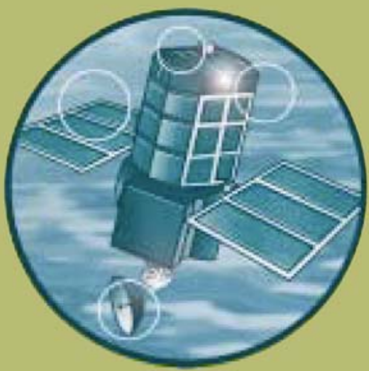
Joint Defra / Environment Agency Flood and Coastal Erosion Risk Management R&D Programme

Risk-based probabilistic fluvial flood forecasting for integrated catchment models – Phase 1 Report

Science Report: SC080030/SR1



Product code: SCHO0809BQWP-E-P

Environment Agency

defra
Department for Environment Food and Rural Affairs

The Environment Agency is the leading public body protecting and improving the environment in England and Wales.

It's our job to make sure that air, land and water are looked after by everyone in today's society, so that tomorrow's generations inherit a cleaner, healthier world.

Our work includes tackling flooding and pollution incidents, reducing industry's impacts on the environment, cleaning up rivers, coastal waters and contaminated land, and improving wildlife habitats.

This report is the result of research commissioned by the Environment Agency's Science Department and funded by the joint Environment Agency/ Defra Flood and Coastal Erosion Risk Management Research and Development Programme.

Published by:
Environment Agency, Rio House, Waterside Drive, Aztec West, Almondsbury, Bristol, BS32 4UD
Tel: 01454 624400 Fax: 01454 624409
www.environment-agency.gov.uk

ISBN: 978-1-84911-097-6

Author(s):
K. Sene, A. Weerts, K. Beven, R.J. Moore, C. Whitlow, P. Young

Dissemination Status:
Released to all regions
Publicly available

Keywords:
probabilistic, flood, forecasting, ensemble, catchment, rainfall runoff, hydrodynamic

Research Contractor:
Atkins Ltd, Woodcote Grove, Ashley Road, Epsom, Surrey, KT18 5BW

Environment Agency's Project Manager:
Stefan Laeger, Rio House, Waterside Drive, Aztec West, Almondsbury, Bristol, BS32 4UD

Collaborator(s):
Deltares, Lancaster University, CEH Wallingford, Edenvale Young

Science Project Number:
SC080030/SR1

Product Code:
SCHO0809BQWP-E-P

Science at the Environment Agency

Science underpins the work of the Environment Agency. It provides an up-to-date understanding of the world about us and helps us to develop monitoring tools and techniques to manage our environment as efficiently and effectively as possible.

The work of the Environment Agency's Science Department is a key ingredient in the partnership between research, policy and operations that enables the Environment Agency to protect and restore our environment.

The science programme focuses on five main areas of activity:

- **Setting the agenda**, by identifying where strategic science can inform our evidence-based policies, advisory and regulatory roles;
- **Funding science**, by supporting programmes, projects and people in response to long-term strategic needs, medium-term policy priorities and shorter-term operational requirements;
- **Managing science**, by ensuring that our programmes and projects are fit for purpose and executed according to international scientific standards;
- **Carrying out science**, by undertaking research – either by contracting it out to research organisations and consultancies or by doing it ourselves;
- **Delivering information, advice, tools and techniques**, by making appropriate products available to our policy and operations staff.

Steve Killeen

Head of Science

Executive summary

Robust forecasts are vital in providing a comprehensive flood warning service to people and businesses at risk from flooding. For fluvial flood forecasting, rainfall–runoff, flow routing and hydraulic models are often combined into model cascades and are run automatically in the Environment Agency's National Flood Forecasting System (NFFS).

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 that the Environment Agency provides.

This report describes the findings from Phase 1 of the project 'Risk-Based Probabilistic Fluvial Flood Forecasting for Integrated Catchment Models', whose main aim is to develop and test practical probabilistic methods to quantify and, where possible, reduce uncertainties around fluvial flood forecasts from sources other than predicted rainfall. The project started in November 2008 and will complete in late 2010. The main objectives of Phase 1 were to perform the following tasks in order to better define the methods and case studies to be investigated during Phase 2 of the project:

- Task 1.1 – To review current Environment Agency (regional and national) and international experience in addressing uncertainties associated with fluvial flood forecasting and consult key stakeholders to refine user requirements.
- Task 1.2 – To review and investigate which additional sources of uncertainty should be considered to gain a fuller (quantified) understanding of uncertainties in the flood forecasting process and to define in which situations/scales this may be beneficial. Particular focus should be placed on the aspects other than rainfall uncertainty, such as uncertainty associated with rainfall–runoff, routing and hydraulic components.
- Task 1.3 – To recommend and test suitable techniques for the probabilistic treatment of the most important sources of uncertainty and combine them into a high-level unified, scalable framework for integrated catchment models.
- Task 1.4 – To investigate the requirements, possibilities and benefits of real-time/state updating of probabilistic hydraulic/hydrological models and the value of different types of data (historical and real-time) in constraining uncertainties.

The consultations took place during December 2008 and January 2009 and involved more than 25 regional and national staff. More than 20 catchments were suggested as potential case studies for the project, and the discussions suggested that the key sources of uncertainty which it would be useful to consider include catchment averaging of raingauge data, the validity of rating curves, and the calibration of rainfall–runoff models. Much useful background material was also obtained on recent experience with integrated catchment models, and ongoing regional and national studies into model uncertainty.

Based on the review and consultations, and a workshop held on 16 March 2009, this report sets out the key proposals for Phase 2 of the project, including the proposed uncertainty framework, and the methods to be developed and tested on the case studies. The high-level version of the framework, which is described in this report, uses the following seven items as key decision points in selection of an appropriate

uncertainty estimation technique: level of risk, lead time requirement (linked to catchment response time), types of models, sources of uncertainty, data assimilation, operational requirements, and model run times. The detailed version of the framework, to be developed during Phase 2, will consist of flowcharts, decision trees and other formats.

The project also has the scope to investigate up to four case studies, consisting of two integrated catchment models, and two simpler models which form the basic building blocks of more complicated models. The following four catchments have been selected for study during Phase 2:

- Upper Calder (rapid response catchment)
- Lower Eden (flow routing reach)
- Ravensbourne (integrated catchment model)
- Upper Severn (integrated catchment model).

The test configurations for these catchments will be trialled on the Environment Agency's NFFS, using the following uncertainty estimation techniques:

- Forward Uncertainty Propagation – specification of a range, or ensemble, of values.
- Data Assimilation – Kalman Filter/Data Based Mechanistic, Ensemble Kalman Filter.
- Conditioning – post-processing of outputs using quantile regression and Bayesian Model Averaging.

The following four approaches to reducing model run times were also reviewed (and will be investigated further during Phase 2): increased computing power, hydrodynamic model reconfiguration, statistical sampling and model emulators. The Lower Eden case study will be used as a test bed for some of these studies, and will also be used for a short investigation into probabilistic inundation mapping using the existing mapping functionality in NFFS.

Contents

List of tables and figures

1. Introduction

1.1 Background

Robust forecasts are vital in providing a comprehensive flood warning service to people and businesses at risk from flooding. For fluvial flood forecasting, rainfall–runoff, flow routing and hydraulic models are often combined into model cascades and are run automatically in the Environment Agency's National Flood Forecasting System (NFFS).

The outputs from these models are currently 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 that the Environment Agency provides.

This R&D project will develop and test practical probabilistic methods to quantify and, where possible, reduce uncertainties around fluvial flood forecasts from sources other than predicted rainfall (which is already being addressed within the 'Hydrological Modelling with Convective Scale Rainfall' EA R&D project). This will provide an overarching framework for assessing uncertainties in fluvial forecasting in a risk-based manner which, for completeness, will also provide the possibility to include rainfall forecasting uncertainty. Some of the specific project objectives are:

- to review current experience and consult key stakeholders to refine user needs;
- to recommend and test suitable techniques for the probabilistic treatment of the most important sources of uncertainty and combine them into an overarching uncertainty framework;
- to assess the possibilities and benefits of real-time/adaptive updating for probabilistic hydraulic/hydrological models;
- to demonstrate and validate the suggested techniques for linked forecasting models through case studies in NFFS;
- to recommend and investigate alternative ways of reducing run times for probabilistic flood forecasts;
- to provide updated guidance on probabilistic fluvial flood forecasting and develop an implementation plan.

The main output from this project will be up-to-date practical guidance on how to use probabilistic techniques in fluvial flood forecasting in order to interpret possible uncertainties around flood forecasts and ultimately allow for improved management of flood events. This will be supplemented by a number of practical case studies which will demonstrate how certain uncertainty techniques can add value to the forecasting process. The main focus of this project is to demonstrate the practicality and benefits of applying probabilistic techniques to fluvial flood forecasting models which are sufficiently robust to be considered for use in an operational environment.

1.2 Scope of report

The project started in December 2008 and will complete in late 2010, and includes the following three main phases:

Phase 1 – review, consultation and scoping (4–5 months)

Phase 2 – development of suitable uncertainty framework and application to case studies (11–12 months)

Phase 3 – best practice guidance (7–8 months).

The research contractors for the project are Atkins (lead), Deltares, Lancaster University, CEH Wallingford and Edenvale Young. This report summarises the work performed during Phase 1 of the project, and includes recommendations for the case studies and techniques to be considered during Phase 2. The main tasks within Phase 1 are shown in Table 1.1.

Table 1.1 Summary of tasks within Phase 1 of the project.

Task	Description
1.1	To review current Environment Agency (regional and national) and international experience in addressing uncertainties associated with fluvial flood forecasting and consult key stakeholders to refine user requirements
1.2	To review and investigate which additional sources of uncertainty should be considered to gain a fuller (quantified) understanding of uncertainties in the flood forecasting process and to define in which situations/scales this may be beneficial. Particular focus should be placed on the aspects other than rainfall uncertainty, such as uncertainty associated with hydrologic, routing and hydraulic components
1.3	To recommend and test suitable techniques for the probabilistic treatment of the most important sources of uncertainty and combine them into a high-level unified, scalable framework for integrated catchment models
1.4	To investigate the requirements, possibilities and benefits of real-time/state updating of probabilistic hydraulic/hydrological models and the value of different types of data (historical and real-time) in constraining uncertainties

The report is based upon work performed during the first 4 months of this 2-year project, and includes the findings from the consultation exercise on the main sources of uncertainty in integrated catchment models, a description of the high-level version of the uncertainty framework which has been developed during Phase 1 of the project, and detailed plans for the case studies and Phase 2 of the project. The report also provides updates to reviews of probabilistic flood forecasting techniques from previous Defra/Environment Agency studies, and some additional studies since the time of those reviews, and provides more detail on data assimilation techniques, model run-time issues, and the propagation of uncertainty in integrated catchment models.

1.3 Layout of report

The remainder of this report is presented as follows:

- Section 2 – **Review and consultations** – describes international and Environment Agency research into probabilistic forecasting, and summarises the main findings from the consultation exercise, which involved 25 regional and 2 area staff and was performed over the period 18 December 2008 to 16 January 2009.
- Section 3 – **Sources of uncertainty** – describes the main sources of uncertainty in integrated catchment models, model-specific considerations, and some possible criteria for risk-based selection of modelling approaches.
- Section 4 – **Uncertainty framework** – describes the main techniques and issues which will need to be considered in the uncertainty framework to be developed during Task 2.1 of the project, and which were considered as possibilities to include in the case studies during Phase 2. A high-level (conceptual) version of the framework is also presented.
- Section 5 – **Phase 2 Recommendations** – presents recommendations for the techniques and case studies to be considered in Phase 2 of the project, and for further development of the uncertainty framework.

In particular, the following tables provide a summary of the case studies and methods to be considered during Phase 2 of this project:

- Table 4.6 – summarises the methods which have been reviewed in this report.
- Table 5.10 – groups the methods in Table 4.6 by type, and summarises the methods which will be considered in Phase 2 of this project.
- Table 5.11 – summarises the sources of uncertainty to be considered during Phase 2 of this project.

The final figure in Section 5 (Figure 5.2) also summarises the case studies which will be considered, and the methods which will be applied to those case studies.

2. Review and consultations

This section describes the main findings under Task 1.1 of the project, whose main aims were:

> *To review current Environment Agency (regional and national) and international experience in addressing uncertainties associated with fluvial flood forecasting and consult key stakeholders to refine user requirements*

Section 2.1 describes experience from recent international research projects and pre-operational testing of probabilistic flood forecasting systems, and includes a brief summary of research into uncertainty estimation within the FRMRC, FREE and FLOODsite research programmes, which have all benefited from Environment Agency funding, and the international HEPEX, COST731, MAP D-PHASE and EFFS projects. Section 2.2 then discusses recent and ongoing Defra/Environment Agency projects and summarises the main findings from the consultation exercise.

2.1 International experience

2.1.1 International research programmes

FRMRC – Probabilistic Flood Forecasting

The first phase of the Flood Risk Management Research Consortium (FRMRC1) ran from 2004 to 2008 and the research programme was divided into nine research priority areas.

The main area which considered real-time flood forecasting was Research Priority Area 3 'Real-Time Flood Forecasting', led by Professor Ian Cluckie. The topics that were considered included weather radar, catchment modelling, artificial intelligence techniques for flood forecasting, and data assimilation and probabilistic flood forecasting.

The data assimilation and probabilistic flood forecasting work was carried out mainly at Lancaster University by Professor Peter Young and Dr Renata Romanowicz, leading to User Report 5 (Young *et al.* 2006) that is available from the FRMRC website (http://www.floodrisk.org.uk/). The main application area for the work was a cascade of adaptive Data Based Mechanistic (DBM) flood forecasting models for the River Severn as reported in a number of scientific publications (Romanowicz *et al.* 2006b, 2008; Beven 2009).

Probabilistic flood inundation modelling was also studied in FRMRC1 under Research Priority Area 9 'Risk and Uncertainty'. Following earlier work carried out in the European Flood Forecasting System (EFFS) project, which is described later (De Roo *et al.* 2003; Pappenberger *et al.* 2005a), this work made use of the GLUE methodology (see Section 4.4) in the sensitivity analysis and calibration of flood inundation models against historical inundation information (e.g. Pappenberger *et al.* 2005a,b, 2006b, 2007a, 2007b, 2008).

The work in this research priority area also led to more general assessments of modelling uncertainties (Pappenberger and Beven 2006) and a decision tree for choosing uncertainty estimation methods (Pappenberger *et al.* 2006a, http://www.floodrisknet.org/methods). Professor Jim Hall at Newcastle University, partly funded by FRMRC1, also led the development of a general software package (Reframe) for formulating uncertainty estimation problems. This was demonstrated for a small part of the Thames Estuary 2100 project.

Some work related to probabilistic flood forecasting is also ongoing under the next phase of FRMRC (FRMRC2), which started in 2008. Super Work Package 1, led by Professor Garry Pender at Heriot-Watt University, is investigating ways of making hydraulic model calculations more efficient so that they can be used more easily in forecasting and dynamic flooding calculations. As part of this work, dynamic model emulation techniques (Young and Ratto 2008) are being developed using the DBM methodology developed at Lancaster University (see Beven *et al.* 2008b; Young *et al.* 2009). The Lancaster DBM flood forecasting methodology is also being applied to the River Eden in North West England, including an investigation of small, low-cost, intelligent GPRS (General Packet Radio Service) linked level sensors, with a view to constraining the uncertainty in flood forecasts (see Leedal *et al.* 2008).

In other parts of FRMRC2, a wide range of flood risk management systems is being considered, including catchment, coastal and urban systems for evaluation of flood risk and coupled atmospheric and flood forecasting for flood warning. Further information about FRMRC activities and products can be found at http://www.floodrisk.org.uk/

FRMRC – Good Practice Guidelines

An additional component of work in FRMRC2 (Work Package 1.7), which is being carried out in collaboration between Lancaster University, Middlesex University and the Environment Agency, is the development of Good Practice Guidelines for the incorporation of risk and uncertainty into flood risk management. Some topics that are being considered include issues about what different types of uncertainty estimation mean; the value of different types of data in constraining uncertainties; and how uncertainty estimates should be interpreted by policy and decision makers. The guidelines will provide a structured process to guide the interaction between scientists and stakeholders in flood risk management as a way of addressing these problems (the translationary discourse of Faulkner *et al.* 2007).

During the first phase of FRMRC, the major application areas for risk and uncertainty estimation were in flood forecasting and flood inundation prediction, including the use of uncertain data for model calibration and updating predictions during flood events (e.g. Beven *et al.* 2005; Pappenberger *et al.* 2006b, 2007a, 2007b; Romanowicz *et al.* 2006a, 2008). There are, of course, many other interesting areas of flood risk management for which uncertainty and risk are important, including:

- flood risk mapping (probabilities and consequences);
- real-time flood forecasting;
- flood frequency estimation;
- infrastructure design and assessment;
- predictions of impacts of land use change;
- coastal hazard assessment;
- joint tidal and river hazard assessment in estuaries;
- joint fluvial and pluvial assessments in urban areas.

These different types of applications might require different types of uncertainty and risk assessments. The aim in this part of FRMRC2, however, is to develop a general framework for the development of guidelines for good practice that might be applied in any of these areas of flood risk management.

The overall aim of Work Package 1.7 is therefore to develop a methodology or route map for the process of developing Guidelines for Good Practice, with a demonstration in one or more application areas. The priority area agreed with the Environment Agency is flood inundation modelling, with a case study on the River Eden at Carlisle (see also Hall and Solomatine 2008). This application has been chosen to complement the work to be performed within the present project under Tasks 2.1 and 3.1 (uncertainty framework and guidelines for probabilistic fluvial flood forecasting).

It has also been agreed that there will be an exchange of ideas between the two projects regarding development of the guidelines. The present programme for FRMRC2 is to develop a first draft of the guidelines for April 2010 and – following a period of consultation and review – to publish the guidelines in June 2011. Task 2.1 of the present project (uncertainty framework) should therefore be completed before publication of the draft FRMRC2 guidelines, which should in turn be available in time to consider when developing the guidelines under Task 3.1, which starts in April 2010.

Work Package 1.7 will also build upon ideas developed during FRMRC1, which included an assessment of the different techniques available for uncertainty estimation for different types of flood risk management (e.g. Pappenberger *et al.* 2006a). A technical report is available on the FRMRC website and a decision tree and Wiki site was developed as a guide for users coming to the subject for the first time (http://www.floodrisknet.org.uk/methods), including initial development of a decision tree for selection of appropriate methods (Figure 2.1).

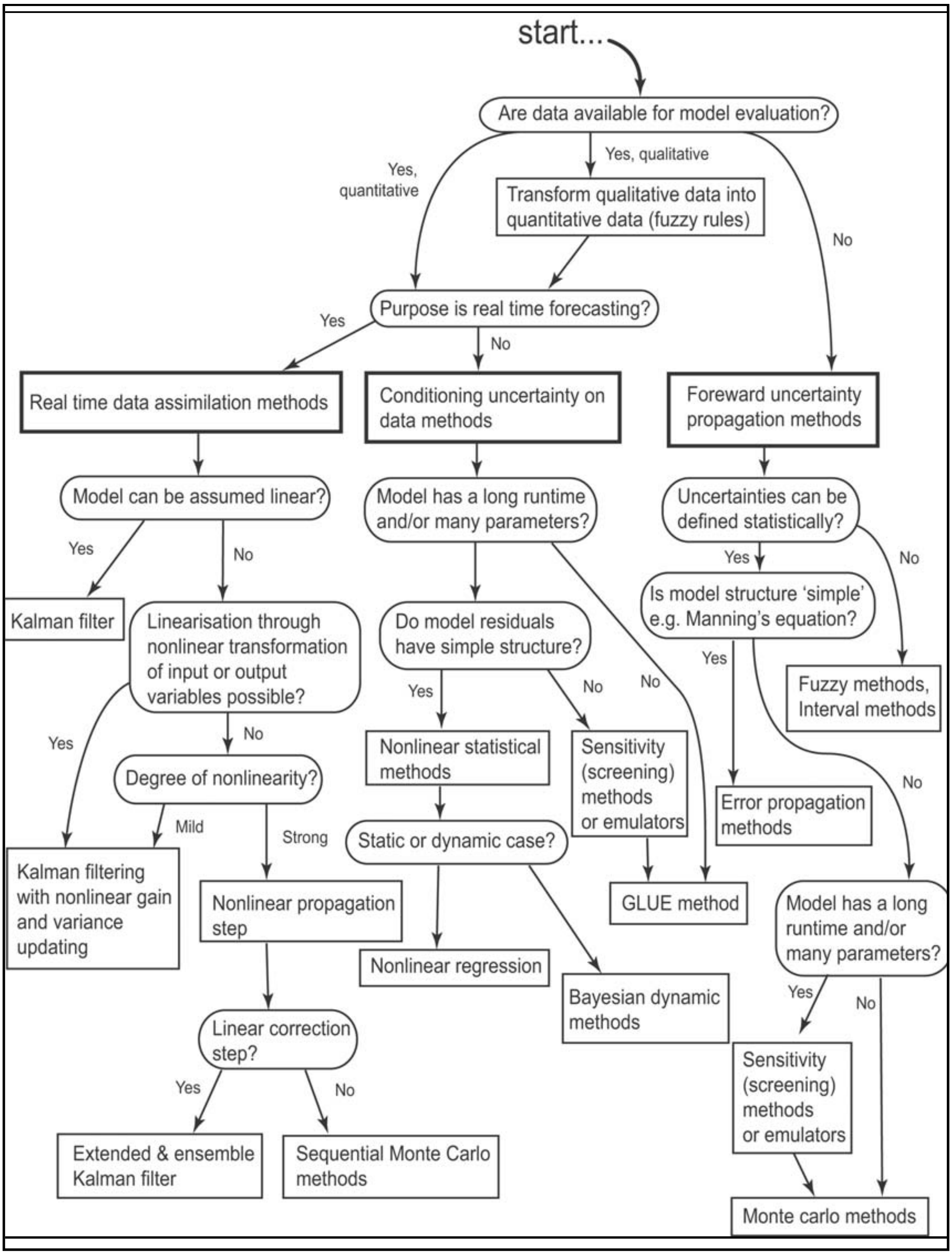


Figure 2.1 FRMRC Research Priority Area 9 method selection example (Pappenberger *et al.* 2006a).

Consideration was also given to the communication of uncertainties between scientists and practitioners (Faulkner *et al.* 2007; McCarthy *et al.* 2007).

Work Package 1.7 will also explore ways of visualising and presenting uncertainties to users. The process of developing a translationary discourse will require a means of interaction between researchers within FRMRC, in other related research projects (e.g. FRACAS within the NERC FREE programme) and with end users from Defra, Environment Agency, SNIFFER and consultancies. Two fundamental research components informing the development of Guidelines for Good Practice will be:

i. a deeper analysis of the degrees and types of uncertainty that the users find most difficult to grasp; and

ii. an exploration of the ways in which that information is conveyed or visualised to users involved in decision making.

FRMRC2 will trial the format of a translational discourse between science, forecasters and relevant decision makers. Research will be undertaken in the form of translational discourse workshops (piloted at the Exeter co-location workshop, March 2006) with observation, interviews and post-workshop questionnaires. Two workshops will be held – an initial workshop in May 2009, and a final workshop in November 2010 – and each will comprise approximately 20 participants consisting of key UK researchers and users. Existing tools will be used to develop a more sophisticated understanding of how 'uncertainty' currently informs flood warning professionals' activities and choices. This will allow a deeper understanding of the role of language and visualisation in assisting the understanding of uncertainty and the relative roles of new methods of communicating uncertainty in flood risk management analysis.

The development of the route map will build upon the experience of implementing uncertainty assessments in the activities of the Netherlands Environmental Assessment Agency (MNP: see, for example, Janssen *et al.* 2003, 2005; van der Sluijs *et al.* 2003, 2005). For example, uncertainty assessments for applications within MNP involve a six-stage checklist as follows:

i. Problem framing

ii. Involvement of stakeholders

iii. Selection of indicators

iv. Appraisal of knowledge base

v. Mapping and assessment of relevant uncertainties

vi. Reporting of uncertainty information

This appears to provide a useful framework for the Guidelines for Good Practice that is consistent with the discourse for conveying uncertainty information to users discussed in Faulkner *et al.* (2007) and McCarthy *et al.* (2007). It is one part of the more general framework discussed by Hall and Solomatine (2008).

Flood Risk from Extreme Events (FREE)

The Flood Risk from Extreme Events (FREE) programme is a 3-year research programme from 2006 to 2009 which seeks to address three central environmental problems associated with flood risk:

- Estimation of the probability, and associated risks, of extreme events leading to flooding occurring in the period from minutes to weeks ahead. Research will be carried out to increase scientific knowledge of: ensemble prediction methods; down/up scaling; aggregation/ disaggregation and propagation of uncertainty through flood forecasting; other statistical methods; warning systems.

- Changes in the intensity and frequency of flooding, and associated weather regimes, resulting from natural and anthropogenic climate change over the next century. Factors dictating our ability to predict the risk of flooding on timescales from seasons to decades will be determined.
- Integrated 'Clouds-to-Catchment-to-Coast' (CCC) flood simulation; involving meteorological, hydrological, and ocean models linked to user products. A coastal zone involves river catchments, an urban conurbation, mixed land use areas, an estuary and adjacent coastal shelf ocean. This modelling framework will be developed and used for holistic flooding scenarios such as arising from combined storm surges and contemporaneous heavy rainfall. CCC is a major output of FREE requiring full integration of the research to be carried out.

The FREE project 'Exploitation of new data sources, data assimilation and ensemble techniques for storm and flood forecasting' is a collaborative project between meteorologists at Reading (the university and the Met Office Joint Centre for Mesoscale Meteorology) and hydrological modellers at CEH Wallingford with support from the Environment Agency. The project has three main themes:

i. data assimilation of new radar measures of the atmosphere into weather models;

ii. construction of physically based ensembles of weather model rainfall forecasts;

iii. probabilistic flood forecasting using improved high-resolution weather model rainfalls in deterministic and ensemble form.

Weather forecast models currently have difficulty capturing the rapid evolution of convective storms leading to flash floods. Assimilating new radar information (using the Met Office variational assimilation system) on the evolving humidity fields (using radar refractivity) and air motions in the boundary layer (using Doppler winds from insects) is being investigated with the prospect of weather models being able to track developing convection before precipitation appears. The analysis fields will be used as initial conditions for ensembles of short-range, high-resolution weather forecasts. This is expected to yield forecasts with improved locations (in space–time) for rainfall events.

Initial condition errors are not the only cause of inaccuracies in high-resolution (1–4 km) weather prediction models. Errors in the lateral boundary conditions and modelling errors also contribute, while the nonlinear nature of convective dynamics also places a limit on deterministic predictability. It becomes important to recognise and determine the uncertainties in the forecast precipitation. A structured approach to ensemble construction is being pursued, accounting for initial condition, lateral boundary and model uncertainties. Perturbations are being designed on the basis of physical insight into convective forcing mechanisms such that the convective-scale ensembles reflect the most significant contributions to forecast uncertainty.

The use of ensemble rainfall forecasts as a means of obtaining probabilistic flood forecasts is being investigated as the third theme of the project. Ensemble rainfall forecasts will be interfaced to hydrological models and probabilistic outputs created. Different scales of application will be investigated, ranging from localised flash flooding of small catchments through to indicative first-alert forecasting with countrywide coverage, including forecasts of discharges to the sea for use in shelf–

sea models. It also aims to assess the impacts of improvements in the Numerical Weather Prediction (NWP) model resolution on flood forecast performance.

The third theme on probabilistic flood forecasting is of most relevance here and progress to date is summarised in the following. A case study of the Carlisle flood using high-resolution NWP model rainfalls in the PDM rainfall–runoff model has demonstrated the improved accuracy of the rainfall forecasts and their potential value for issuing earlier flood warnings (Roberts *et al.* 2009). Use of the CEH Grid-to-Grid (G2G) model for area-wide flood forecasting has involved developing new model initialisation schemes based on steady-state assumptions and novel data assimilation methods (based on empirical state correction) for this distributed grid-based model. The relevance of using a distributed model in conjunction with ensemble rainfall forecasts for convective storms, where storm position uncertainty can dominate, has been illustrated using high-resolution NWP rainfall pseudo-ensembles for the Boscastle storm. Risk maps of flood threshold exceedance that indicate the space–time evolution of flood risk during the event have been developed to support visualisation of the probabilistic flood forecasts (Cole *et al.* 2009).

Another project within FREE is being led by Lancaster University (Professor Keith Beven, Professor Gordon Blair, Dr Paul Smith, Dr Danny Hughes), and is investigating the use of networked sensors to constrain uncertainties in flood forecasts and the calibration of hydraulic models. The work is being carried out in collaboration with the Environment Agency (Steve Mayall, Bangor), Bristol University (Professor Paul Bates) and the Proudman Oceanographic Laboratory (Dr Kevin Horsburgh). The case study for this project is the lower part of the River Dee, including the tidally affected section. The Proudman Oceanographic Laboratory is providing forecasts of surge affected tidal levels and Bristol University is running 2D hydraulic models for part of the river close to the junction with the Alun, where a wireless network of GridStix level sensors has been installed (Smith *et al.* 2008a). Coupled to each of the sensors, a small computer can run the Lancaster DBM flood forecasting software to provide adaptive local forecasts with probabilistic uncertainty estimates.

FLOODsite

FLOODsite is a major research programme within the Global Change and Ecosystems priority of the Sixth Framework of the European Commission. The programme runs from 2004 to 2009 and includes contributions from 37 universities and research institutions.

Research in FLOODsite is being performed across more than 20 tasks, including the following tasks which are of relevance to the present project:

- Task 16 – Real-time guidance for flash flood risk management, which is considering a range of techniques for improving the forecasting of flash floods (led by the University of Padova).
- Task 20 – Development of framework for the influence and impact of uncertainty, which is considering techniques for both flood risk assessment and flood forecasting and warning (led by the University of Newcastle and the UNESCO–IHE Institute of Water Education).

For example, within Task 16, a Bayesian uncertainty extension has been developed based upon the Flash Flood Guidance concept developed by the National Weather Service in the USA. This task has also considered the use of vulnerability indicators (e.g. vehicles damaged) as a guide to flash flood potential (University of Padova).

Within Task 20, perhaps the main area relevant to the present project has been work on the propagation of uncertainty through integrated flood forecasting models, and development of a modelling framework called UNcertainty Estimation based on local Errors and Clustering (UNEEC). A range of computational intelligence and related techniques have been implemented, including artificial neural networks, locally weighted regression, and M5 model trees (Shrestha and Solamatine 2006) based on the following general principles:

- no assumptions about parameter distributions
- model bias is not assumed zero
- localised (i.e. separate) models are built for different input combinations
- model independent

Further details can be found at the programme website http://www.floodsite.net/

HEPEX

The Hydrologic Ensemble Prediction EXperiment (HEPEX) is an international effort that brings together hydrological and meteorological communities from around the globe to build a research project focused on advancing probabilistic hydrological forecasting techniques. The initiative provides a framework for the exchange of ideas and developments by the participating research groups and organisations, although does not directly fund the development of new techniques itself.

The HEPEX mission is to demonstrate how to produce reliable hydrological ensemble predictions that can be used with confidence by the emergency management and water resources sectors to make decisions that have important consequences for economy, public health and safety. The key science issue for HEPEX is reliable quantification of hydrologic forecast uncertainty. HEPEX is addressing the following key questions:

- What are the adaptations required for meteorological ensemble systems to be coupled with hydrological ensemble systems?
- How should the existing hydrological ensemble prediction systems be modified to account for all sources of uncertainty within a forecast?
- What is the best way for the user community to take advantage of ensemble forecasts?

These scientific questions are being treated in separate working groups and test beds. Test beds can be a single basin (and its sub-basins), a region containing multiple basins, or possibly a global collection of basins that facilitate experiments addressing questions over a range of scales and climates. The test-bed projects are exploring a wide range of forward uncertainty propagation, post-processing and data assimilation techniques for a range of timescales (short-term through to seasonal), and several examples are provided elsewhere in this report (e.g. for the US National Weather Service, and the Netherlands).

Regardless of geographical domain, test beds focus on one or more clearly defined HEPEX science questions, have the potential to develop data resources needed for community experiments to address the questions, and are expected to include active user participation. Special workshops are held to exchange ideas between scientists and end users. These have been held in the UK (2004), USA (2005), Italy (2007) and the Netherlands (2008). The next workshop on downscaling of atmospheric forecasts for hydrologic prediction is due to be held from 15 to 19 June 2009 in Toulouse.

Further information on HEPEX can be found at: http://hydis8.eng.uci.edu/hepex/

COST731

The European Science Foundation COST731 project is concerned with 'Propagation of Uncertainty in Advanced Meteo-Hydrological Forecasting Systems'. It is the successor to the COST717 project on the 'Use of Radar in Hydrological and Numerical Weather Prediction (NWP) Models'. The UK representatives on the Management Committee are Dr Sue Ballard (Met Office) and Professor Keith Beven (Lancaster University). COST731 meetings are an opportunity to exchange information and experience of different methods in meteorological nowcasting and data assimilation for precipitation prediction, radar calibration methodologies, hydrological forecasting and the communication of uncertain forecasts to users and stakeholders.

There are three main working groups in COST731:

- WG-1: Propagation of uncertainty from observing systems (radars) into NWP.
- WG-2: Propagation of uncertainty from observing systems and NWP into hydrological models.
- WG-3: Use of uncertainty in warnings and decision making.

The main aims of the project are:

- to provide reports (or peer-reviewed publications) on the key research topics of the working groups;
- to act as a test bed for demonstrating the value and incentives of advanced hydro-meteorological modelling for flood forecasting in Europe;
- to provide a set of well-established and realistic examples (possibly in the form of training software) demonstrating the value and possibilities of probabilistic hydro-meteorological forecasts to potential end users.

As with HEPEX (see above), the COST731 initiative provides a framework for the exchange of ideas and developments by the participating research groups and organisations, although it does not directly fund the development of new techniques itself.

Further information on the COST731 project may be found at: http://cost731.bafg.de/servlet/is/Entry.9691.Display/

MAP D-PHASE

The D-PHASE project (Demonstration of Probabilistic Hydrological and Atmospheric Simulation of flood Events in the Alpine region) is a Forecast Demonstration Project of the WWRP (World Weather Research Programme of WMO). It aims to demonstrating some of the many achievements of the Mesoscale Alpine Programme (MAP), in particular the ability to forecast heavy precipitation and related flooding events in the Alpine region.

The project is addressing the entire forecasting chain ranging from limited-area ensemble forecasting, high-resolution atmospheric modelling (km-scale), hydrological modelling and nowcasting, to decision making by end users. For example, the Hydrology division of the Swiss Federal Office for the Environment (FOEN) is involved in the project as an end user and, for more than two decades, has been running an operational hydrological forecasting service for the High Rhine basin. Initially forecasts were only issued for the most downstream gauging station at Rhein-Rheinfelden, with additional forecasts for the main tributaries having been set up more recently. At present, forecasts for 11 gauging stations are issued to regional authorities and private customers.

