DATA ASSIMILATION IN A COMBINED 1D-2D FLOOD MODEL USING THE ENSEMBLE KALMAN FILTER

JOHAN N. HARTNACK, HENRIK MADSEN, JACOB V.T. SØRENSEN

DHI Water & Environment, Agern Allé 5, DK-2970 Hørsholm, Denmark

ABSTRACT

The combined 1D-2D numerical flood modelling system MIKE FLOOD is considered in a data assimilation setting. Data assimilation facilities have been implemented in MIKE FLOOD for assimilation of water level measurements. The data assimilation system is based on the ensemble Kalman filter methodology. In the ensemble Kalman Filter method the probability density of the model state is represented by an ensemble of model states. In a model forecast each ensemble member is propagated according to the dynamical system subject to a stochastic element in the form of model errors. The resulting ensemble provides estimates of the forecast state and the corresponding variance. When measurements are available, the ensemble may be updated using the standard Kalman filter updating scheme. This paper discusses the implementation of the ensemble Kalman Filter in MIKE FLOOD and demonstrates its use for updating model results in a flooding application.

1. INTRODUCTION

The MIKE FLOOD model developed by DHI Water & Environment (DHI, 2005) is a flexible modelling system that combines one-dimensional and two-dimensional hydrodynamic modelling. The model is based on a dynamic linking between the well-established 1D and 2D numerical modelling systems MIKE 11 and MIKE 21 enhanced with features that are tailored towards modelling of floods.

Data assimilation facilities have been implemented in MIKE FLOOD to adaptively update the state of the system by assimilation of flux and water level measurements in both the river system and in the 2D model domain. The implementation is based on the ensemble Kalman filter (first introduced by Evensen, 1994), which has recently been applied in operational hydrodynamic and hydrological modelling, including the MIKE 11 and MIKE 21 modelling systems (Madsen and Cañizares, 1999; Hartnack and Madsen, 2001; Madsen et al., 2003).

This paper describes the implementation of the ensemble Kalman filter in MIKE FLOOD. A twin test example is presented to demonstrate the implementation.

2. MIKE FLOOD MODELLING SYSTEM

The MIKE FLOOD modelling system combines one-dimensional and two-dimensional modelling in one modelling system. It allows modelling part of the model domain in 1D detail (related to the main river channel) and other parts in 2D, and hence provides a computational efficient approach for dealing with different spatial resolutions and process descriptions in different parts of the model domain. Furthermore, the combined approach facilitates implementation of fine scale structures, such as weirs and culverts, within a 2D domain. MIKE FLOOD can be used to simulate the detailed flooding pattern on floodplains in 2D detail while maintaining an efficient 1D description for the in-channel flow. Because of its dynamic linking it is also well suited for modelling coincident river and storm surge flooding in coastal areas.

Three different linkage types are available for dynamically linking the 1D and 2D models: (i) standard link, (ii) lateral link, and (iii) structure link. The standard link type is a point-to-point coupling between a computational point in the 1D model and a computational point in the 2D model domain. It can, for instance, be used to couple upstream river branches to a downstream flood plain. The lateral link type couples the 1D model with the 2D domain over a part of the river system by a lateral inflow/outflow coupling between the two domains. The flow is described through a weir equation in each linked grid point. The lateral link type can be used for simulation of overtopping of levees and flow into the floodplain. The structure link type uses a 1D model in a 2D domain to model flow through a structure (defined in the 1D model). This link type is useful for modelling small-scale features which cannot be adequately resolved using the typical grid size in the 2D model.

3. ENSEMBLE KALMAN FILTER

The data assimilation system implemented in MIKE FLOOD is based on the Kalman filter. The Kalman filter algorithm uses a stochastic representation of the model and the measurements to take into account the inherent uncertainties in the modelling and assimilation processes. Model uncertainties are related to errors in the physical conceptualisations and mathematical approximations of the governing processes, errors due to the use of non-optimal parameter values, and errors in the model forcing. Measurement uncertainties are related to errors in the measuring devices and representation errors, e.g. caused by using a point measurement to represent a grid value in the numerical model.

