Deltares

Enabling Delta Life

2009.06.11 Improving Operational Flood Forecasting through Data Assimilation

Flood Control 2015

Deltares

Title
2009.06.11 Improving Operational Flood Forecasting through Data Assimilation

Project	**Reference**	**Pages**
1200379-005	1200379-005-ZWS-0005	17

Keywords
Flood Forecasting, Data Assimilation, State Updating, Uncertainty, Research Plan

Summary
The present report summarizes the research plan and progress of the PhD project "Improving Operational Flood forecasting through Data Assimilation" carried out within the Flood Control 2015 program. This project focuses on reduction and quantification of uncertainties in flood forecasting through data assimilation in distributed hydrological models. Currently, most forecast offices use a lumped hydrological model (with deterministic or manual state updating) for forecasting, but there is a clear tendency to move to distributed models for hydrological forecasting. To be ready for the future, strategies should be developed how to perform (ensemble) data assimilation using these models in a real-time setting. This requires detailed knowledge of initialisation uncertainty, model uncertainty and forcing uncertainty. Moreover, this requires detailed knowledge how these uncertainties influences the data assimilation and subsequently the (ensemble) hydrological forecast.

Version	Date	Author	Initials	Review	Initials	Approval	Initials
	dec. 2009	Albrecht Weerts	AW	Martin Ebel	ME	Toon Segeren	

State
final

Contents

1 Introduction

1.1 Background

Robust forecasts (unbiased and skilful) are vital in providing a comprehensive flood warning service to people and businesses at risk from flooding. For fluvial and storm surge flood forecasting, rainfall–runoff, flow routing, 1D-hydraulic and 2D-hydraulic models are often combined into model cascades and are run automatically in operational flood and storm surge forecasting systems.

Currently, the outputs from these models are often deterministic with one model run delivering the flood forecast which is assumed to be the best representation, although Forecasting Duty Officers assess and advise on the uncertainty in forecasts based on experience and judgement. However, it is widely known that the accuracy of flood forecasts can be influenced by a number of factors, such as the accuracy of input data, and the model structure, parameters and state (initial conditions). Having a sound understanding of these modelling uncertainties is vital to assess and improve the flood forecasting service.

Real-time flood forecasting applications often make use of a cascade of inter-linked hydrological and, in some cases, hydrodynamic models, embedded in a data-management environment such as that of FEWS Rivers Rhine and Meuse. Model cascades (or integrated catchment models) are typically run in two principal modes of operation:

i) a historical mode – in which models are forced by hydrological and meteorological observations over a limited time period prior to the onset of the forecast (e.g. to initialise model stores)

ii) a forecast mode – in which models are run over the required forecast lead time, forced by outputs from other models, with the internal model states at the end of the historic run taken as initial conditions for the forecast run

Increasingly, models are forced using meteorological forecasts of precipitation and sometimes other variables, such as air temperature (e.g. where snowmelt is an issue) and evaporation, in addition to the use of forecasts from river locations further upstream.

Figure 1.1 illustrates these different modes of operation, and how they differ from the much longer period of records which are typically used in model calibration.

