A framework for estimating global-scale river discharge by assimilating satellite altimetry

Menaka Revel[1], Daiki Ikeshima[2], Dai Yamazaki[1], and Shinjiro Kanae[2]

[1]University of Tokyo
[2]Tokyo Institute of Technology

January 20, 2023

Abstract

Understanding spatial and temporal variations in terrestrial waters is key to assessing the global hydrological cycle. The future Surface Water and Ocean Topography (SWOT) satellite mission will observe the elevation and slope of surface waters at <100 m resolution. Methods for incorporating SWOT measurements into river hydrodynamic models have been developed to generate spatially and temporally continuous discharge estimates. However, most of SWOT data assimilation studies have been performed on a local scale. We developed a novel framework for estimating river discharge on a global scale by incorporating SWOT observations into the CaMa-Flood hydrodynamic model. The local ensemble transform Kalman filter with adaptive local patches was used to assimilate SWOT observations. We tested the framework using multi-model runoff forcing and/or inaccurate model parameters represented by corrupted Manning's coefficient. Assimilation of virtual SWOT observations considerably improved river discharge estimates for continental-scale rivers at high latitudes (>50°) and also downstream river reaches at low latitudes. High assimilation efficiency in downstream river reaches was due to both local state correction and the propagation of corrected hydrodynamic states from upstream river reaches. Accurate global river discharge estimates were obtained (Kling–Gupta efficiency [KGE] > 0.90) in river reaches with > 270 accumulated overpasses per SWOT cycle when no model error was assumed. Introducing model errors decreased this accuracy (KGE [?] 0.85). Therefore, improved hydrodynamic models are essential for maximizing SWOT information. These synthetic experiments showed where discharge estimates can be improved using SWOT observations. Further advances are needed for data assimilation on global-scale.

A framework for estimating global-scale river discharge by assimilating satellite altimetry

Menaka Revel[1*], Daiki Ikeshima[2*], Dai Yamazaki[1], and Shinjiro Kanae[2]

[1] Institute of Industrial Science, The University of Tokyo, 4-6-1, Komaba, Meguro-ku, Tokyo 153-8505, Japan

[2] Department of Civil and Environmental Engineering, Tokyo Institute of Technology,2-12-1-M1-6 O-okayama, Meguro-ku, Tokyo 152-8552, Japan

[*] these authors equally contributed to the manuscript.

Corresponding author: Menaka Revel (menaka@rainbow.iis.u-tokyo.ac.jp)

Key Points:

- A framework for assimilating satellite altimetry into a global river hydrodynamic model was developed to estimate river discharge globally.
- Virtual experiments for future SWOT satellite suggest discharge in downstream reaches of continental rivers can be accurately estimated.
- Correct hydrodynamic parameterization will enhance the accuracy of river discharge estimates for the upstream reaches of rivers at low latitudes when SWOT observations become available.

Abstract

Understanding spatial and temporal variations in terrestrial waters is key to assessing the global hydrological cycle. The future Surface Water and Ocean Topography (SWOT) satellite mission will observe the elevation and slope of surface waters at <100 m resolution. Methods for incorporating SWOT measurements into river hydrodynamic models have been developed to generate spatially and temporally continuous discharge estimates. However, most of SWOT data assimilation studies have been performed on a local scale. We developed a novel framework for estimating river discharge on a global scale by incorporating SWOT observations into the CaMa-Flood hydrodynamic model. The local ensemble transform Kalman filter with adaptive local patches was used to assimilate SWOT observations. We tested the framework using multi-model runoff forcing and/or inaccurate model parameters represented by corrupted Manning's coefficient. Assimilation of virtual SWOT observations considerably improved river discharge estimates for continental-scale rivers at high latitudes (>50°) and also downstream river reaches at low latitudes. High assimilation efficiency in downstream river reaches was due to both local state correction and the propagation of corrected hydrodynamic states from upstream river reaches. Accurate global river discharge estimates were obtained (Kling–Gupta efficiency [KGE] > 0.90) in river reaches with > 270 accumulated overpasses per SWOT cycle when no model error was assumed. Introducing model errors decreased this accuracy (KGE ≈ 0.85). Therefore, improved hydrodynamic models are essential for maximizing SWOT information. These synthetic experiments showed where discharge estimates can be improved using SWOT observations. Further advances are needed for data assimilation on global-scale.

Plain Language Summary

River discharge is an important indicator for managing the world's freshwater resources. Advances in computing technology have facilitated the development of hydrodynamic models, which can be used to predict river water states and compensate for the lack of in-situ observation facilities. However, these models have inherent limitations, including the simplified physics, forcing errors, and inaccurate parameters. Satellite observations, such as those from the Surface Water and Ocean Topography (SWOT) mission, may be incorporated to improve these models. Because the SWOT satellite is due for launch in 2021, assessing the potential benefits of incorporating SWOT observations into global hydrodynamic models is essential. Therefore, we performed observation assimilation experiments using a technique known as Kalman filtering, which assesses model uncertainty and expected observation errors. Note that SWOT observations are not recorded continuously; therefore, the hydrodynamic model was used to extrapolate water states in time and space. We found that incorporating SWOT observations provided accurate river discharge estimates, in continental-scale rivers. Furthermore, correcting model parameters will considerably improve river discharge estimates. This framework may be used to generate accurate global river discharge estimates when SWOT observations become available. Therefore, these methods can be helpful for mitigating conflicts in transboundary river basins (e.g., Mekong).

1. Introduction

River discharge is a key variable for understanding the global hydrological cycle and assessing water resources (Oki & Kanae, 2006). Networks of *in situ* stream gauging stations are a fundamental data source for estimating spatial and temporal variations in the discharge of major rivers worldwide. However, the numbers of accessible stream gauges are not adequate to fully understand details of the global hydrological cycle, and real-time access to gauged discharge data is usually available only in developed countries. Although remote sensing of river discharge is a challenging research topic, recent advances in satellite observation technology are expected to enhance our understanding of river discharge variation on the global scale (Marcus & Fonstad, 2010).

The Surface Water and Ocean Topography (SWOT) satellite is a next-generation satellite altimetry mission due to launch in 2021 (Durand et al., 2010). This satellite will measure two-dimensional water surface elevation (WSE) across its 120 km wide swath using a Ka-band radar interferometer. The WSE of rivers and lakes will be measured at < 100 m spatial resolution over 5–10 day intervals, depending on the satellite's location during its 21-day repeat-cycle orbit (Biancamaria et al., 2016). The fine spatial resolution will ensure that rivers wider than 50–100 m (Pavelsky et al., 2014) and lakes larger than 1–5 ha (Lee et al., 2010) are included, providing information on surface-water dynamics in unprecedented detail. In addition to WSE data, the high-resolution measurements will also provide accurate information on water -surface slopes across river networks.

Although the SWOT satellite will not measure river discharge directly, algorithms to estimate discharge from variables that will be measured by SWOT (e.g., WSE, slope, and width) have been developed (Durand et al., 2016; Garambois & Monnier, 2015; Gleason & Smith, 2014). These algorithms can estimate river discharge in some ungauged rivers with approximately 35% root mean square error (Bonnema et al., 2016; Durand et al., 2016). However, due to the limited frequency of observations, these satellite-based methods cannot produce spatially and temporally continuous estimates of river discharge. Consequently, recent research has investigated whether SWOT measurements can be integrated into river hydrodynamic models (Andreadis et al., 2007; Biancamaria et al., 2011; Brêda et al., 2019; Pedinotti et al., 2014).

The potential benefits of assimilating future SWOT observations into river hydrodynamic models have been assessed using observing system simulation experiments (OSSEs) (Andreadis et al., 2007). Because the SWOT satellite has not yet been launched, synthetic SWOT observations were generated using a river hydrodynamics model that was assumed to be 'true.' Then, the synthetic observations were assimilated into a corrupted hydrodynamics model. The data assimilation framework was evaluated by comparing the estimated river discharge improved by the assimilation against the 'true' simulation. Some SWOT data assimilation methods have already been developed and tested in several river basins, including a 50 km reach of the Ohio River (Andreadis et al., 2007), the main-stem of the Ob River (Biancamaria et al., 2011), the Niger River (Munier et al., 2015; Pedinotti et al., 2014), the entire Congo basin (Revel et al., 2019), and the Amazon basin (Brêda et al., 2019; Emery et al., 2019). These studies demonstrated that SWOT observations had the potential to improve river hydrodynamic simulations and estimate river discharge and/or hydrodynamic parameters continuously in space and time. However, previous SWOT assimilation studies used regional-scale river models or expensive data assimilation algorithms, which cannot be applied easily on a global scale. In addition, a global-scale hydrodynamic model that can assess WSE measurements rapidly is needed because SWOT WSE measurements will be generated daily. To evaluate the effectiveness of data assimilation in global-scale using SWOT observation, a global-scale study with computationally efficient river model and data assimilation algorithm are essential.

This study evaluated the use of SWOT observations to estimate river discharge on a global scale. We developed a new data assimilation framework for integrating SWOT observations into a global river hydrodynamic model. Using the Catchment-based Macro-scale Floodplain (CaMa-Flood) global hydrodynamic model (Yamazaki et al., 2011) and a highly efficient data assimilation method called the local ensemble transform Kalman filter (LETKF; Hunt et al., 2007), we assimilated global-scale data at a reasonable computational cost. A detailed description of the data assimilation framework is provided in Section 2. The experimental conditions and evaluation methods are presented in Section 3. Results are explained and discussed in Section 4, and Section 5 includes a summary and discussion.

2. Development of Data Assimilation Framework

2.1 SWOT data assimilation framework

We developed a new global-scale data assimilation framework for hydrodynamic modeling to estimate river discharge using SWOT altimetry data. The CaMa-Flood hydrodynamic model formed the core of our global data assimilation framework. This framework was designed to assimilate WSE data gathered by the SWOT satellite. We used the LETKF, an ensemble Kalman filter (EnKF) variant, as our data assimilation algorithm. In addition, we used a physically-based adaptive localization method to utilize as many observations as possible.

Figure 1 shows the workflow for our data assimilation framework. First, from the initial water state in time step T (x_T^a in Figure 1), the water state ($x_{T+\Delta T}^f$ in Figure 1) for time $T + \Delta T$ was simulated using the CaMa-Flood model for the duration ΔT, forced by the land surface runoff data at the corresponding time. Here, multiple forecasted water states were prepared from the different initial water conditions and runoffs. An ensemble of forecasted water states is essential for assessing prior error covariance in LETKF assimilation procedure. The corrected water state at $T + \Delta T$ ($x_{T+\Delta T}^a$ in Figure 1) was derived by combining SWOT observations and the ensemble of forecasted water states using the LETKF algorithm, taking the model variance and observation error into account. The corrected water state was used as the initial water state for the next simulation. At the beginning of the simulation ($T = 0$), ensembles were based on the spin-up simulation from the previous year.

2.2 River hydrodynamics model: CaMa-Flood

We used the CaMa-Flood hydrodynamic model (Yamazaki et al., 2011, 2012, 2013) to form the core of our data assimilation framework. The CaMa-Flood model receives runoff from a land surface model (LSM) as input forcing (i.e., the quantity of water entering a river from a unit land area in mm/day) and simulates river and floodplain hydrodynamics (i.e., river discharge, WSE, inundated area, and surface water storage) on a global scale. The spatial resolution of the CaMa-Flood model, which was set to 0.25° in this study, is coarser than that of two-dimensional flood inundation models (typically <1 km; Bates et al., 2010). Instead of solving two-dimensional floodplain flows at high resolution, the CaMa-Flood model simulates floodplain inundation

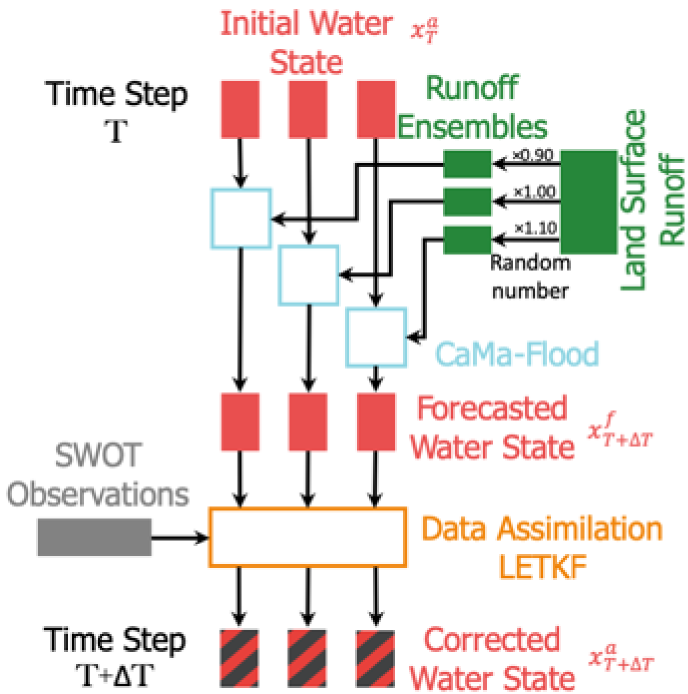

Figure 1: Workflow for the Surface Water and Ocean Topography (SWOT) data assimilation framework.

dynamics using sub-grid topography parameters delineated from fine-resolution topography. Whereas the water mass balance (i.e., surface water storage and river discharge) is calculated at coarse-grid resolution, the complex floodplain inundation is represented by diagnostic sub-grid physics. Therefore, the CaMa-Flood model achieves computationally efficient simulations of global-scale river hydrodynamics. The CaMa-Flood model calculates river discharge using a local inertial flow equation (computationally efficient modification of the shallow water equation) (Bates et al., 2010; Yamazaki et al., 2013). Because the pressure term is included in the local inertial equation, river discharge is estimated based on the water surface slope. This is a key difference between the CaMa-Flood model and conventional global river models, which use a kinematic-wave flow equation that neglects the pressure term. Combining the sub-grid flood inundation scheme and the local inertial flow equation generates a realistic representation of the WSE in river channels and floodplains. A previous study confirmed that WSE measurements obtained from simulations that used the CaMa-Flood model were similar to those observed using satellite altimetry (Yamazaki et al., 2012). Therefore, we chose the CaMa-Flood model to form the hydrodynamics core of our data assimilation framework. In this study, we used the latest version of the CaMa-Flood model (ver. 3.96), which integrates highly accurate state-of-the-art global topography datasets, MERIT DEM and MERIT Hydro (Yamazaki et al., 2017, 2019).

2.3 Input runoff forcing

In this study, we generated ensembles of forecasted water states using CaMa-Flood to calculate the error covariance, in accordance with the LETKF data assimilation algorithm. We generated the ensemble of forecasts using different runoffs (Figure 1) as input forcing for the CaMa-Flood model. The number of ensemble members was set to 18, in accordance with the computational cost and recommended minimum number of ensembles for the LETKF algorithm

CaMa-Flood hydrodynamic model has a higher computational cost than the data assimilation algorithm, although the number of Monte Carlo sampling errors decreases as the number of ensembles increases (Evensen, 2009). In addition, more than 10 ensembles should be used with the LETKF algorithm (Miyoshi et al., 2007). Therefore, we prepared 18 different runoffs from a multi-model runoff project named EartH2Observe “Global Earth Observation for Integrated Water Resource Assessment” (E2O), the tier-2 water resources reanalysis (WRR2) development project (Dutra et al., 2017). The multi-model runoff data were simulated using bilinear interpolated ERA-interim meteorological data with topographic temperature correction and Multi-Source Weighted-Ensemble Precipitation data. There were eight runoff outputs from different LSMs or Global Hydrological Models (GHMs). We used runoff outputs from Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land (HTESSEL) LSM as “virtual truth” in our experiments because combining the HTESSEL and CaMa-Flood models produces reliable results (Dutra et al., 2017). We treated remaining runoff outputs from E2O WRR2 as “corrupted” runoff inputs. PCR-GLOBWB, JULES, LISFLOOD, ORCHIDEE, WaterGAP3, and W3 were used, whereas SURFEX outputs were not used due to incompatibility with the CaMa-Flood hydrodynamic model. In total, 18 ensembles were generated from the runoff outputs of six LSMs/GHMs. We also took variation in meteorological forcing into account; this was assumed normally distributed, with mean = 0 and standard deviation = 0.1 considering the variability of runoffs from twentieth-century atmospheric model ensemble (ERA-20CM: Hersbach et al., 2015). We generated three perturbations from each runoff output. Therefore, there were 18 ensembles in total. Further information regarding runoff forcing is provided in the Supplementary Information (Text S1, Table S1).

2.4 Data assimilation using the LETKF

For the data assimilation method, we used the LETKF (Hunt et al., 2007), which is a variation of the EnKF (Evensen, 2003), an advanced Kalman filter (KF; Kalman, 1960). KF methods estimate future states in time-evolution models by merging model estimates and observations and using weighting procedures that are assessed for their reliability; here, the weighting process is called the Kalman gain. As repeated assimilations progress, better estimates can be made due to the accrual of previous observations. In addition, KF methods assimilate observations and model forecasts using the covariance among pixels. Consequently, a single target pixel is assimilated by combining observations from many pixels. This reduces the likelihood of observation error and also allows locations with no observations to be assimilated. However, KF methods were developed for linear models and cannot be applied to nonlinear models, such as river hydrodynamics models. The EnKF is a variant of the KF that incorporates the Monte Carlo method, enabling data assimilation for nonlinear models. The EnKF calculates different assimilated states using slightly different inputs or initial values. Each of these states is called an ‘ensemble member’ and the set of these members is the ‘ensemble’. The Kalman gain matrix can be calculated from model estimates, even for nonlinear models, using the variation among ensemble members. However, using the EnKF or other KF variants for data assimilation on a global scale is associated with a significant increase in computational cost. Therefore, applying these data assimilation methods to large regions is problematic.

The computational cost associated with using the EnKF can be reduced by using the LETKF. This allows data to be assimilated on a global scale. The LETKF is a type of EnKF that increases computational speed by ignoring the covariance between distant pixels (see Text S2 for a detailed description of the LETKF). For each target pixel, a small group of pixels called a ‘local patch' is considered. The Kalman gain matrix for each target pixel is calculated using observation error and the ensemble variation of forecasted states of the pixels in each local patch. Revel et al. (2019) developed physically-based adaptive empirical local patches for hydrological data assimilation that include spatial correlation among the WSE measurements. Further information regarding empirical local patches is provided in the Supplementary Information (Text S3).

We assimilated WSE measurements from SWOT observations into a forecasted water state using the LETKF (Figure 2) and corrected the initial condition for the next step. For pixels that had no observations within their local patch, the ensemble of forecasted states was used as the corrected state. Note that using the LETKF assimilation algorithm may result in mass balance errors when the local patches are smaller, particularly in upstream river reaches. However, applying the physically-based adaptive empirical local patch considerably decreases mass balance errors. Previous water states were not corrected in the current time step, which may have resulted in mass balance errors. Implementing smoothing data assimilation methods, such as the Kalman smoother, can reduce these errors. However, this greatly increases the computational cost and may raise new uncertainties. Our focus was to provide better river discharge estimates on a global scale. We do not recommend that the outputs from this study are used for precise mass balance calculations. The major advantage of using the LETKF assimilation here was its increased computational speed.

3. Evaluation of the Data Assimilation Framework

3.1 Experimental objectives and assumptions

We performed OSSEs to evaluate the river discharge estimates from the SWOT data assimilation framework. OSSEs are often used to assess the potential of new measurements before these are implemented (Sylvain Biancamaria et al., 2016). In the OSSEs, we generated synthetic observations using a hydrodynamic model and compared them with the assimilated results. The OSSEs consisted of three parts: the 'true simulation' representing the assumed-to-be-true (hereafter true) situation to generate synthetic SWOT observations for data assimilation; the 'corrupted simulation' representing a model forecasted simulation, which is usually separate from the true state; and the 'assimilated simulation' representing data assimilation from a model forecast using synthetic SWOT observations (Figure 2). The object of these experiments was to evaluate our SWOT data assimilation framework and determine whether it can estimate global river discharge accurately with poor land surface runoff forcing data and/or poor model parameter estimates. We made the following assumptions for these experiments:

1. We assumed similar models may be used to generate virtual SWOT observations and forecasted water states. We used the CaMa-Flood hydrodynamic model to represent both true and corrupted/assimilated water states. However, the CaMa-Flood model includes uncertainties in both its physics and model parameters. For example, the CaMa-Flood model assumes that WSE measurements for the river channel and floodplain are the same within each grid box, whereas observation-based studies have suggested that there is a time lag in water-level changes between these two situations (Alsdorf et al., 2005). A uniform water level was assumed for each 0.25° pixel, whereas real WSEs have sub-grid variations. In addition, incorrect topography parameters (e.g., elevation and channel bathymetry) will generate bias in water state forecasts. The CaMa-Flood model uses a global constant value for Manning's coefficient, although these are spatially distinct. Due to the uncertainties described above, using the same hydrodynamic model for true and corrupted/assimilated simulations in OSSEs

may be too optimistic because the hydrodynamics of actual rivers may not well represent in hydrodynamic models. Therefore, we also assessed imperfect model conditions, which are represented by using different Manning's coefficients in true and corrupted/assimilated simulations in this study. Further information regarding the imperfect model experiment is provided in section 3.4.

2. We assumed that synthetic SWOT observations are compatible with the CaMa-Flood model-grid scale (~25 km). The SWOT satellite will observe WSE at 50–100 m resolution, depending on the distance from the satellite (Fjørtoft et al., 2014). Therefore, it is necessary to increase the scale of high-resolution SWOT observations to match the coarse-scale grids of the CaMa-Flood model. This is particularly important in steep river reaches and areas with floodplains, where sub-grid variations in WSE are large. Consequently, the mean WSE cannot be used in the coarse-scale grid for data assimilation under the unit-catchment assumption of the CaMa-Flood model (Yamazaki *et al*., 2011). Therefore, we assumed that average SWOT observations within a certain distance of the unit-catchment mouth can be used for data assimilation in this study. Further information on generating synthetic SWOT observations is provided in section 3.2.1.

