SWOT data assimilation for operational reservoir management on the upper Niger River Basin

S Munier, A Polebistki, C Brown, G Belaud, D. P. Lettenmaier

AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2014WR016157

SWOT data assimilation for operational reservoir management on the upper Niger River Basin

S. Munier[1,2], A. Polebistki[3], C. Brown[4], G. Belaud[5], and D. P. Lettenmaier[2]

[1]Estellus, Paris, France, [2]Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington, USA, [3]Department of Civil and Environmental Engineering, University of Wisconsin-Platteville, Platteville, Wisconsin, USA, [4]Department of Civil and Environmental Engineering, University of Massachusetts Amherst, Amherst, Massachusetts, USA, [5]Montpellier SupAgro, UMR G-EAU, Montpellier, France

Key Points:
- EnKF is used to assimilate SWOT data into a coupled hydrodynamic-reservoir model
- The persistence of the assimilation greatly increases with the use of a smoother
- An automatic controller allows to optimally define reservoir releases

Correspondence to:
S. Munier,
simon.munier@gmail.com

Citation:
Munier, S., A. Polebistki, C. Brown, G. Belaud, and D. P. Lettenmaier (2015), SWOT data assimilation for operational reservoir management on the upper Niger River Basin, *Water Resour. Res.*, *51*, 554–575, doi:10.1002/2014WR016157.

Received 18 JUL 2014
Accepted 19 DEC 2014
Accepted article online 30 DEC 2014
Published online 26 JAN 2015

Abstract The future Surface Water and Ocean Topography (SWOT) satellite mission will provide two-dimensional maps of water elevation for rivers with width greater than 100 m globally. We describe a modeling framework and an automatic control algorithm that prescribe optimal releases from the Selingue dam in the Upper Niger River Basin, with the objective of understanding how SWOT data might be used to the benefit of operational water management. The modeling framework was used in a twin experiment to simulate the "true" system state and an ensemble of corrupted model states. Virtual SWOT observations of reservoir and river levels were assimilated into the model with a repeat cycle of 21 days. The updated state was used to initialize a Model Predictive Control (MPC) algorithm that computed the optimal reservoir release that meets a minimum flow requirement 300 km downstream of the dam. The data assimilation results indicate that the model updates had a positive effect on estimates of both water level and discharge. The "persistence," which describes the duration of the assimilation effect, was clearly improved (greater than 21 days) by integrating a smoother into the assimilation procedure. We compared performances of the MPC with SWOT data assimilation to an open-loop MPC simulation. Results show that the data assimilation resulted in substantial improvements in the performances of the Selingue dam management with a greater ability to meet environmental requirements (the number of days the target is missed falls to zero) and a minimum volume of water released from the dam.

1. Introduction

Large dams are designed to store water that can subsequently be released to meet requirements for uses such as hydropower and water supply, flood control, navigation, recreation, and others. As demands for reservoir storage to meet requirements for these uses increase (and hence water effectively becomes more scarce), it becomes increasingly important to optimize reservoir releases [*Bader et al.*, 2003; *Lautze and Kirshen*, 2007]. In recent decades, environmental requirements for water have increasingly been recognized, and these often take the form of minimum reservoir releases to meet minimum flow requirements downstream of reservoirs. The flow propagation in the river reaches depends on many factors such as the distance, the physical characteristics of the river (section geometry, slope, roughness) or inflows and outflows. In particular, in large river basins, time delays may range from a few days to a few weeks. For real time reservoir management, these factors have to be accounted for to optimize reservoir releases. Adapted automatic controllers can be used to compute optimal reservoir releases [see e.g., *Castelletti et al.*, 2008; *Goedbloed et al.*, 2011; *Nikoo et al.*, 2014; *Raso et al.*, 2014] that address multiple management objectives and constraints, such as maximizing water supply deliveries subject to maintaining minimum streamflow at strategic locations.

Nevertheless, the performance of such models or controllers are limited by our knowledge of the physical characteristics of the river reaches and floodplains downstream of reservoirs and are highly impacted by boundary conditions (e.g., inflows from the reservoir and the ungauged tributaries, as affected by rainfall and evapotranspiration). Two approaches can be taken to reduce the effects of errors in the upstream and lateral boundary conditions. The first one is the calibration of parameters of the hydrology model used to predict reservoir inflows, which can be based on historical observations. The second is to assimilate observations so as to correct the effects of model errors. One common problem is the availability of real-time data

that represent the hydraulic state of the river both upstream and downstream of the reservoir (e.g., water levels, extent, and surface slope). This is particularly true in the large rivers of Africa, which have limited real-time monitoring networks. For instance, in the Niger River Basin, distances between gauges are up to 200 km on the main stem in Mali and up to 600 km in Nigeria. Real-time communication is another issue that can degrade the performance of reservoir regulation. Some monitoring systems include remote communication with in situ sensors; however, at many stations measurements are made manually as is communication with reservoir operators. Furthermore, even in cases with in situ recording and transmission capabilities in remote locations, data transmission may be delayed by days or even weeks. When no data are available in real time, reservoir managers usually use some predefined reservoir rules to determine reservoir releases. Such rules are based on historical time series and manager's experience and may be approximated empirically to be simulated in data-scarce regions [for examples, see *Meigh et al*., 1999; *Coe*, 2000; *Guentner et al*., 2004]. Although reservoir-operating rules are well adapted to seasonal reservoir management, the major limitation in a real-time context is that the actual state of the river is not accounted for, which can lead to insufficient river flows during dry years or excess releases during wet years.

Recent developments in remote sensing techniques are bringing new perspectives to hydrology and water resources management, such as crop monitoring from satellite imagery [e.g., *Lu et al*., 2013; *Sakamoto et al*., 2014] or flood forecasting from altimetry [e.g., *Biancamaria et al*., 2011]. The planned Surface Water and Ocean Topography (SWOT) satellite mission, a collaboration between NASA (National Aeronautics and Space Administration) and CNES (Centre National d'Etudes Spatiales), will further facilitate such applications. The primary SWOT instrument will be a wide swath altimeter that will provide maps of water elevation and extent at an unprecedented spatial resolution (50–100 m) and precision (centimetric vertical accuracy when averaged over areas of 1 km^2, *Durand et al*. [2010]). Potential applications of SWOT data in the hydrological sciences have been demonstrated recently by several authors. *Andreadis et al*. [2007] assimilated virtual SWOT observations into a hydrodynamic model using an Ensemble Kalman Filter (EnKF) to reduce modeling errors due to uncertainties in lateral inflows over a segment of the Ohio River. A similar approach was applied successfully by *Biancamaria et al*. [2011] to correct water surface elevations in the Ob River. Other studies have focused on river characteristics (bathymetric depths and slopes, roughness) estimates by SWOT data assimilation [*Durand et al*., 2008; *Yoon et al*., 2012; *Durand et al*., 2014]. At the global scale, *Biancamaria et al*. [2010] characterized the error of discharge estimates in rivers wider than 50 m using rating curves, with a focus on the SWOT temporal sampling.

Previous studies that have evaluated the potential for SWOT hydrological applications have focused on the characterization of surface water bodies (rivers and lakes) or the improvement of water elevation and discharge estimates. The potential for use of SWOT measurements in an operational context, such as flood forecasting and water resources management, is of interest for two reasons. The first is that opportunistic use of SWOT data will most likely be possible, even in the context of a science mission (consider, for example, the case of extreme events such as major floods or droughts). The second is that SWOT represents a fundamental shift in technology from nadir pointing altimeters (essentially all previous satellite altimeters) to swath altimetry, which has important implications for fields other than hydrology and water resources (e.g., oceanography). It is likely that swath altimetry will become a standard in future missions, and in this sense, SWOT can be considered a proving ground for methods that likely will be applicable to future missions as well. It is important to note that SWOT is not intended to be an operational mission, but rather a scientific mission intended to demonstrate the potential of swath altimetery (which is a new concept). However, some provision will most likely be made for near real-time (which typically can be interpreted as latency of a day or two) release of a small subset of the entire SWOT data stream. Furthermore, if the mission is successful, operational follow-on missions may well be developed, following a research to operations pathway similar to what has occurred for meteorological satellites. Operational satellite missions usually provide data with latency of 24 h or less, which is the basis for the simulations we show below.

Reservoir operations have impacts on downstream and there is increasing concern related to the environmental impacts. However, they often operate in systems that are partially observed and with only partial release controls. While reservoir managers can observe their own reservoir storage and inflows, they are typically unaware of downstream conditions, including inflows from other tributaries (downstream of their facility), abstractions downstream and the operations of other facilities. This study shows how remotely observed data such as will be available from SWOT can be used to partially fill these gaps in knowledge and

Figure 1. The Upper Niger River Basin.

improve operations. Our objective was to investigate the potential use of SWOT observations for estimating reservoir operations and their resultant effects on downstream flows. As in previous studies [e.g., *Andreadis et al.*, 2007; *Biancamaria et al.*, 2011], we use a twin experiment consisting of a "true" simulation, from which synthetic SWOT observations are derived, and a corrupted simulation, in which synthetic SWOT data (including simulated errors) are assimilated. Both simulations are based on a coupled hydrology-hydrodynamic-reservoir model which represents the transformation of rainfall into runoff, streamflow propagation in the main stream channel, and volume variations in the reservoir. Reservoir releases are determined by an automatic controller and implemented in both the true and corrupted simulations. The controller accounts for the state of the river (water level and discharge) and the reservoir (storage volume) as well as the streamflow dynamics between the reservoir and a downstream point where a minimal flow constraint is imposed. We use the Model Predictive Control (MPC) framework [*Garcia et al.*, 1989] with a simplified routing model initialized at each time step by the states of the corrupted coupled model after the assimilation of SWOT observations for both the river and the reservoir. Finally, the simplified routing model is used in a optimization loop to determine the minimum reservoir release that is required to satisfy the downstream minimum flow constraint. Consequently, contrary to previous similar studies, the "true" simulation is not independent of the assimilation procedure in the Model Predictive Control Experiment since it is affected by dam regulation.

We applied the approach outlined above to the upper Niger River Basin (Figure 1), between Banankoro (upstream) and Kirango (downstream). The Selingue reservoir, located on a tributary near Banankoro, is used for hydropower generation and low-flow augmentation at Kirango and for the Niger Inner Delta (located just downstream of Kirango). We represented only a low-flow constraint at Kirango in this exploratory analysis, but other management objectives such as hydropower maximization can be added within the MPC framework. Although in situ data are available in the upper Niger River Basin, such data suffer from gaps and delays in transmission, and remote sensing data are a potential alternative and/or could provide complementary information. In this study (which should be interpreted as a proof of concept), we do use historically available in situ data for model calibration, and the model is assumed perfect. Extending our methodology to ungauged regions would require more attention to model calibration, although we note that new calibration techniques based only on remote sensing data are being developed, and may be applicable to this problem [see e.g., *Durand et al.*, 2008; *Yoon et al.*, 2012].

The remainder of the paper is organized as followed. First, the study domain and available data are described in section 2. Section 3 details the methodology, including the modeling framework, the synthesis of SWOT observations, the assimilation scheme, and the automatic controller. Finally, results are presented and discussed in section 4.

2. Study Domain and Available Data

The Niger River is the third longest river in Africa. Its upper basin (Figure 1) has one large reservoir (Selingue), which has a storage capacity of 2.1 km^3 and has been in operation since 1982 [*Zwarts et al*., 2005]. The mean annual inflow to Selingue is about 8 km^3 (or 250 m^3/s in average). A larger dam (about 6 km^3) is planned in Guinea on a Niger River tributary in order to satisfy increasing demand (hydropower, agriculture) in Guinea and Mali.

