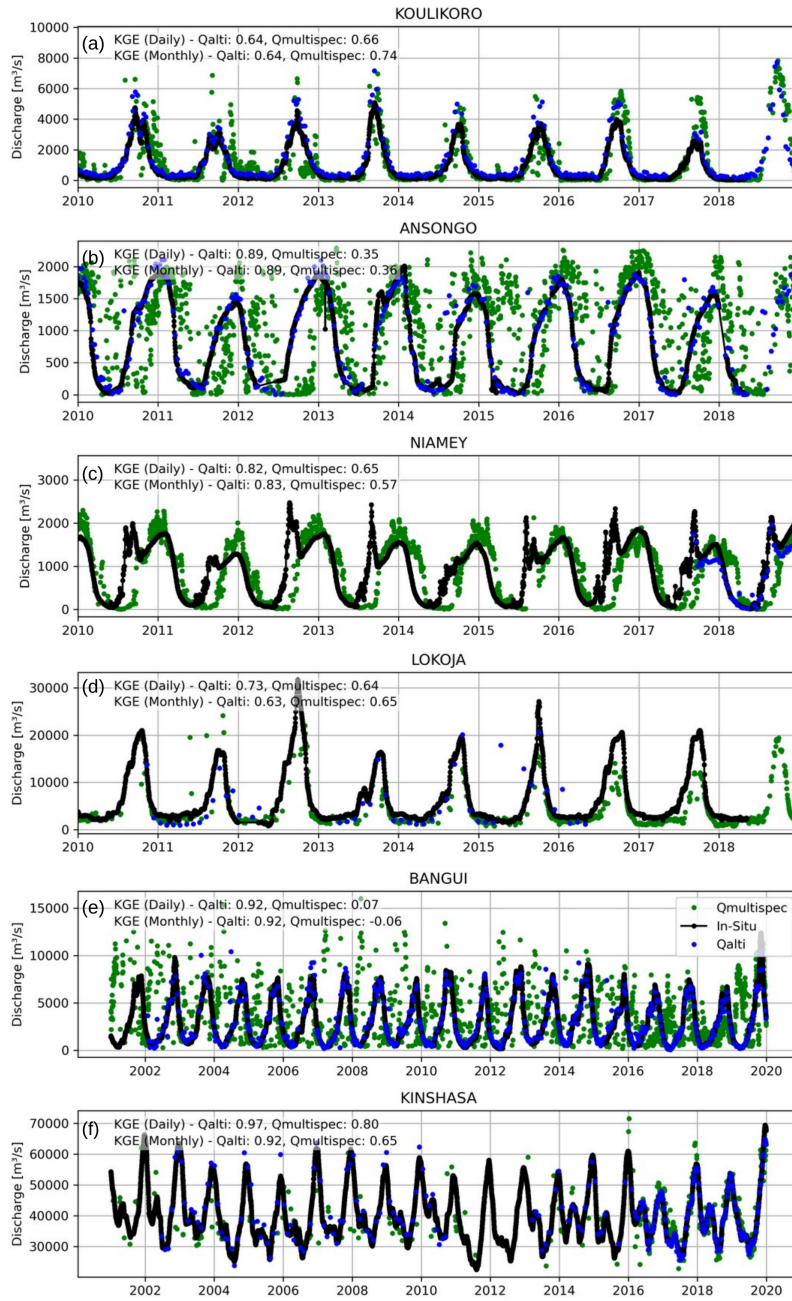


Figure S2. Discharge time series comparing altimetry-based discharge (Qalti, blue), multispectral imagery discharge (Qmultispec, green), and in-situ observations (black) at four stations in the Niger Basin (first four panels) and two stations in the Congo Basin (last two panels).

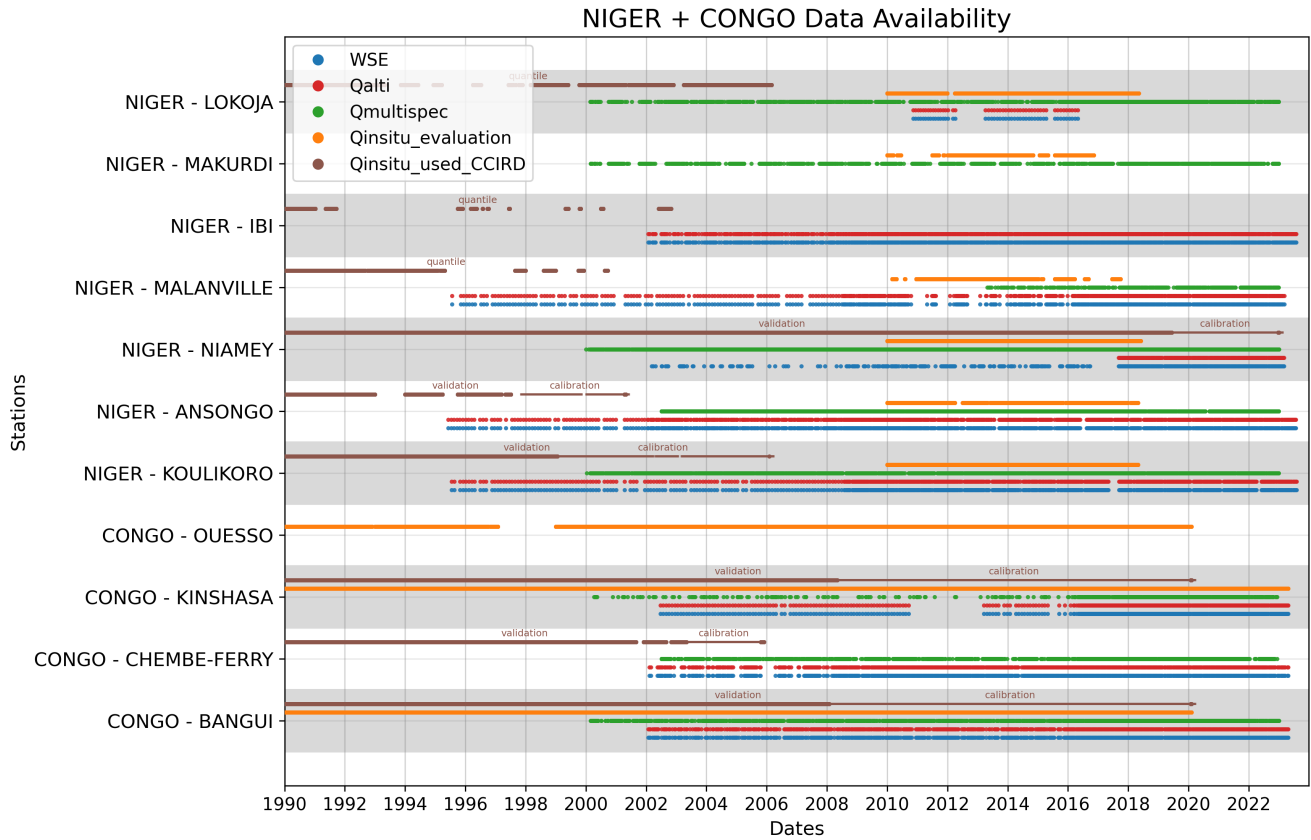


Figure S3. Temporal availability of ESA CCI products and in-situ discharge datasets used for product generation and model evaluation in the Niger and Congo basins. The figure shows, for each station, the availability of WSE, Qalti, Qmultispec, independent in-situ observations used for model evaluation, and in-situ discharge data involved in the CCI discharge rating-curve calibration and/or validation. When applicable, the calibration and validation periods used for rating-curve construction are indicated.

S3 Open-Loop performance

This section presents a detailed evaluation of the Open-Loop simulations performed with the CTRIP and MGB models over the Niger and Congo basins. These baseline runs serve as a reference for assessing the impact of the assimilation of satellite-derived observations. Model outputs are compared to independent in situ discharge records to evaluate how well each model captures seasonal flow patterns, discharge magnitudes, and to analyze performance variability across stations in both basins.

