Proc. International Conference "Innovation, advances and implementation of flood forecasting technology" 17-19 October 2005, Tromsø, Norway.

ENSEMBLE-BASED METHODS FOR DATA ASSIMILATION AND UNCERTAINTY ESTIMATION IN THE FLOODRELIEF PROJECT

M. B. Butts (1), A. K. Falk (1), J. Hartnack (1), H. Madsen (1), A. Klinting (1), T van Kalken (2), D. Cadman (3), and D. Price (3)
(1) DHI Water & Environment, Agern Alle 5, DK 2970 Hørsholm, Denmark,
(2) DHI Water & Environment Ltd, PO Box 300-705, Albany #103 E Centre, Oaklands Rd, Albany, New Zealand
(3) Environment Agency, Anglian Region, Kingfisher House, Goldhay Way, Orton Goldhay, Peterborough PE2 5ZR, United Kingdom

Abstract
Real-time flood management decisions including flood warning must be based on an understanding of the uncertainties and associated risks. Forecast uncertainty requires the estimation of the uncertainties associated with the hydrological model inputs (precipitation observations and forecasts), model structure, parameterisation and calibration, and methodologies that predict how the uncertainties from different sources propagate through the hydrological system. Ensemble-based approaches are attractive because they allow effects of a wide range of uncertainties to be incorporated.

Within the EU 5th framework project FLOODRELIEF, complementary ensemble-based approaches have been developed to address the issue of handling and quantifying forecasting and modelling uncertainties. Firstly a general stochastic framework based on the Ensemble Kalman Filter has been developed for flood forecast modelling. The Kalman filter provides a natural framework for determining how the different sources of uncertainty propagate through the hydrological models and to reduce forecast uncertainty via data assimilation of real-time observations. An evaluation of this framework is presented for two case studies; US NWS study catchment, the Blue river basin and the UK FLOODRELIEF study catchment, the Welland and Glen. The results of this evaluation highlight the fact that one of the major outstanding problems in uncertainty estimation is the characterisation of the sources of uncertainty.

The second approach is the development of an internet-based decision support system designed together with the FLOODRELIEF end-user to support ensemble forecasting including the uncertainty framework presented here. For example, ensemble rainfall forecasts using meso-scale meteorological forecasts (Xuan et al, 2005) downscaled meteorological forecasts (Butts et al., 2005), radar forecasts, best case and worst case forecasts can be used by operational forecasters to models to estimate an uncertainty range. In this manner a direct and intuitive estimate of forecast uncertainties can be achieved to address the issue of how ensemble results can be communicated to flood managers and decision-makers.

Key words: ensemble methods flood forecasting, forecast uncertainty, numerical weather forecasting, weather radar

INTRODUCTION

Uncertainty is inherent in the flood forecasting process and increasingly flood forecasting decision makers are recognising that real-time flood management decisions including flood warning must be based on an understanding of the uncertainties and associated risks, (Cadman et al., 2005). There are a number of potential sources of uncertainty that may contribute to forecast uncertainty. Several authors, (Rajaram and Georgakakos, 1989, Refsgaard 1997, Madsen 2000, Beven 2000, Butts et al., 2004) classify the sources of modelling uncertainty into the following main groups

- Random or systematic errors in the model inputs (boundary or initial conditions);
- Random or systematic errors in the observed data used to measure simulation accuracy;
- Uncertainties due to sub-optimal parameter values;
- Uncertainties due to incomplete or biased model structure.

In the case of forecasting modelling additional contributions to forecast uncertainty and accuracy need to be considered. The most important of these are the uncertainties associated with *forecasted* inputs such as quantitative precipitation forecasts, Butts et al., (2002). In addition the efficiency of the data assimilation may have a significant effect on forecast accuracy, (Madsen et al., 2000). Finally unpredictable effects such as processing or human errors, channel blockage and dyke failure and the like may have significant local impact on forecast accuracy.

An important component of the FLOODRELIEF project is the treatment of uncertainties for flood forecasting. There are a number of challenges in treating these uncertainties

- Quantification of the uncertainty sources;
- Evaluation of the impact of the different sources of uncertainty on the flood forecast accuracy including their propagation;
- Evaluation of the impact of the uncertainty on the decision making and management options;
- Provision of this uncertainty information in a manner that can be understood by operational forecasts and decision makers.