Seasonal variations in the discharge of the Niger River are mainly controlled by the monsoon during JJA, which account for average precipitation of about 1500 mm in the basin upstream of Koulikoro and an annual average discharge of about 1000 m^3/s during the last decade. The low-flow period extends from December to June, with more water released than stored in the Selingue reservoir. Reservoir releases are mainly for power production, but also support irrigation along the river. The effect of reservoir releases on discharge are felt downstream in the ~360 km reach to Markala. The time of travel through this reach is about 30 days during low-flow periods.

A small run-of-river reservoir at Markala, located a few kilometers upstream of Kirango, has a very limited storage capacity, but it raises water levels about 5.5 m above the floodplain so that about 100,000 ha of the cropped lands of the Office du Niger are irrigated by gravity [*Hertzog et al*., 2014]. Between Selingue and Markala, there are many other irrigated plots supplied by the Niger by submersion during floods, and by pumping during low flow. Downstream of Markala, the gauge station at Kirango requires a minimum discharge of 50 m^3/s, an environmental flow that is necessary to preserve aquatic life in the mainstream, before the river enters its inner delta [*Zwarts et al*., 2005].

Maintaining this flow requires releasing water from Selingue in order to compensate for withdrawals in this reach. However, neither the withdrawals nor the inflows from tributaries are gauged, so that water must be released based on river flow estimates. The monitoring network is managed by the Direction Nationale de l'Hydraulique in Mali. Between Selingue and Markala, water levels are read on graduated rods at seven monitoring stations which presently have no automated recording or data transmission capability. Only the station at Koulikoro is equipped with an automatic level sensor and data transmission capability. This absence of real-time spatial observations has implications for the ability of the reservoir system to meet environmental requirements. As we will show, this deficiency potentially can be met in part through the use of SWOT data.

Within the study area, discharge measurements were available from four in situ gauge stations: Banankoro, Selingue, Koulikoro, and Kirango (see Figure1). Data from the Banakoro station were provided by the Global Runoff Data Centre (GRDC), while data from the three other stations were provided by the Institut de Recherche pour le Développement (IRD, France). Discharge measurements for the four sites, shown in Figure 2, were used to calibrate the models. The figure shows that high flows are mainly controlled by the inflow from Banankoro. During low flow periods (January to May), the discharge at Banankoro is lower than 50 m^3/s and the discharge at Koulikoro and Kirango is maintained by the release from the Selingue dam. Also, the discharge at Kirango is far smaller than at Koulikoro, which is mainly due to the diversion at Markala and evaporation losses.

3. Methods

3.1. Overview

In order to assess the potential benefit of SWOT data for operation of the upper Niger reservoir system, we developed a modeling framework designed to schedule real-time reservoir releases depending on a given constraint, with the capability to assimilate synthetic SWOT water level observations. The main model has three components that simulate (1) the inflows from upstream and lateral catchments, (2) the propagation of river discharge in the reach between the upstream dam and the downstream low flow control point, and (3) the dynamics of the reservoir (volume and water level changes).

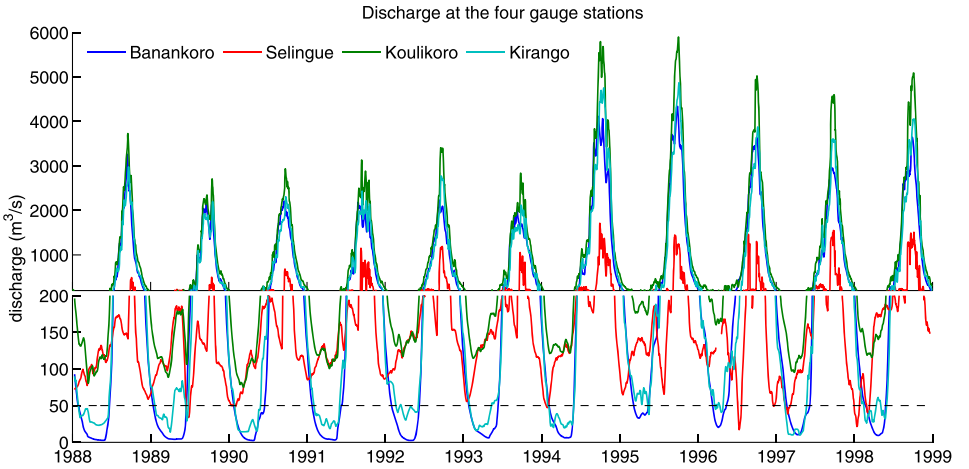

Figure 2. Discharge measurements from in situ gauge stations. The *y* axis has been stretched to enhance low flows.

As no SWOT observations are currently available, we set up a twin experiment based on historical data. The model was run with past meteorological observations to simulate the "true state" which was then used to produce synthetic SWOT observations. An ensemble of corrupted meteorological forcings was generated by adding noise (see section 3.4.2) and applied to the same model to compute an ensemble of model states that represents the best operating estimate. The ensemble of corrupted meteorological forcings was generated by adding some noise (see section 3.4.2). The twin experiment set up also allows simulation of different reservoir release scenarios. It is important to note that only the meteorological forcings were corrupted, but not the model itself (structure or parameters). This implies that model errors are assumed to be negligible compared to input data. Notwithstanding that this is clearly a simplifying assumption, our model calibration results (below) suggest that it may in fact be the case that most model error in this case is associated with errors in the model forcings.

We used an Ensemble Kalman Filter [EnKF, *Evensen*, 1994] to assimilate the synthetic SWOT water levels in order to correct the model state ensemble (water level in the river reach and in the reservoir). Consequently, at each time step (1 day in this study), the manager should have a reliable estimate of the river and reservoir states. This information, combined with potential future rainfall, was used to compute the optimal reservoir release that would satisfy the minimum discharge at the downstream end of the reach. Potential future rainfall may be provided by meteorological centers with associated uncertainties. Since in the upper Niger River basin, there is almost no precipitation during the low-flow period which is considered for dam operations, we used historical observations instead of potential future rainfall.

A first experiment, which we term "SWOT Data Assimilation" focused on different assimilation strategies and was aimed at evaluating the impact of SWOT data assimilation on predicted water levels and discharges. In this experiment, the reservoir was not considered. In a second experiment, called "Operational Reservoir Management," we implemented a Model Predictive Control (MPC) framework [*Garcia et al*., 1989] to compute reservoir releases automatically. Figure 3 shows the general framework of this experiment. On the figure, P represents the meteorological forcings and $\tilde{P}$ the corrupted ones. The discharge simulated by the hydrological model is denoted by $Q_{V\ IC}$; $Q_{release}$, Q_{target}, and Q_{down} stand for the Selingue reservoir release, the downstream required discharge, and the simulated downstream discharge, respectively. Q and H are the discharge and water level simulated by the hydraulic model of the reach downstream of the dam and H_{res} represents the water level in the reservoir. The superscript t is used for the "truth" simulation.

Considering that the meteorological forcings to the large-scale hydrology model can in principle be provided by remote sensing observations and/or atmospheric reanalyses, our general methodology is well adapted to ungauged basins since no in situ discharge or water level data are required. The river reach considered here is in fact at least partially gauged, but we argue that this study constitutes a hypothetical case that allows exploration of the capabilities of SWOT data for operational water management. Furthermore, in

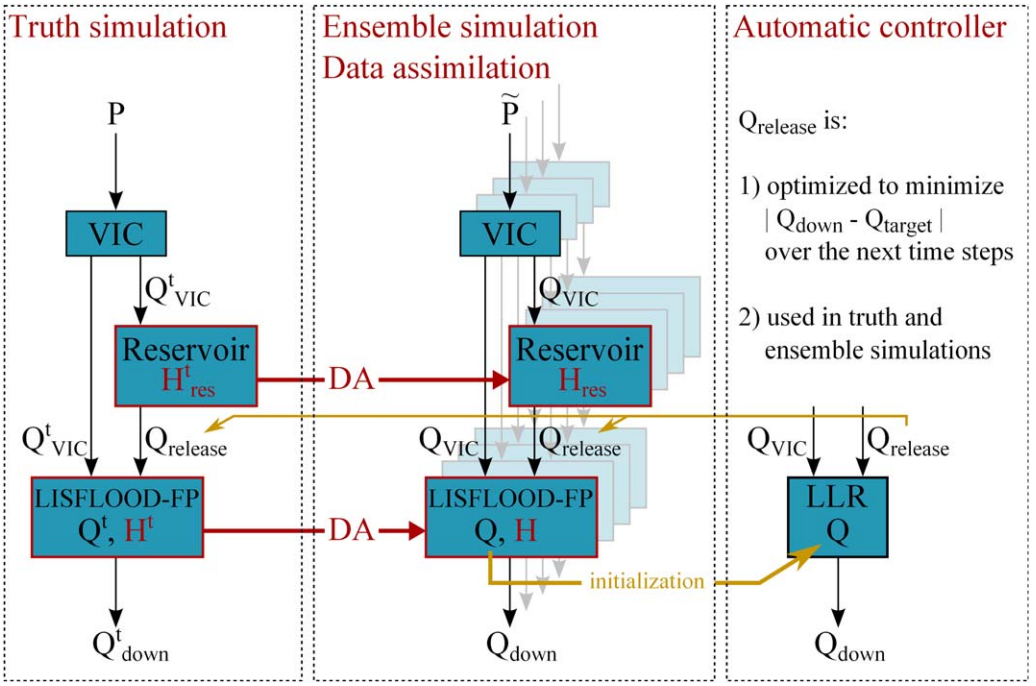

Figure 3. General framework of the Operational Reservoir Management experiment.

order to simplify the model structure, we considered only one operational use of Selingue reservoir: maintenance of low flows. Nevertheless, other constraints, such as hydropower, could be quite easily added in the MPC framework which is well adapted to multiobjective control.

3.2. Modeling Framework

Our basic modeling structure consists of three models that simulate the hydrology of the basin upstream of the dam, that propagate water through the main channel of the river downstream of the Selingue reservoir, and that represent the dynamics of the storage variations in the Selingue reservoir. We summarize each of the three models in the following sections.

3.2.1. Hydrological Model

The hydrological model is the Variable Infiltration Capacity [VIC, *Liang et al*., 1994] model, which simulates the vertical fluxes at the land-atmosphere interface using four primary atmospheric forcings (precipitation, wind speed and minimum, and maximum daily temperature) and three derived forcings (downward solar and longwave radiation, and dew point). The model-predicted runoff is routed through the river network using a simple routing model (kinematic model developed by *Lohmann et al*. [1996]), which produces river discharge into the reservoir. We also apply the VIC model to catchments which provide lateral inflows to the river downstream of the reservoir. These lateral inflows, along with reservoir releases, provide the discharge boundary conditions to the hydrodynamic model (section 3.2.2). The VIC model was run at one-half degree latitude-longitude spatial resolution using forcings from a gridded meteorological data set developed by *Sheffield et al*. [2006]. The model can be run at different time steps. Here a daily time step and the water balance mode were chosen, meaning that the surface temperature is assumed equal to surface air temperature, i.e., as in most hydrology models, there was no iteration for surface energy balance closure. Five VIC model parameters were calibrated using the Shuffled Complex Evolution algorithm [*Duan et al*., 1992] with Nash-Sutcliffe and Root Mean Square Error (RMSE) values (predicted-observed discharge at the four stations: Banankoro, Selingue, Koulikoro, and Kirango) used as objective functions. Calibration was performed over the period 1970–1980 at multiple subbasins in the Upper Niger basin. Results of the calibration are discussed in section 4.

3.2.2. Hydrodynamics Model

The LISFLOOD-FP hydrodynamic model [*Bates and Roo*, 2000] was used to propagate the discharge predicted by VIC at the Banankoro station and the Selingue dam to the Kirango station (downstream).

