lotus

D6.4: Hydrological-hydrodynamic modelling and data assimilation system

Author list

Author	Affiliation
Henrik Madsen	DHI, Denmark
Marc Ridler	DHI, Denmark
Peter Nygaard	DHI, Denmark
Raphael Schneider	DTU Environment, Denmark
Peter Bauer-Gottwein	DTU Environment, Denmark

Acronyms

AR1	First order autoregressive model
CRPS	Continuous ranked probability score
DA	Data assimilation
EnKF	Ensemble Kalman filter
KF	Kalman filter
LRM	Low resolution mode
MKL	Intel Math Kernel Library
NDVI	Normalized difference vegetation index
RMSE	Root mean square error
RR	Rainfall runoff
SAR	Synthetic aperture radar
SARIN	SAR-Interferometric mode
TRMM	The Tropical Rainfall Measurement Mission

1 Table of Contents

2 Executive summary

This report has been prepared as part of the project 'Preparing Land and Ocean Take Up from Sentinel-3 (LOTUS)' Work Package 6 'Applications of new GMES data in value-adding land services', Deliverable 6.4.

A catchment-scale hydrological-hydrodynamic modelling and data assimilation approach has been developed for assimilation of river water level measurements obtained from satellite altimetry data. The approach developed is based on the MIKE 11 hydrological-hydrodynamic modelling system and the general-purpose DHI Data Assimilation library. The data assimilation implementation has been tailored to assimilation of drifting-orbit altimetry data such as Cryosat-2.

The report describes the MIKE 11 hydrological-hydrodynamic modelling system, the DHI Data Assimilation library, and implementation of the library in MIKE 11 for assimilation of Cryosat-2 altimetry data. To demonstrate and verify the modelling and data assimilation system developed two assimilation tests were performed, respectively, (i) assimilation of synthetic drifting-orbit altimetry (Cryosat-2 like) water level measurements in the MIKE 11 hydrodynamic model, and (ii) assimilation of discharge measurements in the MIKE 11 rainfall-runoff model. Both tests successfully verified the data assimilation system implemented in MIKE 11 and demonstrated its value for improving hydrological-hydrodynamic model predictions.

Application of the MIKE 11 modelling and data assimilation system for assimilation of Cryosat-2 data into a hydrological-hydrodynamic model of the Brahmaputra river basin is described in Deliverable 6.5.

3 Introduction

Hydrological-hydrodynamic modelling is a key component in decision support systems for water resources management. Assimilation of data in hydrological-hydrodynamic models can significantly improve model performance and predictive capability. The immensely increasing availability of water and environmental data from different data sources has offered new opportunities for data assimilation (DA). New, efficient multivariate DA methodologies are required to fully utilize and optimally combine these data sources in hydrological-hydrodynamic modelling. This will form the basis for provision of new and improved services within on-line monitoring of water environments and real-time forecasting and early warning.

In the LOTUS project a catchment-scale hydrological-hydrodynamic modelling approach has been developed for assimilation of river water level from satellite altimetry data. The modelling approach is based on DHI's hydrological-hydrodynamic modelling system MIKE 11 (MIKE by DHI, 2014). Data assimilation capabilities have been implemented in the MIKE 11 model using a general-purpose DA framework developed by DHI, which has been tailored to assimilation of drifting-orbit (CryoSat-type) altimetry data.

In the following is described the hydrological-hydrodynamic modelling approach, and the DA framework and its implementation in MIKE 11 for assimilation of Cryosat-2 altimetry data. To demonstrate and verify the MIKE 11 DA modelling system developed results of different assimilation tests are presented. Application of the MIKE 11 DA modelling system for assimilation of Cryosat-2 data into a hydrological-hydrodynamic model of the Brahmaputra river basin is described in Deliverable 6.5.

4 Hydrological-hydrodynamic modelling approach

The MIKE 11 hydrological-hydrodynamic modelling system has been applied to a number of river basins around the world, representing many different hydrological and hydraulic regimes and climatic conditions. In the following is given a brief description of the rainfall-runoff and hydrodynamic model components of MIKE 11 that are used in the project. For a more detailed description see MIKE by DHI (2014).

4.1 Rainfall-runoff model

The hydrological model used in this study is the NAM rainfall-runoff model that was originally developed at the Institute of Hydrodynamics and Hydraulic Engineering at the Technical University of Denmark (Nielsen and Hansen, 1973). The NAM model simulates the rainfall-runoff processes occurring at the catchment scale. NAM forms part of the rainfall-runoff (RR) module of the MIKE 11 modelling system. The RR module describes the rainfall-runoff processes in one or more contributing catchments that generate lateral inflows to a river network. In this manner it is possible to treat a single catchment or a large river basin containing numerous catchments and a complex network of rivers and channels within the same modelling framework.

NAM is a lumped, conceptual model that consists of a set of linked mathematical equations describing in a simplified form the behaviour of the land phase of the hydrological cycle. NAM represents various components of the rainfall-runoff process by continuously accounting for the water content in four interrelated storages that represent different physical elements of the catchment. These storages are:

- Snow storage
- Surface storage
- Lower or root zone storage
- Groundwater storage

In addition, NAM can include man-made interventions in the hydrological cycle such as irrigation and groundwater pumping. The NAM model structure is illustrated in Figure 1.

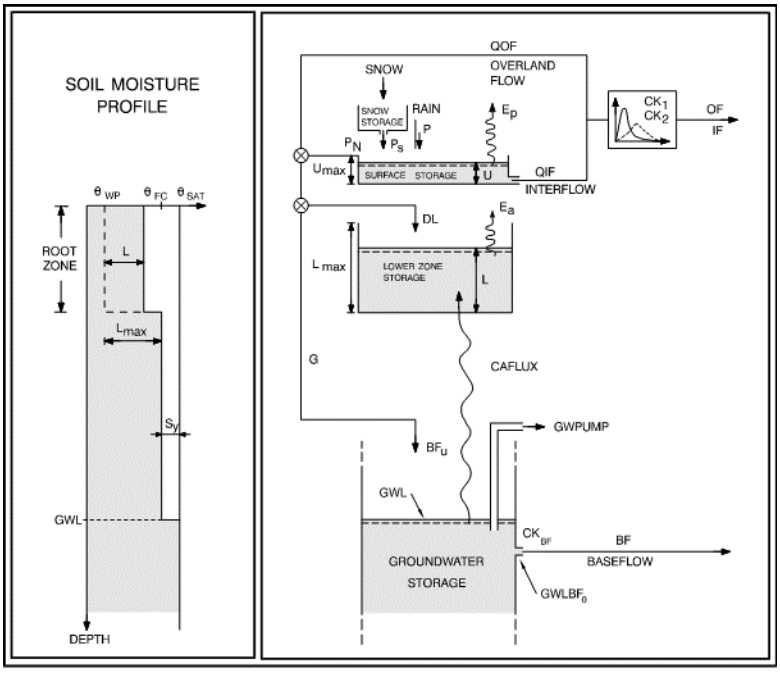

Figure 1: NAM rainfall-runoff model structure.

NAM uses as input meteorological data in terms of time series of rainfall, potential evapotranspiration and temperature (if the snow module is included). If irrigation or groundwater pumping are included in the model, additional input time series of irrigation amounts and groundwater abstraction rates are required. Since NAM is a lumped model, input time series represent catchment average values.

Based on the input data NAM produces catchment runoff as well as information about other elements of the land phase of the hydrological cycle, such as actual evapotranspiration, soil moisture content in the root zone, groundwater recharge, and groundwater storage. The resulting catchment runoff is split conceptually into overland flow, interflow and baseflow components (see Figure 1).

The NAM model includes a number of conceptual parameters that reflect the physical characteristics of the catchment. These parameters cannot be measured directly from measurable quantities of catchment characteristics, and therefore model calibration is needed. In the model calibration, parameters are tuned so that simulated catchment runoff matches the observed runoff as closely as possible. MIKE 11 includes an automatic calibration procedure for estimation of the most important NAM model parameters for the catchments where discharge measurements are available. Typically, not all catchments defined in the MIKE 11 model are gauged, and in this case estimated parameters from calibrated catchments are transferred to ungauged catchments.