At the outset of the FLOODRELIEF end-users evaluated the limitations of existing flood forecasting systems to define end-user requirements. The FLOODRELIEF end-users identified uncertainty as one of the most significant limitations in current forecasting systems, (Butts and Khatibi, 2003; Price et al., 2003). Within the FLOODRELIEF project the main focus has been developing approaches to determine the impact of different sources of uncertainty associated with flood forecasts, each of these challenges is being addressed within different parts of the project.

This paper presents some of our investigations into the propagation of uncertainties from meteorological and hydrological boundaries and the reduction of uncertainties using real-time updating or data assimilation. Firstly a general stochastic framework based on the Ensemble Kalman Filter has been developed for flood forecast modelling. This framework can be used both to estimate the propagation of uncertainties from different sources through the hydrological forecasting models and to reduce forecast uncertainty via data assimilation of real-time observations. Secondly the ability to represent the uncertainty using ensemble predictions either from meteorological model ensembles, from hydrological model ensembles, has been developed with the FLOODRELIEF decision support system.

THE UNCERTAINTY FRAMEWORK

The main aim of this study was to develop a framework for data assimilation and for quantifying uncertainties for flood forecasting. This framework allows evaluation of the propagation of uncertainties using Monte-Carlo approaches and ensemble Kalman filter procedures for data assimilation to reduce forecast uncertainties. Flood forecasting is carried out using hydrological and hydraulic models, in this case the MIKE 11 model, (Havnø et al., 1995).

The Kalman Filter is used for data assimilation purposes as a method for adjusting simulated values towards measured values in a way that is consistent with the dynamics of the system. Based on the available measurements and the given model forecast an updated estimate may be constructed. Ensemble Kalman Filter (EnKF) was initially introduced by Evensen (1994) as a Monte Carlo based Kalman filter for an ocean model. Because of its improved ability to treat non-linearity and model-independent nature, this filter is widely used in oceanography and meteorology and more recently hydrology, (Houtekamer and Michell, 1997, Madsen and Canizares, 1999, Hartnack and Madsen, 2001, Madsen et al., 2003).

The idea of the Ensemble Kalman Filter is to propagate a number of realizations from the initial state and then derive the statistical quantities needed for the Kalman filter update from this ensemble. The spreading of the ensemble is introduced by perturbing the boundary conditions of the model. The Ensemble Kalman Filter also provides a straight forward method for determining the propagation of the uncertainties in the model. After taking each realization one time step ahead the standard deviation of the ensemble can be calculated for each computational point. If measurements are available, the model state and model uncertainty are updated according to the Kalman filter update scheme. If no measurement update is performed at the end of the time step, the calculated standard deviations will show the evolution of the model uncertainty.

BLUE RIVER CASE STUDY

The 1232 km^2 Blue river basin is located in south-central Oklahoma and flows into the Red River at the Texas-Oklahoma border, (Figure 1). The watershed is semi-arid, with significant convective rainfall events. Distributed rainfall data is available in the form of NEXRAD gridded data provided at hourly intervals at a spatial resolution of 4 km by 4 km.

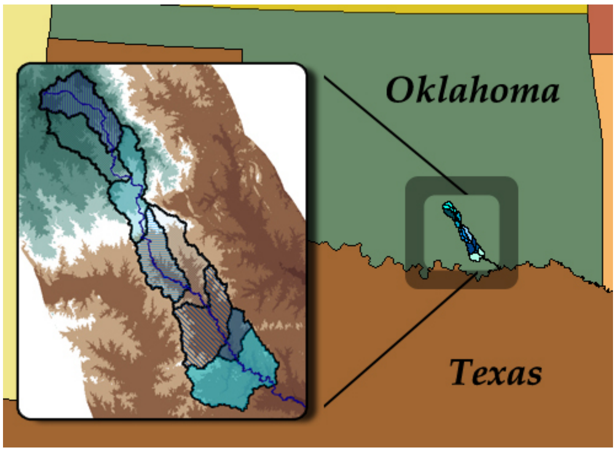

Figure 1 The Blue river catchment location map and subcatchments used for modelling.

The Blue River base is one of the test basins within the Distributed Modeling Intercomparison Project (DMIP) organised by the Hydrology Lab of the National Weather Service (NWS). The purpose of the DMIP study was to evaluate the capabilities of existing distributed models and identify avenues for model improvements, http://www.nws.noaa.gov/oh/hrl/dmip/. It has been the subject of several investigations into model accuracy and uncertainty (Boyle et al., 2001, Butts et al., 2004). The uncertainty framework developed here, (Falk et al., 2005a), has been implemented in the MIKE 11 catchment and river modelling tool, Havnø et al., (1995). A distributed model was developed using the NAM rainfall-runoff model (Madsen, 2000) with 8 subcatchments to represent the rainfall-runoff processes and a single river branch to represent the main river. This model has been calibrated using automatic calibration methods (Madsen 2000, Butts et al., 2004).

