climate change initiative

RIVER DISCHARGE

WP4 Validation

river discharge cci

Laetitia GAL & Maxime VAYRE

User Workshop

Météo-France, Toulouse

03-04 June 2024

Objective

Objective: Compare performance of the different algorithms/product and validate river discharge time series

Validate River Discharge (Lead Hydro Matters)

- **Validation:**
  - With Cal/Val in-situ data over validation period
  - With independent in-situ data
- **Errors – first prior to an end-to-end error budget:**
  - WSE errors between altimetry and in-situ data
  - Quantile approach – time lag between Q and WSE & daily vs monthly
  - Rating curve algorithm

Consistency analysis and round robin (Lead CLS)

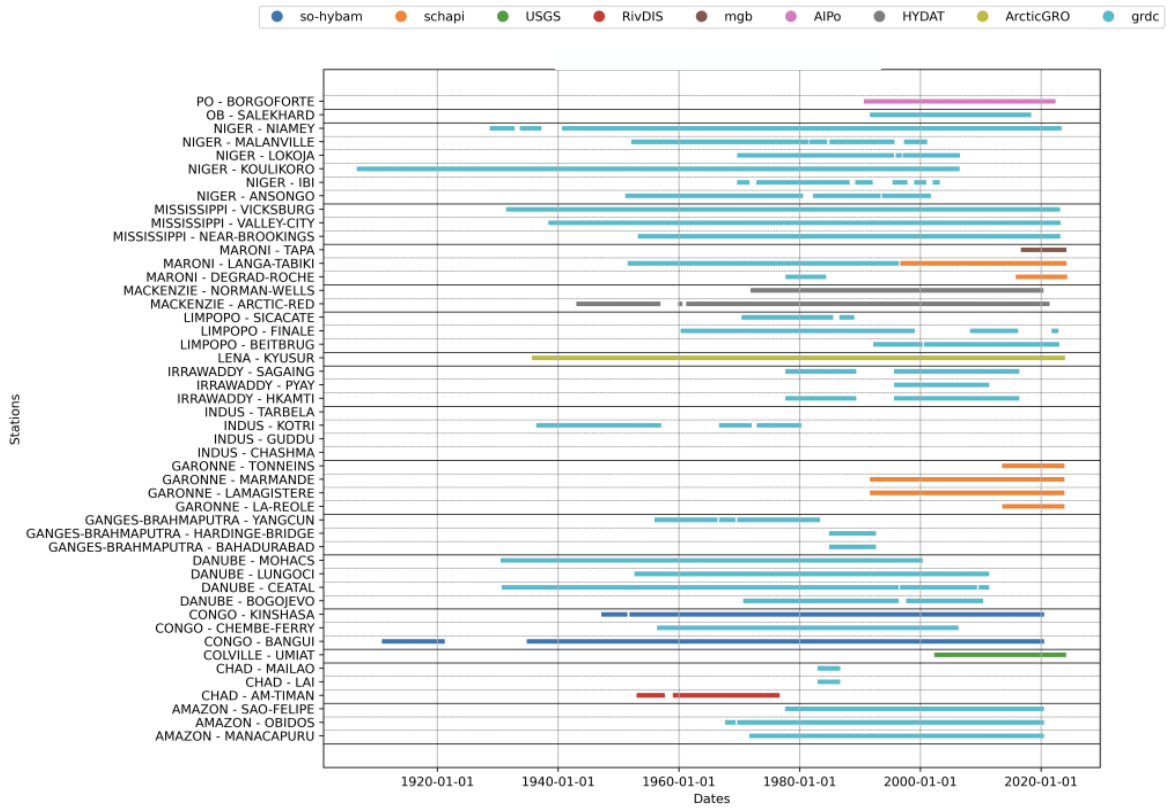
- **Format**: CCI data Standard
- **Time/space resolution**: Completeness and spatial coverage
- **Errors (in situ comparison)**: Discharge products are compared to in situ data (RMSE, Pearson, Bias, Nash, KGE)

Validation – available data

With Cal/Val in-situ data

Available in-situ discharge data for each station **used** to setup satellite-based RD methodology.

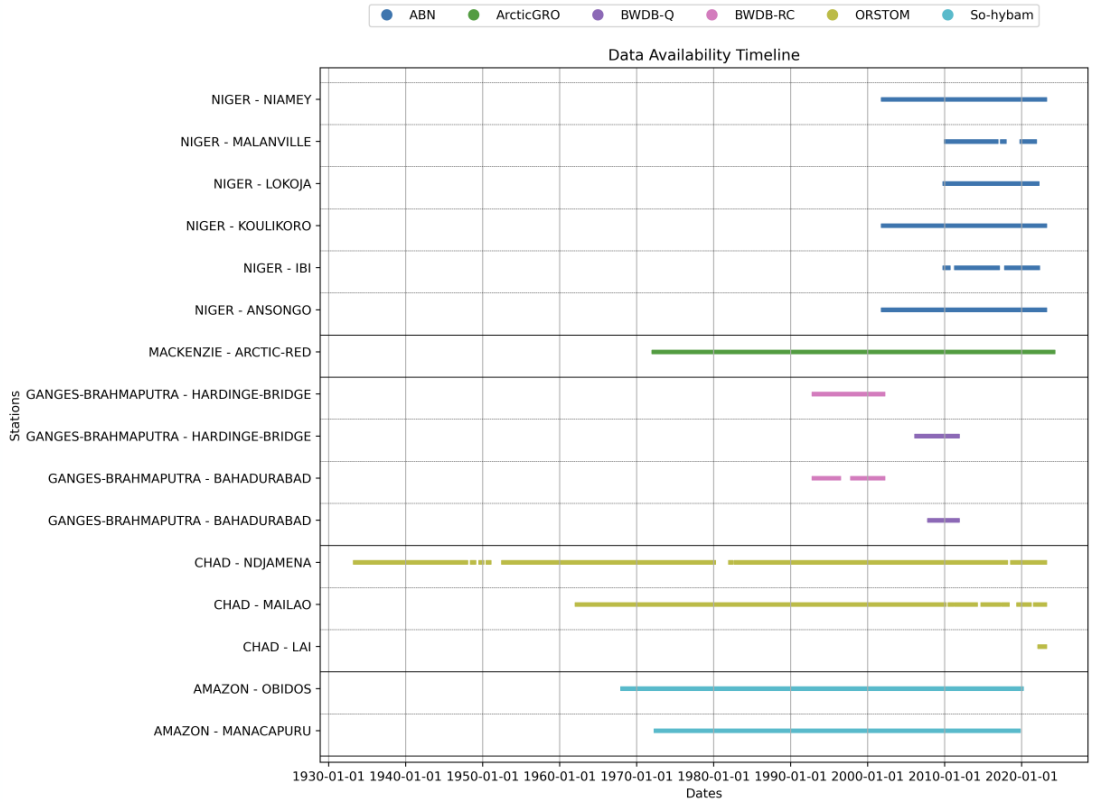
- 9 databases
- 53 stations with in-situ data



With Independent in-situ data

Available in-situ discharge data for each station **not used** to setup satellite-based RD methodology.

- 6 databases
- 16 stations with in-situ data



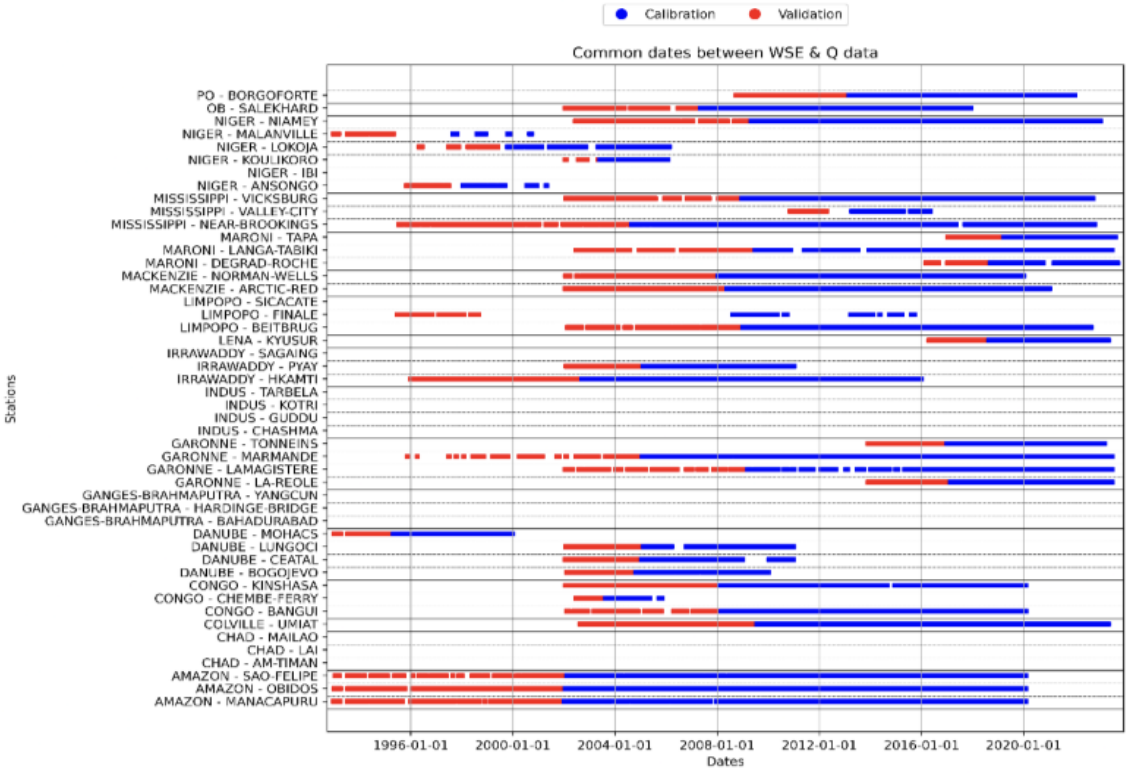
Validation - methodology

With Cal/Val in-situ data

- Identify overlap period between merge WSE from altimeters and in-situ discharge = closest date with time gap < 24H
- Divided this common period into Cal/Val periods
  - First 1/3 part = Validation period (Red)
  - Last 2/3 parts = Calibration period (Blue)

Validation period

All period (cal/val)



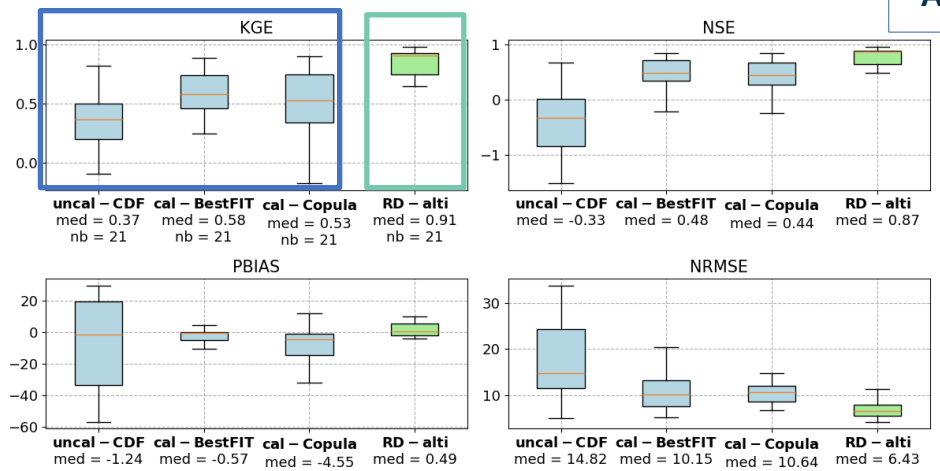
With Independent in-situ data

- Identify overlap period between satellite-based RD products independent in-situ discharge = closest date with time gap < 24H
  - **Over all available stations per products**
  - **Over common stations between products**

Validation - results

With Cal/Val in-situ data

RD-multi *RD-alti*



All period (cal/val)

Over all period (at least 20 years) we observed a **very good efficiency** over calibrated methods with **KGE > 0.5 and NRMSE < 11%**

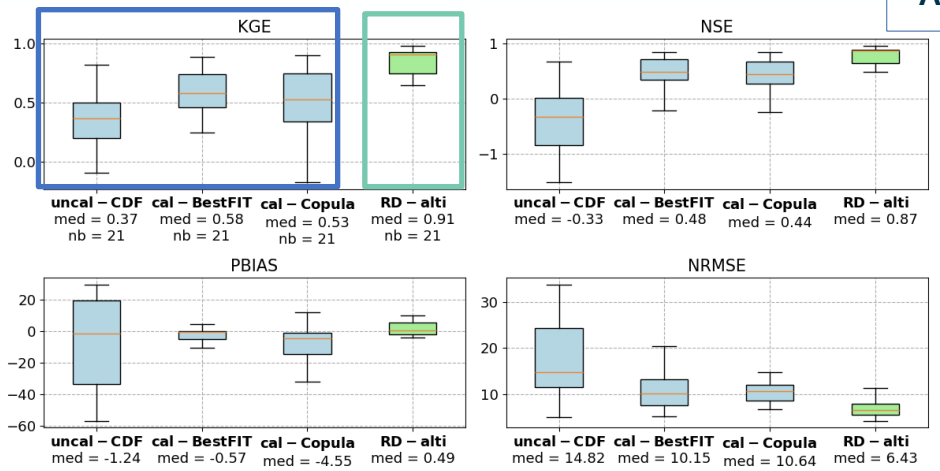
With Independent in-situ data

Validation - results

With Cal/Val in-situ data

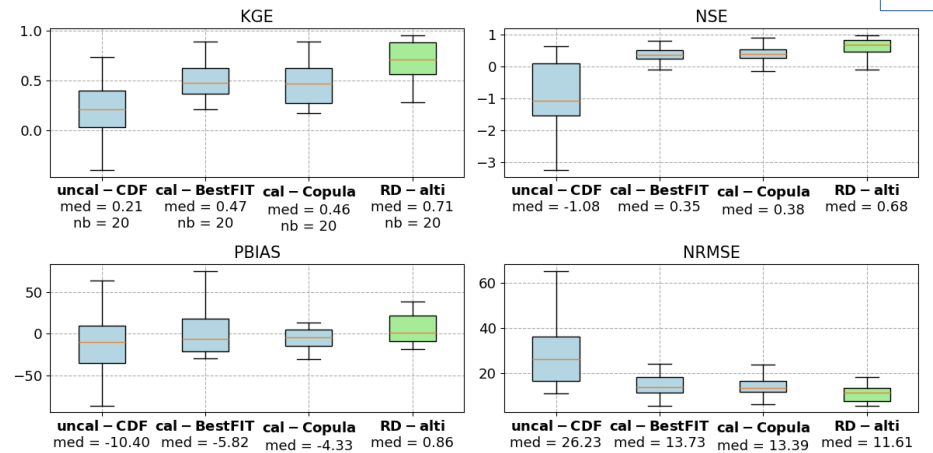
RD-multi *RD-alti*

All period (cal/val)