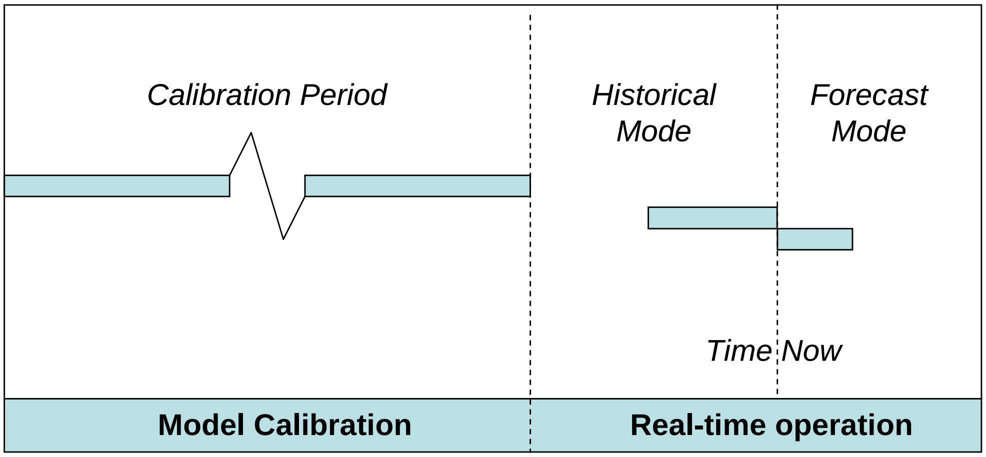

Figure 1.1 *Illustration of historical and forecast modes of operation*

In each of the steps in the model and data processing chain, uncertainties can be attributed to the model inputs, the model structure, internal model states and model parameterisation, with the total predictive uncertainty accumulating in the forecast outputs (e.g. Beven 2009; Pappenberger et al. 2007).

Depending on the lead-time at which forecasts are issued in comparison to the hydrological response time, the dominant uncertainties will lie in the inputs derived from observations, the rainfall-runoff and routing models, or, if applicable, the hydrodynamic models, and from the uncertainty in rainfall and other meteorological forecasts (if used). The various time delays in the warning process also need to be considered, such as the time taken to collect data, age of external forecasts, run models, post-process results, take decisions and issue flood warnings, as discussed later.
The process of making a flood forecast can therefore be subdivided into three problems (e.g. Moll, 1986):

- Estimation of the actual state of the basin at the start of the forecast, which consists of interception storage, soil moisture storage, groundwater storage, other possible storages (e.g. snow storage), and the water levels in rivers, reservoirs, wetlands and lakes;

- Modelling of the movement of water during the period covered by the forecast lead-time through the whole cascade of rainfall-runoff, flow routing, and hydrodynamic models;

- Forecasting of the model inputs during the selected lead-time. These can consist of meteorological inputs, but also inflows from locations further upstream or at other model boundaries (for instance tidal influences, abstractions and discharges).

For an individual model component within an integrated catchment model, these three problems result in uncertainties / errors in the forecast consisting of:

- Initialisation Errors: due to errors in the observations and their spatial and/or temporal interpolation used to estimate precipitation, potential evaporation and temperature, discharge or other boundary conditions in the historical mode of operation;

- Model(ling) Errors: arising from approximating parameterisations/model structures, uncertain model parameters, model resolution limitations, uncertain structure operating/management rules, etc.

- Forcing Errors: errors which occur in the forecast mode of operation when a model component is forced with an input derived from another model with its own Initialisation and Model Errors; for example a Numerical Weather Prediction model or the hydrological or flow routing outputs at a flow forecasting point further upstream.

Note that the initialisation errors are usually not independent from the model errors because normally a model is used to derive the estimate of the actual state of the basins; for instance via data assimilation or just driving the model (cascade) in historical mode until the start of the forecast to estimate soil moisture storage and other variables of interest.

Within an Integrated Catchment Model, these errors combine as illustrated in Figure 1.2. This figure shows the picture for the whole modelling cascade, including the hydrodynamic and coastal components (if relevant). Initialisation and modelling errors occur in each of the different components resulting in forcing errors in the downstream model. The arrows indicate the source of the forcing error, modelling error(s) and initialisation error(s) in the model(s) higher up in the model cascade.