4.2 River model

The NAM catchment models defined for the river basin generate input in terms of lateral inflow to the MIKE 11 river model. MIKE 11 is a 1D flow model for simulating rivers and surface water bodies that can be approximated as 1-dimensional flow. The model can describe sub-critical as well as super critical flow conditions through a numerical scheme, which adapts according to the local flow conditions (in time and space). Computational modules are included for the description of flow over hydraulic structures, comprising possibilities to describe structure operation.

The MIKE 11 hydrodynamic module is based on an implicit, finite difference solution of the 1D St Venant equations. Three different solutions are provided:

1. Dynamic wave approach, which uses the full momentum equation, including acceleration forces, thus allowing the simulation of fast transient flows, tidal flows, and backwater flows.
2. Diffusive wave approach, which only models the bed friction, gravity force, and the hydrostatic gradient terms in the momentum equation. This allows to take downstream boundary conditions into account, and thus simulate backwater effects. It is normally not suitable for tidal flows.
3. Kinematic wave approach, which is based on a balance between the friction and gravity forces. This description is appropriate for modelling relatively steep rivers without backwater effects.

Depending on the type of problem, the user can choose the most appropriate solution. All three approaches simulate branched as well as looped networks.

In addition to the solutions of the governing 1D St Venant equations, MIKE 11 includes hydrological routing descriptions. The implemented methods are:

1. Muskingum method
2. Muskingum-Cunge method

Hydrological routing can be used when less detailed hydrodynamic solutions are needed, and no hydraulic structures are included. Larger time steps can be used to facilitate long-term simulations. It is possible to combine hydrological routing for upstream branches with a St Venant solution for the main river.

For setting up MIKE 11, the river network, geometry of cross-sections of the river channels and flood plains, channel and flood plain roughness, and geometry and hydraulic properties of structures (e.g. weirs, culverts) are specified. The channel roughness is typically calibrated against measured water levels and discharges in the river system.

An example of a MIKE 11 setup with NAM rainfall-runoff catchments defined is shown in Figure 2.

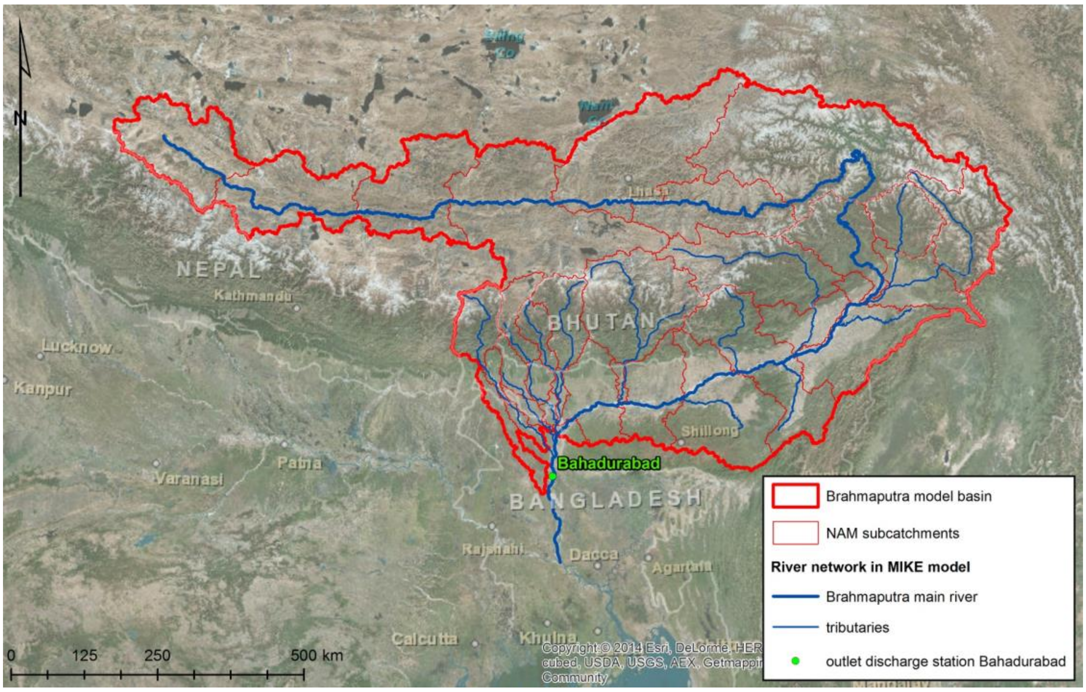

Figure 2: Example of MIKE 11 setup for the Brahmaputra river basin.

5 Data assimilation system

5.1 Data assimilation approach

In general, DA refers to the process of combining model predictions with observations in order to obtain a more accurate estimate of the state of the system. DA is an important part of operational forecasting systems to improve the initialization of the forecast model at time of forecast and thereby improve forecast accuracy.

DA is a feedback process where the model prediction is conditioned to the observations of the modelled system. DA procedures can be classified according to the parts of the forecast system that are modified in the feedback process, see Figure 3. These include updating of the model forcing (input), model states, model parameters, and model outputs. We consider here the DA problem within a general filtering framework. By using this generic formulation different approaches for updating model forcing, state variables and model parameters can be incorporated (corresponding to approach 1-3 in Figure 3). In addition, the filtering framework can be combined with error forecasting in measurement points (approach 4 in Figure 3; Madsen and Skotner, 2005).

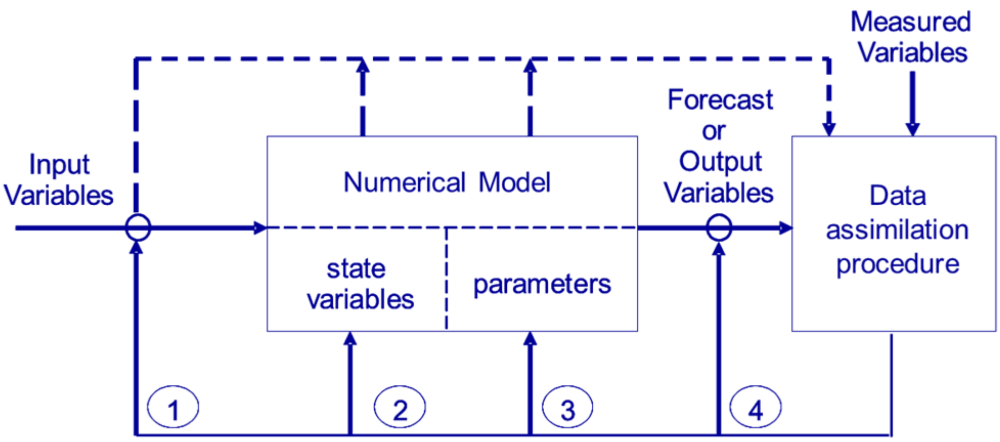

Figure 3: Classification of DA approaches, considering updating of (1) Model forcing, (2) Model states, (3) Model parameters, and (4) Model output.

The filtering framework is based on a sequential predictor-corrector scheme where the model prediction is corrected based on observations. The prediction of the system state can be written

$$x_k = \Phi(x_{k-1}, u_k, \theta_k) \tag{1}$$

where $\Phi(.)$ is the model operator, x_k is the state vector representing the state of the modelled system at time step k, u_k is the forcing of the system, and θ_k represents the model parameters. The observations of the system are related to the system state by

$$z_k = H_k x_k \tag{2}$$

where z_k is the measurement vector, and H_k is a matrix that describes the relation between measurements and state variables (i.e. a mapping of state space to measurement space). It is here assumed that only direct measurements of state variables are available, but, in general, any measurement that can be mapped onto the state space using a linear or non-linear transformation can be used.

The predicted model state and measurements can now be combined using a linear combination

$$x_k^a = x_k^f + G_k(z_k - H_k x_k^f) \tag{3}$$

where x_k^f is the model predicted (forecasted) state, x_k^a is the updated state, and G_k is a weighting matrix. The weighting matrix reflects the relative importance of the model forecast and the measurements, respectively, and the influence of corrections of the state variables at measurement positions to corrections of all state variables in the entire modelling domain.