3. We also assumed that the SWOT satellite can measure the WSE of rivers > 50 m in width at 10 cm error (water area ≥ 1 km^2) accuracy (25 cm for [250 m]^2; water area < 1 km^2). These thresholds were adopted in accordance with the mission goal (Desai *et al*., 2018) and estimated error limits of the SWOT satellite mission (Esteban-Fernandez, 2017). Further information on generating synthetic SWOT observations from these assumptions is provided in section 3.2. The observation capability for each river grid is complex and varies with river width, river length, surrounding topography (Durand *et al*., 2010), and distance from the satellite track (varies between 4~10cm) (Esteban-Fernandez, 2017).

3.2 Experimental design

3.2.1 Overview of OSSEs

The workflow for the OSSEs is shown in the Figure 2. We performed three simulations: 'true', 'assimilated', and 'corrupted'. From the true simulation we derived the synthetic SWOT observations. In the assimilated simulation, we tested our assimilation framework, and we used the corrupted simulation to evaluate our data assimilation framework. We performed the OSSEs over a 1-year period from 1 January 2004 to 31 December 2004. A 1-year spin-up simulation was used to generate initial conditions for the true and corrupted/assimilated simulations.

We used the 'true simulation' to generate the true water state, which was continuous in space and time. In the true simulation, the hydrodynamic model was forced by true input runoff (i.e., HTESSEL LSM runoff output from E2O WRR2), generating the true water state (e.g., river discharge, WSE, and water storage). The initial condition for the true simulation was prepared using a spin-up simulation with the same model settings. The true simulation was used to generate the synthetic SWOT observations and to evaluate the results of assimilation.

To evaluate the assimilation framework, we performed the 'corrupted simulation'. The modified LSM runoff outputs (representing uncertainty in the meteorological data) from E2O WRR2 were used in the corrupted simulation. We used the standard global value of 0.03 for Manning's coefficient (Yamazaki et al., 2011). Using similar input runoffs and model parameters, the water state at the beginning of the simulation period (2004) was prepared by running the CaMa-

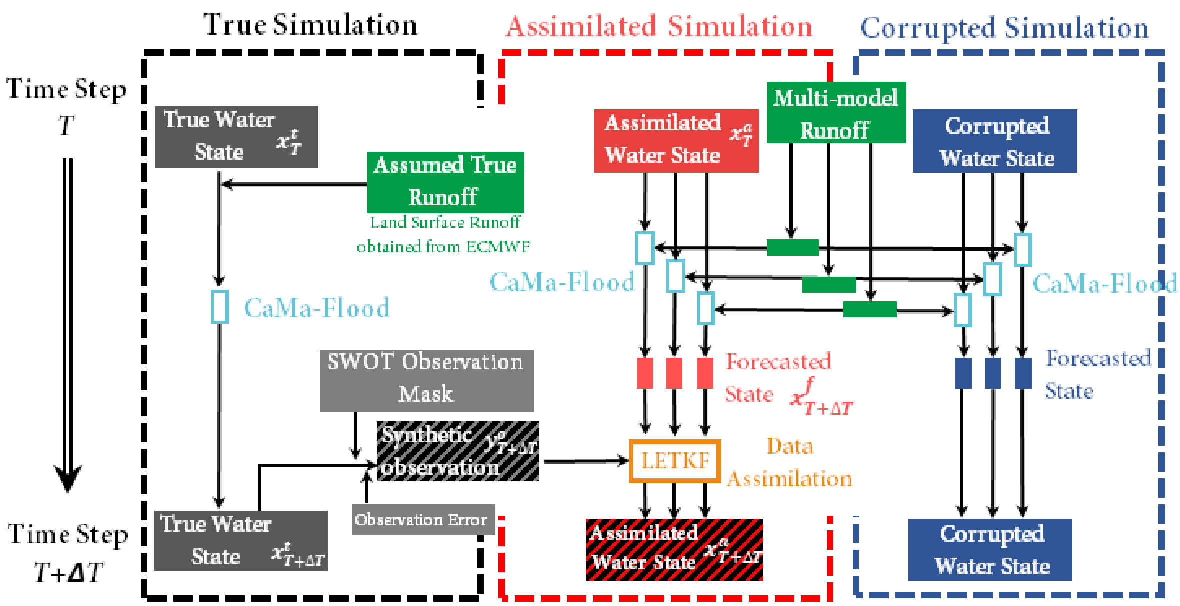

Figure 2: Workflow for the virtual experiment

Flood model for 1 year (2003). The data assimilation procedure was not implemented for the corrupted simulation.

Next, we performed the 'assimilated simulation' to evaluate the use of SWOT observations in estimating global river discharge. We used the same model settings and inputs that were used for the corrupted simulation, but the water state was corrected by assimilating synthetic SWOT observations. We used a physically-based data assimilation technique (Revel et al., 2019) based on the LETKF data assimilation algorithm, as described in section 2.4.

3.2.2 Synthetic SWOT observations

Synthetic SWOT observations were generated by overlaying the SWOT coverage mask onto WSE measurements from the true simulation (Figure 3a). Therefore, we assumed that only part of the true water state (i.e., WSE measurements from SWOT observations) was known, as would be the case if real SWOT satellite observations were being used. The SWOT coverage mask was created using SWOT orbit data (Figure 3a, center panel) available online at the Centre national d'études spatiales web page (CNES, 2015). The orbit data indicate the satellite's path of the 120 km wide observation swath with a 20 km nadir gap for each day, for 21-day orbit cycle. The SWOT coverage mask was created at a resolution of 0.25° to match the grid coordination system of the CaMa-Flood model. If the center of each 0.25° grid was within the observation range of the path data, the grid was considered observed. Because the observed area was different for each day of the orbit cycle, we prepared 21 coverage masks to generate synthetic SWOT observations. Rivers wider than 50 m and within the coverage mask (Figure 3a, right panel) were considered as observed. We also included observation error in the synthetic observations to represent measurement errors. We simulated SWOT observation errors using a mean value of zero and standard deviation of 10 cm, in accordance with the SWOT mission goal of measurement accuracy (Desai et al., 2018; Esteban-Fernandez, 2017). We modeled the measurement error h_{err} as follows:

$$h_{err} = N\left(0, \frac{1}{WL}\sigma_h\right) \quad (1)$$

where W is the river width and L is the river length. In this study, L was set to 1 km to include only observations near the unit catchment mouth, due to internal variation inside the unit-catchment in the CaMa-Flood hydrodynamic model. The term σ_h represents observation error, as described in the SWOT mission goal (Desai et al., 2018; Esteban-Fernandez, 2017), and is equivalent to 10 cm for a water area ≥ 1 km^2 and 25 cm for 1 km^2 > water area ≥ 0.625 km^2. The observation error variance is illustrated in Figure S1 and described in Text S4.

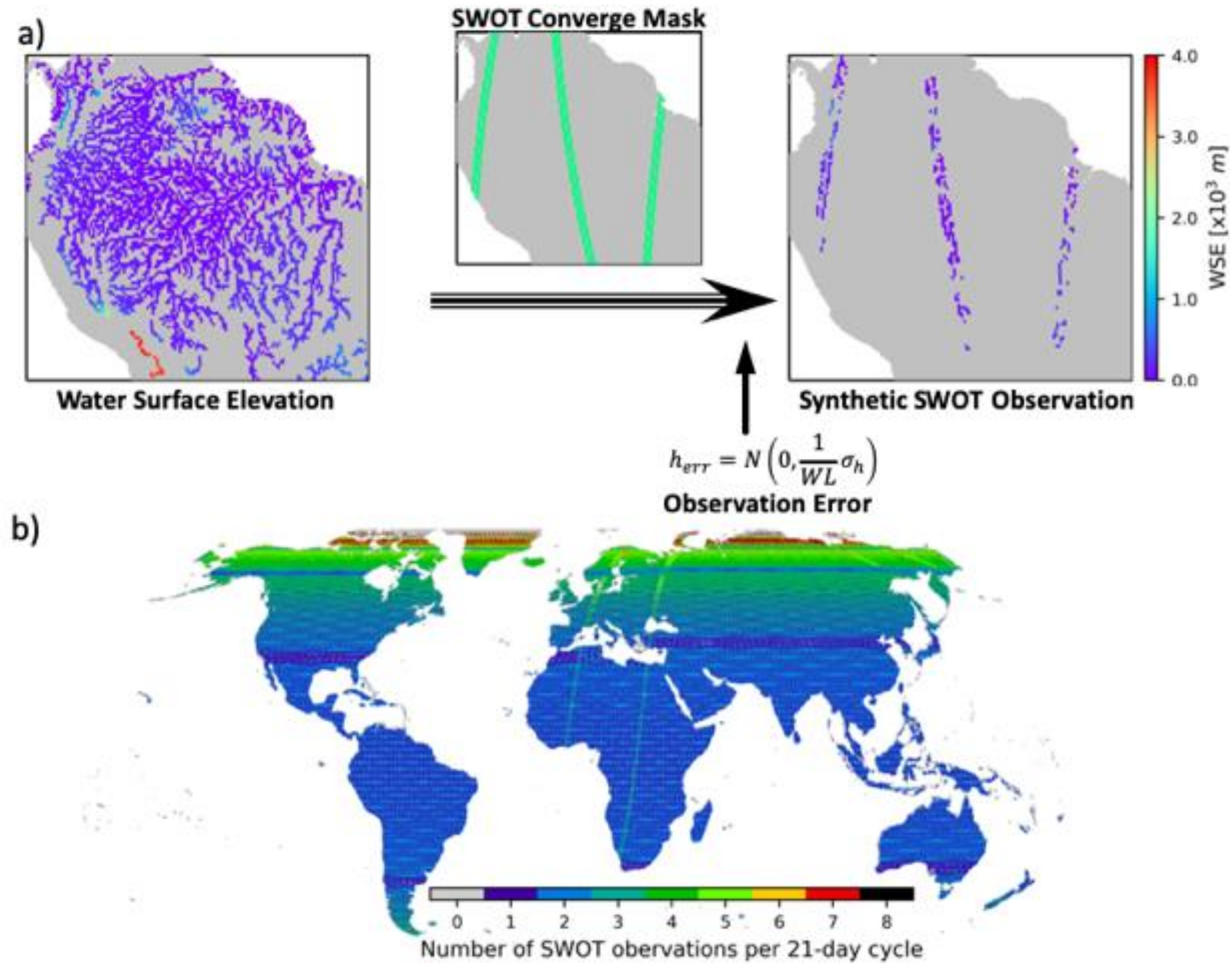

Figure 3: Generation of synthetic SWOT observations. (a) True water surface elevation (left), SWOT coverage mask (center), and synthetic observations (right). (b) Number of SWOT observations within the 21-day cycle. Only ground observations are shown.

3.2.3 Experimental conditions

We performed two experiments with different model settings: perfect and imperfect. The experimental conditions are summarized in Table 1. The perfect model experiment included errors in input runoff forcing, whereas model errors and unrealistic forcing were assumed in the imperfect model experiment.

Table 1: Experimental conditions, including experiment name, Manning's coefficient conditions for true and corrupted /assimilated simulations, as well as input runoff forcing for true and corrupted /assimilated simulations.

Experiment	*Simulation*	*Manning's Coefficient*	*Input Runoff Forcing*
Perfect model experiment	True	Global Constant (0.030)	Runoff forcing from HTESSEL LSM
	Corrupted	Global Constant (0.030)	Modified Ensemble runoff from E2O LSM/GHMs
	Assimilated	Global Constant (0.030)	Modified Ensemble runoff from E2O LSM/GHM s
Imperfect model experiment	True	Spatially varied value depends on the river width	Runoff forcing from HTESSEL LSM
	Corrupted	Global Constant (0.030)	Modified Ensemble runoff from E2O LSM/GHMs
	Assimilated	Global Constant (0.030)	Modified Ensemble runoff from E2O LSM/GHMs

a) Perfect model experiment

To assess the effectiveness of assimilation when a good model is available, we performed a 'perfect' model experiment in which we assumed that there were no errors in the hydrodynamic model. In this experiment, we used the same model parameters (e.g., Manning's coefficient, river channel depth, and river width) for all three simulations. However, the input runoff forcing for true and corrupted/assimilated simulations had different runoff inputs from different LSMs/GHMs.

b) Imperfect model experiment

We also performed an experiment to evaluate river discharge estimates obtained using data assimilation under erroneous model conditions, because modeled river states differ from those of real rivers due to uncertainties in the model physics and parameters. In this experiment, the corrupted/assimilated simulation was performed using a global constant value for Manning's coefficient and corrupted input runoff forcing. We assumed the model error can be represented using the error of the Manning's coefficient values in the hydrodynamic model due the large uncertainty in estimating true Manning's coefficients. To represent model uncertainty, we used different Manning's coefficients for the true and corrupted simulations. For the true simulation, we used a different value for each pixel, depending on the river width. We used a constant value of 0.030 for all river pixels in the corrupted and assimilated simulations. The spatially distributed Manning's coefficient for the true simulation was modeled as described by (Pedinotti et al., 2014), as shown below:

$$n = n_{min} + (n_{max} - n_{min})\left(\frac{W_{max} - W}{W_{max} - W_{min}}\right) \tag{2}$$

where n is the unit catchment average Manning's coefficient at river width W. n_{max} and n_{min} are maximum and minimum Manning's coefficients, which are 0.025 and 0.035, respectively. W_{max} and W_{min} are the maximum and minimum river widths for the river basin in the MERIT hydro river network map (Yamazaki et al., 2019). Figure 4 shows the spatial distribution of Manning's coefficient, modeled using equation (2). Relatively small upstream river sections have relatively large Manning's coefficient values (≈ 0.035), whereas relatively large downstream river sections have relatively small Manning's coefficient values (≈ 0.025). In this study, we ignored the uncertainties associated with the model physics to reduce the complexity of the data assimilation framework.

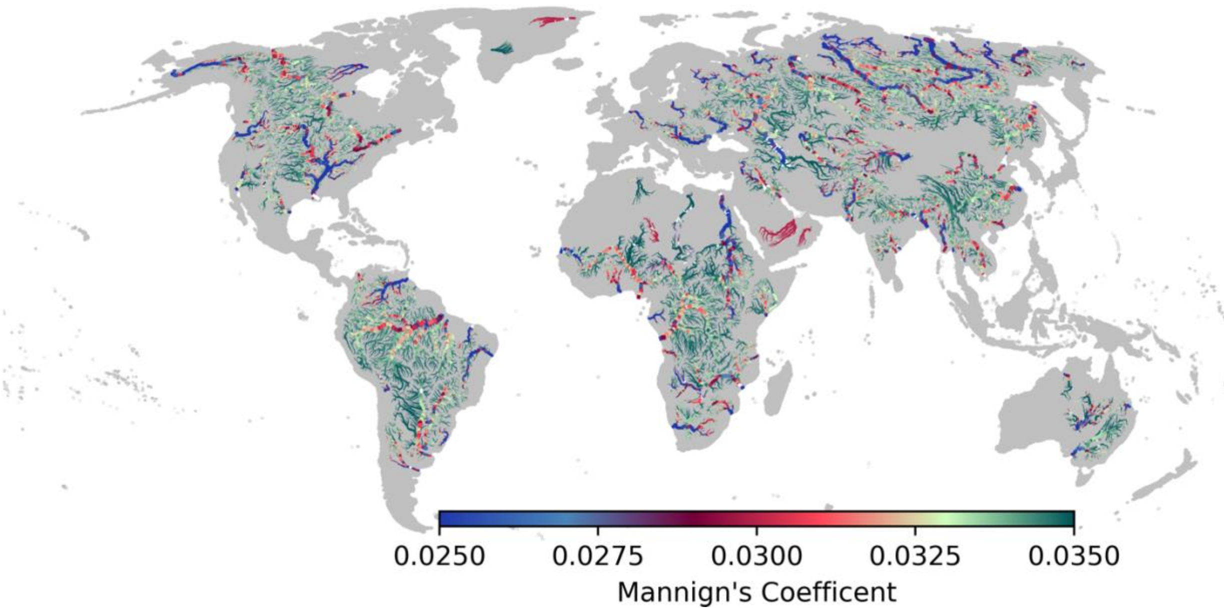

Figure 4: Spatially distributed Manning's coefficient modelled using the equation (2).

3.3 Assimilation diagnostics

The effectiveness of the data assimilation framework was evaluated by comparing the assimilated water state and the true water state. If these two water states were similar, the assimilation framework was considered effective. However, if the assimilated water state was similar to the corrupted water state, then the assimilation framework had not significantly improved river discharge estimates. We used several indices to evaluate the effectiveness of the assimilation framework. The 'assimilation index' (AI) was used to evaluate the instantaneous effect of assimilation on a daily timescale. The Nash–Sutcliffe efficiency (NSE; Nash & Sutcliffe, 1970) based AI (NSEAI) was used to evaluate the effectiveness of assimilation over the entire simulation period of 366 days. In addition, the Kling-Gupta efficiency (KGE; Gupta et al., 2009; Kling et al., 2012), which is often used to evaluate model results against observations, was used to measure the accuracy of assimilated river discharge estimates. AI and NSEAI were used to assess improvements in assimilated river discharge compared to that of true and corrupted simulation results, whereas the KGE was used to evaluate the expected accuracy of assimilated river discharge estimates.

3.3.1 Assimilation index

We introduced a new metric, the AI, to evaluate the effectiveness of data assimilation in a virtual experiment. The AI was calculated from the ratio of instantaneous river discharge error rates in the assimilated and corrupted simulations using equation (3):

$$AI = 1 - \left|\frac{Q_T - Q_A}{Q_T - Q_C}\right| \tag{3}$$

where Q_T, Q_A, and Q_C represent daily discharge from true, assimilated, and corrupted simulations, respectively. Here, Q_A and Q_C represent the mean discharge from ensemble members in each simulation.

The AI describes the similarity between the assimilated and true simulations, compared to the similarity between the corrupted and true simulations. A high AI (maximum of 1) indicates that the assimilated discharge estimate is close to the true discharge, whereas a low AI indicates that the assimilated discharge estimate is not significantly better than the corrupted discharge estimate. An AI value of less than zero indicates that the assimilated simulation includes more error than the corrupted simulation. The AI represents the relative effectiveness of data assimilation and, in contrast to the NSE coefficient (Nash & Sutcliffe, 1970), is not a measure of simulation accuracy. In addition, the AI can be calculated for any time and location during the experiment. Consequently, we can potentially identify when and where the data assimilation framework was effective in estimating river discharge. Therefore, the AI may be used to evaluate the instantaneous effects of our data assimilation strategy.

3.3.2 Nash–Sutcliffe efficiency-based assimilation index

The NSEAI was also calculated to evaluate the effectiveness of assimilation, because when the results of the corrupted and true simulations are similar, the AI may not provide the best assessment of assimilation effectiveness. The NSEAI was calculated by comparing the difference between the NSE values for assimilated and corrupted simulations, as in equation (4):

$$NSEAI = \frac{NS_A - NS_C}{1 - NS_C} \tag{4}$$

where NS_A and NS_C are the NS coefficients for the assimilated and corrupted simulations, respectively. The relative difference between the accuracy of these two simulations is given by equation (4). The NSEAI represents the overall effectiveness of assimilation over the entire simulation period. The NSEAI avoids any over-evaluation that may be due to the corrupted simulation coinciding with the true simulation.

3.3.3 Kling–Gupta efficiency

We used the KGE to compare the accuracy of the assimilated and corrupted simulations. The KGE is an integrated skill metric which is based on a combination of three diagnostically meaningful components of the mean squared error. The mean squared error for the simulated and observed discharge can be separated into three components including the mean, variability, and dynamics (Gupta et al., 2009). These components can be represented by the correlation coefficient (*CC*), bias ratio (*BR*), and relative variability (*RV*). The KEG is given by equation (5):

$$KGE = 1 - \sqrt{(CC-1)^2 + (BR-1)^2 + (RV-1)^2} \tag{5}$$

where;

$$CC = \frac{cov(Q_m, Q_o)}{\sigma_{Q_m}\sigma_{Q_o}} \tag{6}$$

$$BR = \frac{\mu_{Q_m}}{\mu_{Q_o}} \tag{7}$$

$$RV = \frac{\left({}^{\sigma_{Q_m}}/_{\mu_{Q_m}}\right)}{\left({}^{\sigma_{Q_o}}/_{\mu_{Q_o}}\right)} \tag{8}$$

where $Q, \mu,$ and σ are the discharges, the mean of the discharges, and the standard deviation of the discharges, respectively. The subscripts m and o represent the simulated (assimilated/corrupted) and true discharges, respectively. In addition, conventional metrics, such as the percent bias (pBias) of annual mean river discharge, were also used in our evaluations.

4. Results

4.1 Perfect model experiment

Here, we describe the results of the perfect model experiment. First, we evaluated the results obtained for the Amazon basin, the world's largest river basin, to assess the effectiveness of our data assimilation framework in river discharge estimates for continental-scale rivers. Next, we evaluated the potential effectiveness of SWOT observations in river discharge estimates on a global scale.

4.1.1 Amazon River basin

Figure 5a–c shows the temporal variation in simulated river discharge at three locations on the Amazon River: upstream location X (1.125°S,74.875°W), midstream location Y(1.625°S,67.625°W), and downstream location Z(0.875°S,51.125°W), respectively. Black, blue, and red lines represent the discharge for the true, corrupted, and assimilated simulations, respectively. The green lines illustrate the temporal variation in AI defined by equation (3). The AI is marked only for those days in which the true and corrupted discharge showed significant error (>5%) because the AI is much lower, despite the effectiveness of the assimilation, when the two discharges are similar. Green circles indicate the days when the assimilation was performed at each location. In addition, the NSEAI and the transition of the annual mean pBias from a corrupted to an assimilated value is shown in the top right corner of the graph.

