Proceedings of the International Conference "Innovation, Advances and Implementation of Flood Forecasting Technology", Tromsø, Norway, 17-19 October, 2005.

DATA ASSIMILATION IN A COMBINED 1D-2D FLOOD MODEL

Johan Hartnack, Henrik Madsen and Jacob Tornfeldt Sørensen
DHI Water & Environment, Agern Allé 5, DK-2970 Horsholm, Denmark

Abstract

Data assimilation in a combined 1D-2D numerical flood modelling system is considered. The model is based on a dynamic linking between existing and well-established 1D and 2D numerical modelling systems enhanced with new features which are targeted specifically towards modelling of floods. For this combined modelling system data assimilation facilities based on the ensemble Kalman filter methodology have been implemented for assimilation of discharge and water level measurements. The implemented data assimilation system is described and demonstrated on a flood model application in Bangladesh using a twin test experiment. It is demonstrated that the technique can update not only the state but also the forcing applied through the boundary conditions.

Key words: Data assimilation, Ensemble Kalman filter, Flood modelling, Model updating.

INTRODUCTION

In flood modelling 1D river models have traditionally been used. These models may be effective in predicting flood inundation by projecting the model results onto a high resolution digital elevation model. However, when floodplain processes are important and/or more detailed distributed model predictions are required a 2D modelling approach should be adopted. The MIKE FLOOD model developed by DHI Water & Environment (DHI, 2005) is a flexible modelling system that combines 1D and 2D hydrodynamic models especially suited for flood modelling. It allows an efficient 1D modelling of the river system and a detailed 2D modelling of the floodplain. The model is based on a dynamic linking between the well-established 1D and 2D numerical modelling systems MIKE 11 and MIKE 21.

To improve the predictive capabilities of the MIKE FLOOD modelling system by utilising measurements to adaptively update the model state a data assimilation system based on the Kalman filter has been developed. The main strength of the Kalman filter is that it explicitly takes model and measurement uncertainties into account and provides both an updated model prediction and an estimate of the prediction uncertainty. However, a major drawback of the Kalman filter in its original formulation is its huge computational requirements when applied to high-dimensional numerical modelling systems as considered here.

In recent years, cost-effective Kalman filter schemes that are feasible for application in high-dimensional modelling systems have been developed. Such schemes are often referred to as sub-optimal Kalman filter procedures (Todling and Cohn, 1994). State-of-the-art procedures which have been implemented in hydrodynamic modelling systems include the ensemble Kalman filter (EnKF) (Evensen, 1994; Madsen and Cañizares, 1999; Hartnack and Madsen, 2001; Madsen et al., 2003, Sørensen et al., 2004) and the reduced rank square root (RRSQRT) filter (Verlaan and Heemink, 1997; Madsen and Cañizares, 1999). The computational costs of these methods are typically in the order of 100 model simulations (Madsen and Cañizares, 1999) which is orders of magnitudes smaller than required for the full Kalman filter. By applying temporal regularisation (Sørensen et al., 2004) the computational burden can be further reduced.

In this paper the data assimilation system implemented in the MIKE FLOOD model is described. The system is based on the EnKF procedure and allows assimilation of flux and water level measurements in both the river system and the floodplain. For demonstration of the system a flood model application in Bangladesh using a twin test experiment is presented.

MIKE FLOOD MODELLING SYSTEM

By Combining one- and two-dimensional modelling one obtains a highly efficient system as the benefits of both models are utilized while the downsides are eliminated, i.e. detailed modelling and accuracy where needed without sacrificing computational or model development time. With the combined approach it is no longer necessary to compromise between in channel and off-channel spatial resolution. Further the combined approach facilitates implementing fine scale structures within the two dimensional domain e.g. culverts, weirs pumps etc.

The linking of the one-dimensional and two-dimensional models may be achieved through the use of three different linkage types illustrated in *Figure* 1 - 3.