The formulation of the weighting matrix is the most essential part of the filtering scheme, and the different schemes mainly differ from each other in the way this matrix is calculated. The most comprehensive filtering scheme is the Kalman filter (KF) where the weighting matrix (denoted the Kalman gain) is determined based on a least squares minimization of the expected error of the updated state in terms of the errors of both model state and measurements. The main strength of the KF is that it explicitly takes model and measurement uncertainties into account in the updating process and provides an estimate of the uncertainty of the system state. If the model is linear, and the model and measurement errors are independent white noise errors with known covariance matrices, the KF provides the best (with respect to minimum prediction error variance) linear unbiased estimator cf. Eq. (3).

Since hydrological-hydrodynamic forecast systems are, in general, non-linear with complex and possibly biased error structures, the optimality of the KF cannot be guaranteed. In addition, when applied to high-dimensional modelling systems propagation of the covariance matrix may be computationally infeasible. The ensemble Kalman filter (EnKF) that was introduced by Evensen (1994) provides an efficient implementation of the KF for non-linear, high-dimensional systems and has been successfully applied within different modelling disciplines, e.g. numerical weather prediction, operational oceanography, hydrological-hydrodynamic forecasting, and air quality forecasting. In the EnKF the state covariance matrix is represented by an ensemble of size M of possible states that are propagated according to the dynamics of the system. The computational efforts required by the EnKF correspond to approximately M model integrations. A serious disadvantage of the method is that the statistical error decreases very slowly with increasing ensemble size (proportional to $M^{-1/2}$). Different variants of the EnKF have been proposed to optimise computational efforts and improve covariance sampling and propagation, e.g. the Ensemble Transform KF (Hunt et al., 2007) and Deterministic Ensemble KF (Sakov and Oke, 2008).

The state updating scheme in Eq. (3) can be extended to include also update of model forcing and model parameters. The uncertainty in model forcing can be described as coloured noise using a first order autoregressive process. To take the coloured model noise into account an augmented state vector is defined that contains the model state and the forcing error. In this way the forcing error is updated along with the model state itself and enables correction of biased model forcing (see e.g. Sørensen et al., 2006). With a similar approach parameter estimation can be included in the KF by augmenting the state vector with model parameters (e.g. Rasmussen et al., 2015). The augmented state vector approach can also be used to include estimation and correction of biases in the state or in the observations (e.g. Drecourt et al., 2006).

5.2 DHI data assimilation library

The DHI data assimilation library is a set of generic assimilation filters, noise models, observation mapping methods and result analysis tools. It is designed in a modular fashion using object-oriented best practices with interfaces defining the boundaries of each module. The library is coded in C# using the .NET 4.0 Framework. The matrix equations are solved using the system optimized Intel Math Kernel Library (MKL).

The library currently supports different ensemble-based KF algorithms:

- Classical EnKF
- Ensemble Transform KF
- Deterministic Ensemble KF

It includes procedures for localisation, joint state, parameter and model noise estimation, and bias-aware filtering. Furthermore, it supports use of different stochastic error models to describe model and measurement errors.

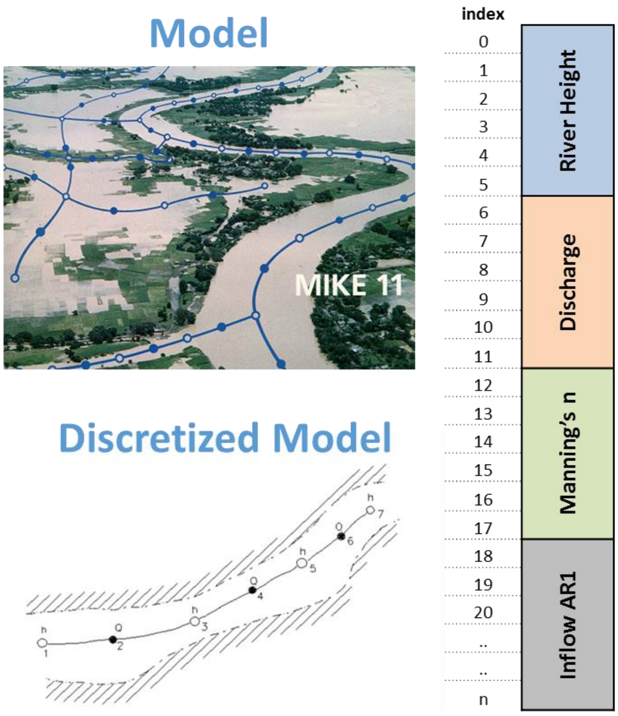

Figure 4: Model discretization and definition of state vector.

A key concept in DA is the "state vector". Once a model is discretized in space (see Figure 4), it is constituted by a number of variables (states, parameters, forcing errors, or biases) of different sizes. With the DA library, we select which variables we are interested in, for example a) river height, b) discharge, c) Manning's n parameter, and d) inflow forcing AR1 error. A state vector is then defined consisting of these variables and arranged into a long vector (of length *n*). Whatever variables are included in the state vector will be updated and corrected during assimilation.

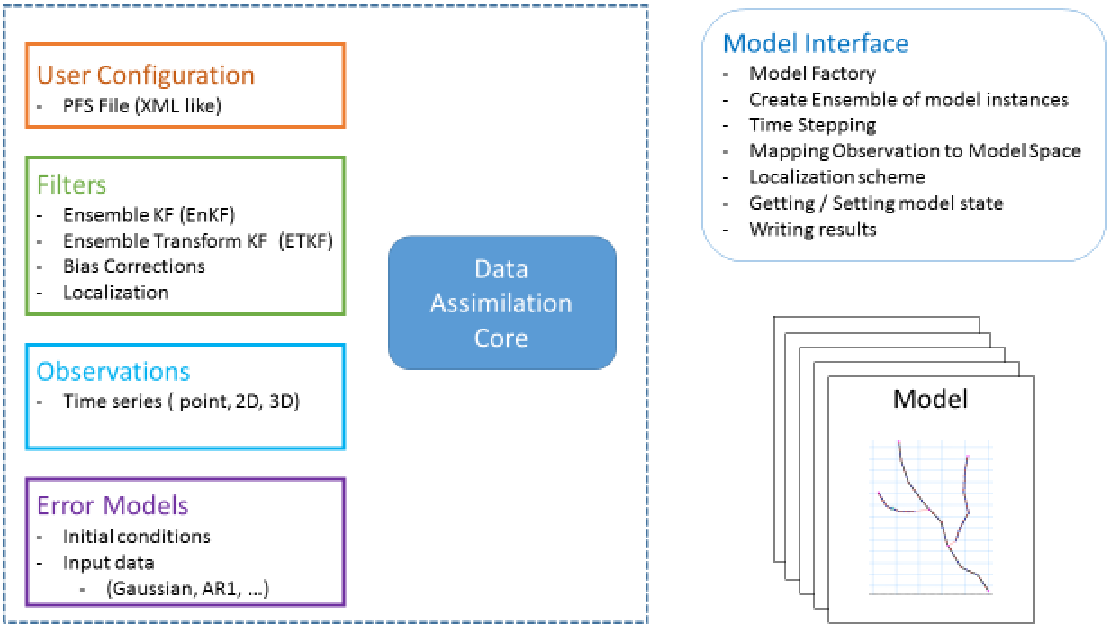

Figure 5: Overview of DA modules (left), and the Model Interfaces (right) required to connect a model to the DA library.

The DA library contains five main modules (Figure 5):

1. User Configuration. Reads a PFS File (XML like) containing details of the assimilation experiment such as, ensemble size, location of the model on disk, location of the observations, type of filter, localization details, noise models, and which variables should be included in the state vector.
2. Filter. Takes abstracted vectors and matrices from the Core and solves the KF equations to calculate the optimal correction to the ensemble of models. The system of equations are solved using the Intel MKL.
3. Observation. Handles the observation variable and how it relates to the model. This module determines based on a measurement location, its corresponding index in the state vector. Different types of observations are grouped together in an ObservationCollection Class.
4. Error Models. Some generic perturbation algorithms for adding uncertainty to the model forcing, initial conditions, state, and observations.
5. Core. Communicate and controls all the modules.

In order to connect a new model to the assimilation library, the Model Interface (Figure 5 right) must be implemented. These interfaces define how the model is created, controlled in time, and variables within the model accessed. The interfaces must be implemented in C# using .NET 4.0.