At downstream location Z, the assimilated discharge was almost identical to the true discharge, although the initial conditions were generated using corrupted runoff. The AI at location Z remained >0.8 during most of the simulation period and was generally stable, regardless of SWOT assimilation availability. There were some low AI values when the discharge error between

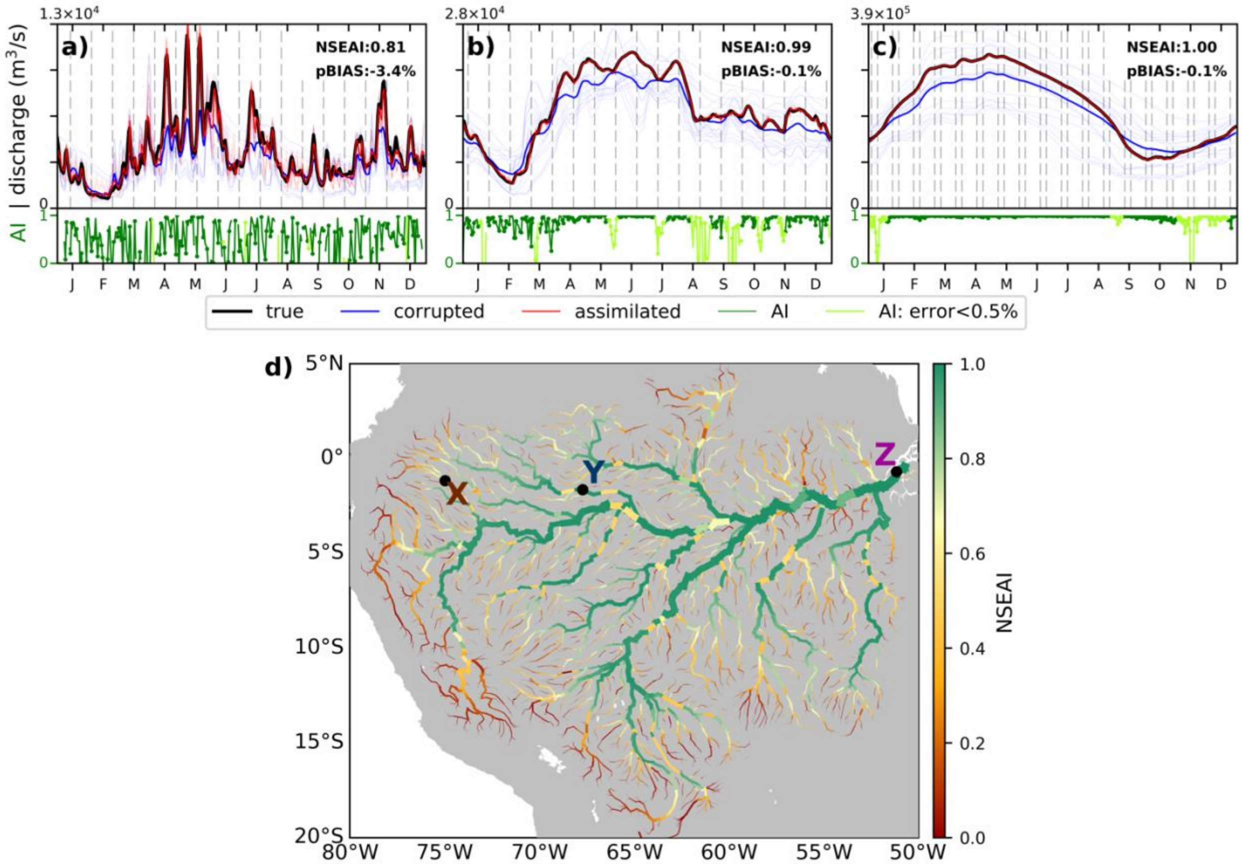

Figure 5 : Simulated river discharge and Nash-Sutcliffe efficiency-based assimilation index (NESAI) in the Amazon basin. Discharge hydrographs at the a) upstream X(1.125°S,74.875°W), b) midstream Y(1.625°S,67.875°W), and c) downstream Z(0.625°S,51.125°W). River discharges of true, assimilated, and corrupted simulations are shown by black, red, and light blue lines, respectively. Dark blue line represents the ensemble mean of corrupted simulation. The dashed-grey vertical lines indicate the times of direct SWOT observations. The assimilation index (AI) (green line in lower panel) is shown for days when the error between the true and corrupted discharges was >5%. Light green line indicates the AI when error was < 5%. Green dots represent the times of data assimilation. The NSEAI and percent bias (pBIAS) of the assimilated simulation are shown in the upper right corner of the hydrographs. (b) map of NESAI. The locations of hydrographs X, Y, and Z are annotated by black dots.

the true and corrupted simulations was < 5% in Figure 5c. At the upstream location X, the overall trend for discharge in the assimilated simulation was similar to that observed for the true simulation (Figure 5a). However, the difference between the two discharges was unstable and varied throughout the year. In addition, AI values fluctuated significantly and were frequently large on assimilation days (i.e., days with green circles on the AI graph) but small thereafter. At the midstream location Y (Figure 5b), discharge in the assimilated simulation generally matched true discharge, as it did at downstream location Z. However, fluctuations observed at location Y were greater than those observed at location Z. The NSEAI value was greatest at location Z and lowest at upstream location X. The NSEAI value at location Y was intermediate.

Figure 5d shows the spatial distribution of NSEAI values throughout the Amazon basin. The NSEAI was calculated for each grid to compare the effectiveness of data assimilation at different locations. Because the runoff error varied over time, AI values also fluctuated. Therefore, the level of agreement among overall discharge values for true and assimilated simulations also varied over time. To obtain the overall discharge correction, we developed the NSEAI (equation 4) to evaluate the effectiveness of assimilation. At downstream location Z, the NSEAI value was greater (1.00) than at upstream location X (0.98) or midstream location Y (0.78). The Amazon basin and associated river branches (i.e., the Amazon, Solimoes, and Madeira Rivers) had high NSEAI values (> 0.8), indicating effective assimilation. Other large river branches (e.g., the Tapajos, Negro, Purus, Juura, and Paura Rivers) also had relatively high NSEAI values (> 0.6), whereas most minor river branches had lower NSEAI values (< 0.4). These results imply that SWOT data assimilation is more effective at improving river discharge estimates in large rivers than in upstream river reaches.

NSEAI values were strongly influenced by two factors: the local state correction and the upstream inflow correction. Local state correction is a water state update derived from SWOT data assimilation at the current location. This is a direct correction based on SWOT observations and occurs only when a SWOT observation is available within its' local patch. Local state corrections were available only once every few days, particularly for upstream river reaches, because the empirical local patches were not large enough to accommodate SWOT observations every day. The effects of local state corrections were most obvious at upstream location X (Figure 5a). Here, the AI rapidly increased when local SWOT observation assimilations occurred (green circles, Figure 5a) Local state corrections were effective at locations where the river water flow from upstream was less than the surface and subsurface runoff. At these pixels, the variation among the ensemble is likely to increase rapidly because the water state is sensitive to a change in external conditions (e.g., runoff forcing). This increase in ensemble variation affected the Kalman gain matrix and decreased model weight while increasing observation weight at data assimilation. In addition, at these pixels, river discharge differences between the assimilated and true simulations increased when assimilation was unavailable because the water state is sensitive to surface and subsurface runoff and river discharge fluctuates frequently. Therefore, AI values decreased when assimilation was unavailable.

The second factor strongly influencing NSEAI values, upstream inflow correction, is caused by previously corrected water states from upstream pixels. Downstream pixels are more accurate when the water state at upstream pixels is corrected, because the upstream water state significantly affects river discharge at downstream locations. This factor is independent of local SWOT observation availability because it has an indirect effect on local data assimilation. The upstream inflow correction was significant at downstream location Z (Figure 5c). Here, AI values remained high (mean AI= 0.96) and the NSEAI reached its maximum value (i.e., 1.00) because the upstream drainage area at this location is large. Consequently, upstream inflow correction was constantly available somewhere within the catchment area. AI values did not fluctuate as much at location Z as they did at locations X or Y. Local state and upstream inflow corrections were both significant factors at midstream location Y (Figure 5b). Upstream inflow correction was observed as increased AI values on days on which no local observations were recorded. However, the AI values showed some fluctuations due to the smaller drainage area at location Y, compared to that at downstream location Z. The increases in AI values at location Y, which were due to local state corrections, were smaller than those observed at upstream location X. At location Y, the combination of upstream inflow and local state corrections meant that the NSEAI was high (0.98). At upstream location X, the upstream inflow correction was smaller and AI values generally increased at the days of assimilation. At location X, the catchment area was smaller (44916.08

km2) than at locations Y or Z, and location X also had a smaller upstream river reach (width > 50 m) where SWOT observations could be made. Therefore, at location X, there was a lower probability of inflow assimilation upstream, which limited the effect of inflow corrections. To evaluate the importance of upstream inflow corrections further, we performed an extra experiment in which upstream inflow corrections were excluded (Text S5 and Figure S2). We found that AI values significantly decreased when upstream inflow corrections were excluded, suggesting that data assimilation must be applied to the entire upstream region to estimate discharge in a large basin accurately.

Figure 6 shows the percentage deviation from true value (pBias) of annual river discharge for the assimilated (Figure 6a) and corrupted (Figure 6b) simulations compared to the true simulation. As with the NSEAI values (Figure 5d), the annual mean river discharge was accurately estimated in the major branches of the Amazon River. The pBias of the assimilated river discharge was almost zero, particularly in the downstream reaches of the Amazon mainstream (50–65°W), which indicates that the virtual assimilation framework succeeded in accurately estimating the total quantity of water flowing from the Amazon basin to the Atlantic Ocean. These results are promising because estimates for the terrestrial water budget within the global hydrological cycle are generally unreliable where sparse distribution of river gauges makes it difficult to estimate the total freshwater discharge from land to the oceans.

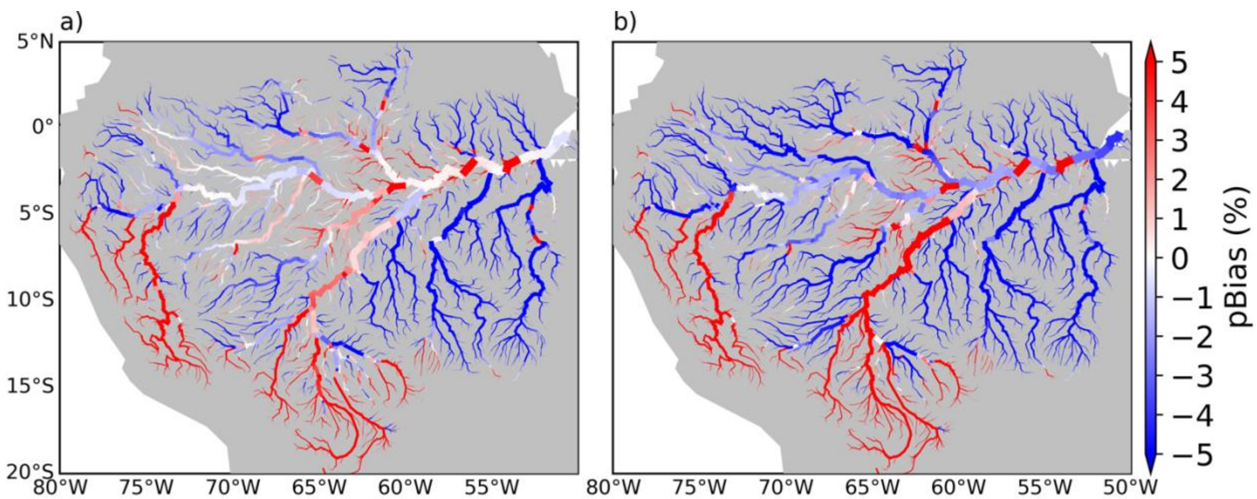

Figure 6 : Percentage bias (pBias) of river discharge in Amazon basin: a) between the assimilated and true simulations (left), and b) between the corrupted and true simulations (right).

4.1.2 Global-scale analysis

Figure 7 shows simulated river discharge hydrographs for 42 locations, which are also listed in Table 2. Table 2 describes the geographical details and assimilation diagnostics of each location. We selected three locations for each river: upstream, midstream, and downstream. Each row of panels in Figure 7 represents one river (e.g., row 1 of Figure 7 shows the hydrographs for the Amazon River: upstream, midstream, and downstream from left to right, respectively). In addition, the geographical location of each hydrograph is marked in Figure 8. The locations are indicated by two upper-case characters in the upper-left corner of each panel in Figure 7 (e.g., PK = Prek Kadam). Furthermore, NSEAI values and the pBias of the assimilated simulation are shown in the upper-left corner of each panel (in the PK panel, these are shown upper center). We deliberately included PK, which is not on the main stream of the Mekong River but situated between the Tonlè Sap and the Mekong, to discuss the effectiveness of our assimilation framework in scenarios involving reverse flow.

Generally, the distribution of NSEAI values was similar to that observed in the Amazon River basin, with high NSEAI values downstream and low values upstream. Among the 14 rivers shown in Figure 7, the Amazon, Ob, Lena, Mississippi, Volga, Ganges–Brahmaputra, Nile, Yangtze, Yukon, Indus, and Irrawaddy generally had higher NSEAI values downstream than upstream. For the Congo River, the Kinshasa (KH) location had a slightly lower NSEAI value (0.74) than upstream locations due to the corrupted simulation result, which was very similar to the true simulation. Similarly, for the Niger River, the downstream Lokoja (LJ) location had a slightly lower NSEAI value (0.93) than upstream locations. In addition, PK had a high NSEAI value (0.97) although PK is not on the main stream of the Mekong River. Most of the other hydrographs shown in Figure 7 had higher NSEAI values at downstream than upstream locations.

Some important hydrodynamic processes (e.g., reverse flow towards Tonlé Sap in the Mekong River, glacial runoff in the Yukon River, and backwater effects in the Amazon River) were captured using our assimilation framework and the CaMa-Flood hydrodynamic model. Reverse flow towards Tonlé Sap in the wet season was well-characterized by the CaMa-Flood model. In addition, negative discharges at PK (Figure 7) were also well estimated by our data assimilation framework, achieving an NSEAI value of 0.97. Glacier runoff from the Llewellyn Glacier at Atlin Lake is the main source for the Yukon River in Alaska, but in LSMs the glacier runoff process may not be well-characterized (Hock, 1998; Hock & Holmgren, 2005; Zhao et al., 2013). Clearly, the timing of peak flow for the true and corrupted simulations was different. Our data assimilation framework can be used to accurately estimate discharge even if glacier runoff is not modeled successfully (e.g., the Dawson [DW], Stevens Village Ak [SV], and Kaltag, alas. [KA] in Figure 7). The discharge along the main stream of the Amazon River was estimated accurately (NSEAI > 0.99). In addition, hysteresis in the stage–discharge relationship due to backwater effects has also been successfully modeled in the Amazon River using the CaMa-Flood hydrodynamic model (Yamazaki et al., 2012). Because we assimilated the WSE by including SWOT observations, the water-surface dynamics were well-characterized. Therefore, important hydrodynamic processes such as reverse flow, glacier-runoff-induced river discharge, and backwater effects were captured using our assimilation framework and the CaMa-Flood hydrodynamic model.

The large lakes and their downstream reaches had relatively low NSEAI values compared to nearby reaches. For the Great Lakes of North America, the NSEAI values were slightly smaller than those of the surrounding river reaches (Figure 8). For such large lakes, the quantity of discharge may be determined from upstream water flow, particularly during the dry season. Due to less rainfall and runoff, the ensemble of forecasted water states is less diverse because the quantity of inflowing runoff is small compared to the total water storage. The small variation in forecasted WSE decreased model variance in the Kalman gain matrix, which hampered data assimilation and decreased NSEAI values. However, lakes situated downstream of these rivers had high NSEAI values due to large upstream inflow correction from upstream reaches. For example, the Caspian Sea, which is situated downstream of the Volga River, had very high NSEAI values. Therefore, upstream river reaches where upstream flow is greater than surface and subsurface runoff showed low assimilation efficiencies, whereas downstream river reaches showed higher assimilation efficiencies.

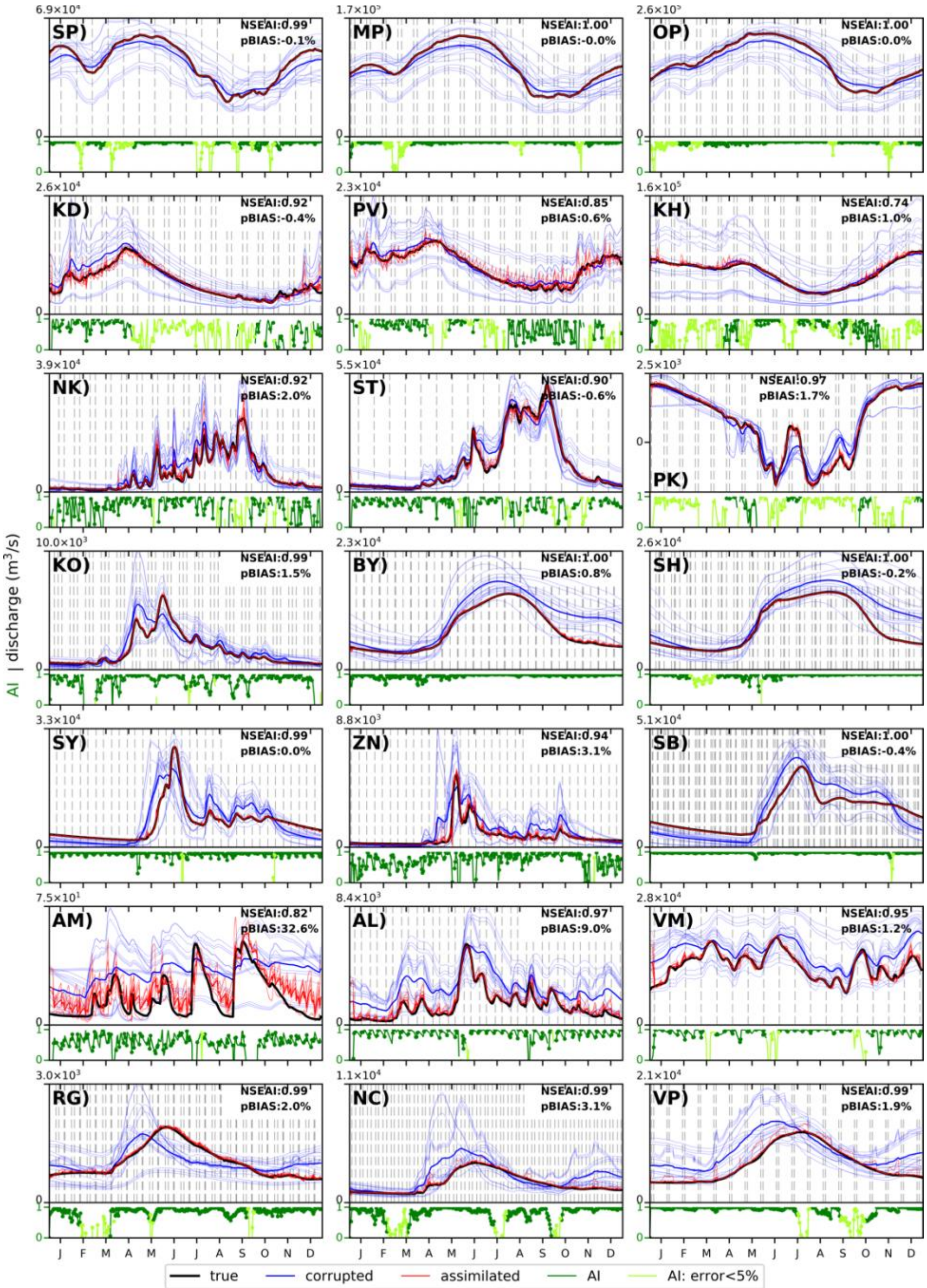

Figure 7 : Simulated discharge for the perfect model experiment on a global scale. The locations are summarized in Table 2. River discharges of true, assimilated, and corrupted simulations are shown by black, red, and light blue lines, respectively. Dark-blue line represent the ensemble mean of corrupted simulation. The dashed-grey vertical lines indicate the times of direct SWOT observations. The assimilation index (AI) (green line in lower panel) is shown for days when the error between the true and corrupted discharges was >5%. Light green line indicates the AI when error < 5%. Green dots represent the times of data assimilation. The colors and key are identical to those shown for Figure 5a-c.