S3.1 Niger Basin

The MGB model demonstrated overall good performance across the Niger Basin in its Open-Loop mode. It generally captured the broad hydrological dynamics of the basin, with all stations—except Nantaka—showing a Nash-Sutcliffe Efficiency (NSE)

score above 0.5. The correlation coefficient exceeded 0.8 basin-wide, with a median of 0.93, indicating the model successfully reproduced flow patterns during both high and low-flow periods. However, the MGB model tends to overestimate discharge during high-water periods, particularly upstream of stations like Koulikoro and Ansongo, with a median bias around 20%. The overestimation is particularly pronounced in the Inner Niger Delta, where it reaches 40% at Nantaka.

In contrast, the CTRIP model performs poorly across most of the basin, with a median NSE score of around -4 and values consistently below 0. This poor performance is primarily due to pronounced overestimations during high-flow periods. In the region between Ke-Macina and Niamey, discharge overestimations reach up to 2.3 times the observed median values, largely due to the model's inability to account for the significant evaporation and complex hydrodynamics of the Inner Niger Delta.

Moreover, CTRIP tends to underestimate discharge during low-flow periods, particularly along the Benue tributary and near the Niger outlet at Lokoja. In contrast, the MGB model accurately captures low-flow conditions, especially in the downstream sections of the basin. However, MGB slightly underestimates internal variability, with a median KGE γ value of 0.9, a discrepancy more evident in the central basin after the river passes through the Inner Delta.

In the Upper Niger (Banankoro, Mandiana, and Koulikoro), both models generally reproduce the basin's hydrological dynamics, capturing the seasonal peaks and low flows with correct timing. However, the key issue lies in the representation of high-flow events. The MGB model performs strongly here, with NSE scores consistently above 0.65 and correlations exceeding 0.9. Meanwhile, the CTRIP model, although showing correlations above 0.8 in these stations, exhibits significant overestimation during high-flow periods, often doubling the observed high flows, as seen in panels (a) and (b) of the discharge time series (Figure S4).

In the Inner Niger Delta (Ke-Macina, Nantaka, Akka, Dire), both models encounter difficulties due to the region's complex hydrodynamics, where the river splits into multiple branches and experiences significant water losses from evaporation. The MGB model overestimates discharge by over 20% during high flows, particularly at Nantaka, where the peak discharges are nearly doubled. The CTRIP model performs even worse, with exaggerated flow biases, particularly downstream of stations like Dire and Ansongo (panels (d,e) Figure S4). The overestimation grows from twice the observed values at Ke-Macina to three times at Ansongo. The failure to represent evaporation and the delta's complex branch dynamics leads to these inflated discharge estimates in CTRIP. Additionally, CTRIP struggles to capture the gradual shift in the high-water season from September to December/January, particularly between Nantaka (panel c) and Niamey (panel f), where the correlation drops below 0.5 as the flow moves downstream.

In the Middle Niger (Ansongo, Alcongui, Niamey, Malanville), the MGB model continues to show strong performance, particularly at Niamey, where it captures the bimodal hydrological cycle driven by the Sirba tributary (panel (c), Figure S5) (NSE = 0.8, r = 0.96). However, MGB overestimates the second flow peak by 20%. The CTRIP model continues to struggle in this region, with difficulties in capturing both the magnitude and timing of high flows. In Alcongui (Gorouol tributary), both models encounter challenges due to intermittent flows driven by seasonal rainfall, with correlation scores of r close to 0.8 in MGB and r close to 0.5 in CTRIP.

In the Lower Niger (Makurdi on the Benue tributary, and Lokoja near the Niger outlet), CTRIP performs relatively well compared to its performance upstream. While it underestimates both high and low flows at Makurdi (panel g, Figure S4), the hydrological cycle is reasonably well reproduced at Lokoja (panel h), despite a systematic overestimation of peak flows. This relative improvement is largely because CTRIP's parameters—such as channel width, slope, and Manning's roughness coefficient—are globally adjusted based on empirical relationships and global datasets to ensure consistent performance at the outlets of large river basins without region-specific calibration (?). However, CTRIP still struggles with high flow overestimation and low flow underestimation. The MGB model performs better at Lokoja (panel d, Figure S5), as it is calibrated for each specific basin to accurately capture local hydrological processes. While the hydrological cycle is accurately reproduced, MGB tends to underestimate discharge during high flows, with about 10% underestimation at Lokoja, likely due to difficulties in fully representing the Benue's high flows.

S3.2 Congo Basin

Considering the significantly limited access to in-situ station data across the Congo Basin, the performance analysis of the MGB and CTRIP models remains constrained. The available in-situ data from the So-HyBam database includes only three stations, which cover two of the four main drainage systems: Ouessou on the Sangha tributary and Bangui on the Ubangui

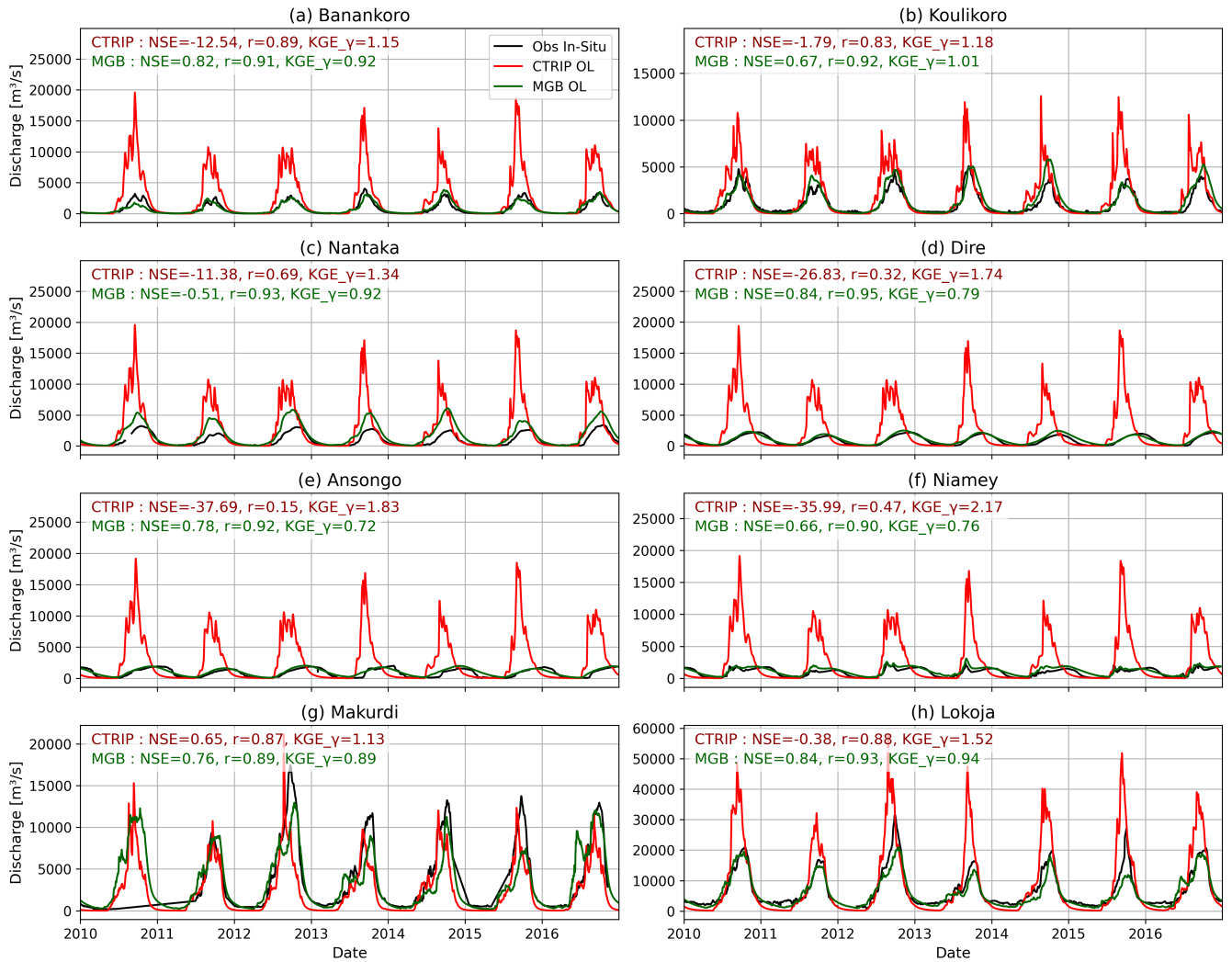


Figure S4. Discharge time series at same 8 stations within the Niger basin, comparing CTRIP (red) to MGB Open-Loop (green) and in-situ observations

tributary. The third station is located at Kinshasa, downstream of the confluence of the four tributaries, on the Congo river. The models are therefore evaluated based on their performance at these three stations.