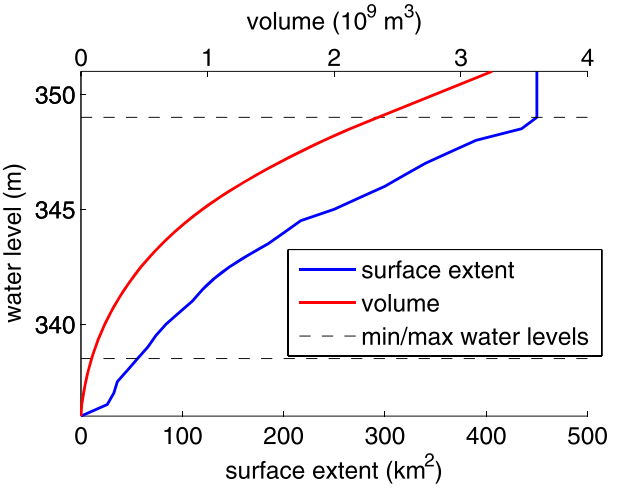

Figure 4. Variations of surface extent and volume with respect to water elevation for the Selingue reservoir.

LISFLOOD-FP uses a coupled 1-D/2-D hydraulic model to simulate flood wave propagation over a delimited domain. Hence it provides water surface elevations and discharge at each time step and at each pixel within a channel (and overbank if so desired) domain. A subgrid model [*Neal et al.*, 2012] was used to improve the simulation of flow propagation through the domain by considering a main channel with specific physical characteristics (width, bottom and bank elevation, roughness, etc.). The *Neal et al.* [2012] study focused on the Niger Inner Delta, which is located downstream of Kirango. In our study, we used the same methodology but we recalibrated the parameters for the upstream portion of the basin which is of interest here. We used floodplain topography from the SRTM-based HydroSHEDS database [*Lehner et al.*, 2008] with a resolution of 910 m. The Manning coefficient was assumed constant and uniform, equal to 0.05 for the floodplain and 0.035 for the river. We ran LISFLOOD-FP at a varying time step (initially 50 s) with the "Acceleration and Adaptive time step" version of the model. Model output was daily. In addition, the calculation of the normal depth (water depth when the surface slope equals the bottom slope, usually computed using Manning's equation) was used as the downstream boundary condition.

LISFLOOD-FP includes a loss term that can represent, for instance, evaporation from the river surface and infiltrations through the bed. The loss term is assumed uniform in space but may vary in time. Note that the loss term could easily be made more complex spatially if desired. This term is quite important in our case as only about two thirds of the volume entering the upstream channel (Banankoro station plus Selingue dam) appears at the Kirango downstream station. These losses appear to be attributable to evaporation and infiltration as well as withdrawals for rice irrigation in the "Office du Niger" agricultural domain. For our purposes, we do not need to partition among these types of losses, but we do need to represent total losses. To quantify total losses, we first ran the model without any losses, then subtracted the observed discharge at Kirango from the simulated discharge, and computed a seasonal loss term (long-term monthly mean). The daily loss term used in LISFLOOD-FP was then obtained via a mass conservative linear interpolation of the seasonal loss term. Validation results are presented in section 4.

3.2.3. Reservoir Model

A reservoir model was used to simulate the variations of volume and water level in the Selingue reservoir. A simple water balance of the reservoir was used at each time step, with volume variations computed from inflows from the upstream catchment and outflows to the river (reservoir releases). A predefined volume-water elevation relationship, elaborated from observations [*Zwarts et al.*, 2005, p. 24], was then used to update the reservoir water elevation (see Figure 4). Minimum and maximum possible water elevations were taken as constraints: when the reservoir water elevation reaches its minimum value, no more water can be released, whereas if it reaches its maximum value, no more water can be stored in the reservoir (thereafter outflows equal inflows so long as the water level condition continues). Additionally, a minimum flow release was predefined to sustain low flows downstream of the dam, except if the minimum water elevation was reached. The minimum flow was set to 10 m^3/s. This minimum flow is used only when no release is required to meet the downstream discharge requirement and when the reservoir is not full. This occurs when the discharge at Banankoro is quite large, usually at the beginning of the high-flow period when the reservoir refills. In practice, the only impact of the minimum flow is that the reservoir refill is slightly slower.

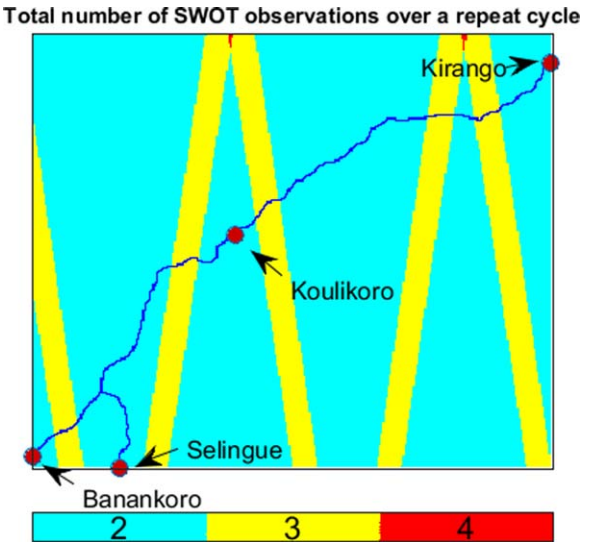

Figure 5. SWOT coverage, number of observations. The considered river reach is shown in dark blue.

3.3. SWOT Observations

The SWOT satellite is planned to be launched in 2020. A full description of the mission and its specifications is given by *Rodriguez* [2009]. Basically, it will provide high-resolution images of water surface elevations over oceans and continental water bodies using a 120 km width swath radar interferometer. The repeat period will be 21 days with global coverage (between latitudes 78°N and 78°S). Figure 5 shows the number of observations over the study domain during each 21 day cycle. Since SWOT data are not yet available, we generated SWOT observations using the reference coupled model forced by historical meteorological data, to simulate the "true" state. SWOT observations were generated using the actual orbit characteristics, as obtained from *Rodriguez* [2009]. For each time step with available SWOT observations, water surface elevations simulated by the hydrodynamic model with the SWOT swath were corrupted with (additive) observation errors. We used a random white noise of 10 cm standard deviation to represent observation errors [*Enjolras et al*., 2006; *Rodriguez*, 2009]. It should be noted that the error magnitude is somewhat optimistic as it does not include satellite position uncertainties and atmospheric effects (which typically are spatially correlated, but temporally uncorrelated). We explored sensitivity to different kinds of SWOT errors in section 4.4.1.

3.4. Assimilation Scheme

The assimilation scheme and method for generating ensemble members are as in *Biancamaria et al*. [2011]. Hence we only give some summary information here, and refer the reader to this article for further details.

3.4.1. Local Ensemble Kalman Smoother

The Kalman filter [*Kalman*, 1960] is a sequential data assimilation scheme widely used in hydrological sciences. At each time step, new observations are combined with the model outputs derived from the simulated state (*forecast*) to compute an update state (*analysis*). The latter is obtained by optimally accounting for observation and model errors. Here the system state includes water levels and discharges, while observations are SWOT water levels. The original Kalman filter, based on linear systems and Gaussian errors, has been extended to nonlinear models. One of the most popular extension is the Ensemble Kalman Filter (EnKF) described by *Evensen* [1994]. It is based on a Monte Carlo approach to approximate the error covariance matrix, which usually is not known:

$$P_e = \overline{(A-\bar{A})(A-\bar{A})^T} \tag{1}$$

where A is the forecast state ensemble, P_e the ensemble covariance matrix, and the superscript T the transposition operator. Each column of matrix A represents a specific ensemble member state.

We implemented the square root analysis algorithm developed by *Evensen* [2004] in order to avoid measurements perturbations. With the same notation as above, the analysis step is given by:

$$A^a = A + P_e H^T \left(HP_e H^T + R\right)^{-1}(D - HA) \tag{2}$$

where A^a is the analysis state ensemble, D the observation matrix (columns are identical and represent the observation vector), H the observation operator, and R the measurement covariance matrix. Note that R is always assumed diagonal, i.e., each observation error is assumed independent.

As in *Biancamaria et al*. [2011], we included in the analysis step a correlation matrix S representing the region of influence of each observation to avoid spurious long-range correlations in the error covariance

matrix. This idea was first introduced by *Hamill et al*. [2001] in order to limit the influence of observations to the surrounding region. Equation (2) then becomes ($\times$ represents the Schur products, i.e., element by element multiplication):

$$A^a = A + \left[S \times (P_e H^T)\right] \left\{H\left[S \times (P_e H^T)\right] + R\right\}^{-1} (D - HA) \tag{3}$$

As suggested by *Hamill et al*. [2001], we chose a fifth order function from *Gaspari and Cohn* [1999] to compute matrix *S*. Then the correlation factor from *S* equals 1 at the point where a SWOT observation is available, and decreases to 0 following a Gaussian curve. The Gaspari function depends on a length-scale relative to the distance where it reaches 0. After some trials, we found satisfactory results without any spurious long-range correlations using a 200 km length scale.

Finally, no assimilation is possible at time steps with no available SWOT observation. When this happens, we also used a smoother in order to keep the benefit from the update operated at the previous time steps. Note that the smoother goes both forward and backward in time (essentially post processing), whereas filters goes only forward in time. The objective was to propagate the error covariance forward in time until a new observation is available. The analysis equation is then replaced by [*Moore*, 1973]:

$$A_j^a = A_j + \left[S \times (P_{eij} H^T)\right] \left\{H\left[S \times (P_{eii} H^T)\right] + R_i\right\}^{-1} (D_i - HA_i) \tag{4}$$

$$P_{eij} = \overline{\left(A_j - \overline{A_j}\right)\left(A_i - \overline{A_i}\right)^T} \tag{5}$$

where *i* and *j* represent two different time steps ($j > i$), with and without SWOT observation, respectively. Note that only past observations are used by the smoother, or in other words, future observations, that are not available at current time, are not used.

In the following, the final assimilation scheme is called *Local Ensemble Kalman Smoother* (LEnKS). It is applied to the LISFLOOD-FP and the reservoir models. The reservoir state (water level) could have been included into the LISFLOOD-FP state, leading to a simultaneous correction of water levels in the river and in the reservoir. However, because of the dam, water levels in the reservoir are quite decoupled from water levels in the river, and we prefer to apply the LEnKS to each model independently.

3.4.2. Ensemble Member Generation

The ensemble of model states is intended to represent model uncertainties, including uncertainties in the meteorological forcings, the model structure, and the parameters. As in *Andreadis et al*. [2007]; *Biancamaria et al*. [2011], only the first category is considered here. To generate the ensemble, the forcings of the hydrological model (precipitation, mean and range temperature and wind) were corrupted using the methodology initially developed by *Auclair et al*. [2003] which consists of perturbing the most statistically significant modes. Each forcing field is first decomposed into Empirical Orthogonal Functions (EOFs), then the first modes that explain 95% of the variance are multiplied by a white noise with a 0.2 standard deviation before the signal is reconstructed. The main advantage of the method is that the corrupted fields have coherent spatial and temporal patterns. To limit computational costs, we used 20 ensemble members, which means that 20 corrupted fields of each model forcing have been generated. The example of the spatially averaged precipitation of the 20 members is shown in Figure 6.

3.4.3. Inflow Correction

At each time that SWOT observations are available, the data assimilation scheme corrects water levels in the river reach resulting from errors in inflow boundary conditions (which come from hydrology model predictions). The smoother is used to propagate this correction in time until the next SWOT observation. The smoother is more efficient if the inflows are corrected simultaneously. *Madsen and Skotner* [2005] and *Andreadis et al*. [2007] used an error forecast model for that purpose.