The DA library performs a number of tasks:

1. Reads a configuration (pfs) file to set up the assimilation system.
2. Creates an ensemble of model instances.
3. Reads observation files and collects them in an ObservationCollection class.
4. Maps the observations to the model indices (constructs the H Matrix cf. Eq. (2))
5. Time steps the ensemble to the time of observation.

6. The Core reads model values and observation values to create the matrices for assimilation.
7. The Filter is called with access to the matrices. The model updates are calculated.
8. The ensemble is updated based on the Filter's results.
9. Steps 5-8 are performed until there are no more observations or the models finish.

5.3 Implementation of data assimilation in MIKE 11

MIKE 11 connects to the DA Library by having implemented the defined interfaces (Figure 5 right). The implementation is designed for speed and flexibility. The ensemble of models is created (using the Model Factory) from a single model instance, and each ensemble member is perturbed during simulation within the MIKE 11 model engine. In this way, the ensemble does not need to be pre-created and stored in separate directories. This speeds up initialization and simplifies testing different perturbation schemes in order to ensure the most realistic uncertainty description. The DA implementation supports update of discharge and water level states in the hydrodynamic model and the internal model states of the NAM rainfall-runoff model.

The implementation is further extended to accommodate assimilation of altimetry data. First, as opposed to traditional in-situ water level or discharge observations, the measurements come intermittently during satellite overpasses, and vary in location. This means that the assimilation library advances in time according to the availability of measurements instead of assimilating every model time step. Second, the along-track measurements are grouped together and linearly interpolated between the two nearest river calculation points (Figure 6). The grouping and projection of the Cryosat-2 data used in LOTUS are described below.

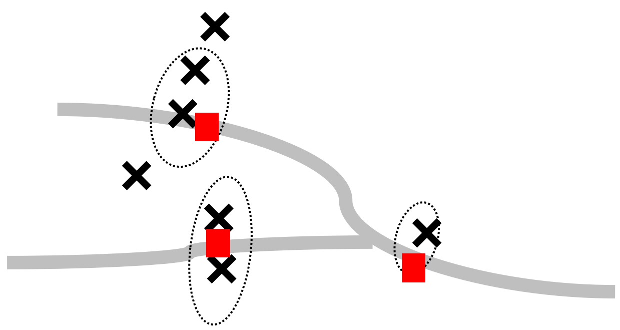

Figure 6: Grouping of altimetry measurements and projection onto the river network.

The implementation provides the option to extend the one measurement to multiple virtual ones in order to increase the impact of each satellite by-pass. The measurement can be extended in time as shown in Figure 7 by increasing the measurement uncertainty.

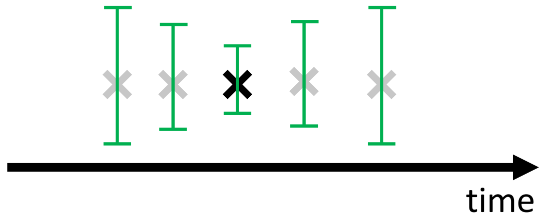

Figure 7: A measurement is extended in time to create additional virtual measurements to increase the impact of each measurement. The measurement (black cross) can be a single measurement or a grouped average. The green bars represent the measurement uncertainty.

Uncertainty in the MIKE 11 hydrodynamic model can stem from the RR inflows into the river. The implementation of the MIKE 11 DA allows perturbation of the RR inflows using a noise model in order to create an ensemble of realizations. The noise model can include both spatially and temporally correlated noise. In the case of a first order autoregressive noise model, model errors can be appended to the state vector (state augmentation) such that the KF corrects the error model during the filter update.

Update of hydrological model states is supported through the implementation of the NAM rainfall runoff model in the DA framework. The NAM model states consist of a number of storages representing the different components of the catchment processes (see Figure 1). The implementation supports the update of 6 states:

- Surface storage
- First overland flow reservoir
- Second overland flow reservoir
- Inter flow first reservoir
- Root zone storage
- Groundwater storage

The states are included in the state vector of the filter and are updated with water level and discharge states of the hydrodynamic model when river water level and discharge measurements are available.

Model uncertainty in the NAM model can be introduced either through state or forcing perturbation using appropriate perturbation methods available in the DA library. Precipitation is the only forcing that can be perturbed at the moment as it was considered the dominating source of uncertainty, but perturbation of temperature and potential evapotranspiration could be included. The following types of noise models are available:

- first order autoregressive normal distribution
- first order autoregressive truncated normal distribution (interval]-1; 1[)
- first order autoregressive log-normal distribution

To account for correlated errors in time and space the methods are all implemented as autoregressive processes, and spatial correlation is included by specifying a spatial correlation matrix. For state perturbation only the first order autoregressive normal distribution method is available, and state errors are considered spatially uncorrelated.

5.4 Filtering and projecting CryoSat-2 data for data assimilation

For DA of water level measurements, CryoSat-2 level 2 data provided by DTU Space within the LOTUS project has been used. The basis for the data is the ESA baseline-b L1b 20 Hz product with retracking applied by Villadsen et al. (2015). Using this data for river altimetry is most attractive over areas that are covered by the high-resolution SARIn and SAR mode of CryoSat-2 with an along-track resolution of about 300 m (see Figure 8). LRM mode only provides an along-track resolution of around 7 km (Wingham et al., 2006). For areas covered in the SARIn mode, an off-nadir correction can be applied to the data to determine the exact ground location of the measurement. Using an estimate of the true location of the ground reflection of the altimeter data instead of the nadir as in SAR mode can increase the amount of usable data. Especially over typically narrow shapes as rivers, this off-nadir correction has shown to be of high importance.

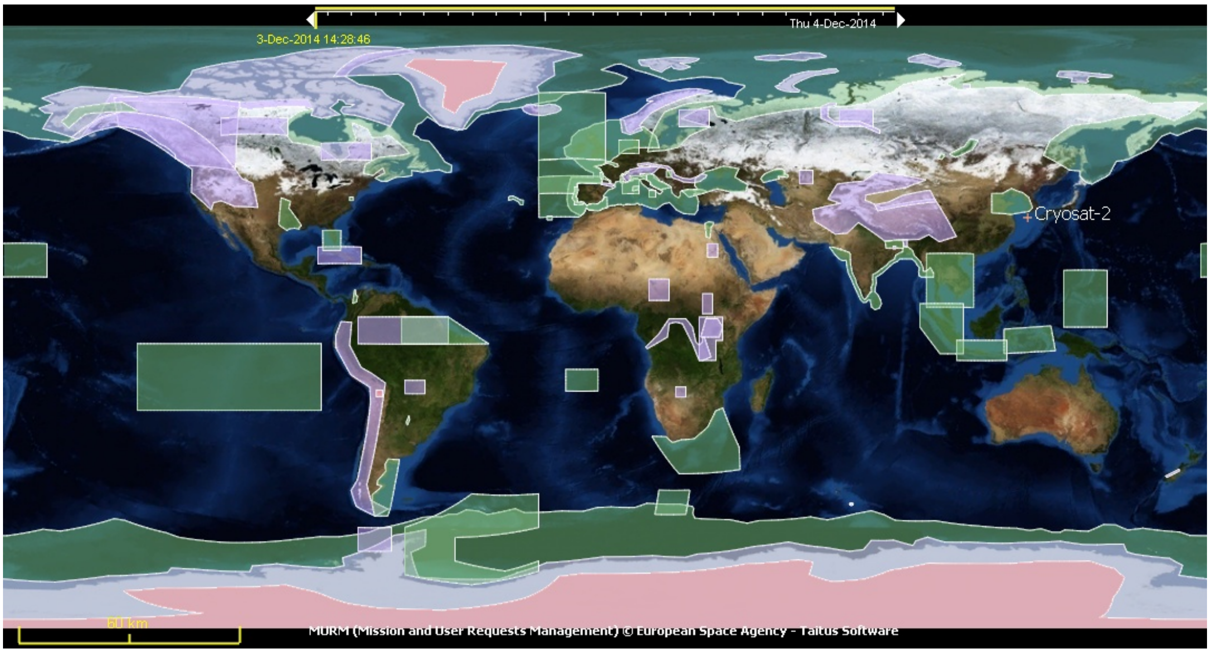

Figure 8: Geographical mode mask version 3.6 for CryoSat-2. SARIn mode: purple. SAR mode: green. LRM mode: red and all other areas. From https://earth.esa.int/web/guest/-/geographical-mode-mask-7107

CryoSat-2 has a drifting orbit, completing a full cycle every 369 days. This high spatial resolution of the data, combined with the fact that the CryoSat-2 data itself does not provide any reliable indicator whether it was acquired over water or land surface require a high-resolution water mask to filter out relevant data points that represent river water level measurements. In this study, a river mask based on Landsat-7 and 8 NDVI imagery with 30 m resolution was chosen. An alternative to optical imagery is the use of SAR imagery; however, SAR imagery with global coverage is only freely available since the start of the Sentinel-1 mission in 2014. SAR imagery can be acquired independent of cloud coverage; an advantage over any optical imagery. Dependent on the dynamics of the river system, a river mask only stays valid for a certain period of time. In the case of the Brahmaputra in the Assam valley, the braided river bed is very dynamic and experiences relevant changes each flood season. Hence, an individual river mask for each year has to be extracted.