In the present operational forecasting system at FOEN a rainfall-runoff model (HBV) is coupled with the COSMO7 meteorological forecast from MeteoSwiss, with the deterministic ECMWF forecast being used as a backup. Hourly data from SwissMetNet as well as from the discharge gauging network of FOEN are used for data assimilation. The integration of all components is realised in the Flood Early Warning System (FEWS).

The increased frequency of flood events in recent years in Switzerland and in neighbouring countries has increased awareness with the public as well as the media and has led to higher expectations of the hydrological forecasting service. In particular, a major flood event in August 2005 and the resulting post-event analysis showed the need for improvements in various respects. In particular, some key factors which were identified were the consideration of uncertainty as well as more frequently updated high-resolution forecasts as input in the hydrological model. The MAP D-PHASE project provided FOEN with an excellent opportunity to benefit from the latest developments in the field of numerical weather prediction and to test new products.

To allow these to be tested during the MAP D-PHASE Operational Period (DOP) (2007–2008), a second forecasting system was set up at FOEN in parallel to the operational forecasting system. The main focus during the DOP was to gain experience with probabilistic and high-resolution models as input for flood forecasting models. For that reason, outputs from COSMO2 were used (which is a refined version of COSMO7), together with COSMO-LEPS and SRNWP-PEPS outputs. These were coupled with HBV in addition to COSMO7 and ECMWF outputs which had already been used in operational flood forecasting. The flood event of 8 and 9 August 2007 was a useful practical test for FEWS MAP D-PHASE.

The additional forecasts have been useful for the dissemination of warnings and – in collaboration with MeteoSwiss and the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) – it is planned to make a full evaluation of the MAP D-PHASE forecasts and to incorporate the results into the operational service.

A second minor focus was testing the use of additional data assimilation techniques; for example, an alternative approach to updating of HBV, and use of radar-derived catchment rainfall at an hourly resolution provided by MeteoSwiss Locarno instead of the spatially interpolated ground-based measurements.

More information on MAP D-PHASE can be found at:
http://www.map.meteoswiss.ch/map-doc/dphase/dphase_info.htm

European Flood Forecasting System (EFFS)

The EFFS was sponsored by the 5th Framework Programme of the European Commission. The project started in 1999 and was completed in 2003 after a 6-month extension period. The project consortium was initially formed by 11 partner institutes. During the third year the consortium was extended via a special amendment to include eight additional institutes from seven Newly Associated States of the European Union.

The goal of the project was the set-up and semi-operational testing of a continental-scale flood forecasting system for major river basins in Europe. The purpose of the forecasting system is to function as a pre-warning system for national and regional forecasting offices across Europe.

The principal research aim was to explore if it is possible to extend the lead time of the warning process up to 10 days into the future. This was to be achieved with the use of numerical weather forecasts. Various deterministic and ensemble forecasts delivered by national and international meteorological services were used within the system to drive a sequence of rainfall–runoff models and hydraulic models for the principal river systems, in particular the rivers Rhine and Po. The weather forecasts were downscaled from a global circulation model to a high-resolution local model. Various semi-operational tests led to the conclusion that – at that time – 10-day forecasting periods were considered non-reliable for most situations due to the high uncertainty in the meteorological forecasts beyond a duration of about 6 days.

An important aspect within the EFFS project was the uncertainty inherent to the forecasting process. The principal sources of uncertainty include the internal model parameter uncertainty and the input uncertainty. Another relevant part of the project was the communication of the forecast results to the forecast end users, including rescue services and expert forecasters. The principal responsibility of the forecaster is to interpret a particular warning issued by a forecasting system and give instructions for an eventual evacuation or other operational response. Forecasters thus need to foresee the consequences of eventual false alarms or of a situation for which an evacuation would have been warranted. In this particular part of the project the various needs and feedback from the end-user community were explored and addressed explicitly.

Figure 2.2 shows the principal working packages of the EFFS project (structured into 10 blocks) and the institutes that led the individual work packages.

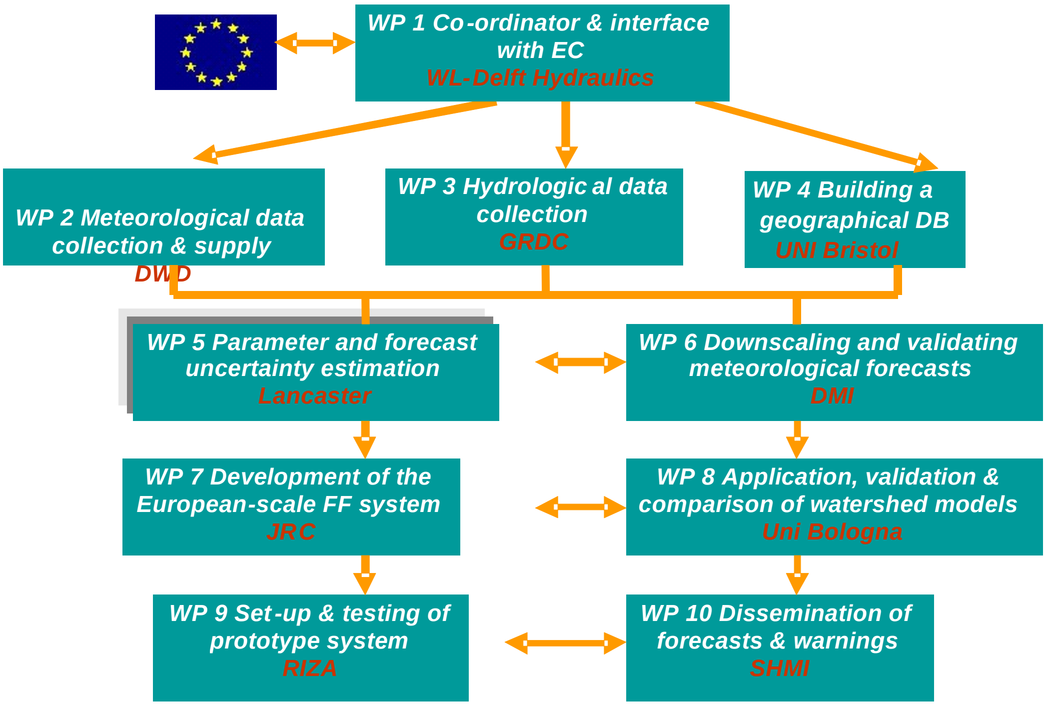


Figure 2.2 EFFS project organogram.

The EFFS project ended in 2003. More information can be found at the following link: http://cordis.europa.eu/data/PROJ_FP5/ACTIONeqDndSESSIONeq112362005919ndDOCeq129ndTBLeqEN_PROJ.htm

EU IMPRINTS project

The EU Framework 7 project IMPRINTS is concerned with flash flood and debris flow risk assessment as well as real-time forecasting and is led by Professor Daniel Sempere-Torres of CRAHI, UPC, Barcelona. One of the project partners is Lancaster University (Professor Keith Beven). Others include the EU Joint Research Centre, Ispra, Italy; MeteoSwiss and WSL from Switzerland; and the French Flood Forecasting Institute, SCHAPI.

The project started on 15 January 2009. There will be a number of case study catchments used in the project in Spain, Switzerland, France and Italy. Stakeholders from these catchments are involved in planning the project. Probabilistic forecasting will be used in both the assessment of risk and in real-time forecasting, making use of ensemble rainfall predictions from ECMWF and COSMO-LEPS, radar projections, raingauge and water level data.

The progress of the project can be followed on the IMPRINTS website: http://www.imprints-fp7.eu/

2.1.2 Pre-operational and operational experience

This section describes the current status, and technical approaches used, within the following systems which are currently undergoing pre-operational testing, or are already in operational use.

National Weather Service

The National Weather Service (NWS) of the USA is currently migrating the current NWSRFS flood forecasting system to the Community Hydrological Prediction System (CHPS), which will incorporate the functionality of the current Advanced Hydrologic Prediction Service (AHPS). AHPS is a critical delivery mechanism for NWS's integrated water services, and provides a web-based suite of information-rich forecast products. This includes the display of uncertainty-quantified forecast information for occurrence of floods and droughts from hours to days and months in advance. These graphical products provide useful information for use as planning tools by water resources and emergency managers. It is noted that 'these new products will enable government agencies, private institutions, and individuals to make more informed decisions about risk-based policies and actions to mitigate the dangers posed by floods and droughts'. (http://www.weather.gov/ahps/about/about.php).

In support of AHPS, the experimental ensemble forecasting system XEFS is being developed for implementation in CHPS. The idea is to develop an integrated short, medium and long-range ensemble streamflow forecasting system that can be developed and implemented at the River Forecasting Centres (RFCs) over the next few years.

A modular approach is being taken to the development of XEFS which consists of several parts that treat uncertainty explicitly:

- Ensemble Pre-Processor (EPP) to deliver unbiased and skilful weather and climate ensembles. EPP is a method to account for temporal-scale-dependent relationships in both forecast errors and precipitation and temperature variability over the entire forecast period (Schaake *et al.* 2007).

- Ensemble Streamflow Prediction system (ESP). The ESP procedure was proposed by Day (1985). In this procedure an empirical ensemble of precipitation and temperature inputs is sampled from the validated long-term time series of catchment average temperature and precipitation. A sample is drawn from each available year in the historical series using the current day as the starting point. The ensemble thus created is a representation of the climatology of catchment temperature and precipitation, and can be run through the forecast model cascade, resulting in a climatology-based seasonal forecast, conditional on the states of the system at the time of forecast. These forecasts are used for water resource planning in the USA.

- Hydrological ensemble post-processors such as HMOS (Hydrological Model Output Statistics). This is a method for providing uncertainty information about hydrological forecasts, specifically short-term forecasts. HMOS should, in theory, cover both meteorological and hydrological uncertainties. With HMOS, a deterministic forecast is transformed into an ensemble based on past performance. See Seo *et al.* (2006) for more information. To develop HMOS a long-term archive of both forecasts and observations is needed. HMOS is currently being tested for operational use by several regional forecast centres in the USA.

- EnsPost (ensemble post-processor). EnsPost is used to post-process ESP forecasts to remove bias and correct the mean forecast.
- Ensemble Verification System (EVS).
- Ensemble Product Generator (EPG) to generate the ensemble products.

These components are being implemented in the CHPS (FEWS) system. Under Phase 1 (2007–2009) of the XEFS programme, the focus is on use of ensemble rainfall and temperature forecasts, while Phase 2 will consider explicit accounting for other sources, including model structure, parameters and states as well as flow regulation. The candidates for potential approaches include, but are not limited to:

- ensemble data assimilation to reduce uncertainty in the initial conditions and to keep track of growth (due to accumulation of errors in time and/or through the forecast system) and reduction (due to newly available observations) of uncertainty;
- a parametric uncertainty processor to reduce and to explicitly account for uncertainty associated with model calibration;
- multi-model ensembles to reduce the effects of and to account for structural errors in models;
- new techniques for modelling of flow regulations and accounting of uncertainties associated with them.

More information can be found at:
http://www.nws.noaa.gov/oh/hrl/chps/XEFS_proj.html

European Flood Alert System

The European Flood Alert System (EFAS) is currently under development at the EU Joint Research Centre, and followed on from the EFFS project, which was described earlier.

The aim of the system is to provide medium to long-term flood forecasts (3–10 days) based on ensemble rainfall inputs (Thielen *et al.* 2004). The system is intended to complement the flood forecasting services operated by national authorities and to assist in disaster prevention, preparedness and damage assessment. The project started with a research phase from 2003 to 2006, during which pilot studies were performed for the Danube and Elbe basins, and has now been extended to a number of other countries and river basins across Europe, including the Rhine and the Po. The project is funded by the European Union and the models are developed and operated at the Joint Research Centre in Italy.

The EFAS models operate on a gridded basis, and use the 51 member ensemble rainfall and temperature forecasts from the European Centre for Medium-range Weather Forecasts (ECMWF) in the UK. Deterministic forecasts from the German Weather Service (DWD) are also included. Raingauge data are also used as a model input.

The hydrological component of the system uses a LISFLOOD rainfall–runoff and flow routing modelling approach on a 5 km grid, coupled to ECMWF ensemble forecasts (rainfall, temperature etc). A higher resolution version (1 km grid), which will become the standard, has been developed for the Danube and Elbe catchments and was used successfully to provide advance warning of flooding during the 2005 and 2006 flood events in those catchments. Other developments are also in progress, and include hydraulic modelling and state updating functionality. The system includes many novel ideas for the assessment of probabilistic forecasts and for the display and interpretation of probabilistic flood forecasts.

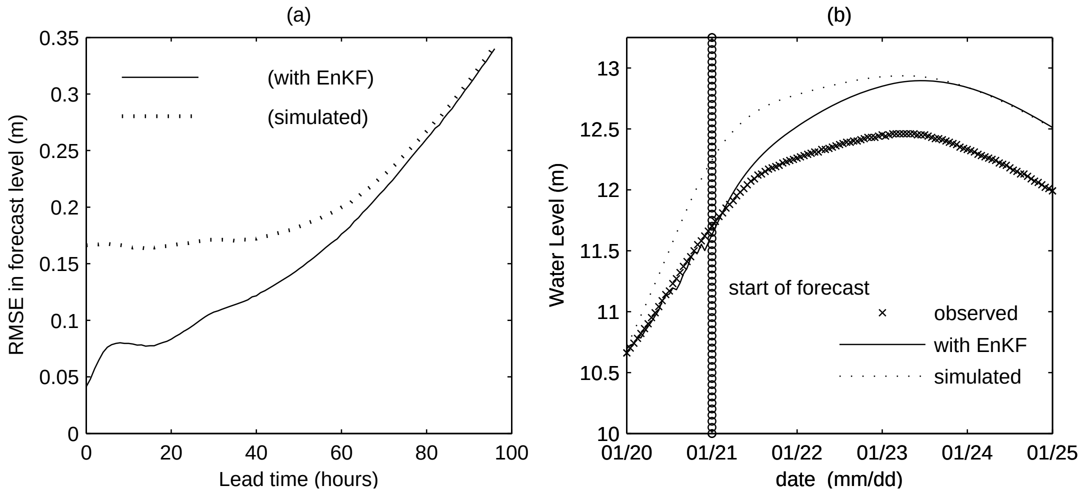
Dutch Centre for Water Management (Rhine and Meuse)

The Dutch Centre for Water Management (WMCN) is responsible for flood warning and daily forecasts for shipping for the rivers Rhine and Meuse that enter the Netherlands from Germany and Belgium respectively. Since 1999, a flood forecasting system called FEWS-NL Rhine & Meuse has been under development. It became operational in December 2008.

Within FEWS-NL, interpolated temperature and rainfall fields, derived from synoptic meteorological measurement stations, are used as inputs to hydrological models of the rivers Rhine and Meuse. Subsequently, discharges calculated by the hydrological model are input into a hydrodynamic model of the Rhine and Meuse.

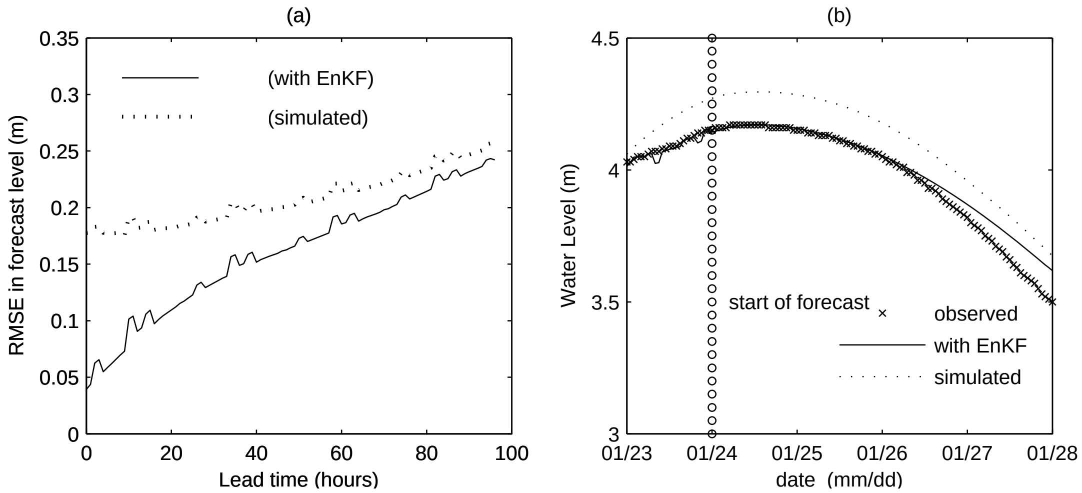
To improve the forecasts made with the FEWS-NL system, technological improvements are introduced when they become available. The Rhine component is also one of the test beds for the HEPEX project, which was described earlier. For example, to constrain uncertainty in the operational forecasting system, error correction and other data assimilation techniques are used. The effect of error correction on improvement of the root mean square errors (RMSE) in forecasts has been investigated as a function of lead time for both the Meuse and Rhine (Broersen and Weerts 2005; Weerts 2007; Werner *et al.* 2009).

Also, in the operational forecasting system for the River Rhine, an Ensemble Kalman Filter (EnKF) using 48 ensemble members has been implemented for the hydraulic SOBEK-RE model of the Rhine using the generic data assimilation module DATools (Weerts 2007, 2008; Werner *et al.* 2009). An example of the results at Lobith is given in Figure 2.3 and at Olst (downstream of Lobith) in Figure 2.4.



(a) Root mean squared error of the water level forecast at the gauge of Lobith on the Rhine with EnKF and without assimilation as a function of lead time determined over a 2-year hindcast (2006 and 2007). (b) Observed water level together with the mean of the EnKF water level forecast and the water level forecast without assimilation at Lobith for an event in January 2007. The HBV-96 – SOBEK-RE model cascade is forced using HIRLAM NWP.

Figure 2.3 Example of Ensemble Kalman Filter outputs for Lobith.



(a) Root mean squared error of the water level forecast at the gauge of Olst on the Rhine with EnKF and without assimilation as a function of lead time determined over a 2-year hindcast (2006 and 2007). (b) Observed water level together with the mean of the EnKF water level forecast and the water level forecast without assimilation at Olst for an event in January 2007. The HBV-96 – SOBEK-RE model cascade is forced using HIRLAM NWP.

Figure 2.4 Example of Ensemble Kalman Filter outputs for Olst.

In addition to this work, EnKF and particle filtering approaches for state updating of conceptual hydrological models have been investigated (Weerts and El Serafy 2006; Weerts *et al.* 2008a, 2009). In 2008, Bayesian Model Averaging (BMA) was also investigated for deriving a probabilistic forecast for the Rhine (Beckers *et al.* 2008).

A new rainfall interpolation scheme has also recently been introduced, which tries to emulate high-quality areal rainfall estimates as much as possible during operational forecasting. The new approach was compared with estimates derived with the current interpolation method used in the operational system and it was found that the areal rainfall estimates derived with the new interpolation procedure emulate the high-quality precipitation data better for most catchments of the Rhine basin, especially in those catchments where orographic influences play a role, such as the Black Forest and the Vosges mountains (Weerts *et al.* 2008b).

Within the FEWS-NL Rhine & Meuse systems, several deterministic and ensemble forecasts (HIRLAM, DWD-LM, DWD-GME, ECMWF-DET, ECMWF-EPS and COSMO-LEPS) are also used to gain insight into the uncertainty caused by meteorological conditions as far as 10–15 days ahead. Post-processing of these forecasts is necessary to remove hydrological and meteorological biases that exist within these forecasts (Regianni and Weerts 2008a, 2008b; Reggiani *et al.* 2009).

Lake Como Decision Support System

Lake Como is in northern Italy and is regulated for irrigation and energy production. However, the small available free storage has resulted in several major flooding incidents downstream following high inflows. A real-time forecasting system has been developed to assist with gate operations at the lake, with assessment of uncertainty, and with forecasts provided for 0–24 hours and 1 to 10 days ahead (Todini 2004). The forecasts and associated uncertainty are then used as part of a stochastic optimisation algorithm to preserve the expected benefits from irrigation and hydropower, while minimising the expected damages from flooding in the town of Como. The system has been operational since October 1997 and has been used successfully during several flood events.

FLOODRELIEF

As part of the EU-funded FLOODRELIEF project, a general stochastic framework was developed based on the Ensemble Kalman Filter with case studies using MIKE11 hydraulic models for the Blue River basin (USA) and the Welland and Glen catchment (UK). The influence of uncertainty was examined by assuming typical magnitudes and distributions of errors in the inputs (Butts *et al.* 2005).

GeoGUI

GeoGUI is a simple, small footprint, flood forecasting system that is based on a combination of the PRTF rainfall–runoff method and the ISIS forecasting model although it will accept inputs from any rainfall–runoff forecasting module. It was originally configured for the River Eden in North West Region, and is believed to have been one of the first operational real-time ISIS hydrodynamic models in the UK. For the time being it is still operational in Thames Region, where it has been combined successfully with their CASCADE forecasting system which gathers data from the Regional Telemetry System and the regional implementation of the RFFS (River Flow Forecasting System) developed originally by CEH Wallingford. It was also used operationally in South West Region, where it was successfully combined with the WRIP (Weather Radar Information Processing) system developed by Plan B UK. A typical database size is less than 10MB, which compares with several hundred megabytes for an NFFS datastore, but it should be remembered that GeoGUI has limited functionality when compared with the full NFFS.

River Nith Forecasting System

The Lancaster University DBM approach to probabilistic flood forecasting was first implemented in an application to the River Nith in Scotland carried out for the Solway River Purification Board and later adopted by SEPA (see Lees *et al.* 1994; Beven 2001).

Building on the nonlinear DBM work of Young and Beven (1991), the Nith model consists of a cascade of rainfall–flow and flow routing elements with a two-input tidal model to forecast the tidal reach of the Nith at Dumfries. The Nith model provides uncertainty estimates for discharge forecasts with lead times of up to 6 hours for all the gauging sites in the catchment and is adaptive when real-time river flow levels are available.

Since that work, further applications, using a more sophisticated method for modelling the effective rainfall nonlinearity, have been made to the River Hodder (Young 2002) and, as part of FRMRC, to the River Severn (Romanowicz *et al.* 2006b, 2008) and River Eden (Leedal *et al.* 2008). More detail on the background to the methods may be found in Young (2002, 2009), Young *et al.* (2006), Romanowicz *et al.* (2006a,b, 2008) and Section 3.3.3 below.

CI-FLOW

The Coastal and Inland Flooding Observation and Warning Project (CI-FLOW) is a 5-year programme of research which began in 2008 and aims to evaluate and test new forecasting techniques for river and coastal floods, using the Tar-Pamlico and Neuse river basins in North Carolina as a test bed. The project is being led by the NOAA National Severe Storms Laboratory. The outcome will be a prototype coupled model system adaptable to any coastal river system to help in freshwater forecasting of floods and flash floods, water management, determination of land use and ecosystem impacts, and coastal storm surge forecasts. The modelling system will consist of ensembles of inland river models and coastal ocean/estuary models, each using inputs from high-resolution numerical weather forecast models and multi-sensor precipitation estimates. Existing models which will form part of the ensemble system include 1D and 2D hydrodynamic models, a distributed rainfall–runoff model, a coastal forecasting model, and a conceptual rainfall–runoff model (Figure 2.5).

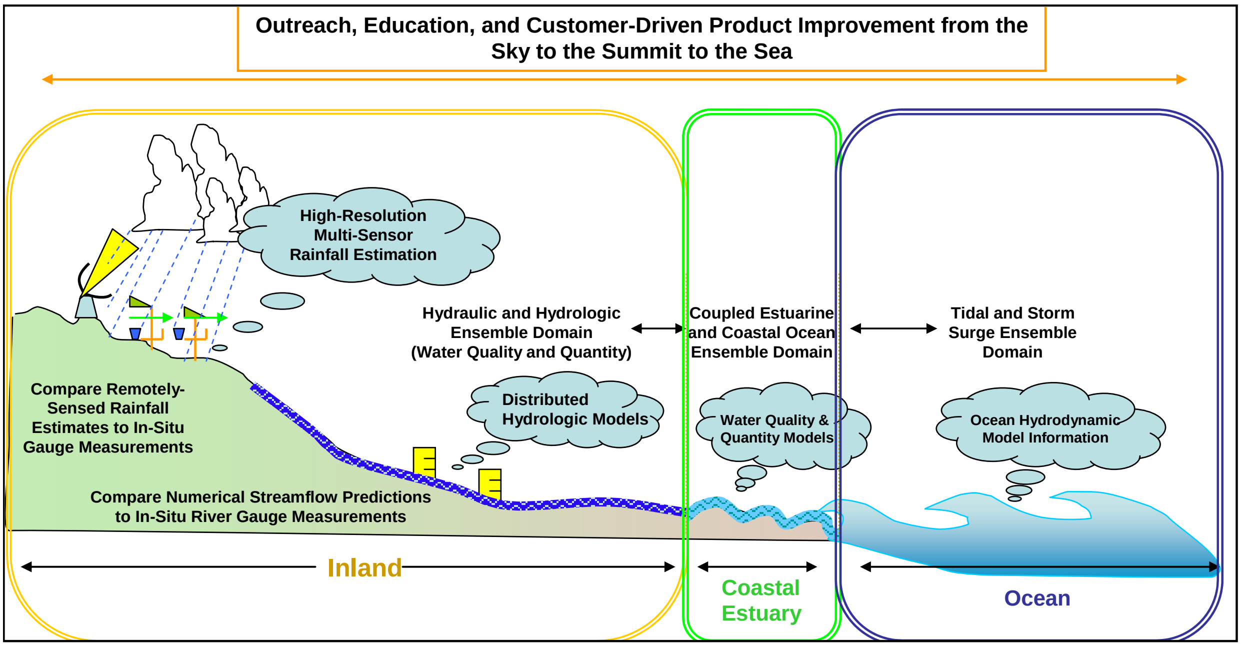


Figure 2.5 The three CI-FLOW project elements (inland, coastal ocean/estuary, and ocean) that are being connected to provide an integrated accounting of water quantity and quality from the sky to the summit to the sea (from the CI-FLOW Project Plan 2008).

Further information can be found on the CI-FLOW website: http://www.nssl.noaa.gov/projects/ciflow

2.2 Environment Agency experience

2.2.1 General guidance

This section provides a brief review of existing guidance and information on assessment of uncertainty provided within current Environment Agency documents which are relevant to this project.

Real-Time Modelling Guidelines

The Real-Time Modelling Guidelines were prepared during 2001–2002 under the guidance of a Project Board which included representatives from most of the regional flood forecasting teams, the (then) National Flood Warning Centre, and the Met Office. The research contractors were Atkins, Edenvale Young and JB Chatterton & Associates. The main outputs from the project were a technical report, the guideline document, and a range of worksheets and templates to assist with using the guidelines.

Based on guidance from the Project Board, the guidelines aimed to offer a structured decision-making framework for selection of appropriate models for fluvial flood forecasting applications, but without being prescriptive, requiring users to apply some judgement supported by local knowledge. It was also recommended that, wherever possible, locally applicable costs, targets and other factors should be used in place of the default values supplied. The approach should be tailored according to the level of detail required, and how much analysis had already been performed for the catchment/Flood Warning Area in previous studies.

The guidelines covered the choice of possible modelling approaches for a given Flood Warning Area, accounting for:

- Environment Agency targets for flood warning systems;
- different physical types of catchment and river, including floodplains and control structures;
- the varying levels of data availability and quality;
- the levels of risk and the consequences of error;
- the cost and time of developing or improving a system.

Table 2.1 shows the types of models which were considered in the guidelines.

Table 2.1 Summary of model types considered in the Real-Time Modelling Guidelines (Environment Agency 2002).

General type	Category	Example
Empirical	Correlation models	Level-level or flow-flow correlation, time-of-travel maps, flood warning contingency table
Rainfall–runoff	Black-box models	Transfer function (linear, nonlinear), unit hydrograph
	Conceptual models	Lumped or distributed rainfall–runoff models, snowmelt models
Routing	Hydrological routing	Muskingum, Muskingum-Cunge and variable parameter versions, some black-box models
	Kinematic routing	Fixed and variable parameter versions
Hydrodynamic	1D, 2D or 3D model	Section 105 or other model converted to real-time use

The issue of lead time requirements was a key consideration in the model selection process, based primarily around Environment Agency targets for a minimum 2-hour lead time for flood warnings (1 hour in Wales). An allowance was also included for decision-making time, dissemination time (e.g. by phone) and other factors. Risk was interpreted using a cost-benefit analysis, in which the benefits derive from the reduction in flood damages which can arise from an accurate, reliable and timely flood warning, taking account of social factors such as the ability to respond to warnings, and awareness of the meaning of flood warnings.

Data availability and quality was also a key consideration, with a discussion of the types of data required for real-time modelling and model calibration, the quality of data, record lengths, rating curves and other factors. The type of information required by operational staff and professional partners was also discussed throughout the document; for example, if just the peak flow is required, a simple correlation might be suitable in some situations whereas, if real-time inundation maps are required, then a hydrodynamic model might be selected.

The guidelines presented information in a wide range of formats, including flowcharts, risk matrices, case studies, and tables of strengths and limitations. Guidance was also provided on issues such as model calibration, the likely times required for model development, key risks and assumptions, and topics such as high flow ratings.

Since being issued, the guidelines have been used on a range of model development projects, and aspects of the guidelines have appeared in work instructions and other internal Environment Agency documents. At a similar time to preparation of these guidelines, national guidelines were also prepared on Rainfall Measurement and Forecasting techniques, Estuary Flood Forecasting techniques, and Coastal Flood Forecasting techniques.

Flood Warning Level of Service

The Flood Warning Level of Service is the required performance of the flood warning service in support of the Flood Warning Investment Strategy. The principles are described in AMS Work Instruction 137_05 (Environment Agency 2009), which explains what the Environment Agency defines as the appropriate Level of Service for each Flood Risk Area and hence Flood Warning Area in the following general categories:

- Identifying areas of risk (Flood Risk Areas)
- Establishing Flood Watch and Flood Warning Areas
- General procedures and organisation
- Detection and observation of flooding
- Forecasting and warning message preparation
- Disseminating the warning message
- Raising awareness of risk, flood warning service and response with the public
- Post-event data collection, reporting and archiving
- Improving service effectiveness after post-flood event review.

In the area of risk assessment, the work instruction includes an updated version of the well-known risk assessment matrix for subdivision of Flood Warning Areas by level of risk, and Figure 2.6 shows the version for fluvial flooding (there is a separate matrix for tidal flooding). Note that, in this figure, the probability reflects the standard of protection for the Flood Warning Area, and is inversely related to the return period values which are shown at the top of the figure (e.g. 1:100 years).

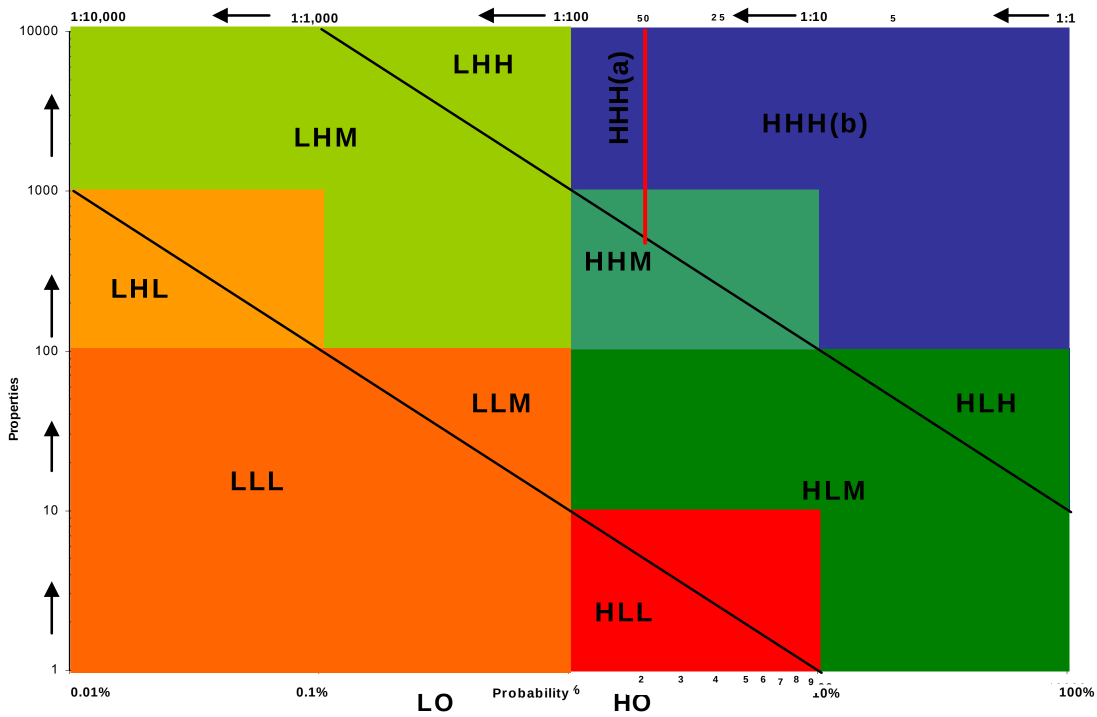


Figure 2.6 Risk matrix for fluvial risk locations (from Environment Agency 2009).

The risk categories (LHL etc) are used to define the required level of service for Detection & Forecasting, Warning Dissemination, and Communicating Flood Risk. For example, the Detection & Forecasting requirements include guidelines for the density of raingauge and river level gauge networks, and applicability of weather radar rainfall estimates in flood forecasting, and also introduce the requirement to estimate flood forecast performance in terms of crossing of threshold levels, using a simple contingency table approach (Table 2.2).

Table 2.2 Contingency table from AMS 137_05 for assessing flood forecast performance (Environment Agency 2006).

	Threshold forecast to be crossed	*Threshold crossing observed*	
		Yes	*No*
	Yes	*a*	*b*
	No	*c*	*d*

The performance measures which are defined are False Alarm Rate $FAR = b / (a + b)$ and Probability of Detection $POD = a / (a + b)$. The assessment of performance is based upon the forecast of a threshold being crossed, and whether this was observed through telemetry data, not whether or not flooding actually occurred. This ensures that the forecasting method is not penalised for incorrect setting of threshold levels. The minimum requirements for POD and FAR, and for instrumentation network densities, are shown in Table 2.3.

Table 2.3 Detection and forecasting requirements (Environment Agency 2009).