The MGB model shows overall good performance in reproducing the hydrological dynamics of the Congo Basin, though certain biases and errors persist. With a median correlation score of 0.88, the model successfully captures the general flow patterns across the three stations. The best performance is observed at Bangui, with a correlation of 0.94, while Ouesso and Kinshasa also perform consistently, with correlation values exceeding 0.87. However, the Nash-Sutcliffe Efficiency (NSE), with a median score of 0.4, suggests moderate accuracy in predicting discharge magnitudes, pointing to areas where the model could improve, especially in terms of magnitude prediction.

On a monthly scale, the MGB model generally performs well, particularly in terms of correlation (r), internal variability (γ), and NSE. This indicates good internal consistency and a strong ability to capture the seasonal and intra-annual variability in discharge. At Bangui and Kinshasa, KGE γ monthly values are close to 1, reflecting the model's strength in reproducing

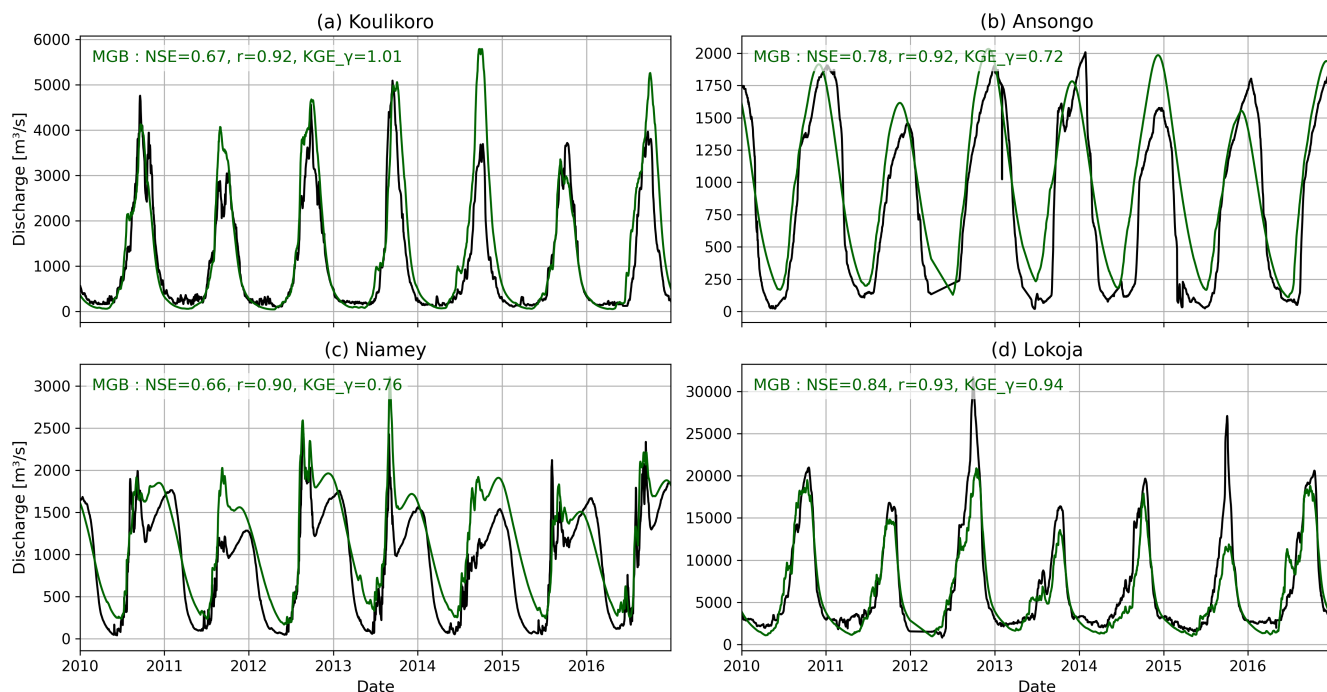


Figure S5. Discharge time series comparing MGB (green) to in-situ observations on Koulikoro, Ansongo, Niamey and Lokoja (Niger basin)

seasonal cycles (panels d,f, Figure S6). However, short-term hydrological fluctuations present more challenges, as evidenced by the increased variability in daily scores, particularly at Ouesso (panel a, Figure S6). Here, the model overestimates internal variability by around 20%, especially in the first peak of the bimodal cycle, highlighting difficulties in predicting short-term fluctuations.

In contrast, the CTRIP model exhibits weaker overall performance, particularly near the outlet at Kinshasa, where NSE score falls below -4. On the tributaries, the model shows slightly better performances, with NSE values ranging between -0.7 and 0.3. The median correlation coefficient of 0.8 suggests that while CTRIP accurately captures seasonal flow patterns, the model struggles with overall discharge magnitudes. Discharge is largely underestimated, with a median bias of 43% across the basin, particularly near Kinshasa (panels e,f, Figure S6).

At the tributary stations (Ouesso and Bangui), both models demonstrate reasonable performance in capturing the evolution of the annual hydrological cycle. The MGB model tends to overestimate discharge during high-flow periods, with overestimations of 17% at Ouesso (Sangha tributary) and 8% at Bangui (Ubangi tributary). For CTRIP, the annual cycle is well represented at both tributaries, but the model consistently underestimates discharge. The relative bias is 0.56 at Bangui and 0.32 at Ouesso, indicating a clear underestimation (panels b, d). Despite this, the correlation coefficients are satisfactory, with values close to 0.8 at both stations, reflecting a fairly accurate representation of the hydrological cycle.

At Kinshasa, the MGB model underestimates discharge by 19% across both high and low-flow periods. This underestimation is most pronounced during the second peak of the bimodal cycle (April to June), which is significantly underrepresented in the model, leading to an NSE score of 0 at Kinshasa. For CTRIP, the underestimation of discharge at Kinshasa is less pronounced compared to the tributaries, especially between February and April, when discharge is reasonably well captured. However, the model consistently underestimates discharge in the following months, with a relative bias of 0.63, likely due to underestimations of rainfall within the catchment area, leading to a shortfall in the observed water input and contributing to the systematic bias in discharge predictions at this station (Figure S6, panels e, f). Additionally, the correlation coefficient at Kinshasa drops to 0.46.