In this study, we assumed that the data assimilation procedure was able to correct the downstream discharge efficiently, and we then compared it with the downstream discharge computed in parallel without applying LEnKS. This resulted in a multiplicative correction factor applied to each inflow. The correction factor was computed for each member and updated each time SWOT observations were available. In order to account for the propagation time in the river reach, the correction factor was averaged over the last 30 days. The correction factor is then calculated from the 30 days prior to the current observation time and is

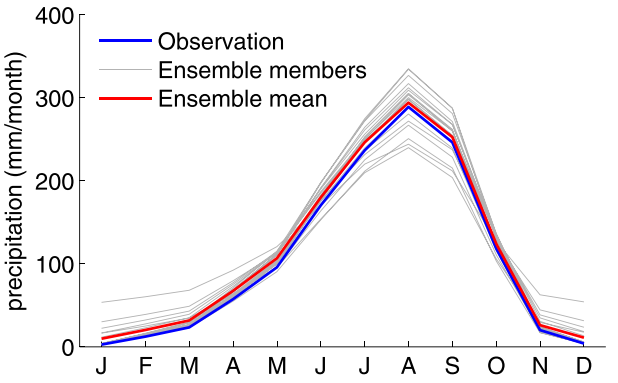

Figure 6. Spatial average of precipitation for the 20 members and comparison with the original measurements.

applied to the inflows until the next SWOT observation becomes available. This technique greatly reduces the biases of each corrupted member.

During the low-flow period, rain events are quite rare, leading to smooth variations of the river flows. Hence, at the first order, the methodology used to corrupt the meteorological forcing may be equivalent to a simple additive bias. The inflow correction we implemented behaves like a smooth time-varying bias correction. Other methods, such as error models or state augmentation, would certainly be more efficient if high flows are also under consideration. In our case, we preferred to use a correction method that was as simple as possible while still being reasonably representative of errors likely to be experiences in practice.

3.5. Model Predictive Control

3.5.1. Principle

A common objective of water managers is to release water from upstream reservoirs in order to satisfy different kinds of constraints downstream. In the interest of simplification, we considered only one constraint: maintaining a minimum flow of 50 m^3/s at Kirango, located just upstream the Niger Inner Delta. The simplest strategy considers that reservoir releases equal the difference between observed and required flows at the same time step. Such an example is presented in the results section which follows, but not surprisingly, the constraint often is not satisfied because the method does not account for evaporative losses as well as for the propagation time between Selingue and Kirango, which is about 30 days during the low flow periods. Consequently, the manager has to anticipate potential downstream flow decreases in order to release the optimum volume of water at the best time. To do that, some key information has to be known: the amount of water stored in the river reach, the propagation dynamics and potential future meteorological forcings.

The concept of Model Predictive Control (MPC) relies on an optimization loop in which different reservoir release scenarios are tested in order to find the optimal one. At the time step k, a propagation model is first initialized and a scenario is chosen and implemented in the propagation model as a boundary condition, as well as the future meteorological forcings. The propagation model computes the downstream discharge for the next time steps to a predefined leading horizon L. The latter has to be larger than the propagation time (here 30 days) else the effects of the reservoir release cannot be observed. Then a penalty function is used to evaluate the scenario, and if the penalty is still too high, the optimization is repeated with another scenario. When the optimum is found, the computed reservoir release is inserted into the reservoir model as a flow requirement and the optimization loop is reinitialized at time step $k+1$.

In this study, we used the ensemble of corrupted meteorological forcings computed for the ensemble assimilation procedure to simulate future meteorological forcings. The integrated model coupled with the SWOT data assimilation framework was used to estimate the system state which includes water level and discharge at any point in the river and in the floodplain. This state was used to initialize the propagation model. Finally, the propagation dynamics were approximated by the previously implemented LISFLOOD-FP model. Nevertheless, the optimization loop is operated at each time step and requires many propagation simulations (for each tested scenario). The number of iterations depends on the observed downstream discharge and the state of the river. It typically ranges from one to several tens. The LISFLOOD-FP model is computationally too expensive to be used in iterative procedures so we instead used a simplified flow propagation model described in the next section.

3.5.2. Simplified Streamflow Propagation Model

To simulate the propagation of streamflows through the river reach, we used a Linear Lag-and-Route model called LLR based on the model developed by *Munier et al*. [2008] in place of the more complex (but computionally expensive) LISFLOOD-FP model. Details of this model can be found in *Munier* [2009]. In essence, the river reach is decomposed into several segments of equal length in which the propagation is simulated by linear reservoirs. The LLR model takes the same inputs as the LISFLOOD-FP model, i.e., VIC and reservoir models outputs. The model state includes discharges at each segment boundary and the output is the downstream discharge at Kirango. This configuration provides a simple but accurate representation of the attenuation and time delay characteristics of flow dynamics [*Munier et al*., 2008]. Note that the LLR model only simulates discharge, and therefore cannot be used to assimilate SWOT water level data. The number of segments and the reservoir time constant are calibrated using a theoretical case applied to the LISFLOOD-FP model. Calibration results are shown in section 4. In contrast to LISFLOOD-FP, LLR does not account for floodplain storage, but this is not a constraint here since only low flows are considered.

In the MPC framework, at each time step, discharges at the segment boundaries are extracted from the LISFLOOD-FP model after the assimilation step, and used to initialize the LLR model. Outputs from VIC and evaporation in the channel, representing hydrological lateral flows, and Selingue reservoir releases from the different scenarios are then used as boundary conditions to compute the downstream discharge at Kirango. Computation times required by the LLR model are small, so that the model is well suited for the MPC optimization loop that operates at each time step.

3.5.3. Optimization Procedure

The objective of the optimization procedure is to find the optimum reservoir release defined as the minimum flow to be released from the Selingue reservoir that is required to maintain a downstream discharge at Kirango greater than the predefined threshold Q_{target} equal to 50 m^3/s. The optimization loop is initialized with a reservoir release ($Q_{release}$) equal to the predefined minimum flow release of 2 m^3/s. The LLR model is first initialized using the analyzed ensemble mean from LISFLOOD-FP (obtained after the assimilation step). Then for each reservoir release scenario, the cost function is computed using the downstream discharge simulated by the LLR model. $Q_{release}$ is iteratively incremented by 0.1 m^3/s until the cost is lower than a predefined cost threshold. The reservoir release is supposed to be constant in the future. Varying reservoir releases could be considered but the computation times would increase consequently. Besides the optimum release is applied only on the next time step since the optimization is reinitialized at the following time step and a new optimum release is computed.

The cost function has to account for downstream flows below the minimum requirement Q_{target} for the next L time steps (L is the leading horizon). It is then defined by:

$$J(Q_{release}) = \sum_{i=1}^{L} \alpha(i) \max\left[Q_{target} - Q_{down}(i), 0\right] \tag{6}$$

where Q_{down} is the downstream discharge simulated by LLR with $Q_{release}$, inflows from the VIC model and evaporation as boundary conditions. Function α is a weighting factor that accounts for the time period when the flow release at the next time step actually impacts the downstream discharge. The first few time steps are neglected because of propagation times. To define the weighting function, we propagated a step input through the LLR model and extracted the time period when the downstream discharge ranges from 0.1 to 0.9 (see Figure 7).

4. Results

This section is divided into three subsections. The first one summarizes the implementation of the VIC and LISFLOOD-FP models, the second one the SWOT data assimilation results, and the third one the results of the MPC applied to the integrated model.

4.1. Model Validation

In this twin experiment, the integrated model does not have to be perfect since the "truth" is simulated and the meteorological forcings are corrupted to simulate observation uncertainties. Modeling errors are not taken into account in this study. Nonetheless, the model dynamics have to be realistic in order to provide

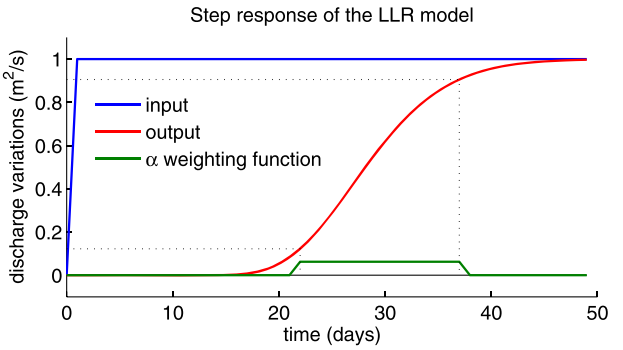

Figure 7. Propagation of a step input through the LLR model and the α weighting function.

confidence in the overall methodology. Therefore, we implemented and evaluated the VIC and LISFLOOD-FP models using real (*in situ*) data.

4.1.1. VIC

The parameters of the VIC model were calibrated using observed discharge at four stations: Banankoro, Selingue, Koulikoro and Kirango. The calibration period is from 1970 to 1980. Figure 8 compares the observed and simulated discharges at these stations. These graphs show that inflows to the target domain is simulated accurately by the VIC model, with Nash-Sutcliffe values ranging from 0.73 to 0.84.

The calibrated VIC model was then applied to provide hydrological inflows to the LISFLOOD-FP model.

4.1.2. LISFLOOD-FP

An initial LISFLOOD-FP simulation was performed with VIC outputs as upstream and lateral inflows. The resulting downstream discharge was greatly overestimated when compared with *in situ* observations. The difference is due to the effects of evaporation of water stored in the river and in the floodplain, as well as to withdrawals for irrigation, especially just upstream the Kirango station where a small diversion dam is used to irrigate rice crops in the Office du Niger area. Groundwater infiltration, although more difficult to estimate, may also play a role. The sum of these three terms was taken into account in the LISFLOOD-FP model by adding an "effective" evaporation term. This term, shown in Figure 9, was computed from the seasonal difference between the total inflows from VIC and the observed downstream discharge. It was then linearly interpolated to the LISFLOOD-FP time step. On average, the evaporation represents about 34% of the total inflow, and 61% of the downstream discharge.

The coupled VIC-LISFLOOD-FP model was run at a daily time step for the period 1988–1998. This period was chosen because very few observations were missing, and it is representative of the long term dynamics of the river (with "typical" high and low flow variations over a range of time scales). For instance, while the

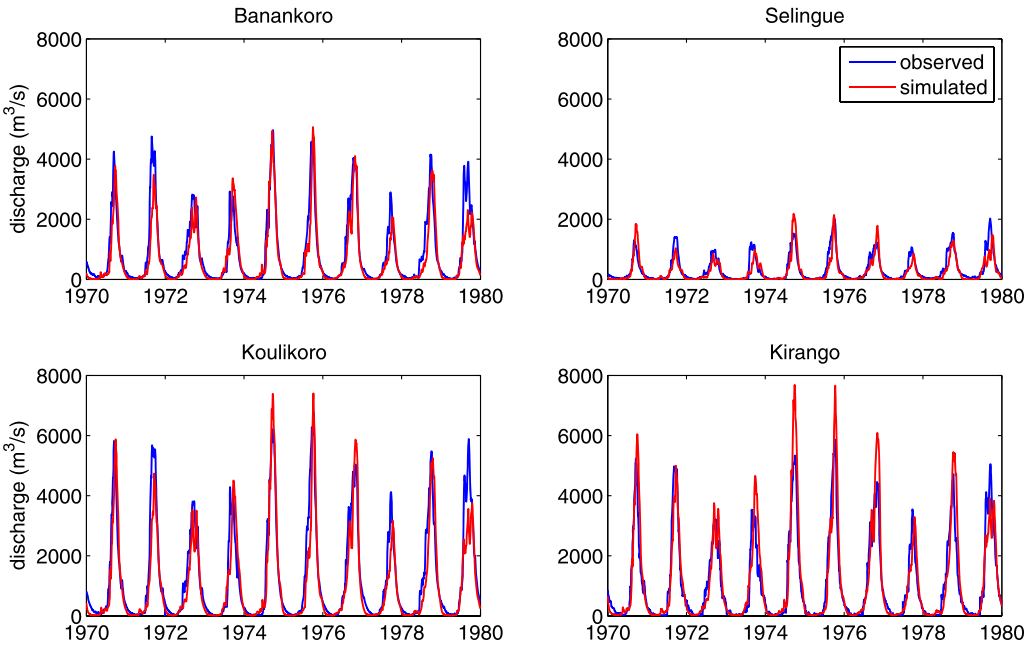

Figure 8. Comparison of the discharge simulated by VIC and observations at Banankoro, Selingue, Koulikoro, and Kirango.