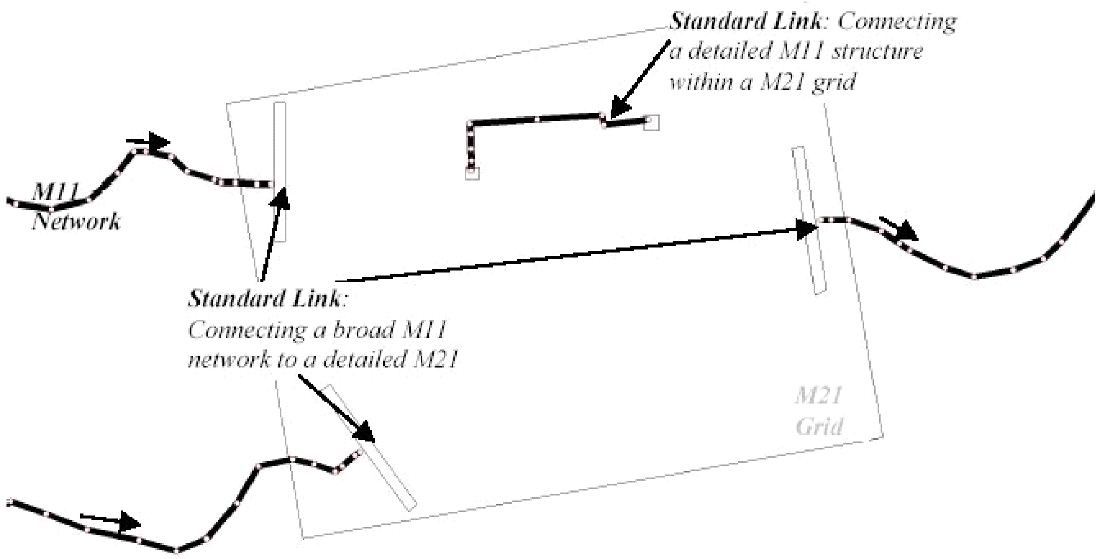

Figure 1 Example of a standard link type between MIKE 11 and MIKE 21.

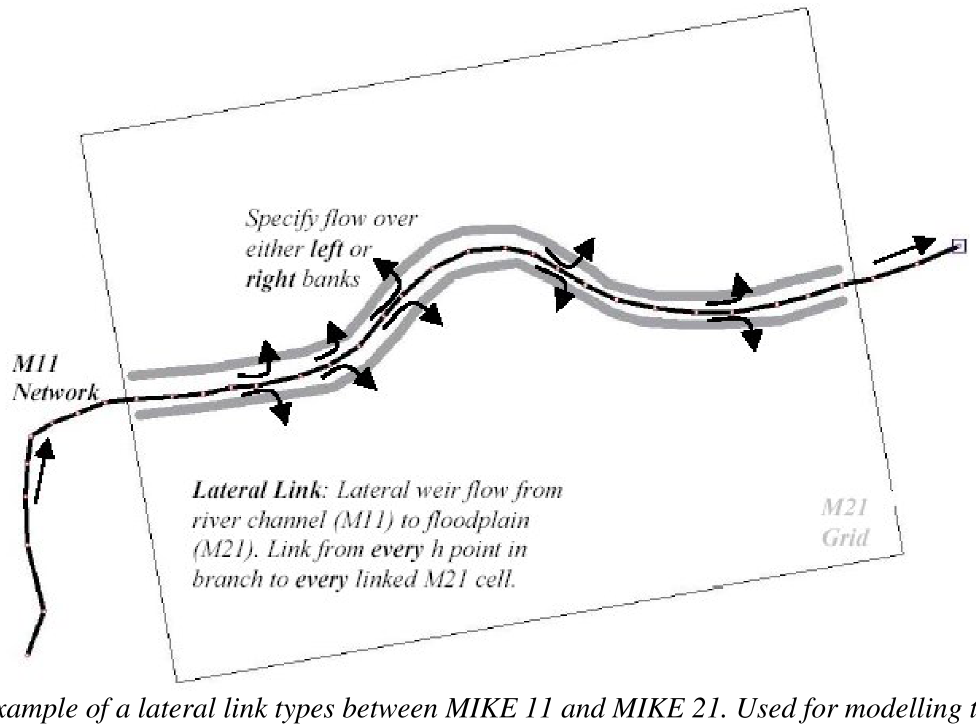

Figure 2 Example of a lateral link types between MIKE 11 and MIKE 21. Used for modelling the overtopping of levees.