572

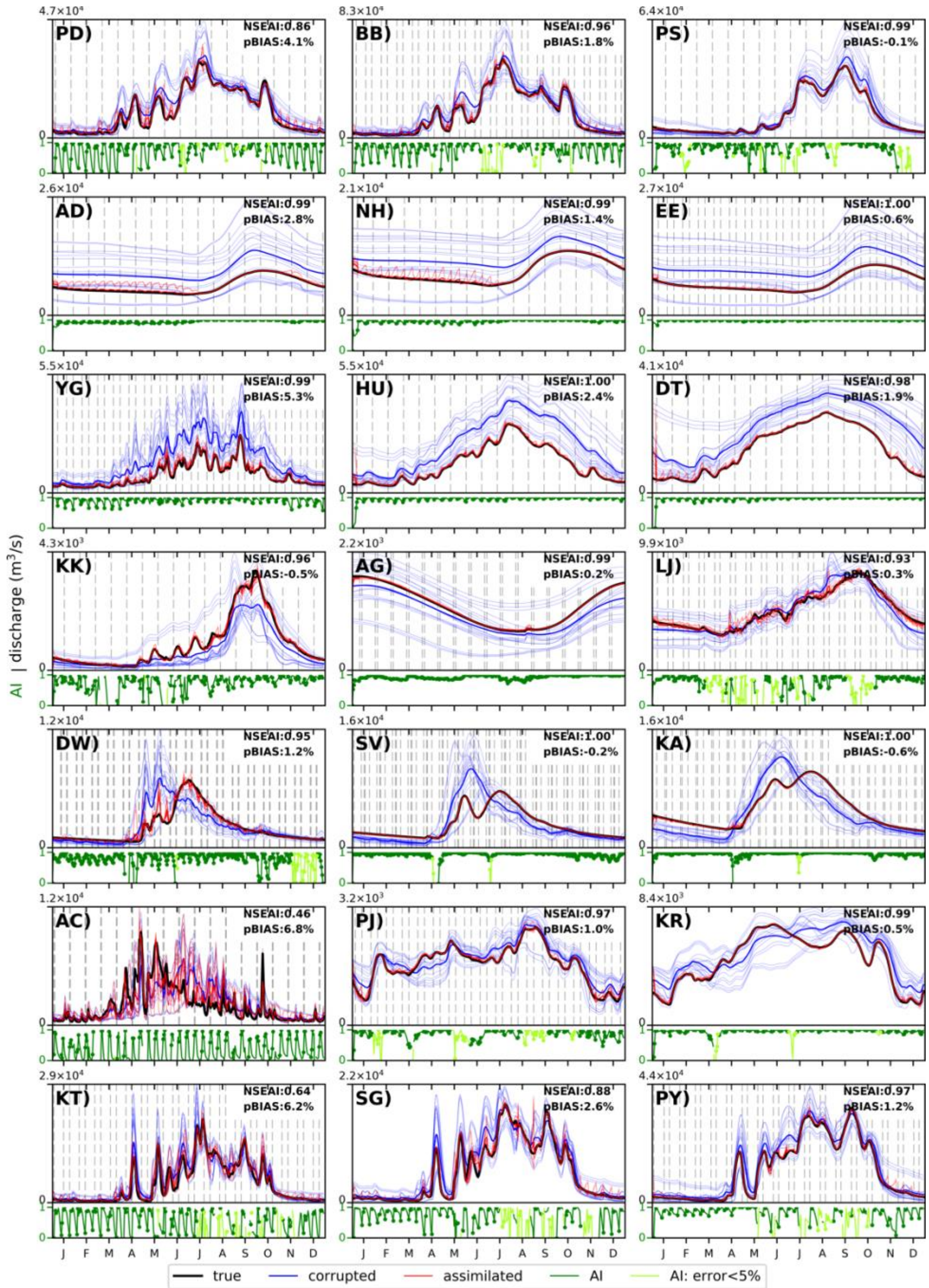

Figure 7. (continued)

Small upstream river reaches located at high latitudes had high NSEAI values, whereas similar river sections at low latitudes had lower NSEAI values. River reaches flowing toward Hudson Bay in North America had high NSEAI values (> 0.7), whereas upstream river reaches in high-mountain Asia had low NESAI values (< 0.4;Figure 8). A similar trend was observed in the hydrographs of most upstream locations of the Indus (Attock [AC]) and Irrawaddy (Katha [KT])

Table 2 : Summary of Figure 7, including river name, station name, location of station, upstream catchment area, number of accumulated overpasses per SWOT cycle, NSEAI, as well as pBias for assimilated and corrupted simulations.

	River	Station	Longitude [°]	Latitude [°]	Catchment Area [km^2]	Accumulated Overpasses	NSEAI	pBais [%] (assimilated)	pBias [%] (corrupted)
SP)	Amazon	Sao Paulo de Olivenca	-68.73	-3.41	1009573	525	0.99	-0.11	-4.26
MP)		Manacapuru	-60.14	-3.33	2204488	1172	1.00	-0.02	-5.06
OP)		Obidos - Porto	-55.50	-1.94	4672153	2490	1.00	0.00	-4.14
KD)	Congo	Kindu	25.87	-2.74	785477	419	0.92	-0.36	8.51
PV)		Ponthierville	25.50	-0.29	932105	496	0.85	0.64	3.75
KH)		Kinshasa	15.25	-4.32	3606257	1972	0.74	0.95	0.47
NK)	Mekong	Nong Khai	102.83	17.95	304814	185	0.92	2.05	16.45
ST)		Stung Treng	105.93	13.50	637963	43	0.90	-0.64	6.28
PK)		Prek Kdam	104.83	11.75	87708	43	0.97	1.67	3.97
KO)	Ob	Kamen'na Obi	81.33	53.87	212392	182	0.99	1.51	9.02
BY)		Belogor'ye	68.50	61.08	2227911	2075	1.00	0.84	31.53
SH)		Salekhard	66.58	66.64	2490615	2439	1.00	-0.16	24.64
SY)	Lena	Solyanka	120.75	60.48	774718	741	0.99	0.00	6.21
ZN)		Zmeinovo	108.56	57.96	139284	127	0.94	3.11	43.46
SB)		Stolb	126.75	72.42	2451204	3089	1.00	-0.40	3.80
AM)	Mississippi	Aitkin, MN	-93.76	46.56	14914	16	0.82	32.64	93.98
AL)		Alton, ill.	-90.34	38.94	443250	270	0.97	9.03	86.23
VM)		Vicksburg, MS	-90.97	32.25	2932631	2048	0.95	1.23	23.18
RG)	Volga	Rybinskaya Ges	39.00	58.04	152867	170	0.99	2.03	-0.81
NC)		Naberezhnyye Chelny	52.25	55.70	369824	0	0.99	3.07	32.84
VP)		Volgograd Power Plant	44.61	48.75	1364059	1315	0.99	1.90	30.96
PD)	Ganges-Brahmaputra	Pandu	91.33	26.10	412713	251	0.86	4.13	18.10
BB)		Bahadurabad	89.59	24.92	512475	312	0.96	1.84	17.46
PS)		Paksey	89.30	23.92	941267	558	0.99	-0.13	9.47
AD)	Nile	Aswan dam	32.90	24.12	2911499	1569	0.99	2.81	52.55
NH)		Nag Hammadi	32.13	26.20	2990389	1615	0.99	1.39	34.97
EE)		El Ekhsase	31.27	29.75	3032569	1642	1.00	0.56	46.78
YG)	Yangtze	Yichang	111.32	30.66	976284	5	0.99	5.31	82.21
HU)		Hankou	114.36	30.65	1441953	93	1.00	2.44	69.23
DT)		Datong	117.74	31.03	1677326	1012	0.98	1.91	50.56
KK)	Niger	Koulikoro	-7.50	12.94	118079	65	0.96	-0.51	-24.22
AG)		Ansongo	0.50	15.53	489888	271	0.99	0.22	-11.12
LJ)		Lokoja	6.75	7.67	1990463	1078	0.93	0.27	-1.39
DW)	Yukon	Dawson	-139.50	64.15	265258	292	0.95	1.24	1.97
SV)		Stevens Village Ak	-149.79	65.88	501116	762	1.00	-0.21	-1.55
KA)		Kaltag, alas.	-158.65	64.20	756273	1177	1.00	-0.55	-11.11
AC)	Indus	Attock	72.24	33.90	199682	132	0.46	6.77	9.90
PJ)		Panjnad	71.00	29.33	278545	169	0.97	0.99	6.31
KR)		Kotri	68.33	25.25	830667	515	0.99	0.48	10.08
KT)	Irrawaddy	Katha	96.27	23.96	84522	51	0.64	6.17	14.54
SG)		Sagaing	95.99	21.87	124739	22	0.88	2.60	15.03
PY)		Pyay	95.16	18.70	360734	206	0.97	1.21	13.61

rivers (Figure 7). These incremental changes in the NSEAI values of small river reaches occur because of the frequent SWOT observations recorded at high latitudes. At low latitudes, observation frequencies are much lower. Within the 21-day orbital cycle of the SWOT satellite (Figure 3b), there will be more than four observations at high latitudes (> 50° N), compared to only one or two observations at low latitudes. Note that assimilation frequencies are high because we used an adaptive empirical local patch in the LETKF data assimilation framework. The number of assimilations in the 21-day orbital cycle was more than 10 days at > 50°N and less than 10 days at < 30°N. However, large rivers, such as the downstream reaches of the Amazon River can be assimilated almost every day. Therefore, the assimilation efficiency of small river reaches depends on assimilation frequency.

Due to the high frequency of SWOT observations, rivers located at high latitudes had high NSEAI values, even at upstream locations. Figure 9 shows the relationship between NSEAI values and upstream drainage area, taking latitude into account. The black dotted line in Figure 9a shows the relationship between NSEAI values and drainage area, whereas the gray curve indicates the 1–σ range. At low latitudes, upstream reaches and small rivers had small NSEAI values (Figure 9b) due to low observation frequencies; however, NSEAI values were higher (> 0.8) at locations with larger drainage areas. The downstream reaches of the Congo River had low NSEAI values (~0.6) because corrupted river discharge values were similar to the true values. At high latitudes (> 60° N; Figure 9d), almost all river reaches had very high NSEAI values (> 0.8). This was because more SWOT observations could be made within the catchment of rivers with larger drainage areas. The mid latitudes (20°–60°) also showed similar NSEAI trends: low values upstream and high values downstream. However, at higher latitudes (>60°), despite the large catchment areas, most of the river pixels showed improved assimilation. There was a weaker relationship between NSEAI values and drainage area in rivers located at high latitudes (Figure 9d). This was due to the increase in SWOT observation frequency at high latitudes. Despite an apparently increasing trend, the relationship between NSEAI values and upstream drainage area remains unclear. The 1–σ range is large up to an upstream drainage area of 10^6 km^2. Therefore, an indicator that combines the effects of observation frequency and upstream drainage area is needed to understand the global variation in assimilation efficiencies.

Figure 10a shows the boxplot of NSEAI with the accumulated overpasses per SWOT cycle The number of. accumulated overpasses for a particular location was obtained by totaling all SWOT overpasses in river reaches upstream of that location. A global map of the total number of overpasses per SWOT cycle is provided in the Supplementary Information (Figure S5). The combined effect of a large upstream drainage area and observation frequency may be expressed using the total number of accumulated overpasses. In the Figure 10a, the accumulated overpasses were divided into 100 size bins to visualization purposes. Overall, it shows increasing trend but the pixels with smaller accumulated overpasses demonstrate a NSEAI high variation (0.0~1.0), especially pixels with accumulated overpasses < 100. The variation at low accumulated overpasses were mostly due to the upstream drainage area (Figure S6a) but that is not much related to the observation frequency (Figure S6b). In the low accumulated overpasses, an internal variation can

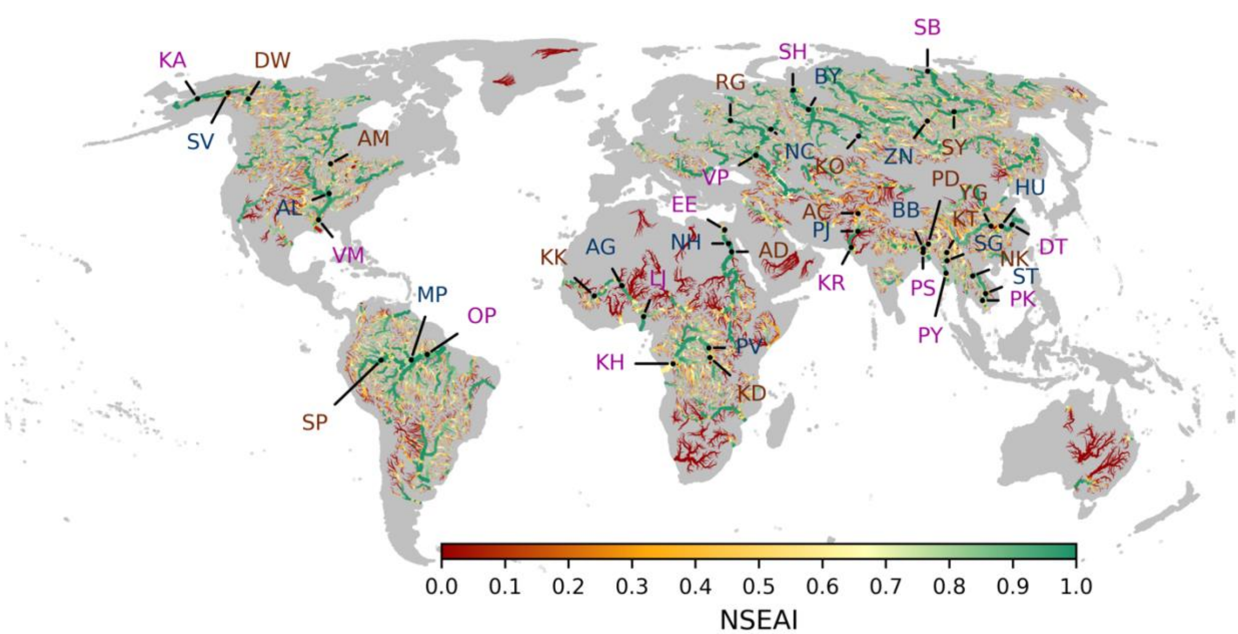

Figure 8: Global NSEAI map of the perfect model experiment. The annotations correspond to the panels shown in Figure 7. Upstream, midstream, and downstream annotations are shown in brown, blue, and violet, respectively.

be seen with the size of upstream drainage area where NSEAI was increasing with the upstream drainage area. According to the Figure 10a, it can be expected NESAI will be > 0.8 in the locations where the accumulated overpasses per SWOT cycle > 100. NSEAI become almost 1.0 in the river pixels with accumulated overpasses per SWOT cycle > 565. However, a slight deviation from high NSEAI was observed in the river reaches where accumulated overpasses are between 1901-2000.

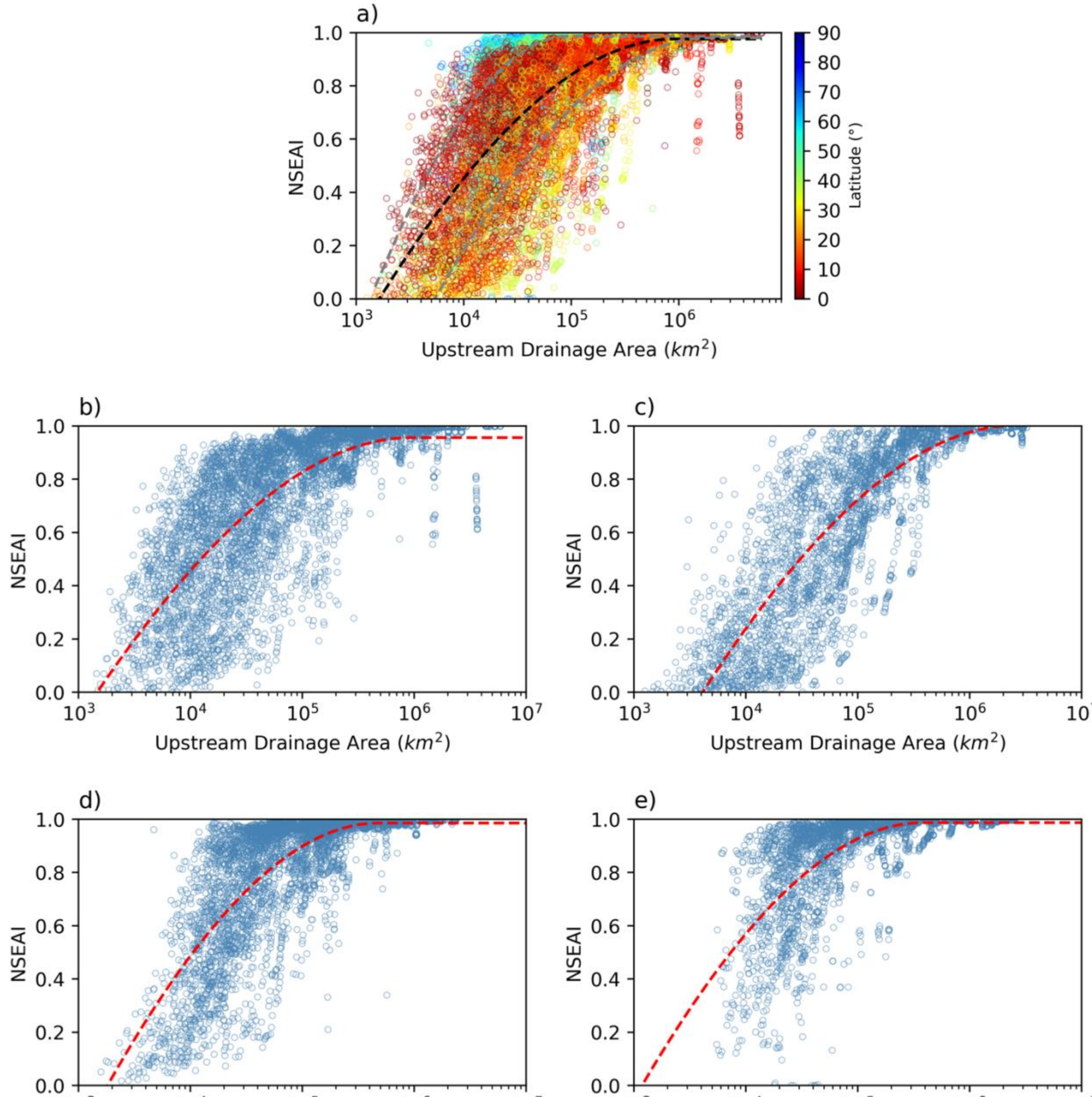

Figure 9 : Relationship between the upstream drainage area and NSEAI: a) for whole globe; b)-e) are for latitudes between 0º-20 º, 21º-40 º, 41º-60 º, 61º-80 º, respectively. The dashed-black line in panel a) and dashed-red lines in panels b)-e) presents the fitted curves for the mean values. The dashed-grey line shows the 1- σ range for the fitted mean curve. Pixels with annual mean discharge > 100 m^3/s are shown. The colors in panel a) show the latitude of each pixel.

Those NSEAI deviation are due to the low NSEAI values of the downstream reaches of Congo river because the corrupted simulation efficiency in the downstream reaches of Congo river was high (NSE>0.8). Hence, higher assimilation efficiency is expected in places where the number of accumulated overpasses per SWOT cycle higher assimilation efficiency is expected in places where the accumulated overpasses are higher in the perfect model scenario.

The assimilation frequency depends on the size of the adaptive empirical local patch. The assimilation frequency was defined as the number of assimilations per SWOT cycle which represents the local correction of each river pixel where higher assimilation frequency tends to produce high NSEAI values (Figure 10b). A global map of assimilation frequency is presented in supplementary information section (Figure S7). Empirical local patches in the upstream reaches is smaller than that of the downstream reaches (Revel et al., 2019). As explained above small upstream river reaches have low NSEAI compared to downstream. The NSEAI values are > 0.8 in the river reaches where the assimilation frequency per SWOT cycle > 11. Large variation in the NSEAI in 8-11 assimilation frequency band can be due to the variation of accumulated overpasses per SWOT cycle. Hence, both accumulated overpasses and assimilation frequency contributes to the assimilation efficiency of our assimilation framework.

The contribution of the accumulated overpasses and the assimilation frequency can be illustrated by the Figure 10c. Mean NSEAI at each assimilation frequency and accumulated overpass is shown as 2-dimensional map where colors indicate the mean NSEAI and contours

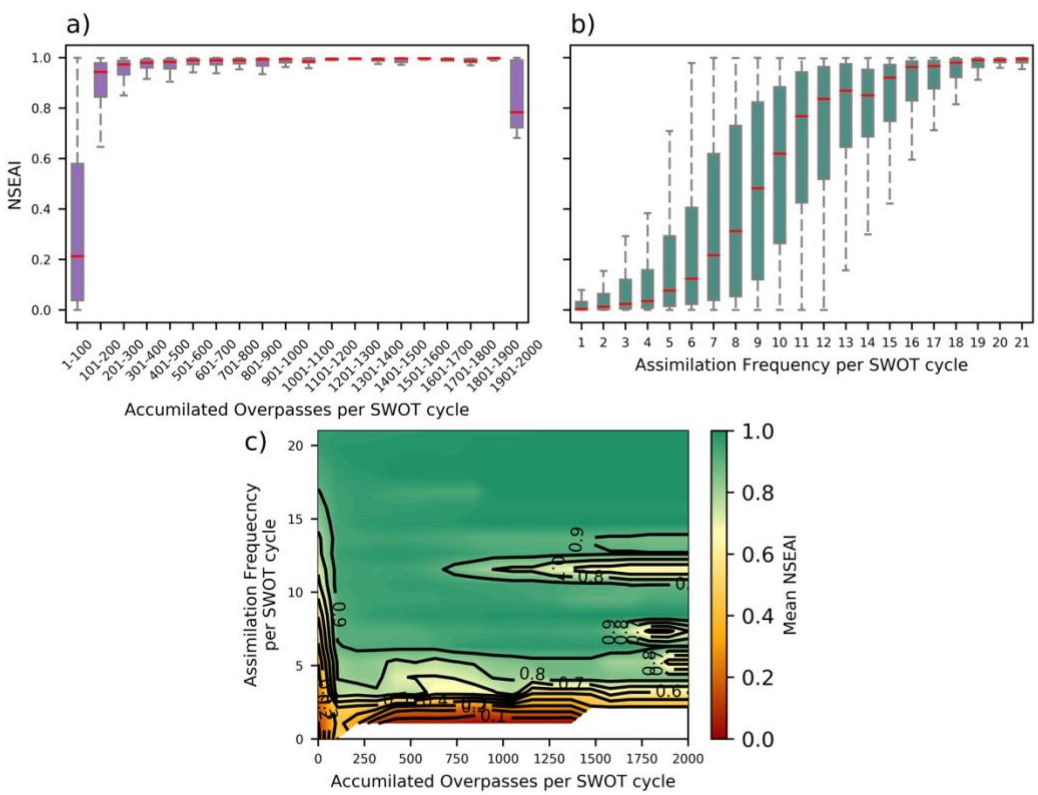

Figure 10: Boxplot of NSEAI with a) accumulated overpasses and b) assimilation frequency per SWOT cycle. c) Mean NSEAI variation with accumulated overpasses and assimilation frequency. Panel a is between 1-2000 accumulated overpasses and the data are presented in 100-overpass category intervals.

from 0.0~1.0 is plotted at 0.1 interval (black lines in Figure 10c). The river reaches with higher assimilation frequency and accumulated overpasses were shown higher NSEAI values (>0.9). Very high assimilation efficiency can be expected in the places where the accumulated overpasses > 100 and assimilation frequency > 7. Therefore, the assimilation efficacy can be explained by the combination of the accumulated overpasses per SWOT cycle (measure of upstream inflow correction) and the assimilation frequency (measure of local correction).