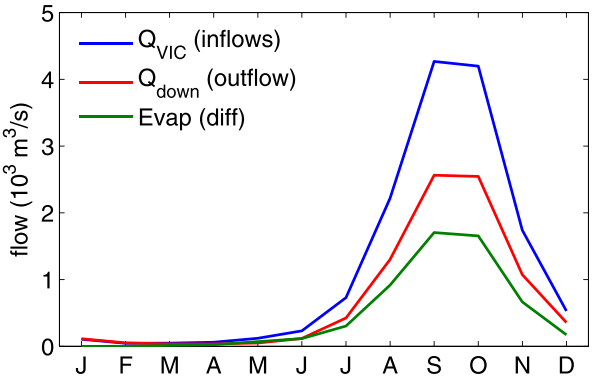

Figure 9. Seasonal mean of inflows simulated by VIC (Q_{VIC}), outflows simulated by LISFLOOD (Q_{down}) and evaporative losses ($Q_{VIC} - Q_{down}$).

period 1988–1993 was quite dry, 1994–1998 was much wetter. Figure 10 compares the simulated and observed discharge at Koulikoro and Kirango. The comparison shows excellent performance of the coupled VIC-LISFLOOD-FP model, with Nash-Sutcliffe values of 0.99 at Koulikoro and 0.97 at Kirango. Channel losses were computed by averaging the difference between the downstream discharge and total inflows for each month. The same losses were then applied for each year. In addition, the VIC model used to compute the inflows for the LISFLOOD-FP model was calibrated over a different time period than used for evaluation. We were therefore especially encouraged by the results shown in Figure 10. In any event, in the context of a twin experiment, it is important to understand the model has to be realistic, but not necessary perfect. For the same reason, the validation of water level is not essential for the twin experiment, although it would be for the application of the proposed framework to the actual operation. Here, given that virtual SWOT observation is generated by LISFLOOD-FP, spatiotemporal variations of water levels should be realistic.

4.2. SWOT Data Assimilation

In this section, we first describe the strategy for correcting water levels when SWOT data are assimilated into the LISFLOOD-FP model. We then summarize the implementation and performance of data assimilation algorithms based on the Local Ensemble Kalman Filter.

4.2.1. Water Level

Figure 11 shows an example of the effect of SWOT data assimilation on water depths at time step $k = 8$ (when the SWOT swath is crossing the study domain). The true water depth is shown in Figure 11(a). We used values located at the intersection between the river and the swath (grey shading) to simulate SWOT observation with a random noise having 10 cm standard deviation. We created an observation vector (columns of matrix D) including these values, as well as a localization matrix S which limits the impact range of

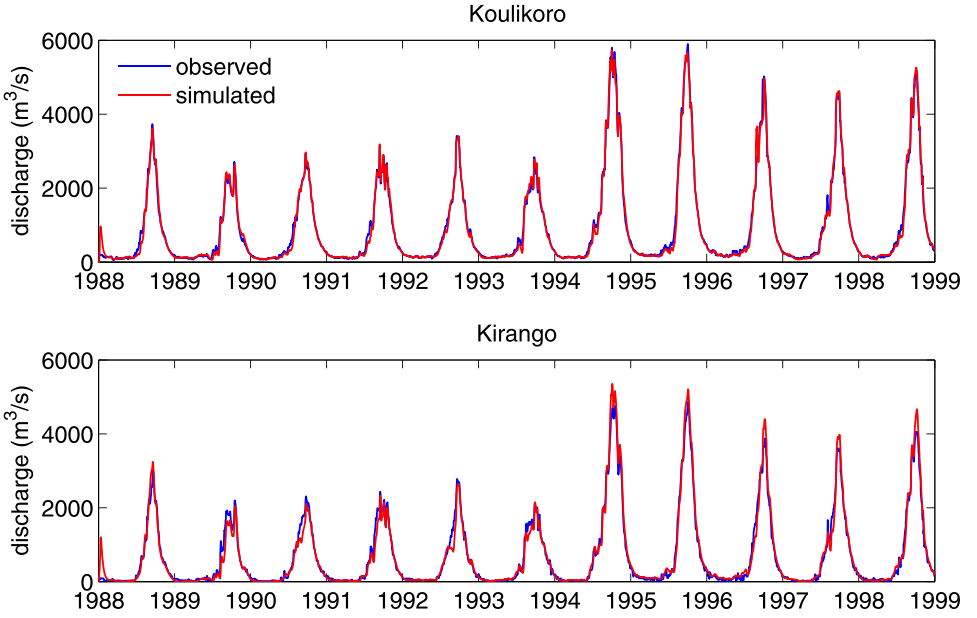

Figure 10. Comparison of the simulated and observed discharge at Koulikoro and Kirango.

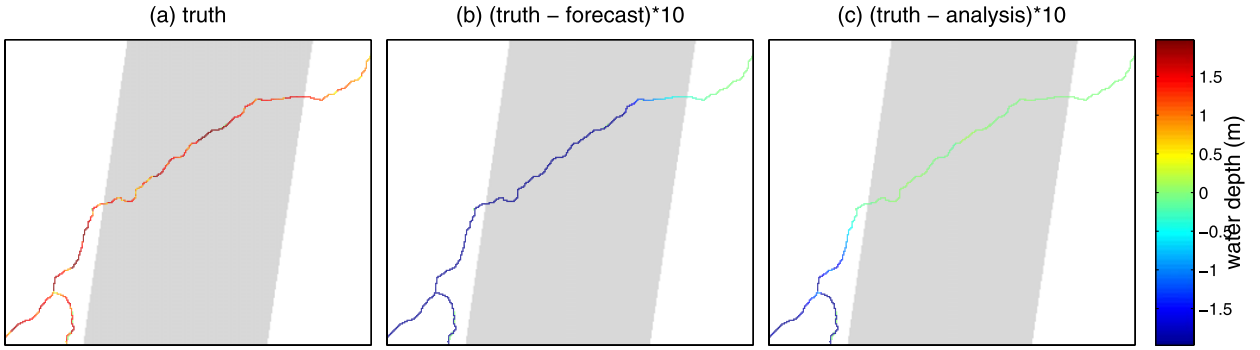

Figure 11. Water depth for time step $k = 8$: (a) truth, (b) truth minus forecast and (c) truth minus analysis. The gray shading represents the SWOT swath at this time step.

data assimilation. Matrices *D* and *S* were then used to correct the model state using equation (3). Figures 11b and 11c show the difference between the truth and the forecast and analysis water depths, respectively. The forecast clearly underestimates water depths under the swath. The LEnKS algorithm was able to correct the state, resulting in differences between the truth and the analysis that are close to zero under the swath. The RMSE for the whole river reach decreased from 52 cm for the forecast to 18 cm for the analysis, while it decreased from 36 cm to less than 1 cm for the region under the swath. Some points upstream of the swath boundary are partly corrected, while the most upstream points are not. This is due to the introduction of the localization matrix *S*. Finally, it is clear that the forecast is very close to the truth in the downstream part of the reach, which is not surprising because the SWOT swath crossed the downstream part of the river at the previous time step. Consequently, the water depths in this reach were corrected, and the correction propagated in time (and in space).

One common problem with filters like the EnKF, especially for small ensemble sizes, is insufficient ensemble variance which can lead to poor performance and filter divergence where the filter no longer responds to the observations [*Anderson*, 2007]. In our case, no divergence was observed. We believe that the good performances of the filter may be explained by the specificity of SWOT observations. When the SWOT swath crosses the study domain, it provides observations for all the pixels within a subreach of the river. Even with the use of the localization (which also tends to improve the filter performance), a lot of observations are used to correct the states corresponding to pixels within or near the subreach. Consequently, the weight of observations in the filter tends to increase, compared to the weight of the ensemble members, even with a low ensemble variance. Furthermore, the smoother also plays an important role in the performances of the filter. We made some simulations including covariance inflation with different inflation factors (not shown). The results showed an impact on the ensemble spread, but not on the filter performances.

4.2.2. Discharge at Kirango

We also analyzed the effects of the SWOT data assimilation on the downstream discharge at Kirango. We compared different algorithms: open loop (OL, no data assimilation), LEnKF, LEnKS, and LEnKS with the inflow correction (LEnKS+IFC). A 1 year simulation was performed (from July 1989 to June 1990), including the wet and dry seasons. Results are shown in Figure 12.

The OL run results show that the meteorological forcing corruption led to a wide spread in the discharge ensemble. The ensemble mean overestimates the truth by 44%. The application of LEnKF efficiently corrects the water levels and consequently the discharge each time SWOT observations were available. But when two consecutive observations were separated by several days, each ensemble member tended to return to its open loop state. This transition lasted a few days, which corresponds to the propagation time in the river reach. After this transition period and before a new observation was available, the system state was no more influenced by SWOT observations, but only by the inflows, which are not corrected. In this case, the persistence of the SWOT data assimilation was only a few days.

If the smoother was applied (Figure 12c), the persistence clearly increased due to the propagation in time of the error covariance until a new observation was available. It can be seen that the ensemble mean was generally close to the truth and the ensemble spread was highly reduced. The spread was even more

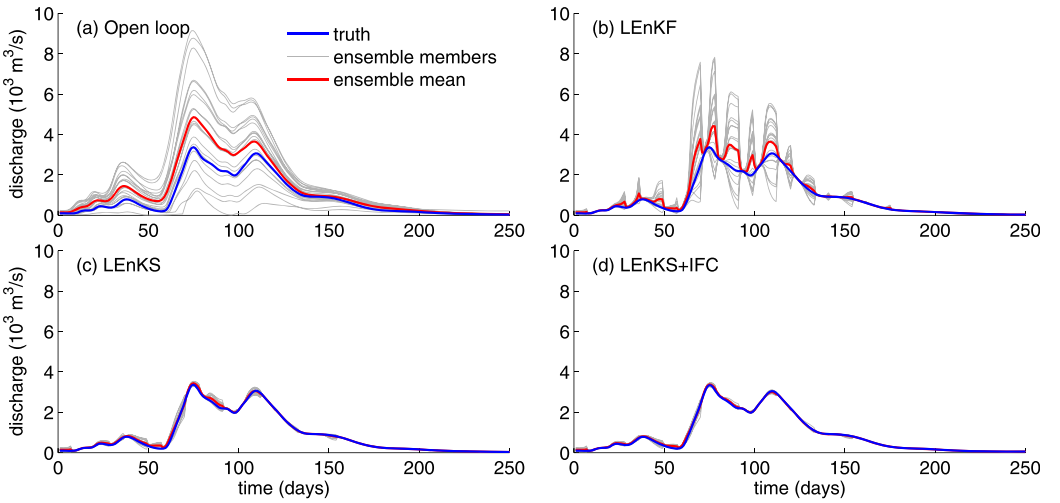

Figure 12. Effect of the SWOT data assimilation on the downstream discharge for different algorithms: (a) open loop, (b) LEnKF, (c) LEnKS, and (d) LEnKS with inflow correction.

reduced by correcting the upstream and lateral inflows (Figure 12d), leading to a slight increase of the persistence of SWOT data assimilation.

Quantitatively, the Nash-Sutcliffe criterion of the ensemble mean was 0.71, 0.91, 0.998, and 0.999 for the open loop, LEnKF, LEnKS, and LEnKS+IFC, respectively, showing the high efficiency of the LEnKS algorithm. The mean ensemble standard deviation was 419, 146, 25, and 18 m^3/s for the four simulations, respectively.

In addition, the assimilation was more persistent during low-flow than high-flow periods (Figure 13), mainly because of longer propagation times (due to lower flow velocity and wave celerity). For instance, the propagation time between Selingue and Kirango is about a week at high flows and about 25 days at low flows. The latter is almost three times larger than the maximum time between two observations (9 days in our case).