Over all period (at least 20 years) we observed a **very good efficiency** over calibrated methods with **KGE > 0.5 and NRMSE < 11%**

Validation period



Validation period **largely affected by old/less accurate missions** (MODIS, LandSAT, T/P, Envisat, ERS2) <u>BUT</u> still **good efficiency** : median NRSME < 15%

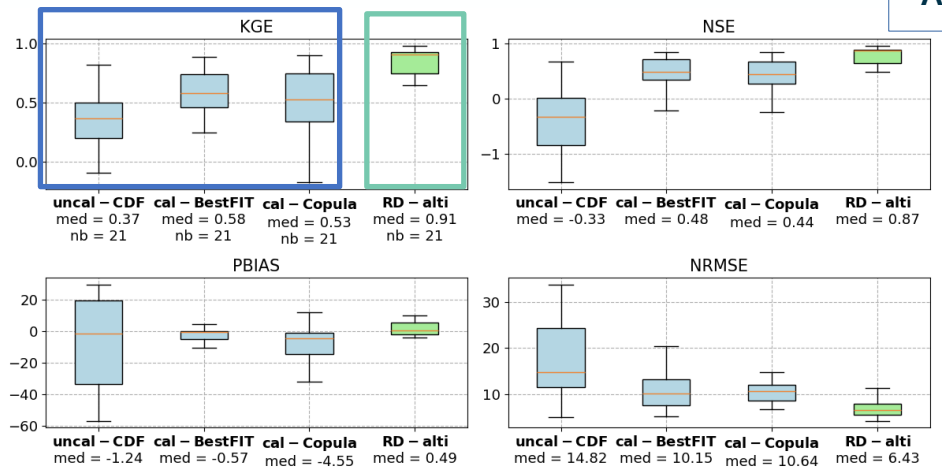
With Independent in-situ data

Validation - results

With Cal/Val in-situ data

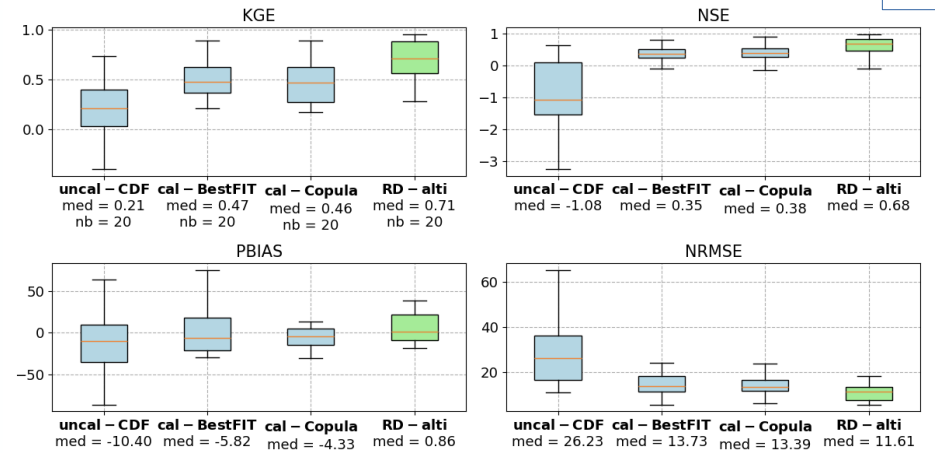
RD-multi *RD-alti*

All period (cal/val)



Over all period (at least 20 years) we observed a **very good efficiency** over calibrated methods with **KGE > 0.5 and NRMSE < 11%**

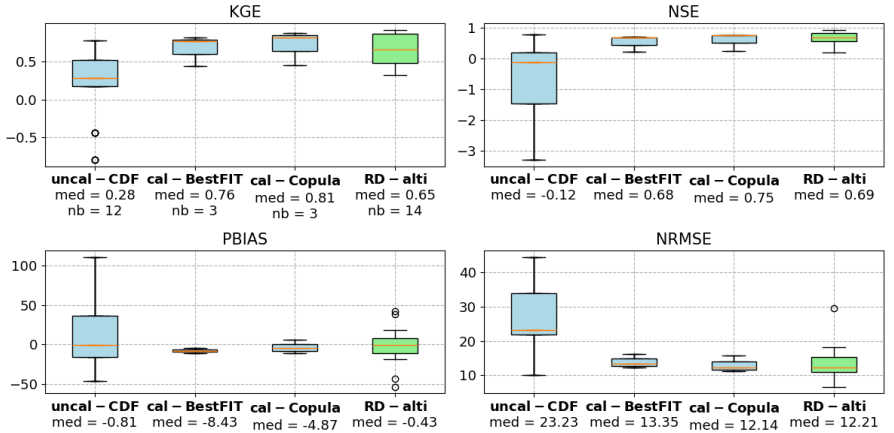
Validation period



Validation period **largely affected by old/less accurate missions** (MODIS, LandSAT, T/P, Envisat, ERS2) BUT still **good efficiency** : median NRSME < 15%

With Independent in-situ data

Over all available stations per products



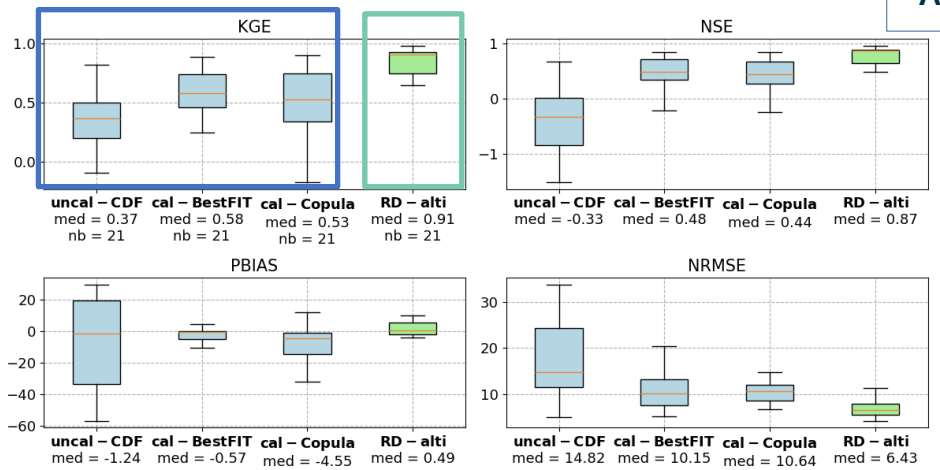
- **High disparity for uncalibrated** method for multi-based RD
- Over calibrated methods: very good efficiency with **KGE >0.75 and NRMSE < 13.5%**

Validation - results

With Cal/Val in-situ data

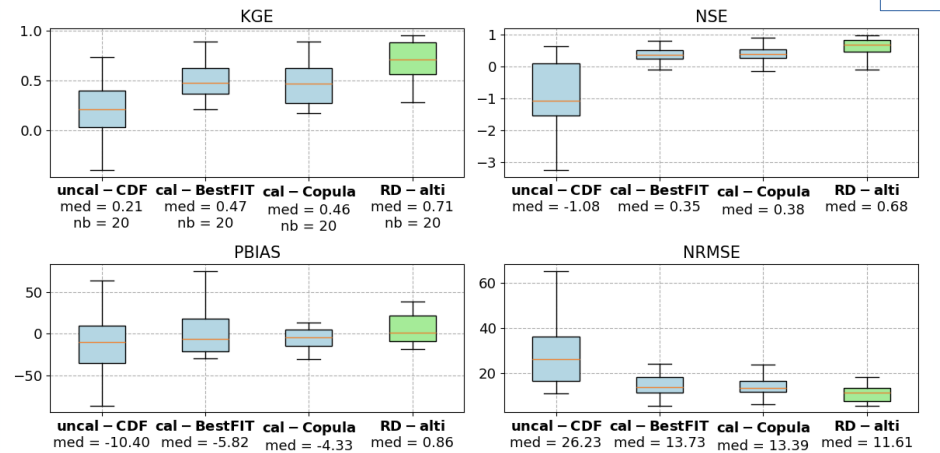
RD-multi *RD-alti*

All period (cal/val)



Over all period (at least 20 years) we observed a **very good efficiency** over calibrated methods with **KGE > 0.5 and NRMSE < 11%**

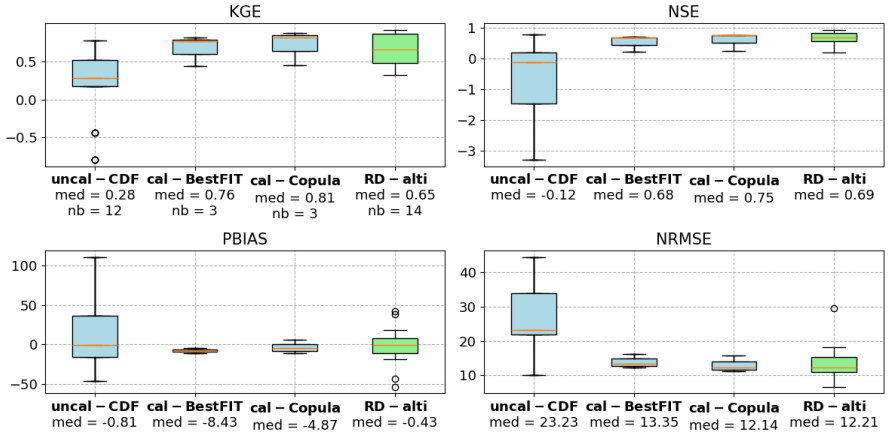
Validation period



Validation period **largely affected by old/less accurate missions** (MODIS, LandSAT, T/P, Envisat, ERS2) <u>BUT</u> still **good efficiency** : median NRSME < 15%

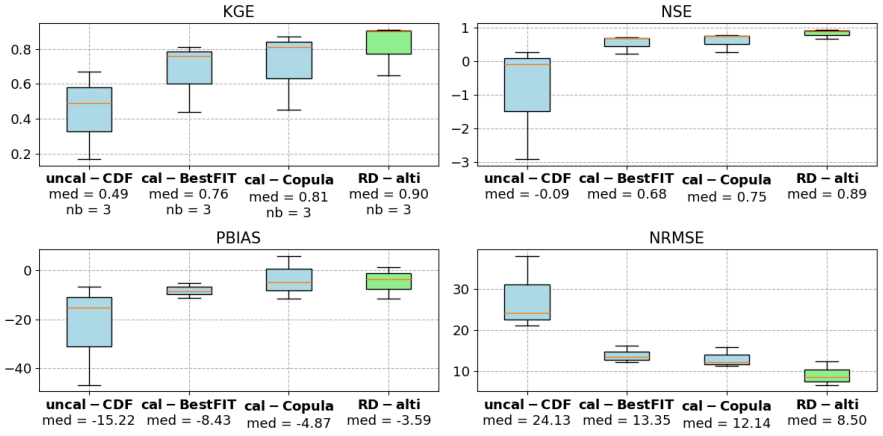
With Independent in-situ data

Over all available stations per products



- **High disparity for uncalibrated** method for multi-based RD
- Over calibrated methods: very good efficiency with **KGE >0.75 and NRMSE < 13.5%**