Requirement type	Level of service		
	Minimum	Intermediate	Maximum
Water level detection requirements (maximum distance from boundary of community)			
Steep rivers, > 1 in 200	1 km downstream 1 km upstream	1 km downstream 1 km upstream	1 km downstream 1 km upstream
Medium steepness, < 1 in 200 and > 1 in 1000	1 km downstream 10 km upstream	1 km downstream 5 km upstream	1 km downstream 1 km upstream
Low steepness, < 1 in 1000 and estuarine	2 km downstream 20 km upstream	2 km downstream 10 km upstream	2 km downstream 2 km upstream
Class of radar data	3A	2A	2A
Raingauge density	1 gauge per 150 km^2	1 gauge per 100 km^2	3 gauges per 100 km^2
False Alarm Rate	< 50%	< 40%	< 30%
Probability of Detection	> 50%	> 60%	> 70%

At present the same criteria are applied to each threshold, but the work instruction notes that, when more data are available, this will be reviewed such that higher thresholds (e.g. Severe Flood Warning) may have more stringent criteria, and lower thresholds have less stringent criteria.

Other measures, in addition to these, are currently being considered by the ongoing Environment Agency Performance Measures, or Skill Score, project, which is due to complete later in 2009. A new performance monitoring module in the NFFS will facilitate the calculation of these measures.

Flood Risk Management Modelling Strategy

As part of this review, an update was obtained on the current version of the Environment Agency's Flood Risk Management Modelling Strategy (FRMMS) to establish whether probabilistic methods in flood forecasting form part of the medium to long-term vision.

The custodian of this document is the Flood Risk Management Policy Team and the status of the document is that it has been subject to external consultation but not yet finally signed off by the Environment Agency. It is understood (personal communication) that the final version of the strategy will contain messages on the move to probabilistic forecasting as a key strategic aim. However, it should be noted that the role of the FRMMS is not to specify modelling techniques in terms of uncertainty estimation or probabilistic modelling.

Performance Measures for Flood Forecasting

An Environment Agency research report, *Performance Measures for Flood Forecasting* (Environment Agency 2005) was written in 2005 by HR Wallingford Ltd and Edenvale Modelling Services.

The primary performance measures used at present by the Environment Agency relate to threshold crossings. There are two of these measures, referred to as the Probability of Detection (POD) and the False Alarm Rate (FAR).

These measures are only meaningful when associated with a predetermined lead time and the NFFS includes a means to automatically calculate them once a threshold has been crossed. It is dependent to some extent on when the forecast was made, relative to when the threshold was crossed, as generally it cannot be guaranteed that a forecast would have been made at exactly the right moment. The current version of the NFFS does not calculate the POD and FAR values exactly according to Environment Agency requirements but it is expected that this issue will be remedied in the next version.

It is important to note that this measure relates to a single forecast and as such is entirely deterministic.

While these measures represent the current position, it is expected that future measures will pertain to the prediction of peak levels and their timing.

An extensive review of rainfall forecast performance measures was also carried out by CEH Wallingford for the Environment Agency and the Met Office in 2003, under a project called 'Development of Rainfall Forecast Performance Monitoring Criteria. Phase 1: Development of Methodology and Algorithms'. The Main Report (Jones *et al.* 2003) was complemented by a User Guide to the assessment tool software that included simple examples of the use of each performance measure (Jones *et al.* 2004). While this focused on the assessment of rainfall forecast products in use by the Environment Agency, the methodology and algorithms reviewed are also very relevant to the problem of assessing the performance of flood forecasts.

The study distinguished between assessment measures in continuous variable form, in categorical form based on Skill Scores, and those in probability form. It also considered the practical relevance of the form of the error (additive or proportional) and the role of transformations, such as logarithms. The advantages and disadvantages of each measure were carefully reviewed in a practical context. The great variety of measures reflected the need to judge different attributes of a forecast (e.g. bias, typical error size, exceedance of a threshold) and to cope with different forms of forecast (e.g. value, probability). The choice of measure thus depends on the form of the forecast and the users' main interests in relation to their practical application in support of flood warning. While all measures have some value, the report selected a small number as being most important.

While these performance measures have generic application, and are as relevant to assessing flood forecasts as to rainfall forecasts, there are additional ones that are especially relevant to flood forecasts. These focus on assessing how well a forecast reproduces the form of the flood hydrograph and encompass such characteristics as the magnitude and timing of the flood peak and the start to rise. As indicated above, the categorical Skill Scores would typically be defined in terms of crossing critical thresholds, such as those relating to flood warning alarm levels.

2.2.2 Defra/Environment Agency R&D studies

PFFW

A national Flood Incident Management project is currently determining approaches and business change required to embed probabilistic flood forecasting and warning (PFFW) into current Environment Agency practices. The present project is closely aligned with this project and will provide some of the required scientific evidence and methods.

Probabilistic Flood Forecasting Scoping Study

The aim of the Probabilistic Flood Forecasting Scoping Study was to assess the current state of knowledge and direction of developments in probabilistic flood forecasting in consultation with external researchers, and to discuss requirements with end users in order to scope a long-term development programme for the introduction of probabilistic forecasting into operational use.

The project was performed during 2006 and 2007 and ran in parallel with the following two related Environment Agency research studies, which started at a similar time:

- Use of Probability Forecasts – Met Office
- Hydrological Modelling with Convective Scale Rainfall – WL/Delft Hydraulics and CEH Wallingford

The project considered both fluvial and coastal forecasting, and related areas such as pluvial forecasting in urban areas, including some of the main operational implications in terms of training, decision support systems, presentation of information etc, with a key aim being to identify research needs and other follow-on projects. The project also considered experience gained by other organisations which are considering including probabilistic information in operational forecasts. Wide-ranging consultations were also held with flood warning, flood forecasting, policy and other staff on questions such as:

- What is a realistic rate of implementation for probabilistic forecasting techniques, and what are the main priorities?
- What opportunities are there for learning from work already under way internationally?
- What research, system, process and policy developments may be required?
- How widely should probabilistic forecasts be disseminated?
- What are the likely public awareness campaign and training requirements?

The initial findings from these consultations were presented at a project workshop on 13 February 2007, which was attended by flood forecasting staff from all Environment Agency regions, and flood warning staff from several area teams. The Technical Report for the project was published later in 2007. The report covers sources of uncertainty in the flood forecasting process, current approaches within the Environment Agency to assessing uncertainty in flood forecasts, international research on ensemble flood forecasting techniques, possible applications of decision-support systems, and risk-based forecasting techniques in other (non-water) sectors.

The outputs from this project will help to inform Defra and the Environment Agency in developing a plan for bringing this important development into operational use over the next few years.

Hydrological Modelling with Convective Scale Rainfall

This research project was initiated in order to investigate and benefit from developments in numerical weather prediction being carried out by the UK Met Office. In 2009, a high-resolution nowcasting system called STEPS will become operational at a 2 km resolution. For longer-term numerical weather prediction, a new system has been developed called MOGREPS, which uses a coarser model resolution of 24 km. Both systems will be run in ensemble mode.

Part of the requirement of the project was to develop operational research into the practicalities of integrating a probabilistic approach into the current NFFS system. The project focuses on ensembles of rainfall inputs and does not address additional areas of uncertainty. Practical implementation guidance is given on how ensembles of rainfall inputs can be relatively easily incorporated into the NFFS system.

A live demonstration system was set up in Delft for North East and Thames regions of the Environment Agency as a proof-of-concept (although it excluded ensembles of hydrodynamic model runs in order to reduce computational burden on the test system). In addition to this, methods for visualisation of probabilistic information such as spaghetti plots, probability plots and persistence tables were developed for the two example regions.

The project also noted some useful preliminary examples of where probabilistic forecasting based on ensemble rainfall forecasts can add value, where drawbacks may exist, and where forecasting based on rainfall ensembles is potentially misleading. The main conclusions were that, for larger, well-calibrated catchments, probabilistic forecasting can be useful in increasing forecast confidence. In small catchments with a chance of high-intensity storms, probabilistic forecasting could lead to a significant increase in the number of false alarms but also an increase in detection rates. In poorly calibrated or ungauged catchments, probabilistic forecasting can give a false sense of the range of possible outcomes and can therefore be misleading. However, in the first half of 2009 the project will run MOGREPS products in hindcast mode to better understand the potential increase in skill derived from the use of probabilistic information.

Within the project, CEH Wallingford is working with the Joint Centre for Mesoscale Meteorology (Met Office) to trial high-resolution NWP model rainfall outputs, using models at 1, 4 and 12 km resolution. A case study of the Boscastle storm has demonstrated the value of the finer resolution model rainfall product for this convective event. Pseudo-ensemble forecast products have been created from the deterministic NWP rainfall forecasts by randomly displacing the rainfall fields in space over a spatial domain of a size reflecting the positional uncertainty. These rainfall forecast ensembles have been used as input to lumped (PDM) and distributed (G2G) hydrological models to produce ensemble flood forecasts. The distributed model outputs have been further processed to produce risk maps of flood exceedance over a forecast horizon of interest, as an illustration of one form of probability forecast product. This work serves to prepare the Environment Agency for using high-resolution NWP rainfall ensembles for probability flood forecasting, when these become operationally available from the Met Office.

The final phase of the project is using the summer 2007 convective storms over the Midlands (focusing on the upper Avon and Tame catchments) as a further case study for investigating the potential value of high-resolution NWP rainfall and ensemble forms of them for probabilistic flood forecasting and warning. It will also provide an opportunity to trial the new STEPS ensemble rainfall product. This regional case study is being complemented by a national (England and Wales) case study which aims to demonstrate and assess the G2G model for area-wide flood forecasting, operating within a test NFFS operational environment. Earlier phases of the project have led to NFFS module adapters for the G2G model and for the HYRAD procedures that derive gridded estimates of rainfall (from radar and/or raingauge data) used as model input. It is planned that the England and Wales implementation of the G2G model will be used by the new Flood Forecasting Centre, providing an indicative country-wide picture of future flood risk complementing the more detailed regional NFFS forecasts.

Use of Probability Forecasts Project

The ‘Use of Probability Forecasts’ joint Environment Agency–Met Office funded project is managed by the Met Office within the Joint Centre for Hydrometeorological Research (JCHMR) and aims to prepare a suite of uncertainty-based (rainfall) forecast products for use by the Environment Agency in operational fluvial flood forecasting and warning. The project started in May 2006.

The types of product to be produced were based on an initial user requirement for uncertainty-based forecast products. This was developed based on feedback from potential users from a questionnaire and workshop. The requirement included an implementation plan for any products identified, including a consideration of training and IT requirements, possible quick wins, and recommendations for follow-on projects to support the further integration of forecast uncertainty into fluvial flood forecasting and warning procedures.

The main phase of the project started in 2007 with the aim to develop a strategy for interfacing probabilistic rainfall forecasts with operational hydrological forecast models and flood warning procedures, drawing upon information from the user requirement study, and the recommendations from related projects. The objectives of this stage included development of a joint Met Office–Environment Agency proposal on the interfacing of probabilistic rainfall forecasts to operational hydrological forecasting models and flood warning procedures.

In 2007/08, a second phase of the project implemented a web-based operational trial of probabilistic precipitation forecasts for the Environment Agency. This trial included the provision of MOGREPS NAE (North Atlantic and European configuration of the Met Office Global and Regional Ensemble Prediction System) based probability of exceedance maps and stacked probability charts for predefined areas and rain accumulation thresholds. Following completion of the trial in December 2007, a workshop was held to review feedback and clarify aspects of the user requirement. A third phase of the project, started in October 2008, will deliver an operational, web-based service from the end of April 2009, providing a range of MOGREPS NAE and STEPS based probabilistic precipitation products (stacked probability charts) similar to those trialled in 2007.

The project is envisaged as the first of a series of developmental steps towards the integration of rainfall forecasting uncertainty into fluvial forecasting models and flood warning procedures.

Probabilistic Coastal Flood Forecasting: Forecast Demonstration and Evaluation

This project was a major study, led by HR Wallingford, and including the Met Office and Proudman Oceanographic Laboratory. The project started in March 2006 and finished in January 2009. The main aims were to review and develop existing methods for coastal flood forecasting, including offshore and nearshore modelling, and considering ensemble forecasting. The project included several modelling elements which were grouped under four headings, any or all of which could be developed further:

- Surge ensemble modelling for all of the UK, run in near operational mode.
- Temporary wave ensemble modelling specific to the South East Irish Sea, for demonstration use.
- Wave transformation and overtopping models specific to the South East Irish Sea, for demonstration use.
- Generic handling of a large number of uncertainties associated with nearshore waves and overtopping.

A real-time demonstration of the system provided distributions of surge, sea level, offshore waves, nearshore waves and overtopping rate at 15-minute intervals, updated 12 hourly.

The project included a pilot study for the area from Fleetwood to the Dee in North West England, consideration of forecast evaluation, and a scoping study for integration into NFFS for which the main tasks included:

- development of nearshore and coastline models for the chosen area, following the recommendations from the model evaluation;
- linking and incorporation of the new models into a pilot forecasting system;
- demonstration of the system at the trial sites.

FD2114: Review of the effects of land use on flood runoff generation

The FD2114 project, led by Professor Enda O'Connell of Newcastle University, provided a comprehensive review of the information available about the effects of land use on flood runoff generation (O'Connell *et al.* 2005). The conclusions of the review were that, while there was evidence of land use and land management effects at small plot scales, taking this information to larger scales involved significant uncertainties such that it was difficult to distinguish effects at larger catchment scales.

The FD2120 project 'Analysis of historical data sets to look for impacts of land use and management change on flood generation' was a direct result of research recommendations made in FD2114 and attempted to identify changes in rural catchment responses using modern data analysis and modelling methods. A group of 10 catchments were identified that had good hydrological data and where soil and agricultural information suggested that they had been subject to change in land use and management and should be susceptible to change. From the analyses carried out, it was suggested that there may well be effects of rural land management on flood runoff generation but that any effects were obscured by the uncertainties in the catchment input and discharge measurements and the long-term variability in climate drivers.

2.2.3 Findings from consultations

Introduction

During the period 18 December 2008 to 16 January 2009, the consultation meetings and telephone conferences listed in Table 2.4 were held as part of this project.

Table 2.4 Summary of consultation meetings and telephone conferences.

Consultees	Date	Location/method
Southern Region	18 December 2008	Telephone conference
Midlands Region	19 December 2008	Solihull
North East Region	22 December 2008	Leeds
Anglian Region	6 January 2009	Telephone conference
Thames Region	8 January 2009	Telephone conference
Thames Barrier	13 January 2009	Telephone conference
EA Wales	13 January 2009	Telephone conference
South West Region	14 January 2009	Telephone conference
North West Region	16 January 2009	Warrington

The consultations were undertaken by Atkins in collaboration with representatives of the Project Board, and built upon information already gathered during 2006 and 2007 from the consultations and workshops on the 'Probabilistic Flood Forecasting Scoping Study', and during 2008 on the 'Hydrological Modelling with Convective Rainfall' project.

The majority of meetings lasted for 2–3 hours, with the shortest lasting just under 2 hours, and the longest lasting about 4 hours. Between 2 and 4 Environment Agency staff participated in each meeting. A detailed agenda was issued in advance of each meeting, for which the main topics for discussions were as follows:

- Sources of uncertainty – views on the relative importance of different sources of uncertainty in different forecasting situations, and for different types of models and forecast lead-time requirements.
- Case studies – the selection criteria for case studies during Phase 2 of the project.
- Relevant studies – ongoing regional studies (if available) on sources of uncertainty in fluvial flood forecasting models.
- Integrated catchment models – experience with the performance of integrated catchment models combining rainfall–runoff, flow routing and/or hydrodynamic components.
- User requirements – for the Real-Time Modelling Guidelines and other project outputs.

This section summarises the main findings from these consultations, while a more detailed summary is provided in Appendix A of this report.

Sources of uncertainty

An early suggestion from the Project Board was that the project could investigate the practicability and benefits of applying probabilistic techniques for the following sources of uncertainty:

- Catchment averaging of raingauge data
- Validity of rating curves
- Model calibration (hydrodynamic models)
- Model calibration (rainfall–runoff models)
- Representation of floodplain storage
- Representation of antecedent conditions
- Representation of ungauged inflows
- Influence of structure operations.

Table 2.5 summarises the main findings regarding each of these sources of uncertainty.

Table 2.5 Summary of key issues raised on sources of uncertainty.

Item	Key issues raised
Catchment averaging of raingauge data	Range of approaches used with varying sources of uncertainty (Thiessen, Thiessen modified, weights, region-wide); radar rainfall data not widely used quantitatively at present; NWP/nowcast outputs required for some fast response catchments
Validity of rating curves	Uncertainty over accuracy of high flow end of ratings for natural sections (and some structures); seasonal influences on ratings (3 regions); mobile river beds (2 regions)
Model calibration (hydrodynamic models)	Determining appropriate roughness coefficients; representing seasonal changes in roughness; tidal influences on calibration (2 regions); datum issues (1 region); wave-speed estimates in flow routing models (2 regions); understanding/improving the performance of 'monolithic' models; deciding where/how/whether to implement real-time updating; representation of abstractions/discharges (if relevant); problems at model boundaries
Model calibration (rainfall–runoff models)	Model structural/conceptualisation/catchment size issues; performance measures used (linking to type of rainfall event); usually a limited number of parameters important for flood flows; performance outside range of calibration; most appropriate performance measures to use
Representation of floodplain storage	Uncertainties in roughness coefficients, flow paths and survey data (however, mainly for the lower reaches of some large catchments); performance for multiple events
Representation of antecedent conditions	Drift in models over time, particularly in the transition from summer to winter months; representation for smaller/moderate events; representation of evaporation (1 region), snowmelt component (2 regions), groundwater component (3 regions)
Representation of ungauged inflows	Role of choice of approach (scaling, parameter transfer etc); catchment conceptualisation/magnitude relative to main channel flows
Influences of structure operations	Parameterisation of control rules; departure of rules from design/published values; lack of real-time telemetry data on structure settings; representation of off-line storage (2 regions); representation/influence of reservoirs (where no model is included); operational problems (gate failures, blockages, flood defence breaches etc)

Following the general discussions on sources of uncertainty, each region was asked, if possible, to rank these sources in order of importance. Table 2.6 summarises the replies which were given, with a score of 1 being the highest, and 8 the lowest.

Table 2.6 Most important sources of uncertainty (1 = highest, 8 = lowest, c = depends on the catchment, forecasting issues, locations of forecasting points, lead-time requirements etc, shading shows higher 3-4 entries).

Source of uncertainty	Anglian	Midlands	North East	North West	Southern	South West	Thames	EA Wales
Catchment averaging of raingauge data	1	1	6	1	2	c	3	1
Validity of rating curves	4	4c	7	3	c	2	2	2
Model calibration (hydrodynamic models)	2/3	5c	5	5	c	c	c	2/3
Model calibration (rainfall–runoff models)	2/3	2	1	4	1	3	1	2/3
Representation of floodplain storage	c	3	4	c	c	c	c	c
Representation of antecedent conditions	c	8c	3	c	c	1	c	4
Representation of ungauged inflows	5	6c	2	2	3	c	c	c
Influence of structure operations	c	7c	8	c	c	c	c	c

The tentative conclusion from these discussions was that the following three sources of uncertainty were considered particularly significant:

- Catchment averaging of raingauge data
- Validity of rating curves
- Model calibration (rainfall–runoff models).

However, consultees were also keen to emphasise how uncertainties associated with hydraulic factors are still an issue in the lower reaches of a catchment, and that there can be specific forecasting situations which are exceptions to the order shown in the table (e.g. the Thames Barrier).

Another general comment was that the choice of an appropriate model structure and catchment conceptualisation can significantly reduce uncertainty. The interplay between model performance, and data assimilation, was also noted, with some assimilation techniques also causing erroneous outputs in some situations.

Case studies

The project includes scope for four case studies and the following general types of case study were discussed at each meeting:

- Two complex linked integrated catchment models, combining rainfall–runoff, flow routing and/or hydrodynamic models (e.g. based on PDM/KW/ISIS, MRCM/DODO/ISIS).
- Two simpler examples to illustrate key concepts: for example a rapid response catchment with rainfall and other sources of uncertainty, and a well-defined river reach including structures, rating curve uncertainty and other factors.

The intention would be that at least one of these case studies should use a data assimilation technique. The general consensus from the consultations seemed to be that the choice of types of case study seemed sensible, and would provide examples which will be of interest and relevance to all regions. There was also a general view that the case studies should be chosen on technical grounds (i.e. specific types of models and forecasting situations), rather than other criteria such as risk, or current model performance regarding level of service. The only additional suggestion for a type of case study was for a confluence flooding problem.

Each region was also asked for suggestions on specific case studies to consider, and these are summarised in Figure 2.7, while Table 3.1 in Appendix A of this report provides more details.

Figure 2.7 Regional suggestions for possible case studies.

Relevant studies

Table 2.7 summarises the main studies which were noted during the consultations.

Table 2.7 Summary of main studies noted during the consultations.

Item	Study
Recent operational studies (past 1–2 years)	Time Lagged Ensembles
	Peak Level Forecast Range Plots
	'Poor Man's Ensembles'
Regional Studies	Rainfall–Runoff Model Sensitivity Tests
	Rainfall Actual Sensitivity Tests
	Performance Measures/Level of Service studies
	Reservoir Influences
National Studies	FEWS User Day 16–17 October 2008
	Error Correction Workshop
	Communication and Dissemination of Uncertainty
	Fast Response Catchments study
	Data Assimilation R&D studies
	Coastal Flood Forecasting project
	Skill Scores project

Copies of reports and samples of outputs were provided as appropriate and have helped to inform the preparation of this report.

Experience with Integrated Catchment Models

All regions reported experience with developing and using integrated catchment models. Table 2.8 summarises the types of models which were described as being currently in use or being tested with a view to implementation operationally (the numbers indicate the number of regions reporting use of each type of model in at least part of the region).

Table 2.8 Indicative summary of current uses of integrated catchment models for catchments where models are available.

Method	All catchments	Many/several catchments	Some catchments
PDM/ISIS	1	2	
PDM/KW/ISIS			2
PDM/KW		1	1
MCRM/DODO		1	
MCRM/DODO/ISIS			1
NAM/MIKE11		1	
PRTF/ISIS		1	1
PRTF/VPMC/ISIS			1
TCM/ISIS	1		

Several regions also described the use of rainfall–runoff models alone for fast response catchments and/or lower risk catchments, and the use of level to level correlations. The types of rainfall–runoff model currently in use included PDM, HYADES, MCRM, TCM, NAM and PRTF models.

The current approaches to real-time updating in each region were also discussed, and Table 2.9 summarises the methods that were reported.

Table 2.9 Indicative summary of current uses of approaches to data assimilation for catchments where integrated catchment models are available.

Method	All catchments	Most/several catchments	Some/a few catchments
FEWS ARMA, CEH ARMA, MCRM AR	1	4	6
PDM state updating		2	1
Manual – snow			1
Manual – groundwater			1
Manual – soil moisture deficit		1	
Manual – gate settings			1
ISIS GAUGE			1
MIKE11 error correction			
MIKE11 state updating			1
PRTF updating		2	

User requirements

The discussions showed a wide range of views on the requirements from the guidelines and uncertainty framework. Some general conclusions were that:

- The guideline document should set out a national approach to how the selected probabilistic techniques should be implemented, with a general feeling (with one exception) that it should not be too prescriptive.
- The document should be linked to, or supplemented by, the NFFS case studies, so that a practical implementation can be viewed, and that this should be combined with generic 'lessons learned'.
- A balance needs to be kept between providing long lists of choices and options, and providing only a small number of approaches (or a single approach) which is not generally applicable to other catchments or regions.
- It was generally agreed that the various options and choices should be guided by the scale of the flood risk they are designed to mitigate, with a risk matrix approach generally favoured (linking into Flood Warning Levels of Service), but with other possibilities including the number of properties at risk, or the presence of a Severe Flood Warning.
- Ideally, regional differences should be considered; for example between regions with an extensive coverage of integrated catchment models, where the current focus is on improving the performance of existing models, and regions with only a few models, where the focus is on extending the coverage of models.

2.3 Summary of chapter

Some key points from this chapter include:

- There is considerable research under way internationally on developing techniques for estimating uncertainty in fluvial flood forecasting models (although the emphasis is often on uncertainty in rainfall forecasts).
- Methods which have been used or trialled operationally in fluvial flood forecasting models include quantile regression (USA), various forms of Kalman filtering (including extended and ensemble approaches), Bayesian Model Averaging (Netherlands), Monte Carlo approaches (Netherlands, Italy), and DBM approaches (Scotland).
- There are several Environment Agency guidelines and work instructions which are relevant to this project, including the Real-Time Modelling Guidelines, the Flood Warning Levels of Service work instruction, and studies on performance measures (although the focus is usually on deterministic forecasting).
- It will also be useful to consider the outputs from other Environment Agency R&D projects on probabilistic forecasting methods (covering rainfall forecasting and coastal flood forecasting). The Flood Risk Management Research Consortium (FRMRC2) Good Practice Guidelines should also provide useful ideas to assist in development of the uncertainty framework.
- The consultation exercise suggested more than 20 potential catchments for use in the case studies, and that the most important sources of uncertainty to consider include catchment averaging of raingauge data, the validity of rating curves, and model calibration (for rainfall–runoff models). Real-time updating is also important to consider.

3. Sources of uncertainty

3.1 Introduction

This section describes the main findings under Task 1.2 of the project, which is defined as follows:

> *To review and investigate which additional sources of uncertainty should be considered to gain a fuller (quantified) understanding of uncertainties in the flood forecasting process and to define in which situations/scales this may be beneficial. Particular focus should be placed on the aspects other than rainfall uncertainty, such as uncertainty associated with hydrologic, routing and hydraulic components*.

Section 3.2 considers findings from previous studies on the main sources of uncertainty in fluvial flood forecasting models. Section 3.3 then considers the most common types of fluvial flood forecasting models which are used operationally by the Environment Agency, and the sensitivity of model outputs to model parameters and other factors (e.g. antecedent conditions).

Initial findings are also presented of the practical aspects of implementing model parameter sampling schemes within NFFS, together with considerations of model run times and other possibilities, such as model emulators.

3.2 Review of previous studies

This section considers findings from a range of previous research and operational studies into the uncertainty in fluvial flood forecasts. Table 3.1 shows the main sources of uncertainty which were identified from discussions during the Probabilistic Flood Forecasting Scoping Study (Environment Agency 2007). However, only some of these factors are amenable to analysis and, following discussions with the Project Board, it was agreed that the following main sources of uncertainty would be considered in this review:

<u>Rainfall–runoff models</u>

- Rainfall forecasts
- Catchment averaging of raingauge data
- Representation of antecedent conditions
- Representation of ungauged inflows
- Model calibration (rainfall–runoff models)

<u>General</u>

- Validity of rating curves

<u>Flow routing and hydrodynamic models</u>

- Model calibration (flow routing and hydrodynamic models)
- Representation of floodplain storage
- Influence of structure operations

Table 3.1 Sources of uncertainty identified during the Probabilistic Flood Forecasting Scoping Study (2006–2007).

Component	Typical sources of uncertainty
Catchment averaging procedures (raingauge inputs)	• Representation of physical processes (topography, elevation etc) • Type of rainfall event (convective, frontal, orographic etc) • Raingauge density and distribution • Instrumental problems at one or more of the rain gauges used
Choice of model type and structure	• Lumped, semi-distributed, distributed rainfall inputs • Representation of catchment runoff processes • River channel and floodplain representation • Under/over parameterisation (parsimony) • Flood defence loading/fragility (if represented) • Gate operations • Representation of ungauged inflows • Representation of abstractions/discharges • Representation of groundwater influences
Model calibration	• Effectiveness of optimisation routines • Choice of optimisation criteria • Availability of sufficient high flow events for calibration • Skill of person calibrating the model
Operational	• Changes in catchment/channel characteristics since model was calibrated • Use of different input data streams from those used in the original model calibration (e.g. radar rainfall or forecasts instead of raingauges) • Events outside the range of the model calibration • Model stability problems • Representation of initial/antecedent conditions • Representation of snowmelt (if applicable) • Instrument/telemetry downtime problems (rainfall)
Real-time updating procedures	• Appropriateness for the type of model used • Sophistication of calibration software • Quality of the high flow data used both for calibration and in real time • Event-specific problems (backwater, bypassing, debris etc) • Instrument/telemetry downtime problems (flows)

Note that rainfall forecasts are included since the uncertainty framework will also consider ensemble rainfall forecasts, although the technical aspects of generating ensemble flood forecasts from these inputs are being considered separately in the 'Hydrological Modelling with Convective Scale Rainfall' project, and will not be considered on this project. Real-time updating (data assimilation) is also an important consideration, but is discussed later.

3.2.1 Rainfall–runoff models

Rainfall forecasts

Rainfall forecasts are used in rainfall–runoff models to extend the lead time of flood forecasts beyond the time at which measured rainfall has an effect on the modelled flood response, and can be a dominant source of uncertainty in flood forecasts at higher lead times. Particularly at longer lead times, rainfall forecasts can have significant magnitude and spatial location errors associated with them and these will be different for the main storm types: convective, orographic and frontal.

At lead times greater than the response time from observed rainfall, uncertainties in the prediction of rainfall and temperature will dominate the uncertainties in the response, increasing in importance as the lead time of the forecast increases. It is clear that how important each of these sources of uncertainty is will depend on the relation between the desired warning lead time and the response times of each of the contributions (e.g. Lettenmaier and Wood 1993).

For the forecaster to usefully use forecast values at longer lead times, it is important that the influence of the uncertainty in the inputs is quantified. A logical approach to this quantification is to explore the sensitivity of the forecast values to variable inputs. There are several operational approaches used, ranging from *ad hoc* application of differing input scenarios to ensemble forecasting, and these are described later in this report.

Environment Agency

The rainfall forecast products currently used with rainfall–runoff models by the Environment Agency are STEPS, NWP and MOGREPS. STEPS is a radar extrapolated forecast of rainfall that is blended with NWP rainfall, the latter gaining greater weight with increasing lead time. It is a 2 km product out to a lead time of 6 hours. An ensemble form of STEPS is planned to be released in 2009.

NWP rainfall is now produced using a 4 km resolution weather model and the Environment Agency receives this out to 1½ days as 15-minute rain accumulations (and rain-rates); there are four NWP model runs a day.

MOGREPS is a rainfall ensemble product produced operationally by the Met Office. It is being trialled operationally for use in probabilistic flood forecasting within the ‘Hydrological Modelling with Convective Scale Rainfall’ project. A drawback is its coarse resolution (24 km) but it allows the Environment Agency to prepare for future higher resolution ensemble rainfall products and their use for probability flood forecasting.

The operational availability of high-resolution NWP model outputs has been a significant advance in recent years, moving from 12 to 4 km and with 1.5 km planned. This has been accompanied by studies demonstrating the improvements in accuracy and how these might benefit flood forecasting and warning in the future: for example see the case study of the Carlisle flood reported by Roberts *et al.* (2009).

The ‘Hydrological Modelling with Convective Scale Rainfall’ project (see Section 2) is assessing the 1.5 km product for the 2007 summer floods using lumped and distributed rainfall–runoff models for catchments across the Avon and Tame in the Midlands Region. Pseudo-ensembles are being generated to emulate (at least at a functional level) the future availability of ensembles at this resolution, so that experience can be gained in probabilistic flood forecasting via an ensemble approach.

International experience

There have also been many other studies of the use of ensemble rainfall forecasts with hydrological models. For example (see Section 2), one of the first such studies (EFFS) considered an ensemble of 52 forecasts up to 10 days ahead, with the ensembles treated as having equal probability when used to drive hydrological models (such as LISFLOOD in the EFAS system). Some attempts have also been made to make use of ensemble inputs in updating for probabilistic flood forecasting by constraining the outputs of hydrological models by real-time flow information within the GLUE methodology (e.g. Pappenberger *et al.* 2005a). However, it is difficult to constrain individual rainfall ensemble members based on rainfall or radar observations because the resulting weights cannot be carried over to a new set of members when the ensemble is rerun at the next forecast interval.

Table 3.2 summarises a range of approaches used internationally for considering input (rainfall forecast) uncertainties in operational flood forecasting systems. All the approaches are constructed as part of the process of making a forecast. Where a standardised set of scenarios is combined in an empirical ensemble, this is generally sampled in one of the first steps of the data processing cascade, with the following steps being configured to simply loop over the available samples.

Table 3.2 Examples of empirical methods for considering input uncertainties applied in operational forecasting systems.

Approach	Forecasting system	Description
What-if scenarios	Several	Support of what-if scenarios is provided in almost all operational forecasting systems to investigate the influence of uncertain (usually meteorological) inputs. Scenarios are implemented during the forecast process and typically apply multipliers to the forecast rainfall, or allow the forecaster to input a user-defined forecast rainfall profile based on expert judgement
Standardised multipliers on meteorological and hydrological inputs	FEWS-Rhine Federal Office for the Environment, Switzerland (FOEN)	In the current operational forecasting system used for the Rhine basin in Switzerland (Bürgi 2002), a standardised what-if scenario is applied. The normal forecast derives temperature and precipitation inputs from the MeteoSwiss 7 km NWP model for the shorter lead times, falling back to the ECMWF deterministic forecast for the longer lead times. A mini ensemble is created by setting user-defined multipliers on forecast precipitation and temperature. An additional multiplier is set for post-processing the hydrograph to allow for an indication of increasing error with lead time when there is neither snow nor precipitation. Figure 3.1 gives an example of the results of this three member ensemble
Combinations of meteorological input products	National Flood Forecasting System (NFFS) Environment Agency, UK	In NFFS two deterministic rainfall forecast products are available: a radar-based nowcast product and a NWP forecast product. Under normal forecasting conditions these time series are merged, with the radar-based estimates having priority out to the maximum lead time available, then falling back to the NWP forecast and finally to a zero rainfall profile. To explore the influence of each of these inputs, a set of standard what-if scenarios has been defined, using different combinations of input products. The combinations explored include the default merged rain profile, a radar-only forecast, an NWP-only forecast, as well as a forecast using zero future rainfall (Figure 3.2)
Re-sampling best guess and 5% and 95% confidence interval precipitation forecast	HPC QPF Sampling NCRFC, National Weather Service, USA	The Hydro-meteorological Prediction Centre (HPC) in the USA produces both best guess quantitative precipitation forecasts and also 5% and 95% confidence interval forecasts. These are sampled using the first 24, first 48 and first 72 (60 in the case of Confidence Interval (CI) forecasts) hours of precipitation data to obtain a nine-member empirical ensemble precipitation input. This is used in the operational forecast as an input to a nine-member hydrological ensemble forecast (Halquist 2006)

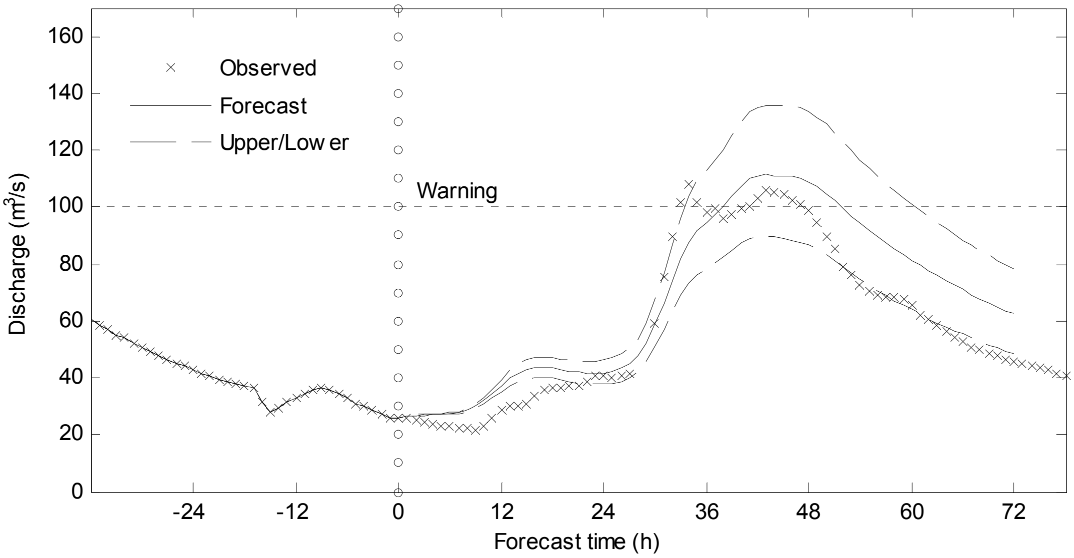


Figure 3.1 Forecast for the Emme at Emmenmat in Switzerland, showing the default forecast, as well as upper and lower scenarios using user-defined multipliers on flow, precipitation and temperature.