The black line in Figure 13 represents the minimum discharge requirement at Kirango (Q_{target}). The graph clearly shows that if no water is released from the Selingue dam, the management constraint is not satisfied.

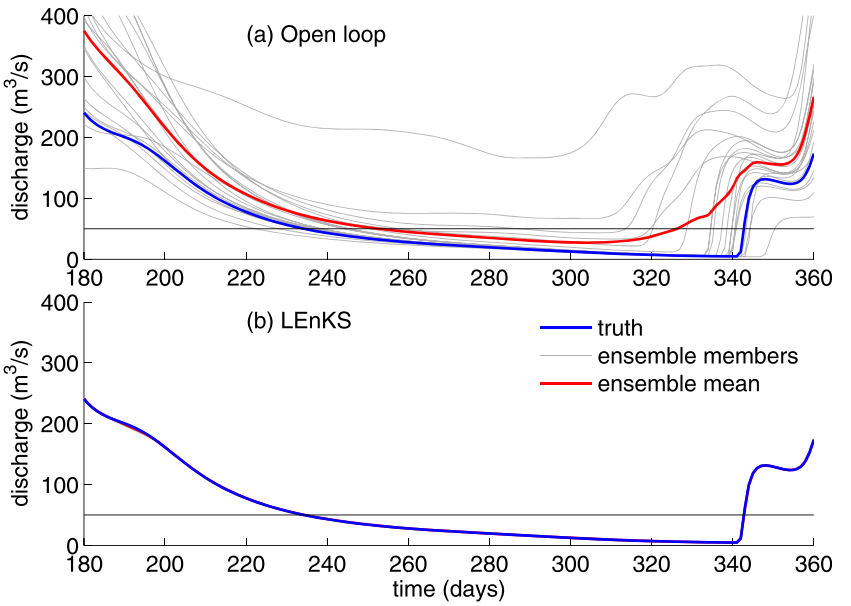

Figure 13. Downstream discharge for open loop and LEnKS during a low flow period.

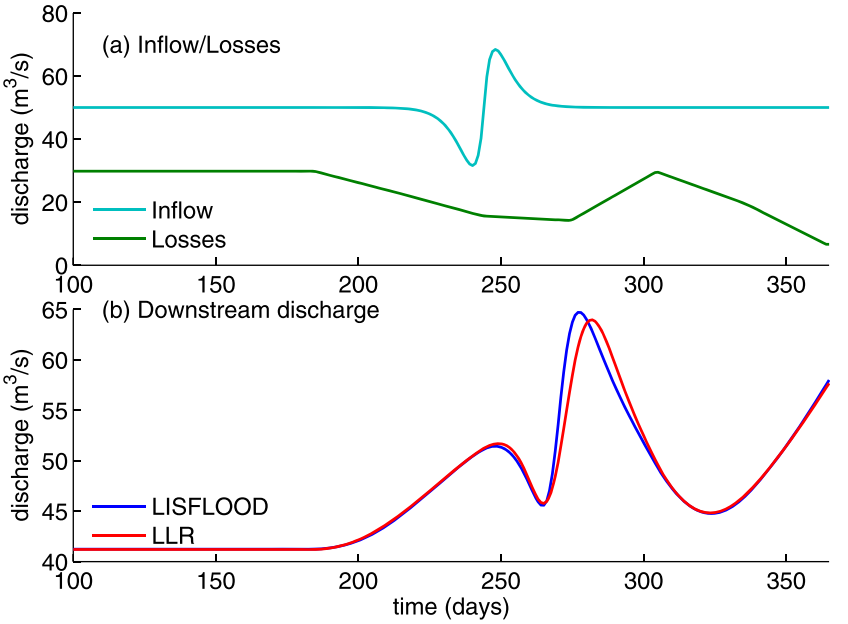

Figure 14. Calibration of the LLR model with a theoretical scenario: (a) inflows (discharge around 50 m^3/s at Selingue and 0 anywhere else) and outflow (evaporation), and (b) downstream discharge simulated by LISFLOOD-FP and calibrated LLR.

In this example, the downstream discharge is lower than Q_{target} for 100 days, with a minimum near zero at the end of the dry season. This motivates the need to manage the Selingue dam to respect the downstream low-flow constraint, which can be done with the MPC framework as described in the next section.

4.3. Operational Reservoir Management

In the previous section, we analyzed the efficiency of different data assimilation strategies and showed that the best one is LEnKS with inflow correction. This strategy is used in this section, where we present the results of the automatic control of the Selingue Dam with the objective of maintaining a minimum flow at the downstream station (Kirango).

In order to apply the MPC, we first have to set up the LLR model. Two parameters are calibrated: the number of reservoirs and the time constant. A theoretical scenario is used with the LISFLOOD-FP model, which simulates the propagation of an upstream discharge injected at Selingue. Evaporation previously computed is also simulated. The upstream discharge represents a flow variation around the predefined minimum discharge requirement Q_{target} of 50 m^3/s. Since the LLR model is linear, it will correctly simulate the flow dynamics for small variations around Q_{target}, which is the regime where it will be used in the MPC framework. The downstream discharge at Kirango simulated by LISFLOOD-FP is used as a reference to calibrate the LLR model. The best results were obtained with 14 reservoirs and a time constant of 1.70 days. Results are shown in Figure 14. The LLR model is clearly able to reproduce the flow propagation between Selingue and Kirango. The discharge simulated by the LLR model is slightly delayed during the peak flow, which can be explained by the nonlinearities of the flow dynamics: the flow velocity and the wave celerity increase with the discharge, which is not taken into account in the LLR model.

With the LLR model in place, it is possible to run the entire simulation including MPC and SWOT data assimilation. Figure 15 shows the results of three simulations, in terms of (a) downstream discharge and (b) volume in the Selingue reservoir. The first one corresponds to the case without regulation (no reservoir). This simulation (as shown also in Figure 10) clearly shows the need for reservoir releases for low-flow augmentation. The MPC was implemented in the two other simulations, coupled with and without SWOT data assimilation (DA on the figure). The effects of reservoir releases, which increase the downstream discharge during critical periods, are accounted for in both simulations. In general, the corrupted inflows overestimate the truth on average. Hence, if SWOT data are not assimilated, which means that the operator does not correct the model states using observations, the amount of water in the river reach is overestimated, especially at the downstream location. This results in reservoir releases that are lower than are necessary to meet the

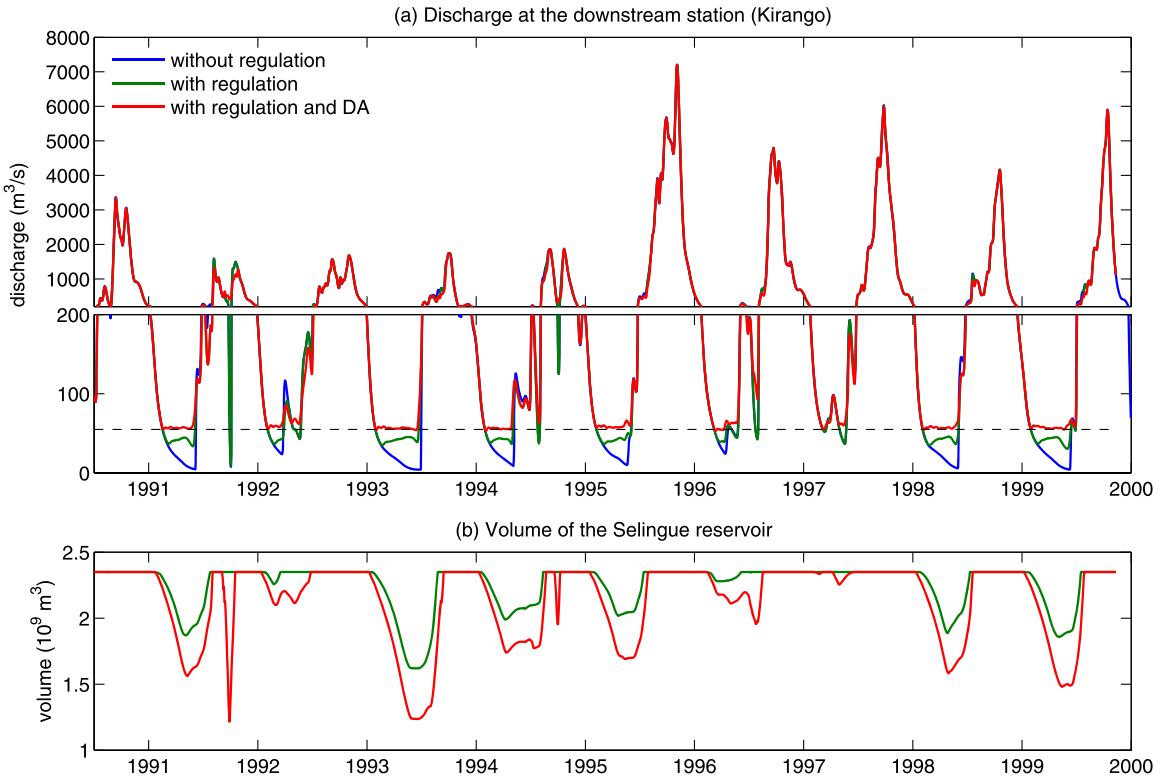

Figure 15. (a) Downstream discharge with and without regulation and (b) Selingue reservoir volume. The *y* axis on graph (a) has been stretched to enhance low flows.

downstream requirement and a downstream discharge lower than Q_{target}. In contrast, if the corrupted inflows tended to underestimate the truth, more water would be released from the dam and the downstream discharge would be greater than Q_{target}. In this case, the management constraint would be satisfied, but not optimally since the reservoir releases would not necessarily be minimum, and water would be released from the Selingue reservoir that otherwise could be saved for future use.

In contrast, if SWOT data are assimilated into the LISFLOOD-FP model, the model state is efficiently corrected (as shown previously) and the model is able to represent accurately the hydraulic state of the river. This better representation leads to a better estimation of required reservoir releases. As a result, the downstream discharge is never lower than Q_{target} and the management constraint is satisfied every year. Moreover, during critical periods, Q_{down} is very close to Q_{target}, which ensures the optimality of the method. Additionally, the minimum water volume in the reservoir is never reached, even during the driest years (e.g., 1993).

It is interesting to consider the very dry event during the wet season of 1991 (in October), which lead to a downstream discharge lower than the requirement for six consecutive days. The MPC anticipated this event and released a large amount of water during this short period. However, Figure 15 shows that resulting downstream discharge reached a minimum of 400 m^3/s, far greater than Q_{target}, which suggests that excess water was released in this particular case. The reason is that the LLR model is linear and has been calibrated around a reference discharge of 50 m^3/s. A nonlinear extension of the LLR model is currently in progress and should improve the results in conditions like this. Another option would be to stop the control procedure during the wet season. Meanwhile, as for low-flow periods, the simulation without data assimilation overestimated the downstream discharge and did not forecast that Q_{down} was about to decrease to such low values.

4.4. Discussion

4.4.1. SWOT Errors

One of the main limitations of this study is the assumption as to the nature of errors. These may include errors in meteorological forcings, SWOT observations, and model structure and parameters. Here errors in meteorological forcings are accounted for in the ensemble members generation step, and SWOT errors are assumed zero-mean Gaussian with a 10 cm standard deviation (std). Other SWOT errors, such as

troposphere reduction or orbit determination, are neglected, as well as modeling errors. Errors due to the troposphere or to the orbit would likely have time-varying bias, i.e., spatially correlated (almost uniform over such a small area) and temporally uncorrelated, with an std not exceeding 10 cm for a 1 km^2 pixel (SWOT Science Requirement Document, available at https://swot.jpl.nasa.gov/files/swot/SWOT_Science_Requirements_Document.pdf). Modeling errors are more difficult to characterize, as they are a combination of errors in the structure and in the parameters. Nevertheless, SWOT observations can also be used in a data assimilation framework to correct bathymetric errors [*Durand et al*., 2008; *Yoon et al*., 2012].