Over common stations between products = 3 stations

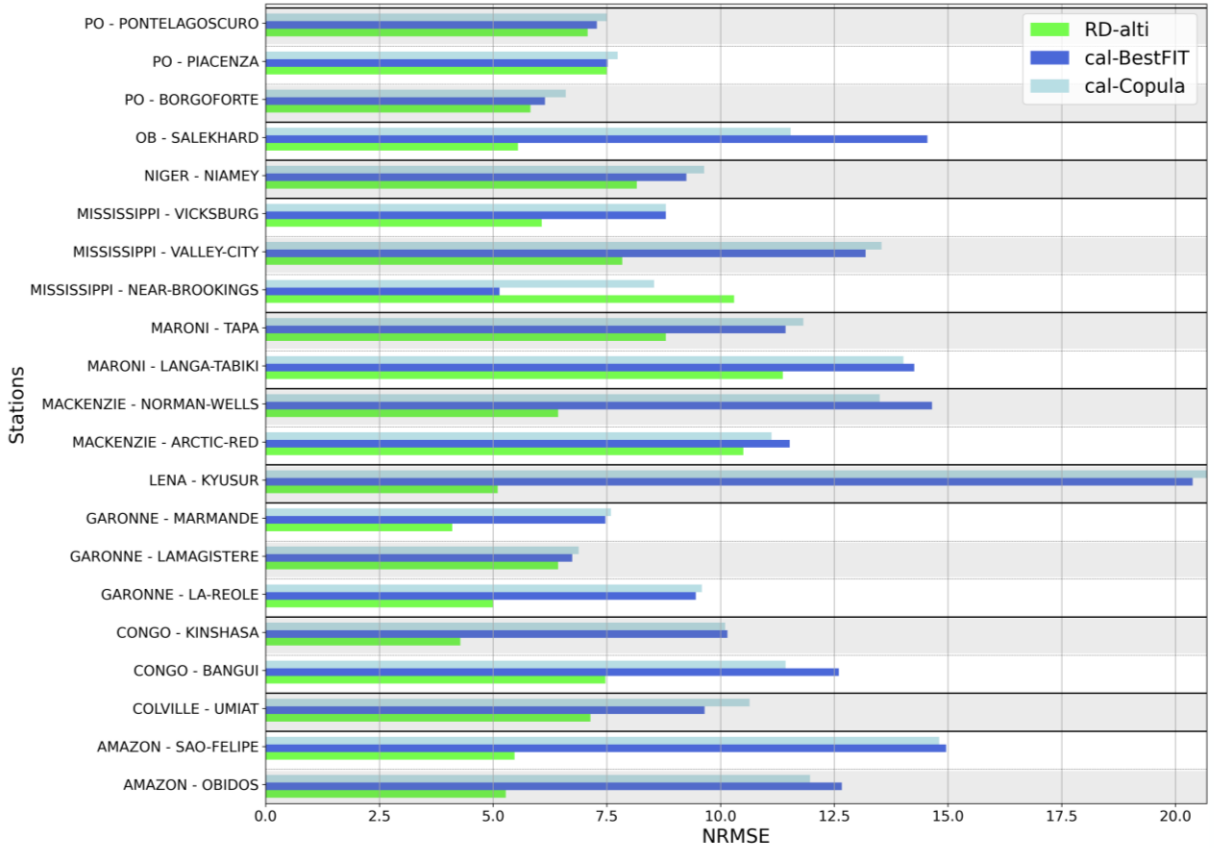


- For the 3 common stations, the same analyse can be made than before

Validation - results

With Cal/Val in-situ data

All period over common stations between calibrated RD products



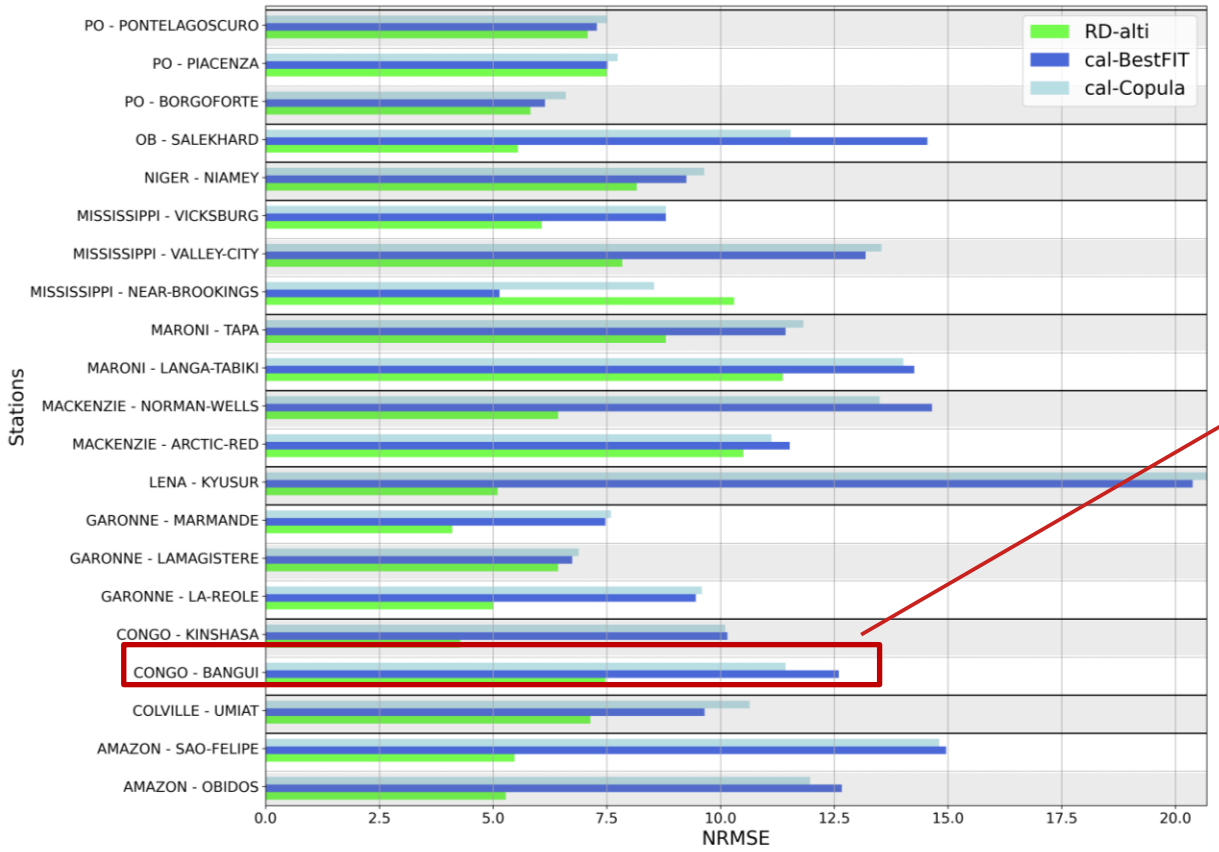
- Over satellite-based RD-cal products (21 stations)

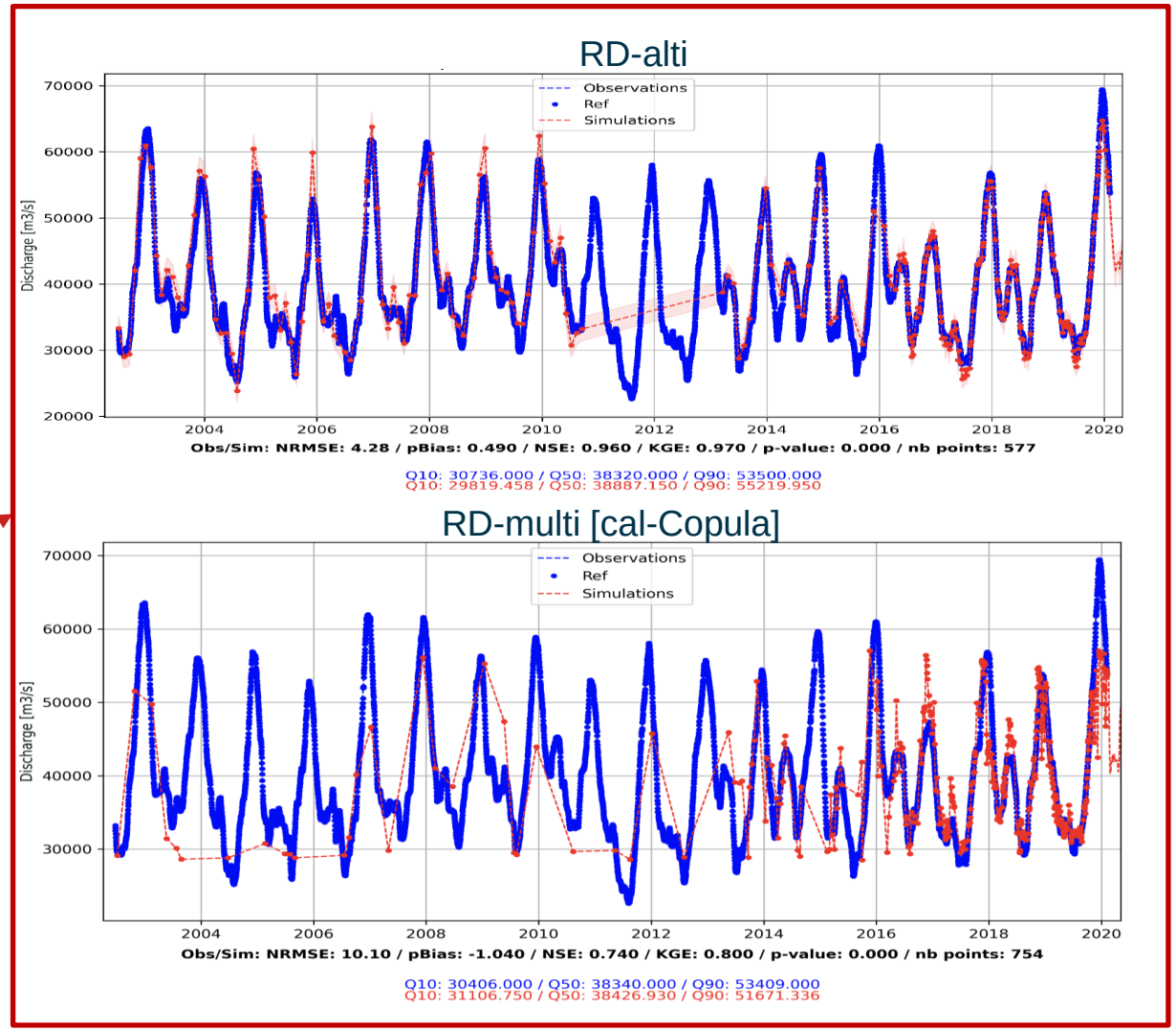
With Independent in-situ data

Validation - results

With Cal/Val in-situ data

All period over common stations between calibrated RD products



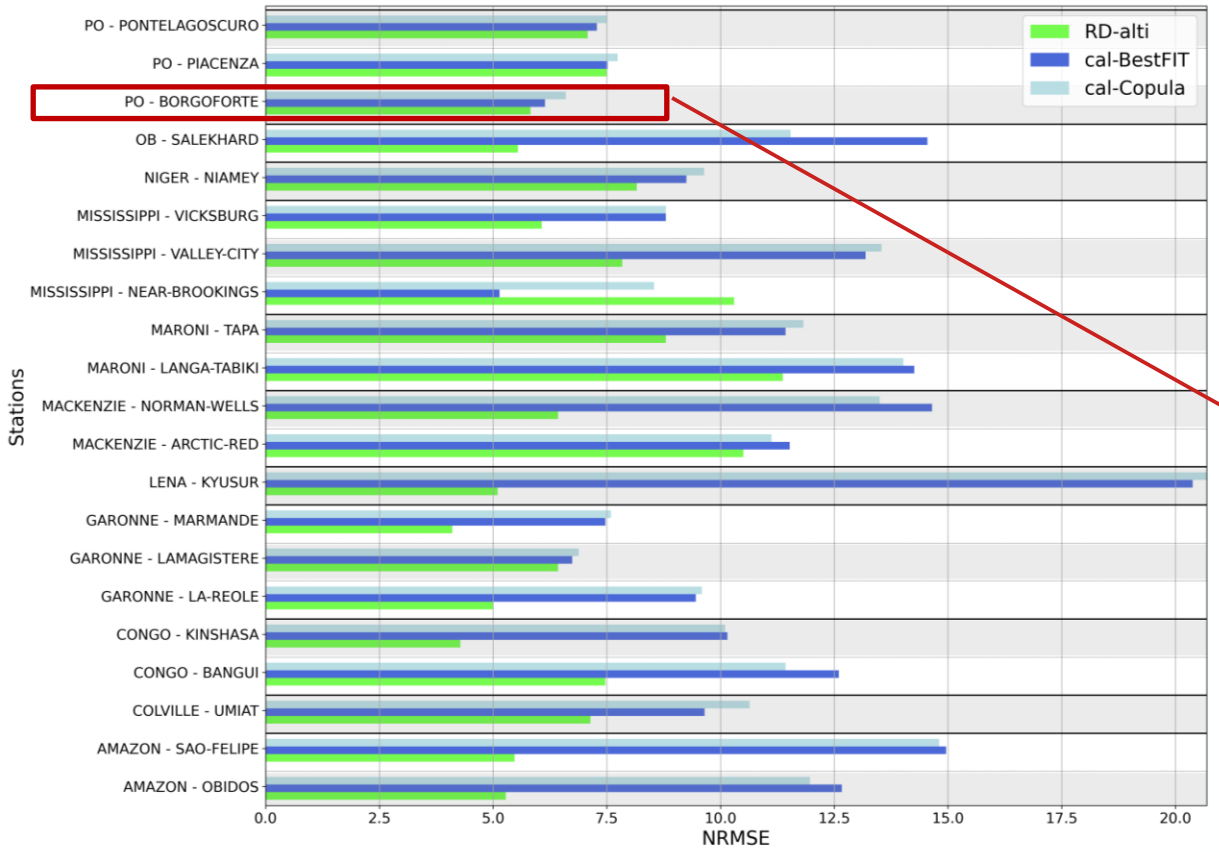


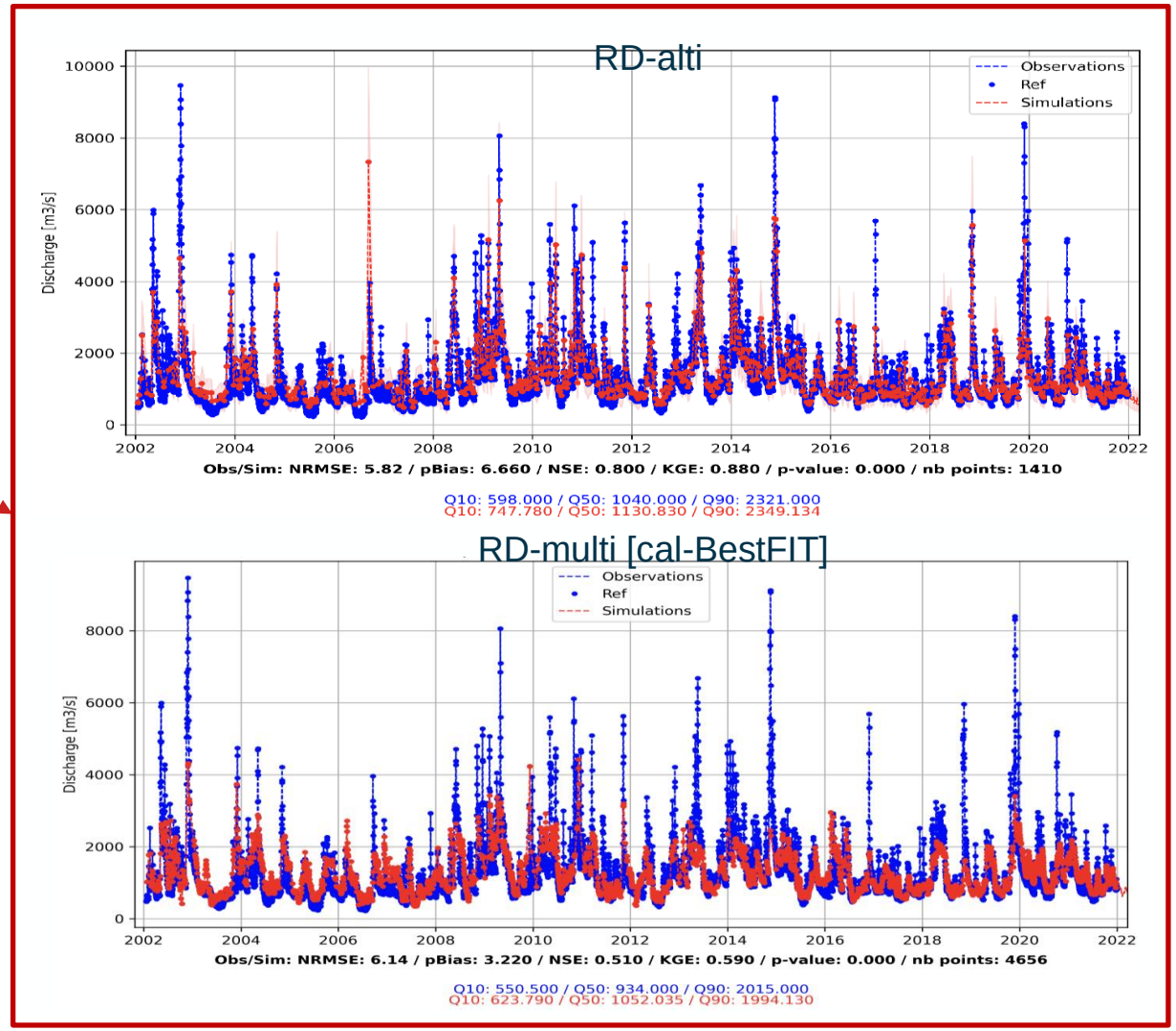
- Over satellite-based RD-cal products (21 stations) : **NRMSE < 15 %**
- RD-multi able to **add some points where alti is not available**