Continental-scale rivers show high KGE values after synthetic SWOT observations have been assimilated. Figure 11 shows a boxplot of KGE values for assimilated discharge in continental -scale rivers. Here, we show the five rivers with the largest catchment areas for each latitude band. All the rivers shown in Figure 11 have median KGE values > 0.6. KGE values provide diagnostic insight into the performance of our assimilation framework in river discharge estimates. The KGE value combines correlation, relative bias, and variability to reproduce temporal dynamics while preserving flow durations. The median KGE values of large low-latitude rivers were high (KGE > 0.9). The Yangtze River had the lowest KGE value among the 15 rivers shown in Figure 11 due to low KGE values in the small tributaries south of the river. However, almost all global rivers showed improved KGE values in assimilated compared to corrupted simulations (Text S6, Fig. S3). Therefore, the assimilated discharge estimates represented the hydrodynamics of the rivers accurately, reaching KGE values of > 0.9 for continental-scale rivers.

4.2 Imperfect model experiment

Here, we describe the results of the imperfect model experiment. Model error was represented using a spatially distributed Manning's coefficient in the true simulation, whereas a global constant Manning's coefficient of 0.03 was used in the assimilated and corrupted

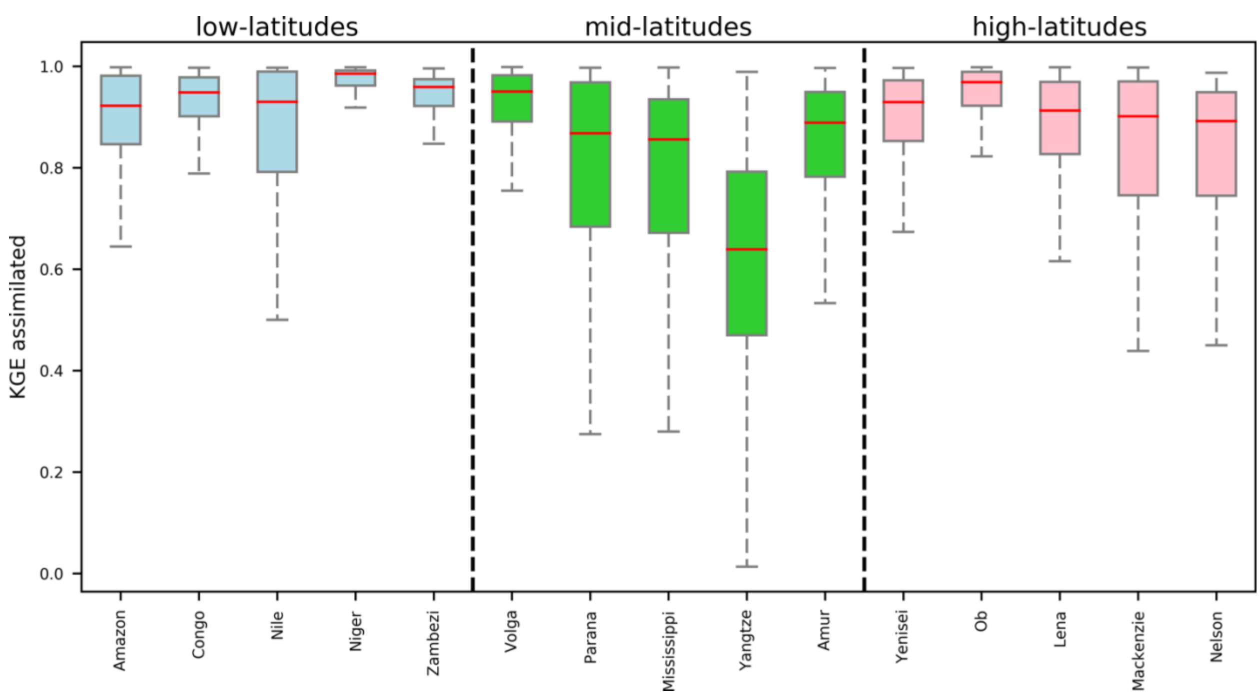

Figure 11: Kling-Gupta efficiency (KGE) of assimilated discharge in the perfect model for continental-scale rivers at low-latitudes (light blue), mid-latitudes (light green), and high-latitudes (pink). River pixels with mean river discharge > 100 m^3/s were used to create the boxplot.

simulations. Figure 12 shows hydrographs from 42 locations representing upstream, midstream, and downstream locations in 14 rivers. Figure 12 shows a global map of NSEAI values with annotated locations corresponding to those shown in Figure 7. Table 3 summarizes the assimilation efficiency diagnostics and the Manning's coefficient values used in the true simulation for the locations shown in Figure 12.The assimilation efficiency was improved using our data assimilation framework even when the model was imperfect. Most of the hydrographs shown in Figure 12 have high NSEAI values (> 0.7), except for those representing Sao Paulo de Olivenca (SP), Ponthierville (PV), KH, Paksey (PS), Ansongo (AG), AC, KT, and Sagaing(SG). Differences in peak flow timing between the true and corrupted simulations are observed in the hydrographs (Figure 12). These were due to differences in the Manning's coefficients used in each simulation. Wave propagation in the true and assimilated/corrupted simulations differed due to differences in the Manning's coefficient values used and in the local inertial equation (Bates et al., 2010; Yamazaki et al., 2011). By assimilating SWOT observations, we were able to improve peak discharge timing estimates (Figure 12). However, estimates of peak discharge differed slightly from the true values due to changes in the local stage discharge relationship which, in turn, were caused by differences in the Manning's coefficient values used for the true and was not corrected instantaneously. In conclusion, NSEAI values were high at most locations shown corrupted/assimilated simulations. It should be noted that we assimilated WSE, river discharge in Figure 12, even in the imperfect model.

Although WSE was incorporated into the CaMa-Flood hydrodynamic model, the magnitude of river discharge differed between assimilated and true simulations. At the Obidos–Porto (OP) location, peak flow was underestimated in the assimilated simulation, whereas low flow estimates were more accurate (Manning's coefficient = 0.0295, Table 3). Peak discharge was also underestimated for the Salekhard (SH), Solyanka (SY), Stolb (SB), Alton, ill (AL), Vicksburg, MS (VM), PS, Aswan Dam (AD), Nag Hammadi (NH), El Ekhsase (EE), KA and Kotri (KR) locations because the values used for Manning's coefficient in the true simulations (Table 3) were lower than those used in the assimilated/corrupted simulations. Peak discharge was slightly overestimated at the Rybinskaya Ges (RG), AG, and Panjnad (PJ) locations (Manning's coefficient > 0.030 in the true simulation). Nonetheless, discrepancies in peak river discharge may not affect NSEAI values if Manning's coefficient values vary significantly (e.g., at the SH, SY, AL, PS, NH, and EE locations). This is because the seasonal hydrodynamics of river flow play important roles in determining NSEAI. Therefore, overestimation and underestimation of peak discharge in the assimilated simulations highlights a major limitation of our assimilation framework, which must be addressed.

Major discrepancies were observed in NSEAI values at some locations (e.g., SP, PV, KH, SB, PS, AG, AC, KT, and SG) where corrupted simulations were coincidentally similar to true simulations. The NSEAI value at the AG location on the Niger River was extremely low (0.04) compared to other continental-scale rivers and had a Manning's coefficient of 0.0325 in the true simulation. This decrease in assimilation efficiency occurred because the corrupted river discharge was similar to the true discharge value. The NSE values for the corresponding corrupted and assimilated simulations were very similar (0.93). Therefore, improvements due to the assimilation of SWOT observations were not apparent. In addition, NSEAI values do not reflect assimilation efficiency if the NSE values in the corrupted simulations are close to 1.0 because this is the maximum value. The corrupted simulations were very similar to the true simulations and had NSE values > 0.9 at the SP, PV, KH, PS, KT, and SG locations. In contrast, there were low NSEAI

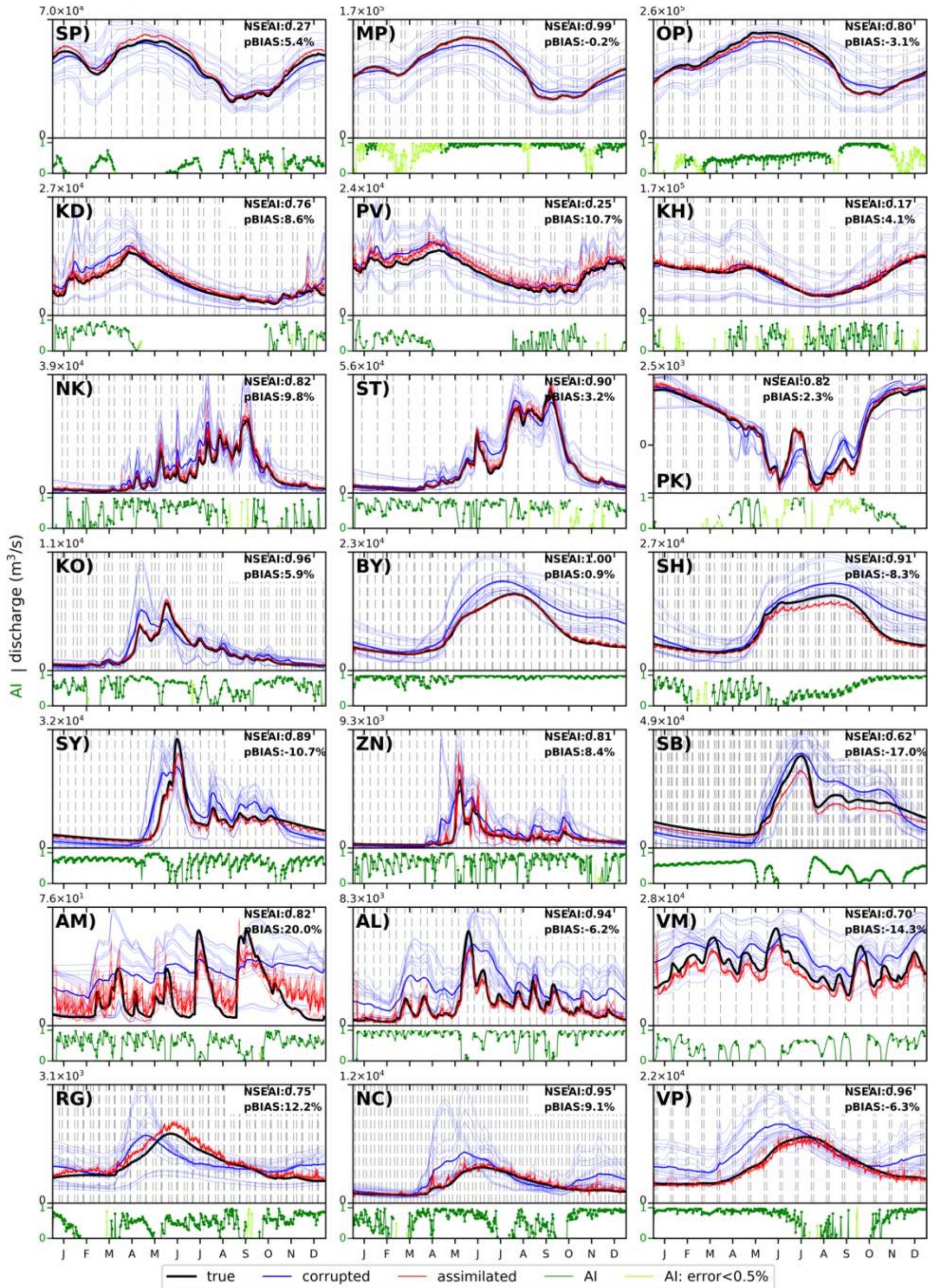

Figure 12 : Simulated discharge for the imperfect model experiment on a global scale. The locations are summarized in Table 3. River discharges of true, assimilated, and corrupted simulations are shown by black, red, and light blue lines, respectively. Dark-blue line represent the ensemble mean of corrupted simulation. The dashed-grey lines indicate the times of direct SWOT observations. The AI (green line in lower panel) is shown for days when the error between the true and corrupted discharges was >5%. Light green line indicates the AI when error was < 5%. Green dots represent the times of data assimilation. The colors and key are identical to those shown for Figure 5a-c.

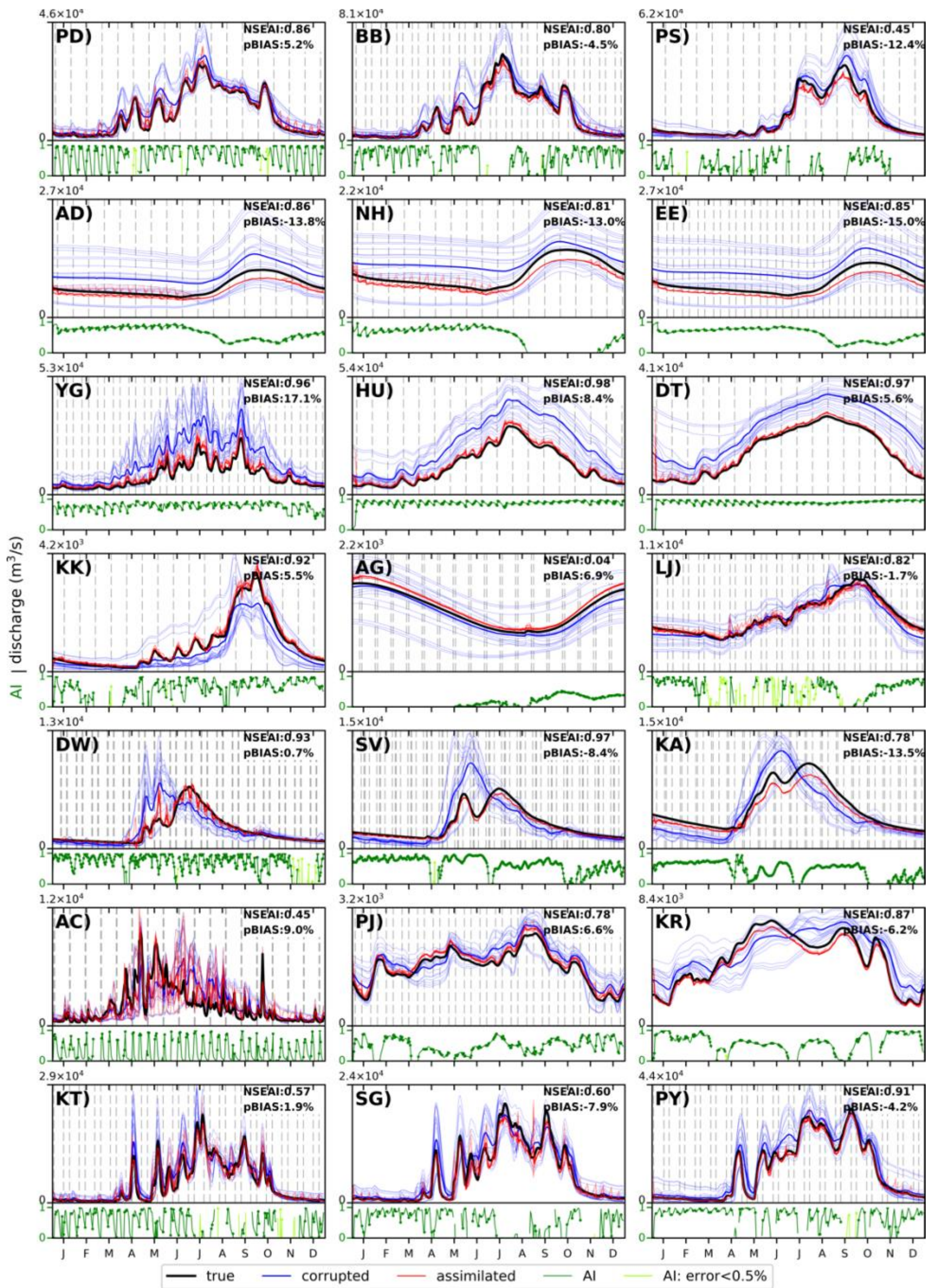

Figure 12. (continue)

values at the SB (0.62) and AC (0.45) locations. Assimilated discharge was underestimated at the SB location on the Lena River due the difference in the Manning's coefficient in true simulation (0.025). On the other hand, few observations were available for the AC location, situated in the upstream reaches of the Indus River (Manning's coefficient in true simulation is almost 0.03). However, all locations shown in Figure 12 had positive NSEAI values, indicating that assimilated discharge estimates were more similar to the true simulation than the corrupted simulation. Therefore, low assimilation scores in large rivers were due to corrupted and true discharge values being coincidentally similar.

Table 3: Summary of Figure 12, including station name; Manning's coefficient for true simulation, NSEAI, as well as pBias for assimilated and corrupted simulations.

	Station	Manning's Coefficient $[s/m^{2/3}]$ (true simulation)	NSEAI	pBais[%] (assimilated)	pBias [%] (corrupted)
SP)	Sao Paulo de Olivenca	0.0328	0.27	5.45	-1.53
MP)	Manacapuru	0.0276	0.99	-0.24	-3.23
OP)	Obidos - Porto	0.0295	0.80	-3.10	-3.78
KD)	Kindu	0.0322	0.76	8.57	10.48
PV)	Ponthierville	0.0340	0.25	10.71	10.12
KH)	Kinshasa	0.0281	0.17	4.06	2.26
NK)	Nong Khai	0.0324	0.82	9.80	18.58
ST)	Stung Treng	0.0275	0.90	3.23	8.14
PK)	Prek Kdam	0.0326	0.82	2.31	8.50
KO)	Kamen'na Obi	0.0292	0.96	5.88	10.74
BY)	Belogor'ye	0.0273	1.00	0.87	33.01
SH)	Salekhard	0.0250	0.91	-8.32	26.88
SY)	Solyanka	0.0250	0.89	-10.68	7.80
ZN)	Zmeinovo	0.0292	0.81	8.36	46.07
SB)	Stolb	0.0250	0.62	-17.00	4.23
AM)	Aitkin, MN	0.0343	0.00	19.99	90.09
AL)	Alton, ill.	0.0250	0.94	-6.21	89.38
VM)	Vicksburg, MS	0.0250	0.70	-14.29	24.55
RG)	Rybinskaya Ges	0.0334	0.75	12.24	4.99
NC)	Naberezhnyye Chelny	0.0318	0.95	9.12	36.04
VP)	Volgograd Power Plant	0.0326	0.96	-6.29	34.19
PD)	Pandu	0.0313	0.86	5.23	20.83
BB)	Bahadurabad	0.0262	0.80	-4.46	16.24
PS)	Paksey	0.0250	0.45	-12.43	10.47
AD)	Aswan dam	0.0250	0.86	-13.80	42.82
NH)	Nag Hammadi	0.0259	0.81	-12.98	31.19
EE)	El Ekhsase	0.0256	0.85	-14.97	42.32
YG)	Yichang	0.0329	0.96	17.08	85.29
HU)	Hankou	0.0296	0.98	8.35	71.90
DT)	Datong	0.0318	0.97	5.56	53.06
KK)	Koulikoro	0.0326	0.92	5.52	-22.58
AG)	Ansongo	0.0325	0.04	6.92	-6.69
LJ)	Lokoja	0.0284	0.82	-1.71	-3.38
DW)	Dawson	0.0285	0.93	0.74	4.03
SV)	Stevens Village Ak	0.0282	0.97	-8.40	0.36
KA)	Kaltag, alas.	0.0261	0.78	-13.51	-11.50
AC)	Attock	0.0299	0.45	8.95	11.81
PJ)	Panjnad	0.0323	0.78	6.60	11.15
KR)	Kotri	0.0290	0.87	-6.17	8.72
KT)	Katha	0.0268	0.57	1.91	14.74
SG)	Sagaing	0.0251	0.60	-7.95	12.39
PY)	Pyay	0.0315	0.91	-4.19	15.86

Figure 13 shows global map of NSEAI values for the imperfect model experiment. In general, the level of variation among NSEAI values in the perfect and imperfect model experiments was similar. Continental-scale rivers showed high NSEAI values (> 0.8) in downstream reaches at low latitudes and all river reaches at high latitudes. Small upstream river reaches at high latitudes showed higher NSEAI values (> 0.8) than those at low latitudes (NSEAI <0.7; e.g., upstream reaches and tributaries of the Amazon and Congo rivers). However, almost all of these rivers had positive NSEAI values. Therefore, discharge estimates in all river reaches were improved by implementing our data assimilation framework. Consequently, global analysis (Figure 12) showed that our assimilation framework improved the accuracy of river discharge

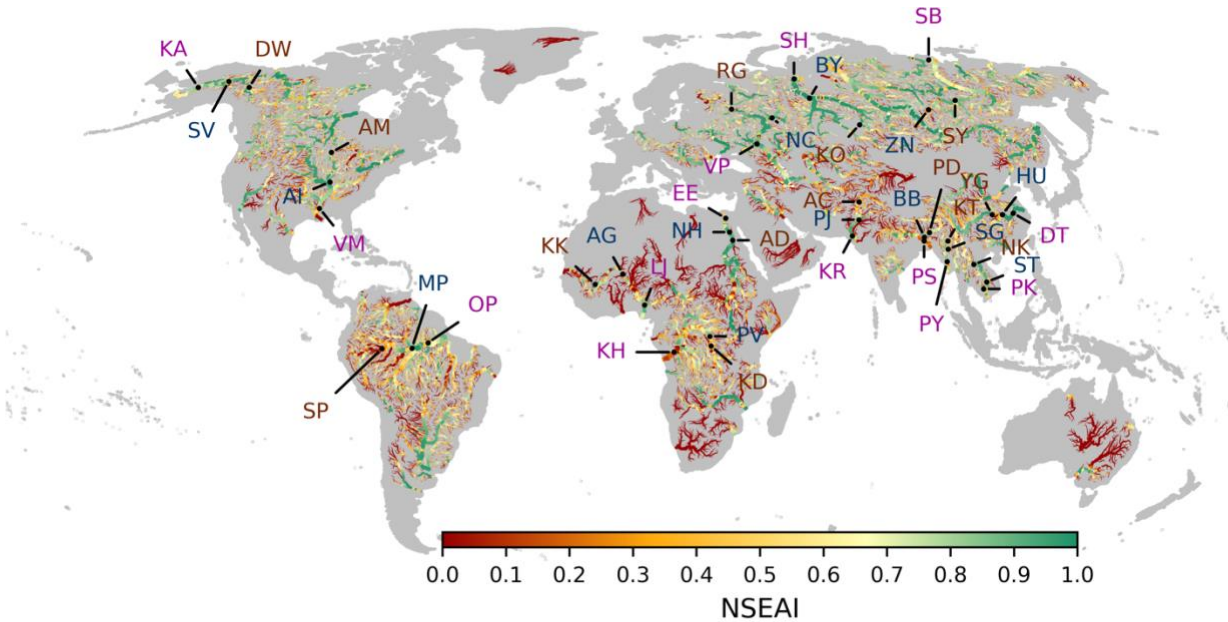

Figure 13: Global NSEAI map for the imperfect model experiment. The annotations correspond to the panels shown in Figure 12. Upstream, midstream, and downstream annotations are shown in brown, blue, and purple, respectively.

estimates in continental-scale rivers. Figure 13 shows that NSEAI values were generally lower in the imperfect model experiment than in the perfect model experiment. The NSEAI was designed to evaluate the effectiveness of data assimilation in a perfect model with unrealistic input runoff forcing. However, here we used the NSEAI to assess the assimilation efficiency of the imperfect model experiment. The midstream reaches of large rivers exhibited higher NSEAI values than downstream and upstream reaches (Table 3). A Manning's coefficient of 0.030 was used for corrupted/assimilated simulations, whereas different Manning's coefficients were used in the imperfect model experiment for true simulations, depending on river width. Discrepancies in NSEAI values corresponded to the magnitude of model error, as discussed below. However, assimilation efficiencies for the downstream reaches of large rivers exhibited greater discrepancies in the imperfect than in the perfect model experiment (e.g., the Congo, Ob, and Lena rivers; Figure 12).