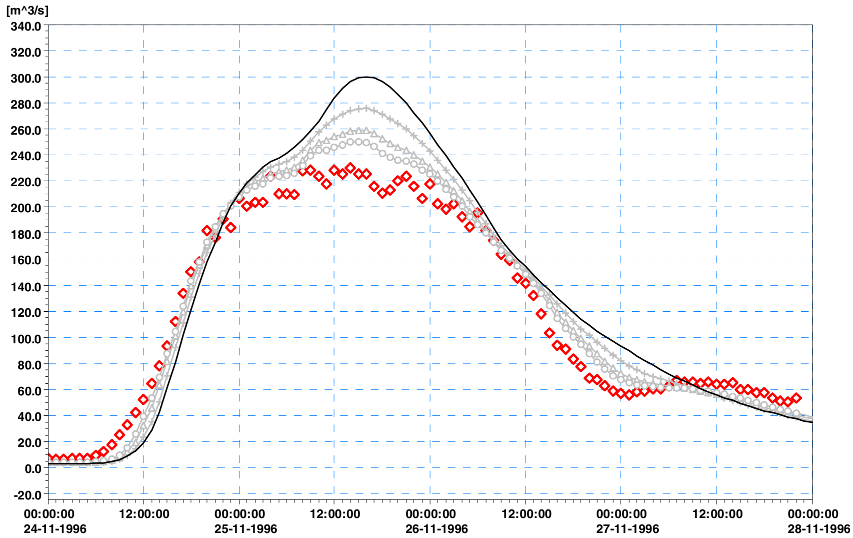

Figure 2 Impact of variations in the uncertainty in lateral inflows from the NAM catchment models. The deterministic reference is shown as a (black) solid line and the measured data as (red) diamonds. Three simulations with data assimilation using different relative standard deviation on runoff error, 10% (crosses), 20% (triangles), 30% (circles), are shown and the errors are uncorrelated in time (white noise).

The Blue River has an unusual aspect ratio requiring both distributed routing and distributed catchment modelling where the timing and volume of the catchment inflows can be important for predicting discharge at the outlet. The uncertainty framework can be used to look at different formulations of the Ensemble Kalman Filter. A series of investigations have been carried out to evaluate the efficiency of this framework and to determine the sensitivity of the filter formulation to a range of parameters and conditions (Falk et al., 2005b). A subset of these results is presented here for a single peak event (22NOV1996 to 29NOV1996) using discharge measurements taken at the outlet of the catchment.

In the first case the inflows from the different subcatchments were calculated deterministically and then an uncertainty assigned to these inflows, Figure 2. Data assimilation is carried out by updating the river state to model this peak. The magnitude of the inflow uncertainty will obviously affect the magnitude of the downstream discharge. As shown in Figure 2 as the uncertainty increases the data assimilation has greater freedom to construct an updated state closer to the measurements. In the Kalman filter framework the uncertainty in the discharge measurements is also taken into account. As the measurements themselves are uncertain the assimilated values will match the observed discharge to within the uncertainty bounds. In this case the uncertainties in the measured discharge are assumed to be normally distributed white noise with zero mean. The standard deviation is assessed to be 5% relative to the measured value (Butts et al., 2004).