Validation - results

With Cal/Val in-situ data

All period over common stations between calibrated RD products



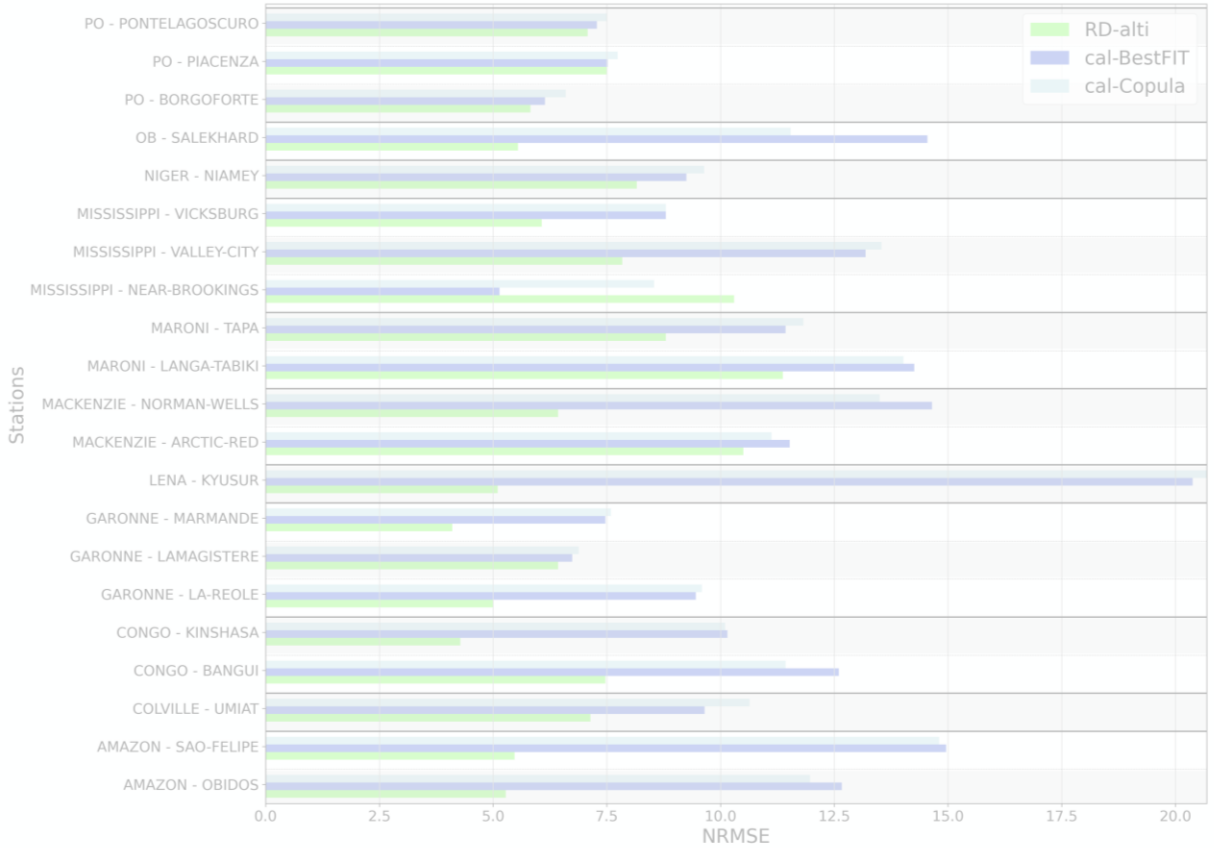


- Over satellite-based RD-cal products (21 stations) : **NRMSE < 15 %**
- RD-multi able to **add some points where alti is not available**
- RD-alti able to **better catch the high variability**

Validation - results

With Cal/Val in-situ data

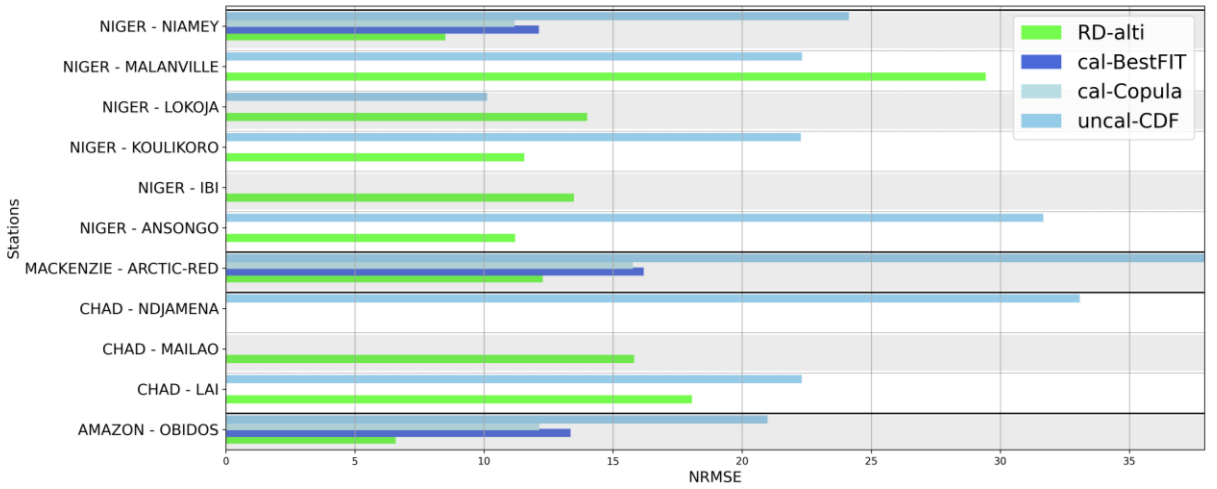
All period over common stations between calibrated RD products



- Over satellite-based RD-cal products (21 stations) : **NRMSE < 15 %**
- RD-multi able to **add some points where alti is not available**
- RD-alti able to **better catch the high variability**

With Independent in-situ data

All period over available stations for all RD products

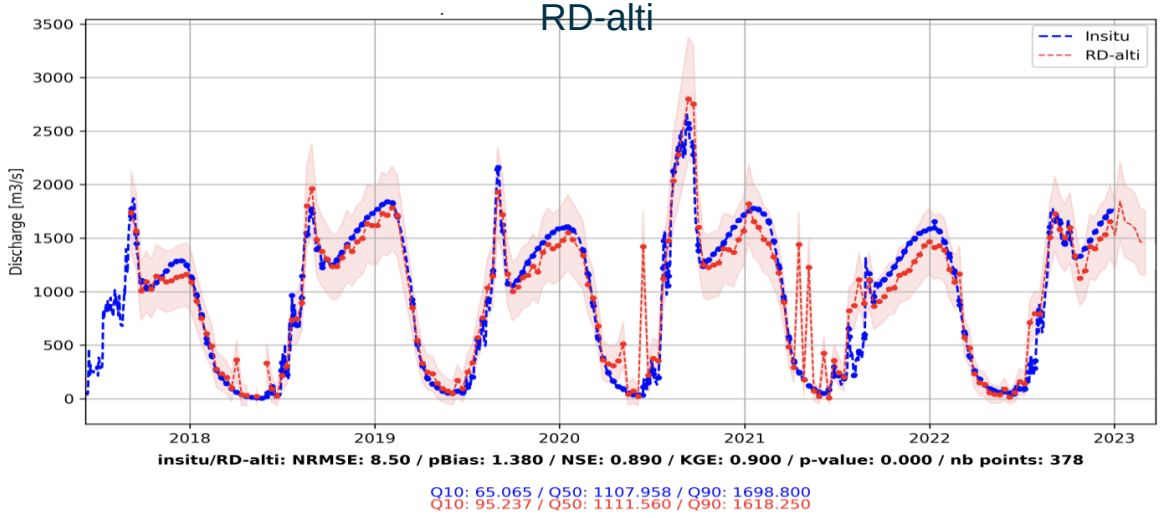


- Validation with independent in-situ data: (11 stations)

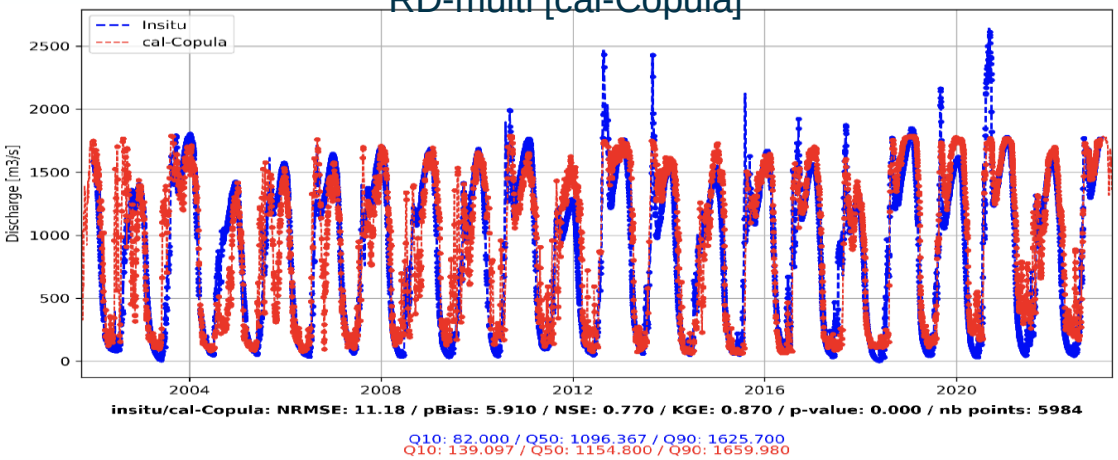
Validation - results

With Independent in-situ data

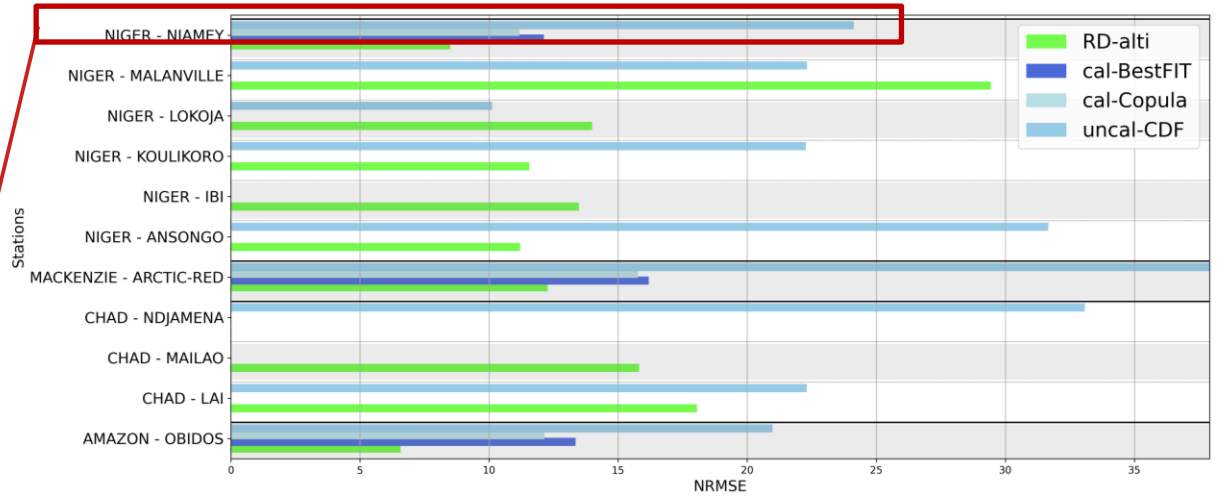
RD-alti



RD-multi [cal-Copula]



All period over available stations for all RD products

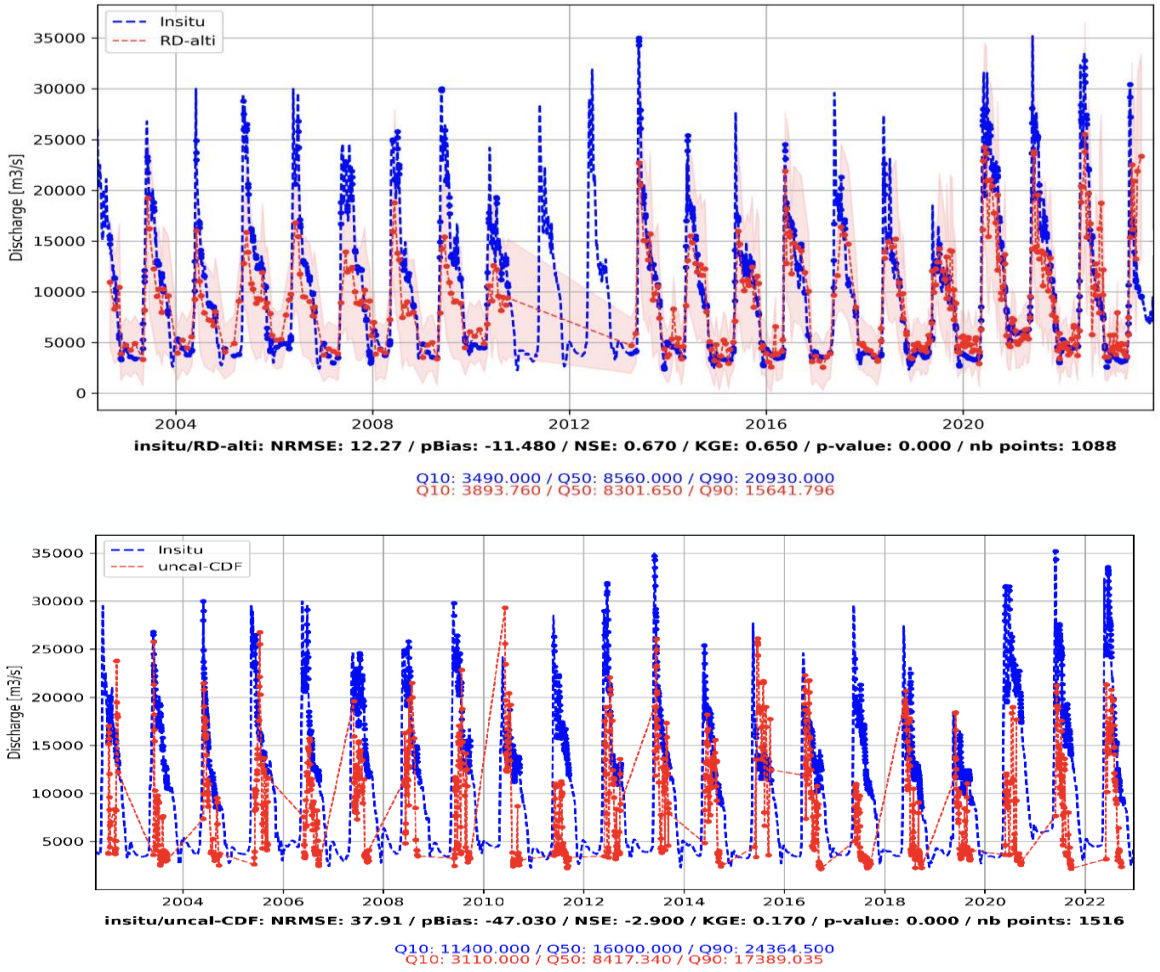


- Validation with independent in-situ data: (11 stations) : **NRMSE < 30 %**
- **RD-alti** able to provide a **good estimation of the temporal variability** with the flood events **but there is still outliers**
- **RD-multi less efficient** than RD-alti and do not catch the extreme events over the same period **but can provide more years of observation**