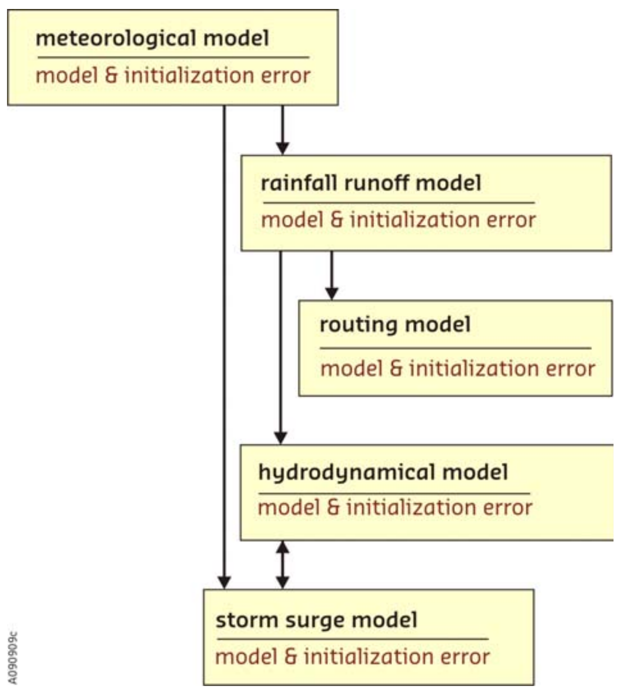

Figure 1.2 *Flood and storm surge forecast model cascade indicating the sources of the errors in the forecasts. The arrows indicate the errors in the forcing (when looking at it from the viewpoint of the receiving model cascade component) or model output (when looking at it from the viewpoint of the producing model cascade component)*

The errors in the forecast model cascade need to be quantified and/or reduced for the following reasons:

- to provide more accurate forecasts;
- to provide accurate information regarding the uncertainty of the forecast (and, if possible, unbiased and skilful estimates);

Quantification can be seen as providing a description/method to quantify the uncertainties and typically involves the use of forward uncertainty propagation techniques to give an idea of the uncertainty in the forecast. Of course, where suitable information is available, it is preferable to first reduce the uncertainties. This can be achieved by two key approaches (1) making use of recent observations (data assimilation) and (2) applying adjustments based on the historical performance of the forecasts made using the model cascade (forecast calibration).

Data assimilation is a feedback system in which the forecast is conditioned on all available information that is available at the time the forecast is made (the forecast origin or 'time now'). This includes information on the current state of the system, but also entails past performance of the forecast system (possibly further conditioned on secondary information such as the time of year, synoptic situation etc.).

Often, the term data assimilation is used to describe the use of real-time recent data to improve forecasts, whilst the term conditioning (on historical data) or forecast calibration is used to describe methods for improving forecasts based on the historical performance (i.e. not taking account of any real-time data which may be available). Alternative terms in flood forecasting include real-time updating or real-time adaptation.

Figure 1.3 shows schematically where these different approaches interact with the forcing errors, the rainfall-runoff, flow routing and hydrodynamic models, and the forecast produced by the end-to-end model cascade.

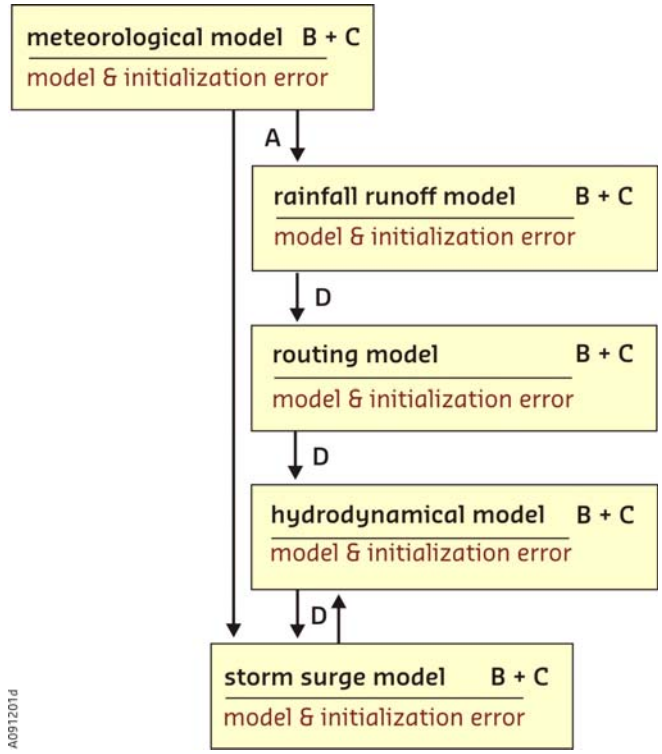

Figure 1.3 *Flood and storm surge forecast model cascade indicating the sources of the errors in the forecasts. The arrows indicate the errors in the forcing (when looking at it from the viewpoint of the receiving model cascade component) or model output (when looking at it from the viewpoint of the producing model cascade component), and where and with which methods uncertainty can be reduced and quantified in the model cascade (A) Meteorological Forecast Calibration, (B) State Updating, (C) Parameter Updating, (D) Forecast Calibration/Output Updating*

These four operational approaches to updating which are described in the figure are more general forms of the widely used terminology in hydrological forecasting of input updating, state updating, parameter updating and output updating (e.g. Refsgaard, 1997; Serban and Askew 1991).