To gain some insight into the effects of such errors on the MPC performance, we performed some additional simulations with different SWOT measurement error characteristics. We compared four cases, each one including an uncorrelated random error (zero-mean Gaussian) with different std plus an optional bias: (a) random error (10 cm std) without bias, (b) random error (50 cm std) without bias, (c) random error (10 cm std) with a uniform but randomly time-varying bias (10 cm std), and (d) random error (10 cm std) with a constant bias (10 cm). Case (a) is the one presented in the previous section; case (b) is identical except that it has a much larger standard deviation. Case (c) simulates an additive error potentially due to tropospheric reduction or orbit errors; the 10 cm std for both random error and bias are higher than expected, i.e., this case is quite pessimistic. Case (d) represents what would be expected with some errors in the model parameters. More specifically, an additive 10 cm constant bias is equivalent, to first order, to an overestimation of the channel width (985 m instead of 800 m) or an underestimation of the Manning roughness coefficient (0.028 instead of 0.035). Figure 16 shows the simulation results for the first low-flow season. First, despite higher errors in (b) than in (a), the LEnKS algorithm is able to satisfactorily correct water levels and the downstream discharge is very close to the target when the errors are uncorrelated with no bias: under such conditions, the algorithm is very effective at filtering errors of considerable magnitude. This is because the uncorrelated nature of the errors makes it possible for the water level of all pixels observed by SWOT to be used to correct the water level at each pixel thanks to the local matrix. Since the error is assumed zero-mean Gaussian and spatially uncorrelated, the LEnKS is able to average out these errors. Results of case (c) show that the added bias may have a nonnegligible impact on the performances. Since the bias is temporally uncorrelated, the water levels are sometimes underestimated and sometimes overestimated, leading to oscillations of the downstream discharge around the target. As stated previously, this is a pessimistic case. Finally, in the last case, water levels are systematically overestimated by SWOT. Consequently, although the ensemble mean discharge remains close to the target, reservoir releases are always too low and the true downstream discharge is about 10 m^3/s bellow the target. These results shows that EnKF performs well when no bias is considered, whatever the error amplitude, while in contrast, biases may significantly impact the performances.

4.4.2. SWOT Data Latency

Another potential limitation of the approach is data latency, i.e., how close to real time will SWOT data be available? As noted above, SWOT is a research mission; however, some data will be available in near real-time, and furthermore, if SWOT is successful operational follow-on missions are a strong possibility. In either case, data latency order 1 to 2 days is realistic, given experience with "fast release" of data from other research satellites (e.g., the rapid response products from NASAs Earth Observing System satellites; see https://earthdata.nasa.gov/data/near-real-time-data/rapid-response). In consideration of 1 to 2 day latency, we showed that using the smoother with the ensemble Kalman filter substantially increases the persistence of the assimilation. More specifically, Figure 12d shows that the ensemble mean is close to the truth, even during periods when there is no SWOT data over the study domain, when the smoother (rather than filter) is used. For instance, we showed that the persistence is greater than 9 days, which is the maximum number of days without any SWOT observation. This is true at both high and low flows. Considering that the propagation time between Seligue and Kirango is about three times greater at low flows than at high flows, we can expect a persistence far greater than 9 days at low flows. To support this, we performed a new simulation with the assumption of only one SWOT observation for each 21 days cycle. The resulting downstream discharge is shown in Figure 17. Despite the fact that the assimilation performance is slightly worse at high flows, results clearly show that the persistence is greater than 21 days at low flows. Consequently, the impact of 1 to 2 days SWOT data latency on MPC performances would be negligible.

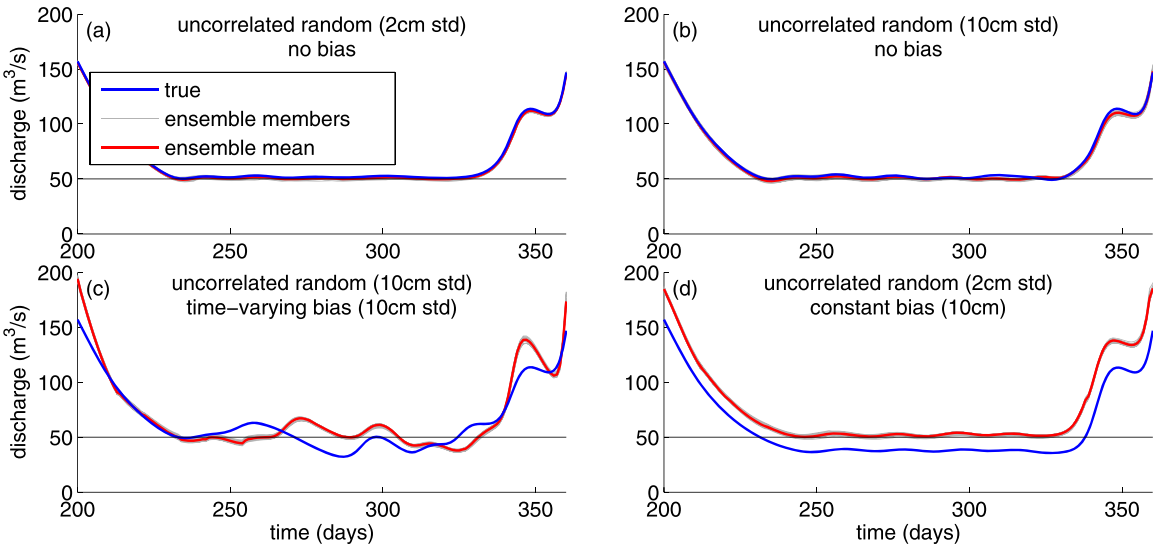

Figure 16. True and ensemble downstream discharge for different SWOT errors.

4.4.3. Calibration of Model Parameters

One of the main advantages of satellite observations is their global coverage. SWOT data thus will have potential value for water resources management, especially in ungauged regions. Nevertheless, our methods require the calibration of the model, which has been done here using historical observations. In cases where historical data exist, at least at the downstream end of the river reach, the method can be applied as developed in the paper (i.e., the model can be calibrated), and SWOT data can be very valuable, especially if in situ observations no longer exist (which is quite frequent in Africa). If the method is to be applied in completely ungauged regions, the model calibration would require more attention. Several teams worldwide are currently working on algorithms to retrieve river characteristics (width, bathymetry, roughness) from remote sensing data, including future SWOT data [see, e.g., *Durand et al.*, 2008; *Yoon et al.*, 2012; *Durand et al.*, 2014]. Even though some key challenges still remain, such studies are promising and we expect that when the SWOT satellite will be launched in 2020, they will provide a basis for model calibration.

That said, uncertainties (errors) in the model parameters do have an impact on our results. The hydrology model is used to simulate inflows into the reservoir and the river reach. The model was calibrated in order to provide realistic inflows for the "truth" simulation. Corrupting the meteorological forcings to generate the ensemble strongly impacted the reservoir and river inflows (Figure 12a which shows the downstream

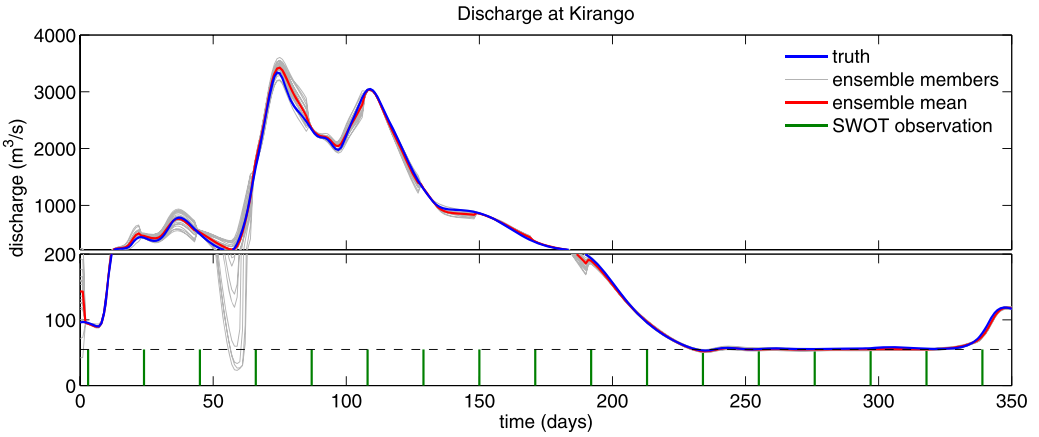

Figure 17. True and ensemble discharge at Kirango with only one SWOT observation for each 21 days cycle.

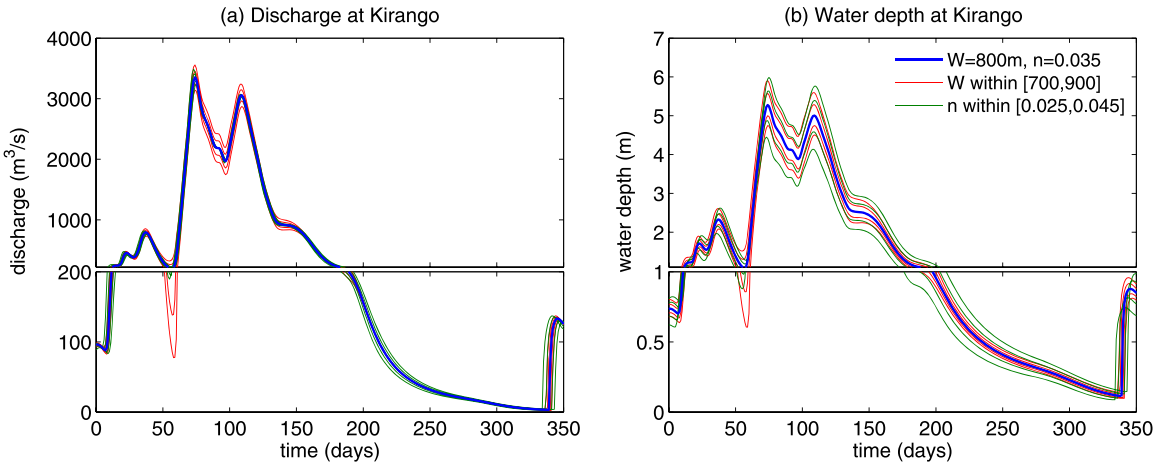

Figure 18. Discharge (a) and water depth (b) at Kirango with different values of the Manning coefficient (*n*) and of the river width (*W*). Same inflows are used for all the LISFLOOD-FP simulations; the SWOT data assimilation and MPC algorithms are not used here.

discharge obtained with different corrupted inflows). Uncertainties in the hydrology model parameters would have similar impacts on the inflows.

With respect to the hydrodynamic model, input uncertainties dominates parameters uncertainties in our case. To support this point, we ran the LISFLOOD-FP model with different values of the river width (from 700 m to 900 m versus 800 m truth) and of the Manning coefficient (from 0.025 to 0.045 versus 0.035 truth). The results, shown in Figure 18, demonstrate that the impact of these parameters variations on both the discharge and the water depth is far lower than the meteorological uncertainties shown in Figure 12a. Nevertheless, it is possible to integrate the parameter uncertainties into the Kalman filter algorithm, for instance by performing a dual state-parameter estimation, and this could be the topic of future work.

5. Conclusions

Our main objective was to investigate the potential of water level data from the future SWOT mission for operational water resources management. To that end, we implemented a modeling framework with data assimilation and an automatic control algorithm to define optimal reservoir releases from the Selingue Reservoir in the Upper Niger River Basin.