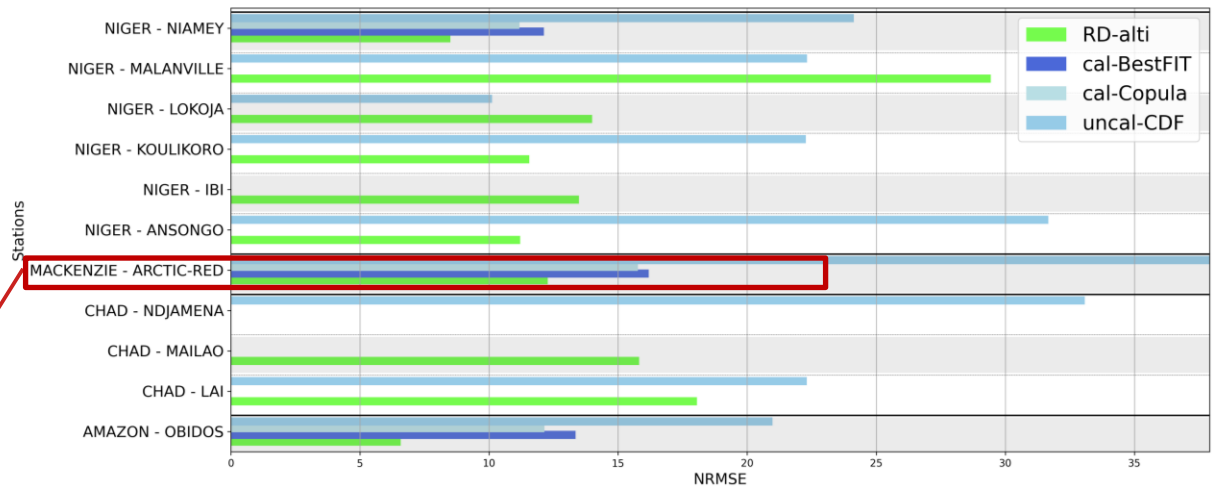
ABN database

Validation - results

With Independent in-situ data



All period over available stations for all RD products



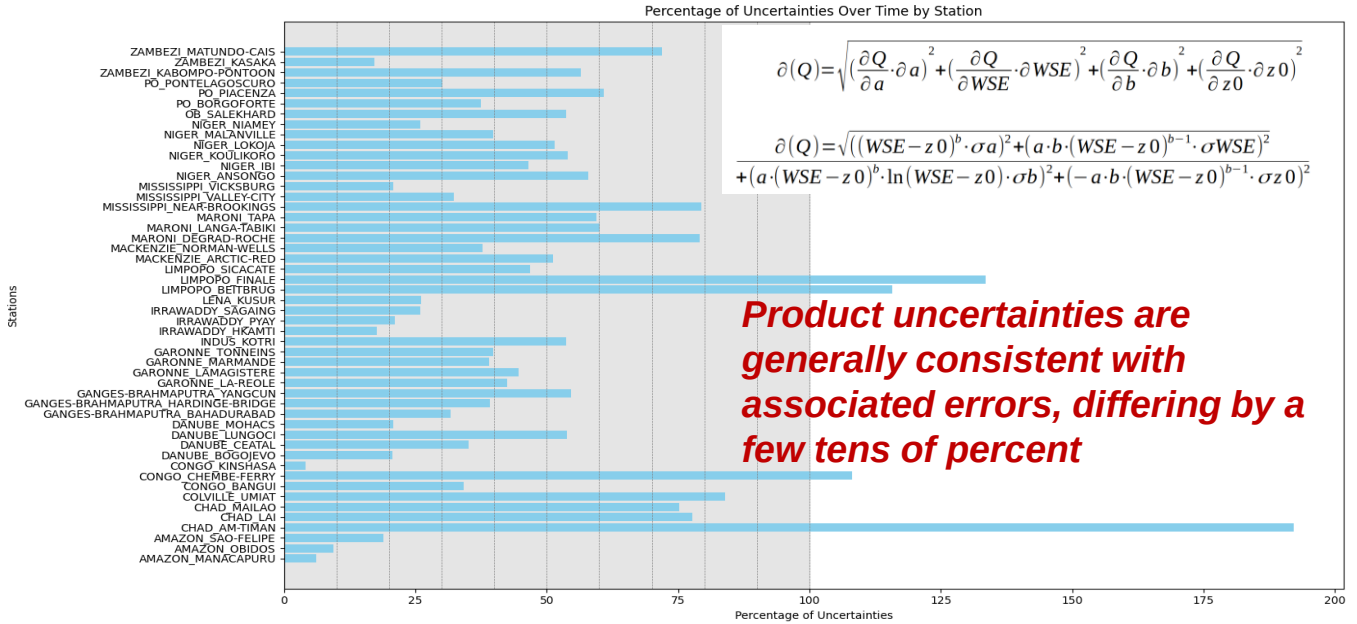
- Validation with independent in-situ data: (11 stations) : **NRMSE < 30 %**
- **RD-alti** able to provide a **good estimation of the temporal variability** with the flood events **but there is still outliers**
- **RD-multi less efficient** than RD-alti and do not catch the extreme events over the same period **but can provide more years of observation**
- **RD-alti** able to provide a **good estimation** of the RD **over the arctic basin** especially if we take into account the **associated uncertainty**
- **RD-multi [uncal-CDF] difficulty to observed frozen period** masked out in the multi indices calculation - probability of snow by MODIS

ArcticGRO database

Uncertainty

Uncertainty propagation

- Essential for assessing the **reliability of RD estimations**
- **Method**: Gaussian error propagation quantifies uncertainties in parameters a, WSE, b, and z0.
- **Assumptions**: Assumes parameter uncertainties are independent and based on linearization.
- **Average Uncertainty:**
  - Sensor changes over time.
  - Misinterpretation of altimeter data.
  - Challenges with rating curves and spatial disparities.
  - Increased sensitivity during extreme flow events.

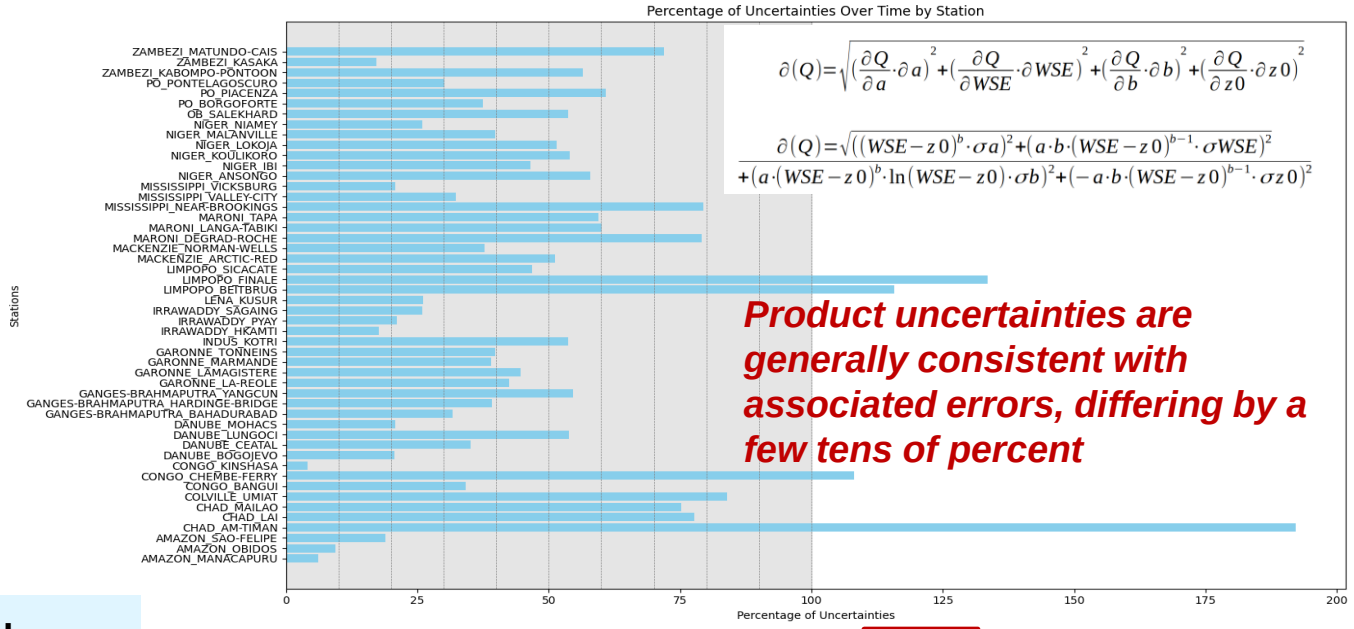


Uncertainty

Uncertainty propagation

- Essential for assessing the **reliability of RD estimations**
- **Method**: Gaussian error propagation quantifies uncertainties in parameters a, WSE, b, and z0.
- **Assumptions**: Assumes parameter uncertainties are independent and based on linearization.
- **Average Uncertainty:**

  - Sensor changes over time.
  - Misinterpretation of altimeter data.
  - Challenges with rating curves and spatial disparities.
  - Increased sensitivity during extreme flow events.



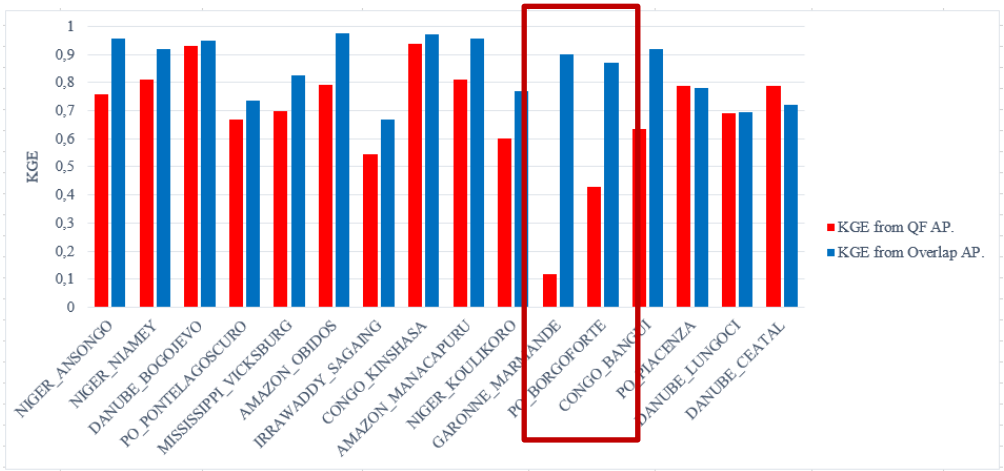
$$\partial(Q)=\sqrt{\left(\frac{\partial Q}{\partial a}\cdot\partial a\right)^2+\left(\frac{\partial Q}{\partial WSE}\cdot\partial WSE\right)^2+\left(\frac{\partial Q}{\partial b}\cdot\partial b\right)^2+\left(\frac{\partial Q}{\partial z0}\cdot\partial z0\right)^2}$$

$$\partial(Q)=\sqrt{\left((WSE-z0)^b\cdot\sigma a\right)^2+\left(a\cdot b\cdot(WSE-z0)^{b-1}\cdot\sigma WSE\right)^2}$$
$$+\left(a\cdot(WSE-z0)^b\cdot\ln(WSE-z0)\cdot\sigma b\right)^2+\left(-a\cdot b\cdot(WSE-z0)^{b-1}\cdot\sigma z0\right)^2$$

Product uncertainties are generally consistent with associated errors, differing by a few tens of percent

Error from using Quantile approach vs. Overlap approach

- RD estimates using the quantile function (non-overlap) approach have **higher uncertainties** compared to the overlap approach over the same period:
  - Non-Overlap Approach: Median KGE = 0.62 , NRMSE = 14.0%
  - Overlap Approach: Median KGE = 0.90 , NRMSE = 9.9%
- **Larger time gaps** (> 10years) between Q and WSE data lead to **decreased statistical performance**, particularly in rivers with high variability
- Quantile approach = s**ensitive to temporal distribution of hydrological events:** leading to variability in performance across different stations and periods.



Objective

Objective: Compare performance of the different algorithms/product and validate river discharge time series

Validate River Discharge (Lead Hydro Matters)

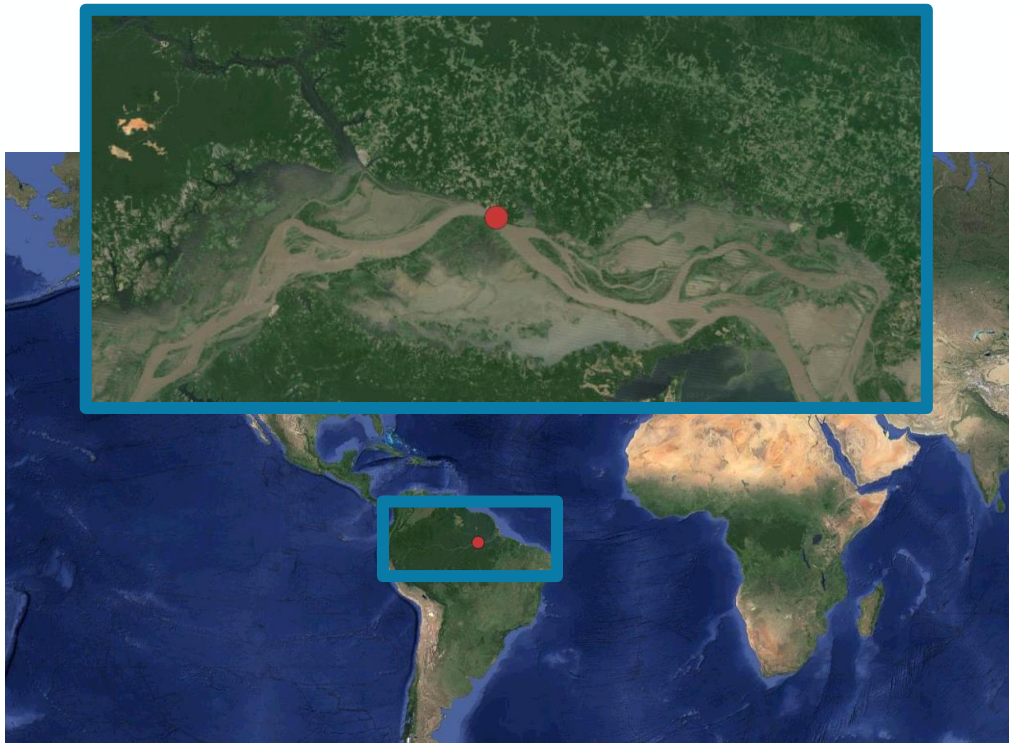
- **Validation:**
  - With Cal/Val in-situ data over validation period
  - With independent in-situ data
- **Errors – first prior to an end-to-end error budget:**
  - WSE errors between altimetry and in-situ data
  - Quantile approach – time lag between Q and WSE & daily vs monthly
  - Rating curve algorithm