These four approaches to uncertainty quantification and reduction are further described in (Weerts and Beckers, 2009; Weerts et al., 2010).

1.2 Description of topic and objectives of the FC2015 PhD project

This project, carried out within the FC2015 program, focuses on reduction and quantification of uncertainties in flood forecasting through data assimilation in distributed hydrological models. Currently, most forecast offices use a lumped hydrological model (with deterministic or manual state updating) for forecasting, but there is a clear tendency to move to distributed models for hydrological forecasting. To be ready for the future, strategies should be developed how to perform (ensemble) data assimilation using these models in a real-time setting. This requires detailed knowledge of input uncertainty, model uncertainty and forcing uncertainty and how these influence the data assimilation and subsequently the (ensemble) forecast.

Key goals of this project are therefore (1) to identify and quantify the sources of the input, model, and forcing uncertainties, (2) to show how these uncertainties propagate through the distributed hydrological model(s) used for forecasting, and (3) how (distributed) measured data can be used to reduce the uncertainty and how it affects the hydrological forecast (i.e. skill and bias) and (4) to determine which and how data assimilation methods can be best used in an operational forecast setting.

2 Scientific objectives

Initialisation uncertainty is caused by uncertainty in the model input and the model itself during the historical mode. Within an ensemble data assimilation environment, given the uncertainties of the model input, the uncertain model itself (i.e structure and/or parameters) and the uncertain measurements used for updating, the model states are adjusted to approximate the true state of the physical system. As a result, the initialization uncertainty will be reduced and an estimate of the uncertainty of the model forecasts can be provided.

2.1 Input Uncertainty Specification

It is clear that in ensemble data assimilation the spread in the ensemble members is determined by the specified errors in the model structure, the historical input data and discharge data. One needs to make sure that realistic assumptions for these errors are made. For average areal rainfall derived in operational flood forecasting systems (with a limited number of rainfall stations), uncertainties can vary of the order of ±0–50% (Willems, 2002), but this range is often chosen as 0–15% (E. Todini, personal communications). The use of areal precipitation estimates derived from radar might give a better handle on specifying realistic rainfall uncertainties especially when using distributed hydrological models.

Emerging Questions:
-How can the input uncertainty (during historical mode, see Figure 1.1) be specified for both lumped and distributed hydrological models?
- Can realistic error bounds be derived for radar rainfall fields?
-Is there added value of radar rainfall data versus precipitation gauge data in a data assimilation scheme using lumped and distributed hydrological models?

2.2 Model Uncertainty Specification

Hydrological models can be divided according to Wagener et al. (2004) and Dooge and O'Kane (2003) into three groups, as follows:

(1) Metric (empirical, black box): Metric models are purely based on the information derived from the data and no prior knowledge about catchment behaviour is needed. Artificial Neural Networks (ANN) and Transfer Functions (TF) represent an example of this type of models. These models are usually spatially lumped, in other words the catchment is looked up on as a single unit.

(2) Parametric (conceptual, grey box): Parametric models aim to satisfy the water balance of the catchment, which is represented by a system of reservoirs. The storages are filled by fluxes such as precipitation, infiltration, percolation and emptied through evapotranspiration, discharge, drainage. The model parameters are often not measurable in the field and needs to be derived in a calibration procedure. These models are either lumped, quasi-lumped (i.e. segmentation of catchment onto smaller sub-catchments) or fully distributed.

(3) Mechanistic (physically based, white box): Mechanistic models are described by the conservation of mass, momentum and energy equations. They became practically applicable in 1980s with computer power availability. Nevertheless, data demand,

scale-related problems and over parameterization are drawbacks of mechanistic models.

Historical overview of rainfall runoff models ranging from empirical to physically models is given by Todini (2007). A detailed description of more than 20 hydrological models, which are used all over the world is summarized in Singh (1995).