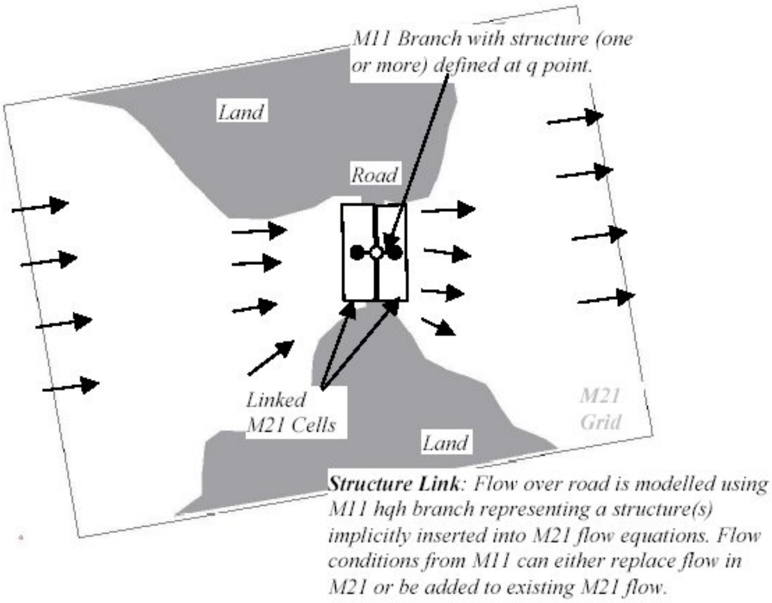

Figure 3 Example of a structure link types between MIKE 11 and MIKE 21. Useful for modelling small scale features in the topography which cannot be resolved using the typical mesh size in the 2-D model.

DATA ASSIMILATION SYSTEM

Kalman filter update scheme

The Kalman filter data assimilation algorithm is based on a stochastic representation of the model dynamics and the measurements. The numerical scheme of the MIKE FLOOD modelling system can be written in a state-space form

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

where $\Phi(.)$ is the model operator representing the numerical scheme of MIKE FLOOD, x_k is the state vector representing the state of the modelled system at time step k, and u_k is the forcing of the system. 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 potential flood point of the grid representing the flood plain. The forcing includes different boundary conditions for the river model (water level, discharge, rainfall-runoff, rating curves) and possible wind forcing.

For modelling the uncertainty of the system, it is assumed that model errors are related to errors in the forcing terms

$$x_k = \Phi(x_{k-1}, u_k + \varepsilon_k) \tag{2}$$

where ε_k is the model error. The model error process is described by defining the model error variance as well as the temporal and spatial correlation structure. If a correlated system noise is defined, an augmented state vector formulation is adopted in the Kalman filter. This formulation provides not only an update of the model state x_k but also an update of the (erroneous) model forcing (Madsen and Cañizares, 1999).

The data assimilation system allows assimilation of flux and water level measurements in the river system and on the floodplain. 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 (a mapping of state space to measurement space), and η_k is the measurement error. This error is assumed unbiased with covariance matrix R_k.

The Kalman filter is a succession of two steps. First, the model is employed to issue a forecast, and then the observed data are meld with the forecast to provide an updated state (analysis step). Denote by x_k^f a one-step ahead forecast of the state of the system, according to the model operator $\Phi(.)$, cf. Eq. (1) and P_k^f the covariance matrix of this forecast. The Kalman filter update of the state vector and the error covariance matrix are given by

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

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

where K_k is the Kalman gain matrix

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

which serves as a weighting function of model forecast and measurements. In Eqs. (4)-(6) superscripts f and a refer to, respectively, forecast and analysis (or update).