As in previous studies [*Andreadis et al.*, 2007; *Biancamaria et al.*, 2011], we first demonstrated the positive effect of SWOT data assimilation on the reservoirs ability to meet water level and discharge targets with an Ensemble Kalman Filter scheme. In addition, we evaluated the persistence of the assimilation, which describes the duration of the effect of assimilating SWOT observations a key concern, as the satellites average revisit time will be about 12 days in the vicinity of the Niger basin. We showed that using EnKF with a smoother much more effectively propagated the error correction forward in time (as contrasted with an algorithm that was forward looking only), and this substantially increased the persistence of the assimilation scheme, which is greater than 21 days under low flow conditions. A correction applied to the corrupted hydrological inflows also improved the persistence.

We compared the performances of the MPC with and without SWOT data assimilation, and showed that the assimilation of SWOT data into the hydrodynamic model greatly improved the results and led to optimal reservoir releases. Indeed, during critical (low flow) periods, the downstream discharge was very close to the minimum flow requirement, which suggests that the scheme is able to save water in the Selingue Reservoir.

In our view, the most important contribution of this proof-of-concept study is that it demonstrates that SWOT data can potentially be used for real time water management in an ungauged river basin. Nevertheless, we identified several key limitations that suggest future investigations. First, even though the most important errors in hydrological predictions generally originate from the meteorological forcings, other

sources of error, such as those related to the SWOT observation (see section 4.4.1) or to the model structure and its parameters (see section 4.4.3), may have a nonnegligible impact and should be considered in further studies. Second, only one management constraint was considered here (low flow sustainability), and considering others (such as hydropower) is certainly desirable. We note, though, that the MPC framework is well suited to such more general problems, since other management constraints can be quite easily added into the penalty function. Third, it is important to note that SWOT will be a scientific mission, and other constraints such as delays between receiving data and providing final products to the end-user will inevitably affect its operational use (however, provisions are being considered for some limited release of data in near real-time). Finally, in this study, no in situ data were assimilated, as our interest was in evaluating the potential capabilities of SWOT data. Incorporation of in situ data into the assimilation procedure (which is easily done in principle) could improve the results, especially in smaller basins where the SWOT revisit time will be large compared to the propagation time (persistence).

Although the SWOT mission is intended primarily for scientific rather than operational purposes, the entire mission can, in a sense, be considered a proof of concept that if successful, will likely motivate future missions that may well be more operational in nature. For this reason, understanding the operational implications of SWOT data is important not only for SWOT, but for potential future missions. The potential is seemingly greatest in those parts of the world where in situ data are sparse, which is a further motivation for our strategy of evaluating performance of the Selingue Dam in the absence of in situ data.

Acknowledgments

This work was funded in part by NASA grant NNX13AD98G to the University of Washington, and by the University of Washingtons Robert and Irene Sylvester Professorship. We also would like to acknowledge Jeffrey Neal (University of Bristol) and Konstantinos Andreadis (JPL, NASA) for their help on the LISFLOOD-FP model and the Ensemble Kalman Filter scheme, respectively. Data from the Banakoro station were provided by the Global Runoff Data Centre (GRDC), while data from the three other gauging stations were provided by the Institut de Recherche pour le Development (IRD, France). The meteorological data set was developed by *Sheffield et al.* [2006]. For data access, please contact the corresponding author (S. Munier, simon.munier@gmail.com).

References

Anderson, J. L. (2007), An adaptive covariance inflation error correction algorithm for ensemble filters, *Tellus Ser. A*, *59*(2), 210–224.

Andreadis, K. M., E. A. Clark, D. P. Lettenmaier, and D. E. Alsdorf (2007), Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model, *Geophys. Res. Lett.*, *34*, L10403, doi:10.1029/2007GL029721.

Auclair, F., P. Marsaleix, and P. De Mey (2003), Space-time structure and dynamics of the forecast error in a coastal circulation model of the Gulf of Lions, *Dyn. Atmos. Oceans*, *36*, 309–346.

Bader, J. C., J. P. Lamagat, and N. Guiguen (2003), Management of the Manantali Dam on the Senegal River: Quantitative analysis of a conflict of objectives, *Hydrol. Sci. J.*, *48*(4), 525–538.

Bates, P. D., and A. P. J. De Roo (2000), A simple raster-based model for flood inundation simulation, *J. Hydrol.*, *236*, 54–77.

Biancamaria, S., K. M. Andreadis, M. Durand, E. A. Clark, E. Rodriguez, N. M. Mognard, D. E. Alsdorf, D. P. Lettenmaier, and Y. Oudin (2010), Preliminary characterization of SWOT hydrology error budget and global capabilities, *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, *3*(1), 6–19, doi:10.1109/JSTARS.2009.2034614.

Biancamaria, S., M. Durand, K. M. Andreadis, P. D. Bates, A. Boone, N. M. Mognard, and E. A. Clark (2011), Assimilation of virtual wide swath altimetry to improve Arctic river modeling, *Remote Sens. Environ.*, *115*(2), 373–381, doi:10.1016/j.rse.2010.09.008.

Castelletti, A., F. Pianosi, and R. Soncini-Sessa (2008), Water reservoir control under economic, social and environmental constraints, *Automatica*, *44*(6), 1595–1607.

Coe, M. T. (2000), Modeling terrestrial hydrological systems at the continental scale: Testing the accuracy of an atmospheric GCM, *J. Clim.*, *13*(4), 686–704.

Duan, Q., S. Sorooshian, and V. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, *28*(4), 1015–1031.

Durand, M., K. M. Andreadis, D. E. Alsdorf, D. P. Lettenmaier, D. Moller, and M. Wilson (2008), Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model, *Geophys. Res. Lett.*, *35*, L20401, doi:10.1029/2008GL034150.

Durand, M., E. Rodriguez, D. E. Alsdorf, and M. Trigg (2010), Estimating river depth from remote sensing swath interferometry measurements of river height, slope and width, *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, *3*(1), 20–31, doi:10.1109/JSTARS.2009.2033453.

Durand, M., J. Neal, E. Rodriguez, K. M. Andreadis, L. C. Smith, and Y. Yoon (2014), Estimating reach-averaged discharge for the River Severn from measurements of river water surface elevation and slope, *J. Hydrol.*, *511*, 92–104, doi:10.1016/j.jhydrol.2013.12.050.

Enjolras, V., P. Vincent, J. C. Souyris, E. Rodriguez, L. Phalippou, and A. Cazenave (2006), Performances study of interferometric radar altimeters: From the instrument to the global mission definition, *Sensors*, *6*, 164–192.

Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics, *J. Geophys. Res.*, *99*(C5), 10,143–10,162.

Evensen, G. (2004), Sampling strategies and square root analysis schemes for the EnKF, *Ocean Dyn.*, *54*(6), 539–560.

Garcia, C. E., D. M. Prett, and M. Morari (1989), Model predictive control: Theory and practice: A survey, *Automatica*, *25*(3), 335–348.

Gaspari, G., and S. E. Cohn (1999), Construction of correlation functions in two and three dimensions, *Q. J. R. Meteorol. Soc.*, *125*(554), 723–757.

Ghile, Y. B., M. U. Taner, C. M. Brown, and J. G. Grijsen (2014), Bottom-up climate risk assessment of infrastructure investment in the Niger River Basin, *Clim. Change*, *122*(1-2), 97–11.

Goedbloed, A., S. Galelli, and D. Schwanenberg (2011), Assessing the effectiveness of a real-time control method for Marina Reservoir management, edited by F. Chan, D. Marinova, and R. S. Anderssen, pp. 1652–1658, *MODSIM2011, 19th International Congress on Modelling and Simulation*, Modelling and Simulation Society of Australia and New Zealand, December 2011.

Guentner, A., M. S. Krol, J. C. D. Araújo, and A. Bronstert (2004), Simple water balance modelling of surface reservoir systems in a large data-scarce semiarid region, *Hydrol. Sci. J.*, *49*(5), 901–918.

Hamill, T. M., J. S. Whitaker, and C. Snyder (2001), Distance-dependent filtering of background error covariance estimates in an Ensemble Kalman Filter, *Mon. Weather Rev. 129*(11), 2776–2790.

Hertzog, T., J. C. Poussin, B. Tangara, I. Kouriba, and J. Y. Jamin (2014), A role playing game to address future water management issues in a large irrigated system: Experience from Mali, *Agric. Water Manage.*, *137*, 1–14.

Kalman, R. E. (1960), A new approach to linear filtering and prediction problems, *Trans. Appl. SME J. Basic Eng.*, *82*(Series D), 35–45.
Lautze, J., and P. Kirshen (2007), Dams, health, and livelihoods: Lessons from the Senegal, suggestions for Africa, *Int. J. River Basin Manage.*, *5*(3), 199–206.
Lehner, B., K. Verdin, and A. Jarvis (2008), New global hydrography derived from spaceborne elevation data, *Eos Trans. AGU*, *89*(10), 93, doi: 10.1029/2008EO100001.
Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges (1994), A simple hydrologically based model of land-surface water and energy fluxes for general-circulation models, *J. Geophys. Res.*, *99*(14), 415–428.
Lohmann, D., R. Nolte-Holube, and E. Raschke (1996), A large-scale horizontal routing model to be coupled to land surface parametrization schemes, *Tellus Ser. A*, *48*(5), 708–721.
Lu, L., C. Wang, H. Guo, and Q. Li (2013), Detecting winter wheat phenology with SPOT-VEGETATION data in the North China Plain, *Geocarto Int.*, *29*(3), 244–255.
Madsen, H., and C. Skotner (2005), Adaptive state updating in real-time river flow forecasting: A combined filtering and error forecasting procedure, *J. Hydrol.*, *308*, 302–312.
Meigh, J. R., A. A. McKenzie, and K. J. Sene (1999), A grid-based approach to water scarcity estimates for eastern and southern Africa, *Water Resour. Manage.*, *13*(2), 85–115.
Moore, J. B. (1973), Discrete-time fixed-lag smoothing algorithms, *Automatica*, *9*, 163–173.
Munier, S. (2009), Modélisation intégrée des écoulements pour la gestion en temps réel d'un bassin versant anthropisé (Integrated flow modeling for real time management of an anthropized basin), PhD thesis, 241 pp., AgroParisTech, Montpellier.
Munier, S., X. Litrico, G. Belaud, and P. O. Malaterre (2008), Distributed approximation of open-channel flow routing accounting for backwater effects, *Adv. Water Resour.*, *31*(12), 1590–1602, doi:10.1016/j.advwatres.2008.07.007.
Neal, J., G. Schumann, and P. Bates (2012), A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas, *Water Resour. Res.*, *48*, W11506, doi:10.1029/2012WR012514.
Nikoo, M. R., R. Kerachian, A. Karimi, A. A. Azadnia, and K. Jafarzadegan (2014), Optimal water and waste load allocation in reservoir river systems: A case study, *Environ. Earth Sci.*, *71*(9), 4127–4142.
Raso, L., D. Schwanenberg, N. C. van de Giesen, and P. J. van Overloop (2014), Short-term optimal operation of water systems using ensemble forecasts, *Adv. Water Resour.*, *71*, 200–208.
Rodriguez, E. (2009), SWOT science requirements Document, JPL document, Initial release, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, Calif. [Available at http://swot.jpl.nasa.gov/files/SWOT_science_reqs_final.pdf.]
Sakamoto, T., A. A. Gitelson, and T. J. Arkebauer, (2014), Near real-time prediction of US corn yields based on time-series MODIS data, *Remote Sens. Environ.*, *147*, 219–231.
Sheffield, J., G. Goteti, and E. F. Wood (2006), Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, *J. Clim.*, *19*, 3088–3111.
Yoon, Y., M. Durand, C. J. Merry, E. A. Clark, K. M. Andreadis and D. E. Alsdorf (2012), Estimating river bathymetry from data assimilation of synthetic SWOT measurements, *J. Hydrol.*, *464–465*, 363–375, doi:10.1016/j.jhydrol.2012.07.028.
Zwarts, L., P. van Beukering, B. Kone, and E. Wymenga (2005), *The Niger, A Lifeline, Effective Water Management in the Upper Niger Basin*, 169 pp., RIZA, Lelystad.