It is here assumed that model uncertainties are caused only by errors in the model forcing. In this case the stochastic representation of the MIKE FLOOD model can be written

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

where $\Phi(.)$ is the model operator that represents the numerical scheme of MIKE FLOOD, 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 ε_k is the model error. The state vector includes discharges and water levels in the computational grid of the river system as well as depth-averaged x and y velocities and water levels in each computational point of the 2D grid. The forcing includes different boundary conditions for the river model (water level, discharge, rainfall-runoff, rating curves) as well as open boundary conditions (prescribed water levels or fluxes) and wind forcing of the 2D modelling part.

Model errors are assumed to be correlated in both time and space. Time correlation is implemented using a first-order autoregressive model

$$\varepsilon_k = A\varepsilon_{k-1} + \zeta_k \tag{2}$$

where $A = \text{diag}[\alpha_1,\alpha_2,,]$ is a correlation matrix with autoregressive coefficients α_i, and ζ_k is white noise. The state is augmented with the correlated model errors, which are updated as part of the filtering process along with the model state. The coloured noise formulation is very effective in correcting erroneous model forcings (Madsen and Cañizares, 1999), including correction of model bias (Drécourt et al., 2006). Spatial correlation of errors in the wind forcing and open boundary conditions is implemented using exponential correlation models.

For updating the model, the data assimilation system utilises water level and flux measurements at specific locations in the river system and 2D model domain. This is formulated in the measurement equation

$$z_k = C_k x_k + \eta_k \tag{3}$$

where z_k is the vector of measurements, C_k is a matrix that describes the relation between measurements and state variables (mapping of state space to measurement space), and η_k is the measurement error, which is assumed unbiased.

In the Kalman filter, the model forecast x_k^f and corresponding covariance P_k^f are corrected based on the measurements to get an updated state x_k^a and covariance P_k^a

$$x_k^a = x_k^f + K_k (z_k - C_k x_k^f) \tag{4}$$

$$P_k^a = P_k^f - K_k C_k P_k^f \tag{5}$$

where K_k is the Kalman gain matrix

$$K_k = P_k^f C_k{}^T \left[C_k P_k^f C_k{}^T + R_k \right]^{-1} \tag{6}$$

and R_k is the covariance matrix of the measurement errors. It is assumed that measurement errors are uncorrelated. In this case a computationally efficient and robust sequential algorithm that process one measurement at a time can be used for the Kalman filter update (Chui and Chen, 1991). This avoids explicit calculation of the covariance matrix P_k^f and the matrix inversion for the calculation of the Kalman gain in Eq. (6).

The Kalman filter is implemented in MIKE FLOOD using the ensemble Kalman filter (EnKF) method, first introduced by Evensen (1994). In the EnKF the stochastic model equation is represented by an ensemble of state vectors. In the model forecast, each of these states are propagated according to the model dynamics and forced by model errors, cf. Eq. (1). This provides an estimate of the covariance matrix P_k^f. In the measurement update, each state vector is updated according to the Kalman filter update scheme. The resulting updated ensemble provides an estimate of the average state and the corresponding covariance matrix P_k^a. It should be noted that in the implementation the covariance matrices are never calculated explicitly.

The EnKF has been widely applied in oceanography and meteorology and more recently in hydrodynamic and hydrological modelling (e.g. Hartnack and Madsen, 2001; Madsen et al., 2003; Sørensen et al., 2004; Drécourt et al., 2006). The success of the EnKF is mainly due to its ease of use and very flexible implementation. The EnKF can be easily implemented in a numerical engine, basically by making a loop around the one-time-step-ahead model prediction, and it allows definition of noise in any component of the numerical model (model parameters, model forcings, or the model state itself). However, one of its drawbacks is related to the rather slow convergence of the sample approximation to the true covariance. For a small ensemble size, spurious correlations may be introduced, resulting in unrealistic updates and potential model instabilities. Temporal and spatial regularisation techniques have been introduced to solve these problems (Sørensen et al., 2004), hence making the EnKF computationally feasible for real-time applications.