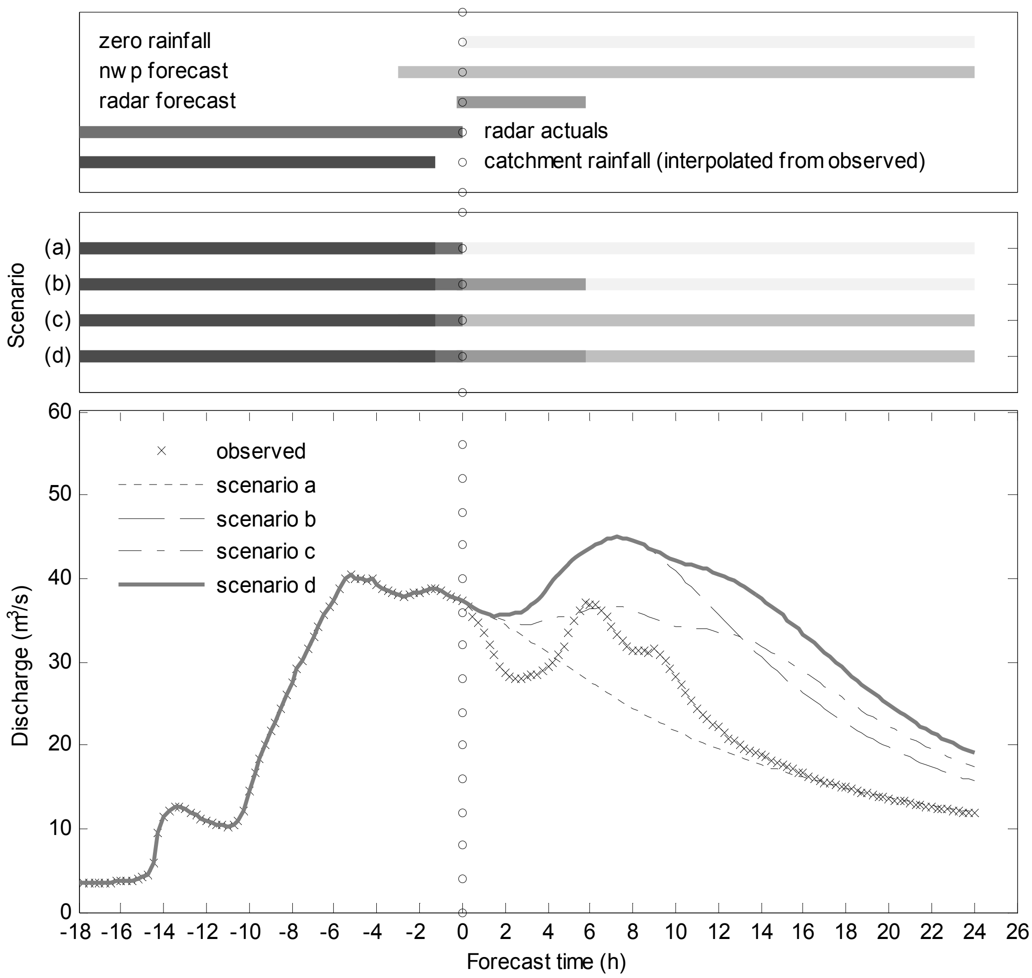


Figure 3.2 Standard set of scenarios used by the Environment Agency. The different forecast inputs are shown in the top figure. These include observed catchment rainfall, radar actuals, radar forecasts (lead time 6 hours), the NWP forecast (lead time 36 hours) and a zero rainfall profile. The middle figure shows four scenarios created using differing combinations of these inputs. The lower figure shows the response to these scenarios at the gauge of Gargrave on the River Aire for a forecast on 7 January 2005 at 23:00.

While the use of numerical weather prediction model outputs for deriving boundary conditions in the forecast has been a significant development in extending the lead time of hydrological forecasts, it is recognised that there are also considerable uncertainties in these weather predictions. To address these, ensemble prediction systems have been established, with the first of these systems becoming operational in the early 1990s. Examples are the ECMWF-EPS system operated by the European Centre for Medium-range Weather Forecasts (Molteni *et al. 19*96) and the GEFS system operated by the US National Centre for Environmental Protection (Tracton and Kalnay 1993). These are global ensemble prediction systems, predicting the evolution of the weather with an emphasis on medium-term predictions (5–15 days lead time). Ensemble forecasts are generated by perturbing the initial conditions, assumed *a priori* to be equally likely, and computing the evolution of the meteorology due to these perturbed initial conditions.

Table 3.3 provides some examples of ensemble prediction systems used in driving the hydrological ensemble forecast from selected operational forecasting systems utilising the Delft FEWS framework. This shows that, in several cases, multiple ensembles are considered. The first two of these EPS are relatively straightforward, with all ensemble members being uniform in length, and with uniform spatial and temporal resolution. The SRNWP-PEPS ensemble is an exception to this. This is in effect a multi-model ensemble, with differing spatial and temporal resolutions, as well as differing domains for each ensemble member. This variability is quite challenging to run, with the run properties for each member being adapted dynamically by Delft FEWS depending on the properties of that provided. The variability in lead time also creates difficulties in interpreting the results using standard statistical parameterisations as the number of ensemble members to consider will differ with lead time.

Table 3.3 Overview of meteorological ensemble applied in operational forecasting systems using the Delft FEWS framework.

Ensemble system	Forecasting system	Description
ECMWF-EPS	FEWS-NL Rhine & Meuse Catchments, Institute for Inland Water Management and Waste Water Treatment, the Netherlands	Global Ensemble Prediction System (EPS). The current EPS has a horizontal resolution of about 40 km, and has 51 members, of which the first member is the control run. The lead time of the re-sampled EPS used here is 240 hours at a resolution of 12 hours
COSMO-LEPS	FEWS-NL Rhine & Meuse See above BfG, Rhine Catchments, Federal Institute of Hydrology, Germany Po, ARPA-SIM, Bologna, Italy FOEN (experimental), Federal Office for the Environment, Switzerland	Limited-area Ensemble Prediction System. This 16-member EPS is obtained by running a non-hydrostatic limited area model, nested on the members of the ECMWF-EPS ensemble. The ECMWF-EPS ensembles used in providing the 16-member forecast are obtained through a cluster analysis of the full 51 EPS members for three ensuing forecasts. The resulting 10 km resolution ensemble is much better suited to resolving severe weather at small scales (Marsigli *et al.* 2005)
SRNWP-PEPS	FEWS-Rhine (experimental) See above	This ensemble of short-range NWP products is actually a multi-model, or poor man's ensemble (Quiby and Denhard 2003). The ensemble is constructed using the deterministic high-resolution NWP models from participating Meteorological agencies across Europe, with up to 21 ensemble members being available at any one time (depending on how many contributing deterministic forecasts are available). The lead time of each member varies, as well as the resolution and the spatial domain
MOGREPS	NFFS (T46, experimental)	The Met Office Global and Regional Ensemble Prediction System (MOGREPS) is an ensemble system that produces uncertainty information for short-range forecasts, up to 2 days ahead. It focuses on aiding the forecasting of rapid storm development, wind, rain, snow and fog

Figure 3.3 provides an example of outputs from both the ECMWF-EPS ensemble and the COSMO-LEPS ensemble for the Rhine at Maxau for the same forecast base time. In the upper two plots the raw ensemble outputs are shown, that is 16 members for the COSMO-LEPS ensemble and 50 members + control run for the ECMWF-EPS ensemble. The lower plots show the parameterised ensemble results, showing the minimum and maximum of all members, the median and the interquartile range. The ensemble results can be seen to be quite different for these two NWP ensembles. These differences are attributed to scale effects, as this gauge is towards the upper end of the catchment. For gauges further down in the catchment these differences are less pronounced. There is for both also an under-representation of spread at the shorter lead times. This is an obvious consequence of considering only the meteorological uncertainties, as described through the meteorological ensemble forecast.

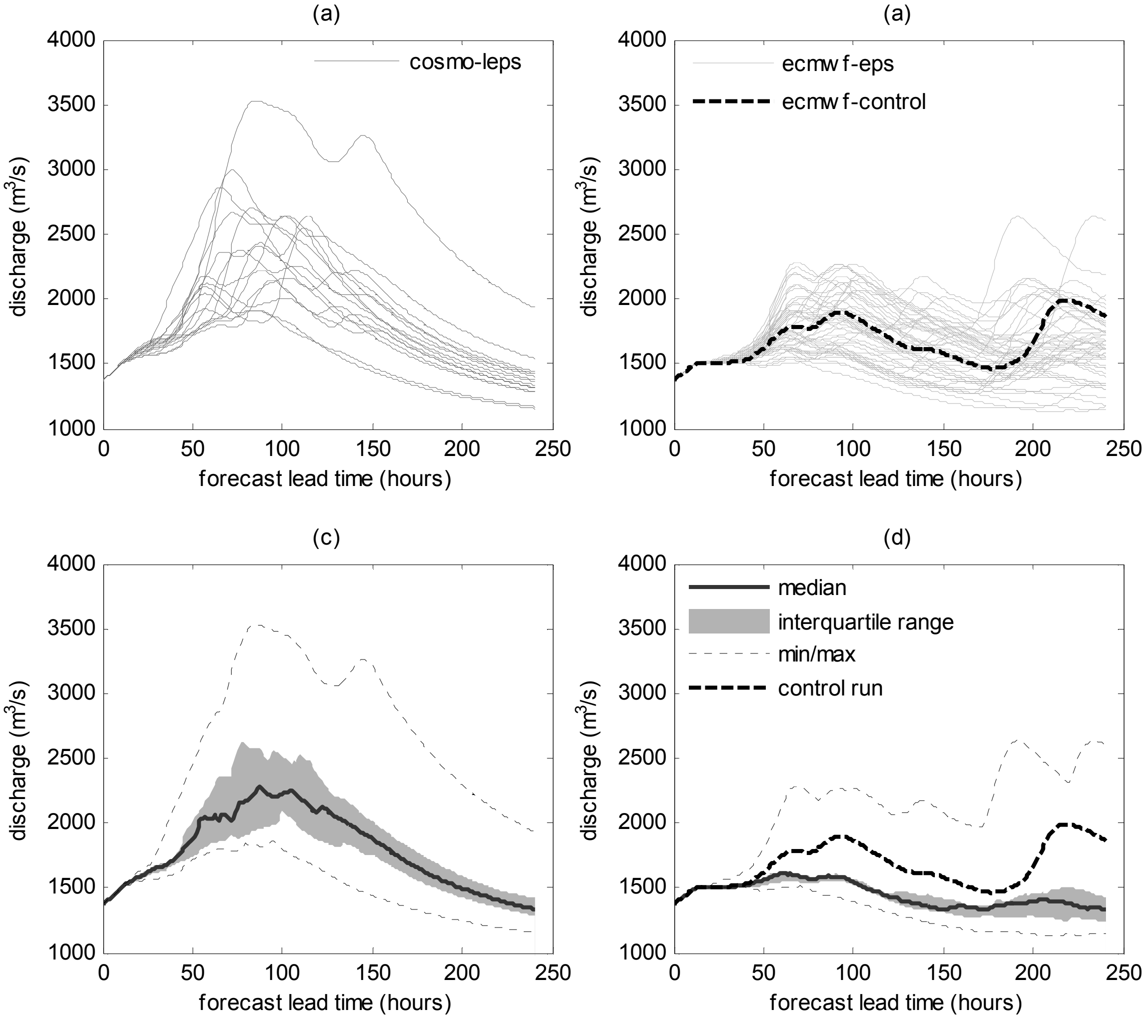


Figure 3.3 Example of two NWP ensemble forecasts for the River Rhine at Maxau, COSMO-LEPS and ECMWF-EPS. Both forecasts have the base time of 06:00 UTC 1 March 2007 ; (a) and (b) show the raw ensemble outputs, while (c) and (d) shows the parameterised ensemble outputs.

Catchment averaging of raingauge data

The estimation of catchment average rainfall from raingauge data is a well-researched topic. Methods range from simple linear weighting methods, such as Thiessen polygons and its variants, to more complex multiquadric surface-fitting and Kriging methods that encompass uncertainty estimates. The project consultation meetings have highlighted catchment rainfall as the most significant source of uncertainty, serving to confirm this from a practioner's viewpoint.

Previous research has also shown that lumped conceptual rainfall–runoff models are very sensitive to having good estimates of catchment average rainfall as input. Also, the uncertainty in catchment average rainfall need not be consistent between events, and this is one reason why data assimilation methods are important in compensating for errors in observed inputs in any particular event as it happens.

The question of how best to estimate catchment average rainfall becomes broader when combining raingauge data with radar estimates of rainfall. For example, transient errors in the radar data can make the use of radar data in isolation problematic.

Recent research on this issue by Cole and Moore (2008, 2009a) discusses these problems and presents one way of combining the good spatial pattern information that radar provides with the better point estimates of rainfall magnitude provided by raingauge networks. The benefits to flood forecasting are demonstrated using both lumped (PDM) and distributed (G2G) rainfall–runoff models. Raingauge-only as well as merged radar and raingauge estimators are developed and compared for use with rainfall–runoff models. It is important to note that the multiquadric methods employed are available both within HYRAD, and in module adapter form for use within NFFS. Also of note is the link this research makes between integrated multiquadrics and the popular Thiessen weighting methods for deriving rainfall estimates averaged over grid-square and catchment areas.

Further insight into rainfall estimation accuracy can be gained through looking at studies using dense experimental networks of raingauges, such as that installed on the HYREX (HYdrological Radar Experiment) project, which was funded by the Environment Agency as part of their contribution to this NERC Special Topic. The HYREX network consisted of 50 tipping-bucket raingauges within the 135 km^2 River Brue catchment in Somerset, and research studies based on this dataset include the studies by Wood *et al.* (2000a,b) and more recent work by Villarani *et al.* (2008).

Similarly, empirical results are provided by Moore (2002) concerning the effect of raingauge density on the accuracy of rainfall estimation using the Environment Agency's dense raingauge network over London and the Thames Basin.

Representation of antecedent conditions

It is well known that antecedent conditions are important in flood runoff generation, which has a highly nonlinear relationship to input rainfall that is dependent on the antecedent state of a catchment. Even in flash flood situations in the UK, the importance of antecedent conditions is crucial. Although both the Lynmouth and Boscastle flood events occurred in summer (on 16 August in 1952 and 2004 respectively) the amount of runoff generated was increased by the fact that the catchment areas involved had already been wetted by rainfall in the preceding week. In larger catchments as well, the highest floods occur when heavy rainfalls fall onto already saturated catchments (as in the January 2005 Carlisle flood when runoff coefficients were estimated at 73% and in the summer 2007 floods that followed the wettest 3 months in England since records began).

This suggests that knowledge of antecedent conditions is needed in flood forecasting so that runoff coefficients in a major event can be reflected in the modelled flows more accurately. There are many different ways of assessing the antecedent state of a catchment. The first is by continuous simulation. This forms the basis of the conceptual rainfall–runoff models used widely by the Environment Agency within NFFS: namely PDM, TCM, MCRM and NAM. These are continuous soil moisture accounting models that aim to represent the changing responsiveness of catchments through water balance principles. This is also the approach that is adopted by the European Flood Alert System (EFAS) run at the Joint Research Centre (see Section 2). In this case the LISFLOOD model employed is run at coarser time and space scales for the whole of Europe, switching to shorter time-steps when a potential flood event is detected.

The Met Office MORECS/MOSES-PDM land surface schemes also provide estimates of soil moisture state at a spatial scale across the whole of the UK (Hough 2003). MORECS is updated every day and uses a 40 km grid. MOSES-PDM (Smith *et al.* 2006) now runs within the UKPP (UK Post Processing) suite at an hourly time-step and for a 2 km grid. Outputs from these models might be used as an index of antecedent wetness for a catchment, but should be used with care because the simple hydrological components of MORECS/MOSES-PDM may not be directly compatible with the hydrological models used in flood forecasting.

The MORECS approach is also used in the South West Region of the Environment Agency, but has been enhanced to include more raingauges in the estimation procedure, and to operate at a finer spatial resolution, typically using catchment-based polygons with areas in the range 50–400 km^2. These values are used in conjunction with PRTF rainfall–runoff models, and recent studies suggest that these locally derived values improve upon both the standard MORECS outputs, and MOSES values. The Catchment Wetness Index (CWI) is also used rather than the soil moisture deficit since this includes allowance for recent rainfall (API5) and appears to be more closely related to catchment runoff. Observed baseflows are also used as a surrogate for antecedent conditions in catchments where this has been shown to improve model performance.

A further possibility is to measure antecedent wetness directly. This is generally problematic because of the large spatial heterogeneities in soil saturation that are expected to be found in catchments that vary spatially in geology, soil type, land use and management, and the way that topography affects saturation. However, a useful index of antecedent conditions, at least in perennial streams, is the river flow itself.

River discharge is used to initialise the subsurface drainage stores of both lumped and distributed hydrological models. But this still leaves scope for variability in the near surface soil storage that will have an effect on the model flows being able to reflect the runoff coefficients observed.

An interesting use of the river flow itself as an index of wetness arises in the transfer function (TF) and DBM forecasting methods (e.g. Moore 1982; Young 1993, 2002; Young and Beven 1994; Romanowicz *et al.* 2006a, b, 2008). Data analysis methods were used in calibration to determine what the effective rainfall might be, and it was found that this was then strongly, but nonlinearly, related to discharge. This relationship can then be used in the prediction of effective rainfall in forecasting.

Representation of ungauged inflows

The mass balance issue associated with ungauged inflows imposes similar problems to the catchment averaging of raingauge data. Again, models with gains that are identified directly from historical data can allow for any consistent differences between upstream and downstream discharges (or levels, see Romanowicz *et al.* 2008) while data assimilation can compensate for variations from past events using real-time adaptation techniques.

The problem of representing ungauged lateral inflows to flow routing and hydrodynamic models has recently been reviewed for the Environment Agency under the ‘Rainfall–runoff and other modelling for ungauged/low-benefit locations’ project (Moore *et al.* 2008). Simple scaling and transposition methods working from nearby gauged locations and applying factors, based for example on area and Standard Average Annual Rainfall (SAAR) weightings, are commonly used and can be effective. They rely on similarity assumptions relating to runoff response and the forcing rainfall; these impact on accuracy.

Model parameter transfer provides a more sophisticated option and relaxes the former assumption, as the transferred rainfall–runoff model can be forced by the rainfall estimated over the ungauged catchment. Methods for transfer commonly require that model parameters are related to catchment properties, for example through regression or use of site-similarity measures; these approaches include estimates of model simulation accuracy.

A new approach pursued within the project is to formulate a conceptual–physical grid-based model where the model structure and properties are linked directly to spatial datasets on elevation, soil, geology and land cover; only a small number of area-wide parameters are left to be calibrated across the model domain. This approach led to the development of the Grid-to-Grid (G2G) Model, currently undergoing pre-operational NFFS trials within the Environment Agency. The capability to forecast 'everywhere' on a 1 km grid across England and Wales means that the model can be used to provide estimates of ungauged inflows in real time that reflect the morphology of the ungauged catchment and the forcing rainfall. The accuracy of the approach, while not being as good as for a model calibrated to a gauged catchment, is encouraging especially for catchments where flood response is strongly controlled by topography. Further developments and accuracy assessments are ongoing and planned. The flood forecasts from the G2G Model could be used as ungauged lateral inflows required by flow routing and hydrodynamic models employed to represent flow through a river network in NFFS.

Model calibration (rainfall–runoff models)

The principles for calibration of rainfall–runoff models are well established. For example, the manuals to the PDM and TCM (PSM) rainfall–runoff models (CEH Wallingford 2005a,b) contain practical user guides that aim to provide clear advice on model calibration. This includes the strategy for record selection for calibration and assessment as well as the sequential grouped approach to parameter estimation; daily as well as 15-minute time-steps can be used in the calibration process. A good understanding of the model is a prerequisite and practice is required to develop an acceptable level of skill. Emphasis is placed on intelligent manual adjustment of parameters supported by the hydrograph interactive visualisation facility, using automatic optimisation as a later stage refinement. This impacts on the quality of the final calibration of the simulation model.

The model time-step can also be a consideration, particularly for model performance as levels rise towards and through flood warning threshold levels. For a large, slowly responding river, an hourly time-step may be sufficient, while a shorter time-step is required on flashier catchments. Also, the frequency of updating can significantly affect the accuracy of updated forecasts for such catchments (Moore *et al.* 1993). With current Environment Agency telemetry systems, the minimum time-step which is feasible at present is 15 minutes (although some radar rainfall products are available at 5-minute intervals).

In contrast, the calibration of the updating parameters (state correction or ARMA error prediction) is a largely automated task employing a modified form of the Simplex automatic optimisation procedure. Constraints to ensure admissible ARMA model parameters are applied. The emphasis is on practical use, so more advanced facilities for plotting the objective function surface to explore parameter interdependence and associated dot and dash plots are not highlighted. Performance measures provided focus on RMSE and R^2 efficiency (and the error mean and variance) along with peak timing and magnitude statistics. But visualisation of the modelled hydrograph against the observed hydrograph is paramount when judging performance. The different models share a common generic calibration shell sometimes referred to as TSCAL (Time Series CALibration). A quick guide to objective functions is also provided so a model calibration can be tailored to different uses other than flood forecasting. There are no uncertainty estimates provided for the calibrated parameters.

Probabilistic flood forecasting introduces some interesting issues for model calibration. In the past, models have been calibrated by trial and error or by optimisation using a chosen performance measure without taking account of uncertainties in either inputs or output estimates, except in an informal way. It is now more widely recognised that uncertain inputs and outputs, as well as uncertain model structures, will affect the calibration process. In fact, except in the case where the model is fully identifiable from the available data (Young 2009), there may not be a clear optimal model parameter set, but rather many different sets that give acceptable simulations in calibration (e.g. Beven 2005, 2009).

This is currently a topic of considerable debate in the hydrological literature between those who promote Bayesian statistical parameter estimation methods (e.g. Mantovan and Todini 2006; Liu and Gupta 2007; Stedinger et al. 2008) and those who suggest that the modelling process involves epistemic as well as statistical errors, so that methods of statistical inference might give misleading results (e.g. Beven 2006, 2009; Beven et al. 2008c).

This debate is not really important, however, to the forecasting problem when the objective is to minimise the variance of the *n*-step ahead forecast. Clearly having a well-calibrated model might help towards this aim, but it is highly likely that the characteristics of the next flood event will be rather different to any of the events in the calibration period. Thus data assimilation and updating processes assume a much greater importance than in the case of simulation (at least while telemetry of sensor data or direct observation remains available during an event).

In forecasting, it is also possible to forecast river levels directly if a model is formulated to do so (e.g. Romanowicz *et al.* 2006b, 2008). This has the advantage of making direct use of the level measurements, for which uncertainty is generally small and, therefore, not introducing the uncertainty associated with the rating curve into the forecasting process. The disadvantage of this approach is it does not invoke any mass balance constraints, although this is difficult to justify in flood situations anyway, when we cannot be sure of a mass balance match between measured inputs and measured outputs for either rainfall-flow or flow routing problems because of measurement uncertainties. It also does not produce flow estimates where these are required to estimate velocities or the upper boundary conditions for a hydrodynamic routing model. In both cases, however, a rating curve (together with its uncertainty) could be introduced at this point to convert levels to velocities or discharge. In models that are constrained to maintain mass balance, a similar effect can be achieved by conditioning rainfall multipliers in applications to historical events (e.g. Kuczera *et al.* 2006).

3.2.2 Validity of rating curves

Rating curves are of critical importance in flood forecasting at least in terms of the operational models existing in the current regional NFFS configurations. They are used to calculate flows up to 'time now', which is then the start point for subsequent forecasts when error correction is applied

They are also crucial in the calibration of rainfall–runoff models and incorrect rating curves can give rise to wildly inaccurate flow values. A very good example of this occurred in the January 2005 event for Carlisle when the Environment Agency's existing rating curve was used to calculate flows at Temple Sowerby which is the primary upstream site from which forecasts for Carlisle were issued. The Environment Agency's rating curve (extended far beyond any spot flow gaugings) yielded a flow of 382 m^3/s, which gave rise to predicted levels in Carlisle that did not overtop the defences, underpredicting the actual observed peak by approximately 1 m. After local remodelling of this gauge using 1D/2D methods (ISIS/TUFLOW) it was determined that the calculated flow should have been 926 m^3/s, which did give rise to the correct levels in Carlisle. This situation was exacerbated by there being entrainment banks local to the gauge which overtopped during the event. It was of course inappropriate to extend the rating curve beyond the bank levels in this case.

Incorrect rating curves can also give rise to effective percentage runoff values above 100% or indeed below 10% which would nullify the benefits of an accurate historical or forecast rainfall distribution. However, they may be less relevant when considering explicit relationships between recorded levels or rainfall totals and levels at forecasting points but such models do not exist within the current NFFS regional configurations. Rating curves are also important in general terms for current hydrological analysis such as the FEH (Flood Estimation Handbook) statistical approach, which is dependent on historical annual maximum flow (AMAX) series that in most cases are calculated via rating curves.

If it is required to forecast levels even in extreme events up to 0.1%AEP (as specified in 'Making Space for Water'), it is required that the associated rating curve is appropriate for the measured water level at that return period.

It must be emphasised that uncertainty associated with rating curves can always be reduced by simulating the extension of the relevant rating curve using hydraulic models as described above for Temple Sowerby. This process is described at length in the Environment Agency R&D Report W6-061 (Ramsbottom and Whitlow 2003), which considered simple hydrological techniques through to complex three-dimensional (3D) modelling. Best practice within the Environment Agency is to use 1D methods where the site is less critical and hybrid 1D–2D models when the site is designated as critical to accurate prediction of threshold crossings at high risk locations. For example, very good results have been obtained on the Thames floodplain at Sutton Courtenay and Farmoor using this approach even when more than half the flow bypasses the station.

3.2.3 Flow routing and hydrodynamic models

Model calibration (flow routing models)

The principles of model calibration are well established for flow routing models. For example, the manual to the KW channel-flow routing model (CEH Wallingford 2005c) contains a practical user guide that aims to provide clear advice on model configuration and calibration. This discusses configuration of a routing reach in terms of subdivision into sections, the location of ungauged lateral inflows and the form of scaling used, the form and location of overflows, and the form of wave-speed curve.

Once a configuration has been decided and parameters affecting the mass balance constrained on physical grounds, there is much more reliance on automatic optimisation than for rainfall–runoff modelling. The TSCAL calibration environment is used, as described previously in relation to PDM and TCM. For data assimilation ARMA error prediction is used with the parameters calibrated automatically and constrained to be admissible.

Both DODO used in Midland Region and the ISIS flow routing models (Muskingum and VPMC) rely on manual calibration. The DODO model was documented for the Environment Agency (then the National Rivers Authority) by CEH (then the Institute of Hydrology): see Wallingford Water (1994). Further details of the ISIS flow routing models are provided in the user manual to the software.

Routing models have also been developed that can work directly with water levels (e.g. Romanowicz *et al.* 2006a, b, 2008). As noted above, the observations of levels are generally of low uncertainty at all stages (as long as the sensor continues to provide information during a flood). Experience suggested that this usefully controls (but does not eliminate) the heteroscedasticity of the forecast variance (Romanowicz *et al.* 2008).

Model calibration (hydrodynamic models)

In general terms for almost all practical modelling studies, the calibration of hydraulic models for design purposes is reliant on comparison between peak levels, and a figure often quoted (e.g. for flood risk mapping applications) is that the recorded and simulated values should be within 250 mm. The timing of the simulated peak levels is usually of secondary consideration in this case, whereas it assumes much more importance for flood forecasting.

Experience tells us that a well-calibrated forecasting model should achieve peak level agreement within 150–200 mm without error correction and better than this (less than 100 mm) when error correction on level is performed.

It is in the context of practical ARMA-based error correction on either flow or level that the importance of time accuracy in the original calibration can be seen. For example, this has been particularly apparent in the real-time simulation of flows on the Bristol Avon.

Model calibration of hydraulic models is often done by physical argument (e.g. assumptions about roughness coefficients), followed by expert evaluation of the results in comparison to any observational data available. In part, this is because of the computer requirements of running optimisation codes with distributed hydrodynamic models (especially fully 2D models); in part, because of a lack of calibration data on spatial patterns of inundation; and, in part, because hydrodynamic models are assumed to be more secure in their physical representation of the system than in the case of rainfall–runoff models.

Representation of floodplain storage

Representation of floodplain storage and conveyance can be an important issue in flood forecasting particularly when a substantial amount of flow is conveyed or stored beyond bank top. It is often the case when older ‘design’ hydrodynamic models have been converted for flood forecasting purposes that LiDAR data were not available during the construction of the model which made representation of floodplain flow and storage particularly difficult. LiDAR coverage is now relatively extensive, which enables direct simulation of floodplain flows or storage via a 2D model that can be linked dynamically with an existing channel model.

It is not standard practice at present to use 2D or 1D/2D models within the NFFS with the notable exception of the Proudman Oceanographic Laboratory's Continental Shelf (CS3) model. The implication of this is that information gained from a 2D model would have to be used to inform the re-schematisation of the 1D forecasting model. The conclusion of this is that, as with rating curves, uncertainty in floodplain storage or conveyance can be reduced by additional modelling work.

Indeed, floodplain modelling is an active area of research within FRMRC and other research programmes, including comparisons of model outputs with observed inundation extents (e.g. Romanowicz *et al.* 1996; Aronica *et al.* 1998; Romanowicz and Beven 1998; Bates *et al.* 2004; Pappenberger *et al.* 2005b, 2006b, 2007a). The model performance can depend upon a range of factors, and uncertainties can arise from a number of sources; for example, in observations of flood extent (e.g. Pappenberger *et al.* 2005b); the accuracy of LiDAR data; the representation of the floodplain geometry and flow paths (especially in 1D models or coarse grid 2D models); the representation of the effects of floodplain infrastructure on flows (walls, buildings, culverts etc); and from spatial differences in effective roughness.

Influence of structure operations

Structure operations can also significantly affect the forecasting of levels either due to their failure to operate as expected or due to blockage, or because they are controlled by a third party such as British Waterways or a riparian owner (such as often occurs at mills).

In practical situations the influence of structures is often greater when backwater lengths associated with them extend further upstream as occurs often during low-gradient watercourses.

The operation of structures is often critical in terms of operational use of washlands or sacrificial storage and there are many examples of these systems such as the Sale and Didsbury storage basins on the Upper Mersey and the extensive network of washlands upstream of Lincoln in the Witham catchment. The operational rules associated with these structures can be immensely complicated and embroiled with historical arrangements made with riparian landowners such as exist for the Upper Mersey flood basins.

Also, in some cases the structure does not operate as envisaged due to manual intervention of operators for whatever reason and this can cause particular difficulties for flood forecasting.

The potential role of structure blockage as a source of flooding cannot be underestimated. To illustrate this point, Edenvale Young has undertaken a large number of pre-feasibility (250) and project appraisal studies for Powys County Council (40) and Caerphilly Borough Council. Of the 27 first stage project appraisals for Powys County Council, 22 or (81%) of the sites have blockage as the primary or secondary flooding mechanism in conjunction with high rates of flow. These studies are all on ordinary watercourses and a large number of the sites include trash screens, culverts and medium-sized bridges which are vulnerable to blockage and cause flooding. On a nationwide basis, the Welsh Assembly Government has calculated that approximately 60% of all flooding problems on ordinary watercourses in Wales relates to the blockage of culverts. This percentage seems to be holding true for subsequent groups of projects and for sites in Caerphilly Borough. In terms of blockage to culverts (not bridges) the percentage is slightly lower at 63%, which reflects the national average.

The implications of a prescribed amount of blockage at a particular structure can be simulated directly by a hydraulic model but it is unfeasible to simulate all blockage scenarios that may occur.

The conclusion of this analysis is that for some forecasting models – particularly those covering urban and heavily culverted areas – uncertainty in structure operation or blockage can never be eliminated and it is a factor that could usefully be simulated using probabilistic methods

3.3 Model-specific considerations

This section considers some typical sources of uncertainty for the types of rainfall–runoff, flow routing and hydrodynamic models which are currently used within the Environment Agency.

In developing the scope for this study the Project Board decided that event-based rainfall–runoff models – which cannot run in continuous mode – such as PRTF should not be considered, although the findings on uncertainty in catchment averaging of rainfall and rating curves will be very relevant to PRTF models. However, some broader discussion of the PRTF model is included here for completeness.

Section 3.3.1 discusses rainfall–runoff models while Section 3.3.2 discusses flow routing and hydrodynamic models. The focus of the discussion is on rainfall–runoff models, since the consultations identified this topic as a key source of uncertainty. Finally, Section 3.3.3 describes some practical considerations with the implementation of models on NFFS.

3.3 1 Rainfall–runoff models

Rainfall–runoff models of conceptual lumped form and based on continuous water accounting principles have many commonalities but differ importantly in their detail. Thus some general remarks can be made at the outset on typical sources of error affecting such models.