Currently, most forecast offices use lumped parametric models (PDM, HBV-96, SAC-SMA), but there is a clear tendency to move towards fully distributed parametric models (like PDM-G2G, HLRDM-SACSMA, LARSIM, WASIM, etc). Until now the operational data assimilation schemes used are either manual or deterministic. The question is how ensemble data assimilation can be facilitated using these types of distributed models. Model uncertainty must be taken into account and the model uncertainty should reflect the uncertainty given the knowledge of the modeled area and model structure used.

Emerging questions:
- How can the model uncertainty (given a model calibration) be derived and specified for both lumped and distributed hydrological models?
- How can the model uncertainty be formulated separately from initialisations uncertainty?
- Can a generic process be formulated to derive the model uncertainty?

2.3 Data Assimilation Methods

Data assimilation is a technique to merge measurements of any type with estimates from geophysical models (Reichle, 2008). It can be seen as an update of the model state with externally measured variables (Pauwels and De Lannoy, 2006; Clark et al., 2008), which also quantifies the errors, uncertainties in input data, model structure and observations (Weerts and El Serafy, 2006; Salamon and Feyen, 2009; Clark et al., 2008).

The data assimilation methods used in hydrology can be divided into two classes, as follows: (1) sequential methods and (2) variational methods. The most well known sequential method is the Kalman filter. The Kalman filter (KF) is originally developed for linear systems (Kalman, 1960). Since the hydrological processes are rather non-linear, it further developed into the Extended Kalman filter (EKF), e.g. (Georgakakos,1986). The major drawbacks of Extended Kalman Filter (EKF) are the high computational demand for the propagation of the background error covariance (especially for large system state vectors), and the neglection of higher order derivatives for the background error covariance propagation and the mapping of the observational information (the observed discharge) to the model state variables (Pauwels and De Lannoy, 2009; Salamon and Feyen, 2009). The Ensemble Kalman filter (EnKF) (Eversen, 2003; Evensen, 2009) represents another variation of KF. The EnKF propagates an ensemble of model realizations (generated from model perturbations) through time, and estimates the background error covariance matrix from the ensemble statistics. EnKF is computationally efficient, but is limited to Gaussian distributions (Weerts and El Serafy, 2006;Pauwels and De Lannoy, 2009; Salamon and Feyen, 2009). Particle Filter (PF) is another form of a recursive Bayesian filter based on Monte Carlo simulation. Particles, associated with weights, are used to approximate the posterior probability distribution functions (Weerts and El Serafy, 2006). The advantage of this technique is that no assumptions on the form of the prior probability density function (pdf) of the model states are necessary and that the full prior pdf is being used, in contrast to EnKF. In theory, this would mean that particle filtering is more sensitive to the tails of the prior pdf, a property which

maybe of vital importance in flood forecasting, although this maybe at the cost of a lot more simulations (Weerts and El Serafy, 2006).
Variational methods have been widely used in data assimilation for numerical weather prediction as a means of dealing with a very large number of observations to be assimilated in a computationally efficient way. The technique depends on defining the adjoint model, which provides local gradient terms for any predicted variable that can be matched to an observable. These gradients will vary in space and time, depending on the nonlinearity of the model. Linear extrapolation is then used to adjust model predicted variables towards the observed values, depending on an estimate of the covariance matrix. In this, it is similar to the Extended Kalman Filter(EKF) but, unlike the EKF, does not update the covariance matrix as the data assimilation proceeds. A hydrological forecasting applications in which variational methods are investigated was presented by Seo et al. (2003, 2009).

Emerging questions:
- Which method (sequential or variational) is more suitable for operational flood forecasting both in terms of hydrological performance and in terms of operational performance? (This work maybe carried out together with NOAA-NWS)
- Which method is best in handling delays between modelled states and streamflows measurements (rainfall-runoff+routing)?

2.4 Operational Forecasting

During the forecast mode (see Figure 1.1), there are no measurements to adjust the states to approximate the true state of the modelled system anymore. The ensemble forecast resulting from the ensemble data assimilation also gives an estimate of the forecast uncertainty. Given proper error specification the initialisation uncertainty, the model uncertainty and the forcing uncertainty the forecasted error band should result in meaningful probabilities (tested over longer periods).

Emerging questions:
- How can forcing uncertainty (forecast mode) be specified?
- How sensitive are the resulting forecasts for the specification of input and model uncertainties? Does the data assimilation using proper uncertainty specifications result in unbiased model forecasts?
- Or is forecast calibration of the forecast made using data assimilation still necessary?