Assimilation efficiencies were correlated with the magnitude of model error. Figure 14 shows a scatter plot of NSEAI values with upstream drainage areas (Figure 14a) and the total number of accumulated overpasses per SWOT cycle (Figure 14b) compared to the magnitude of model error in the imperfect model experiment. Here, we define the magnitude of model error as the absolute difference in the Manning's coefficient for the true and corrupted/assimilated simulations. In Figure 14, the magnitude of model error is separated into five categories, based on absolute Manning's coefficient errors. Mean curves for each category are shown in Figure 14a and b. River reaches with large upstream drainage areas had low NSEAI values when the model error was large, and vice versa. However, in small river reaches (upstream drainage area < 1×10^5 km^2), variations in NSEAI values were not reflected by the magnitude of model error. As in the perfect model experiment (Figure 10a), NSEAI values reached a maximum level after 1000 accumulated overpasses in each model. The fitted curves show that the assimilation score was high when the

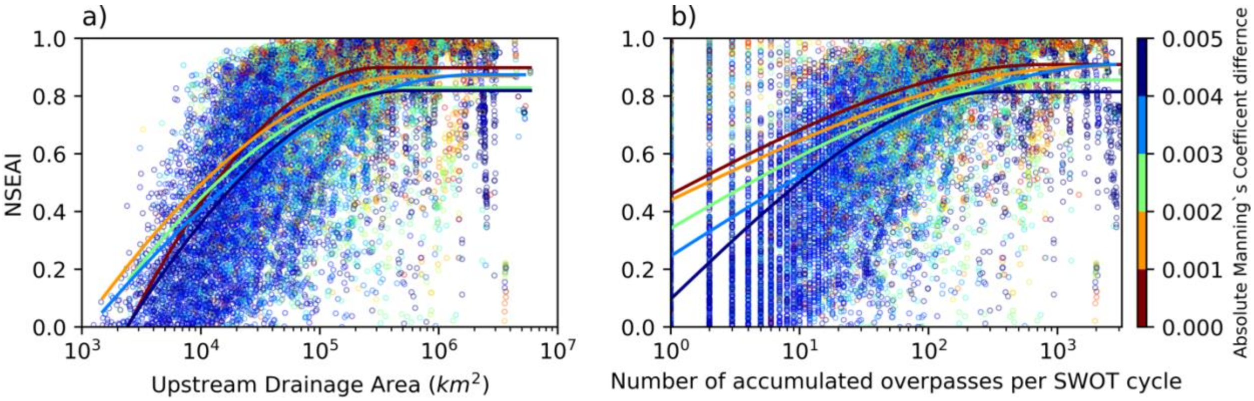

Figure 14: Scatter plot comparing NSEAI with a) upstream drainage area and b) accumulated overpasses per SWOT cycle based on the Manning's coefficient error for the imperfect model experiment.

Manning's coefficient error was low (shown in red in Figure 14), and when the model error was high the assimilation efficiency was low (shown in violet in Figure 14). However, it is difficult to discern a clear relationship from Figure 14. However, mean variation in NSEAI values is inversely related to assimilation efficiency and model error. Therefore, the magnitude of model error has an important impact on assimilation efficiency.

Figure 15 is a boxplot showing KGE values for 15 continental-scale rivers at low, mid, and high latitudes. All of these continental-scale rivers had median KGE values of > 0.6, which corresponds to good model efficiency. Large rivers at low latitudes showed good assimilation, with median KGE values > 0.8. Similarly, rivers at high latitudes also had high median KGE values (> 0.8). Rivers at mid latitudes had slightly lower median KGE values. Although the Yangtze River had the lowest median KGE value (~0.6), the main stem of the river was well-characterized by our framework (Figure S4). Small tributaries in the south of Yangtze River showed low KGE values (~0.5) which caused reduction of median KGE. River discharge estimates in most of the continental-scale rivers were better in the assimilated simulation than in the corrupted simulation (Text S6 and Figure S4). Therefore, river discharge estimates for continental-scale rivers were generally good.

4.3 Perfect model vs. imperfect model experiments

In the perfect model experiment, we assumed that model error was avoidable, whereas in the imperfect model experiment, we assumed that model error was included in Manning's coefficient. Our data assimilation framework produced good results in both experiments. In general, the perfect model experiment had higher assimilation scores (i.e., NSEAI values) than the imperfect model experiment. Figure 16 shows a boxplot of KGE values for the assimilated simulation with different numbers of accumulated overpasses per SWOT cycle for the perfect (Figure 16a) and imperfect (Figure 16b) model experiments. KGE values offer diagnostic insight into the performance of our assimilation framework in river discharge estimates. The KGE value combines correlation, relative bias, and variability to reproduce temporal dynamics while preserving flow durations. The graph in Figure 16 shows the total number of 1–600 overpasses on the horizontal axis in 20-overpass category intervals. In both experiments, the KGE values were > 0.6 in almost all the river reaches in the assimilated simulation. There were large variations in both experiments when the total number of accumulated overpasses per SWOT cycle was < 250. However, in the perfect model experiment, median KEG values were ≥ 0.95 in river reaches with

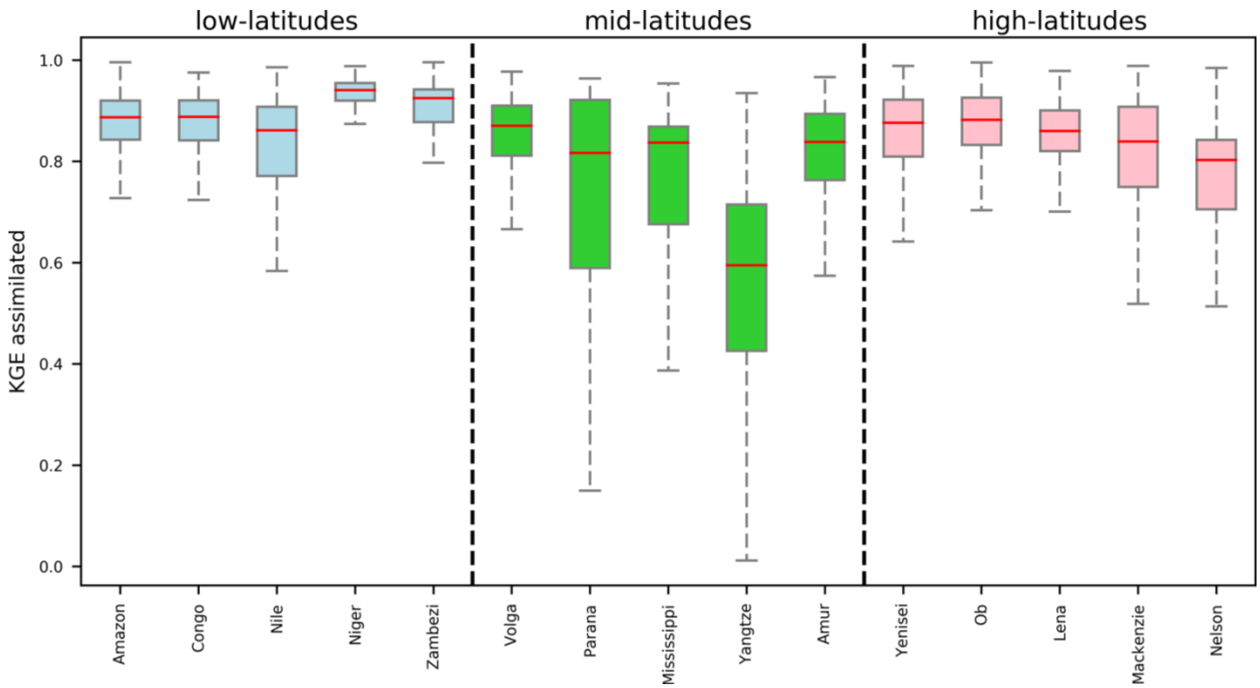

Figure 15: KGE of assimilated discharge in the imperfect model experiment for continental-scale rivers at low-latitudes (light blue), mid-latitudes (light green), and high-latitudes (pink). River pixels with a mean river discharge > 100 m^3/s were used to create the boxplot.

> 270 accumulated overpasses per SWOT cycle. The KGE values varied little in places where there were > 270 accumulated overpasses. In the imperfect model experiment, variation in KGE values was lower when there were > 270 overpasses per SWOT cycle. Therefore, minimizing model errors (i.e., in Manning's coefficient) is important for accurately assimilating SWOT observations.

5. Summary and Discussion

In this study, we developed a framework for global data assimilation using physically-based empirical localization parameters and the LETKF algorithm. We generated synthetic SWOT observations using simulated WSE measurements from the CaMa-Flood global hydrodynamic model, satellite orbit information, and expected observation errors. We evaluated the effectiveness of data assimilation on global river discharge estimates using OSSEs. The effectiveness of assimilation was evaluated using a perfect model in which the hydrodynamic model was error free and an imperfect model in which model error was included in Manning's coefficient. In the perfect model experiment, we used similar parameters for the true, corrupted, and assimilated simulations, whereas different runoff forcing was applied in the true and corrupted/assimilated simulations. In contrast, different model parameters (e.g., Manning's coefficient values) were used in the true and corrupted/assimilated simulations, and different runoff forcing was applied in the imperfect model experiment.

The perfect model experiment was performed using different runoff forcing in true and corrupted/assimilated simulations. River discharge simulations were significantly improved by data assimilation at most continental-scale river locations, particularly those at high latitudes (>50°) and in downstream river reaches at low latitudes. Discharge at upstream locations was well-

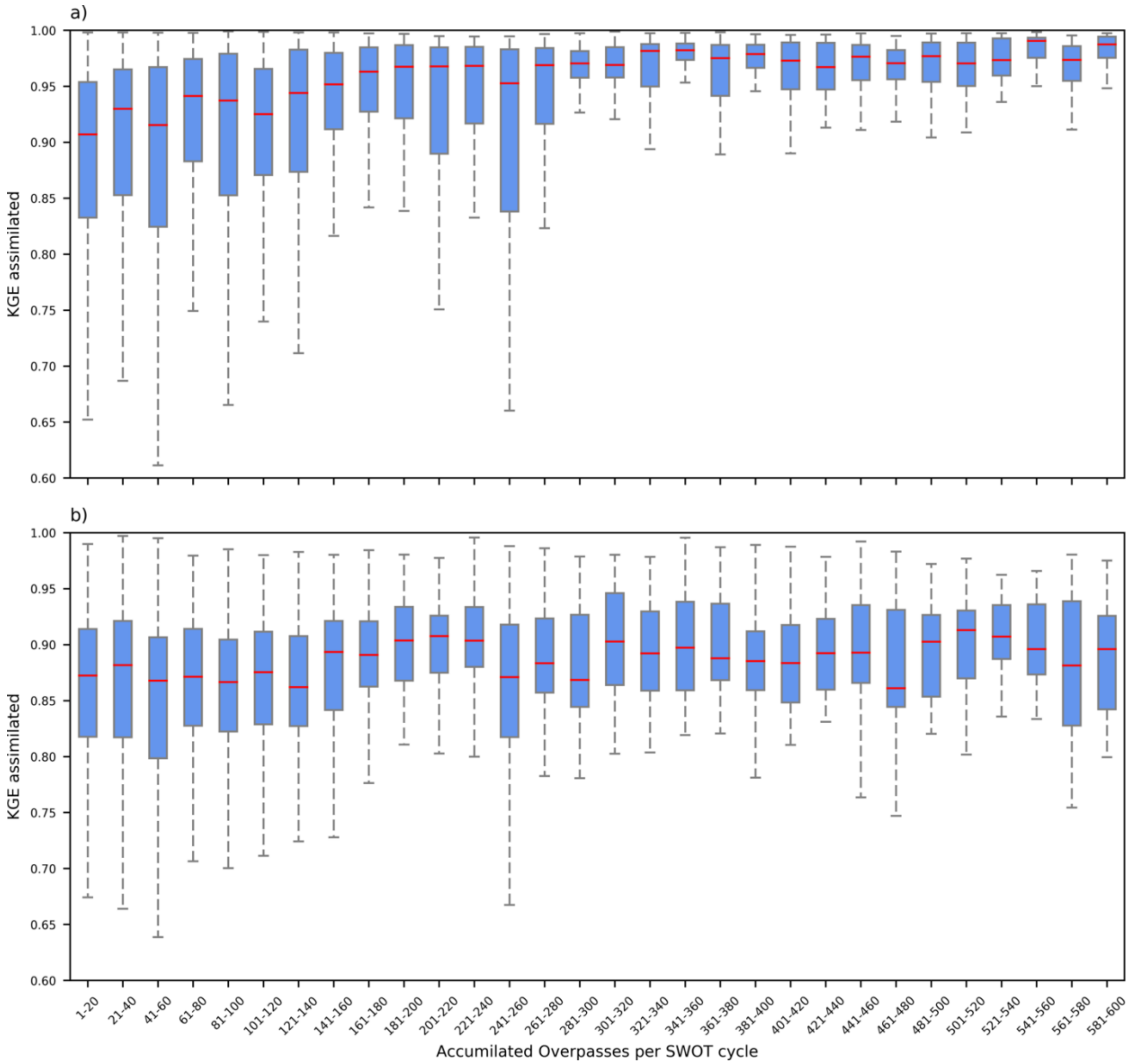

Figure 16: Boxplot of KGE values for assimilated simulation of a) perfect model and b) imperfect model experiments compared with the accumulated overpasses per SWOT cycle. There were between 1-600 accumulated overpasses. The data are presented in 20-overapass category intervals.

characterized on days on which local observations are available. However, the assimilation efficiency decreased on days on which there was no local observations. Nonetheless, at downstream locations, the assimilation efficiency was consistently high even on days in which there were no local observations. Rivers located at high latitudes also had high assimilation efficiencies in most river reaches, including upstream locations. Therefore, river discharge estimates in continental-scale rivers were improved by assimilating the SWOT observations.

The size of the empirical local patch and the number of upstream observations contributed to the low assimilation efficiency of upstream river reaches (Figure 10). Local patches in upstream river reaches were smaller than those in downstream river reaches (Revel et al., 2019), which decreased the number of assimilations per SWOT cycle. If the empirical local patch is reduced in size, the water state of an assimilated pixel may not be consistent with that of adjacent pixels, resulting in sudden changes in both WSE measurements and discharge at the target pixel. Moreover, the inflow correction from upstream is low in upstream river reaches. At upstream locations, differences between assimilated and true discharge were due to fewer local and upstream observations (Figure 5a). We used adaptive empirical local patches (Revel et al., 2019) to incorporate information from distant pixels and reduce inconsistencies in WSE measurements after assimilation. However, inconsistent WSE measurements may occur when target pixels are found in upstream river reaches. The coverage of SWOT observations may be improved by interpolating SWOT observations in upstream river reaches (Yoon et al., 2013).

The imperfect model experiment, which assumed the presence of errors, implied that river hydrodynamics, such as annual mean flow or flood peak timing, may be evaluated by assimilating SWOT observations even when realistic forcing data and parameters are unavailable. Instantaneous corrections from SWOT observations could not be directly applied to river discharge estimates because we assimilated WSE measurements only. Generally, assimilation efficiency was lower in the imperfect model experiment than in the perfect model experiment. However, global river discharge estimates were improved by incorporating SWOT observations even when model parameter errors and unrealistic runoff forcing were included. In addition, the KGE value for assimilated discharge was > 0.6 in most river reaches. As the model error is strongly associated with data assimilation effectiveness, improving our model is essential for generating more accurate river discharge estimates using SWOT observations and data assimilation. However, some of the downstream reaches of large rivers (e.g., the Congo River) showed decreased assimilation efficiencies because the corrupted discharge was coincidentally similar to the true discharge.

Further studies are needed to define the geographical parameters for hydrodynamic modeling using satellite altimetry (Durand *et al*., 2010; Yoon *et al*., 2012; Pedonetti *et al*., 2014; Revel *et al*., 2018; Emery *et al*., 2019; Breda *et al*., 2019). Before implementing a global assimilation framework, topography parameters for river hydrodynamic models must be carefully defined. Increasing the accuracy of global digital elevation models (DEMs) (O'Loughlin *et al*., 2016; Yamazaki *et al*., 2012, 2017, 2019) is essential because all river hydrodynamic models use DEMs for their baseline topography data. Realistic representations of channel cross-sections (i.e., width and bathymetric depth) are also needed because these are used to determine flow conveyance capacity. Recently, global-scale river-width datasets have been developed using high-resolution satellite images (Allen & Pavelsky, 2015; Yamazaki et al., 2014). Methods to estimate channel bathymetric depth using SWOT observations have also been proposed (Brêda *et al*., 2019; Durand *et al*., 2008; Revel *et al*., 2018; Yoon *et al*., 2012). Estimating other hydrodynamic parameters, such as Manning's coefficient, is also essential and may be achieved using SWOT observations (Emery et al., 2019; Pedinotti et al., 2014). River hydrodynamic models will be improved by accurately defining topography parameters. In addition, we must be able to quantify the error associated with data assimilation frameworks that is caused by topographic uncertainties.

In this study, we were able to assimilate SWOT observations into a global-scale river model. Furthermore, we demonstrated that this model could be used to estimate river discharge in large river basins. However, to apply this data assimilation framework to real SWOT observations in the near future, further studies on the model physics and ancillary topography data development are needed. In addition, further studies are required to determine how to apply our data assimilation framework to real SWOT observations. The following items must be addressed to ensure our method can be applied to real SWOT observations on a global scale:1. A method is needed to transform fine (< 100 m) SWOT observations into a form that can be applied to a coarse CaMa-Flood model grid (~25 km) (discussed in section 3.1), 2. An interpolation method is needed to smooth WSE measurements at assimilated locations and ensure these are consistent with adjacent pixels (discussed in section 4.1.1), 3. The accuracy of runoff estimates must be improved (discussed in section 4.1.2), 4. The accuracy of river models must be improved by precisely defining geographical parameters such as Manning's coefficient (discussed in section 4.2).

6. Acknowledgments

Data available from the authors upon request. This research was supported by JSPS Grant-in-Aid for Scientific Research (16H06291, 18H01540, and 20H02251).