The main function of a rainfall–runoff model is usually to transform catchment average rainfall, and an estimate of potential evaporation (PE), to river flow at the catchment outlet, and models of this type are therefore associated with the following sources of uncertainty:

i. input

ii. model structure

iii. model parameter

iv. model state

v. output.

A dominant source of input uncertainty is in the catchment average rainfall estimated from raingauges, and possibly radar, and for future times from extrapolated forms of these estimates and from NWP models. This issue has been discussed previously.

The PE estimate is another source of model input uncertainty. The estimate used may range in complexity from a simple sine curve, through to annual (and maybe diurnal) profiles derived from MORECS, to the use of MOSES PE estimates. Experience suggests that rather simple approximations may suffice. This is possibly because of the way that modelled soil moisture deficits are replenished to a saturation level that wipes out the previous history of PE. It is also partly 'calibrated out' in the model process of converting from PE to actual evaporation as a function of soil moisture deficit.

A very recent and quite detailed operationally focused review of PE estimation has been carried out by CEH for Southern Region as part of an investigation to apply an extended form of the PDM rainfall–runoff model to groundwater catchments (Cole and Moore 2009b). This highlights differences in MORECS and MOSES PE estimates and also considers PE profiles in use with rainfall–runoff models within the NFFS. Recommendations are made on the PE estimates to be used in Phase 2 of the project. The sensitivity of rainfall–runoff performance to the PE estimate will be included in the Phase 2 assessment.

While rainfall and PE are the principal time series inputs to rainfall–runoff models, there may be other inputs deserving of inclusion in the modelled water budget. For example, the extended PDM (Moore and Bell 2002) for groundwater catchments can utilise time series of pumped abstractions; however, this is not yet available in the NFFS adapter form of the PDM. A similar capability has been developed for the TCM (in PSM) but also remains in research code form.

Snowmelt models can be embedded or linked to rainfall–runoff models and these have their own forms of input uncertainty. The most important input uncertainty is the estimation of the snowfall input, which is very problematic at the catchment scale and a major source of snowmelt flood uncertainty. Also atmospheric variables that control the melt of snow, particularly air temperature, are important in the maturation and depletion of the snowpack: fortunately, temperature is more accurately and easily estimated over catchment areas. CEH has carried out reviews and assessments of snowmelt models for the Environment Agency: these include the MCRM snowmelt model and the PACK model (standalone but used by the Environment Agency with the PDM) used in NFFS. For further discussion, see Moore *et al.* (1996, 1999), Bell and Moore (1999, 2000) and Bell *et al.* (2000).

Model structure uncertainty will always exist as a model is by definition a highly simplified mathematical representation of reality. This is particularly the case for a rainfall–runoff model as the processes operating within a catchment controlling the transmission and storage of water are complex and largely non-observable. This leads to rather simple representations of the water storage and transmission processes operating in a catchment. A particularly convenient approximation is the catchment-average formulation that lumped rainfall–runoff models assume. In particular the pattern of rainfall within the catchment is not taken into account: this is best thought of as model structure source of error.

Parameter uncertainty is intimately linked to model structure but is best associated with the uncertainty of the model calibration process, whether by manual or automatic means. In the same way that there is no 'valid' or 'correct' model structure, these terms also do not apply to the parameter set being sought. Thus 'model validation' is a misnomer and model assessment is preferred. The parameter set chosen will depend on the purpose of the model forecasts and in turn to the objective: this may be couched informally (e.g. visually acceptable fit) or formally (using objective functions which can be equivalent to performance measures). It is common for a 'good' rainfall–runoff model structure to suffer from problems of model parameter interdependence making the search for an 'optimal' parameter set (or sets) challenging. This leads to manual calibration being the preferred option for some rainfall–runoff models, at least in the initial stages, possibly using automatic optimisation as a refining and support tool. The modeller's skill level is likely to impact on parameter uncertainty, along with the dataset being used for calibration (in terms of the record length, the range of floods encompassed and quality).

The states of a model (such as the contents of stores representing soil-, ground- and channel-water) can be prone to uncertainty, most clearly due to errors in the rainfall input aggregating as errors in the store water contents. There is interplay between model structure, model parameter and rainfall input in the way this uncertainty in the state manifests itself. It is possible to correct for errors in the states through data assimilation, commonly using river flow observations. Similar principles apply when initialising a model for the first time. State correction is most successful at low flows when flows largely derive from a 'slow store' (typically representing groundwater) and a direct relation between storage and water release applies. The flow observation itself is subject to uncertainty and this 'output uncertainty' may be a major source, particularly as the river goes out of bank. Output uncertainty also impacts on model parameter uncertainty through the model calibration process.

Having broadly reviewed the sources of uncertainty in a general rainfall–runoff model context, attention will be turned to consider some specific research of particular relevance to the Environment Agency operational interests in this area.

The conceptual rainfall–runoff models used within the NFFS are the PDM, MCRM, TCM and NAM. These models were detailed and reviewed by CEH (then the Institute of Hydrology) for the Environment Agency under the R&D project 'Comparison of rainfall–runoff models for flood forecasting' (Moore and Bell 2001). The second part of the project compared the performance of these models, with the exception of NAM, on nine catchments of varied character spread throughout the regions of the Agency (Bell *et al.* 2001). An overview of the project is provided in Moore *et al.* (2000. Performance was assessed in simulation and updating mode using hydrograph plots complemented by R^2 efficiency and Threshold CSI (Critical Success Index) performance measures. The latter skill score measure judged the efficacy of a model to correctly forecast the exceedance of a set of flow thresholds, particularly relevant to the use of a forecast to trigger an alert level of a given severity. For some catchments radar estimates of rainfall were available: this allowed a comparison to be made of model performance as impacted by the radar, raingauge or merged radar–raingauge estimator used for catchment average rainfall. The project results provide important quantitative insight into comparative model accuracy and how this is affected by the rainfall estimate used as input and the nature of the catchment.

Discussions specific to each model used in NFFS follow, primarily relating to ease of use and drawn from the 'Comparison of Rainfall–Runoff Models' project. These discussions are particularly relevant to uncertainty due to model structure and parameters for each model.

PDM

Calibration of the PDM (Probability Distributed Model) is usually straightforward for any given model structure, but the variety of options among the model components means that it may not be easy to determine the best structure. A practical guide now supplied with the PDM provides clear guidance and recommends a standard structure. The Pareto distribution of soil storage depths has been found to provide a simple yet flexible description of soil moisture storage for most catchments. This should be the first choice, only experimenting with other distributions if problems are encountered. Partitioning of rainfall between soil storage and fast and slow response paths is generally achieved through a direct runoff to the fast path with simple recharge to the slow (groundwater) path. For some catchments (e.g. the Witham), direct runoff to the fast response path with demand-moderated recharge to the slow (aquifer) path can prove more successful.

A cascade of two identical linear reservoirs is usually appropriate for the fast response path. A quadratic or, more usually, cubic storage should be used for the slow response path. In general state correction is preferred to error prediction for forecast updating. The soil moisture storage, evaporation, recharge and runoff generation mechanisms in the PDM are interlinked and highly nonlinear, and the effect of changes in the associated model structures and parameter values can be difficult to predict. Use of the PDM at a daily time-step can be useful in determining slow response model parameters; this is now less useful due to improved graphics and support of very long datasets at a 15-minute time-interval. In general the PDM is relatively insensitive (robust) to initial soil moisture conditions. Although an initial period of warm-up is beneficial to the MCRM, the PDM is less demanding.

MCRM

The MCRM (Midlands Catchment Runoff Model) has a large number of parameters: 22 in the main rainfall–runoff part excluding the snow and reservoir components. However, once the initial conditions for an event are set up correctly the model can be more straightforward to calibrate than the number of parameters might suggest. Several of the parameter values should lie within a narrow range and so can be set to standard values initially. Calibration of the model can be divided into five parts: the groundwater, the soil store, the timing, the smoothing, and, finally the interception store. Model parameters can be identified by taking each of these parts in turn and using a mixture of manual and automatic optimisation (the TSCAL environment was used in the research project); in practice, this procedure is normally done iteratively.

The performance of the MCRM model can be very sensitive to the initial soil moisture deficit; it is therefore vital to have a good understanding of the antecedent conditions before calibrating the model. This might involve including a 'warm-up' period before the event begins. Operationally, initial conditions are based on previous runs and adjusted weekly using MORECS data (at the time of the research study).

TCM

Calibration of the TCM (Thames Catchment Model) is relatively difficult and time-consuming, requiring an orderly approach starting from a physically based structure and parameter set and proceeding via judicious optimisation of selected parameters. Automatic methods of parameter estimation are not very useful, except as a last-stage refinement.

The structure of the TCM is based on subdivision of a basin into different response zones representing differing types of land use, soil, geology and topology, for example representing runoff from gravel, clay, aquifer and riparian areas. Identification of these zones can be achieved using the IHDTM (Integrated

Hydrological Digital Terrain Model) in conjunction with spatial datasets on urban area, 100-year flood extent and WRAP (Winter Rain Acceptance Potential). The zonal responses should be sufficiently different both to avoid excessive parameter interaction and also because each zone should have a hydrological justification. This process produces proportions of the catchment covered by differing hydrological response zones. These proportions are multiplied by the area of the catchment to give an initial value for the size of each response zone. Final values for the area of each zone may differ from these initial values, as zone size is a parameter in the TCM and can be adjusted to give optimum model performance. In addition, the zone size can act as a multiplicative rainfall factor, adjusting for the representativeness of the raingauges used. Therefore the total area of the zones may differ from the size of the catchment after calibration, and also their relative sizes compared to each other may change. For some catchments, it may suffice to think simply in terms of a 'slow response' zone and a 'fast response' zone, analogous to the slow and fast response paths of the PDM.

The TCM requires a large number of parameters, but most of these can be left at their default, physically based values unless absolutely necessary. Specifically, for each zone, parameters such as γ, Φ, R_c, q_c and a can often be fixed at standard values for a particular zone type. The main parameters to optimise for each zone are the time constants k and K and the zone area A. Depending on the type of zone they represent, there are recommended starting values for k and K; the starting value for A can be found from the IHDTM used in conjunction with the spatial digital datasets. Optimisation should start with the baseflow zone first and then subsequent zones with faster response times, although often this is an iterative process. The channel flow routing component of the TCM provides delay and attenuation of the combined outflow from the zonal components when running at a sub-daily time-step. However, much of this behaviour can be represented through adjustment of the pure time delay parameter and the time constants of the zonal storages. Therefore, the number of reaches is commonly set to zero. It is possible that where a satisfactory calibration cannot be obtained, experimentation with different numbers of reaches can be carried out, bearing in mind that each reach introduces a delay equal to the model time-step. Consequently, here the final part of calibration after finding the individual zone parameters (particularly the storage time constants k and K and area A) is to estimate the time delay, τ_d. This can often be done using the automatic Simplex optimisation.

NAM

The NAM model employs upper and lower soil zone and groundwater storages and parallel routing stores in its formulation. It has 16 parameters. This model was not included in the Phase 2 assessment so there was no experience gained on ease of use and how this relates to model structure and parameter uncertainty.

PRTF

The PRTF (Physically Realisable Transfer Function) is implemented as an event-based model within the NFFS, and has been used in South West and North West regions for many years within the operational forecasting systems for those regions (Han 1991; Yang and Han 2006).

A standard transfer function (TF) model expresses the output as a weighted combination of r past outputs and s current and past lagged inputs, with the weights termed the r autoregressive and s moving average parameters. In a rainfall–runoff context the output may be river flow and the input rainfall, for example. The PRTF is a special form of TF model that aims to ensure that its impulse response function has a physically realistic form: that is it should be positive and not exhibit oscillatory behaviour (it should be stable). The PRTF approach reparameterises the r autoregressive parameters (typically r is equal to 2 or 3) so they are defined by a single parameter. Moore and Bell (2001) observe that this can be referred to as an ‘equal root’ parameterisation and gives a stable impulse response function for positive values of the parameter. It should be noted that the ‘pure’ TF formulation (without an ARMA error model) used in the PRTF means that observations of flow (for present and past times) are used in forecast construction: a form of ‘full state correction’ that is inherent to the model structure.

In the original formulation, when using total rainfall as an input, the PRTF model applied directly to runoff and rainfall data does not take account of the actual nonlinear nature of catchment flood response to storm rainfall, notably the dependence on antecedent catchment wetness (Han 1991). In response to this adjustment factors were introduced to alter the volume, shape and time response. The volume adjustment scales the moving average parameters while the shape adjustment shifts the peak and is realised through analytical expressions for adjusting the (equal root) moving average and moving autoregressive parameters. A time shift simply changes the pure time delay between rainfall and runoff. These three adjustments are made by operators of the flood forecasting system as a flood event is identified and develops in real time. The procedure involves judgement of catchment conditions and of forecasts made as the flood develops and as new flow observations are received via telemetry. This form of procedure does not fit well within a framework that aims to largely automate forecast construction for all times, not just during flood events, and which in the future will aim to provide uncertainty estimates for such forecasts.

Some other approaches to accommodate the effect of varying antecedent conditions in PRTF models include:

- Use of different parameter sets for different storm characteristics and catchment wetness conditions: however, this can bring difficulties of operational implementation.
- Recursive parameter estimation, often concentrating on the model gain (here, the runoff coefficient) to track the parameter variation.
- Transformations of the rainfall input to an ‘effective rainfall’ that accounts for antecedent wetness conditions and its nonlinear influence on flood response using a Catchment Wetness Index (CWI) approach. Observed flow itself may also be used as a surrogate index of catchment wetness.

The consultation exercise indicated an interest in catchment state and its uncertainty, particularly for regions using PRTF models with ‘effective rainfall’ as input. One option in this case would be to also explore rainfall–runoff models of the conceptual, continuous water accounting type that incorporate catchment state explicitly through their model formulation. This would be a useful topic for investigation but is outside the scope of the present project.

3.3.2 Flow routing and hydrodynamic models

The types of flow routing and hydrodynamic models which are currently used in NFFS are:

Flow Routing

- KW
- DODO
- VPMC

Hydrodynamic

- ISIS
- MIKE11

In general, for all types of model, uncertainties can arise in gauged or ungauged inflows, and in the accuracy of level and/or flow data used in the model calibration, and in real-time operation.

For flow routing models, issues can arise with specification of the wave-speed–flow relationship (if required), and the general conceptualisation of the river reach (number of sections, tributary inflows etc).

For hydrodynamic models, there are additional uncertainties in key parameters such as the hydraulic roughness of open channel sections or culverts, and afflux associated with structures such as bridges, weirs or sluices. For out-of-bank situations this also applies to discharge coefficients and the stage-area relationships for storage areas.

Hydraulic roughness can play an important role in flood forecasting as seasonal variations can be considerable within unmaintained channels. Generally the range of appropriate roughness values affecting forecasts should be possible to estimate given local knowledge of a particular channel even if the precise value at a given time may be difficult to establish. Information to aid this process may be obtained from ultrasonic or ADCP (Acoustic Doppler Current Profiler) gauges in open channels.

Initial conditions are also a source of uncertainty for hydrodynamic models. This is particularly important for storage areas represented as reservoir units which have no means of representing the evapotranspiration process unless by associating a rainfall boundary which is not generally done. Another issue is that gravity drainage of storage areas via flapped outfalls for example is often neglected in design models as it has limited impact on peak levels.

It should be noted that, as for rating curves and floodplain storage, these issues can be addressed by further modelling effort rather than necessarily implementing probabilistic methods as the sole solution.

Section 4.3.3 provides background on approaches to model updating in ISIS and MIKE11, while the following subsections describe the DODO and KW flow routing approaches (with similar issues to consider for the VPMC approach).

DODO

The DODO model is based on a form of Muskingum storage function which relates the volume of water stored in a river reach at a given time to the reach inflow and outflow. The reach inflow is lagged in time with the lag decreasing as a power function of the reach inflow, but limited to a minimum lag value. The component of reach inflow above the bankfull discharge is routed through a parallel, second Muskingum storage, after accounting for an initial contribution to static floodplain storage. On the recession, water in static storage drains out of the reach, initially slowly, but then freely below a critical return bankfull storage as a power function of the volume of water in static floodplain storage. Lateral inflows to the reach are divided equally between the reach inflow and the reach outflow; a downstream input can also be added to the routed outflow to give the final reach outflow. The DODO model has a total of twelve parameters, six representing in-bank routing and six representing out-of-bank routing. It was originally developed for use within the Severn-Trent Flood Forecasting System (ST-FFS) (Douglas and Dobson 1987) and now is used in the Midlands Region forecasting system for hydrological flow routing. A detailed documentation of the model was undertaken by the Institute of Hydrology (now CEH Wallingford) and reported in Wallingford Water (1994).

KW

The KW model is a generalised form of kinematic wave model which makes allowance for wave-speeds to vary with discharge magnitude. In addition, storage functions are provided to represent flow into washlands to complement the modelling of in-bank flows. The basic form of the model is presented in Moore and Jones (1978) and Jones and Moore (1980). Water movement down a river channel is approximated by the kinematic wave equation with lateral inflow. This equation is expressed in finite difference form such that attenuation of the flood wave is controlled by the space discretisation (number of sub-reaches) and wave-speed (dimensionless) for a given time-step. A time-varying wave-speed is allowed, changing as a function of the observed flow. The choice of functions available include a piecewise linear function over three or four segments as well as cubic and exponential parametric functions. An auxiliary threshold storage function can be applied, either at selected model nodes to represent overflow into the floodplain, or to observed lateral inflows to compensate for errors in the rating relationship, especially for out-of-bank flows. A number of forms of parameterised threshold function are available. The use of a variety of parametric functions to define the model form is particularly useful for real-time application to large, complex river basins where the use of survey data might be expensive in time or survey data may not be available. Calibration of the parametric model functions is accomplished using the Model Calibration Facilities of the TSCAL (Time Series CALibration) environment. For a more extensive description of the KW model the reader is referred to the KW user guide (CEH Wallingford 2005c) and Moore (1999).

3.3.3 NFFS-specific considerations

Many of the probabilistic techniques described in this report require multiple model runs at each time-interval (although the number required can be modest, in many cases). The following sections describe some options for performing runs of this type in NFFS, and for either speeding up model run times (with a focus on hydrodynamic models), or for emulating the performance of models.

Scenario runs

The NFFS system is designed to run in a fully automated manner. For this reason most of the parameterisation of models is contained within a fixed file structure which is used for every model run. It has long been the case that one 'optimum' solution to model calibration is sought as the best representation of the physical system.

However, within NFFS there is the possibility to use alternative parameter sets within what-if scenarios. These parameter sets can be predefined in a local directory in PI XML format (ModuleParFiles). The ability to handle varying parameters per run (as opposed to just time series of inputs) is handled by the model adapter. The adapter interprets the PI XML parameters and must incorporate these values into the required native format prior to the run. Alternatively the native format parameter or model file can be predefined and stored in a local directory (ModuleDataSetFiles) for use in the model run as part of a what-if scenario. In this case NFFS has no way to interact with the exported parameters.

Currently, only the MCRM model adapter in NFFS is designed to incorporate PI XML parameter adjustments. However, it should be noted that ISIS data comprises a single modular text file which could be converted readily to XML format if it was regarded as important to investigate parameter updating or uncertainty as part of the current project.

Event-based parameter sets and the ability to modify parameters in real time are also planned for use for FEWS applications in the USA and Australia (available summer 2009). However, one general point to note is that, for deterministic models, such as the conceptual rainfall–runoff models employed in NFFS, the need for multiple parameter sets may be the consequence of a shortcoming in the model structure or of its calibration, where improvement should be first sought. Similarly parameter adjustment (parameter updating) is not generally advised as a form of data assimilation. However, this functionality has an important role to play in uncertainty estimation and this point is discussed later.

Model configuration changes

Model configuration changes, particularly to hydrodynamic models, can give rise to enormous benefits in terms of run times and Table 3.4 presents some startling statistics for a range of models of different levels of complexity (cross-sections, structures, floodplain representation etc).

Table 3.4 Examples of run time improvements for ISIS v3.1 models.

Model	Run duration (hours)	Run time for real-time model (seconds)	Equivalent run time for design model (seconds)	Run time improvement ratio
Oxford Thames	312	24	3600	150
Thames	312	12	1500	125
Lower Colne	312	71	1800	25
Lower Thames	312	25	3600	144
Ravensbourne	39	79	3600	46
Lower Lee	78	35	24000	686

It should be noted that these improvements to run time were not achieved by converting sections to hydrological routing models or by wholesale simplification of the models. As a measure of this, the converted models could still be used in principle for real-time inundation mapping if required.

It must be noted that to achieve run-time improvements of this magnitude requires not inconsiderable skill and experience from the modeller and great care is also required in the subsequent configuration process.

It may be interesting to establish what the minimum run time may be to enable some direct use of hydrodynamic models in a probabilistic context rather than by emulation alone.

In terms of minimising run times, one option may be to minimise run durations for the forecasting runs. This could be achieved by running the models only as far as the target lead time (or just beyond) from 'time now'. Even if this approach is adopted, there is always a need to run the model from the time of the last 'state run'. State runs are re-performed once a day in almost all cases but to reduce run durations it may be possible to increase the frequency of these as appropriate to the model and target lead time.

It is worth noting that in some situations flow routing models can sometimes perform as well as (and run faster than) hydrodynamic models, although the focus of this study is on probabilistic approaches rather than comparing alternative types of deterministic model (and there is a considerable body of research on the strengths and limitations of each approach, e.g. see the Real-Time Modelling Guidelines). Also, the aim in this study is to explore a representative range of all model types, including situations where hydrodynamic or other models/emulators are clearly more appropriate (e.g. where there are backwater or tidal influences).

Model emulators

Emulators provide another option for improving model run times: the DBM models (Young 2002; Taylor *et al.* 2007) can be used for the dynamic emulation of computationally intensive models, such as distributed hydraulic or hydrological models (Young and Ratto 2008). Here, for a specified set of the high order model parameter values, a 'nominal' DBM emulation model is identified and estimated on the basis of data generated by the high order model. This nominal emulation is normally able to explain over 99% of the flow or level variations of the high order model. If this exercise is repeated over a whole range of the high order model parameters, then it is possible to map the relationship between these parameters and the parameters of the DBM model, as shown schematically in Figure 3.4.

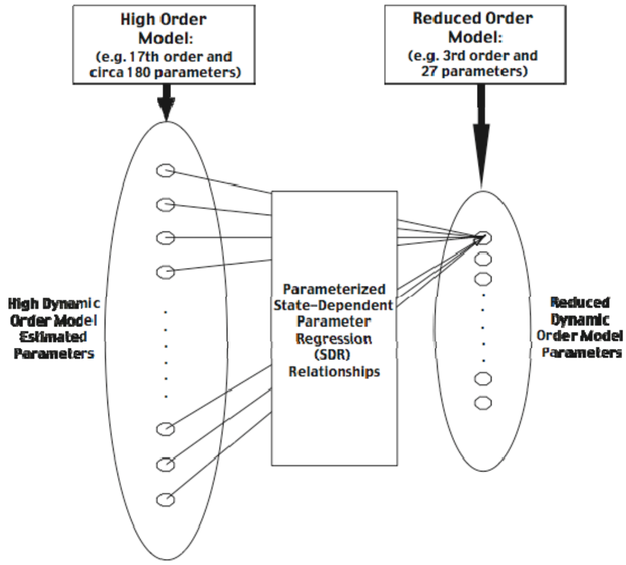


Figure 3.4 The process of dynamic emulation model synthesis.

In this form, the DBM emulator can be used to replace the high order model for functions such as sensitivity analysis, uncertainty estimation or forecasting. This approach to emulation is still the subject of research and development but it holds great promise. For example, Figure 3.5 shows recent results (Beven *et al.* 2008b; Young *et al.* 2009) from a DBM emulation of the HEC-RAS hydrodynamic routing model: here, the estimated nominal emulation model is validated against new data generated by the HEC-RAS model with a new set of upstream inputs.

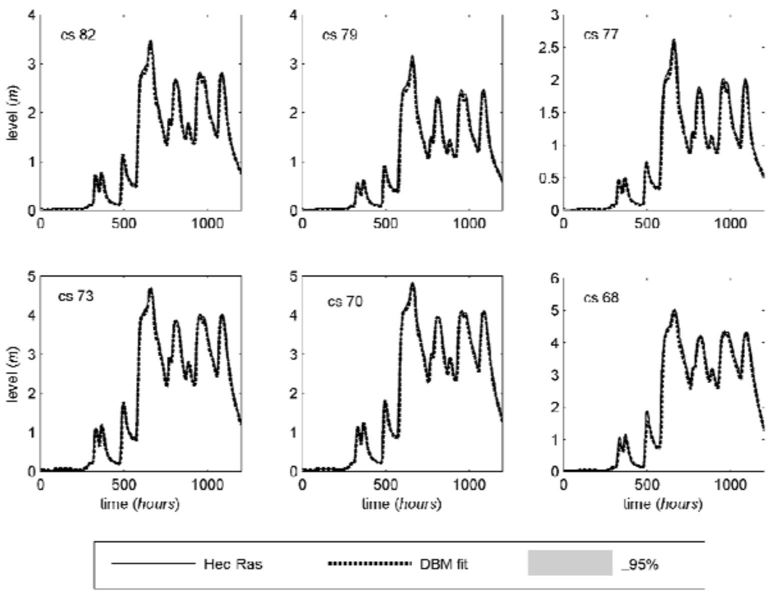


Figure 3.5 Validation of a DBM emulator for the HEC-RAS model at six downstream sites based on a new set of upstream level inputs.

3.4 Summary of chapter

Some key points from this chapter include:

- A detailed summary is provided of the key uncertainties in rainfall–runoff, flow routing and hydrodynamic models. This review is based both on international research, and the experience of the project team.
- The main sources of uncertainty which are discussed include uncertainties in model calibration, estimation of ungauged inflows, rating curves, antecedent conditions, the influence of structures and floodplain storage, and from rainfall forecasts and actuals.
- A similar review has also been performed for the main types of rainfall–runoff, flow routing and hydrodynamic models used within the Environment Agency for flood forecasting, including the PDM, MCRM, NAM, DODO, KW and ISIS models. However, the focus of this review has been on the types of models selected for consideration in the case studies on this project.
- Other topics which are discussed include the existing functionality in NFFS for performing ensemble and scenario runs, and the potential for using model reconfiguration and model emulators to improve model run times when multiple model runs (ensembles) are required at each forecasting time-step.

4. Uncertainty framework

4.1 Introduction

This section describes the review of techniques under Tasks 1.3 and 1.4 of the project, and the high-level version of the uncertainty framework which will form the basis for the more detailed framework to be developed during Phase 2 of the project. The aim of Task 1.3 was:

> *To recommend and test suitable techniques for the probabilistic treatment of the most important sources of uncertainty and combine them into a high-level unified, scalable framework for integrated catchment models.*

while the aim of Task 1.4 was:

> *To investigate the requirements, possibilities and benefits of real-time/state updating of probabilistic hydraulic/hydrological models and the value of different types of data (historical and real-time) in constraining uncertainties.*

In Work Package 9 of the first phase of the Flood Risk Management Research Consortium (FRMRC), the following general classification scheme for techniques for estimating uncertainty was proposed:

- Forward uncertainty propagation methods – in which the uncertainty is assumed to be known or specified in advance, and is used to determine the likely range or distribution of model outputs. Methods include Monte Carlo or ensemble analyses, various statistical techniques (e.g. expectation analyses, analyses of historical data), and fuzzy set methods.
- Conditioning approaches – in which the model outputs are conditioned on current and historical observations, including Recursive Instrumental Variable methods, Bayesian Analytical Methods, Bayesian MCMC Methods, and the GLUE methodology (sometimes known as pre- and post-processing approaches).
- Real-time data assimilation methods – such as the Kalman Filter (and extended and ensemble versions) and particle filters.

Section 4.2 presents a review of forward uncertainty propagation techniques), while Section 4.3 reviews data assimilation and pre- and post-processing approaches. Section 4.4 then considers a range of generic uncertainty estimation tools. Finally, Section 4.5 discusses development of the high-level uncertainty framework.

4.2 Forward uncertainty propagation techniques

The following forward uncertainty propagation techniques are discussed here:

- Multi-model techniques
- Monte Carlo methods
- Fuzzy set methods.

Ensemble approaches are also another example of this technique, and are most usually associated with rainfall forecasts. They are discussed in detail in Section 3.2.1.

In simple situations, analytical approaches to the forward uncertainty propagation problem can sometimes be taken under very strong assumptions that the system is linear, and (most often) that the sources of error can be characterised as normal distributions with known constant variance (e.g. Beven 2009). For example, analytical propagation of uncertainty, with these strong assumptions, is often used for time-stepping of uncertainties out to the required lead time in adaptive real-time forecasting (see later).

Multi-model techniques

Model structural uncertainty is often neglected in hydrological forecasting systems. An approach to deal with this uncertainty is to use multi-model ensemble techniques, which are also used within the atmospheric forecasting community. For chaotic systems like the atmosphere, considering these model uncertainties may be more important than in systems that are not chaotic. Models with different model structure are used to get a handle on the model structural uncertainty. These model results can be combined into one forecast with uncertainty bounds using (Bayesian) model averaging (BMA) techniques (see later).

For example, in Italy, ARPA-SIM have developed a flood forecasting system for the Po river (FFS-PO) composed of three modelling chains that simulate the entire Po basin and river system behaviour starting with observed data and forecast meteorological data. The models that should be implemented in the full version of FFS-PO are:

- Hydrological models: NAM, HEC-HMS, TOPKAPI
- Hydrodynamic models: MIKE11, HEC-RAS, PAB.

The modelling system is shown in Figure 4.1.

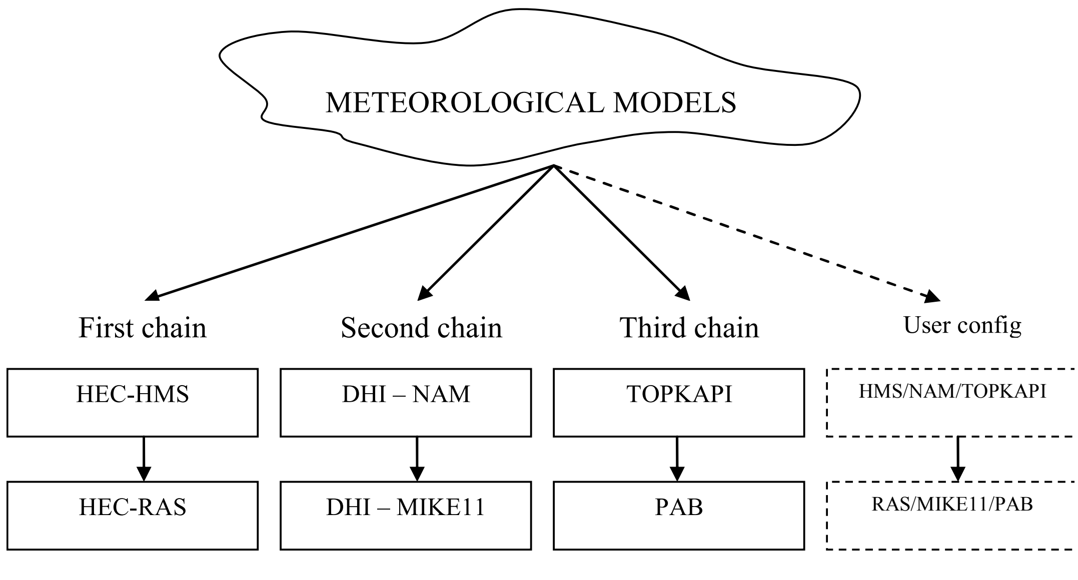


Figure 4.1 Illustration of the modelling system for FFS-PO.

According to the ARPA-SIM specification these three model chains will work in parallel performing forecasts on an hourly basis, so that the overall system will include the following model combinations:

- 11 meteorological models (LAMI + ESEMBLE)
- 3 hydrological models (NAM, HEC-HMS, TOPKAPI)
- 3 hydrodynamic models (MIKE11, HEC-RAS, PAB).

The full system will be configured for 33 model chains + 1 user. An evaluation of forecasts made or the experience gained with this system is not yet available.

A similar approach was also adopted on the MAP D-PHASE project, which is described in Section 2, with multiple rainfall–runoff models running in parallel, with outputs compared to a common set of thresholds at a catchment basis to determine the likelihood of flooding.

Monte Carlo approaches

Monte Carlo simulation is a powerful methodology for the forward propagation of uncertainty through nonlinear systems, particularly when the dimensions of the sources of uncertainty are small. As in any forward uncertainty estimation methodology, it is first necessary to define the character of each source of uncertainty (often in the form of prior assumptions about the distributions of inputs or parameters) and any correlations or interactions between sources of uncertainty.

Given those prior distributions, random samples are taken and used to run the model. Care must be taken with selecting a random number generator, particularly where large numbers of samples are to be taken. For forward uncertainty propagation, the random samples are usually generated using a form of importance sampling, such that samples are chosen in a way consistent with the prior distributions, so as to be of equal weight in representing the posterior output distribution. This can be achieved by sampling uniformly on the probability scale for each source of uncertainty. There are also more efficient, but more approximate sampling techniques, such as the Latin hypercube (Beven 2009).

Correlation between sources of uncertainty is also important. While it is often the case that we know little about whether different sources of uncertainty interact, ignoring correlation can lead to overestimation of the real uncertainty. A modern method of representing multivariate interactions that can be used with arbitrary distributions is the use of copula sampling (Beven 2009).

Fuzzy set methods

As noted earlier, not all uncertainties in the modelling of environmental systems can be treated properly using probabilistic methods, although it is a common strategy to use probabilistic methods as an approximation, even in cases where errors are not easily represented using a simple stochastic error model. It is more acceptable to do so in the case of forecasting in real time when the deficiencies of the error model can be compensated by updating during data assimilation. In simulation, however, the recognition that not all uncertainties are probabilistic and that some error may result from epistemic errors has led to a variety of alternative approaches to error propagation (Beven 2002, 2005).

One of these approaches is the use of fuzzy sets, originally due to Zadeh (1965), where uncertainty of a variable is represented as a membership function (normally in the range 0–1) to a fuzzy set. Such sets can then be combined using a choice of methods, such as set union and set intersection, giving more flexibility than within a formal statistical framework. More detail on fuzzy set methods can be found in Klir and Folger (1988), Klir (2006) and Beven (2009).

Fuzzy methods have also been incorporated into the Generalised Likelihood Uncertainty Estimation (GLUE, see Section 4.4 below) methodology as a way of calibrating models subject to non-probabilistic uncertainties. (see, for example, Beven 2006, 2009; Page *et al.* 2007; Li *et al.* 2009), including flood routing models (Pappenberger *et al.* 2007a, 2007b).