Figure 9 shows an example of the filtering and projecting process applied to the CryoSat-2 level 2 data points over a section of the Brahmaputra river. Only data points above the river mask for the respective year are used, i.e. considered to represent the river water surface. In our case, a 1D hydrodynamic river model was used (the model river representation can be seen in Figure 9), and in this case the CryoSat-2 observation points have to be projected onto the river line to obtain the observation locations in model coordinates.

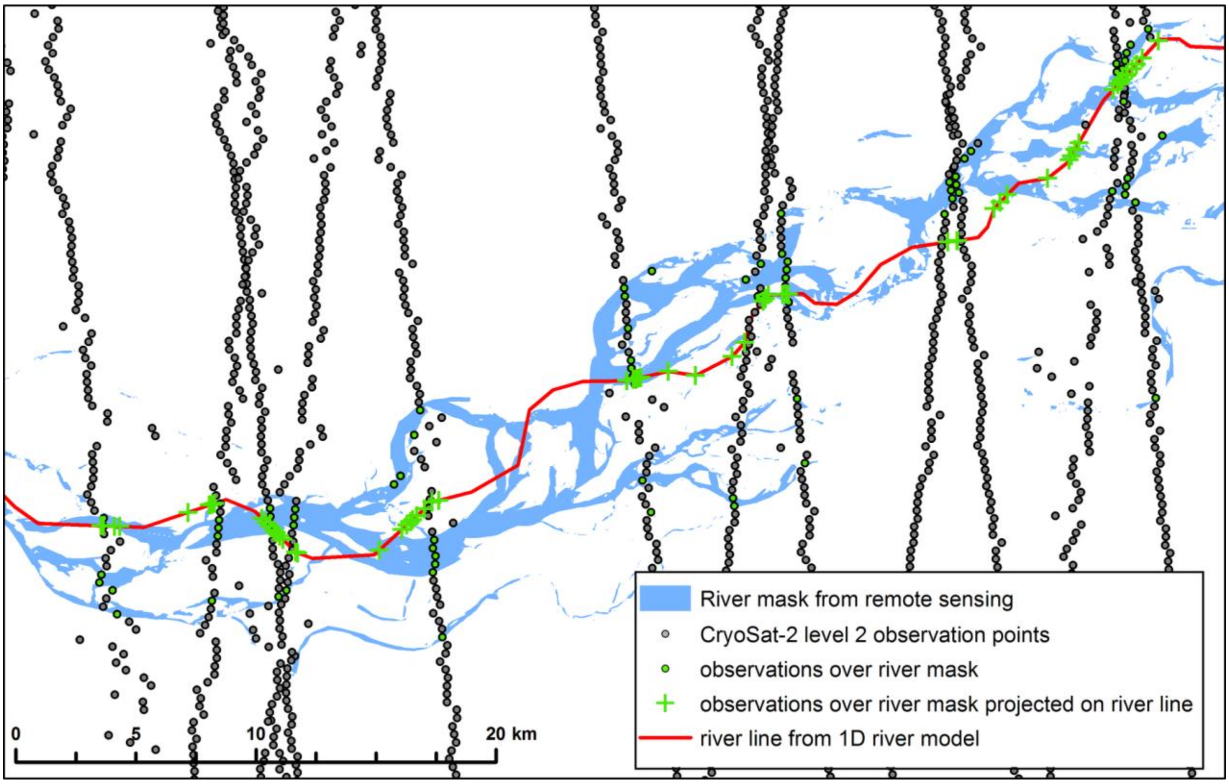

Figure 9: Exemplary section of the Brahmaputra in the Assam valley showing the river mask, the CryoSat-2 observations and their mapping to the 1D river model, all for the year 2013.

In Figure 9, 12 transects of CryoSat-2 crossing the Brahmaputra river can be seen, each with several observations over the river mask. The data of one river transect is – at least for hydrologic modelling – obtained at the exact same time. Assimilating multiple observations at the same model time step at (almost) the same location in the model does not add additional information. Hence, for DA purposes, it was decided to not use all single observations, but to aggregate observation points of one transect, using mean location and elevation of the single observations. This also makes the process more robust to observation errors and outliers. The most simple approach is to aggregate all points of one transect into one observation. This, however, only works well if the satellite transect is more or less perpendicular to the river as in Figure 9. Otherwise, one transect can span a long section along the river, and aggregating it into one observation might reduce information that can be obtained from the observations. Such cases are illustrated in Figure 10.

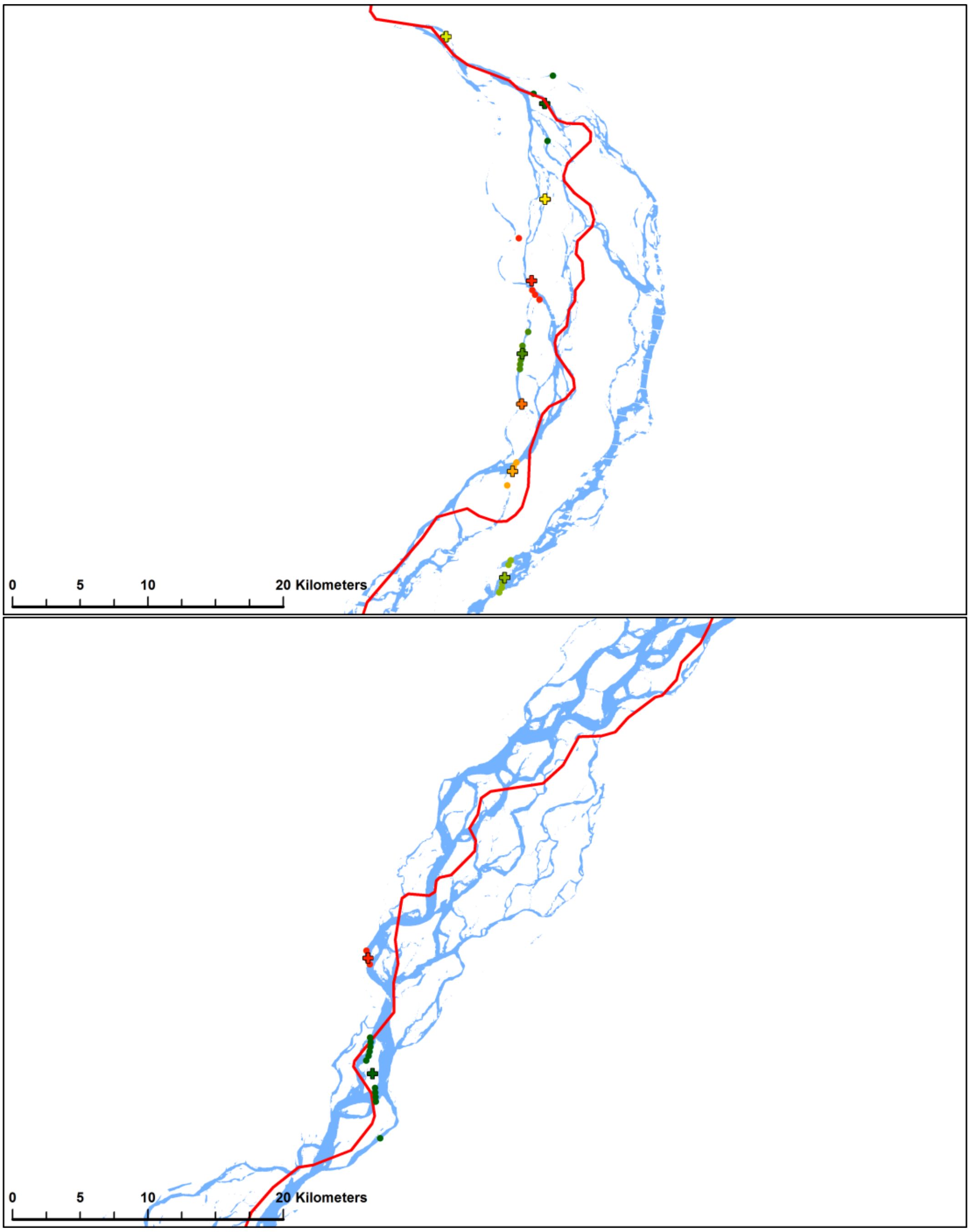

Figure 10: Exemplary result of k-means clustering of CryoSat-2 observation points along the Brahmaputra. Original points over the river mask are displayed as dots; the cluster means as crosses.