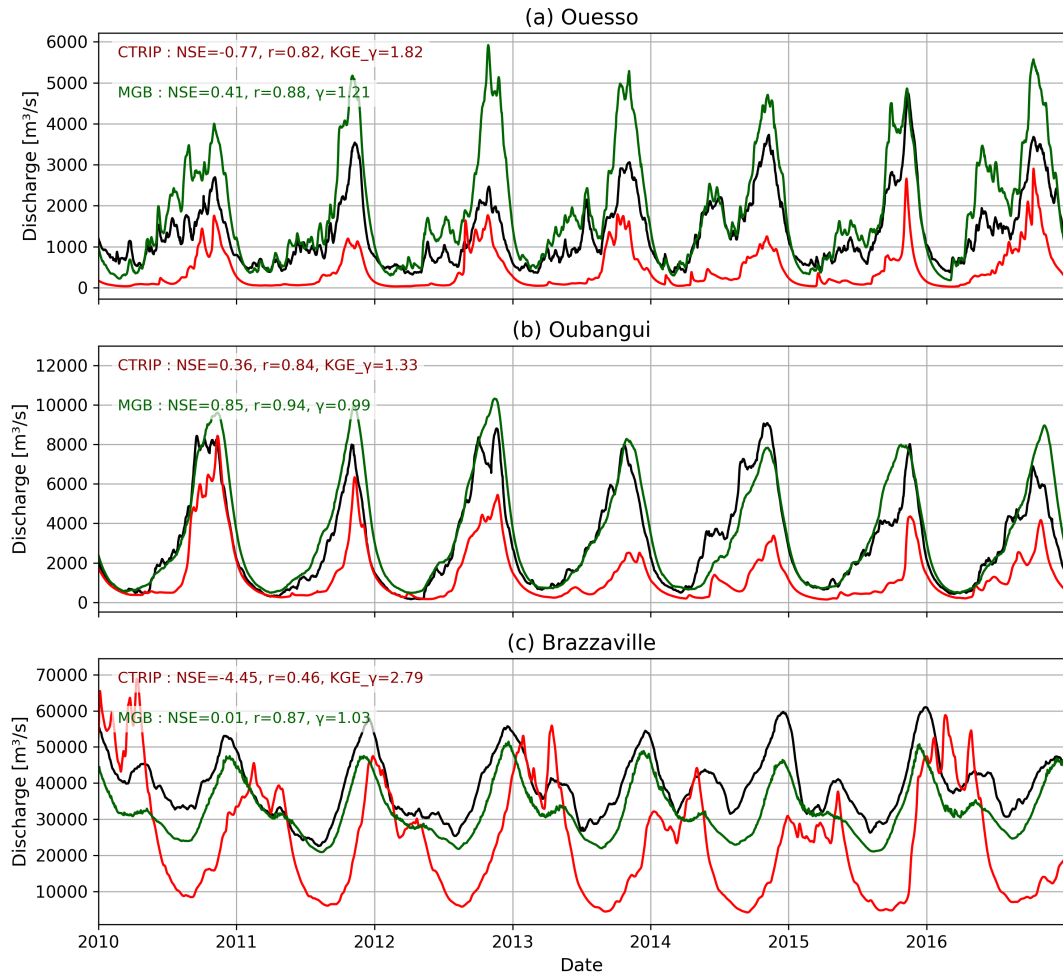


Figure S6. Discharge time series over the period 2010-2017 and climatologies comparing CTRIP (red), MGB (green), and in-situ observations (black) at three stations: Ouessou (a, b), Bangui (c, d), and Beach Brazzaville/Kinshasa (e, f)

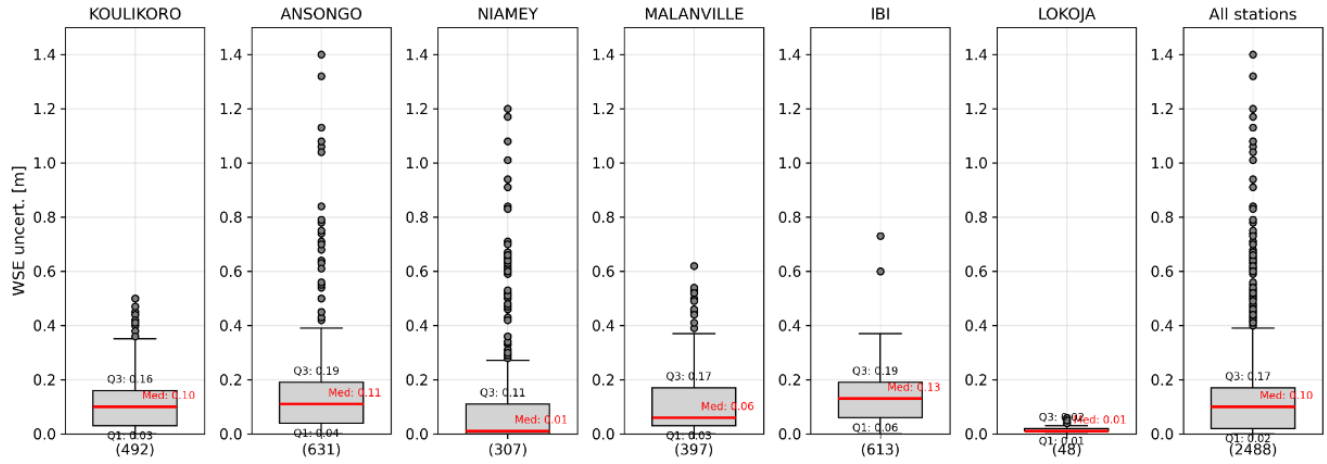
145 S4 Uncertainties of CCI products used in the study

The observation uncertainties used in the assimilation experiments were derived from the ESA CCI products, based on selected time series from 2000 to 2020. The boxplots below summarize the distributions of uncertainty estimates for both water surface elevation anomalies (WSE, in meters) and altimetry-based discharge (Qalti, relative error), computed at the station level.

Figure S7 shows that WSE uncertainties typically range from 0.01 to 0.85 m, with lower median values in the Niger basin (e.g., Koulikoro: 0.10 m, Ibi: 0.13 m) and slightly higher ones in the Congo basin (e.g., Bangui: 0.34 m, Kinshasa: 0.57 m). The 0.2 m and 0.4 m values used in our experiments were chosen to represent typical (median to third quartile) and upper-bound scenarios.

Figure S8 shows that Qalti relative uncertainties span a wide range across stations, from 4% to over 100%. While basin-scale medians are around 62% in Niger and 79% in Congo, stations such as Kinshasa, Bangui, Niamey, and Malanville exhibit much lower uncertainties. The 30% and 50% test values used in assimilation were selected to represent optimistic and conservative scenarios respectively, ensuring effective weighting in the Kalman filter.

Niger basin



Congo basin

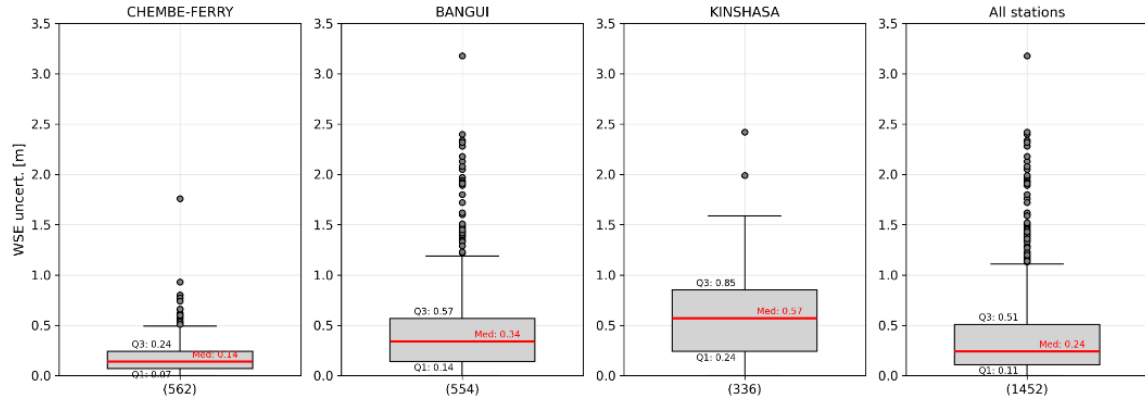


Figure S7. Distribution of computed WSE uncertainties (in meters) from 2000 to 2020 for selected Niger stations (Koulikoro, Ansongo, Niamey, Malanville, Ibi, Lokoja) and Congo stations (Chembe-Ferry, Bangui, Kinshasa), along with aggregated results for all stations.