3 Progress 2009

The FC2015 project named “Operational Data Assimilation using Distributed Hydrological Models” started April 2009. To be able to reach the key and scientific goals this project was formulated as a PhD project under the FC2015 project. The project is carried out as a joint research project between Deltares and Wageningen University. Supervisor of the project is prof.dr.ir. Remko Uijlenhoet (WUR, chair of the Hydrology and Quantitative Water Management Group) and co-supervisor is dr.ir. Albrecht Weerts (Deltares, Inland Water Systems-Operational Water Management).The cooperation is formalized in a contract between Deltares and WUR. A detailed preliminary research proposal has been put together.

In July/August 2009, the PhD position was advertised. Due to the holiday period, the selection period was postponed to early September 2009. Oldrich Rakovec (MSc) started the PhD project November 1 2009. Since November 2009, the PhD candidate worked on improving the preliminary research proposal reported here which will be finalized 6 months after the appointment. A detailed organization and planning schedule of the PhD project is available in Chapter 4. The PhD candidate is also involved in the research school SENSE.

4 Organization & Planning

4.1 Matching relations

4.1.1 Deltares

This FC2015 PhD project (2009.06.11-1200379) has relations with the following two Deltares projects

Project	Titel	PhD candidate	Project Leader
1200322.009	PhD Postprocessing hydrological forecasts	J. Verkade (TUD)	P. Reggiani
1200433.003	PhD Real-time decision support in water systems under uncertainty	L. Raso (TUD)	D. Schwanenberg

4.1.2 Wageningen University

Prof. dr. ir. Remko Uijlenhoet (Wageningen University) will function as promotor and therefore the PhD student will obtain his title from Wageningen University. Prof. dr. ir. Remko Uijlenhoet holds the Hydrology and Quantitative Water Management chair.

4.2 Planning and Reviewing

4.2.1 Review

Name	Role
Uijlenhoet	promotor
Weerts	co-promotor
PhD committee	PhD committee

4.2.2 Planning

Phase	Product	By	Quality control	Approved	Accepted by
Phase 1	PhD research plan	PhD Student	Uijlenhoet/Weerts	Uijlenhoet/Weerts	Uijlenhoet/Weerts
Phase 2	Article 1 Input Uncertainty	PhD Student	Uijlenhoet/Weerts	Peers	Journal
Phase 3	Article 2	PhD	Uijlenhoet/Weerts	Peers	Journal

Phase	Product	By	Quality control	Approved	Accepted by
	Model uncertainty	Student			
Phase 4	Article 3 Forcing Uncertainty/ Operational Forecasting	PhD Student	Uijlenhoet/Weerts	Peers	Journal
Phase 5	Article 4 Case Studies	PhD Student	Uijlenhoet/Weerts	Peers	Journal
Phase 6	PhD Thesis	PhD Student	Uijlenhoet/Weerts	Uijlenhoet/Weerts	PhD Committee

4.2.3 Timeline

The PhD project lasts 4 years (48 months), the following schedule is proposed

Phase 1: start – 6 months;

Phase 2-5 : 6 - 42 months (4*9 months);

Phase 6: 39-48 months;

In the beginning of the research project the Ourthe catchment (tributary Meuse) is used as study area because of the data availability (see Appendix A). Later in the project, other case studies might be considered. The Ourthe case study will at some point also be included in the Demonstrator Flood Control (using only data that is available operationally).

5 Literature

BEVEN, K J (2009) Environmental Modelling: An Uncertain Future?, Routledge: London.

CLARK M.P., D. E. RUPP, R.A. WOODS, X. ZHENG, R. P. IBBITT, A. G. SLATER, J. SCHMIDT, M.J. UDDSTROM (2008), Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model, Advances in Water Resources 31,1309–1324.

DOOGE, I., J. C., AND J.P. O'KANE (2003). Deterministic Methods in Systems Hydrology. A.A.Balkema Publishers.

EVENSEN, G. (2009). Data Assimilation: The Ensemble Kalman Filter. Springer.

EVENSEN, G. (2003). The ensemble kalman filter: theoretical formulation and practical implementation, Ocean Dynamics, 53 , 343-367.