4. TWIN TEST EXPERIMENT

To evaluate the performance of the data assimilation system implemented in MIKE FLOOD a twin test experiment has been carried out. The model being investigated consists of a flood plain with a main river and a tributary (see Figure 1). The 2D modelling domain covers an area of 13 x 15 km. The rivers have a slope of about 0.025 % and cross sections with a width of about 400 meters from side bank to side bank. The main river length is 23 km and the tributary length is 11 km.

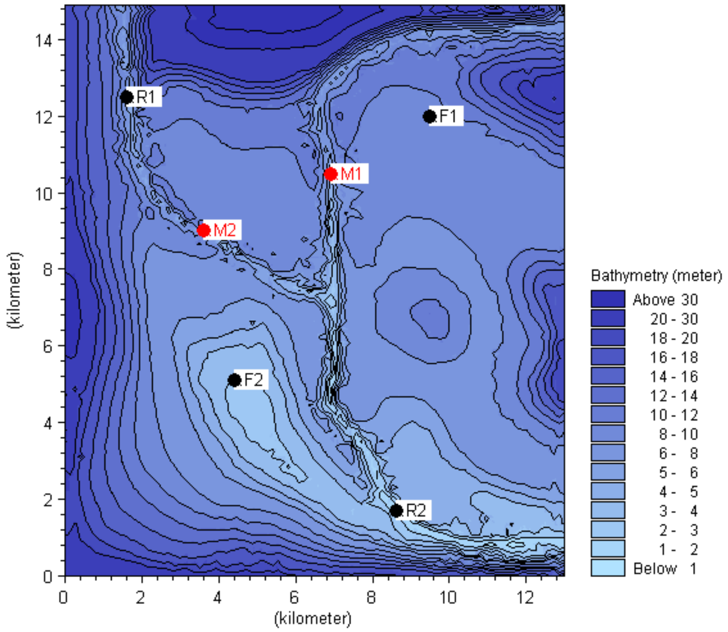

FIGURE 1. Bathymetry of the MIKE FLOOD modelling domain. Marked points are the two measurement points in the river system (M1 and M2), two validation points in the river system (R1 and R2), and two validation points on the floodplain (F1 and F2).

The main river and the tributary where the flow direction is well defined are modelled using the 1D part, whereas the remaining floodplain is modelled using the 2D component. The 1D and the 2D domains are linked using lateral inflow/outflow links. The use of lateral links ensures a physically sound approach to modelling the water entering the floodplain through the overtopping of the river embankments. The 2D domain is completely closed, i.e. no open boundaries are defined. Thus, the only source of flow in the 2D domain is through the spilling from the 1D component into the 2D area. The 1D component has two upstream boundaries consisting of a discharge hydrograph, one at the main river and one at the tributary. At the downstream end of the main river a rating curve has been implemented.

For the twin test experiment, a flood event was considered with flood hydrographs defined at the two upstream boundaries. A reference MIKE FLOOD simulation was made by using the reference hydrograph shown in Figure 2 at both boundaries. From this reference run, water level time series were extracted at two locations in the river system, respectively, 8.8 km downstream of the inflow boundary at the main river, and 7.3 km downstream of the inflow boundary of the tributary (see Figure 1). These time series were subsequently used as measurements for updating a model that was forced with erroneous flood hydrographs at the upstream boundaries (see Figure 2). The ensemble Kalman filter was implemented using an ensemble size of 50. Model errors were introduced at the upstream boundaries using a coloured noise description with a resulting standard deviation of 10 % of the instantaneous discharge at the boundaries.

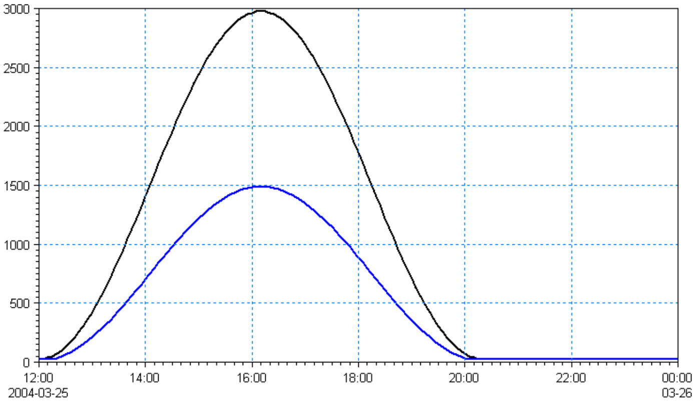

FIGURE 2. Hydrographs [m^3/s] applied at the two upstream boundaries in the river system. The upper curve is used in the reference simulations, and the lower curve is used as the erroneous boundary conditions.