In Figure 10, all the dots in each plot represent the water observation points of one transect of CryoSat-2. It can be seen that they are spread over a long distance along the river. For a meaningful automated aggregation of observation points, a k-means clustering was applied. Each transect's river points were considered separately. First, the total distance spanned by the river points of the respective transect was determined. Then, the number of clusters was determined as the total distance divided by a defined maximum distance of points within one cluster. In the example in Figure 10 the maximum distance was chosen to 5 km. The resulting clusters are indicated by the different colours of the dots in Figure 10. The crosses represent the average location of each cluster's points. These average locations per cluster, together with their respective average elevation are the data that finally are used for assimilation in the hydrological-hydrodynamic model.

6 Data assimilation experiments

For verification and demonstration of the DA procedure implemented in MIKE 11 different tests have been performed, considering (i) assimilation of water level measurements in the hydrodynamic model using synthetic satellite track data, and (ii) assimilation of discharge measurements in the NAM rainfall-runoff model for one of the catchments in the Brahmaputra basin.

6.1 Assimilation of water level measurements in river model

For testing the assimilation of altimetry measurements in the MIKE 11 hydrodynamic model, a small synthetic test model was created. An overview of the model setup is shown in Figure 11. The synthetic test model consists of two branches, with runoff generated by two NAM catchments that each are connected to upstream river branches. Two simulations were performed

- a base run with original precipitation forcing on the NAM models; and
- a synthetic truth run with changed precipitation forcing.

The synthetic truth run is considered to represent the unknown truth from which observations are extracted. When assimilating these observations to the base run model, it is expected that the model will be corrected towards the synthetic truth run. To investigate the effect of observations at different points in the model, the observations were extracted at the points indicated in Figure 11. In reality, satellite altimeter measurements will be distributed over the entire model domain as well. The results were mainly analyzed in terms of discharge at the outlet of the model (red cross in Figure 11). This also reflects real case scenarios, as for example the Brahmaputra model (see Deliverable 6.5) where one is interested in discharge forecast at some (downstream) point of the model.

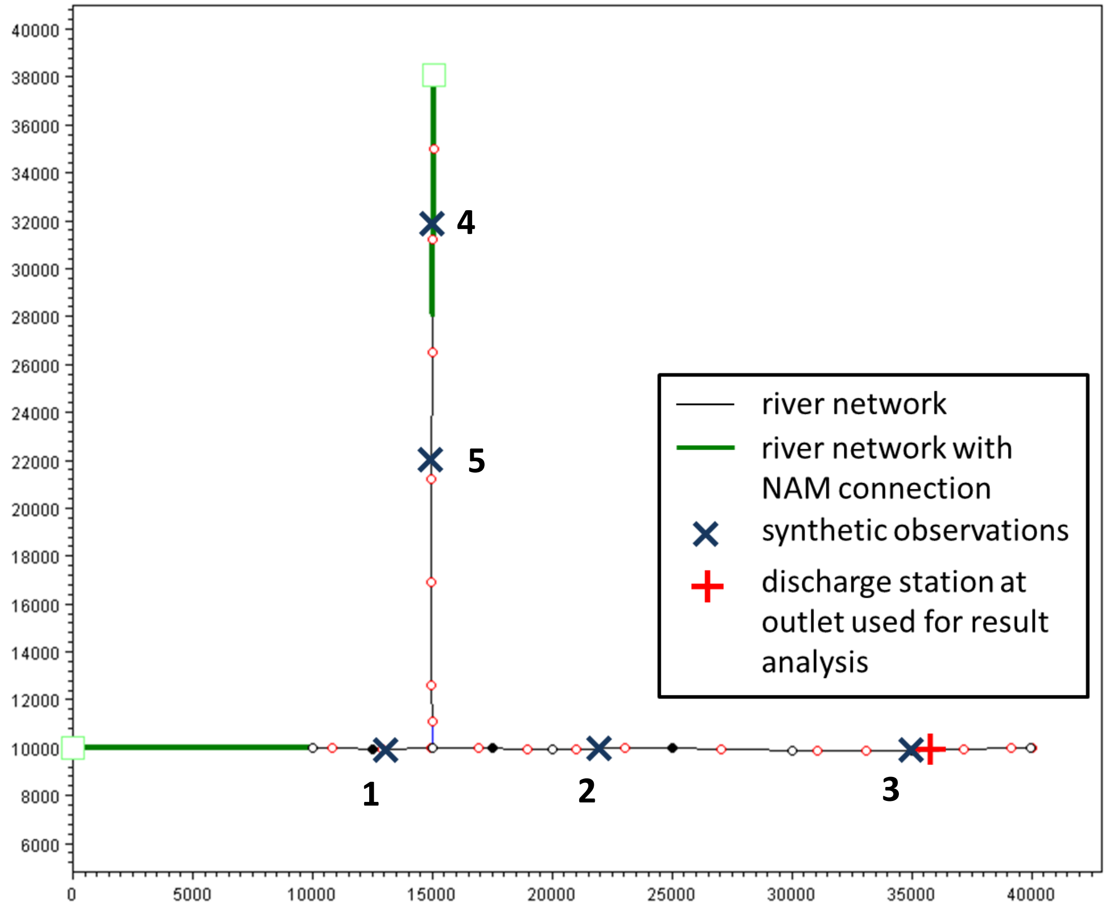

Figure 11: Sketch of the synthetic MIKE 11 hydrodynamic test model. Flow direction is towards the discharge station in the lower right corner. Coordinates are given in metres.

Using ensemble filters one has to describe the model error, usually by creating an ensemble of runs with perturbations of forcings, parameters or states. The model error was considered to be dominated by the runoff error from the NAM catchments, hence the ensemble spread was generated by perturbing these runoffs. To generate a realistic description of model error, especially for larger real-world models with many subcatchments, a spatio-temporal error correlation model was applied. The runoff perturbation was modelled with a relative autoregressive AR1 Gaussian noise model with cross-correlation between the catchments' errors.

The model simulation period was from 01/09/1974 to 30/12/1974. Observations to be assimilated were extracted every five days, at points in the model as indicated in Figure 11 and Table 1.

Table 1: Times and positions (compared to the numbers given in Figure 11) of synthetic observations that were assimilated.

Observation time	06/09	11/09	16/09	21/09	26/09	01/10	06/10	11/10	16/10	21/10	…	30/12
Observation point	1	2	3	4	5	1	2	3	4	5	…	4

The assimilation experiment was run using the Ensemble Transform KF. The filter was used without any additions, i.e. without inflation of the ensemble, damping of the updates, or localization. The most important filter parameters are summarized in Table 2.

Table 2: Filter parameters of the synthetic assimilation experiment.

Ensemble size	50
AR(1) coefficient of NAM runoff perturbation [-]	0.9885
Relative standard deviation of perturbation error [-]	0.2
Spatial correlation of perturbation error [-]	0.75
Observation uncertainty, standard deviation [m]	0.25

The results of the test run are shown in Figure 12. Discharge at the outlet of the model is shown for the base run, synthetic truth run, and all ensemble members. It can be seen that, in general, the assimilation main run is closer to the synthetic truth than the base run. Furthermore, each update reduces the ensemble spread, which represents the model prediction uncertainty.