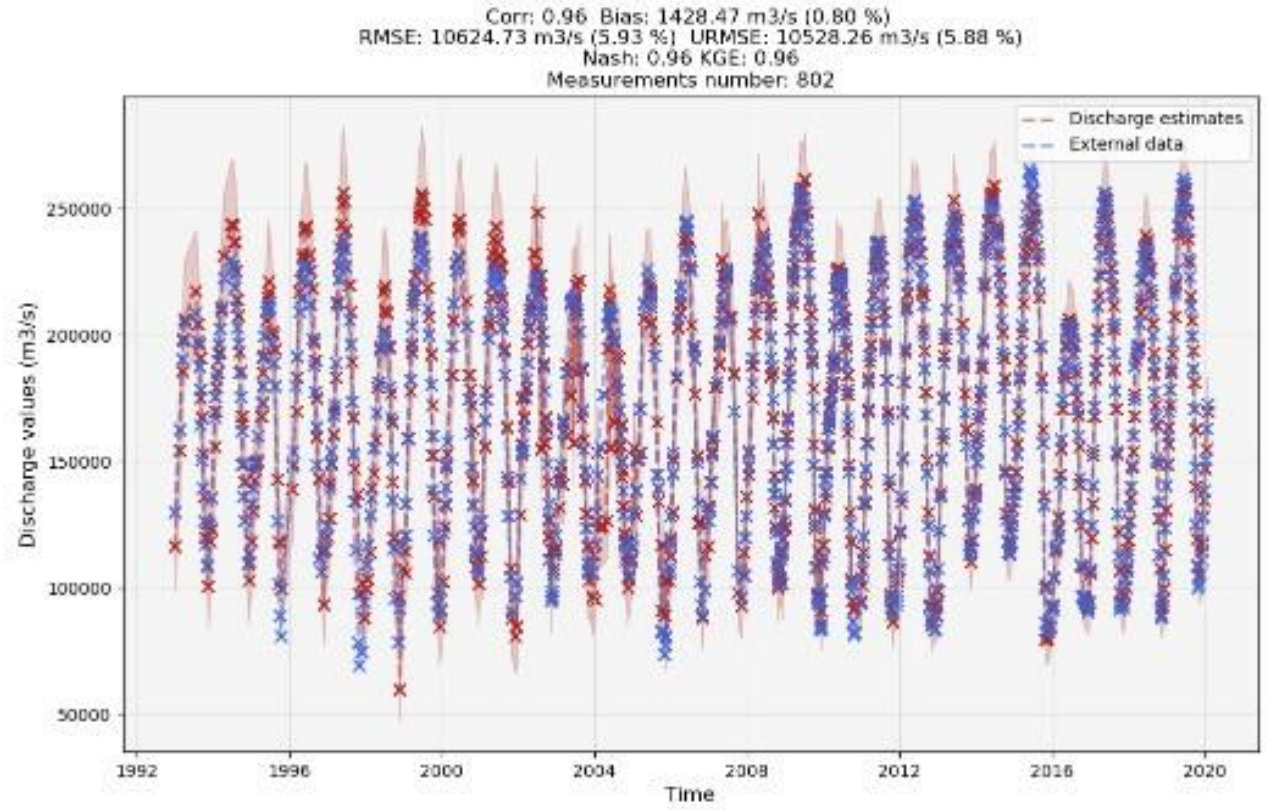
Consistency analysis and round robin (Lead CLS)

- **Format**: CCI data Standard
- **Time/space resolution**: Completeness and spatial coverage
- **Errors (in situ comparison)**: Discharge products are compared to in situ data (RMSE, Pearson, Bias, Nash, KGE)

Validation - methodology

- Metrics computation from discharge estimates and in situ timeseries
- Example with the Obidos station from RD-alti product

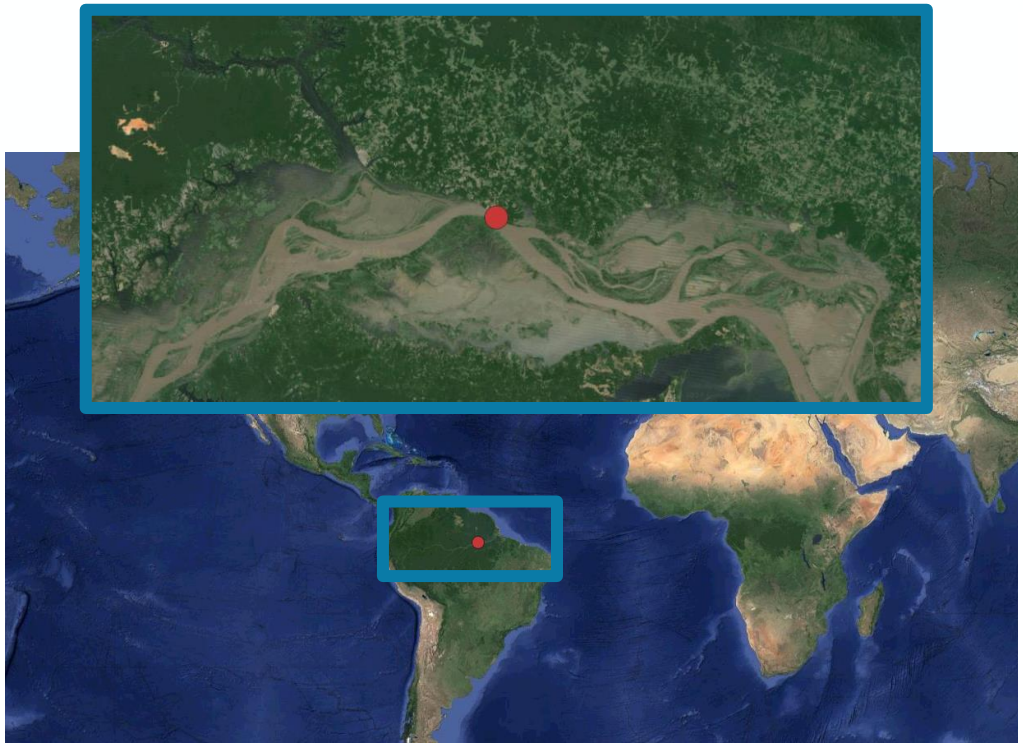


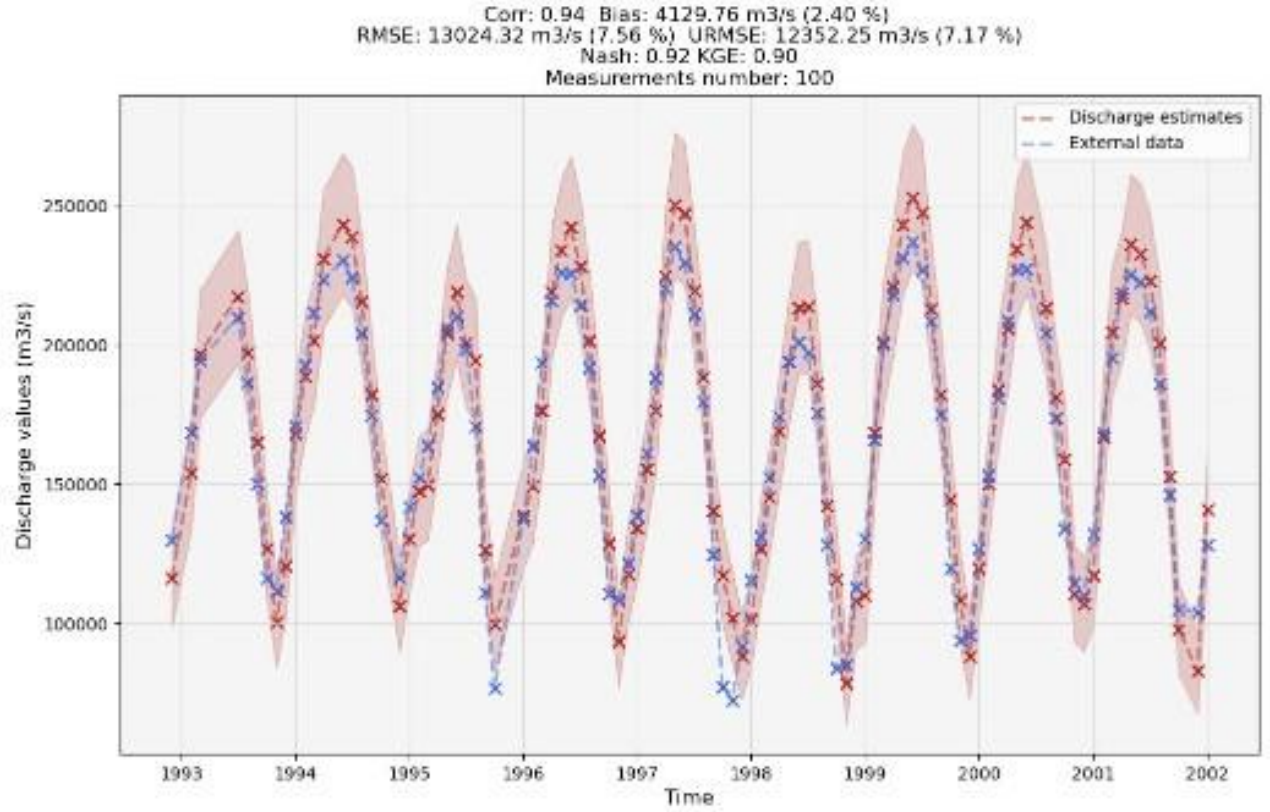


Obidos full period example. RD-Alti

Validation - methodology

- Metrics computation from discharge estimates and in situ timeseries
- Example with the Obidos station from RD-alti product



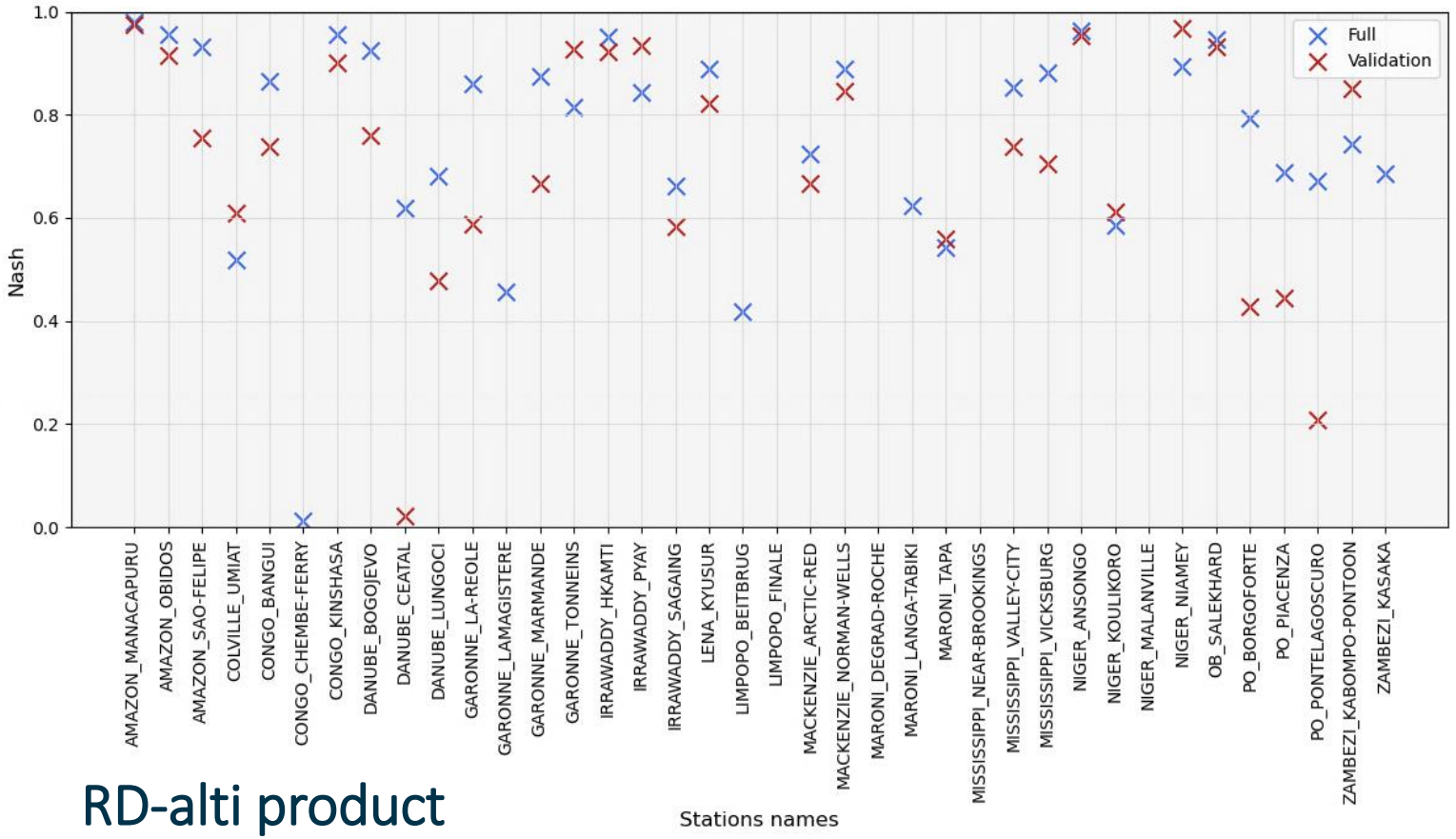


Obidos validation period example. RD-Alti

Validation - results

- Nash coefficient results

$$NSE = 1 - \frac{\sum_{i=1}^{n} (O_i - S_i)^2}{\sum_{i=1}^{n} (O_i - \bar{O})^2}$$