GEORGAKAKOS, K. P. (1986). A generalized stochastic hydrometeorological model for flood andflash-flood forecasting. Water Resources Research, 22 (13), 2083-2095.

MOLL J.R. (1986) Short range Flood Forecasting on the River Rhine. In: River Flow Modelling and Forecasting, Water Science and Technology Library, D. Reidel Publishing Company,Kluwer, Dordrecht, Eds Kraijenhoff and Moll

PAPPENBERGER, F., BEVEN, K.J., FRODSHAM, K., ROMANOVICZ, R. AND MATGEN, P. (2007). Grasping the unavoidable subjectivity in calibration of flood inundation models: a vulnerability weighted approach. Journal of Hydrology, 333, 275–287.

PAUWELS, V. R. N., AND G.J.M. De LANNOY (2009). Ensemble-based assimilation of discharge into rainfall-runoff models: A comparison of approaches to mapping observational information to state space. Water Resources Research, 45 , 1-17.

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

REICHLE, R. H. (2008). Data assimilation methods in the Earth sciences. Advances In Water Resources, 31 (11, Sp. Iss. SI), 1411-1418.

SALAMON, P., and L. FEYEN (2009). Assessing parameter, precipitation, and predictive uncertainty in a distributed hydrological model using sequential data assimilation with the particle filter. Journal of Hydrology, 376 , 428-442.

SEO, DJ., L. CALINA, R. CORBY, T. HOWIESON (2009) Automatic state updating for operational streamflow forecasting via variational data assimilation, Journal of Hydrology, Volume: 367 Issue: 3-4 Pages: 255-275.

SEO, D. J., V. KOREN, V., and N. Cajina (2003). Real-time variational assimilation of hydrologic and hydrometeorological data into operational hydrologic forecasting. Journal of Hydrometeorology, 4 (3), 627-641.

SERBAN P, ASKEW A J (1991) Hydrological forecasting and updating procedures In: Hydrology for the Water Management of Large River Basins, IAHS Publication No. 201.

SINGH, V. P. (1995). Computer models of watershed hydrology. Highlands Ranch, CO : Water Resources Publications.

TODINI, E. (2007). Hydrological catchment modelling: Past, present and future. Hydrology and Earth System Sciences, 11 (1), 468-482.

WAGENER, T., H.S. WHEATER, and H.V. GUPTA (2004). Rainfall-Runoff modelling in Gauged and Ungauged Catchments (, Ed.). Imperial College Press.

WEERTS, A.H. AND J. BECKERS (2009). 2009.06.01 Uncertainty Framework for Operational and Storm Surge Forecasting, Deltares / FC2015 Research Report, 1200379.001.

WEERTS, A.H. AND EL SERAFY, G.Y.H., 2006. Particle filtering and Ensemble Kalman filtering for state updating with conceptual rainfall–runoff models. Water Resources Research, doi:10.1029/2005WR004093.

WEERTS A H, SEO D J, WERNER M, SCHAAKE J (2010). Operational hydrological ensemble forecasting. To appear in Applied uncertainty analysis for flood risk management edited by K Beven and J Hall.

WILLEMS, P. (2002). Stochastic description of the rainfall input errors in lumped hydrological models. Stochastic environmental research and risk assessment, 15 (2), 132-152.

A Overview Available Data Ourthe

The Ourthe River is a tributary of the Meuse River. Most of the available data has been provided by MetSethy and KMI to Wageningen University and is only available for this PhD project.

Data availability Ourthe:

Daily:

Streamflow: 1987-2005, gauges in Ortho, Tabreux and Mabompre
Precipitation: 1968-2005 gauge in St. Hubert
Temperature (Tavg, Tmax, Tmin): 1990-2009 in Rochefort
Potential evapotranspiration: 1990-2005 in St. Hubert and Rochefort, 2003-2006 in Humain

Hourly:

Streamflow: 1990-2005 for river gauges at Tabreux, Nisramont, Mabompre, Ortho, 1992-2005 at Hotton, 1995-2005 at Durbuy.
Precipitation: 1990-2005 for rain gauges in Ortho, Ouffet, Marche, Bastogne,
1995-2005 in Rechamps and Tailles, 1997-2005 in Somme-Leuze, 1998-2005 in St. Hubert,1998-2005 in Flamierges, 2000-2005 in Erezee
Temperature:1990-2009 in St. Hubert, 2003-2009 in Humain