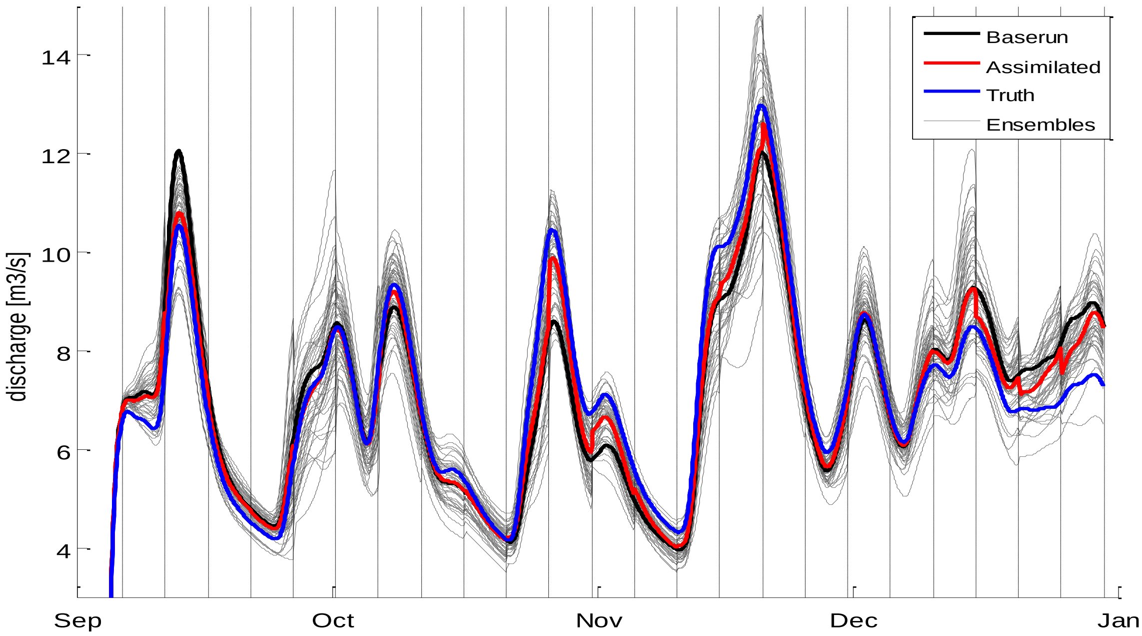

Figure 12: Results of the altimetry assimilation synthetic test run. Observation times are indicated by the vertical dashed lines.

Some performance indicators that show the improvement of the discharge forecast due to DA are displayed in Table 3. Here the results are compared to an open loop run, i.e. a run without DA, only applying the perturbation of the runoff forcing from the NAM models and otherwise unchanged setup. For all indicators we see an improvement for the run with DA over the run without DA. Sharpness and reliability are both given in relation to the 95% confidence intervals of the ensemble results, i.e. sharpness is the average width of the confidence interval, whilst reliability is the portion of synthetic truth values that fall within the confidence band.

Table 3: Performance indicators for synthetic data assimilation experiment in terms of discharge at the outlet.

	Open loop	With DA
RMSE between ensembles and synthetic truth [m^3/s]	0.7263	0.4820
CRPS of ensembles vs. synthetic truth [m^3/s]	0.3769	0.2566
Sharpness [m^3/s]	2.4440	2.0252
Reliability [-]	0.8972	0.9486

This shows that the developed DA framework is successfully able to assimilate distributed altimetry measurements into a 1D hydrodynamic river model for improving discharge forecasts.

6.2 Assimilation of discharge measurements in rainfall-runoff model

The Sankosh catchment is one of the few gauged sub-catchments in the Brahmaputra basin and has therefore been used for testing the DA of the hydrological model (see Figure 13). A large uncertainty is expected to stem from the precipitation as the model is forced by satellite-measured precipitation obtained from the Tropical Rainfall Measurement Mission (TRMM).

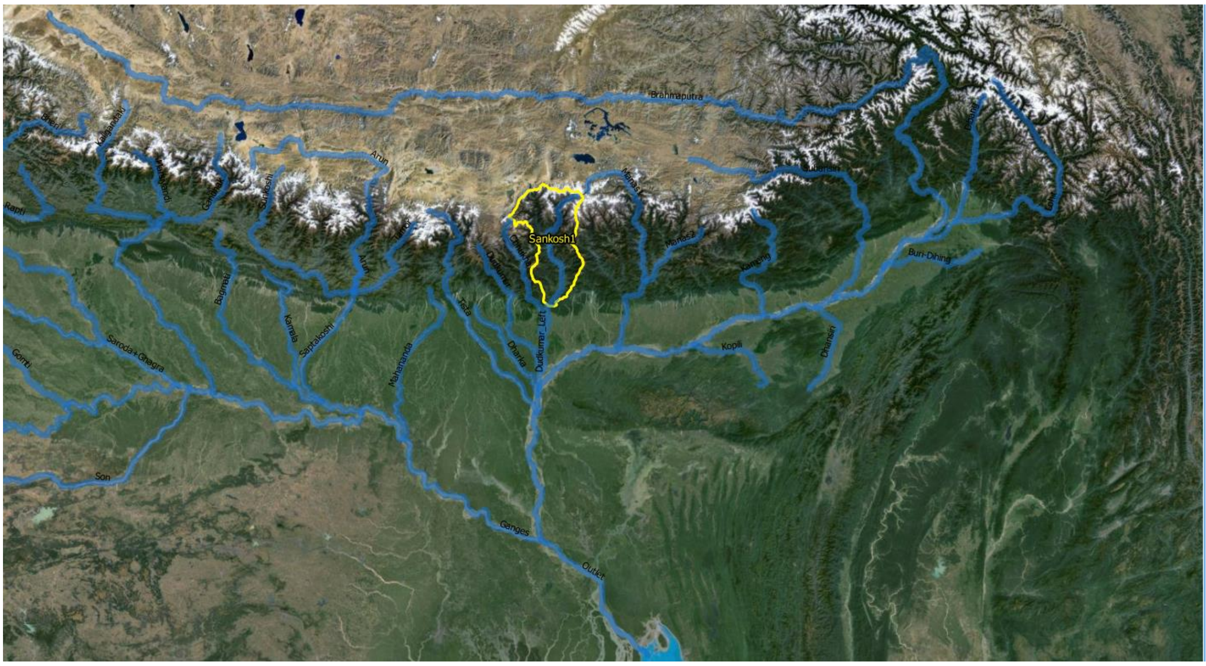

Figure 13: The Sankosh sub-catchment marked with yellow.

There are many degrees of freedom when specifying the model and observation errors for the KF, which make it a challenge to specify optimal error models. The error model parameters could be calibrated to find an optimum but in the below example they have been chosen by trial and error and serve for demonstration purpose. The precipitation forcing and the two model states Groundwater depth and Root zone storage were perturbed.

Daily discharge observations were available, but to assess the model performance after DA the state update only took place every third day, and the following two days where used for validation. Figure 14 shows the results of the DA over a 5-year period. The model states of the NAM model are shown with red and black lines for simulations with and without DA (Open loop), respectively. Figure 15 and Figure 16 show the same as the bottom graph in Figure 14 but zoomed in on a specific year. It is clear that the DA in these two examples most of the time improves the rainfall-runoff simulation after the state update by being closer to the observed runoff.