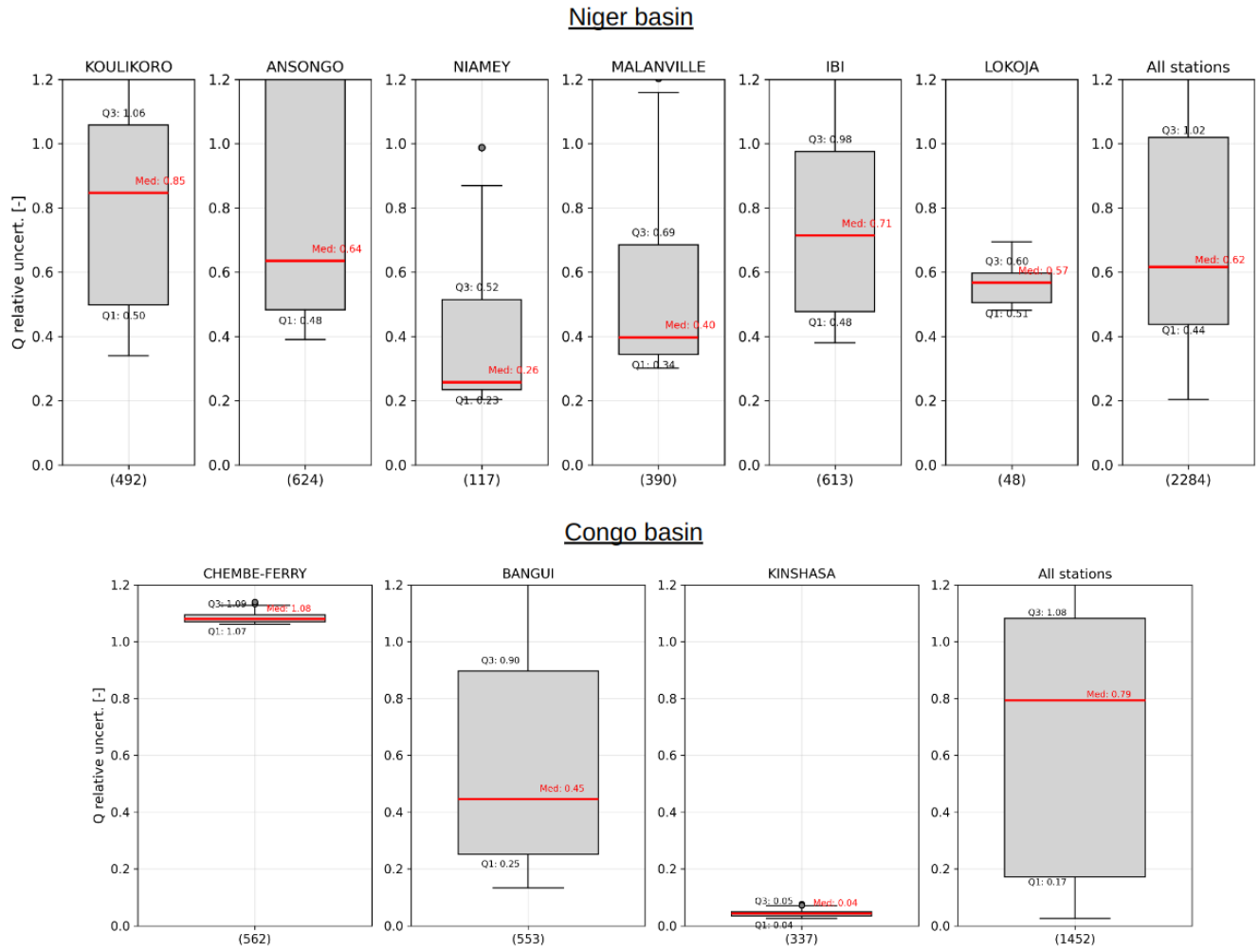


Figure S8. Distribution of relative computed uncertainties for Qalti from 2000 to 2020 across selected Niger stations (Koulikoro, Ansongo, Niamey, Malanville, Ibi, Lokoja) and Congo stations (Chembe-Ferry, Bangui, Kinshasa), along with aggregated results for all stations.

S5 Local patches used for MGB experiments in the Niger basin

Figure S9 illustrates the spatial localization strategy based on adaptive distances (semi-variograms) for the six key stations in the Niger basin. This approach was used to define local patches of influence during the data assimilation experiments.

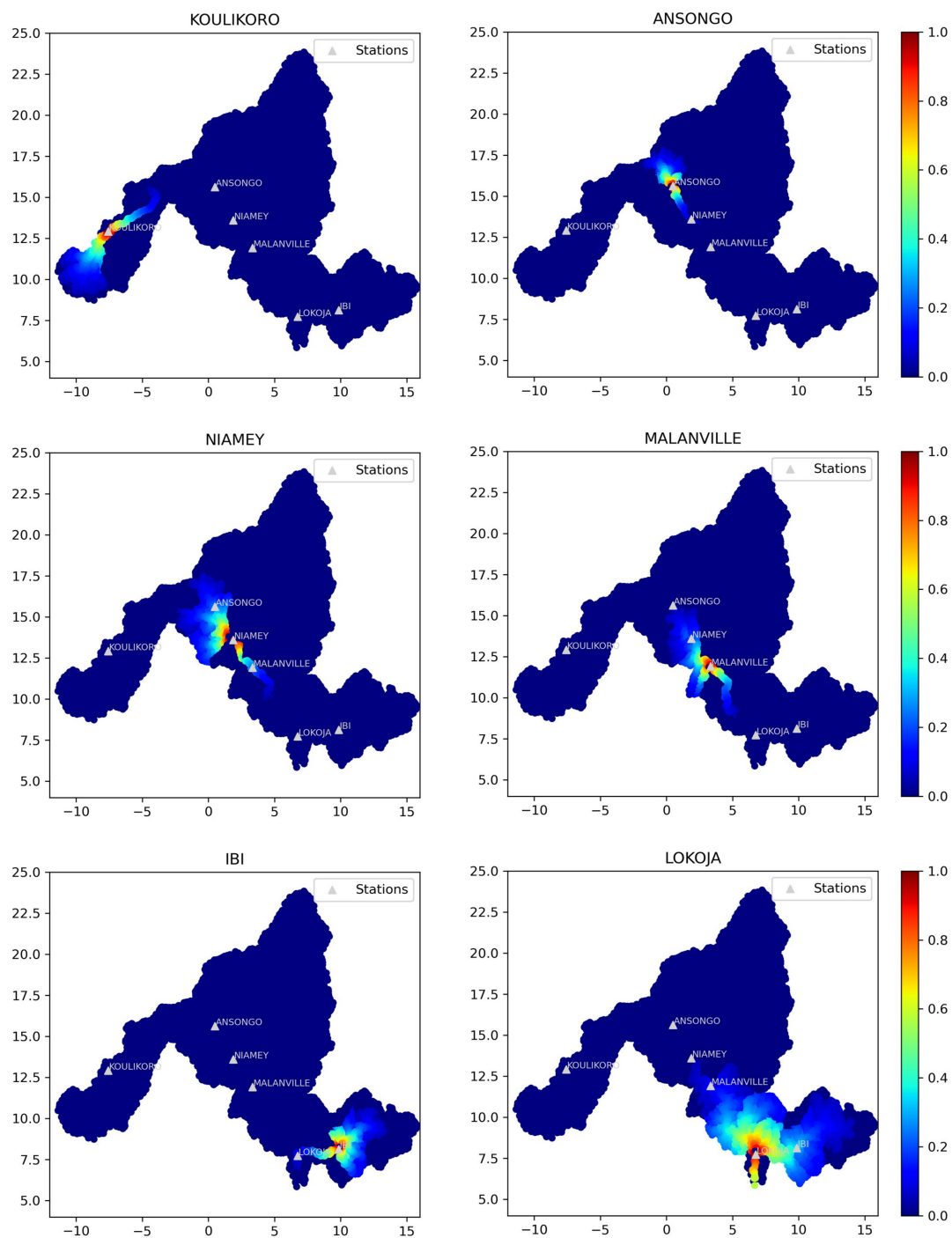


Figure S9. Local patches for Koulikoro, Ansongo, Niamey, Malanville, Ibi and Lokoja stations based on adaptive distance derived from empirical semi-variograms (Niger Basin, MGB-HYFAA model).