Ensemble Kalman filter

The forecast error covariance matrix P_k^f is a combination of time propagation of P_{k-1}^a and forcing of the system by model errors. For large, non-linear systems, the propagation of the errors is the main bottleneck, imposing an unacceptable computational burden. The EnKF proposed by Evensen (1994) approximates the error covariance propagation that significantly reduces the computational burden.

In the EnKF the statistical properties of the state vector are represented by an ensemble of possible state vectors. Each of these vectors is propagated according to the dynamical system subjected to model errors cf. Eq. (2), and the resulting ensemble then provides estimates of the forecast state vector and the error covariance matrix. In the measurement update, the Kalman gain matrix obtained from Eq. (6) is applied for each of the forecast state vectors. To account for measurement errors, the measurements are represented by an ensemble of possible measurements (Burgers et al., 1998). The resulting updated ensemble provides estimates of the updated state vector and the associated error covariance matrix.

When the data assimilation is based on in-situ measurements that are sparsely represented in space, the error covariance matrix needs not to be calculated explicitly. In this case, the measurement matrix C_k has only a few non-zero elements and only the columns in P_k^f that correspond to the non-zero elements in C_k have to be calculated. Furthermore, assuming that measurement errors are uncorrelated, an efficient sequential updating algorithm that processes one measurement at a time can be implemented, which avoids the matrix inversion in Eq. (6) (e.g. Chui and Chen, 1991).

APPLICATION EXAMPLE

To evaluate the performance of the data assimilation system implemented in MIKE FLOOD a twin test experiment has been carried out. First the model was set-up and used in a deterministic setting to generate reference results. From this reference run water level time series were extracted. These extracted water level time series were subsequently used as measurements for the second run where the ensemble Kalman filter was applied. For the ensemble run model errors were introduced through the boundary conditions.

The model being investigated consists of a flood plain that is cut through by a main river and a tributary. The bathymetry is illustrated in Figure 4. The rivers in the system have a slope around 0.25 ‰ and cross sections with a width in the range of 400 meters from side bank to side bank. The main river length is 23000 meters and the tributary length 11000 meters.