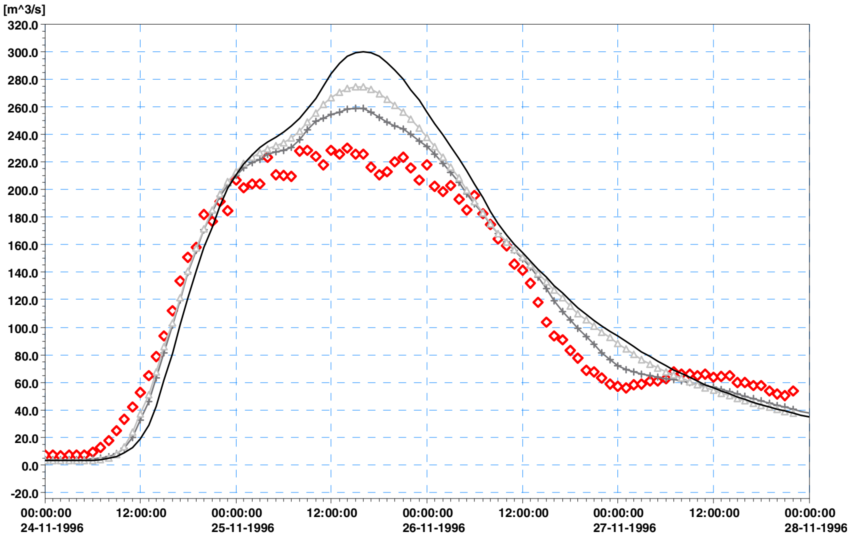

Figure3 Assimilation results using 10% standard deviation on precipitation errors (triangles) with assimilation in the river model compared with 20% standard deviation on runoff errors (crosses) and assimilation in river model. The deterministic reference is shown as a (black) solid line and the measured data as diamonds.

A similar analysis was carried out where instead the uncertainty was assigned to the precipitation. The uncertainty in the lateral inflows is derived by propagating the rainfall uncertainty through the NAM rainfall-runoff model and updating is carried out in the river. Uncertainty analyses showed that uncertainties of around 10% in the rainfall lead to uncertainties of 20% in lateral inflow. There are differences in the assimilated results using these approaches but it is difficult to determine which performs better, Figure 3.

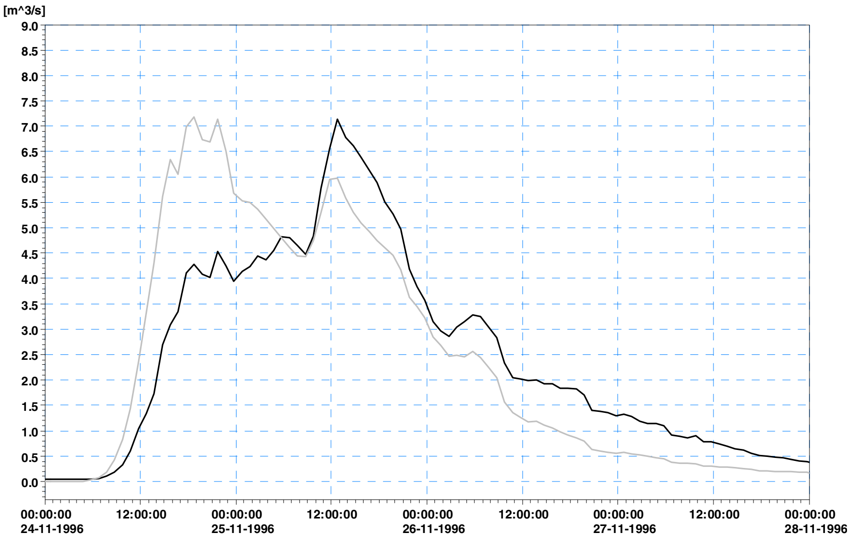

Figure 4 Standard deviations in the simulation of uncertainty propagation for two cases. Case 1, with 20% standard deviation on runoff errors, (dark line), Case 2 with 10% standard deviation on precipitation errors (.light line).

Interestingly the behaviour of the uncertainty estimated in the two cases, shown in Figure 4 as standard deviations, is slightly different with the estimated uncertainty being larger in the early part of the peak in the case of precipitation uncertainty. This means that when updating is carried out in both the NAM rainfall-runoff model and the river model assimilation is better able to represent the early rise of the flood hydrograph when compared to updating in the river only, Figure 5. Differences in the forecasting accuracy of these two cases are currently being investigated.