Figure S10 shows the rating curves at Ibi and Lokoja stations comparing open-loop MGB simulations and ESA CCI-derived observations (dH and Qalti). These comparisons support the interpretation of assimilation results and shows the alignment between simulated and observed rating relationships.

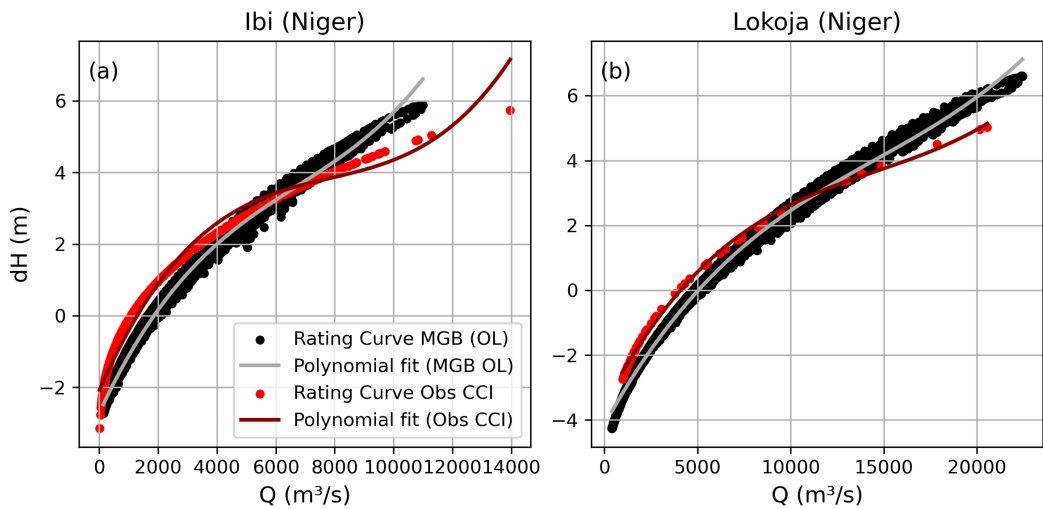


Figure S10. Rating curves from MGB (Open-Loop) versus ESA CCI-derived observations (dH, Qalti) at (a) Ibi and (b) Lokoja stations in the Niger basin.

S7 Discharge climatologies at all Niger stations – MGB model

Figure S11 presents discharge climatologies at all available stations in the Niger basin, ordered from upstream to downstream. Observed discharges are compared with MGB open-loop simulations and the full set of assimilation experiments (dH, Qalti, and Qmultispec), with different uncertainty levels shown via shaded colors for each experiment.

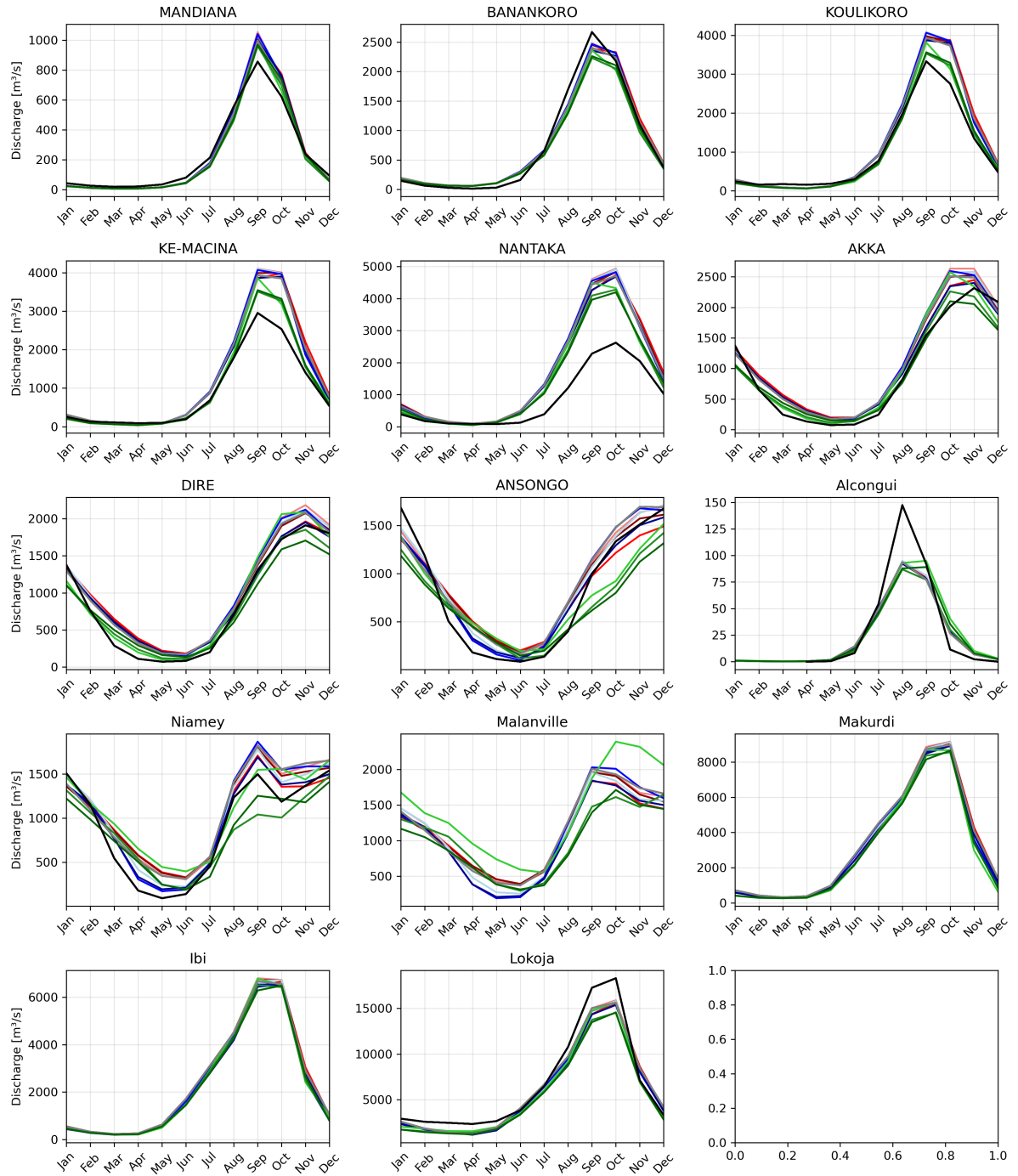


Figure S11. Discharge climatologies at all available stations in the Niger Basin, sorted from upstream to downstream. Shown are in-situ observations (black), MGB open-loop simulations (grey), and data assimilation experiments: dH assimilation (red), Qalti (blue), and Qmultispec (green), with different shades indicating uncertainty levels.

S8 Tables of assimilation results scores

170 Tables 1 and 2 summarize the distribution of performance metrics obtained for all data assimilation experiments conducted with the MGB and CTRIP models over the Niger and Congo basins, respectively. For each experiment (open-loop and assimilation runs), the tables provide the minimum, first quartile (Q1), median, third quartile (Q3) and maximum values of the key evaluation scores: Nash-Sutcliffe Efficiency (NSE), correlation (KGE r), bias (KGE β) and internal variability (KGE γ).

These quantitative statistics complement Figures 5 and 6 of the manuscript.