4.3 Data assimilation and conditioning techniques

Data assimilation techniques can help to both constrain and quantify uncertainty (although for the techniques used operationally at present in the Environment Agency only the former approach is used). Conditioning techniques aim to better express the residual uncertainty in probabilistic terms, based on recent or historical model performance, and can be used with or without a data assimilation procedure. Hence, for data assimilation, the focus is often on reducing uncertainty (although uncertainty estimates can also be provided with some techniques) while, for conditioning, the focus is on quantifying uncertainty.

Note that, for some conditioning techniques, the model performance must be evaluated over long periods which requires access to a database of historical model forecasts for the models and data inputs *in their current state*, or to regenerate those values. This requirement is similar to that for deterministic models when considering models driven primarily by raingauge and/or upstream river flow values; indeed, it is standard practice in many model calibration studies, including sometimes improving the historical data if necessary to account for changes to rating curves, instrument locations etc (or restricting doubtful periods from the analysis).

For radar rainfall data, the data held by the Environment Agency date back to about the year 2000, placing some limits on the periods which can be considered. For rainfall forecasts, conditioning is more problematic at present, although the situation will improve over time. This is because at present there is no long-term hindcast or reanalysis dataset available for the current forecast products. However, the situation is perhaps analogous to that when the Environment Agency started to archive radar rainfall data; over time an archive is developed which can be used in model development (again taking account of any significant changes to nowcasting or NWP models).

This section reviews a range of data assimilation and conditioning techniques which are available internationally (Sections 4.3.1 and 4.3.2), and then Section 4.3.3 discusses examples of techniques which are currently used in NFFS. Finally, Section 4.3.4 discusses some initial considerations on the use of data assimilation techniques in integrated catchment models (which will be considered further in Phase 2 of the project).

4.3.1 Data assimilation techniques

Introduction

Flood forecasting systems are typically constructed from a range of models that are applied to the prediction of discharges and levels in the river system, as a function of observed meteorological conditions and possibly short- or medium-term rainfall forecasts. These models are employed in a real-time environment, where models take information on the current and past states of the system, and forecasts are made for a certain period of time into the future as a function of boundary inputs on the system (Refsgaard 1997).

Data assimilation or updating is a feedback system where the modelling process is conditioned using the information on the current state of the system. These process models can be considered as a set of equations containing parameters and state variables (Refsgaard 1997), where state variables are transient in time, and the parameters are generally (but not always) held constant at some value determined in the calibration of the model prior to application in the real-time environment.

Data assimilation and updating procedures are often categorised into four different approaches (Refsgaard 1997; WMO 1992). Figure 4.2 shows schematically where these different approaches interact with the model. In all cases the updating procedure is applied as a consequence of the comparison of model outputs and observed values. The evolution of the model outputs is a consequence of the input variables.

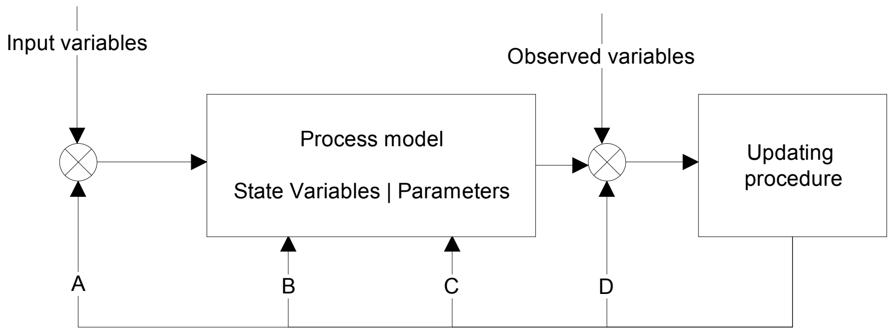


Figure 4.2 Schematic diagram of different updating procedure approaches (from Refsgaard 1997).

The four approaches to updating described in the figure are as follows:

Option A. Updating of input variables

The input variables are typically considered the dominant source of error. Precipitation is often the primary input variable for rainfall–runoff processes, and errors in the estimation of the precipitation can be significant. The input variables are therefore adjusted based on the residuals of the comparison of model outputs and measured variables. Automatic procedures for this approach are scarce, and when applied it is typically carried out manually by experienced modellers. There is also a problem of disaggregation, in that it is often not clear where in the time series of the input variables the error correction is to be applied. The problem of equifinality may also arise, in that error correction in several different areas of the input variables may lead to the desired result. One example of input updating (PT updating) applied in operational systems is the input correction procedure available in Sweden within the HBV-96 model of the Swedish Meteorological and Hydrological Institute (SMHI) based on optimisation of precipitation and temperature over a certain time-window.

Option B. Updating of model state variables

Based on the observed residuals, the model state variables are adjusted. A number of approaches can be followed ranging from simple to complex statistical filters. Simple methods include the correction of the capacity of conceptual reservoirs in a conceptual hydrological model according to the model bias after a complete model run. Another methodology is the Kalman Filter that is well proven for linear systems, and the Extended Kalman Filter for nonlinear systems. Kalman filters adjust the state variables in a physically consistent way (encapsulated in the model transfer functions) using statistical assumptions about the spatial and temporal correlation of model errors. Kalman filters are typically applied in-line with the model, where the model filter step is carried out at the time-step at which the observed data are available. Simple methods are typically off-line procedures, where the state variables are adjusted after the model run.

Option C. Updating of model parameters

Updating of the model parameters according to the observed residuals is in many ways similar to the updating of model states. Kalman filters can be theoretically applied for this purpose, but examples of actual application are scarce. Model parameters such as roughness coefficients and runoff coefficients can be considered for adjustment. For the more physically based approaches to flood forecasting, the procedure could be seen as a constant recalibration of the model, and is therefore somewhat questionable, except perhaps in the case of fully parameterised models. As Kachroo (1992) notes 'it is intrinsically difficult to accept the operation of any hydrological system can change significantly over such a short interval as the observation time'. However, parameter sampling has a role in the estimation of uncertainty in certain types of models.

Option D. Updating of model outputs, or error correction

Error correction (also known as output correction) is one of the oldest and most versatile methods of data assimilation (Madsen *et al.* 2000; Broersen and Weerts 2005). Rather than adjustments of the state variables of the model as done in Option B (and effectively in Option A) the errors are corrected there where they are observed. Often the correlation of model residuals is strong both temporally and spatially, and in this option a model of this residual is developed. A typical example is of ARMA(X) models that are calibrated using the observed residuals. The error model is applied to adjust the model output, and can be equally applied to both the model outputs over the period where observed values are available (update period) as the forecast period. In the latter part, where the adjustment of the model output is propagated into the forecast period, the adjustment will slowly dissipate (depending on the error model structure). The following sections provide more detailed background on some of these techniques.

Kalman and particle filtering

Sequential data assimilation techniques provide a general framework for explicitly taking into account input uncertainty, model uncertainty and output uncertainty. One of the best known sequential data assimilation techniques is the Kalman Filter (KF) (Kalman 1960), which was developed for linear systems. It was later extended to nonlinear models, resulting in the Extended Kalman Filter (EKF). Both became popular in hydrological modelling in the late 1970s and early 1980s (Kitanidis and Bras 1980; Georgakakos 1986a, 1986b). However, the linearisation in the EKF is notoriously inaccurate if the nonlinearities are strong. A possible way to circumvent these problems is by letting the errors evolve with the nonlinear model equations by performing an ensemble of model runs. This has led to the development of the well-known Ensemble Kalman Filter (EnKF) (Evensen 1994; Burgers *et al.* 1998).

In a separate line of research, filter methods for non-Gaussian nonlinear dynamical models have been developed. These sequential Monte Carlo methods, such as sequential importance resampling (SIR) and residual resampling (RR), also known as particle filtering, originate from the research area of object recognition, target tracking, financial analysis, and robotics. So far they have received little to no recognition for the application to hydrological models (Moradkhani *et al.* 2005a; Weerts and El Serafy 2006).

The central idea of EnKF and particle filters is to represent the state probability density function (pdf) as a set of random samples. The difference between the EnKF and particle filters lies in the way of recursively generating an approximation to the state pdf.

When applying EnKF, in the analysis step, only the first two moments of the prior pdf of the model states are used to obtain the posterior pdf of the model states. Applying it to highly nonlinear processes (like rainfall–runoff) makes it unlikely that the prior density is Gaussian distributed, and therefore it is likely that the posterior is determined not only by the mean and variance of the prior density, but also by the whole density.

The advantage of particle filters is that no assumptions on the form of the prior pdf of the model states are necessary and that the full prior density is being used, in contrast to EnKF. In theory this would mean that particle filtering is more sensitive to the tails of the prior distribution, a property which may be of vital importance in flood forecasting, although this may be at the cost of a much larger number of simulations.

For flood forecasting, the potential of EnKF was proven successfully by comparing it with EKF on a calibrated hydrodynamic Rhine application (El Serafy and Mynett 2004) and applying it within FEWS-NL Rhine & Meuse (Weerts 2008); and input uncertainty estimation was recently analysed through the use of EnKF and particle filters for the HBV model (Weerts and El Serafy 2006). EnKF is now used operationally in the FEWS-NL Rhine system to update the states of the hydrodynamic model SOBEK-RE of the Rhine.

EnKF is in general less sensitive than particle filters to mis-specification of the model and of the input uncertainties. Application of both of these data assimilation techniques together in flood forecasting systems is feasible, although the computational burden might still be an obstacle. It is clear that a trade-off between accuracy and computational burden exists. The results of applying the filters to real data showed that the particle filters are more sensitive to the choice of the model error and measurement error (Weerts and El Serafy 2006). This makes EnKF more robust and therefore one may prefer EnKF in an operational flood forecasting setting. For research purposes, where the computational burden is less important and accuracy more important, one may prefer particle filters (Weerts and El Serafy 2006).

Variational methods

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. More detail on the method can be found in Beven (2009).

Data Based Mechanistic methods

The Data Based Mechanistic (DBM) approach provides a range of techniques both for forecasting river levels and flows, and for assessing and updating the uncertainty in the estimates within a data assimilation framework. As noted earlier, the Lancaster University DBM approach to probabilistic flood forecasting was first implemented in providing a forecasting system with data assimilation and probabilistic uncertainty estimation for Dumfries and the Nith catchment in 1991 (Lees *et al.* 1994; Beven 2001). However, this section focuses on the uncertainty and data assimilation aspects of the approach.

Developments in the methodology have been described by Young (2002, 2009) and Romanowicz *et al.* (2006a,b, 2008). The DBM approach uses various computational routines available in the CAPTAIN Toolbox for Matlab (see Taylor *et al.* 2007 and http://www.es.lancs.ac.uk/cres/captain/), or their equivalent, to identify and estimate models directly from rainfall and flow/level data. The resulting model has a physical interpretation in terms of interconnected hydrological stores in series and parallel and, as such, can be compared with similar conceptual models, such as IHACRES, PDM and HYMOD (Figure 4.3). The advantage of the DBM model is that its structure, including the nonlinear input transform (Figure 4.4), and dynamic order are identified statistically from the data and so it is not prone to over-parameterisation. In addition, it is a stochastic model and so the uncertainty associated with its parameters and additive noise components are quantified during the estimation (calibration) phase of modelling. These are then available when the model is incorporated in an appropriate forecasting system.

The simplest and most obvious stochastic forecasting environment for DBM models used in the forecasting environment is a version of the Kalman Filter (Kalman 1960). This incorporates the input nonlinearity and can include advanced options, such as adaptive gain and heteroscedasticity elements that exploit recursive parameter estimation (see Young 2002, 2009; Romanowicz *et al.* 2006; Young *et al.* 2006). The primary advantages, however, are the *inherent* real-time state updating and stochastic error correction (equivalent to ARMA error correction), together with standard error estimates on the forecasts that are a function of the flow or level state-dependent uncertainty estimation, as illustrated in Figures 4.3 to 4.5. These show the results of applying the DBM methodology to adaptive forecasting of the 2005 Carlisle flood using a cascade of only two rainfall-level and one level-to-level components, with identified nonlinearities as shown in Figure 4.4 (but without an error correction model in this case).

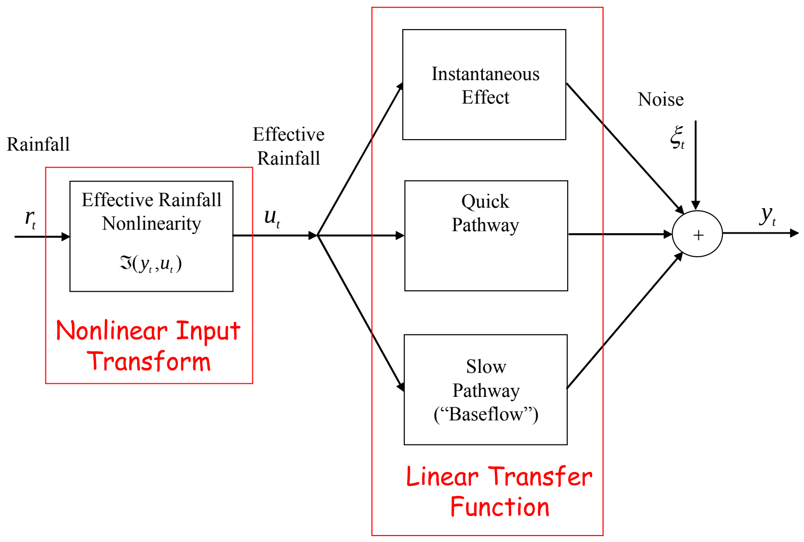


Figure 4.3 Structure of a DBM model.

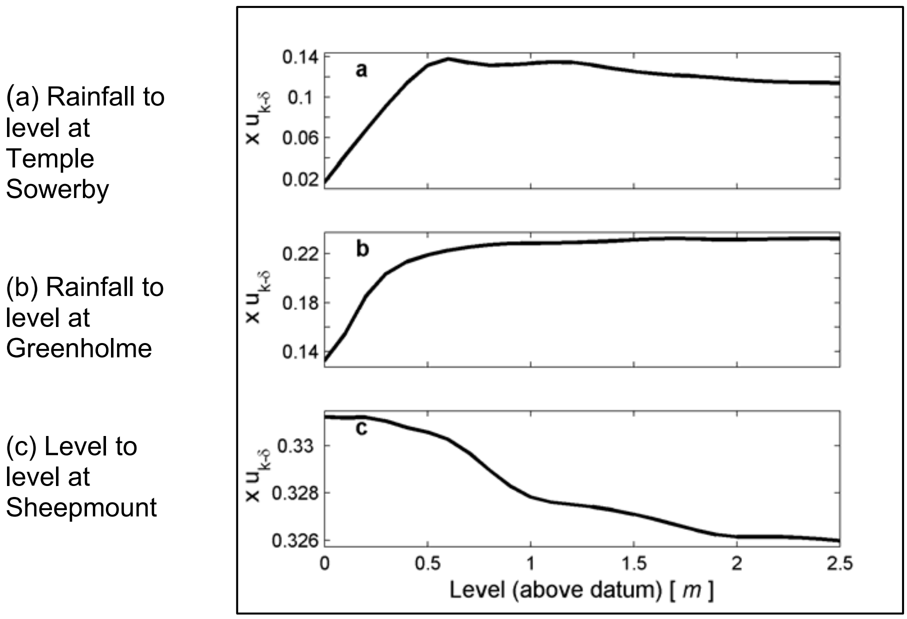


Figure 4.4 Identified nonlinearities in the DBM model for rainfall to level and level-to-level models on an initial application to the River Eden (after Leedal *et al.* 2008).

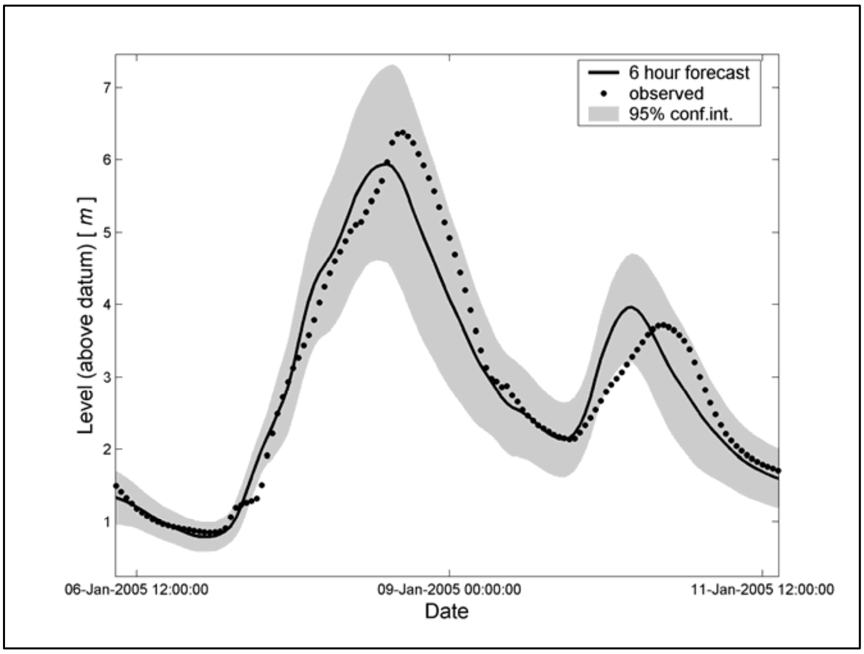


Figure 4.5 Adaptive 6-hour ahead forecasts at Sheepmount, Carlisle during the January 2005 flood event in an initial application of the DBM methodology to the River Eden (after Leedal *et al.* 2008).

4.3.2 Conditioning approaches

Quantile regression

Regression is a widely used statistical technique for investigating the relationship between variables.

Quantile regression is a related technique for estimating functional relationships between variables for all portions of a probability distribution (Koenker and Bassett 1978). Based on the historical performance of a model, the method provides adjustment factors which can be applied to a real-time forecast to improve estimates of uncertainty.

A similar technique is percentile matching (Hashino *et al.* 2007) in which the percentile of a simulated flow $F_s(Q)$ is used to extract a corresponding flow from the inverse of the Cumulative Density Function (CDF) of the observations F_0^{-1}:

$$Q_{bcs} = F_o^{-1}(F_s(Q))$$

where F_s is the cumulative distribution function of the simulated flow Q.

This correction adjusts both the mean and variance (and higher moments) of the simulation outputs to match those of the observed climatology. The simulated and observed CDFs are ideally based on the same period. These techniques have been used in pre-operational testing by the National Weather Service in the USA, for example (Wood and Schaake 2008); see Section 2.1.2.

Bayesian Model Averaging

In a situation of several competing forecast models, a multi-model approach can be adopted to use information from every model, instead of choosing a single optimal model. Even if one model is known to perform generally better than the others, it is not certain which model is optimal for a particular forecast situation. The choice for a forecast model is therefore uncertain.

The Bayesian Model Averaging (BMA) method (Raftery *et al.* 2003, 2005) is an example of a multi-model technique. An overall probabilistic forecast is generated from a linear combination of normally distributed predictive probability density functions (pdfs), based on forecasts from the individual models in the ensemble. The weight of each pdf is the posterior probability that the corresponding model is correct. The individual distributions are bias-corrected conditional pdfs, given that the corresponding model gives the best forecast. The total uncertainty is thus split into two components, corresponding to the model choice and the uncertainty given the optimal choice. The bias-correction is essential for good performance in practical applications.

The Bayesian weights and the sigmas of the individual pdfs are calculated from the performance of each model over a training period, using recent historical observations and model forecasts. For example, Raftery *et al.* (2003) vary the length of the training period to optimise the calibration of the overall probabilistic forecast. For hydrological applications, the optimal training period is at least as long as the time of concentration of a catchment, thereby including a number of hydrological runoff events (Beckers *et al.* 2008).

Observations and forecasts from multiple sites can be included in the training if the performance of the individual models can be assumed similar for each site. This can improve the estimates for the weights and sigmas of the individual models.

A major advantage of the BMA method in an operational setting is that it is relatively robust. Because of the relatively short training period, the method naturally adopts itself to changing environments, such as missing data, model updates that improve the performance of one of the forecast models and addition of a new model to the ensemble.

The BMA method assumes that the ensemble of forecast models represents the range of all plausible models and that it includes, for each forecast situation, the correct forecast. This is never the case in the real world. Therefore, the probabilistic forecast that results from the BMA technique is only an estimate of the true predictive uncertainty. In case of a small number (say less than five) of competing models, or for an ensemble of very similar models, the BMA method inevitably breaks down. If the ensemble of forecast models is too limited, this can sometimes be overcome by using different combinations of meteorological and hydrological components. This approach seems to be promising for at least one application in the Netherlands (Beckers *et al.* 2008).

The BMA method requires that the likelihood of the model choice can be based on the training period. This is no longer valid in rapidly changing conditions, such as dry to wet transitions. It has been suggested that the training should be based on 'similar conditions' instead of a preceding period. However, this has not been tested yet.

In its original formulation by Raftery *et al.* (2003) the uncertainty of the individual model forecasts is assumed to be normally distributed. Vrugt and Robinson (2007) extended the method for Log-normal and Gamma distributed variables. Their results indicate that this generalisation is to be preferred for skewed predictands, such as river discharges or water levels.

Todini (2008) has proposed some modifications to the BMA method, including a transformation to normal space. This avoids the assumption of normally distributed uncertainties and produces good results for the Po River.

Finally, Vrugt and Robinson (2007) point out that the Expectation-Maximization (EM) algorithm that is used to find the weight and sigma of each pdf (Raftery *et al.* 2003) need not necessarily converge to the optimal values. Vrugt and Robinson (2007) propose some alternative optimisation tools. Todini (2008) uses a Newton–Raphson optimisation method.

Vrugt and Robinson (2007) also compared BMA to EnKF and concluded that the EnKF was superior. However, we believe that this depends very much on the specific application.

An Open Source BMA software package in R is available from:
http://cran.r-project.org/web/packages/BMA/index.html

Bayesian Processor of Output (BPO)

A comprehensive approach to gain a description of the overall uncertainty of a flow forecast available is through the application of Bayesian revision. Krzysztofowicz (1999) proposed the Bayesian Forecasting System (BFS) theory for streamflow predictions. This theory constitutes a general framework for Bayesian inference on the uncertainty of a flow forecast, while using deterministic hydrological and/or hydraulic models. The basic concept of the BFS is to derive the uncertainty of a forecast by 'revising' prior knowledge on the behaviour of the system over a historical period of operation. If performed correctly, the result of the revision process, referred to as posterior density function, represents a reliable assessment of the uncertainty of the forecast, which is conditional on a whole range of information available at the beginning of a forecast. It is referred to as predictive uncertainty (Krzysztofowicz 2001; Todini 2007).

A simple Bayesian uncertainty processor for a flood forecasting system, based on the application of Bayes theorem, can be formulated in terms of the random variables s_n, h_n, h_0 (Krzysztofowicz and Kelly 2000) as follows:

$$\phi_n\left(h_n \mid s_n, h_0\right) = \frac{f_n\left(s_n \mid h_n, h_0\right) g_n\left(h_n \mid h_0\right)}{k_n\left(s_n \mid h_0\right)}$$

where s_n is flow or stage forecasted at lead time n (expressed in number of hours or days), h_n is the flow or stage which has been observed for day n, and h_0 is the flow or stage observed at the onset of the forecast. In this example, the conditioning is limited to the random variables h_n and h_0. However, the formulation can be arbitrarily expanded to include additional conditioning variables, if necessary.

The conditional probability density f_n is a likelihood of actual discharges or stages, given a model forecast and conditioning observation(s). The conditional density g_n is a prior probability density function on the flow predicted for day n, conditional on the observation. The denominator is the expected density on the forecasted flow given by the total probability law:

$$k_n(s_n \mid h_0) = \int_{-\infty}^{\infty} f_n(s_n \mid h_n, h_0)\, g_n(h_n \mid h_0)\, dh_n$$

The prior knowledge on the system, which is stochastically described by the conditional probability density function g_n is based on assumed probability distributions on discharges or stages. A proper specification of the prior density and likelihood function is essential in obtaining an informative posterior. Failing to do so may compromise the performance of the processor. The determination of an adequate prior function constitutes a challenge, especially in the absence of sufficiently long historical time series of observations, as is the case in poorly monitored basins.

Examples of the BPO applied to small and medium catchments (short lead times) have been given by Krzysztofowicz and Kelly (2000) and Verkade (2008). Examples of applying the BPO to the Rhine basin (long times) have been presented by Reggiani and Weerts (2008b) and Reggiani *et al.* (2009).

4.3.3 Examples from the Environment Agency

Examples of approaches to data assimilation within NFFS are provided for the following: FEWS ARMA error correction, CEH ARMA error prediction, PDM state correction, and the ISIS GAUGE algorithm.

Approaches to updating in MIKE11 can also be found in Rungo *et al.* (1989), Refsgaard (1997), Madsen *et al.* (2000) and Butts *et al.* (2005), and include an Ensemble Kalman Filter approach, which is not currently used within the Environment Agency. Further information on the technique described by Rungo *et al.* (1989) is provided in the section on ISIS GAUGE, and this option is used operationally in some of the MIKE11 models in Anglian Region.

Section 3 also describes some aspects of state updating approaches within the rainfall–runoff models used on NFFS. An extensive discussion of possible future approaches to forecast updating, at both gauged and ungauged sites, is also given in Section 7 of the Environment Agency R&D Report *Rainfall–Runoff and Other Modelling for Ungauged/Low-Benefit Locations* (Moore *et al.* 2007). A particular priority identified for further research is to investigate a ‘two-pass state-correction’ approach to forecast updating. The approach is intermediate between error prediction and state correction and can be used to (i) deal with serial-correlation in errors from normal state correction and (ii) correct states forwards in time from the time-origin of the forecast.

FEWS ARMA error correction

The FEWS ARMA error correction module is an implementation of the ARMAsel algorithm of Broersen and Weerts (2005) and Broersen (2007). A description of the implementation is given on the Deltares Wiki pages (Schellekens 2008a) along with the FEWS configuration documentation (Schellekens 2008b).

The module actually comprises two different models: the autoOrderMethod and the fixedOrderMethod. The autoOrderMethod provides model identification and calibration routines, but currently does not permit the use of moving average (MA) terms. The fixedOrderMethod does include MA terms but does not allow model order identification. MA terms will, however, be included in the autoOrderMethod in the next release of FEWS.

The module includes several pre- and post-processing routines:

- mean (bias) correction
- Box–Cox transformation
- interpolation of small gaps.

Differencing to improve stationarity is not available through the module.

FEWS can carry out on-the-fly or off-line calibration of the error correction module. On-the-fly calibration is easily configured, with a relative view period of the observed and simulated time series that are presented to the module determining the length of the calibration period. The ARMA calibration routine in the FEWS error correction module runs very quickly, so the length of the calibration period does not significantly affect workflow run times. Off-line calibration is also possible and can be carried out on the stand-alone NFFS.

CEH ARMA error prediction (PDM, KW, TCM, ISIS,…)

State correction techniques have been developed based on adjustment of the water content of conceptual storage elements in the belief that the main cause of the discrepancy between observed and modelled runoff will arise from errors in estimating basin average rainfall, which in turn accumulate as errors in water storage content. Rather than attribute the cause directly and devise empirical adjustment procedures we can analyse the structure of the errors and develop predictors of future errors based on this structure which can then be used to obtain improved flow forecasts.

A feature of errors from a conceptual rainfall–runoff model is that there is a tendency for errors to persist so that sequences of positive errors (underestimation) or negative errors (overestimation) are common. This dependence structure in the error sequence may be exploited by developing error predictors which incorporate this structure and allow future errors to be predicted. Error prediction using an ARMA (AutoRegressive Moving Average) dependence structure is now a well-established technique for forecast updating in real-time (Box and Jenkins 1970; Moore, 1982, 1999). Error prediction is available as an alternative to empirical state correction in the PDM and TCM (PSM) rainfall–runoff models; it is also provided as the updating technique for use with the KW channel flow routing model.

Predictions of the error are added to the deterministic model prediction to obtain the updated model forecast of flows. In contrast to the state correction scheme, which internally adjusts values within the model, the error prediction scheme is wholly external to the deterministic model operation. The importance of this is that error prediction may be used in combination with any model, be it of transfer function, conceptual or ‘physics-based’ form, and for representing rainfall–runoff or channel flow processes.

The ARMA error prediction module developed by CEH Wallingford for use within the NFFS is most commonly used as an updating technique with the PDM and TCM (PSM) rainfall–runoff models and the KW channel flow routing model. In these cases an ARMA model structure is used to forecast future values of a single time series of model simulation errors. Then the forecast error at each time-step in the future can be added to the corresponding model simulation value to obtain the updated forecast. The forecast error is constructed as a weighted combination of past simulation model errors (and forecasts of them for future times) and past one-step-ahead forecast errors if a moving average (MA) component is included. When a model has multiple time series of errors, for example associated with different observation locations within a hydrodynamic model like ISIS, then the ARMA module can also accommodate this. Both the usual additive errors can be modelled or, through log transformation, multiplicative errors: for further details see Moore (2007).

A quality flag associated with each time series value can be used to invoke interpolation of missing values arising either from intermittent missing observations or those at the end of the time series corresponding to future times. A best estimate of a missing value is obtained as a linear weighted combination of neighbouring observed values, with the weights chosen using a minimum error-variance criterion. The approach involves use of the covariance function of the ARMA model rather than just the ARMA coefficients employed in the normal recursive forecast calculation. Further details are provided in a note prepared for the Environment Agency Thames Region (CEH Wallingford 2005d).

The NFFS module adapter form of the CEH ARMA model algorithm is documented in CEH Wallingford (2004).

The ARMA model is calibrated off-line using a historical dataset. CEH's TSCAL (Time Series CALibration) generic Model Calibration Shell environment used with PDM, TCM (PSM) and KW models provides facilities to automatically estimate the ARMA model parameters for a given model structure. A simplex direct search procedure (Nelder and Mead 1965), modified following suggestions made by Gill *et al.* (1981), is used with constraints that ensure the resulting ARMA model is admissible.

Two practical points should be noted. First, the ARMA model is applied assuming the time series has a zero mean (no constant term is included or sample mean subtracted). This ensures that the updated flow forecasts asymptote to the simulation model forecasts with increasing lead time. In turn, the simulation model flows (e.g. PDM) will decay to zero with no forcing rainfall in accord with the mass balance (unless a constant background flow is included in the simulation model to represent the artificial effects of abstractions/returns). Second, the ARMA model parameters are applied in real time at their fixed values obtained through off-line calibration. This is judged 'safer' than recursively estimating them in real time, especially if there is a risk of telemetry data corruption.

PDM state updating

The term 'state' is used to describe a variable of a model which mediates between inputs to the model and the model output (Szollosi-Nagy 1976). In the case of the PDM rainfall–runoff model the main input is rainfall, and basin flow is the model output. Typical state variables are the water contents of the surface and groundwater stores and of the probability distributed soil storage. The flow rates out of the conceptual stores can also be regarded as state variables: examples are the flow out of the surface storage and the flow out of the groundwater storage. When an error occurs between the modelled and observed value of basin runoff it would seem sensible to 'attribute the blame' to mis-specification of the state variables and attempt to 'correct' the state values to achieve concordance between observed and modelled flow. Mis-specification may, for example, have arisen through errors in rainfall measurement which, as a result of the model water accounting procedure, are manifested through the values of the store water contents, or equivalently the flow rates out of the stores. A formal approach to state correction is provided by the Kalman Filter algorithm (Jazwinski 1970; Gelb 1974). This provides an optimal adjustment scheme for incorporating observations, through a set of linear operations, for linear dynamic systems subject to random variations which may not necessarily be Gaussian in form.

For nonlinear dynamic models (such as the PDM), extended forms of the Kalman Filter that invoke a linearisation approximation may be used but are no longer optimal in the state adjustment provided. The implication of this is that simpler, intuitive adjustment schemes can be devised. These potentially provide better adjustments than the more complex and formal extensions of the Kalman Filter which accommodate nonlinear dynamics through approximations. Moore (1999) called schemes which make physically sensible adjustments 'empirical state adjustment schemes'.

A simple example is the apportioning of the error between the surface and groundwater stores of the PDM in proportion to their contribution to the total flow. The basic equation for state adjustment is mathematically similar to the Kalman Filter: namely the updated state estimate is given by the current state estimate plus an adjustment given by a (Kalman) gain parameter times the error. However, in this case the gain is defined empirically rather than statistically through the relative uncertainty (variance) in the estimates of the current state and observation.

The choice of empirical gain to use is guided by physical insight. For example, when apportioning the error adjustment to the fast and slow flows (or stores) of the PDM this is done in proportion to their relative contribution to the totalled modelled flow (their sum). Two gains applied to the proportion adjustment for each state update are treated as parameters to be estimated through off-line optimisation using historical records. If these gains equal unity then a 'full state update' is realised such that the modelled flow equals the observed flow after the state adjustments have been made. When not equal to unity the gains act as 'relaxation coefficients' providing partial (under or over) adjustments towards the observed flow. Thus the gain used for empirical state updating is given by a physical apportionment rule multiplied by a gain factor parameter.

A range of variants of empirical state updating have been developed for use with the PDM: these are described in Moore (1999, 2007) and CEH Wallingford (2005a). Note that the adjustment is carried out sequentially at every time-step. The variant normally used employs a weighting on the proportional adjustment so that more of the error adjustment is apportioned to the fast (surface) store when active; the slow store will have been adjusted outside flood events.

State correction is essentially a form of negative feedback and, although often very effective, can sometimes give rise to over- or under-shooting behaviour characterised by high accuracy at short lead times but with degraded accuracy at moderate lead times before a recovery in accuracy at longer lead times. This behaviour appears to be associated with a combination of some or all of the following: large gain factors, time lags between the correction of a state value and the appearance of an effect on the modelled flow, and rapid increases in the model error (often due to timing errors on the rising limb). The latter is also a problem for ARMA error prediction schemes.

Both the state correction and ARMA error prediction schemes provided by CEH do not embrace the estimates of forecast uncertainty required by this project. Uncertainty estimates for ARMA models have been available for many years (e.g. Box and Jenkins 1970) but the assumptions underlying these mean they only provide rather crude approximations for errors deriving from hydrological models. Since the empirical state correction methods are based on a Kalman Filter structure, it might be possible to develop approximate estimates of uncertainty through analogy with the Kalman Filter variance update equations. This has not been tried to date. The uncertainty estimation methods discussed elsewhere provide possible approximate solutions to both ARMA error prediction and empirical state correction methods applied to different hydrological models.

Error Forecast Model (MCRM and DODO)

The Environment Agency Midlands Region approach to forecast updating is referred to as the 'Error Forecast Model' or EFM. It examines the difference between observed and simulated outflows over the last m time-steps of the hindcast period (6 hours for the hourly operational model). A judgement is made on how predictable future errors are and forecast outflows are adjusted accordingly. The approach is described in detail in Wallingford Water (1994). The approach extended to work for a variable time-step (from the original hourly one to say 15 minutes) is outlined in CEH Wallingford (2009). It is available for use with the MCRM rainfall–runoff model and the DODO hydrological flow routing model.