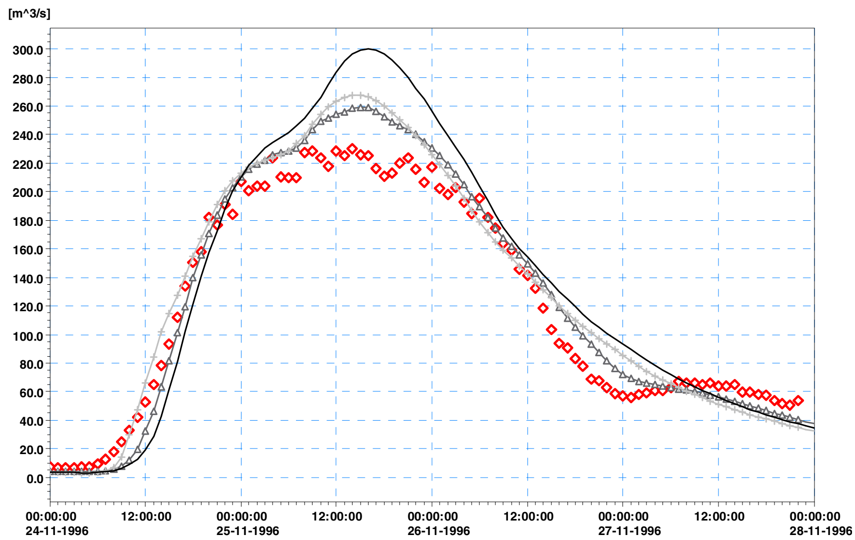

Figure 5 Comparison of data assimilation within the river model (triangles) and within both the river model and the rainfall-runoff model (crosses). The deterministic reference is shown as a solid line and the measured data as diamonds.

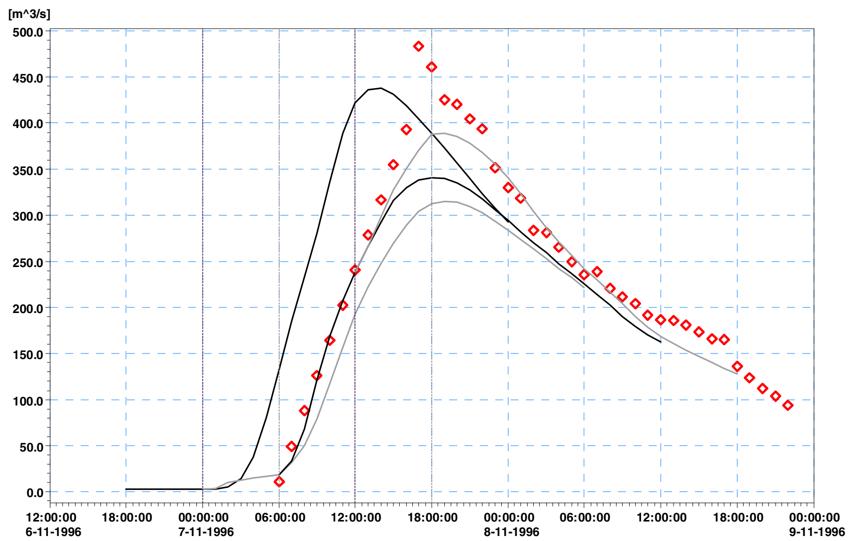

Figure 6 Four forecasts of discharge on 07NOV1996 (TOF: 00:00:00, 06:00:00, 12:00:00, 18:00:00) together with the observed discharge (diamonds). The Time of Forecast (TOF) is indicated by a vertical dashed line.

The analyses shown thus far use data assimilation at each time step. In operational forecasting, simulations are carried out in forecast mode where there is no data available in the forecast period, i.e., data are assimilated until the Time of Forecast (TOF), after TOF the uncertainty is propagated without assimilation. Figure 6 shows a sequence of forecasts. Each forecast (including the uncertainty estimation) uses the previous forecast as initial conditions.

The figure shows four forecasts of discharge for a second major peak in November, 1999. The uncertainties are estimated assuming 10% standard deviation (≈5% variance) in the rainfall and using an ensemble size of 100 members. In figure 7, the forecasted uncertainties are shown as 95% confidence intervals around the ensemble mean.

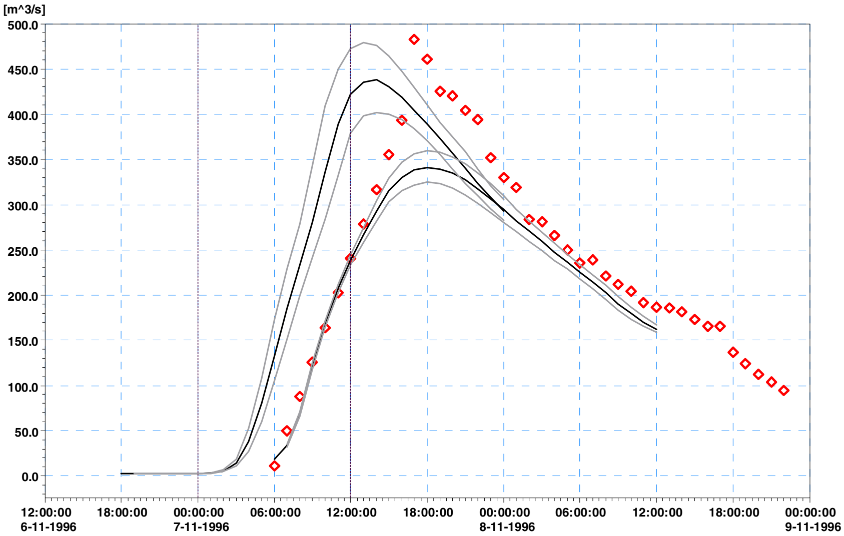

Figure 7 Confidence intervals for two of the forecasts shown in Figure 6 together with the observed discharge (diamonds).