Simulation results are shown in Figures 3-5. In the reference simulation overtopping of the river embankments are primarily occurring at three locations in the river system (see Figure 3). Overtopping first occurs in the upstream part of the main river, and the floodplain at the left bank parallel to the main river (right part of the model domain) is flooded. In the tributary, overtopping occurs at the left bank about 4 km from the upstream boundary, inundating an area in the upper part of the model domain between the tributary and the main river. Overtopping also occurs at the right bank in the tributary close to the confluence with the main river, resulting in large flooding in an area parallel to the downstream part of the main river. When the model is forced with the erroneous boundary conditions, much less flooding is seen. Besides minor flooding close to the river in the upstream part of the main river and at a location at the left bank of the tributary, flooding only occurs in the downstream part due to overtopping of the river embankments in the tributary close the confluence point (see Figure 3).

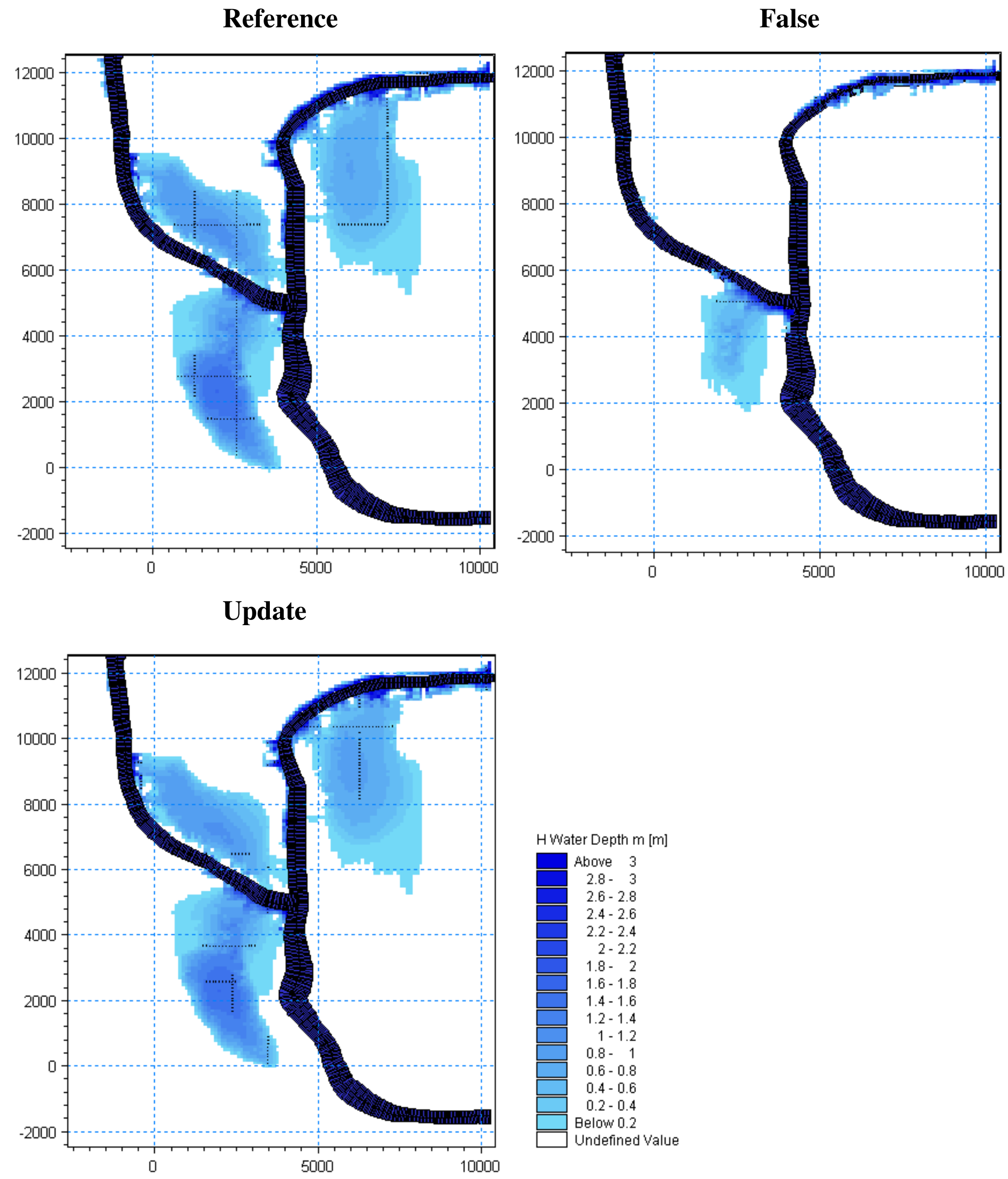

FIGURE 3. Simulated water depth at the flood plain at 2004-03-25 18:00 from, respectively, the reference simulation, the simulation using erroneous boundary conditions (false), and the Kalman filter update.