In outline, the updating scheme is based on forecasting the error differences, using an exponentially fading weighted average of m-1 (normally five) past observed and/or forecast error differences. The forecast error is then the old error plus the forecast error difference. Adding this to the simulation model forecast gives the required updated forecast. The updating procedure for a period t+1 to 2t, asymptotes the updated forecasts back to the raw forecast at a lead time of 2t. If the updated forecast is negative the forecast reverts to that at the previous time-step. In the event that missing data occur in the hindcast period no updating is attempted and the simulated flows are used over the whole forecast period.

Thus the basic approach is to predict error differences from one (hourly) time-step to the next using an exponentially weighted average of five past error differences. It is equivalent to an autoregressive model structure, of order five, operating on the error differences, and using predefined autoregressive coefficients equal to 0.36, 0.21, 0.12, 0.07 and 0.04. While the error model is stationary in the error differences it is non-stationary in the errors themselves. The model is referred to as a non-stationary AutoRegressive (AR) model within the so-called ARIMA class of models (Box and Jenkins 1970). Its use in the present context stems from its ability to project forward a trend. Forecasts with increasing lead time asymptote to a level determined by the autoregressive coefficients and the five previous error differences. Therefore the non-stationary autoregressive model produces forecasts which are stable, asymptoting to a level (only when used for simulation does the integration, or summation process, characteristic of this model become unstable.)

The EFM contrasts with standard forms of ARMA error correction where the updated forecasts are assured of asymptoting to the simulation model forecast with increasing lead time, provided a bias adjustment is not invoked. This is not generally the case for the integrated form of AR model, which asymptotes to a variable level rather than to zero. Return to the simulation forecast is imposed in the EFM by switching to a second updating scheme after a duration t, which assures a linear approach to the simulation model forecast, meeting it a time 2t. This can create a discontinuity in the updating scheme leading to unrealistic forecast hydrographs.

It is conjectured in Wallingford Water (1994) that the motivation for the EFM approach may stem from abrupt changes in observations arising from observation error: then straightforward application of ARMA models can lead to an error model response function having an abrupt rise and a long tail resulting in unrealistic forecasts. It is suggested that this might be overcome using an AR model with equal roots that makes allowance for observation error, yielding a model with a lower less abrupt peak response but still with a long decay to zero. CEH has not been asked to pursue this further although the method has been coded. A variance ratio of the observation and model errors controls the deflation of the peak response.

For state updating, the techniques used within MCRM and TCM are as follows:

- MCRM state updating – There is no formal algorithm for state updating in the MCRM rainfall–runoff model used operationally. However, the facility exists to manually edit the model states. It is believed that this is primarily restricted to resetting the soil moisture store with reference to Met Office estimates of soil moisture deficit. The potential to reset the states of the snowmelt and reservoir balance model components also exists.
- TCM state updating – The forms of 'empirical state correction' developed by CEH for use with the PDM model (see above) have also been applied to the TCM (PSM) model. The principles are similar but state adjustments are made to the zonal flows (equivalent to adjusting the quadratic saturated stores) of the TCM representing different response zones within the catchment being modelled (rather than the fast and slow components of the PDM flows). Similar proportional adjustments are made in relation to each zone's contribution to the total flow. Further details are given in the PSM Rainfall–Runoff Model User Guide (CEH Wallingford 2005b).

ISIS updating

In real-time forecasting, there is a need to improve existing models in real time by an updating process which enables improvements to the outputs caused by the model failing to meet observed data for whatever reason (e.g. geomorphology, inability for one model to accurately represent every situation). It also provides an updated state from which to begin a forecast.

Since version 2.4.1, ISIS has contained a new unit referred to as the GAUGE unit which facilitates internal updating of flows. This was motivated by functionality required by the Anglian Region flood forecasting team which was regarded as particularly important during procurement of the Anglian Flow Forecasting Modelling System (AFFMS), ultimately developed by the Danish Hydraulic Institute but recently replaced by the Anglian Regional implementation of the NFFS. To date, this unit has only been used at South Bridge in Northampton within the suite of River Nene flood forecasting models.

The GAUGEUnit is a method of specifying a time series of observed water level or discharge at a given node (or nodes) so that the model can self-adjust to meet these conditions; it can also be used to project the error to use for the model forecast.

As with all updating methods, it is not intended to be a substitute for model calibration; such methods are reliant on a well-calibrated model to produce more accurate results.

It is planned that the ISIS inputs used in updating will in the future be made generic in order that different updating methods can be introduced and feed the correction to the ISIS simulation at run time. Currently, one updating method has been introduced, which serves as the ISIS default.

The general response is simple and independent of the eventual update method used – a discharge-time series is applied at the location of the GAUGE unit, acting as a ‘correction flow’ to be added or removed from the model. Where the updating is based on stage, rather than level, rating information must be supplied within the GAUGE unit to associate stage with discharge.

There are two processes to be modelled – updating and forecasting. Updating can occur up to the time of forecasting and involves the process of altering model parameters to bring the simulated results into line with observed results by way of generating an additional flow into the system. Forecasting uses the error estimated during the update period to project into the future. Users may select whether they wish to update only or to forecast (which involves updating during the pre-forecast period) as well.

In either case, a simulation involves one simulation up to the time of forecast with no effect from the GAUGE Unit, from which the model derives the relevant parameters for updating. A further simulation is then performed using the additional flow derived from the updating parameters during the pre-forecast period, projected into the forecast period if appropriate. Thus the updating/forecasting run consists of the original model simulation augmented by extra inflows at the gauged site(s).

If there are multiple gauged sites in one model, then it is possible that these will not be independent, and thus updating each site simultaneously would lead to incorrect updating due to double-counting. The method therefore involves running a number of iterations, updating each site simultaneously. The number of extra iterations should be at least equal to the maximum number of non-independent sites. The user is able to specify the number of updating iterations (defaulting to the number of updating sites + 1) in the gauge control file.

The default method that is currently used for ISIS updating is based on the approach by Rungo *et al.* (1989). No estimates of uncertainty are provided as part of this procedure. The updating procedure involves estimation of a phase and amplitude error of the measured quantity during the pre-forecast period. The estimation of both these components of the error helps distinguish between those errors; for example, if one was correcting the error purely on amplitude, then a phase error could be interpreted as a large amplitude error.

As the above methodology describes, there is a run-time cost associated with this approach to error correction since it involves a minimum of one extra model run and in general one additional run per internal updating location. For larger and slower models, this can be a very high price to pay.

4.3.4 Data assimilation in integrated catchment models

The major constraint in uncertainty propagation through cascades of flood runoff and flood routing models is the nonlinearity of the component models. This means that, in general, analytical propagation of errors will not be possible and that recourse must be made to approximate numerical methods. It is then important to differentiate between two cases: situations where data assimilation is possible and situations where data assimilation is not possible. When, during an event, telemetry fails at a site, then the first case may revert to the second.

Data assimilation allows, for any site having transmitted observational data, the updating of forecasts at that point and the uncertainty associated with the forecasts, conditional on the inputs from the model component upstream (ultimately estimates of rainfall input). These inputs will be uncertain, but in both model calibration and data assimilation, it is possible to condition both the forecast and uncertainty estimate on the best estimate of the upstream input. Uncertainty propagation at each site is then only required in time out to the required lead time for the forecast. This is a very simple and computationally efficient way of handling the propagation of uncertainty in integrated models. It will not necessarily underestimate the forecast uncertainties because of the way in which the data assimilation allows the uncertainty estimates to be updated, dependent on the forecast innovation at each time-step. It can be used regardless of what data assimilation methodology is being used.

Alternatively, and more accurately (at least for well-behaved cases), techniques such as the Ensemble Kalman Filter and particle filters can allow for input uncertainties, propagated through a model component within the data assimilation framework. There has, to our knowledge, been little research to compare these approaches within hydrological forecasting situations.

For the case of propagation without data assimilation, this will generally be more computationally demanding because the uncertainty in an upstream component will become the input to an uncertain component downstream. Unconstrained by data assimilation, the uncertainty in the forecasts will therefore grow as propagation proceeds through more downstream components. Propagation then needs to proceed sequentially. Monte Carlo sampling, with a sufficient number of samples, will be the most accurate method of propagation through nonlinear components, but will be computationally expensive. Approximate sampling methods can also be used, such as Latin hypercube sampling, which might lead to considerable run time savings, particularly in the case of multiple correlated inputs to a component. Simpler techniques, in which the Monte Carlo sampling is performed off-line, might also be considered.

4.4 Generic tools and techniques

This section describes three 'toolkit' approaches to support the implementation of data assimilation and/or uncertainty estimation techniques. The methods which are described are DATools, GLUE and CAPTAIN, and Table 4.1 summarises the techniques which are implemented within each approach.

Table 4.1 Summary of functionality of the generic tools and techniques.

Tool or technique	Forward uncertainty propagation	Data assimilation and uncertainty estimation
DATools	Via specification of boundary condition, state and/or parameter uncertainties	Ensemble Kalman Filter, and residual resampling filter
GLUE	Via specification of boundary condition, state and/or parameter uncertainties	Updating of likelihood weights for ensemble of behavioural models
CAPTAIN	Via specification of boundary condition, state and/or parameter uncertainties	Predictor corrector filters, including Kalman Filter

DATools

DATools is a generic software package for data assimilation (El Serafy *et al.* 2007; Weerts 2007; Weerts *et al.* 2009). DATools can be used standalone or within Delft-FEWS. DATools is completely configurable via XML configuration. Using DATools, it is possible to apply data assimilation methods for existing and new models. The focus of DATools lies in enabling data assimilation methods for operational forecasters using Delft-FEWS, but it can also be used for academic studies (standalone version). This means that the PI interface of Delft-FEWS is supported, although some additional requirements with respect to state exchange have been added.

Configuration of DATools can be done via XML configuration. This makes the system very flexible and easy to understand. Within the DATools software it is also possible to perform uncertainty analysis using a module called UATools. This works in a similar fashion to DATools via XML configuration. In the near future model calibration will also be included in DATools.

To be able to use DATools the model needs to be linked via a model adapter to DATools. Existing adapters for FEWS normally only exchange input time series and output time series. To be able to update states or change parameters (for uncertainty analysis) more information needs to be exchanged between the model and DATools. Therefore, the existing adapters need to be adapted/extended to handle exchange of model states and parameters.

DATools is used operationally in FEWS-NL (Weerts 2007, 2008). The hydrodynamic model SOBEK-RE of the Rhine is updated every 2 hours in the historical run. To limit the runtime (forecast runs are much longer than update runs), the mean state is used in the forecast run (although it is also possible to run the full ensemble in the forecast). UATools has also been used for performing uncertainty analyses in several projects by Deltares.

GLUE

GLUE (Generalised Likelihood Uncertainty Estimation) is a methodology for the representation of complex uncertainties in modelling through conditioning of Monte Carlo realisations of one or more model structures. Each realisation is compared with any available observations and given a likelihood measure that is used to weight the prediction of the model realisation when used in prediction. Models that do not give acceptable predictions of the observations are rejected as non-behavioural and given a likelihood of zero. Although GLUE is general, in the sense that formal statistical likelihoods can be used where there is a belief that a simple statistical model is an adequate representation of the errors (see, for example, Romanowicz *et al.* 1996), it is based on a quite different philosophy of model error that accepts that uncertainties may be epistemic rather than probabilistic and accepts the potential for equifinality in model results (Beven 2002, 2006).

GLUE can use a variety of different formal and informal likelihoods, including fuzzy measures, and is flexible in the way in which they are combined as new observational data become available (e.g. Beven and Freer 2001; Beven 2006; Page *et al.* 2007; Smith *et al.* 2008c; Li *et al.* 2009). Uncertainty in the observables used in model calibration can easily be incorporated into the conditioning process. GLUE has been used as an uncertainty estimation methodology for forecasting with data assimilation (e.g. the flood inundation forecasting example of Romanowicz and Beven1998), and can be viewed as a form of particle filtering (e.g. Smith *et al.*

2008b) but, because of its basis in Monte Carlo sampling, it will generally be computationally expensive for forecasting. It has also been used in conditioning a variety of rainfall–runoff models (e.g. Beven and Binley 1992; Beven and Freer 2001; Page *et al.* 2007; Li *et al.* 2009) and flood routing models (e.g. Aronica *et al.* 1998; Bates *et al.* 2004; Pappenberger *et al.* 2005b, 2007a, 2007b).

More information on the GLUE methodology may be found in Beven (2009, Section 4.5, Box 4.4).

CAPTAIN

The Computer Aided Program for Time-series Analysis and Identification of Noisy systems (CAPTAIN) is a toolbox of computational routines (m-files) for use within the Matlab-Simulink software environment (see http://www.es.lancs.ac.uk/cres/captain/ and Taylor *et al.* 2007). Almost all of the routines are numerical recursive algorithms that can be used within a flood forecasting and warning environment. Moreover, by combining them in a customised manner, if necessary with a graphical user interface, it is possible to synthesise a complete flood forecasting and warning system. For example, the quasi-distributed forecasting system for the River Severn, developed at Lancaster University under the aegis of the FRMRC, was implemented in this manner (see Romanowicz *et al.* 2006b; Young *et al.* 2006).

The algorithms in CAPTAIN can be divided into the following four categories that are relevant within the present flood forecasting and warning context (an additional category is concerned with digital control system design). In almost all cases, the various routines are able to handle and infill missing data.

i. Recursive estimation algorithms for time series analysis. These include the Kalman Filter (KF) for data assimilation and forecasting (see Section 4.3.1), which can be made adaptive by the incorporation of the recursive parameter estimation algorithms mentioned below under category ii; the associated Fixed Interval Smoothing (FIS) algorithms for the off-line estimation, interpolation and smoothing of state variables in state-space models; related KF/FIS algorithms for the optimal estimation and forecasting of ‘unobserved component models’ for univariate time series, such as rainfall and flow series, that include trends and harmonic/quasi-harmonic components (as in tidally affected data). The KF algorithm has obvious application in the construction of flood forecasting systems based on state-space models; and the FIS algorithms provide a very flexible tool for (a) the off-line decomposition of time series into physically meaningful components; (b) infilling series by interpolation over gaps; and (c) the on-line forecasting of the components, separately or combined.

ii. Iterative and recursive-iterative algorithms for identifying and estimating linear and nonlinear (state-dependent parameter) transfer function (TF) models in either discrete or continuous-time. The discrete-time TF is implemented as a difference equation model and the continuous-time TF as a differential equation model, both of which are used widely in conceptual hydrological models for the representation of model stores. The data-based, multi-order TF models can be linear or nonlinear and are the major tool in DBM modelling (see Section 4.2), where they are normally decomposed into physically interpretable serial and parallel connections of hydrological stores (generalisations of PRTF models in the simplest linear case). The recursive options for these algorithms allow for their use on-line in real time for applications such as data assimilation and adaptive forecasting.

iii. State-Dependent Parameter (SDP) identification and estimation of nonlinear models. The SDP algorithm investigates whether any parameters in linear models are nonlinearly dependent on other measured variables and estimates the nature of this dependency, initially in non-parametric, graphical form. This converts the linear model into a widely applicable nonlinear form: typical hydrological examples are the DBM rainfall–flow and nonlinear flow routing model, where the finally parameterised SDP nonlinearities have a clear hydrological interpretation and simply convert the input into an 'effective' series (e.g. effective rainfall).

iv. A large collection of other routines and algorithms for a wide variety of tasks, including data analysis (correlation analysis, spectral analysis etc), model order and structure identification (criteria such as AIC, SIC, YIC etc), TF model decomposition, and model diagnostics (Nash-Sutcliffe Efficiency, time-variable parameter estimation etc).

4.5 High-Level uncertainty framework

4.5.1 Introduction

As part of Phase 1 of this project (Task 1.3), a start has been made on developing a generic unified framework for assessing the uncertainty of a forecast in NFFS (i.e. predictive uncertainty). The development of the framework is being performed in three phases:

- Phase 1 – High-level framework – what needs to be considered (Task 1.3).
- Phase 2 – Detailed framework – the choice of approach for each combination of circumstances (Task 2.1).
- Phase 3 – Guidelines for these different situations (Task 3.1).

Table 4.2 shows the items which are being considered in these different phases of the project.

This section focuses on the development of the high-level framework (Task 1.3, Phase 1), which will be developed to a more detailed level in Task 2.1 and beyond once the high-level concept has been agreed with the Project Board. The objectives of the framework are summarised in Section 4.5.2 while Section 4.5.3 describes the key components. Finally, Section 4.5.4 summarises the uncertainty estimation techniques which might be considered for inclusion in the framework.

Table 4.2 Indicative summary of the key considerations in development of the uncertainty framework and guidelines.

Item	Task 1.3 – High-level framework	Task 2.1 – Detailed framework	Task 3.1 – Guidelines
Level of risk	Identify the preferred approach to estimating risk, and the key factors in the framework which will be linked to risk. Also how this links to the complexity of approach	Define the choices and options that are available in each situation	Provide descriptions and tools (e.g. flowcharts, risk matrices) to help users with choosing an appropriate method
Lead time requirements	Identify the key areas in which lead time requirements will influence the choice of uncertainty estimation techniques	Define how lead time requirements relate to catchment response time, and hence to choice of an appropriate uncertainty estimation technique	Provide guidance on how to estimate forecast lead time requirements and relate these to catchment response times, and hence to choice of technique
Types of models	List the uncertainty estimation methods that are potentially suitable for general types of model (rainfall–runoff, flow routing, hydrodynamic etc) and the specific types of models (PDM, ISIS etc) which will be considered in the framework	Define which methods are applicable to each general and specific type of model, taking account of model-specific issues, the capabilities of NFFS, likely model run times in probabilistic mode etc	Provide guidance on applying the recommended methods to each type of model
Sources of uncertainty	Define the sources of uncertainty that will be considered in the framework, and how these relate to lead time requirements, and list possible methods that can be used for each source	Define possible methods that might be used in a forward uncertainty analysis for each source and general type of model considered in the framework, taking account of catchment-specific issues (tidal influences, control structures, reservoirs etc)	Provide guidance on identifying the key sources of uncertainty in models based on forecast lead time requirements, catchment response times to forecasting point(s), catchment-specific issues, data availability, requirements for hindcasts etc
Data assimilation	Describe the main factors to consider when deciding whether to use data assimilation, and how these relate to the level of risk	Describe the techniques that are appropriate in each modelling situation	Provide guidance on the data requirements for each technique described in the framework and the consequences of data errors/instrument failures etc
Operational requirement	Define the operational situations that might influence the choice of uncertainty estimation technique and which will form decision points in the framework (and how these link to lead time requirements)	Define which uncertainty estimation techniques are likely to be appropriate in each situation, taking account of lead time requirements	Provide guidance on common types of operational requirement in the Environment Agency (with examples), and the techniques which could possibly be used operationally
Run times	Define the choices for reducing model run times that will be considered in the framework	Define the methods that are appropriate for the different modelling situations described in the framework	Provide guidance on how each technique might be implemented

4.5.2 Objectives of the framework

The framework will provide guidance on the methods to be used to quantify and reduce the uncertainty of the individual components as well as to quantify and reduce the uncertainty of a complete model cascade, and will categorise each approach by factors that determine its effectiveness including:

- type of model (e.g. rainfall–runoff, flow routing, hydrodynamic);
- lead times (targets, technically feasible);
- availability of good quality river gauge telemetry data for data assimilation;
- complicating factors (e.g. control structures, storage).

The framework will also form the basis for some of the methodologies to be incorporated in the guidelines to be produced within Task 3.1 of this project.

Although the focus is on the types of conceptual rainfall–runoff, flow routing and hydrodynamic models currently used operationally within the Environment Agency (both individually, and in end-to-end integrated catchment models), the framework is being developed in a way that is robust and flexible enough to allow for the inclusion of additional uncertainty techniques later on (outside this project), and aims to form a common framework to assess the uncertainty in flood forecasting models from end-to-end, which is robust, scaleable and risk-based.

The framework should then allow users to meaningfully combine the probabilistic treatment of important sources of uncertainty to gain an understanding of the overall uncertainty in the flood forecast. Numerical Weather Prediction (NWP) rainfall ensembles, as tested in the 'Hydrological Modelling with Convective Scale Rainfall' project, are also being considered within the framework (although consideration of how to generate these ensembles is outside the scope of the present project).

The following section describes some of the key issues which are highlighted in Table 4.2, and the discussion proceeds in the same order as in that table.

4.5.3 High-level framework

Level of risk

Defra and the Environment Agency are moving towards a risk-based approach to flood risk management, and it has been agreed that the framework should take account of the level of risk when selecting appropriate techniques.

Clearly, at high-risk locations (high probability of flooding, and large potential damage or risk to life) the effort of setting up an advanced probabilistic forecast system is more justified than for low-risk areas. For example, if the flood damage can be reduced by taking timely response measures and the expected damage reduction exceeds the investment in the forecast system, then this adds to the economic case for improvements to the forecasting system. In many practical applications, this may amount to the following criteria:

- The probability of flooding and the potential damage and/or loss of life in the area of interest are relatively high.
- The flood damage can be reduced significantly by timely response actions.
- A probabilistic component to the forecasting system is technically feasible.

In development of the guidelines and uncertainty framework in later tasks (e.g. Tasks 2.1, 3.1), a number of risk-based approaches could be considered to assess which are the most important sources of uncertainty in different situations. Examples could include:

- estimates for economic damages;
- measures of vulnerability on the floodplain;
- simple multi-criteria approaches.

For example, models could possibly be optimised based on vulnerability measures (Pappenberger *et al.* 2007a) in which one set of models might be suitable for predicting the risk of flooding in one settlement, while another set of models might be suitable for predicting the risk of flooding in another settlement or at a road junction critical to evacuation plans. Local level sensors might also be installed to allow data assimilation algorithms to be implemented to improve the location forecasts for that location, with sufficient lead time to allow any damage mitigation measures to be effective.

Other simpler indicators could also be considered, such as the number of properties at risk, information derived by regional teams on failure to meet Environment Agency forecasting targets (POD/FAR), or existence of a Major Incident Plan for a location. The National Flood Risk Assessment (NaFRA) dataset was also proposed by a reviewer, but the Project Board felt that this was too broad-brush for this study in its current form.

To assess the level of risk, an approach based on the Environment Agency's Flood Warning Level of Service Risk Matrix is proposed (see Figure 2.6), in which risk is based on the combination of the number of properties at risk, and the probability of flooding (expressed using the standard of protection at the site of interest). However, this approach is aimed primarily at a single forecasting location, whereas for an integrated catchment model forecasts may be provided for several locations, each with different numbers of properties and standards of protection. During Phase 2 of the project, some thought will therefore be required on how to apply this approach to multiple forecasting locations.

Having assessed the level of risk, this can be used as a guide to which approach to use, both in choice of techniques, and the level of complexity of analysis work to come to an optimal modelling solution. For example, in the Real-Time Modelling Guidelines (see Section 2.2.1), two levels of detail were used depending on the level of risk:

- Method A – a purely qualitative approach suitable for a rapid first assessment of potential modelling solutions.
- Method B – the main model selection approach, which aimed to arrive at a reasonable compromise between technical, cost, benefit and other considerations.

A similar approach might be adopted here, with simple and more complex uncertainty estimation approaches (although cost-benefit considerations are outside the scope of this project). The level of risk might also influence the reliance placed on data assimilation, with a trade-off between requiring higher accuracy at high-risk locations (e.g. city centres), to which data assimilation can contribute, and the risks if a key instrument fails during a flood event.

Lead time requirements

In the Real-Time Modelling Guidelines (Section 2.2.1), the lead time requirement was a key aspect of the model selection process, where it was noted that 'The choice of catchment model generally involves a trade-off between accuracy and minimum warning time requirements, with rainfall–runoff models using rainfall forecasts giving the longest lead time, and lowest accuracy, and at-site triggers often giving the most accurate results, although with the shortest lead time'.

Table 4.3 illustrates some of the typical trade-offs between forecast lead time and accuracy with – in general terms – lead time decreasing, and accuracy increasing, moving down the table.

Table 4.3 Indication of the relationship between input data and the type of forecast model (Environment Agency 2002).

Input	Type of forecast model
Rainfall forecasts	Rainfall–runoff
Rainfall measurements	Rainfall–runoff
Forecast inflows	Routing
Measured inflows	Routing
At-site levels and flows	Local model
Levels at or near the site	None (triggers)

The combination of accuracy and lead time defines the quality of the forecast and thereby the added value in terms of potential reductions in flood damage and loss of life. Generally, the potential damage reduction increases with increasing lead time and accuracy up to a point of diminishing returns, beyond which any additional lead time is of little further benefit. Also, as described in the following section, in many situations there are often minimum lead times required for specific response actions.

In practice, the choice of model type and input data will depend on a balance between the operational requirement for lead times and the catchment response time, as described in the following subsections.

The forecasting lead time requirement depends primarily on the lead times required by flood warning duty officers, local authorities and the emergency services. Often, these will be times to the crossing of a flood warning (Action or Response) threshold. Typically, shorter lead times may be used in the actual operational flood warning, while forecasts at longer lead time are used as guidance.

For example, the consultations and review have suggested that the following typical lead times are required:

- 1 or 2 hours – national target (Wales and England respectively).
- 4–6 hours – initiating Major Incident Plans/Severe Flood Warnings.
- Several hours – installation of temporary/demountable barriers.
- Several hours to days – planning staff rotas, early warnings to professional partners, flood watches etc.

These times need to be related to typical catchment response times to the forecasting point(s) of interest as discussed in the following subsection.

In real-time use, the model cascades within NFFS are run in two principal operational modes:

i. a historical mode

ii. a forecast mode.

In the first mode the models are forced by hydrological and meteorological observations over a limited time period prior to the onset of the forecast in order to update the model states. In the second mode, the models are run over the required forecast lead time, and the internal model states at the end of the historical run are taken as initial conditions for the forecast run.

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 rainfall forecasts, rainfall observations, and rainfall–runoff, flow routing and hydrodynamic models (as appropriate). As the process of forecasting is geared primarily towards providing timely and accurate information for the flood warning duty officer when deciding whether to issue a flood warning, the most important uncertainties within the process at that lead time will need to be considered.

Within NFFS two cases can be distinguished:

- Type 1: Forecasts can be made with sufficient lead time directly from measured rainfall and/or radar and/or uncertain radar projections and/or measured or forecast flows at a station further upstream.
- Type 2: Forecasts where lead times are not sufficient using measured data so that rainfall forecasts are critical. There are two sub-cases:
    - small catchments – in this case, NWP (ensemble) forecasts are unlikely to be useful at small catchment scale except for general severe event warnings;
    - large catchments where long lead times (24 hours) are required for decisions about staff mobilisation etc (normal warnings would come under Type 1 cases in large catchments).

In England and Wales the Type 1 case dominates. However, flow forecasts based on NWP rainfall forecasts (Type 2) are increasingly used within the Environment Agency to provide guidance on likely mobilisation and other requirements in advance of a flood event, and the issuing of Flood Watches, although at present are not used directly for issuing Flood Warnings or Severe Flood Warnings.

In the Real-Time Modelling Guidelines, for the indicative method (Method A), the following rule of thumb (adapted from Reed 1984) was proposed based on catchment response time (Tp):

- Tp ≤ 3 hours – Rainfall–runoff modelling based on rainfall forecasts (STEPS/radar-only).
- 3 ≤ Tp ≤ 9 hours – Rainfall–runoff modelling based on rainfall actuals (raingauge/radar) (or rainfall–runoff modelling combined with routing in the lower reaches).
- Tp ≥ 9 hours – Flow routing.

For the more detailed method (Method B), the various time delays in receiving telemetry data, deciding to issue a warning and issuing that warning, were also considered in relation to the catchment response time (i.e. to the locations where forecasts are required). These approaches might be adapted to advise on the type of uncertainty estimation and data assimilation techniques which are most appropriate for different lead times.

Alternatively, an approach which was suggested in the Probabilistic Flood Forecasting Scoping Study (Environment Agency 2007) might be adopted and developed further. This was based on a view of catchment response zones proposed by Lettenmaier and Wood (1993).

The method is based upon a set of criteria which compare the desired warning lead time T_w to the hydrological response time T_P at the location for which the forecast is to be provided, ignoring any time delays in the detection, forecasting and warning aspects of the system (Figure 4.6). This hydrological response time is further sub-divided into the time that water needs to flow through the river channel (T_r) and the time that the water needs to flow from the land phase into the river (T_s). The division between the land phase and the river channel is somewhat arbitrary, but generally the river channel is considered to be the main river (system), while the response of the land phase is the response of (sub)catchments before the water flows into the main river system. The following four situations are defined:

i. $T_w < T_r$ or $T_s << T_r$. The warning will be issued on the basis of water that is already in the main river channel; or the time the water needs to flow from the land phase into the river is insignificant compared to the time the water needs to flow through the main river. This may be the case for forecast point VII in the figure, assuming that catchments E and F have only a minor contribution.

ii. $T_w < T_p$ and $T_s \approx T_r$. The warning will be issued on the basis of water that is still on the land phase and the response time is determined by the time this water needs to flow from the land phase into the river channel as well as by the time the water needs to flow through the main river. This may be the case for forecast point IV in the figure.

iii. $T_w < T_p$ and $T_s >> T_r$. The warning will be issued on the basis of water that is still on the land phase and the response time is mainly determined by the time this water needs to flow from the land phase into the river channel. This may be the case for forecast point I in the figure.

iv. $T_w > T_p$. The desired lead time is such that a warning may be issued on the basis of water that has not yet fallen as rain. In this case also a rainfall forecast is needed for a timely forecast.

Cases i to iii are typically applied for short-range forecasting in medium and larger basins. Case iv is typically applied in either medium to long-range forecasting in larger river basins or for forecasting in small (flashy) river basins. Of course, for the longer lead time situations, forecasts may rely in the early stages of the event primarily on rainfall forecasts or observations, and may exhibit significant reductions in uncertainty as input data streams switch as the event progresses (e.g. from rainfall forecasts to rainfall observations to a flow routing approach). The magnitude of these changes in uncertainty may depend on whether a single model parameter set is used, or whether calibrations have been performed separately for each type of input data stream.

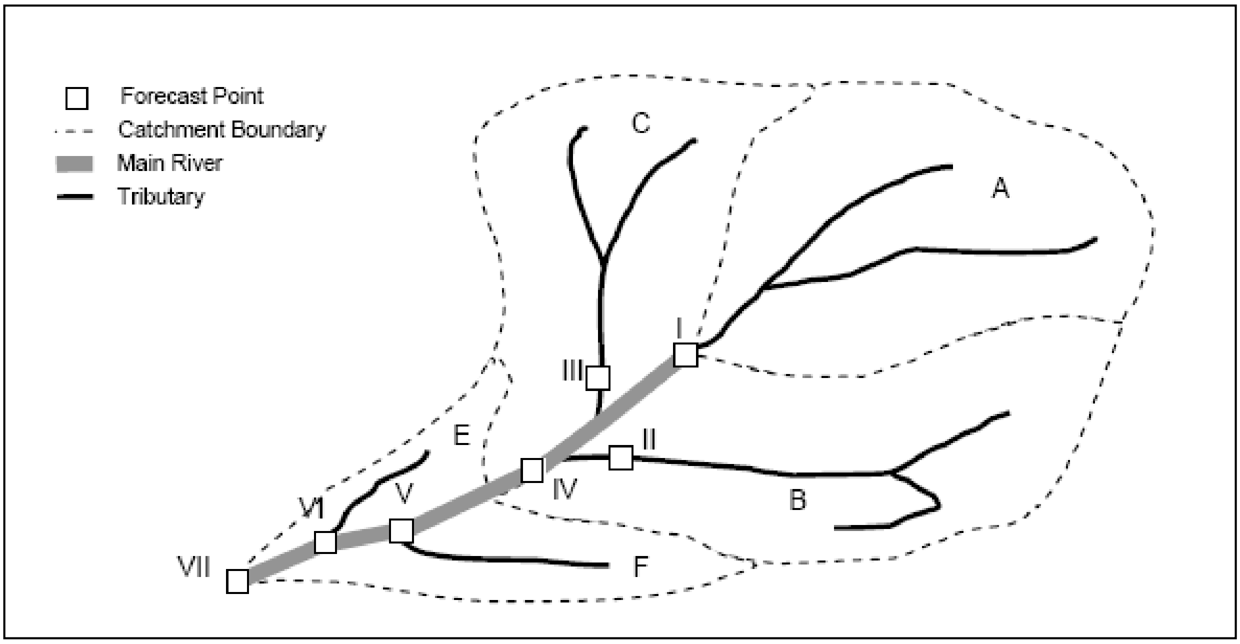


Figure 4.6 Schematic layout of a catchment, including the main river, tributaries and catchments.

If this approach was used, then a tentative classification scheme (which could be developed further) might be as shown in Table 4.4 for some of the more common sources of uncertainty in catchment modelling.

Table 4.4 A possible uncertainty classification scheme for fluvial flood forecasting using the Lettenmaier and Wood (1993) classification scheme (Environment Agency 2008).

Type	Primary sources of uncertainty	Secondary sources of uncertainty
1	- High flow ratings - Hydraulic/routing model parameters - River channel/floodplain survey	Likely - Abstractions/discharges - Runoff from lateral catchments (Type 3 or 4) Depends on catchment/flood risk area - Tidal boundary - Washland operations - Tidal barrier operations - River control structures - Flood defence geometry and condition
2	- A combination of Types 1 and 3	- A combination of Types 1 and 3
3	- Rainfall actuals - Rainfall–runoff model parameters - Antecedent conditions	Likely - River levels (if updating) - High flow ratings (if updating) Depends on catchment/flood risk area - Snowmelt - Reservoir operations - Flood defence geometry and condition
4	- Rainfall forecasts - Rainfall–runoff model parameters - Antecedent conditions	Likely - River levels (if updating) - High flow ratings (if updating) Depends on catchment/flood risk area - Snowmelt - Reservoir operations/state - Flood defence geometry and condition

Types of models

The possibilities for estimating uncertainty depend on the type of model both in general terms (e.g. rainfall–runoff model) and the specific modelling package (PDM, ISIS). Some factors to consider include:

- the main sources of input data (or forecasts) to the model;
- typical model run times;
- the way that the model is implemented in NFFS (e.g. PI XML files);
- the sensitivity of the model to individual model parameters.

The consultations and review have suggested that, based on the current types of integrated catchment models used within the Environment Agency, at least the following types of model may need to be considered in the framework (although will not necessarily all be considered in the case studies):

- PDM – conceptual rainfall–runoff model
- MCRM – conceptual rainfall–runoff model
- TCM – conceptual rainfall–runoff model
- NAM – conceptual rainfall–runoff model
- DODO – flow routing model
- VPMC – flow routing model
- KW – flow routing model

- ISIS – hydrodynamic model
- MIKE11 – hydrodynamic model.