FLOODRELIEF STUDY CATCHMENT: WELLAND AND GLEN

The catchments of the rivers Welland and Glen are located in the East of England and together drain an area of approximately 1150km^2. The River Welland drains a catchment which extends from its headwaters near Market Harborough at around 100 m above sea level to the Wash estuary. The River Glen drains a catchment which rises east of Grantham and joins the Welland north east of Stamford. There are a number population centres within the catchment including Oakham, Stamford, Spalding, Market Harborough and the northern fringes of Peterborough.

Figure 7 Location map and river system for the FLOODRELIEF study catchment, Welland and Glen.

The Welland and Glen system is a complex river system and therefore a comprehensive model was developed using the MIKE 11 river and catchment modelling tool. For flood forecasting continuous rainfall runoff modelling is carried out using the NAM model for a total of 55 subcatchments, of which 20 are pumped. A detailed hydrodynamic model was required in part because of the extensive tidal influences and the number of other artificial influences. The hydraulic model contains more than 200 weirs and culverts and around 50 hydraulic control structures. Also extensive modelling of the flood plains is performed. The model is described in Cadman et al. 2005, Butts et al. 2001. The results presented in the following sections are from the flood event that occurred in January 1998 for the "Kate's Bridge" station which is a key updating point in the operational Anglian Flow Forecasting Modelling System (AFFMS).

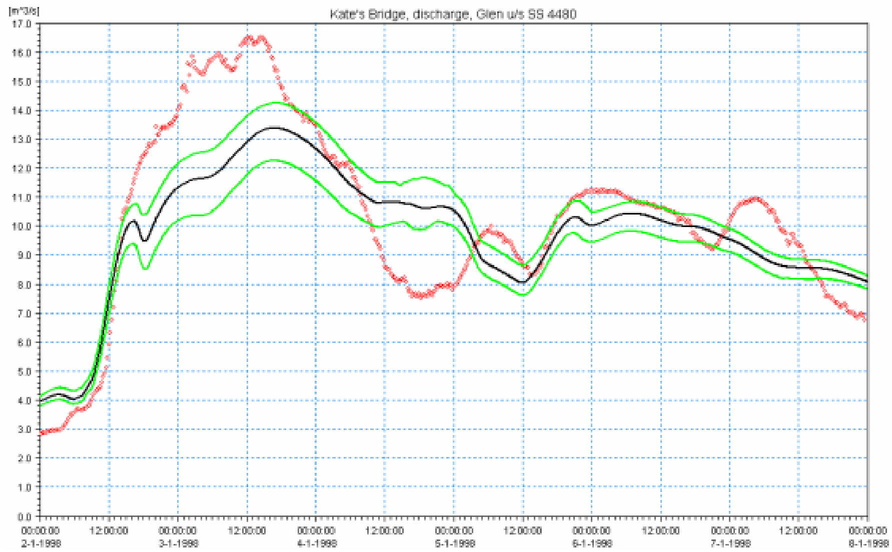

Figure 8 The simulated discharge at Kate's Bridge. The measured discharge is shown with red diamonds.The center (black) curve is the ensemble mean and the (grey) envelope curves are the limits of the 95% confidence interval.