7. References

Allen, G. H., & Pavelsky, T. M. (2015). Patterns of river width and surface area revealed by the satellite-derived North American River Width data set. *Geophysical Research Letters*, *42*(2), 395–402. https://doi.org/10.1002/2014GL062764

Alsdorf, D., Dunne, T., Melack, J., Smith, L., & Hess, L. (2005). Diffusion modeling of recessional flow on central Amazonian floodplains. *Geophysical Research Letters*, *32*(21), 1–4. https://doi.org/10.1029/2005GL024412

Andreadis, K. M., Clark, E. A., Lettenmaier, D. P., & Alsdorf, D. E. (2007). Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. *Geophysical Research Letters*, *34*(10), 1–5. https://doi.org/10.1029/2007GL029721

Bates, P. D., Horritt, M. S., & Fewtrell, T. J. (2010). A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology*, *387*(1–2), 33–45. https://doi.org/10.1016/j.jhydrol.2010.03.027

Biancamaria, S., Durand, M., Andreadis, K. M., Bates, P. D., Boone, A., Mognard, N. M., et al. (2011). Assimilation of virtual wide swath altimetry to improve Arctic river modeling. *Remote Sensing of Environment*, *115*(2), 373–381. https://doi.org/10.1016/j.rse.2010.09.008

Biancamaria, Sylvain, Lettenmaier, D. P. ., & Pavelsky, T. M. . (2016). The SWOT Mission and Its Capabilities for Land Hydrology. *Surveys in Geophysics*, *37*(2), 307–337. https://doi.org/10.1007/s10712-015-9346-y

Bonnema, M. G., Sikder, S., Hossain, F., Durand, M., Gleason, C. J., & Bjerklie, D. M. (2016). Benchmarking wide swath altimetry-based river discharge estimation algorithms for the Ganges river system. *Water Resources Research*, *52*(4), 2439–2461. https://doi.org/10.1002/2015WR017296

Brêda, J. P. L. F., Paiva, R. C. D., Bravo, J. M., Passaia, O. A., & Moreira, D. M. (2019). Assimilation of Satellite Altimetry Data for Effective River Bathymetry. *Water Resources Research*, *55*(9), 7441–7463. https://doi.org/10.1029/2018WR024010

CNES. (2015). SWOT orbit. https://doi.org/https://www.aviso.altimetry.fr/en/missions/future-missions/swot/orbit.html

Desai, S., Rodriguez, E., Fernandez, D. E., Peral, E., Chen, C. W., Bleser, J.-W. De, & Williams, B. (2018). *Surface Water and Ocean Topography Mission (SWOT) Science Requirements Document. Rev B*. California. Retrieved from https://swot.jpl.nasa.gov/docs/D-61923_SRD_Rev_B_20181113.pdf

Durand, M., Andreadis, K. M., Alsdorf, D. E., Lettenmaier, D. P., Moller, D., & Wilson, M. (2008). Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model. *Geophysical Research Letters*, *35*(20), 1–5. https://doi.org/10.1029/2008GL034150

Durand, M., Rodríguez, E., Alsdorf, D. E., & Trigg, M. (2010). Estimating river depth from remote sensing swath interferometry measurements of river height, slope, and width. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *3*(1), 20–31. https://doi.org/10.1109/JSTARS.2009.2033453

Durand, M., Fu, L. L., Lettenmaier, D. P., Alsdorf, D. E., Rodriguez, E., & Esteban-Fernandez, D. (2010). The surface water and ocean topography mission: Observing terrestrial surface water and oceanic submesoscale eddies. *Proceedings of the IEEE*, *98*(5), 766–779. https://doi.org/10.1109/JPROC.2010.2043031

Durand, M., Gleason, C. J., Garambois, P. A., Bjerklie, D., Smith, L. C., Roux, H., et al. (2016). An intercomparison of remote sensing river discharge estimation algorithms from measurements of river height, width, and slope. *Water Resources Research*, *52*(6), 4527–4549. https://doi.org/10.1002/2015WR018434

Dutra, E., Gianpaolo, B., Jean-Christophe, C., Munier, S., Burke, S., Fink, G., et al. (2017). *Report on the improved water resources reanalysis Deliverable*. Retrieved from http://earth2observe.eu/files/Public Deliverables/D5.2 - Report on the Improved Water Resources Reanalysis (WRR tier 2).pdf

Emery, C., Biancamaria, S., Boone, A., Ricci, S., Rochoux, M., Pedinotti, V., & David, C. (2019). Assimilation of wide-swath altimetry observations to correct large-scale river routing model parameters. *Hydrology and Earth System Sciences Discussions*, (June), 1–40. https://doi.org/https://doi.org/10.5194/hess-2019-242

Esteban-Fernandez, D. (2017). *SWOT mission performance and error budget*. *JPL Publication* (Vol. 2018-April). Retrieved from https://pdms.jpl.nasa.gov/

Evensen, G. (2003). The Ensemble Kalman Filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, *53*(4), 343–367. https://doi.org/10.1007/s10236-003-0036-9

Evensen, G. (2009). The ensemble Kalman filter for combined state and parameter estimation. *IEEE Control Systems*, *29*(3), 83–104. https://doi.org/10.1109/MCS.2009.932223

Fjørtoft, R., Gaudin, J. M., Pourthié, N., Lalaurie, J. C., Mallet, A., Nouvel, J. F., et al. (2014). KaRIn on SWOT: Characteristics of near-nadir Ka-band interferometric SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(4), 2172–2185. https://doi.org/10.1109/TGRS.2013.2258402

Garambois, P. A., & Monnier, J. (2015). Inference of effective river properties from remotely sensed observations of water surface. *Advances in Water Resources*, *79*, 103–120. https://doi.org/10.1016/j.advwatres.2015.02.007

Gleason, C. J., & Smith, L. C. (2014). Toward global mapping of river discharge using satellite images and at-many-stations hydraulic geometry. *Proceedings of the National Academy of Sciences*, *111*(13), 4788–4791. https://doi.org/10.1073/pnas.1317606111

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Hersbach, H., Peubey, C., Simmons, A., Berrisford, P., Poli, P., & Dee, D. (2015). ERA-20CM: A twentieth-century atmospheric model ensemble. *Quarterly Journal of the Royal Meteorological Society*, *141*(691), 2350–2375. https://doi.org/10.1002/qj.2528

Hock, R. (1998). *Modelling of glacier melt and discharge*. ETH Zurich. https://doi.org/https://www.research-collection.ethz.ch/handle/20.500.11850/143428

Hock, R., & Holmgren, B. (2005). A distributed surface energy-balance model for complex topography and its application to Storglaciären, Sweden. *Journal of Glaciology*, *51*(172), 25–36. https://doi.org/10.3189/172756505781829566

Hunt, B. R., Kostelich, E. J., & Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, *230*(1–2), 112–126. https://doi.org/10.1016/j.physd.2006.11.008

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, *82*(1), 35. https://doi.org/10.1115/1.3662552

Kling, H., Fuchs, M., & Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, *424–425*, 264–277. https://doi.org/10.1016/j.jhydrol.2012.01.011

Lee, H., Durand, M., Jung, H. C., Alsdorf, D., Shum, C. K., & Sheng, Y. (2010). Characterization of surface water storage changes in Arctic lakes using simulated SWOT measurements. *International Journal of Remote Sensing*, *31*(September 2017), 14–3931. https://doi.org/10.1080/01431161.2010.483494

Marcus, W. A., & Fonstad, M. A. (2010). Remote sensing of rivers: The emergence of a subdiscipline in the river sciences. *Earth Surface Processes and Landforms*, *35*(15), 1867–1872. https://doi.org/10.1002/esp.2094

Miyoshi, T., Yamane, S., & Enomoto, T. (2007). Localizing the Error Covariance by Physical Distances within a Local Ensemble Transform Kalman Filter (LETKF). *Sola*, *3*(1), 89–92. https://doi.org/10.2151/sola.2007-023

Munier, S., Polebistki, A., Brown, C., Belaud, G., & Lettenmaier, D. P. (2015). SWOT data assimilation for operational reservoirmanagement on the upper Niger River Basin S. *Water Resources Research*, *51*(1), 554–575. https://doi.org/https://doi.org/10.1002/2014WR016157

Nash, J. E., & Sutcliffe, J. V. (1970). River Flow Forecasting Through Conceptual Models Part I-a Discussion of Principles*. *Journal of Hydrology*, *10*, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

O'Loughlin, F. E., Paiva, R. C. D., Durand, M., Alsdorf, D. E., & Bates, P. D. (2016). A multi-sensor approach towards a global vegetation corrected SRTM DEM product. *Remote*

Sensing of Environment, *182*, 49–59. https://doi.org/10.1016/j.rse.2016.04.018

Oki, T., & Kanae, S. (2006). Global Hydrological Cycles and World Water Resources. *Science*, *5790*(313), 1068–1072. https://doi.org/10.1126/science.1128845

Pavelsky, T. M., Durand, M. T., Andreadis, K. M., Beighley, R. E., Paiva, R. C. D., Allen, G. H., & Miller, Z. F. (2014). Assessing the potential global extent of SWOT river discharge observations. *Journal of Hydrology*, *519*(PB), 1516–1525. https://doi.org/10.1016/j.jhydrol.2014.08.044

Pedinotti, V., Boone, A., Ricci, S., Biancamaria, S., & Mognard, N. (2014). Assimilation of satellite data to optimize large-scale hydrological model parameters: a case study for the SWOT mission. *Hydrology and Earth System Sciences*, *18*(11), 4485–4507. https://doi.org/10.5194/hess-18-4485-2014

Revel, Ikeshima, Yamazaki, & Kanae. (2019). A Physically Based Empirical Localization Method for Assimilating Synthetic SWOT Observations of a Continental-Scale River: A Case Study in the Congo Basin. *Water*, *11*(4), 829. https://doi.org/10.3390/w11040829

Revel, M., Yamazaki, D., & Kanae, S. (2018). Estimating Global River Bathymetry by Assimilating Synthetic SWOT Measurements. *Journal of Japan Society of Civil Engineers, Ser. B1 (Hydraulic Engineering)*, *74*(4), I_307-I_312. https://doi.org/10.2208/jscejhe.74.I_307

Yamazaki, D., Kanae, S., Kim, H., & Oki, T. (2011). A physically based description of floodplain inundation dynamics in a global river routing model. *Water Resources Research*, *47*(4), 1–21. https://doi.org/10.1029/2010WR009726

Yamazaki, D., Baugh, C. A., Bates, P. D., Kanae, S., Alsdorf, D. E., & Oki, T. (2012). Adjustment of a spaceborne DEM for use in floodplain hydrodynamic modeling. *Journal of Hydrology*, *436–437*, 81–91. https://doi.org/10.1016/j.jhydrol.2012.02.045

Yamazaki, D., Lee, H., Alsdorf, D. E., Dutra, E., Kim, H., Kanae, S., & Oki, T. (2012). Analysis of the water level dynamics simulated by a global river model: A case study in the Amazon River. *Water Resources Research*, *48*(9), 1–15. https://doi.org/10.1029/2012WR011869

Yamazaki, D., De Almeida, G. A. M., & Bates, P. D. (2013). Improving computational efficiency in global river models by implementing the local inertial flow equation and a vector-based river network map. *Water Resources Research*, *49*(11), 7221–7235. https://doi.org/10.1002/wrcr.20552

Yamazaki, D., O'Loughlin, F., Trigg, M. A., Miller, Z. F., Pavelsky, T. M., & Bates, P. D. (2014). Development of the Global Width Database for Large Rivers. *Water Resources Research*, *50*(4), 3467–3480. https://doi.org/10.1002/2013WR014664

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., et al. (2017). A high-accuracy map of global terrain elevations. *Geophysical Research Letters*, *44*(11), 5844–5853. https://doi.org/10.1002/2017GL072874

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G., & Pavelsky, T. (2019). MERIT Hydro: A high-resolution global hydrography map based on latest topography datasets. *Water Resources Research*, 2019WR024873. https://doi.org/10.1029/2019WR024873

Yoon, Y., Durand, M., Merry, C. J., Clark, E. A., Andreadis, K. M., & Alsdorf, D. E. (2012). Estimating river bathymetry from data assimilation of synthetic SWOT measurements.

Journal of Hydrology, *464–465*, 363–375. https://doi.org/10.1016/j.jhydrol.2012.07.028

Yoon, Y., Durand, M., Merry, C. J., & Rodriguez, E. (2013). Improving temporal coverage of the SWOT mission using spatiotemporal kriging. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *6*(3), 1719–1729. https://doi.org/10.1109/JSTARS.2013.2257697

Zhao, Q., Ye, B., Ding, Y., Zhang, S., Yi, S., Wang, J., et al. (2013). Coupling a glacier melt model to the Variable Infiltration Capacity (VIC) model for hydrological modeling in north-western China. *Environmental Earth Sciences*, *68*(1), 87–101. https://doi.org/10.1007/s12665-012-1718-8

The English in this document has been checked by at least two professional editors, both native speakers of English. For a certificate, please see:

http://www.textcheck.com/certificate/248SrS

AGU PUBLICATIONS

Water Resources Research

Supporting Information for

A framework for global-scale river discharge estimation by assimilating satellite altimetry

Menaka Revel[1], Daiki Ikeshima[2], Dai Yamazaki[1], and Shinjiro Kanae[2]

[1]Institute of Industrial Sciences, The University of Tokyo
[2] Department of Civil and Environmental Engineering, Tokyo Institute of Technology

Contents of this file

Text S1 to S6
Figures S1 to S7
Tables S1

Introduction

In this supporting information, we provide the supplemental text, figures and table for the development of the framework for estimating river discharge by assimilating satellite altimetry. Here we introduce the generating input runoff, data assimilating procedure, empirical localization parameters, estimating SWOT observation error, upstream inflow correction, evaluation of river discharge accuracy using KGE, map of accumulated overpasses per SWOT cycle, variation of NSEAI with accumulated overpasses, and map of assimilation frequency.

Text S1. Generating input runoff forcing

We used HTESSEL (ECMWF) model runoff out form E2O WRR2 (Dutra et al., 2017) for true simulation. The original runoff was used for simulations in true simulation. The remaining runoff outs were perturbed by multiplying a random number to generate 18 ensembles. Table S1 presets the random values used for each LSM/GHM runoff output for generation ensembles for perfect model experiment and imperfect model experiment.

Table S1: Generation of input runoff forcing for perfect model experiment

Simulation	LSM/GHM	Random Value (Perfect Model experiment	Random Value (Imperfect Model experiment
True	HTESSEL	Original Runoff is used	Original Runoff is used
Corrupted/Assimilated	PCR-GLOBWB	0.93	0.95
		0.97	1.00
		1.04	1.08
	JULES	0.79	0.91
		0.96	0.96
		0.98	1.01
	LISFLOOD	0.88	0.91
		0.94	0.99
		1.08	1.02
	ORCHIDEE	0.81	0.93
		1.03	1.01
		1.12	1.06
	WaterGAP3	0.93	0.85
		0.97	0.97
		1.00	1.18
	W3	0.91	0.85
		0.99	1.01
		1.08	1.05

Text S2: Data assimilation procedure

The LETKF (Hunt et al., 2007; Ott et al., 2004) algorithm was used in this study to efficiently perform data assimilation in global scale. Here, we used the SWOT-observed water surface elevation as the 'observed variable' of the data assimilation procedure. The model forecasts were propagated using CaMa-Flood hydrodynamic model. Then the assimilated water state was diagnosed using LETKF algorithm and update the initial conditions for next days` simulation. The water state of the proceeding step (i.e., initial water storage) was computed with data assimilation using LETKF with Equation (A1):

$$x^a = \bar{x}^f + E^f \left[\tilde{P}^a (HE^f)^T \left({}^{R}/_{w} \right)^{-1} (y^o - H\bar{x}^f) + \sqrt{m-1} \left(\tilde{P}^a \right)^{\frac{1}{2}} \right] \quad \text{(S1)}$$

where x^a is the assimilated WSE; x^f is the forecasted WSE of each parallel CaMa-Flood ensemble, $\bar{x}^f$ is the mean forecasted value of ensemble members; E^f is the model forecast error covariance matrix, which consist of perturbations which calculated using;

$$E^f = x^f - \bar{x}^f \tag{S2}$$

$\tilde{P}^a$, and $\left(\tilde{P}^a\right)^{\frac{1}{2}}$ were calculated in Equations (S3) and (S4), respectively:

$$\tilde{P}^a = VD^{-1}V^T \tag{S3}$$

$$\left(\tilde{P}^a\right)^{\frac{1}{2}} = VD^{-\frac{1}{2}}V^T \tag{S4}$$

where,

$$VDV^T = (m-1)\, {}^{I}\!/_{\Delta} + (HE^f)^T \left({}^{R}\!/_{w}\right)^{-1} HE^f \tag{S5}$$

where m is the number of ensemble members (= 20), I is an identity matrix, Δ is the covariance inflation parameter (estimated adaptively using innovative statistics following Miyoshi, (2011), with background variance of 0.04^2), H is the observation operator which is linearly related to the observation and the state, R is the observation error covariance matrix, which is a diagonal matrix having observation error variances in the diagonal (explained in Appendix C); and w is the observation localization weightage (explained in Appendix B). LETKF applied using the equation (A1) to a 'empirical local patch' (explained in Appendix B), which is a small domain around each observation point where the observation has correlations with model state variables. The state variables are independently updated within each empirical local patch.

Text S3: Empirical Localization Parameters

Empirical localization parameters were derived using the spatial auto-correlation of simulated WSEs adaptively. We developed physically-based local patch using CaMa-Flood modelled WSE using runoff simulated by minimal advanced treatments of surface interaction and runoff (MATSIRO: Takata et al., 2003) LSM forced by S14 (Iizumi et al., 2017). The empirical local patches where derived by defining a threshold to the spatial dependency weights calculated by conducting semi-variogram analysis on transformed WSE data. Transformation of WSEs involved three steps: (1) removing trends, (2) removing seasonality, and (3) standardizing. Then, we derived the observation localization weights using Gaussian function using localization lengths corresponds to the threshold defined to the spatial dependency weight to bound the empirical local patch. For further information on deriving physically based adaptive empirical localization parameters, please refer to Revel et al., (2019, 2018b).

Text S4: The SWOT observation error

The SWOT mission sets a goal of 10cm accuracy for water area >= 1km^{2}at the WSE measurement. However, the actual accuracy of future distributing observation data is unclear since it varies with river width, river length, surrounding topography (Durand et al., 2010) or even distance from the satellite track (varies between 4~10cm) (Desai et al., 2018). In this study, we modelled observation error to be normally distributed with zero mean and variance of σ_h.

$$\sigma_h = \begin{cases} \frac{1}{WL}0.10 & ,WL \geq 1.0\ km^2 \\ \frac{1}{WL}0.25 & ,1.0\ km^2 > WL \geq 0.625\ km^2 \\ 0.25 & ,WL < 0.625\ \ km^2 \end{cases} \quad \text{(S6)}$$

where W and L are river width and river length, respectively. We adopt L to be 1km as we assume only the observations near the outlet of the unit-catchment can be used for data assimilation because CaMa-Flood unit-catchments show internal variability in WSE, especially in steep upstream reaches. We used $\sigma_h$as the diagonal components in the observation error covariance matrix in LETKF.

Figure S1 presents the global map of observation error variance calculated using the equation (S6). Most of the upstream reaches where $W < 625\ m$ are having observation error covariance of 0.25m. Downstream of large rivers such as Amazons, Congo, Ob, Lena, etc. show smaller variances below 0.10m. Therefore, the observation error variance demonstrates a spatial variability.

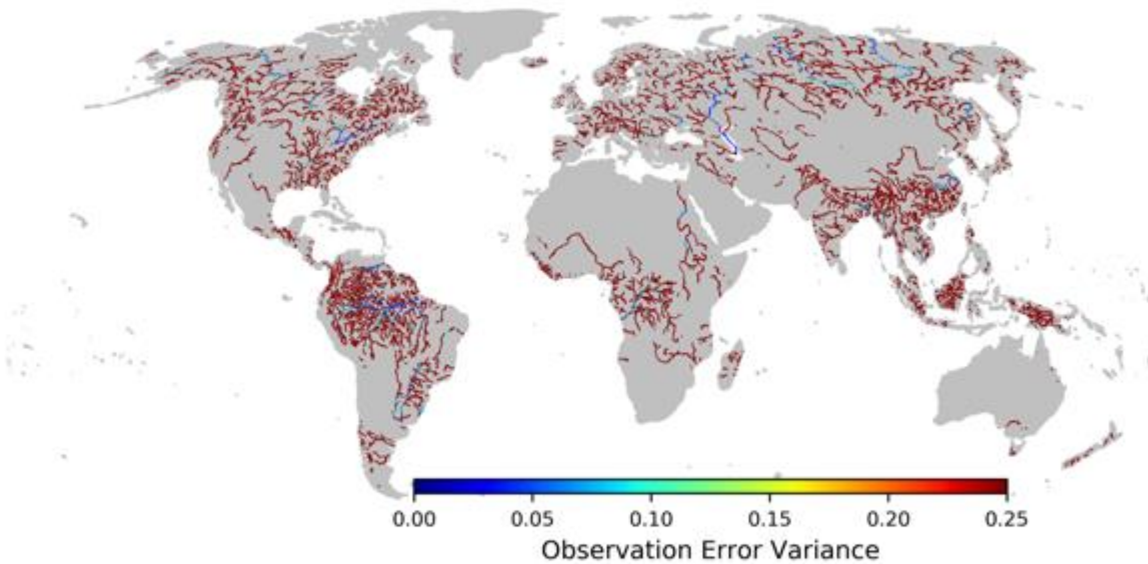

Figure S1. Observation error variance calculated using the equation (S6). Pixels with mean discharge > 100m3/s were used for visualization purposes.

Text S5: Importance of upstream inflow correction

To evaluate the necessity of upstream inflow correction, we performed a 'partially observed experiment,' simulating a situation where only part of the observation is available. In this experiment the inflow correction from the upstream was not corrected at all. Here, we used the settings similar to perfect model experiment and set the whole Amazon River basin as an experimental target but disabled the observations westward (upstream) of the midstream location Y (Figure S2b). Therefore, the location Y received local state correction due to its local SWOT observation, but the inflow from upstream was not corrected (Figure S2b). Aside from the observation area, the data assimilation process was identical to that of the perfect model experiment for the whole Amazon River basin. In the partially observed experiment, the improvement by data assimilation at the midstream location Y was mostly lost. In the partially observed experiment, the assimilated discharge was similar to the corrupted discharge at the location X (Figure S2a). The assimilated discharge was slightly improved at the location Y due to the local assimilations (Figure S2b). The AI was very low, reaching only ~0.5 even on days with local observations. On the other hand, the assimilation in the location Z was very successful, because that location is situated sufficiently far away from the unobserved area and the local patch is large enough to receive local observations every day. The spatial distribution of the NSEAI showed that decrease in assimilation efficiency up river reaches around 100km downstream of location Y. But the far downstream reaches were not affected by the unavailability of the observations in upstream. This suggests the propagation of corrected discharges from upstream pixels (i.e., upstream inflow correction) is important. Data assimilation should be applied to the entire upstream region to achieve reasonable estimations of discharge in continental-scale rivers with large drainage areas.