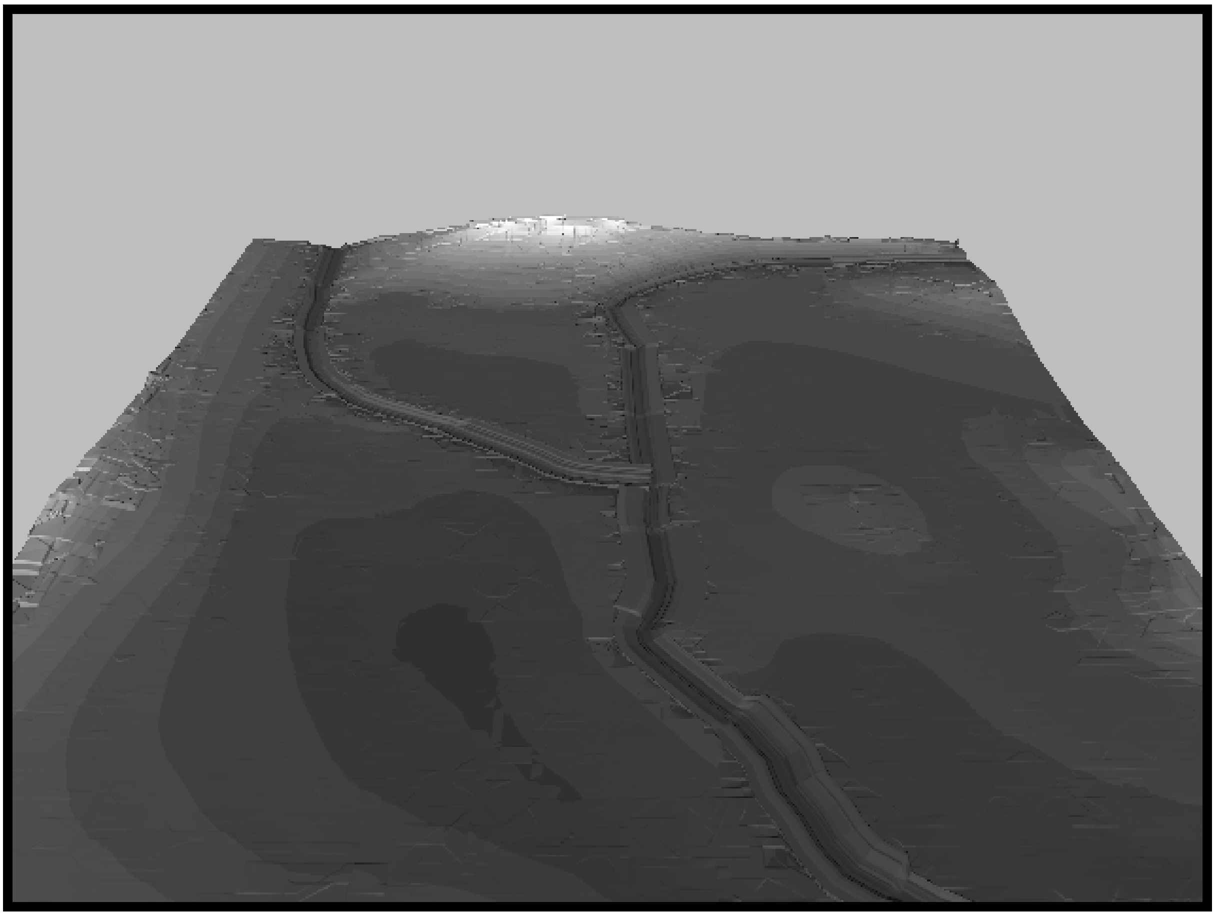

Figure 4: The bathymetry used for the model set-up. The main river runs from the upper right hand corner in a crescent shape to the lower right hand corner. The tributary originates in the upper left hand corner and intersect the main river in the middle of the figure.

The river and the tributary where the flow direction is well defined are modelled using the 1-D part of MIKE Flood whereas the remaining flood plain is modelled using the 2-D component . The linking between the 1-D and the 2-D domain is done using lateral links. The use of lateral links ensures a sound approach to modelling the water entering the floodplain through the overtopping of the river banks in the system.

The 2-D domain is completely closed i.e. no open boundaries. Thus the only source of flow in the 2-D domain is through the spilling of the 1-D component into the 2-D area. The 1-D component on the other hand has two upstream boundaries one at the main river and one at the tributary consisting of a discharge hydrograph. Figure 5 displays the hydrograph used along with the erroneous hydrograph used for the Kalman filter run. At the downstream end of the main river a rating curve has been implemented.

The ensemble Kalman filer run uses an ensemble size of 100 to ensure a proper estimate of the covariance matrix. Model errors are introduced through the upstream boundary conditions which are taken as having a standard deviation equal to 10 percent of the instantaneous discharge. At the peak values this may give rise to very large model errors which in turn can cause numerical instabilities. To mitigate the latter an upper limit of 100 m^3/s was placed on the standard deviation.