RD-alti product

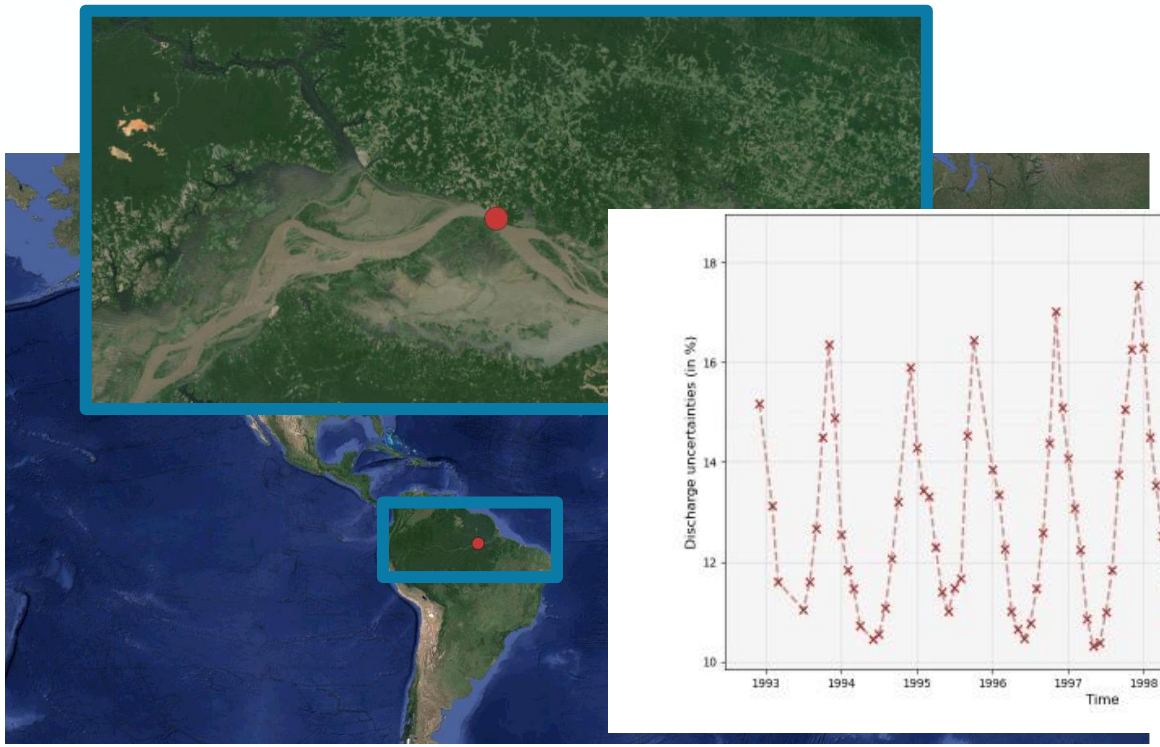
- **Better results over the full period (NSE median of 0.79) than the validation period (NSE median of 0.60)**
- **Validation over the calibration period shows greater results**
- **Recent period was used for calibration. Past period for validation**
- **Altimetry data over ERS or Envisat period is less accurate than recent period with Jason-3, Sentinel3A/B and Sentinel6A**

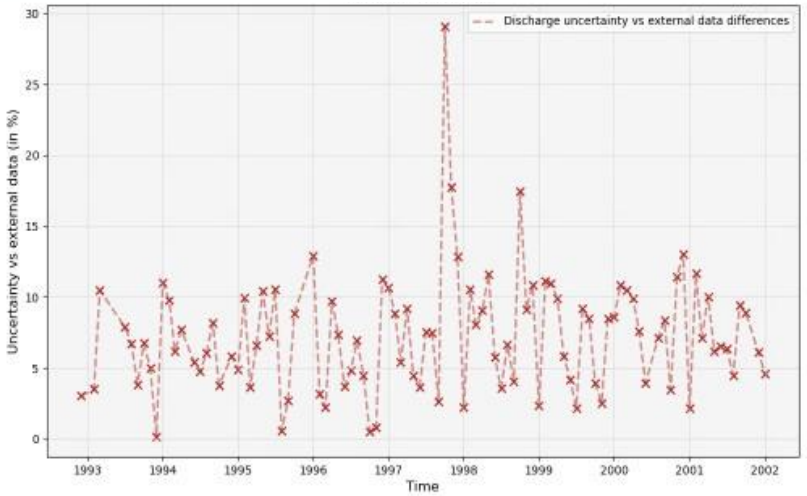
Validation - methodology

- **Uncertainties w.r.t errors**
- Example with the Obidos station from RD-alti product

$$\sqrt{\frac{1}{n} * \sum_{i=0}^{n} \left(\frac{1}{O_i} * (U_i - |O_i - S_i|)\right)^2} * 100\%$$

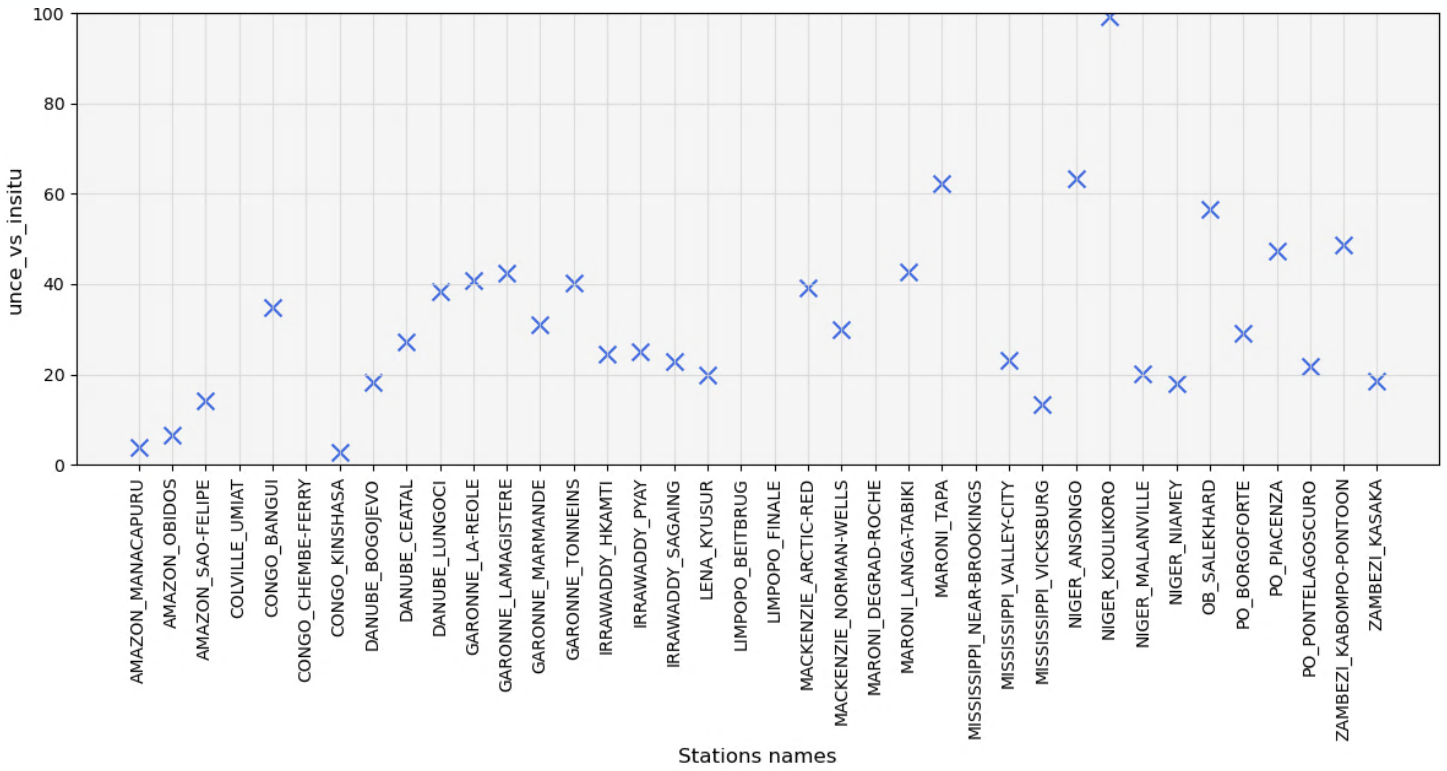
U_i the simulation uncertainty -- S_i the simulation -- O_i is the observation





Uncertainties evolution at Obidos station for RD-alti discharge estimates (left panel, 12% in average) and differences w.r.t in situ (right panel, 7% in average)

Validation - results



- Products uncertainties (RD-alti) quite consistent w.r.t error
- Differences about a few tens of percent (median value of 30%)
- Uncertainties w.r.t errors are more consistent during the calibration period

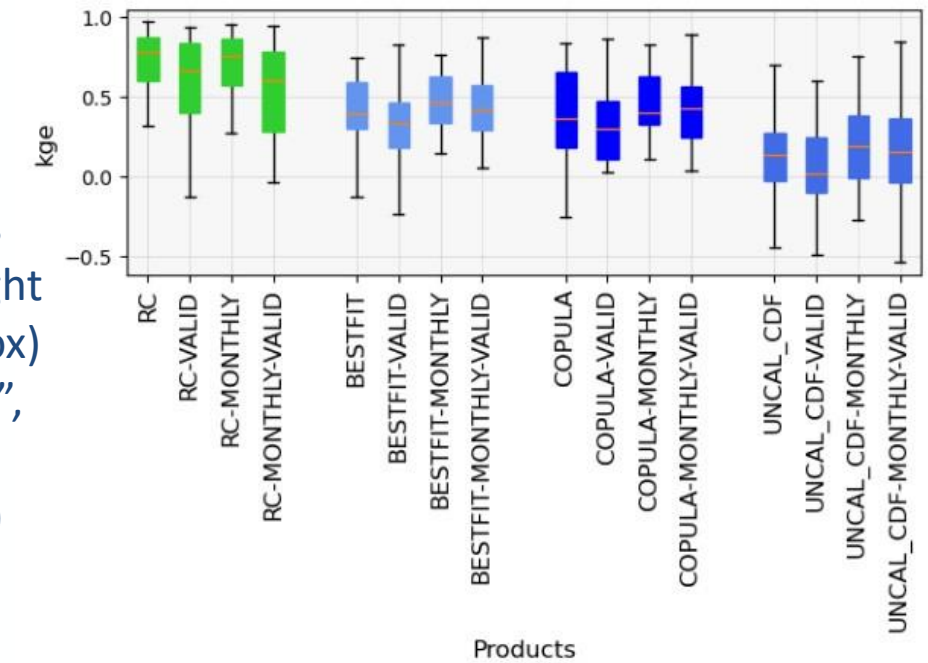
- The differences between errors and uncertainties are correlated with the uncertainty values

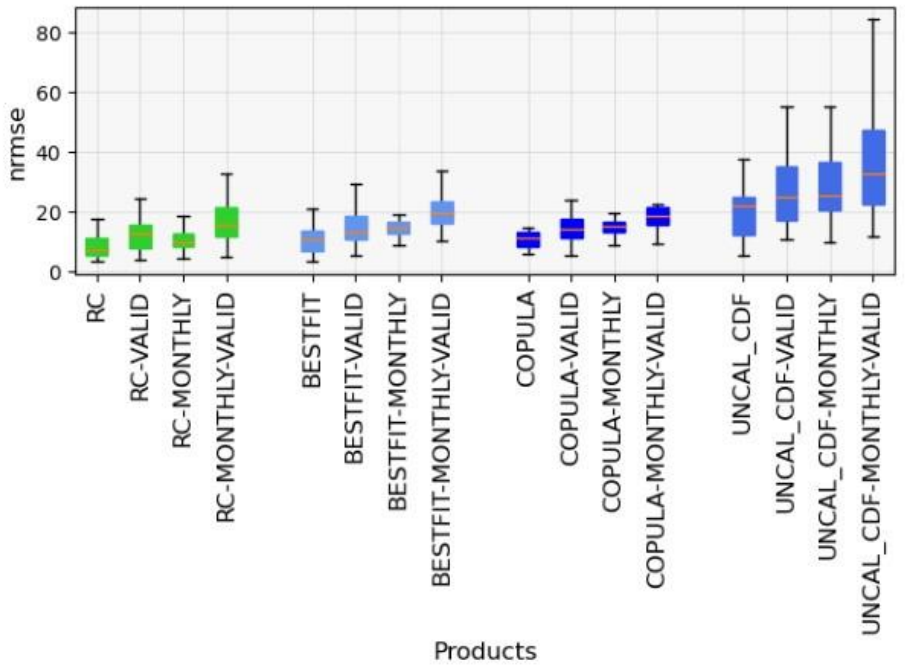
  => the greater the uncertainty, the greater the difference w.r.t the errors

Comparisons between uncertainties and errors of RD-alti discharge estimates

Validation - results

KGE coefficients boxplots (left panel) and NRMSE (right panel) for RD-alti (green box) and RD-multi (“cal-BestFit”, “cal_Copula” and “uncal_CDF” as blue box)





- **Datasets**
  - RD-alti
  - RD-multi (BESTFIT, COPULA, UNCAL)
- **Period:**
  - Full period
  - Validation period
- **Monthly average**

- RD-alti: KGE 0.78 and NRMSE 7.3% as median values with 38 stations. Results are slightly worse with monthly averages (~5% decrease in KGE, 37% increase in NRMSE) => need for better temporal sampling
- RD-multi: KGE 0.4 and NRMSE 10.8% as median values with 24 stations. Monthly averaging improves results (~15% increase in KGE, ~35% decrease in NRMSE for cal-BestFit) => need for noise reduction

- **RD-alti and RD-multi offer complementary benefits, with RD-multi's better temporal sampling and noise reduction with monthly averaging enhancing climate study discharge time series**

Conclusion

- The **CCI River Discharge Products** (CRDP) demonstrate a **high level of accuracy and reliability** compared to other satellite-based and modeled discharge time series

- **Better results for RD-alti than RD-multi when comparing to in situ data (NSE, NRMSE, KGE …)**

- **RD-alti limitations**: the non-overlap method used for estimation introduces some level of uncertainty. Main sources of uncertainty should be highlighted (oldest altimeter data, bias resolution methods). Need for better temporal sampling
- **RD-multi limitations:** difficulties separating land, vegetation, and water signals. Algorithms could be improved and other ancillary data sources (e.g. temperature data) should be used. Need for noise reduction

- **Uncertainty:**
  - RD-alti: Uncertainties are available. Quite good consistency between errors and uncertainties. Ongoing tasks to provide “end to end” error budget
  - RD-multi: Need to be implemented