Table S1: Summary statistics (min, Q1, median, Q3, max) of model performance metrics for the Niger basin.

Model	Score	Experiment	Min	Q1	Median	Q3	Max
MGB	Bias (KGE β)	OL	0.87	1.03	1.21	1.25	1.93
MGB	Bias (KGE β)	dh_0.2	0.83	1.06	1.14	1.24	1.95
MGB	Bias (KGE β)	dh_0.4	0.82	1.07	1.18	1.25	1.95
MGB	Bias (KGE β)	dh_cci	0.83	1.07	1.20	1.27	1.96
MGB	Bias (KGE β)	qalti_0.3	0.83	1.06	1.15	1.26	1.92
MGB	Bias (KGE β)	qalti_0.5	0.82	1.04	1.10	1.22	1.90
MGB	Bias (KGE β)	qalti_cci	0.84	1.06	1.17	1.26	1.94
MGB	Bias (KGE β)	qmultispec_0.3	0.81	0.95	1.04	1.13	1.68
MGB	Bias (KGE β)	qmultispec_0.5	0.77	0.90	1.00	1.02	1.60
MGB	Internal Variability (KGE γ)	OL	0.60	0.76	0.92	1.01	1.16
MGB	Internal Variability (KGE γ)	dh_0.2	0.60	0.72	0.90	0.98	1.17
MGB	Internal Variability (KGE γ)	dh_0.4	0.59	0.75	0.93	1.00	1.18
MGB	Internal Variability (KGE γ)	dh_cci	0.61	0.79	0.94	1.00	1.20
MGB	Internal Variability (KGE γ)	qalti_0.3	0.61	0.81	0.93	1.01	1.17
MGB	Internal Variability (KGE γ)	qalti_0.5	0.59	0.79	0.92	1.00	1.18
MGB	Internal Variability (KGE γ)	qalti_cci	0.61	0.82	0.93	1.01	1.18
MGB	Internal Variability (KGE γ)	qmultispec_0.3	0.62	0.94	1.00	1.14	1.62
MGB	Internal Variability (KGE γ)	qmultispec_0.5	0.60	0.89	1.00	1.14	1.76
MGB	Correlation (KGE r)	OL	0.82	0.91	0.93	0.94	0.95
MGB	Correlation (KGE r)	dh_0.2	0.84	0.90	0.92	0.94	0.95
MGB	Correlation (KGE r)	dh_0.4	0.83	0.90	0.92	0.93	0.94
MGB	Correlation (KGE r)	dh_cci	0.84	0.92	0.93	0.94	0.95
MGB	Correlation (KGE r)	qalti_0.3	0.85	0.93	0.94	0.94	0.96
MGB	Correlation (KGE r)	qalti_0.5	0.83	0.93	0.94	0.95	0.96
MGB	Correlation (KGE r)	qalti_cci	0.84	0.92	0.94	0.94	0.96
MGB	Correlation (KGE r)	qmultispec_0.3	0.39	0.85	0.89	0.92	0.94
MGB	Correlation (KGE r)	qmultispec_0.5	0.43	0.85	0.89	0.93	0.95
MGB	NSE	OL	-0.51	0.57	0.77	0.82	0.84
MGB	NSE	dh_0.2	-0.48	0.61	0.76	0.82	0.84
MGB	NSE	dh_0.4	-0.63	0.60	0.74	0.79	0.85
MGB	NSE	dh_cci	-0.72	0.62	0.74	0.80	0.86
MGB	NSE	qalti_0.3	-0.51	0.64	0.78	0.83	0.86
MGB	NSE	qalti_0.5	-0.34	0.65	0.83	0.86	0.88
MGB	NSE	qalti_cci	-0.62	0.64	0.78	0.83	0.85
MGB	NSE	qmultispec_0.3	-2.85	0.58	0.68	0.80	0.86
MGB	NSE	qmultispec_0.5	-1.69	0.58	0.72	0.80	0.90
CTRIP	Bias (KGE β)	OL	0.00	1.56	2.29	2.81	3.10
CTRIP	Bias (KGE β)	dh_0.2	0.00	1.53	2.15	2.69	2.82

CTRIP	Bias (KGE β)	dh_0.4	0.00	1.53	2.17	2.70	2.84
CTRIP	Bias (KGE β)	dh_cci	0.00	1.53	2.18	2.70	2.84
CTRIP	Bias (KGE β)	qalti_0.3	0.00	1.53	2.20	2.76	2.88
CTRIP	Bias (KGE β)	qalti_0.5	0.00	1.54	2.24	2.77	2.92
CTRIP	Bias (KGE β)	qalti_cci	0.00	1.53	2.24	2.78	2.92
CTRIP	Bias (KGE β)	qmultispec_0.3	0.00	1.16	1.78	2.45	2.58
CTRIP	Bias (KGE β)	qmultispec_0.5	0.00	1.23	1.81	2.46	2.63
CTRIP	Internal Variability (KGE γ)	OL	0.86	1.15	1.43	1.63	2.17
CTRIP	Internal Variability (KGE γ)	dh_0.2	0.86	1.15	1.40	1.61	2.12
CTRIP	Internal Variability (KGE γ)	dh_0.4	0.86	1.15	1.40	1.62	2.13
CTRIP	Internal Variability (KGE γ)	dh_cci	0.86	1.14	1.38	1.61	2.07
CTRIP	Internal Variability (KGE γ)	qalti_0.3	0.86	1.16	1.43	1.63	2.21
CTRIP	Internal Variability (KGE γ)	qalti_0.5	0.86	1.16	1.43	1.63	2.20
CTRIP	Internal Variability (KGE γ)	qalti_cci	0.86	1.16	1.43	1.63	2.19
CTRIP	Internal Variability (KGE γ)	qmultispec_0.3	0.86	1.19	1.46	1.70	2.52
CTRIP	Internal Variability (KGE γ)	qmultispec_0.5	0.86	1.19	1.46	1.70	2.47
CTRIP	Correlation (KGE r)	OL	0.15	0.44	0.69	0.83	0.88
CTRIP	Correlation (KGE r)	dh_0.2	0.15	0.44	0.69	0.81	0.87
CTRIP	Correlation (KGE r)	dh_0.4	0.15	0.44	0.69	0.82	0.87
CTRIP	Correlation (KGE r)	dh_cci	0.15	0.45	0.69	0.81	0.87
CTRIP	Correlation (KGE r)	qalti_0.3	0.15	0.44	0.69	0.83	0.87
CTRIP	Correlation (KGE r)	qalti_0.5	0.15	0.44	0.69	0.83	0.87
CTRIP	Correlation (KGE r)	qalti_cci	0.15	0.44	0.69	0.83	0.87
CTRIP	Correlation (KGE r)	qmultispec_0.3	0.15	0.42	0.69	0.83	0.86
CTRIP	Correlation (KGE r)	qmultispec_0.5	0.16	0.43	0.69	0.83	0.87
CTRIP	NSE	OL	-37.69	-22.97	-3.49	-0.78	-0.33
CTRIP	NSE	dh_0.2	-29.64	-20.12	-3.02	-0.78	0.02
CTRIP	NSE	dh_0.4	-30.24	-20.48	-3.05	-0.78	-0.01
CTRIP	NSE	dh_cci	-29.55	-20.06	-3.03	-0.78	0.03
CTRIP	NSE	qalti_0.3	-32.46	-21.99	-3.28	-0.78	-0.06
CTRIP	NSE	qalti_0.5	-33.21	-22.32	-3.35	-0.78	-0.10
CTRIP	NSE	qalti_cci	-33.21	-22.34	-3.35	-0.78	-0.07
CTRIP	NSE	qmultispec_0.3	-22.33	-11.12	-2.81	-0.78	0.24
CTRIP	NSE	qmultispec_0.5	-23.14	-11.78	-2.91	-0.78	0.22

Table S2: Summary statistics (min, Q1, median, Q3, max) of model performance metrics for the Congo basin.