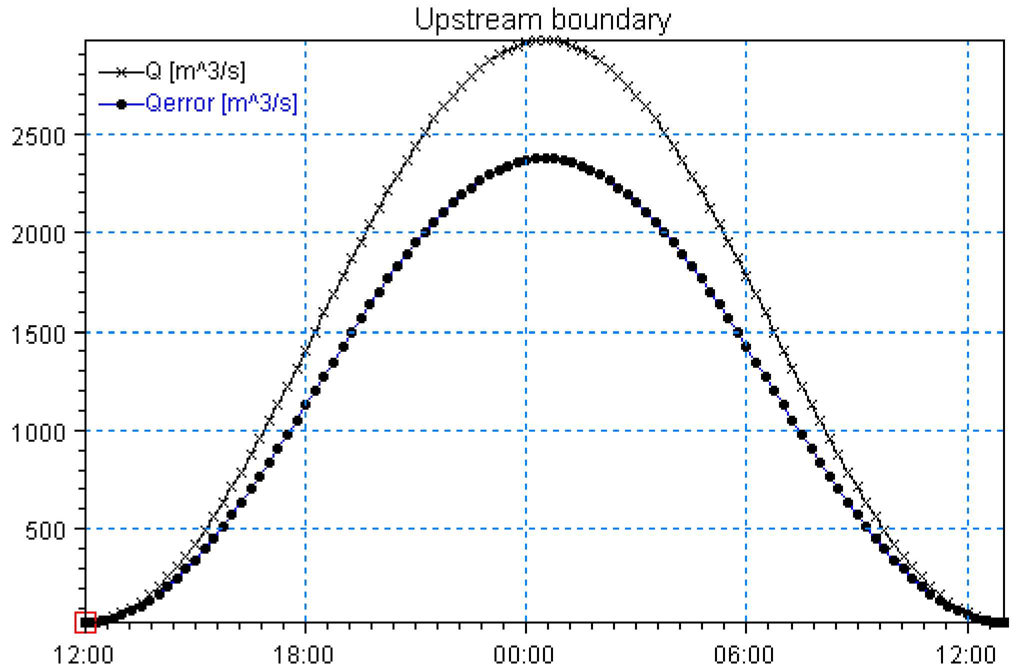

Figure 5 The upstream boundary conditions applied. The upper hydrograph was used for the reference simulation and the lower for the Kalman filter run.

From the reference run two water level time series were extracted one in the main river and one in the tributary. The one in the main river was located 8800 meters downstream of the upstream boundary and the measurement location in the tributary was 7300 meters downstream of the corresponding boundary in the tributary.

To illustrate the effectiveness of the Kalman filter Figures 6 and 7 show the water level in the main river 3000 meters from the upstream boundary and at the downstream boundary condition. Three curves are shown: the reference simulation, the deterministic run with the erroneous boundary conditions and finally the updated water level using the Kalman filter. As is evident from these figures the updating mechanism in the Kalman filter corrects the water level not only at the updating points but throughout the system.

The errors (ε) at the upstream boundaries are described through an autoregressive function of first order.

$$\varepsilon_{k+1} = a\varepsilon_k + \zeta_{k+1} \tag{7}$$

where ζ is white noise. This first order auto regressive expression generates a forecast of the model error one time step ahead. This in turn makes it possible to update the boundary conditions as well. The updated boundary condition is illustrated in Figure 8 along with the reference boundary condition and the erroneous boundary condition. As is evident the Kalman filter does a fine job of correcting the boundary.