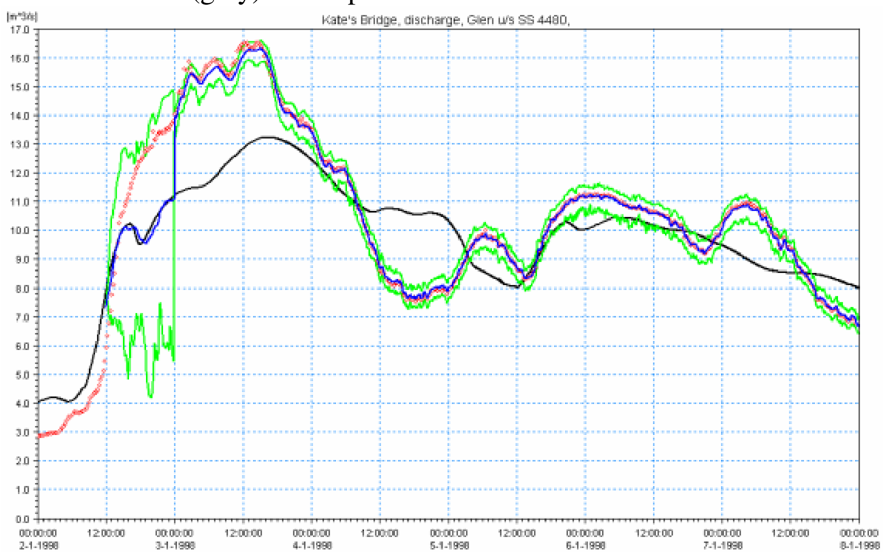

Figure 9. Assimilation of discharge data from Kate's Bridge. The measured discharge is shown with diamonds. The centre (black) curve is the original simulation (ensemble mean) for reference. Starting at 02JAN1998 12:00:00 an ensemble simulation with 20% standard deviation in runoff errors is shown. Assimilation is then initiated 12 hours later. The (grey) envelope curves are the limits of the 95% confidence interval.

The Ensemble Kalman Filter framework is currently being applied to this highly developed and complex river system. Figure 8 illustrates the use of the uncertainty framework for uncertainty propagation. Uncertainty due only to precipitation is propagated through the model assuming a standard deviation of the precipitation uncertainties of 10%. The ensemble mean and the 95% confidence intervals for the simulation are shown. It is clear from this figure that the uncertainty in rainfall in this case is not sufficient to explain the discrepancies between the simulated and observed values. Other sources of uncertainty such as biases in the rainfall, ungauged inflows or uncontrolled structure operation may play a role in this case.

The impact of data assimilation on the forecast and on the uncertainty bounds is illustrated in Figure 9. Initially (the first 12 hours) a deterministic simulation is shown. Over the next 12 hours the impact of assuming 20% standard deviation in the lateral inflows is shown with the corresponding confidence intervals. Finally data assimilation is applied updating the model using the observed data. The model simulations are dramatically improved and the estimated uncertainty is also significantly reduced. The FLOODRELIEF decision support system has been designed to use the results from this uncertainty framework so this and other ensemble-based forecasts can be used in the flood forecasting management (Butts et al. 2005)

CONCLUSIONS

Uncertainty is inherent in the forecasting process. Quantifying this uncertainty can provide important information to the decision making processes associated with flood forecasting. However, because of the detrimental and often catastrophic impacts of flooding accurate forecasts are required. In this paper we present a stochastic framework that allows both uncertainty estimation and data assimilation based on the ensemble Kalman filter.

The framework developed can be used to evaluate the impact of uncertainties in the hydrological and meteorological boundaries on hydrological and hydraulic simulations of flow and water level. A series of investigations have been carried out to evaluate the efficiency of this framework and to determine the sensitivity of the filter formulation to a range of parameters and conditions. Only a part of these results have been presented here but these data will provide guidelines for use in the actual forecasting studies. Intuitively we expect updating of both the catchment model and river models to improve forecasting accuracy as indicated here but further investigations are required. This framework is evaluated for two case studies; US NWS study catchment, the Blue river basin and the UK FLOODRELIEF study catchment, the Welland and Glen. One of the results of performing this evaluation has been to highlight the fact that one of the major outstanding problems in uncertainty estimation is the accurate quantitative characterisation of the sources of uncertainty in hydrological inputs and measurements.

REFERENCES

Beven K.J., (2000): *Rainfall-Runoff Modelling: The Primer*. John Wiley & Sons, England.

Boyle, D.P., Gupta, H.V., Sorooshian, S., Koren, V., Zhang, Z., Smith, M. (2001), *Toward improved streamflow forecasts: value of semidistributed modeling*. Water Resources Research 37(11), pp. 2749–2759.

Butts, M.B., Klinting, A., van Kalken, T., Cadman, D., Fenn, C. and Høst-Madsen, J. (2001) *Design and development of an internet-based flood forecasting system using real-time rainfall, radar, and river flow data*. Proc. of River Basin Management 2001, Cardiff, (R.A. Falconer and W.R. Blain, Eds.) WIT Press, pp. 139-148

Butts M.B., Hoest-Madsen, J. and J.C. Refsgaard J.C: (2002), *Hydrologic Forecasting*, Encyclopaedia of Physical Science and Technology 3rd Edition, 2002.