Model	Score	Experiment	Min	Q1	Median	Q3	Max
MGB	Bias (KGE β)	OL	0.80	0.81	1.08	1.17	1.20
MGB	Bias (KGE β)	dh_0.2	-2.00	-1.74	-1.48	-0.20	1.09
MGB	Bias (KGE β)	dh_0.4	0.43	0.58	0.72	0.90	1.08
MGB	Bias (KGE β)	dh_cci	0.17	0.46	0.76	0.92	1.09
MGB	Bias (KGE β)	qalti_0.3	0.81	0.94	1.06	1.12	1.18
MGB	Bias (KGE β)	qalti_0.5	0.81	0.94	1.06	1.12	1.18
MGB	Bias (KGE β)	qalti_cci	0.82	0.94	1.06	1.11	1.16
MGB	Bias (KGE β)	qmultispec_0.3	0.78	0.81	0.84	1.01	1.18
MGB	Bias (KGE β)	qmultispec_0.5	0.79	0.84	0.89	1.03	1.16
MGB	Internal Variability (KGE γ)	OL	0.99	1.02	1.03	1.21	1.24

MGB	Internal Variability (KGE γ)	dh_0.2	-537.69	-368.18	-198.68	-98.86	0.95
MGB	Internal Variability (KGE γ)	dh_0.4	0.95	40.59	80.23	215.49	350.75
MGB	Internal Variability (KGE γ)	dh_cci	0.95	24.79	48.63	522.31	995.98
MGB	Internal Variability (KGE γ)	qalti_0.3	1.00	1.01	1.02	1.12	1.21
MGB	Internal Variability (KGE γ)	qalti_0.5	1.00	1.01	1.03	1.12	1.21
MGB	Internal Variability (KGE γ)	qalti_cci	0.99	1.01	1.04	1.12	1.21
MGB	Internal Variability (KGE γ)	qmultispec_0.3	0.96	0.97	0.98	1.10	1.22
MGB	Internal Variability (KGE γ)	qmultispec_0.5	0.97	0.98	0.98	1.10	1.22
MGB	Correlation (KGE r)	OL	0.86	0.87	0.88	0.94	0.94
MGB	Correlation (KGE r)	dh_0.2	0.00	0.01	0.02	0.49	0.96
MGB	Correlation (KGE r)	dh_0.4	0.02	0.02	0.02	0.49	0.96
MGB	Correlation (KGE r)	dh_cci	0.03	0.03	0.04	0.50	0.96
MGB	Correlation (KGE r)	qalti_0.3	0.87	0.88	0.88	0.91	0.94
MGB	Correlation (KGE r)	qalti_0.5	0.87	0.87	0.88	0.91	0.94
MGB	Correlation (KGE r)	qalti_cci	0.88	0.88	0.88	0.92	0.95
MGB	Correlation (KGE r)	qmultispec_0.3	0.86	0.87	0.88	0.89	0.90
MGB	Correlation (KGE r)	qmultispec_0.5	0.86	0.87	0.88	0.90	0.92
MGB	NSE	OL	-0.05	0.01	0.41	0.85	0.86
MGB	NSE	dh_0.2	-6.35e+5	-3.96e+5	-1.57e+5	-7.87e+4	0.90
MGB	NSE	dh_0.4	-2.31e+4	-1.32e+4	-3.38e+3	-1.69e+3	0.90
MGB	NSE	dh_cci	-2.98e+4	-1.56e+4	-1.34e+3	-6.71e+2	0.90
MGB	NSE	qalti_0.3	0.01	0.20	0.40	0.64	0.87
MGB	NSE	qalti_0.5	-0.00	0.19	0.38	0.62	0.86
MGB	NSE	qalti_cci	0.11	0.28	0.44	0.66	0.88
MGB	NSE	qmultispec_0.3	-0.22	0.08	0.38	0.57	0.76
MGB	NSE	qmultispec_0.5	-0.19	0.12	0.43	0.63	0.83
CTRIIP	Bias (KGE β)	OL	0.32	0.44	0.56	0.59	0.63
CTRIIP	Bias (KGE β)	dh_0.2	0.32	0.48	0.65	0.66	0.66
CTRIIP	Bias (KGE β)	dh_0.4	0.32	0.47	0.63	0.64	0.64
CTRIIP	Bias (KGE β)	dh_cci	0.32	0.48	0.64	0.64	0.64
CTRIIP	Bias (KGE β)	qalti_0.3	0.32	0.45	0.59	0.61	0.63
CTRIIP	Bias (KGE β)	qalti_0.5	0.32	0.45	0.57	0.60	0.63
CTRIIP	Bias (KGE β)	qalti_cci	0.32	0.46	0.60	0.63	0.66
CTRIIP	Bias (KGE β)	qmultispec_0.3	0.32	0.43	0.54	0.58	0.63
CTRIIP	Bias (KGE β)	qmultispec_0.5	0.32	0.42	0.53	0.58	0.62
CTRIIP	Internal Variability (KGE γ)	OL	1.34	1.58	1.82	2.31	2.81
CTRIIP	Internal Variability (KGE γ)	dh_0.2	1.16	1.49	1.82	2.22	2.63
CTRIIP	Internal Variability (KGE γ)	dh_0.4	1.22	1.52	1.82	2.24	2.67
CTRIIP	Internal Variability (KGE γ)	dh_cci	1.18	1.50	1.82	2.24	2.66
CTRIIP	Internal Variability (KGE γ)	qalti_0.3	1.31	1.56	1.82	2.28	2.75
CTRIIP	Internal Variability (KGE γ)	qalti_0.5	1.33	1.57	1.82	2.30	2.78
CTRIIP	Internal Variability (KGE γ)	qalti_cci	1.33	1.57	1.82	2.20	2.58
CTRIIP	Internal Variability (KGE γ)	qmultispec_0.3	1.30	1.56	1.82	2.30	2.78
CTRIIP	Internal Variability (KGE γ)	qmultispec_0.5	1.31	1.56	1.82	2.31	2.80
CTRIIP	Correlation (KGE r)	OL	0.46	0.63	0.80	0.81	0.81
CTRIIP	Correlation (KGE r)	dh_0.2	0.55	0.68	0.80	0.83	0.87
CTRIIP	Correlation (KGE r)	dh_0.4	0.54	0.67	0.80	0.83	0.86
CTRIIP	Correlation (KGE r)	dh_cci	0.53	0.67	0.80	0.83	0.85
CTRIIP	Correlation (KGE r)	qalti_0.3	0.48	0.64	0.80	0.82	0.83

CTrip	Correlation (KGE r)	qalti_0.5	0.47	0.64	0.80	0.81	0.82
CTrip	Correlation (KGE r)	qalti_cci	0.52	0.66	0.80	0.82	0.84
CTrip	Correlation (KGE r)	qmultispec_0.3	0.46	0.63	0.80	0.80	0.80
CTrip	Correlation (KGE r)	qmultispec_0.5	0.46	0.63	0.80	0.80	0.80
CTrip	NSE	OL	-4.21	-2.44	-0.68	-0.16	0.36
CTrip	NSE	dh_0.2	-3.44	-2.06	-0.68	-0.06	0.56
CTrip	NSE	dh_0.4	-3.59	-2.13	-0.68	-0.08	0.52
CTrip	NSE	dh_cci	-3.60	-2.14	-0.68	-0.08	0.52
CTrip	NSE	qalti_0.3	-4.05	-2.36	-0.68	-0.12	0.43
CTrip	NSE	qalti_0.5	-4.13	-2.41	-0.68	-0.14	0.40
CTrip	NSE	qalti_cci	-3.43	-2.06	-0.68	-0.11	0.46
CTrip	NSE	qmultispec_0.3	-4.18	-2.43	-0.68	-0.19	0.30
CTrip	NSE	qmultispec_0.5	-4.25	-2.47	-0.68	-0.19	0.30