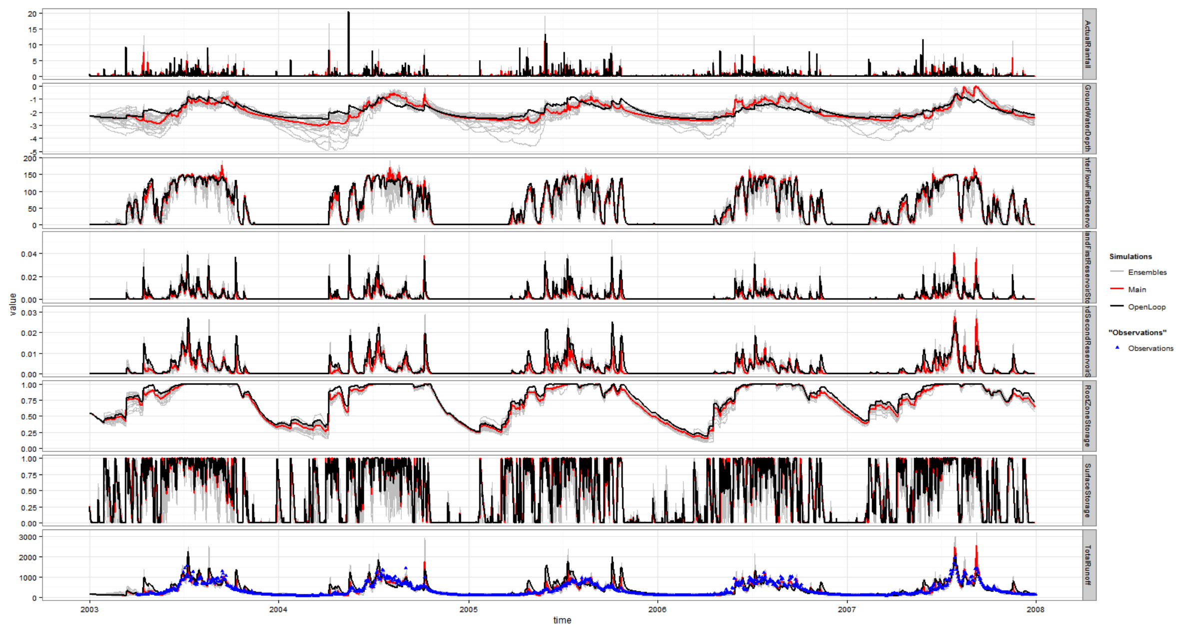

Figure 14: Data assimilation result. From top to bottom: Precipitation, Groundwater depth, First overland flow reservoir storage, Second overland flow reservoir storage, First interflow reservoir storage, Root Zone Storage, Surface storage, and Rainfall runoff. Grey: Ensemble, Red; Main model, Black: Open loop model, Blue: observations.

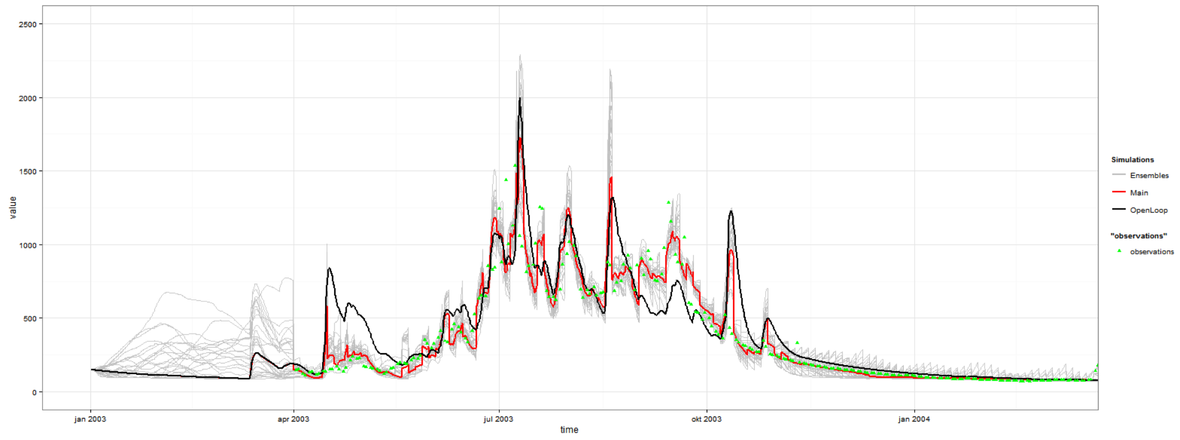

Figure 15: Rainfall-runoff simulation, 2003. Grey: Ensemble, Red: Main model, Black: Open loop model, Green: observations.

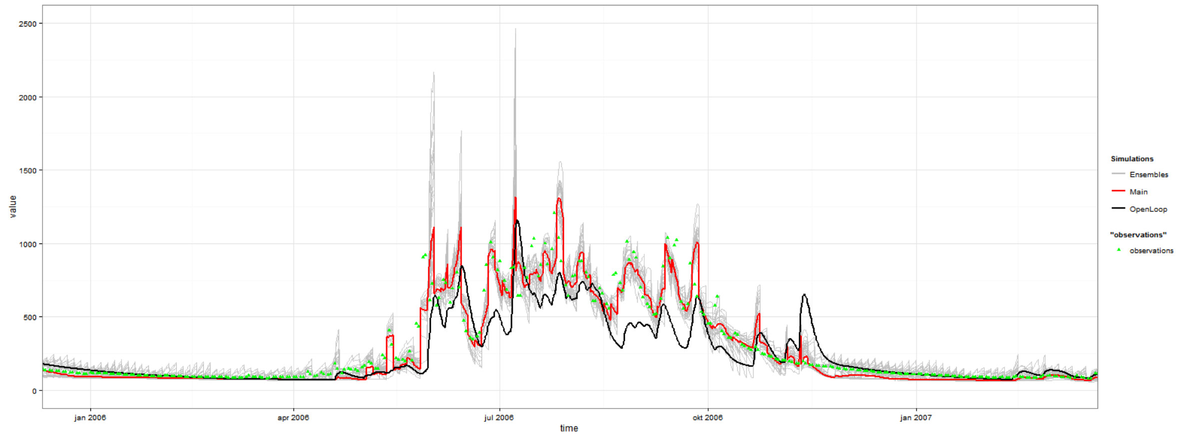

Figure 16: Rainfall runoff simulation, 2006. Grey: Ensemble, Red: Main model, Black: Open loop model, Green: observations.

7 Conclusions

A hydrological-hydrodynamic DA framework has been developed for assimilation of river water level measurements from satellite altimetry data. The framework is based on the MIKE 11 hydrological-hydrodynamic modelling system and the general-purpose DHI DA library, which has been tailored to assimilation of drifting-orbit altimetry data.

To demonstrate and verify the MIKE 11 DA modelling system developed two assimilation tests were performed, respectively, (i) assimilation of synthetic Cryosat-2 like water level measurements in the MIKE 11 hydrodynamic model, and (ii) assimilation of discharge measurements in the MIKE 11 NAM rainfall-runoff model. Both tests successfully verified the MIKE 11 DA system and demonstrated the value of DA for improving hydrological-hydrodynamic model predictions.

References

Drécourt, J.P., Madsen, H. and Rosbjerg, D., 2006. Bias aware Kalman filters: Comparison and improvements, Adv. Water Res., 29, 707-718.

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

Hunt, B.R., Kostelich, E.J., Szunyogh, I., 2007. Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter, Phys. D Nonlinear Phenom., 230(1–2), 112–126.

Madsen, H., and Skotner, C., 2005. Adaptive state updating in real-time river flow forecasting - A combined filtering and error forecasting procedure, J. Hydrol., 308, 302-312.

MIKE by DHI, 2014. MIKE 11 - A Modelling System for Rivers and Channels - Reference Manual.

Nielsen, S.A., Hansen, E., 1973. Numerical simulation of the rainfall runoff process on a daily basis. Nord. Hydrol., 4, 171–190.

Rasmussen, J., Madsen, H., Jensen, K.H., Refsgaard, J.C., 2015. Data assimilation in integrated hydrological modeling using ensemble Kalman filtering: evaluating the effect of ensemble size and localization on filter performance, Hydrol. Earth Syst. Sci., 19, 2999–3013.

Sakov, P., Oke, P.R., 2008. A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters, Tellus A, 60(2), 361–371.

Sørensen, J.V.T., Madsen H. and Madsen H., 2006. The parameter sensitivity of three Kalman filter schemes for the assimilation of tide gauge data in coastal and shelf sea models, Ocean Model., 11, 441- 463.

Villadsen, H., Andersen, O. B., Stenseng, L., Nielsen, K., & Knudsen, P., 2015. CryoSat-2 altimetry for river level monitoring - Evaluation in the Ganges-Brahmaputra basin. Remote Sensing of Environment, 168, 80–89.

Wingham, D. J., Francis, C. R., Baker, S., Bouzinac, C., Brockley, D., Cullen, R., … Wallis, D. W., 2006. CryoSat: A mission to determine the fluctuations in Earth’s land and marine ice fields. Advances in Space Research, 37(4), 841–871.