As illustrated in Figure 3, data assimilation (using in this case water level measurements from only two locations in the river system) is very effective in correcting the model state in the entire model domain. The updated flood map is very close to that obtained from the reference run. More detailed simulation results are shown in Figures 4-5 from two locations in the river system and two locations on the flood plain (locations are shown in Figure 1). Results are shown for two locations upstream of the measurement points (R1 on the upstream part of the tributary and F1 on the upstream part of the flood plain) and two locations downstream (R2 on the downstream part of the main river and F2 on the floodplain downstream of the tributary). At the downstream locations virtually a perfect update is obtained by the Kalman filter. At the upstream locations the update is less accurate, but still a significant improvement is obtained compared to the simulation forced with erroneous boundary conditions. Note that at the upstream validation point on the floodplain, no flooding is simulated with the erroneous boundary conditions.

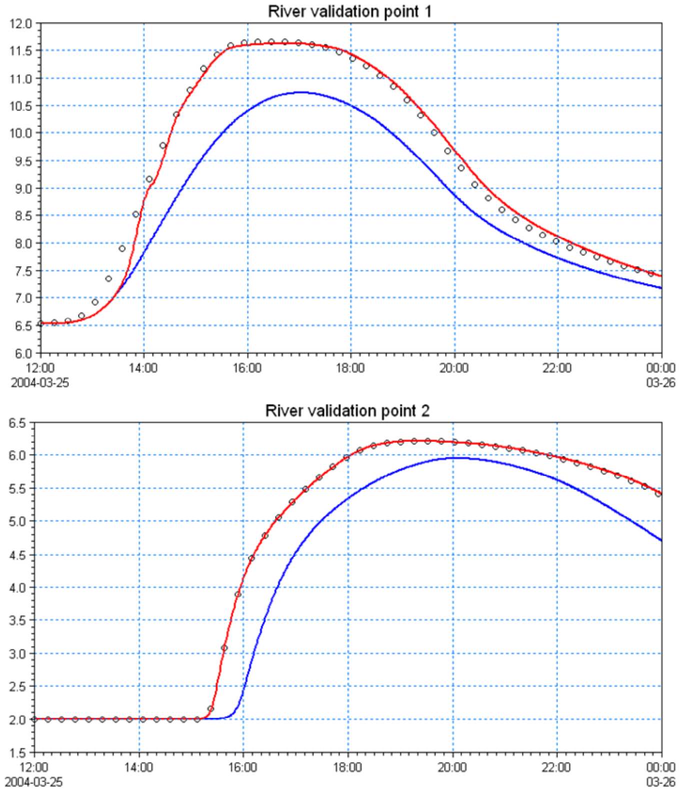

FIGURE 4. Simulated water level [m] at two locations in the river systems from, respectively, the reference simulation (circles), the simulation using erroneous boundary conditions (lower blue curve), and the Kalman filter update (upper red curve).