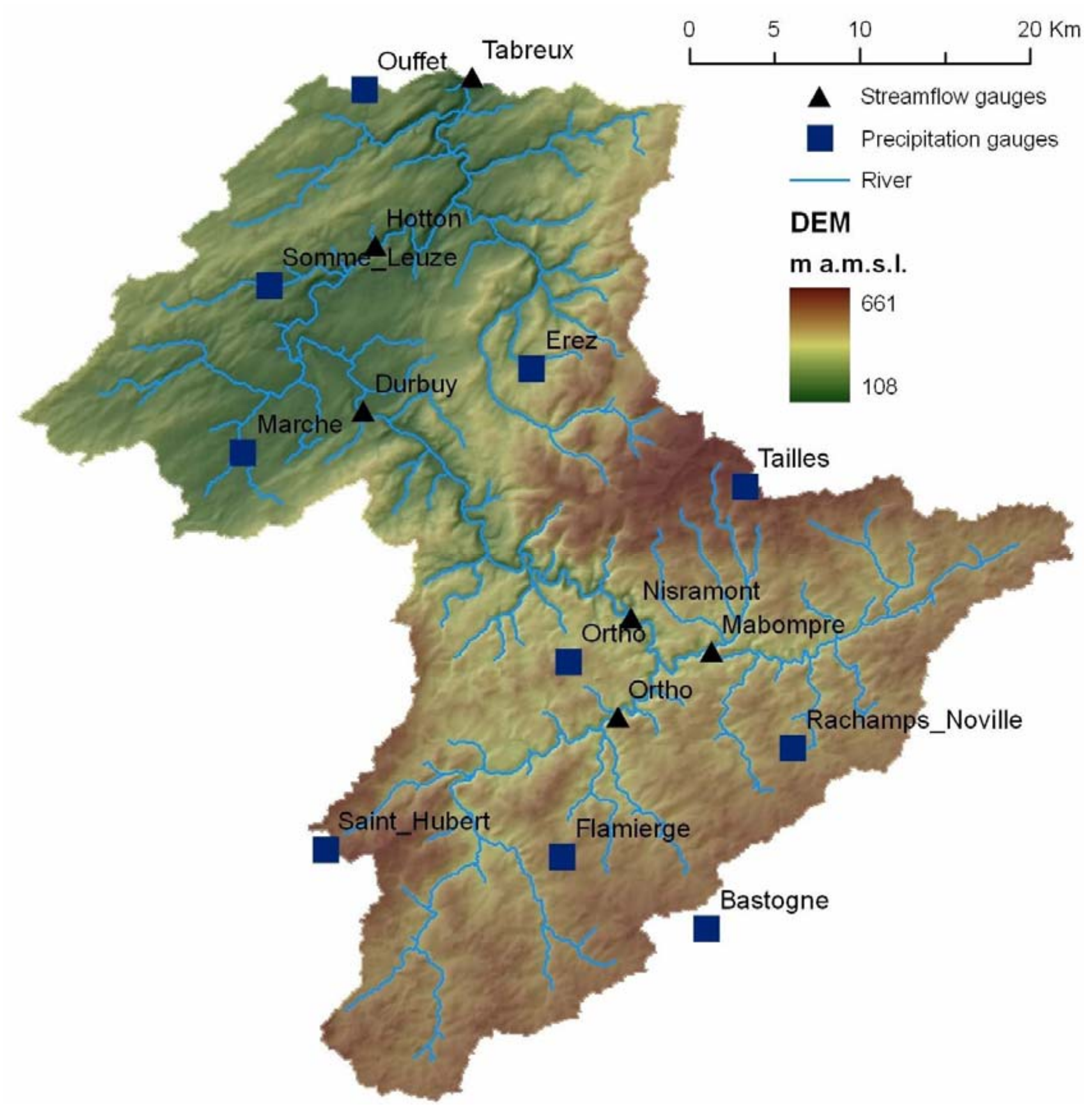

Figure A.1 *Digital elevation model for the Ourthe catchment and precipitation gauges*