Butts, M.B. and Khatibi, R., (2003), *Workpackage 7: Operational End-user Requirements – General Functionality Deliverable 7.1 –Part 1*, (FLOODRELIEF project) [WWW] http://projects.dhi.dk/floodrelief/

Butts, M. B., J. T. Payne, M. Kristensen and H. Madsen. (2004). *An evaluation of the impact of model structure and complexity on hydrological modelling uncertainty for streamflow prediction*. Journal of Hydrology, 298, pp 242-266.

Butts, M. B., Klinting, A. ,Ivan, M., Larsen, J.K., Brandt, J., Christiansen, J.H., Skjøth C.A., Frohn, L.M. Geels, C., Mengelkamp, H-T., Pestel, J., (3),. Johnsen, K-P, Cluckie, I.D:, Han D., Xuan Y.Q., Dubicki, A., Slomska A., Szalinska, W., Strońska, K., Zawiślak T., Lewandowski, A., Olszewski, A., Kolerski, T., Khatibi, R., Price D. (2005), *The FLOODRELIEF internet-based flood forecasting decision support system*, Proceedings of Innovation, advances and implementation of flood forecasting technology, 17-19 October, Tromsø, Norway

Cadman, D, D. Price and M.B. Butts (2005) *Flood Forecasting in the Anglian Region: User-driven Development towards Forecasting Flood Risk*. In: Flood Risk Management in Europe: Innovation in Policy and Practice, Advances in Natural and Technological Hazards Research Series, accepted for publication.

Evensen, G. (1994), *Sequential data assimilation with a nonlinar quasi-geostrophic model using Monte Carlo methods to forecast error statistics*, Journal of Geophysichal Research, 99 (C5), pp. 10143-10162.

Falk., A.K., Butts, M.B., Madsen, H., Hartnack, J., Xuan, Y., Cluckie, I.D., Han, D. (2005a), *Workpackage 6: Uncertainty Framework Deliverable 6.1*, (FLOODRELIEF project) [WWW] http://projects.dhi.dk/floodrelief/

Falk., A.K., Butts, M.B., Hartnack, J., Madsen, H.. (2005b), *Workpackage 6: Prototype Implementation of Uncertainty Framework Deliverable 6.2*, (FLOODRELIEF project) [WWW] http://projects.dhi.dk/floodrelief

Houtekamer, P. L., Michell, L. (1997), *Data Assimilation Using an Ensemble Kalman Filter Technique*, Monthly Weather Review, Vol. 126

Madsen, H and Canizares, R. (1999): *Comparison of extended and ensemble Kalman Filters for data assimilation in coastal area modelling*, Int. J. Num. Meth. Fluids 31, pp. 961-981.

Havnø, K., Madsen, M. N and Dørge J. (1995), *MIKE 11 – A Generalized River Modelling Package,* In: Computer models of Watershed Hydrology, Singh, V.P. Ed., Water Resources Publications, Colorado, USA, pp. 733-782.

Hartnack, J and Madsen, H. (2001), *Data assimilation in river flow modelling*, Proceedings of the 4th DHI Software Conference, Helsingør, Denmark, 2001.

Madsen, H. (2000), *Automatic calibration of a conceptual rainfall-runoff model using multiple objectives*, Journal of Hydrology, Volume 235(3-4), pp. 276-288.

Madsen, H., Butts, M.B., Khu, S.T. and Liong, S.Y. (2000), *Data assimilation in rainfall-runoff forecasting*, Hydroinformatics 2000, 4th International Conference on Hydroinformatics, Cedar Rapids, Iowa, USA, 23-27 July, 2000.

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, 2003. pp. 75–81.

Price, D., Strońska, K., Zawiślak, T., Butts, M.B., and Klinting, A. (2003), *Workpackage 7: Operational End-user Requirements – User Interface Functionality Deliverable 7.1 –Part 2*, (FLOODRELIEF project) [WWW] http://projects.dhi.dk/floodrelief/

Rajaram, H. and Georgakakos, K.P. (1989). *Recursive parameter estimation of hydrologic models*. Water Resour. Res., 25(2), pp. 281–294.

Refsgaard, J.C. (1997). *Validation and intercomparison of different updating procedures for real-time forecasting*. Nordic Hydrology 28, pp. 65–84

Xuan, Y., Cluckie, I.D. and Han, D. (2005) *Uncertainties in Application of NWP-Based QPF in Real-Time Flood Forecasting*. Proceedings of Innovation, advances and implementation of flood forecasting technology, 17-19 October, Tromsø, Norway