5. CONCLUDING REMARKS

A data assimilation system has been implemented in the MIKE FLOOD modelling system and has been tested using a twin test experiment. The data assimilation system is based on the ensemble Kalman filter and allows assimilation of water level and flux measurements in both the 1D and 2D model domains. The twin test experiment illustrates that the data assimilation system is very effective in updating the entire model domain using measurements from only two locations in the river system.

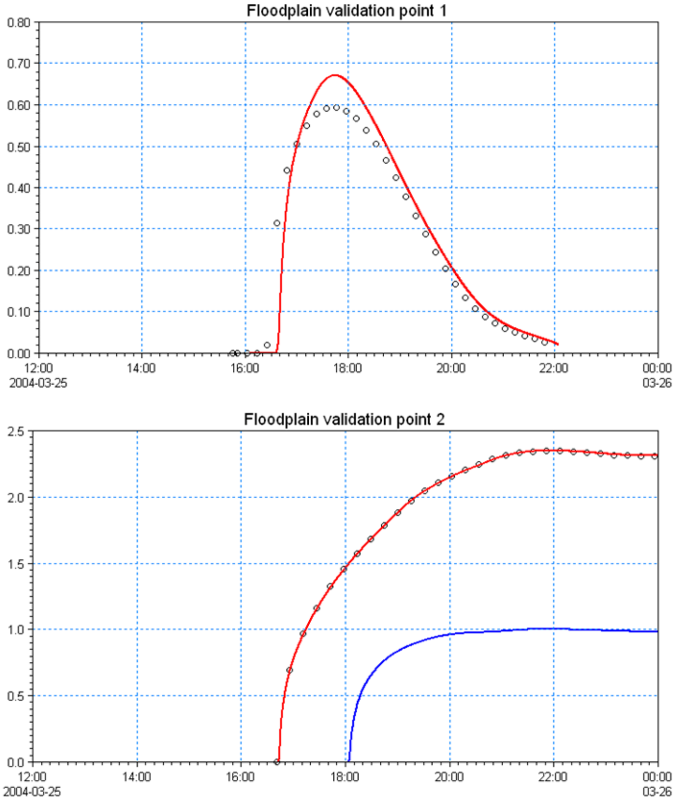

FIGURE 5. Simulated water depth [m] at two locations on the floodplain from, respectively, the reference simulation (circles), the simulation using erroneous boundary conditions (lower blue curve), and the Kalman filter update (upper red curve).

REFERENCES

Chui, C.K. and G. Chen (1991), *Kalman Filtering with Real-Time Applications*, Springer Series in Information Sciences no. 17 (ed. Huang, T.S.), Second edition, Springer-Verlag, Berlin Heidelberg, Germany.

DHI (2005), *MIKE FLOOD User Manual*, DHI Software 2005, DHI Water & Enviroment, Denmark.

Drécourt, J.P., H. Madsen and D. Rosbjerg (2006), Bias aware Kalman filters: Comparison and improvements, *Adv. Water Resour.*, 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.

Hartnack, J. and H. Madsen (2001), Data assimilation in river flow modelling, 4th DHI Software Conference, 6-8 June, 2001, Scanticon Conference Centre, Helsingør, Denmark.

Madsen, H. and R. Cañizares (1999), Comparison of extended and ensemble Kalman filters for data assimilation in coastal area modelling, *Int. J. Numer. Meth. Fluids*, **31**, 961-981.

Madsen, H., D. Rosbjerg, J. Damgård and F.S. Hansen (2003), Data assimilation in the MIKE 11 Flood Forecasting system using Kalman filtering, Water Resources Systems - Hydrological Risk, Management and Development (Proceedings of symposium HS02b held during IUGG2003 at Sapporo, July 2003). IAHS Publ. no. 281, 75–81.

Sørensen, J.V.T., H. Madsen and H. Madsen (2004), Efficient sequential techniques for the assimilation of tide gauge data in three dimensional modeling of the North Sea and Baltic Sea system, *J. Geophys. Res.*, **109**, 10.1029/2003JC002144.