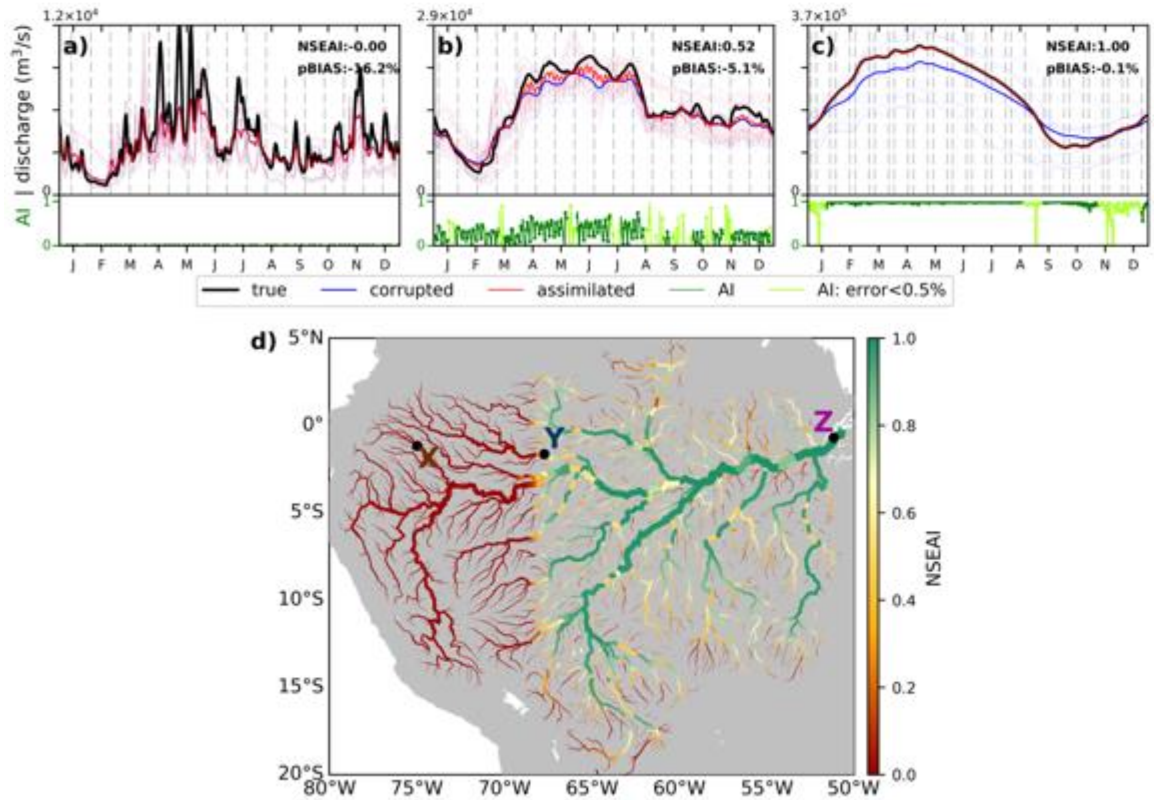

Figure S2. a)-c) Hydrograph for locations X, Y, and Z as in section 4.1.1; and d) NSEAI map for partial observed experiment. Rules are similar to Figure 5**.**

Text S6. Global River Discharge Estimation Accuracy

a) Perfect model Experiment

To further evaluate the assimilation effectiveness, we compare the KGE metric of assimilated and corrupted simulations at global scale (Figure S3). KGE offers diagnostic insights into the performance of our assimilation framework on estimating river discharge. KGE is a combination of correlation, relative bias, and variability which presets the ability to reproduce of temporal dynamics with preserving flow durations. KGE of assimilated simulation (Figure S3a) results are similar to the global map of NSEAI (Figure 7m), large rivers in low latitudes (i.e. Amazons, Congo, Nile, Mekong, Niger, Mississippi) and rivers in higher latitudes show higher KGE values (>0.8). But the relatively small rivers in south-east Asia, Europe and East Coast of Northern America shows slightly lower KGE values (0.8~0.6). On the other hand, KGE values of the corrupted simulation has values around 0.4~0.6 due to the difference of the runoff forcing from the true simulation. Almost all the global rivers are having >0.1 difference in KGE between assimilated and corrupted simulation. The KGE difference is also similar to NSEAI (Figure 7m) large low latitude rivers and higher latitude rivers demonstrate large difference (≥0.4) and smaller rivers in low latitudes shows lower KGE difference (≈0.1). The high KGE difference here means that the data assimilation can benefit the hydrodynamic model, under the assumption that core hydrodynamic model has the correct water physics and river routing system. Therefore, assimilated river discharge shows better simulation efficiency than the non-assimilated simulation (corrupted simulation) in most of the global rivers.

b) Imperfect model Experiment

The KGE-statistic was calculated for evaluating the results of imperfect model experiment and illustrates the insights into the performance of our assimilation framework. Figure S4 shows the global extent of the KGE at imperfect model experiment; KGE of river discharge at assimilated simulation (Figure S4a), corrupted simulation (Figure S4b), and the difference between two simulations (Figure S4c) are presented. Similar to the result of NSEAI (Figure 13), the KGE in the assimilated simulation was large at midstream and downstream locations of the large-scale rivers. Although the upstream locations had a small KGE value, the value rises in the downstream and it becomes almost 1.0 in the downstream. Furthermore, KGE of assimilated simulation are higher than that of corrupted simulation in almost all the global rivers (Figure S4c), difference of KGE was positive at most locations. This denotes that data assimilation of SWOT observation has the potential to correct the simulation even when the model has erroneous geographical parameters (i.e., Manning's Coefficient) and inaccurate runoff forcing. However, an important characteristic of this result must be pointed out: Some locations near river mouth of large rivers had a high KGE value even in the corrupted simulation (Figure S4b). The high KGE values at downstream of the large rivers are due to the coincidence of the true and corrupted discharge well agrees with each other (true and corrupted) (i.e., Congo). In addition, the KGE evaluates the prediction power of model, by focusing on seasonal variation in terms of correlation, relative bias, and variability. As a consequence, KGE was able to remain high at downstream reaches of larges rivers. Those locations tend to have similar seasonal trend (i.e. high-water season happened in the same time) between true and assimilated/corrupted simulation, or have a long period when seasonal trend is almost the same (i.e. discharge at winter season was almost same). Therefore, the high KGE here only means how the data assimilation can benefit the model, under the assumption that core hydrodynamic model error is included in Manning's Coefficient. Hence, the data assimilation is very effective to improve global river discharge under such assumptions. To make data assimilation effective under the real operation of SWOT satellite, hydrodynamic model uncertainties need be decreased.

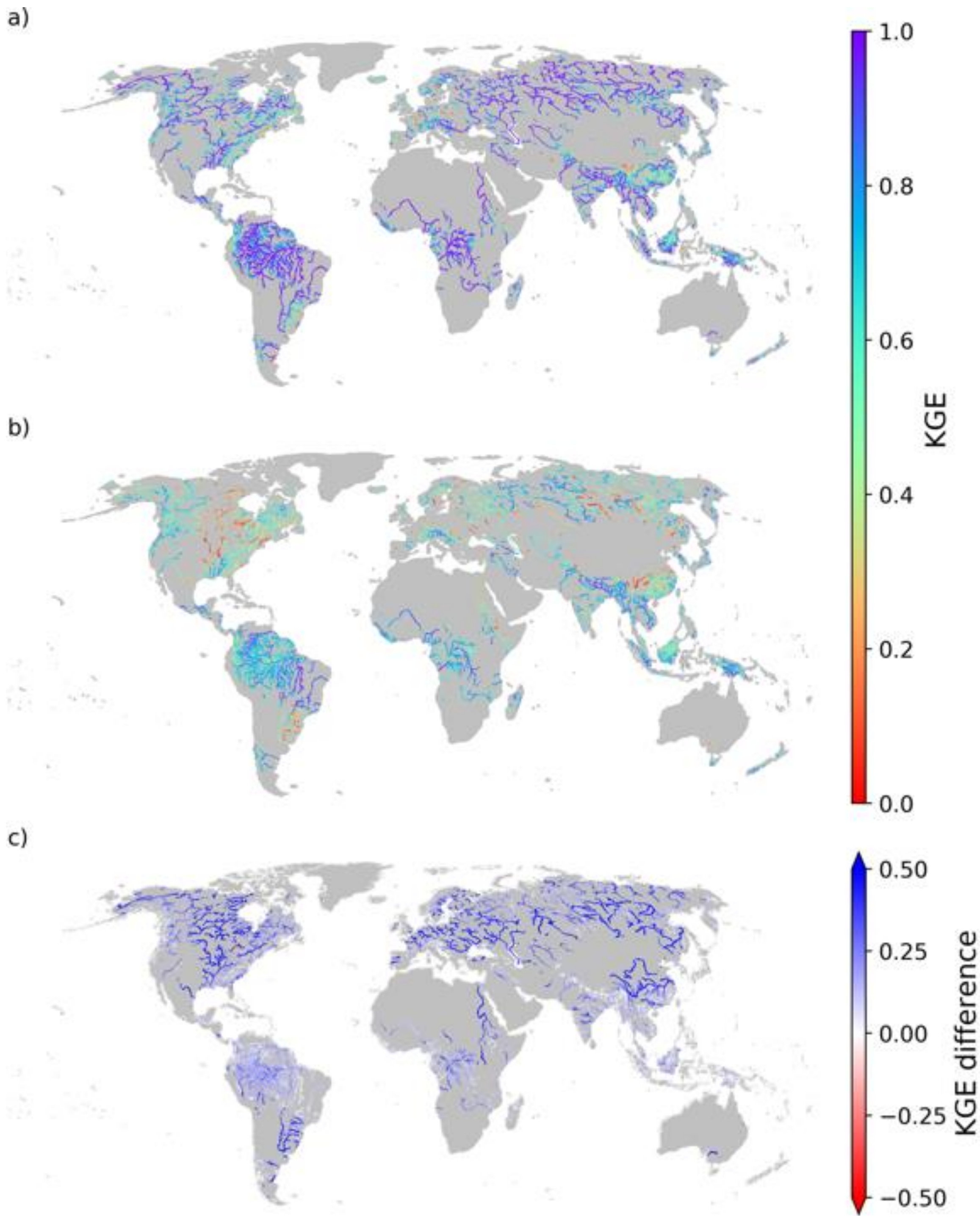

Figure S3. KGE coefficient coefficients of river discharge of a) assimilated and b) corrupted simulations. c) Difference in KGEs for assimilated and corrupted simulations for perfect model experiment. Pixels with mean discharge > 100 m3/s were shown for visualization purposes.

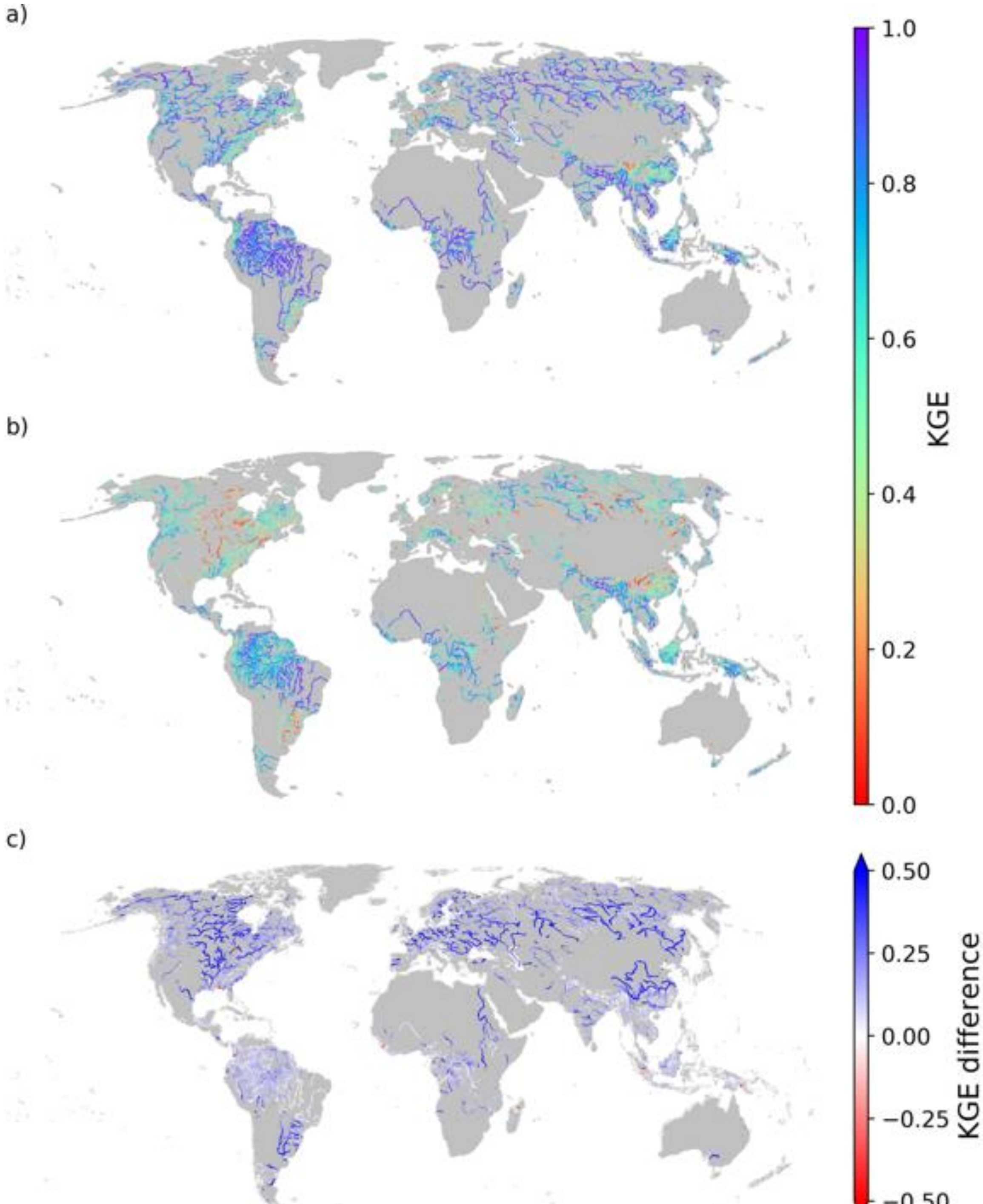

Figure S4. KGE coefficient coefficients of river discharge of a) assimilated and b) corrupted simulations. c) Difference in KGEs for assimilated and corrupted simulations for imperfect model experiment. Pixels with mean discharge > 100 m3/s were shown for visual purposes.

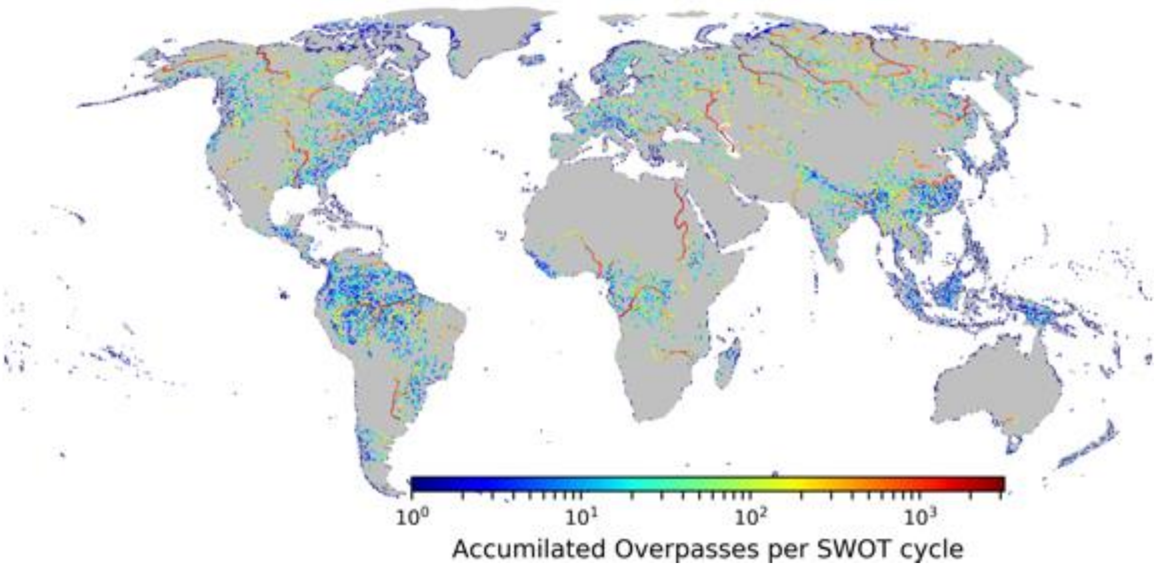

Figure S5: Accumulated Overpasses per SWOT cycle. Pixels with mean discharge > 100 m3/s were shown for visual purposes. Color bar is presented in log scale.

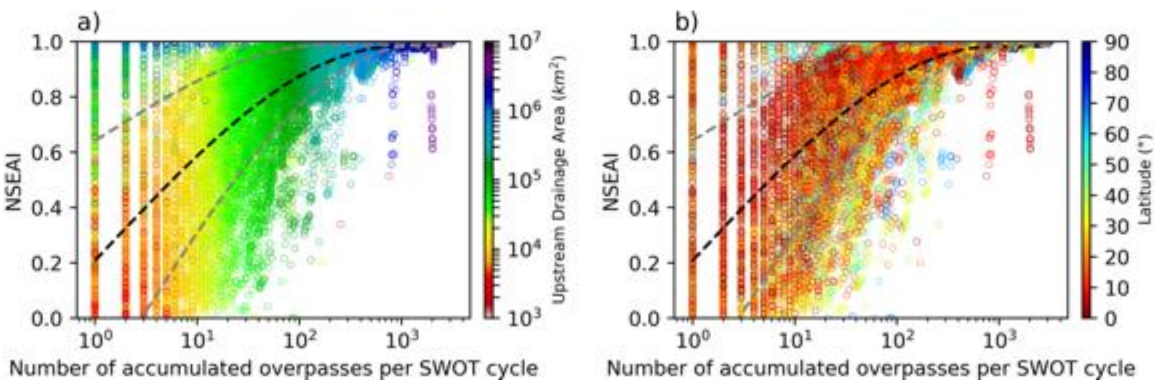

Figure S6: Relationship between accumulated overpasses with NSEAI according to a) upstream drainage area and b) latitude. b) Upstream drainage area with number of accumulated overpasses according to the latitude. The colors represent the upstream drainage area of each pixel in a) and the latitude of each pixel in b) and c).

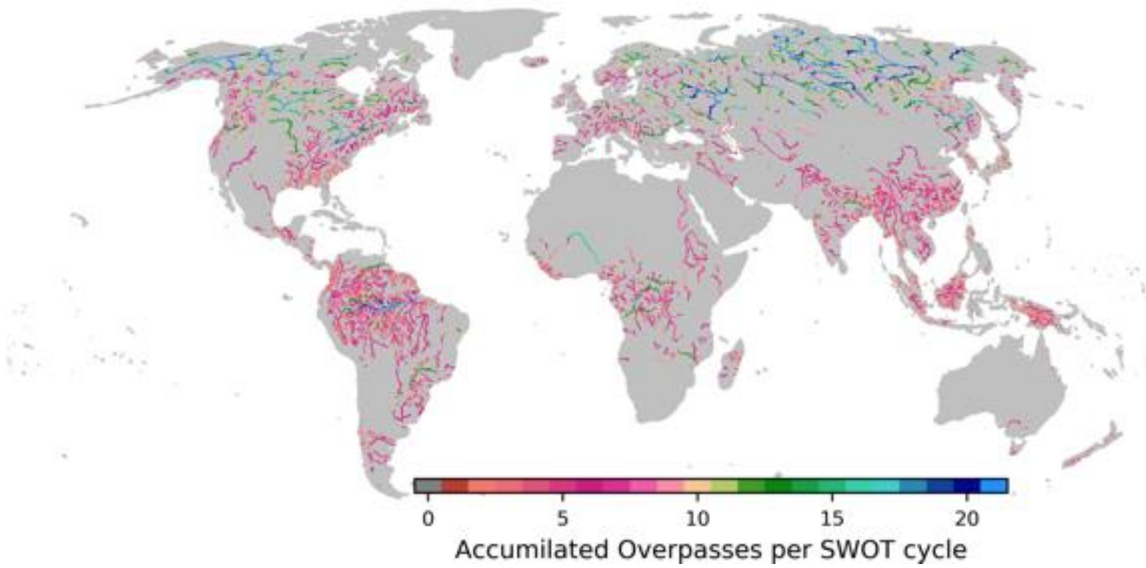

Figure S7: Assimilation Frequency per SWOT cycle. Pixels with mean discharge > 100 m3/s were shown for visual purposes.

Reference

1. Desai, S., Rodriguez, E., Fernandez, D. E., Peral, E., Chen, C. W., Bleser, J.-W. De, & Williams, B. (2018). *Surface Water and Ocean Topography Mission (SWOT) Science Requirements Document. Rev B*. California. Retrieved from https://swot.jpl.nasa.gov/docs/D-61923_SRD_Rev_B_20181113.pdf
2. Durand, M., Fu, L. L., Lettenmaier, D. P., Alsdorf, D. E., Rodriguez, E., & Esteban-Fernandez, D. (2010). The surface water and ocean topography mission: Observing terrestrial surface water and oceanic submesoscale eddies. *Proceedings of the IEEE*, *98*(5), 766–779. https://doi.org/10.1109/JPROC.2010.2043031
3. Dutra, E., Gianpaolo, B., Jean-Christophe, C., Munier, S., Burke, S., Fink, G., et al. (2017). *Report on the improved water resources reanalysis Deliverable*. Retrieved from http://earth2observe.eu/files/Public Deliverables/D5.2 - Report on the Improved Water Resources Reanalysis (WRR tier 2).pdf
4. Hunt, B. R., Kostelich, E. J., & Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, *230*(1–2), 112–126. https://doi.org/10.1016/j.physd.2006.11.008
5. Iizumi, T., Takikawa, H., Hirabayashi, Y., Hanasaki, N., & Nishimori, M. (2017). Contributions of different bias-correction methods and reference meteorological forcing data sets to uncertainty in projected temperature and precipitation extremes. *Journal of Geophysical Research*, *122*(15), 7800–7819. https://doi.org/10.1002/2017JD026613
6. Miyoshi, T. (2011). The Gaussian Approach to Adaptive Covariance Inflation and Its Implementation with the Local Ensemble Transform Kalman Filter. *Monthly Weather*

Review, *139*(5), 1519–1535. https://doi.org/10.1109/ICCSCE.2012.6487150
7. Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., et al. (2004). A local ensemble Kalman filter for atmospheric data assimilation. *Tellus A: Dynamic Meteorology and Oceanography*, *56*(5), 415–428. https://doi.org/10.3402/tellusa.v56i5.14462
8. Revel, Ikeshima, Yamazaki, & Kanae. (2019). A Physically Based Empirical Localization Method for Assimilating Synthetic SWOT Observations of a Continental-Scale River: A Case Study in the Congo Basin. *Water*, *11*(4), 829. https://doi.org/10.3390/w11040829
9. Revel, M., Yamazaki, D., & Kanae, S. (2018). Model Based Observation Localization Weighting Function for Amazon Mainstream. *Journal of Japan Society of Civil Engineers, Ser. B1 (Hydraulic Engineering)*, *74*(5), I_157-I_162. https://doi.org/10.2208/jscejhe.74.5_I_157
10. Takata, K., Emori, S., & Watanabe, T. (2003). Development of the minimal advanced treatments of surface interaction and runoff. *Global and Planetary Change*, *38*(1–2), 209–222. https://doi.org/10.1016/S0921-8181(03)00030-4