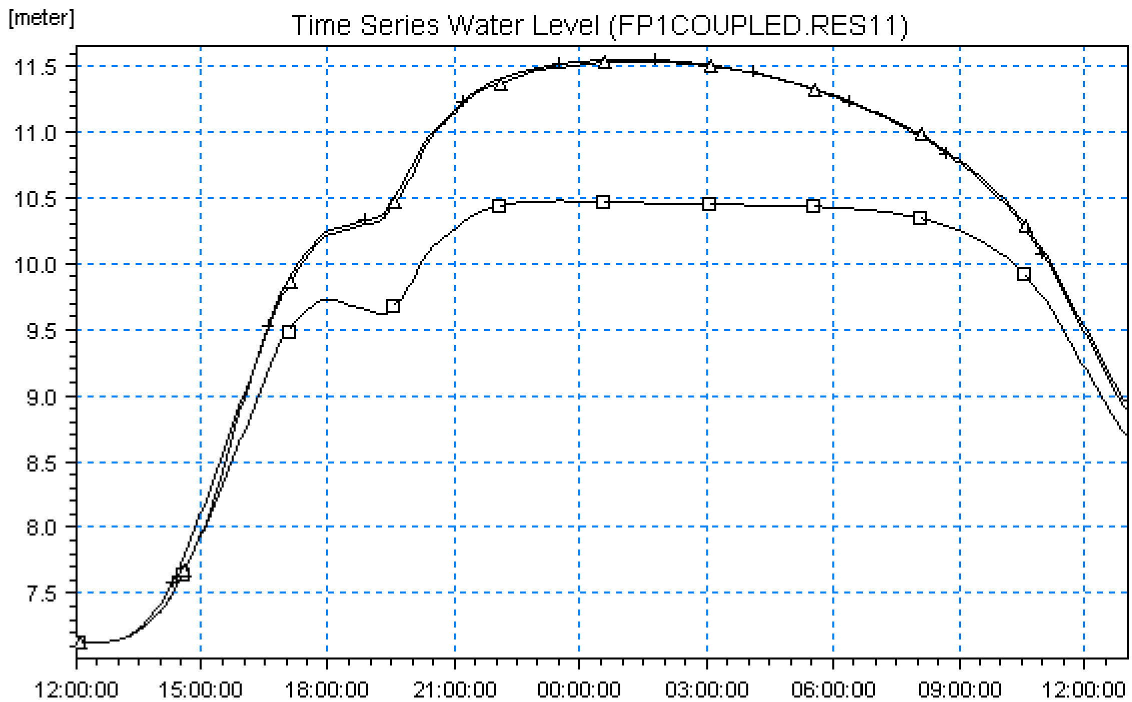

Figure 6: The water level at in the main river 3000 meters downstream of the upstream boundary condition:. (+) *reference simulation,* (□) *erroneous simulation and* (Δ) *Simulation with Kalman filter updating.*

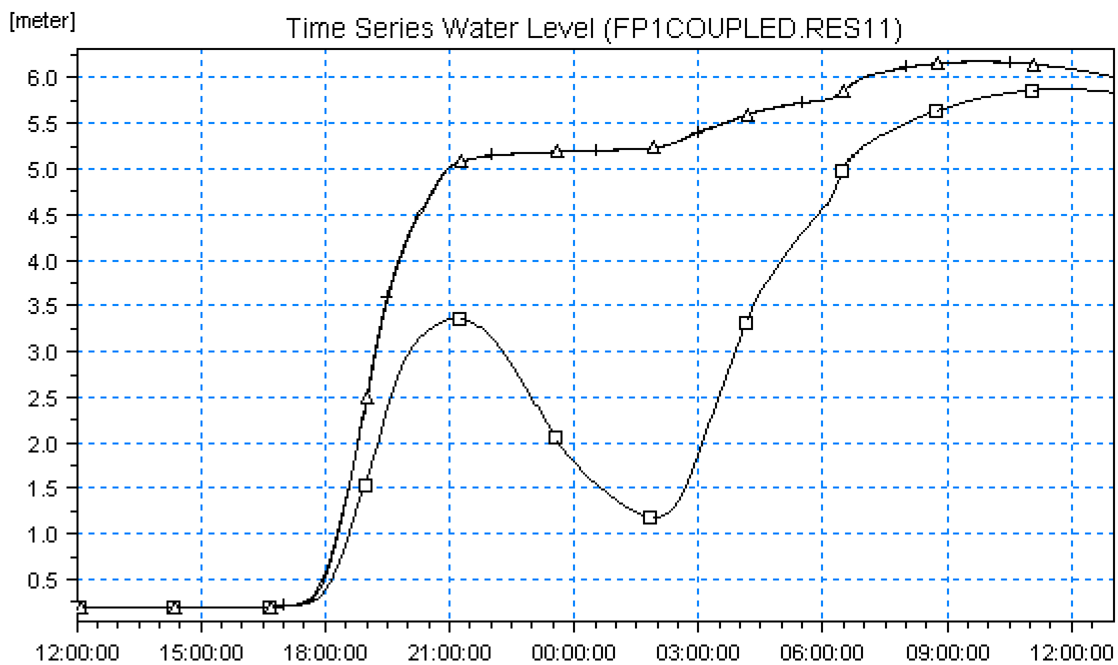

Figure 7: The water level at the downstream boundary condition:. (+) *reference simulation,* (□) *erroneous simulation and* (Δ) *Simulation with Kalman filter updating.*