- **RD-alti and RD-multi: leading options for studying river dynamics and for water resource management at global and regional scales**
- **Ongoing tasks to provide a merged dataset (with RD-alti and multi) with the latest products versions**

river discharge cci

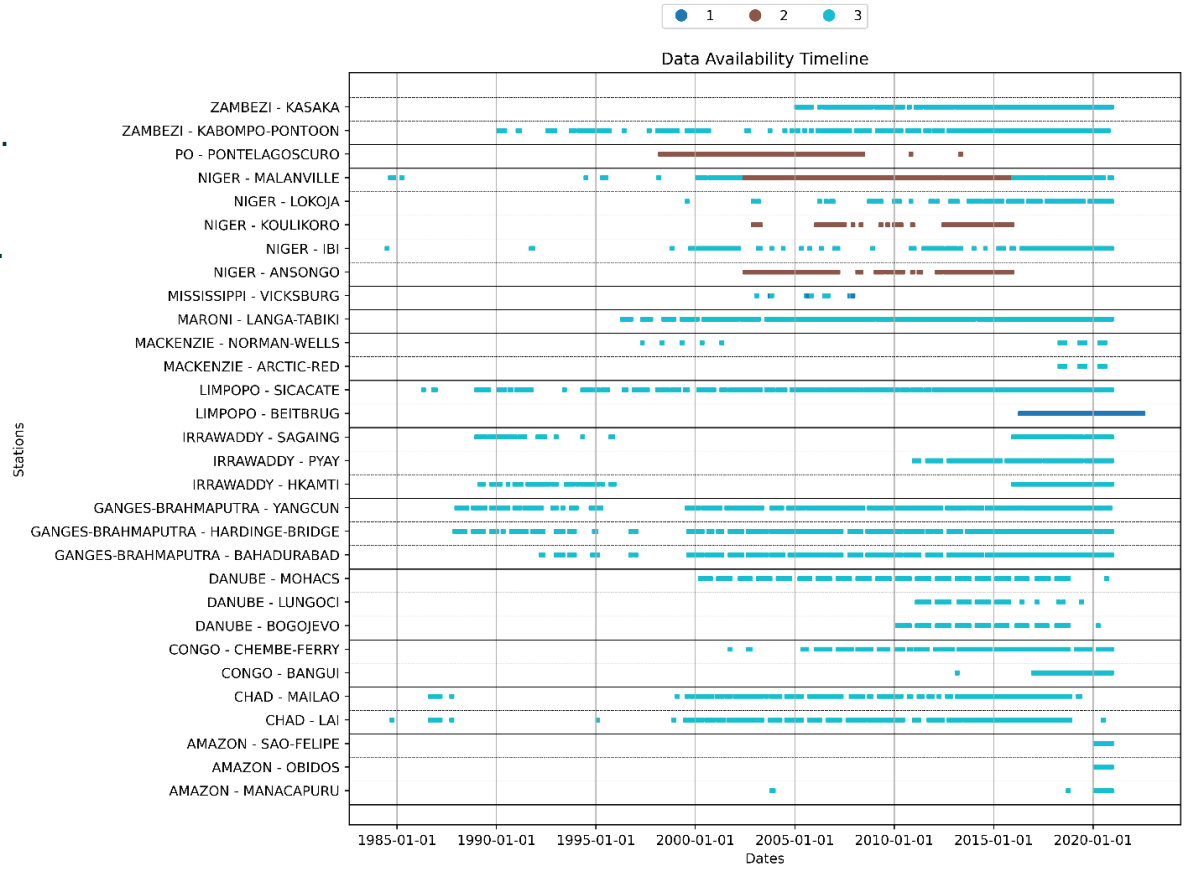
climate.esa.int/projects/river-discharge

Annexes – Intercomparison

With RSEG

DATA

RSEG Comparison: Only satellite-based discharge data considered (flags 1, 2, 3).
Time Series Issues:
- Short Series: Some stations, like the Amazon, have limited satellite data.
- Data Gaps: Some stations end earlier, not always due to GRDC data availability.



Annexes – Intercomparison

With RSEG

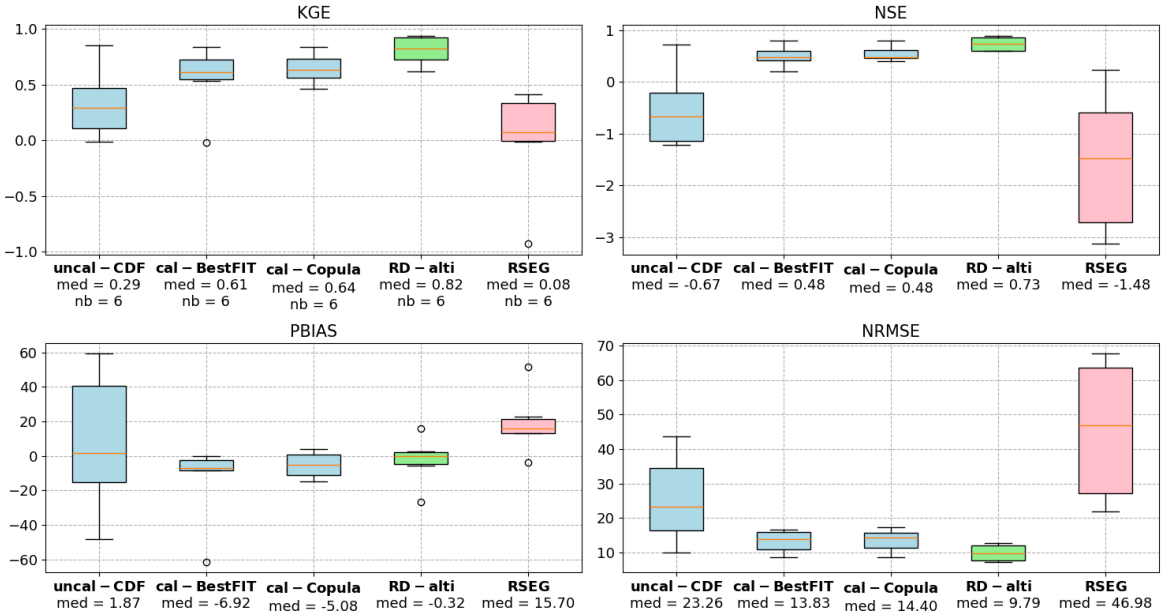
DATA

RSEG Comparison: Only satellite-based discharge data considered (flags 1, 2, 3).
Time Series Issues:
- Short Series: Some stations, like the Amazon, have limited satellite data.
- Data Gaps: Some stations end earlier, not always due to GRDC data availability.

RESULTS

- All methods in the CCI+ RD project show **better efficiency** compared to the global RSEG database (monthly res)

- **Reduced Disparity**: Methods exhibit less disparity in results
- **Calibrated Versions: Show the most significant improvements**



Annexes – Intercomparison

With RSEG

DATA

RSEG Comparison: Only satellite-based discharge data considered (flags 1, 2, 3).
Time Series Issues:
- Short Series: Some stations, like the Amazon, have limited satellite data.
- Data Gaps: Some stations end earlier, not always due to GRDC data availability.

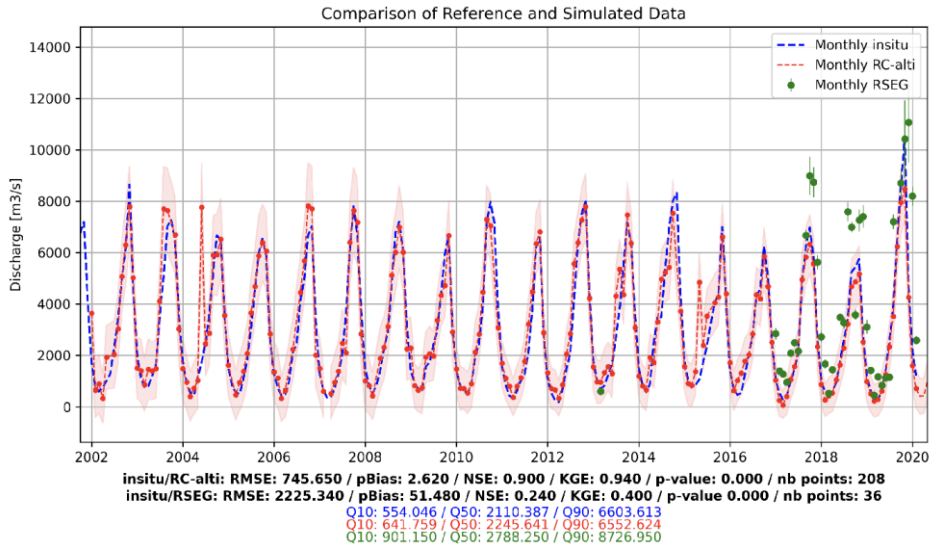
RESULTS

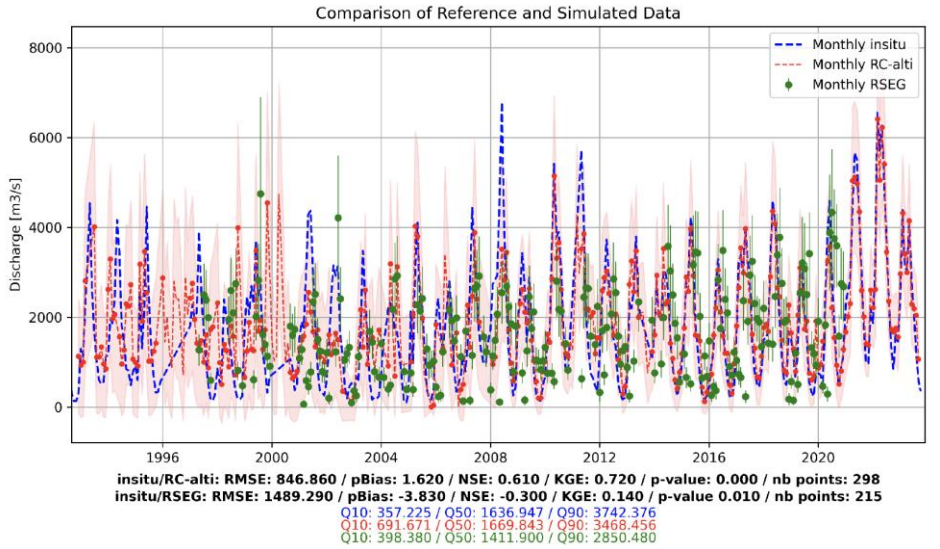
- All methods in the CCI+ RD project show **better efficiency** compared to the global RSEG database (monthly res)

- **Reduced Disparity**: Methods exhibit less disparity in results
- **Calibrated Versions: Show the most significant improvements**

- Comparison RD-alti vs RSEG

- **Better Accuracy**: RD-alti demonstrates higher accuracy in matching in-situ discharge data compared to the RSEG database
- **Consistent Performance**: RD-alti consistently outperforms RSEG across different stations and time periods, indicating its reliability in estimating river discharge





Annexes – Intercomparison

With RSEG

DATA

RSEG Comparison: Only satellite-based discharge data considered (flags 1, 2, 3).
Time Series Issues:
- Short Series: Some stations, like the Amazon, have limited satellite data.
- Data Gaps: Some stations end earlier, not always due to GRDC data availability.

RESULTS

- All methods in the CCI+ RD project show **better efficiency** compared to the global RSEG database (monthly res)

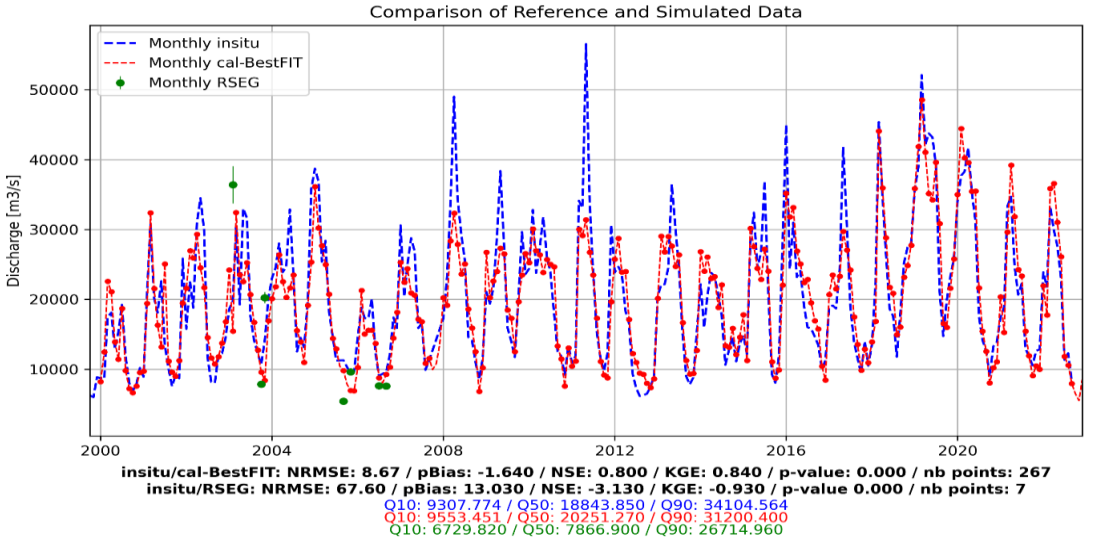
- **Reduced Disparity**: Methods exhibit less disparity in results
- **Calibrated Versions: Show the most significant improvements**

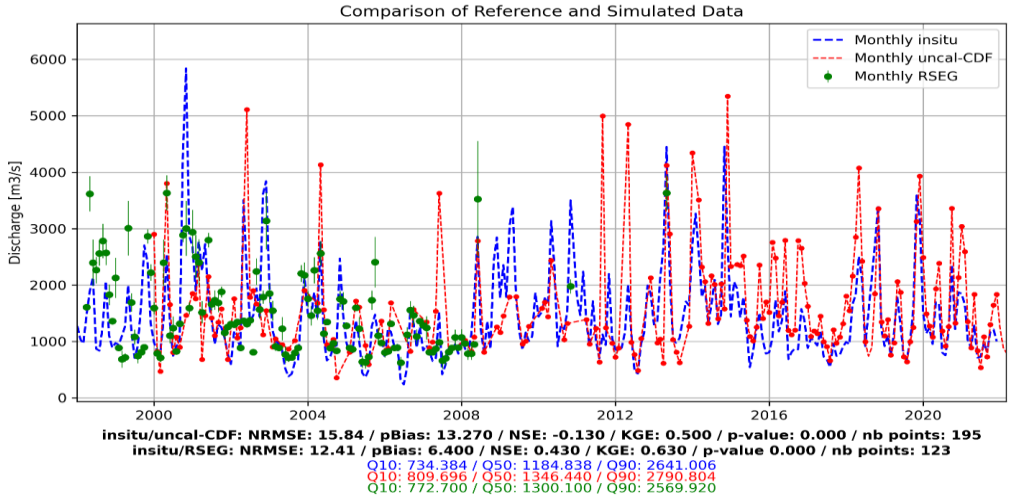
- Comparison RD-alti vs RSEG

- **Better Accuracy**: RD-alti demonstrates higher accuracy in matching in-situ discharge data compared to the RSEG database
- **Consistent Performance**: RD-alti consistently outperforms RSEG across different stations and time periods, indicating its reliability in estimating river discharge

- Comparison RD-multi vs RSEG

- **Better Performance**: RD-multi consistently outperforms RSEG data when compared with in-situ observations = higher accuracy in estimating RD
- **Calibrated Approach Enhancement**: Calibrated RD approaches = better performance compared to uncalibrated ones, indicating their advantage in providing more accurate estimations, especially during peak RD events





Annexes – Intercomparison

With GloFAS

- **GloFAS Overview**: GloFAS, part of Copernicus CEMS, detects global floods using LISFLOOD model with meteorological data.
- **Results Analysis**:
  - **- Discrepancies**: Some stations show discrepancies between RD products and GloFAS, indicating inconsistencies in flood detection.
  - **- Outliers**: Significant differences observed at certain stations suggest limitations in RD product accuracy.
  - **- RD-alti Superiority**: RD-alti outperforms RD-multi, showing potential for improved flood monitoring.
  - **- Enhanced Monitoring**: RD-alti and RD-multi complement GloFAS, enhancing flood prediction for better early warning systems.