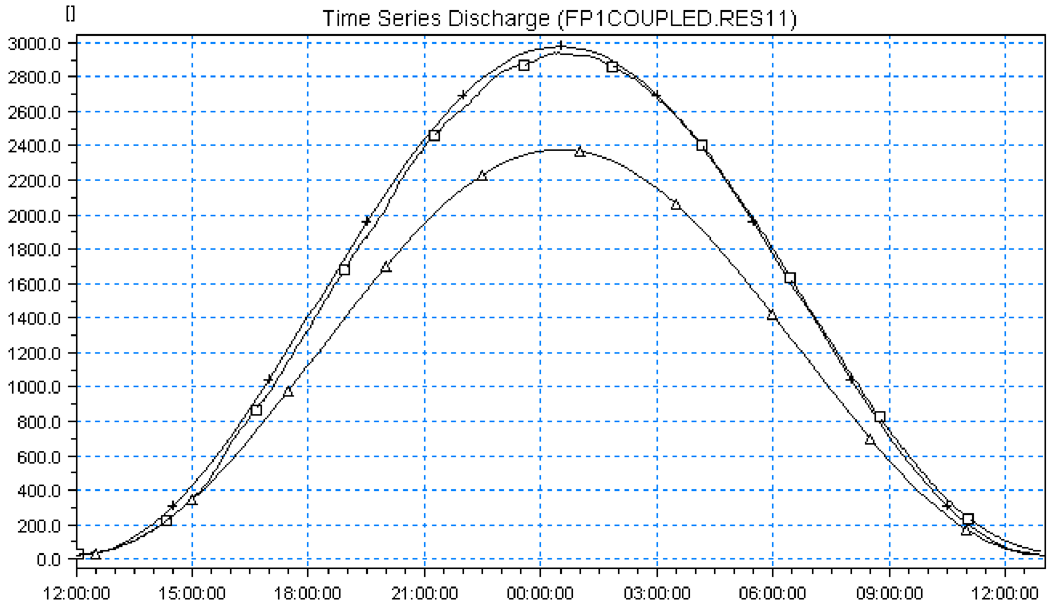

Figure 8: The upstream boundary condition in the main river:. (+) *reference boundary,* (Δ) *erroneous boundary and* (□) *boundary with Kalman filter updating.*

CONCLUSIONS

The Kalman filter has been implemented in a combined 1-D/2-D model and subsequent used for updating an erroneous model. It was shown that not only does the Kalman filtering technique update the state variables in the system (water level and discharges) but by introducing a simple forecast of the model error the technique may also be used for updating the boundary conditions as well.

REFERENCES

Burgers, G., Van Leeuwen, P.J. and Evensen, G., 1998, On the analysis scheme in the ensemble Kalman filter, *Monthly Weather Rev.*, **126**, 1719-1724.

Chui, C.K. and Chen, G., 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.

Evensen, G., 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, 1994.

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

Madsen, H. and Cañizares, R., 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., Rosbjerg, D., Damgård, J., Hansen, F.S., 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., Madsen H. and Madsen H., 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.

Todling, R. and Cohn, S.E., 1994, Suboptimal schemes for atmospheric data assimilation based on the Kalman filter, *Mon. Wea. Rev.*, **122**, 2530-2557.

Verlaan, M. and Heemink, A.W., 1997, Tidal flow forecasting using reduced rank square root filters, *Stochastic Hydrol. Hydraul.*, **11**, 349-368.