The Real-Time Modelling Guidelines (Section 2.2.1) also noted that the following catchment-related issues could be factors in model selection, and similar considerations might also apply to the choice of uncertainty estimation method (to be investigated):

- fast response catchment
- floods can occur on a permeable or dry catchment
- flood response can vary depending on spatial variations in rainfall
- groundwater influences
- large lowland chalk or clay catchment
- urban catchments
- ungauged catchment
- snowmelt
- simple river reach
- flat, lowland river
- simple floodplain
- embanked floodplain
- levels only at reach ends
- reservoirs
- natural lakes, bogs and wetlands
- mobile river bed
- tributary inflows
- fan-shaped flow networks
- flow control structures (sluices, barrages etc)
- off-line storage, abstractions, discharges and diversions (e.g. washlands, pumps, flood relief channels) during flooding conditions
- event-specific problems.

Sources of uncertainty

The major sources of uncertainty that need to be addressed for integrated catchment models operated in NFFS include:

- uncertain boundary conditions (current and future);
- uncertain initial conditions;
- structural errors in each model of the model cascade;

- uncertain model parameters;
- uncertain structure operating rules.

In general, there will be sources of uncertainty at each level and step in the model cascade, and this section highlights (at a high level) the issues which will need to be considered further during Task 2.1 of the project. In particular, the consultations suggested that the following sources of uncertainty are most important in the models which are currently operated by the Environment Agency:

- catchment averaging of raingauge data
- validity of rating curves
- model calibration (hydrodynamic models)
- model calibration (rainfall–runoff models)
- representation of floodplain storage
- representation of antecedent conditions
- representation of ungauged inflows
- influence of structure operations.

The issue of uncertainty in rainfall forecasts (STEPS, MOGREPS etc) was also identified as important, although outside the scope of this project (other than for consideration in the uncertainty framework).

The sources of uncertainty affect the uncertainty in the forecast in different ways, and the relative contributions depend on the lead time and locations of interest. As an example, Figure 4.7 shows indicative estimates for the contributions to the total uncertainty for a downstream location in a large river. For this location, the uncertainty of the precipitation forecast has a limited influence for a very short lead time but becomes dominant at longer lead times.

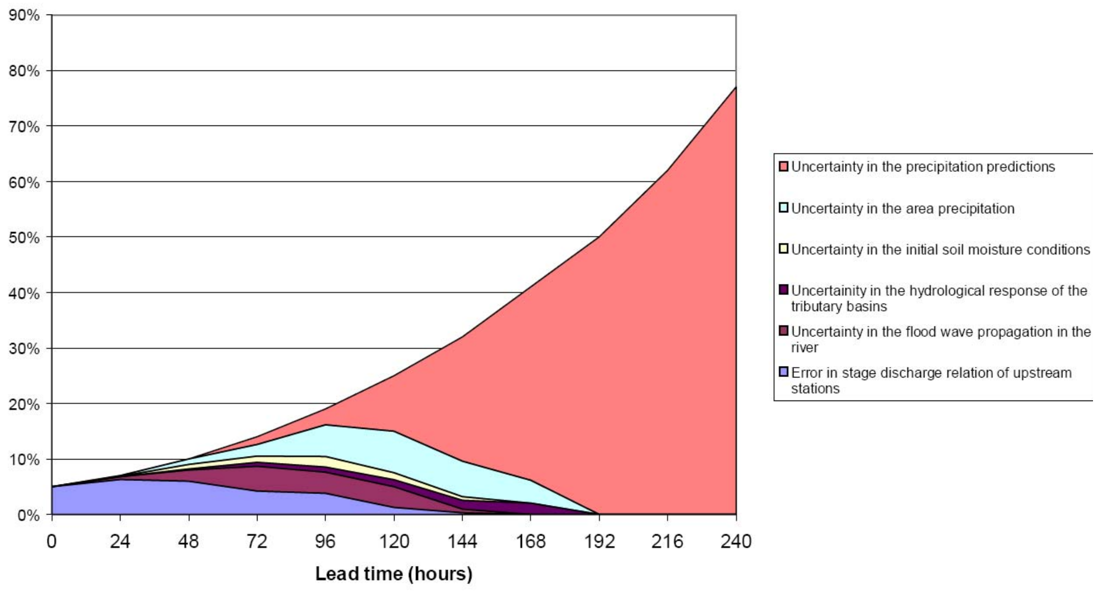


Figure 4.7 Illustrative example of contributions to the total uncertainty as a function of lead time for a specific downstream location in a large river.

However, for locations further upstream, the situation would be very different from that depicted in the figure.

One route to estimating uncertainty and its relative contributions is to use a forward uncertainty analysis, in which the likely uncertainty in model parameters, input data, model state etc are hypothesised, and fed through the model (or chain of models) to estimate the overall uncertainty in flows and/or levels. Some forward uncertainty propagation techniques include:

- analytical approaches
- ensemble techniques
- multi-model techniques
- Monte Carlo approaches
- fuzzy set methods.

Table 4.5 provides some examples of forward uncertainty propagation techniques for the sources of uncertainty shown above.

Table 4.5 Examples of forward uncertainty techniques.

Source of uncertainty	Examples of techniques
Catchment averaging of raingauge data	Monte Carlo sampling of raingauge weights, radar rainfall fields, catchment average rainfall, sampling from a range
Validity of rating curves	Decomposition of error sources, addition of variances (ISO/British Standard procedures), Monte Carlo sampling, evaluation of historical performance, sampling from a range
Model calibration (hydrodynamic models)	Monte Carlo sampling of roughness coefficients, evaluation of historical performance
Model calibration (rainfall–runoff models)	Monte Carlo sampling of key model parameters and/or model initial conditions, multi-model comparisons, evaluation of historical performance, sampling from a range
Representation of floodplain storage	Quasi 2D flood modelling using Monte Carlo sampling or sensitivity analysis based on uncertain model parameters, expert judgement or comparison with observations, Monte Carlo sampling of states (storage volume) or parameters (storage properties)
Representation of antecedent conditions	Monte Carlo sampling of states (e.g. soil moisture deficit) or parameters (runoff coefficient)
Representation of ungauged inflows	Monte Carlo sampling of parameters (e.g. scaling factors on nearby gauged flows), expert judgement, sensitivity analyses
Influence of structure operations	Expert judgement, sensitivity analyses

A key task in Phase 2 of the project will be to review which of these (or other) techniques are appropriate for the types of models currently used in NFFS (PDM, MCRM, TCM, NAM, KW, DODO, ISIS etc). Also, for some techniques, an archive of both data and forecasts is required to assess past forecast accuracy, which requires historical records of sufficient length (with an appropriate number of historical flood events), and the ability to regenerate (hindcast) forecasts, which can be an issue for some types of forecast (e.g. rainfall–runoff model forecasts based on ensemble NWP inputs, or radar rainfall fields).

The influence of each source of uncertainty depends on its contribution to the total uncertainty for the lead times and locations of interest. Therefore, to apply the framework, a starting point could be to perform a qualitative analysis for the catchment of interest to produce a shortlist of uncertainties that are relevant and which it might be useful to propagate through the model cascade, using methods such as those shown in the table above.

Data assimilation

Data assimilation (or real-time updating) can improve the accuracy of forecasts, provided that the data used are of sufficient quality and reliability. In many cases, an error correction or state updating algorithm is used to help to allow for differences between the observed and forecasted values, based on performance evaluated in the historical mode of operation, or using parameters calculated off-line. The consultation exercise showed that the following automated methods are currently in use within NFFS:

- Delft-FEWS ARMA error correction
- CEH ARMA correction (used with PDM, KW, TCM)
- MCRM error correction
- PDM state updating
- MIKE11 error correction
- ISIS GAUGE.

An indirect form of updating is also employed for the two event-based types of model used within NFFS (PRTF and HYADES), in which the most appropriate parameter set to apply is selected on the basis of current flows and/or catchment conditions.

For some approaches to data assimilation, such as the Kalman Filter, an assessment of uncertainty can also be provided. The following general factors should also be considered:

- Quantification and reduction of the forecast uncertainty is only possible when measurements are available. Note that the measurements themselves are also uncertain (rating curves, measurement errors etc). This should be taken into account (if possible) when reducing and quantifying uncertainties.
- Ideally, uncertainties are quantified and treated/reduced closest to their origin. For the Type 1 case above, data assimilation is crucial and should be used wherever possible. The major issues with this case are the propagation of uncertainty and best methods of data assimilation in cascades of models. For Type 2, data assimilation may still be crucial depending on the catchment response and lead time requirement.
- Uncertainty quantification of the forecast is possible at each step in the processing chain (for a single model or for the complete model cascade) by making the current forecast conditional on past performance (given

that the lengths of historical and forecast records are long enough). This holds for both the Type 1 and Type 2 case above.

In some cases, post-processing of ensemble or probabilistic outputs may be desirable to improve the mean and spread of estimates (e.g. using quantile regression or Bayesian uncertainty techniques).

In addition to data quality (high flow performance etc), another factor to consider is the resilience of the data assimilation process and the level of risk, for example if an instrument fails. Also, under current procedures in NFFS, duty officers can view forecasts both with and without updating to help to identify whether any problems have occurred in the updating process (e.g. due to timing errors in model outputs). One question to consider in Task 2.1 is therefore the importance of this type of comparison since it may not be possible/meaningful for some types of data assimilation algorithm.

If a chain or cascade of models is used (i.e. an integrated catchment model), there is also the issue of how the outputs from individual models are to be combined, and whether the outputs are to be updated at each boundary between models (i.e. at several points, moving down the catchment), or only as a final step at the forecasting point(s) for which the forecast is required. There may also be issues to consider regarding whether values are updated in terms of levels or flows, with various factors to consider which may affect the updating procedures, such as backwater influences, tidal influences, the influence of structure operations, and the suitability of high flow ratings.

Operational requirements

Another factor to consider in the framework may be the intended operational use of the forecasts. For example, for some applications a simple estimate or visualisation of the spread of estimates may be sufficient. However, if the probabilistic inputs are being used directly in major decisions (e.g. evacuating a city), or as inputs to another organisation's decision support system, then more formal estimates of probability may be required (perhaps with conditioning or post-processing of outputs).

The question of how the forecasts are evaluated is also important, and links to skill scores and other performance measures. Ideally, improvements in the forecast system (i.e. reduction of the uncertainty) must be quantified in terms of performance through diagnostic forecast verification. Diagnostic verification is concerned with assessing different attributes of (ensemble) forecasts, such as reliability, skill, resolution, discrimination etc, to diagnose the performance of the forecast system and process so that cost-effective improvements may be made. The choice of measures may therefore sometimes influence the choice of modelling approach.

In some situations, a cost or damage function can also be constructed to quantify the effect of making better decisions, based on more accurate probabilistic forecasts. However, this is a developing area, so for practical applications a subjective estimate must often be made of the effect of a higher quality forecast on the decision-making process. In the end, the benefit of improved decision making should outweigh the investment in the forecast system.

Task 2.4 of this project may also consider real-time inundation mapping, in which case the question arises of whether a procedure is required to assimilate data at the hydrodynamic model nodes between telemetry sites (e.g. by interpolation) or whether a state updating procedure can be used instead which avoids the need for this.

Run times

Model run times will probably be a factor to consider in the detailed framework, since some approaches may not be feasible within the time constraints of providing operational forecasts in NFFS. The framework may need to indicate factors to consider, and different choices for different modelling situations, for example:

- Computational improvements – e.g. parallel processing, faster processors.
- Model configuration changes – e.g. nested models, model simplification/rationalisation.
- Statistical approaches – e.g. sampling or grouping of ensembles.
- Model emulators – e.g. simpler models to emulate the behaviour of more complex models.

Also, it might be possible for an initial (more computationally intensive) assessment be performed off-line, as required in some approaches. This would, of course, depend on suitable data (in terms of quality, record length, availability of hindcasts etc) being available to support the approach.

4.5.4 Uncertainty estimation techniques

Earlier sections in this report describe a wide range of approaches to estimation of uncertainty which have been used in research, pre-operational and operational flood forecasting studies. It is the intention that a selection of these techniques will be applied in the case studies to be performed in Phase 2 of the project (see Section 5), while the uncertainty framework might also consider a wider range of possible choices.

Table 4.6 provides an overview of the methods which have been discussed in this report using the following classification scheme:

- Forward uncertainty propagation – methods which consider the likely uncertainty from individual sources, and propagate that uncertainty through a cascade of models.
- Data assimilation techniques – methods which can help to improve the quality of a forecast, while also giving an estimate of uncertainty.
- Conditioning approaches – pre- and post-processing techniques which can help to improve and quantify the model outputs in probabilistic terms.

The table shows that some types of method (e.g. ensemble QPF) have been used more widely than others (although this particular technique is being considered on the ‘Hydrological Modelling with Convective Scale Rainfall’ project, so is outside the scope of the present study).

Table 4.6 Overview of uncertainty estimation techniques for flood forecasting applications described in this report.

Section	Item	Methods
Forward uncertainty propagation		
2.1.1	MAP D-PHASE	Multi-model ensemble QPF Multi-model hydrological models
2.1.1	EFFS	Multi-model and ensemble QPF
2.1.1	EU IMPRINTS	Multi-model and ensemble QPF
2.1.2	EFAS	Ensemble QPF
2.1.2	FEWS-NL	Multi-model/ensemble QPF
2.1.2	Lake Como	Monte Carlo rainfall estimation
2.1.2	CI-FLOW	Multi-model ensembles (QPF, hydrological, surge)
2.2.2	T46	Ensemble, and pseudo-ensemble, QPF
3.3.3	FRMRC2	Data based mechanistic emulators
4.2.1	General	Analytical approaches
4.2.1	ARPA-SIM	Multi-model ensemble (QPF, hydrological, hydrodynamic)
4.2.1	General	Monte Carlo, Latin hypercube, copula sampling
4.2.1	General	Fuzzy sets
4.5.3	General	Simpler techniques (e.g. sampling from a range)
Data assimilation		
2.1.1	FEWS-NL	Ensemble Kalman Filter (hydrodynamic model) Error correction Particle filtering
2.1.1	FLOODRELIEF	Ensemble Kalman Filter
2.1.1	FRMRC2	Adaptive DBM models
2.1.1	FREE	Adaptive DBM models at smart sensor nodes
2.1.1	FLOODsite	Machine learning techniques (ANN, M5 etc)
2.1.1	River Nith	Adaptive DBM models
2.1.1	FRMRC2	Adaptive DBM models
4.3.1	General	Error correction/prediction
4.3.1	General	State updating
4.3.1	General	Parameter updating
4.3.1	General	Kalman filtering Particle filtering
4.3.1	General	Variational methods
4.3.1	General	DBM models
Conditioning approaches (pre- and post-processing)		
2.1.1	NWS XEFS	Ensemble Pre-Processing (EPP)
2.1.1	NWS XEFS	Hydrological Model Output Statistics (HMOS)
2.1.1	NWS XEFS	Ensemble Post-Processing
2.1.1	FEWS-NL	Bayesian Model Averaging (BMA)
2.1.1	FEWS-NL	Bayesian Processor of Output (BPO)
4.3.2	General	Quantile regression
4.3.2	FRMRC	Generalised Likelihood Uncertainty Estimation (GLUE)
4.3.2	General	Bayesian Model Averaging
4.3.2	General	Bayesian Processor of Output

QPF = Quantitative Precipitation Forecast

Table 4.7 provides a shorter summary in which each particular technique is only listed once.

Table 4.7 Overall summary of uncertainty estimation techniques for flood forecasting applications described in this report.

General technique	Methods
Forward uncertainty propagation	Ensemble techniques Multi-model techniques Monte Carlo sampling Analytical approaches Latin hypercube Copula sampling Fuzzy sets Emulators Simpler techniques (e.g. sampling from a range)
Data assimilation	Error correction State updating Parameter updating Particle filtering Kalman Filter Ensemble Kalman Filter Adaptive Data Based Mechanistic (DBM) Machine learning techniques (ANN, M5 etc) Variational methods
Conditioning approaches	Hydrological Model Output Statistics (HMOS) Quantile regression Ensemble Pre-Processing (EPP) Ensemble Post-Processing Bayesian Model Averaging (BMA) Bayesian Processor of Output (BPO)

It is proposed that this general classification scheme is adopted for use in further development of the uncertainty framework.

4.6 Summary of chapter

The main points from this chapter include:

- A detailed review is provided for three main approaches to estimating and/or constraining uncertainty: forward uncertainty propagation, conditioning and data assimilation.
- The forward uncertainty propagation techniques which are discussed include multi-model techniques, Monte Carlo methods, and fuzzy set methods. The data assimilation and conditioning techniques which are discussed include updating of input variables, updating of model state variables, updating of model parameters and updating of model outputs (i.e. error correction).
- The issues with data assimilation and conditioning in integrated catchment models are also discussed, together with generic tools and techniques which could be used to implement them (such as DATools, GLUE and CAPTAIN).
- Particular data assimilation techniques which are described include Kalman filtering, particle filtering, variational methods and Data Based Mechanistic approaches, while conditioning approaches which are considered include quantile regression, Bayesian Model Averaging and Bayesian Processor of Output methods.
- The main data assimilation techniques currently used within the Environment Agency are also reviewed, and the review suggests that, at present, there is no existing data assimilation technique implemented in NFFS which allows for an assessment of uncertainty.
- The high-level features are introduced for the uncertainty framework to be developed further in Phase 2. These include the following key decision points in selection of an appropriate uncertainty estimation technique: level of risk, lead time requirements, types of models, sources of uncertainty, data assimilation techniques, operational requirements and model run times.

5. Phase 2 Recommendations

5.1 Introduction

This section describes proposals for Phase 2 of the project, for which the main tasks to be considered are:

- Task 2.1 – To further develop the high-level unified uncertainty framework for quantifying uncertainty for the major sources of uncertainty identified in Task 1.2 in integrated catchment models and suggest suitable validation measures.
- Task 2.2 – To recommend and investigate alternative ways of reducing run times for real-time probabilistic models without significantly increasing uncertainty.
- Task 2.3 – To demonstrate and validate the probabilistic treatment of uncertainties for selected integrated catchment models through case studies and test configurations in NFFS.
- Task 2.4 – To briefly test how uncertainties in flood forecasting may affect flood extent and depths and to make recommendations for future research on how to carry forward the outputs of probabilistic flood forecasting to the generation of probabilistic flood warning maps (extent, depths).

At the time of writing, the scope for Task 2.4 is still under consideration, and the time and budget available for that task may be transferred to other tasks within Phase 2 of the project. This task is therefore not discussed further here, although one of the case studies has been selected as a possible candidate for consideration on this task (based on current understanding of the requirements for this task).

The main outputs from Phase 2 will consist of the test configurations in NFFS and the Phase 2 report, which will describe the detailed uncertainty framework and approaches to reduce run times with their advantages and disadvantages, and will include the case study fact sheets, a description of the work performed on Task 2.4, and a draft structure for the guidance and implementation plan to be prepared in Phase 3 of the project.

This section describes the proposed approach for the remaining tasks in Phase 2 of the project. Section 5.2 discusses proposals for further development of the uncertainty framework, while Section 5.3 describes the approach to selection of possible case studies, and recommendations for potential case study catchments and associated techniques.

5.2 Further development of the uncertainty framework

During Phase 2 of the project, detail will be added to the framework in terms of specific techniques, types of models, ways of reducing run times etc. For example, Table 5.1 shows some initial ideas on possible key decision points and dependencies which could underpin the logical structure of the framework (although it is not proposed that this format is used in the actual framework). For example, under the entry ‘Level of Risk’, a link could be provided to the ‘Operational Requirement’ entry showing that post-processing (conditioning) of outputs is justified if the risk is high, but possibly not required if the risk is low. These inter-relationships between entries will be explored further during Phase 2 of the project, using the case studies as a starting point for the analysis.

Decisions will also be taken on the most operationally useful formats to use for the framework and guidelines, which could possibly include flowcharts, decision trees, risk matrices, and tables summarising the strengths and limitations of each approach. Additional ideas for the format of the framework and guidelines may also arise during work on Work Package 1.7 within FRMRC2 which is proceeding over the same timescales as this project (see Section 2.1).

The framework will also consider good practice issues such as the need to evaluate model performance after flood events (and recommended approaches), continuing model maintenance, and appropriate performance measures for assessing different aspects of model performance, such as skill scores, skill of the mean forecast, the spread-skill relationship of the probabilistic forecast, the Brier score, ranked probability score, vulnerability based measures, and contingency measures (POD, FAR etc).

However, the focus of the review of performance measures will be on a small number which could be used to assess the probabilistic outputs on the case studies (and which would be practical to implement operationally). As with the Real-Time Modelling Guidelines, the framework will provide clear guidance on the data requirements, benefits and limitations of each approach.

The benefits of new technologies being developed within ongoing research programmes (e.g. FRMRC), such as adaptive sensor networks, could also be discussed as a low cost route to obtaining real-time data at key locations (e.g. critical infrastructure) to assist with model evaluation and updating. Any approaches recommended would where feasible comply with general Environment Agency policy on Flood Incident Management/Forecasting (e.g. a risk based approach). For consistency, if possible, criteria for deciding on level of forecasting service will also be considered.

Table 5.1 Summary of possible decision points in the uncertainty framework.

LEVEL OF RISK	
Lead time requirements	Complexity of method used to estimate lead times
Types of models	
Sources of uncertainty	Whether a forward uncertainty analysis is justified
Data assimilation	The data quality/reliability required to support the approach The need for backup approaches in case an instrument fails
Operational requirement	The need for post-processing to assess probability distributions
Run times	The target run time specified for model runs
LEAD TIME REQUIREMENTS	
Level of risk	
Types of models	
Sources of uncertainty	The sources of uncertainty which are important to consider Possible requirements for ensemble rainfall forecasts
Data assimilation	Possible post-processing requirements (e.g. for ensemble QPF)
Operational requirement	The potential applications of the forecast
Run times	Number of upstream components in the integrated catchment model
TYPES OF MODELS	
Level of risk	Not applicable
Lead time requirements	Not applicable
Sources of uncertainty	The sources of uncertainty which are necessary to consider The uncertainty estimation techniques which can be used
Data assimilation	Possible post-processing requirements (e.g. for ensemble QPF)
Operational requirement	
Run times	Run times required for deterministic model
SOURCES OF UNCERTAINTY	
Level of risk	
Lead time requirements	
Types of models	Sources of uncertainty which may need to be considered Uncertainty estimation techniques which are practical to use
Data assimilation	Data assimilation techniques which are practical to use
Operational requirement	Possible post-processing requirements to improve probability estimates
Run times	Can depend on the number of sources considered
DATA ASSIMILATION	
Level of risk	
Lead time requirements	
Types of models	Influences the types of methods which can be used Decisions required on whether to update at model boundaries
Sources of uncertainty	
Operational requirement	
Run times	Can depend on the method selected e.g. EnKF, ARMA
OPERATIONAL REQUIREMENT	
Level of risk	
Lead time requirements	Directly influences the lead time requirement
Types of models	
Sources of uncertainty	Possible need for post-processing to derive estimates of probability Performance measures/skill scores used
Data assimilation	Whether estimates with and without data assimilation are required Possible requirement for real-time mapping
Run times	The time available in which to calculate and post-process forecasts
RUN TIMES	
Level of risk	
Lead time requirements	
Types of models	
Sources of uncertainty	Techniques which are feasible to use in real-time
Data assimilation	Techniques which are feasible to use in real-time
Operational requirement	The time available for pre-processing, model runs and post-processing

5.3 Selection of case studies and techniques

5.3.1 Introduction

The consultation exercise (see Section 2.2 and Appendix A) confirmed that the following choice of case studies would provide examples which would be of interest and relevant to all regions in the Environment Agency:

- Two complex linked integrated catchment models, combining rainfall–runoff, flow routing and/or hydrodynamic models: for example, based on PDM/KW/ISIS or MRCM/DODO/ISIS.
- Two simpler examples to illustrate key concepts: for example a rapid response catchment with rainfall and other sources of uncertainty, and a well-defined river reach including structures, rating curve uncertainty and other factors.

The intention would be that at least one of these case studies should use a data assimilation technique. There was also a general view that the case studies should be chosen on technical grounds (i.e. specific types of models, and forecasting situations), rather than other criteria such as risk, or current model performance regarding level of service.

The proposed high-level uncertainty framework (see Section 4.5) suggests the following key criteria for the choice of appropriate uncertainty estimation methods:

- level of risk
- lead time requirements
- types of models
- sources of uncertainty
- data assimilation
- operational requirement
- run times.

To assist in development of the framework, it would be useful if the case studies could illustrate a range of issues related to these requirements and this aspect was given although, as noted above, excluding the level of risk from the selection criteria. A number of other factors could also have been considered, such as the availability of previous research and model performance evaluation reports for the catchments, but it was considered that the criteria above were more relevant to this study.

The following sections describe how the recommended case studies were chosen to fit within this framework (Section 5.3.2) together with – for the remaining examples – some practical considerations related to the timescale and resources available to perform the case studies (Section 5.3.3). Finally, Section 5.3.4 describes the techniques which it is provisionally proposed to use on the chosen case studies during Phase 2 of the project.

The case studies would be conducted in such a way that the techniques/methods/NFFS workflows/codes developed on this project could be easily transferred to other cases, or reused by experienced personnel, once any considerations of prior rights for existing software etc (licences, permissions, royalties etc) have been resolved. All files/workflows/scripts etc for the case studies would be provided to the Environment Agency in a project record spreadsheet for the test configurations.

5.3.2 Illustrating use of the uncertainty framework

Lead time requirements

Section 4.5 discusses several possible categorisation schemes by lead time, and Table 5.2 illustrates how the suggested case studies could be classified using the following proposed approach:

- Type 1 – Flow routing model (mainly forecasting from an upstream gauging station).
- Type 2 – Integrated catchment model (combination of rainfall–runoff and flow routing models).
- Type 3 – Rapid response catchment (primarily rainfall–runoff modelling using observed rainfall).

Note that the Type 4 case, using forecast rainfall, is excluded from the case studies because it is being considered as part of the Hydrological Modelling with Convective Scale Rainfall project. Also, note that some of the Integrated Catchment Models also have rapid response and flow routing models included within them.

Table 5.2 Possibilities for the case studies from the consultation meetings.

Region	Type 1 Rapid Response Catchment	Type 2 Integrated Catchment Model	Type 3 Flow Routing Model(s)
Anglian		Cam	
Midlands		Upper Severn, Warwickshire Avon	Lower Trent, Soar
North East	Upper Calder	South Tyne, Wansbeck	Derwent
North West	Upper Ribble, Douglas, Greta	Derwent	Lower Eden, Lower Ribble
Southern		Medway, Sussex Ouse, Adur	
South West		Bude/Neet, Bristol Avon	
Thames		Ravensbourne	Thames Barrier
EA Wales		Tawe, Taff, Solva	

Types of models

The consultations also showed general agreement that the case studies should illustrate a range of model types and cascades of model types, applicable to a number of regions.

However, a Project Board requirement was also NOT to consider event-based models which cannot run in continuous mode, such as PRTF (although related hydrological or hydrodynamic routing reaches could still be considered).

Table 5.3 summarises the potential case studies which remain if these selection criteria are applied.

Table 5.3 Potential case studies taking account of types of models.

Region	Type 1 Rapid Response Catchment	Type 2 Integrated Catchment Model	Type 3 Flow Routing Model(s)
Anglian		Cam	
Midlands		Upper Severn Warwickshire Avon	Lower Trent Soar
North East	Upper Calder	South Tyne Wansbeck	Derwent
North West	Upper Ribble Douglas Greta	Derwent	Lower Eden Lower Ribble
Southern		Medway Sussex Ouse Adur	
South West			Bude/Neet Bristol Avon
Thames		Ravensbourne	Thames Barrier
EA Wales		Tawe Taff Solva	

Sources of uncertainty

The Project Board, and the consultation exercise, also suggested that the following sources of uncertainty are most important in the models which are currently operated by the Environment Agency:

i. Catchment averaging of raingauge data
ii. Validity of rating curves
iii. Model calibration (hydrodynamic models)
iv. Model calibration (rainfall–runoff models)
v. Representation of floodplain storage
vi. Representation of antecedent conditions
vii. Representation of ungauged inflows
viii. Influence of structure operations.

From the consultations, the top three entries in this list were as follows:

- Catchment averaging of raingauge data
- Validity of rating curves
- Model calibration (rainfall–runoff models)

With only a few catchment-specific exceptions, the issue of structure operations was generally towards the bottom of the list for most regions.

Also, the review of previous studies has not revealed any general methods for dealing with uncertainty across the range of types of structure used within the Environment Agency (which is not to say that such a method could not be developed).

It has therefore been decided to exclude consideration of this topic from the case studies.

The following potential case studies have been excluded on this basis:

- Medway – many structures and rating curve issues relating to data quality; not just rating curve uncertainty?
- Soar – significant ungauged inflows and complex structure operations; forecasting model currently under development.
- Solva – flood alleviation reservoir with a culverted outflow of a type which is not widely used within the Environment Agency except in small catchments.

The Thames Barrier example, although of great interest within the Environment Agency, is also a very specific forecasting issue which is unlikely to be of general application nationally.

The consultations also suggested that a study combining ensemble flow forecasts with ensemble surge forecasts would also be of interest, but this is outside the scope of the present project.

For the remaining sources of uncertainty, all of the integrated catchment models illustrate these issues to some extent; however, the following potential case study has been excluded on the basis that it includes complicating factors which may not be widely applicable to other catchments:

- Cam – catchment-specific chalk influences which are not necessarily captured using current rainfall–runoff modelling techniques.

For the rainfall–runoff models, again all are good possibilities, but the following model has been excluded due to the complicating factor of a VPMC flow routing model reach down to the key Flood Warning Area:

- Greta – flow routing reach in addition to the rainfall–runoff model component.

As discussed earlier, there is also a requirement to demonstrate the application of uncertainty estimation methods to a flow routing situation, where the influence of structures and floodplains etc might dominate, but the hydrological aspects are not so important. On these grounds, the following potential case studies have been excluded from the model selection, leaving the choice shown in Table 5.4:

- Bristol Avon – significant gauged and ungauged inflows, in addition to the routing issues.
- Derwent (North East) – proposed as a flow routing example but actually a full integrated catchment model.
- Lower Ribble – major gauged inflow from the River Darwen.
- Taff – reservoirs for which it is understood that, at present, there is no mechanism for updating levels in real time (although this is planned for the future).

Table 5.4 Potential case studies taking account of sources of uncertainty.

Region	Type 1 Rapid Response Catchment	Type 2 Integrated Catchment Model	Type 3 Flow Routing Model(s)
Anglian			
Midlands		Upper Severn Warwickshire Avon	Lower Trent
North East	Upper Calder	South Tyne Wansbeck	
North West	Douglas Upper Ribble	Derwent	Lower Eden
Southern		Sussex Ouse Adur	
South West			Bude/Neet
Thames		Ravensbourne	
EA Wales		Tawe	

Remaining issues

The remaining items in the high-level uncertainty framework are:

- data assimilation
- operational requirement
- run times.

All of the remaining possibilities are suitable to illustrate approaches to data assimilation, and it is understood that the data availability and reliability would be sufficient to support this. Similarly, there are no particular features regarding the operational requirement to select one case study rather than another.

For the integrated catchment models, the issue of run time is important and, to illustrate different potential approaches to assessment of uncertainty, it would be desirable to choose models which use a hydrodynamic component, due to the longer run time compared to a hydrological flow routing method alone (e.g. DODO, KW, VPMC). The Wansbeck catchment can therefore be excluded on these grounds, leaving the catchments shown in Table 5.5.

Table 5.5 Potential case studies taking account of remaining issues.

Region	Type 1 Rapid Response Catchment	Type 2 Integrated Catchment Model	Type 3 Flow Routing Model(s)
Anglian			
Midlands		Upper Severn Warwickshire Avon	Lower Trent
North East	Upper Calder	South Tyne	
North West	Douglas Upper Ribble	Derwent	Lower Eden
Southern		Sussex Ouse Adur	
South West			Bude/Neet
Thames		Ravensbourne	
EA Wales		Tawe	

5.3.3 Practical considerations

A number of practical considerations also need to be taken into account in the choice of case studies.

Firstly, it was noted during the consultations that any proposed models should already be configured onto NFFS so that it was not necessary to spend time doing this during Phase 2 of the project. At present, only two of the models shown in Table 5.5 – the Upper Ribble and the Ravensbourne – are currently not configured although it is understood that configuration of the Ravensbourne is almost completed. This case study has therefore been retained, while the Upper Ribble has been excluded.

Another requirement was that the case studies should be selected so that the experience gained will be of benefit to most, if not all, other regions. For the remaining potential case studies, two regions – Midlands and Thames – use rainfall–runoff model types which are unique to the region, but this was felt to be too broad a criterion to exclude these examples (since much can still be learned about catchment averaging and rating curve issues, and approaches to parameter sampling).

To share knowledge between regions, it would also be desirable to choose at most one category of model per region and, for the case studies which involve hydrodynamic models, to maximise the use of existing expertise in the project team, it would also be desirable to choose models that we have worked on recently, or are currently working on. The following catchments meet this criterion:

- Adur – model development (Atkins)
- Lower Eden – model development (Edenvale Young, Atkins), FREE (CEH), FRMRC (Lancaster)
- Lower Ribble – model development (Atkins)
- Lower Trent – model development (Edenvale Young)
- Ravensbourne – model development ongoing (Edenvale Young)
- Severn – model development (Edenvale Young), FRMRC studies (Lancaster)
- Sussex Ouse – model development (Atkins, Edenvale Young)
- Upper Calder – model development (CEH)
- Warwickshire Avon – T46 case study catchment (CEH).

If this criterion is applied, then Table 5.6 summarises the potential case studies which would remain.

Table 5.6 Potential case studies taking account of project team catchment knowledge.

Region	Type 1 Rapid Response Catchment	Type 2 Integrated Catchment Model	Type 3 Flow Routing Model(s)
Anglian			
Midlands		Upper Severn Warwickshire Avon	Lower Trent
North East	Upper Calder		
North West	Douglas		Lower Eden
Southern		Sussex Ouse Adur	
South West			
Thames		Ravensbourne	
EA Wales			

A further consideration is that – to share expertise among Environment Agency regions – it would be desirable to select only one general category of model (Rapid Response etc) from the remaining selection. Considering the Severn and Warwickshire Avon, the Severn has perhaps been studied more widely, within FRMRC and elsewhere, and would be the preferred choice. For the Sussex Ouse and Adur, the Adur is perhaps of less general interest due to the extensive tidal floodplain in the lower reaches, with the main flood risk areas in the upper and middle reaches of the catchment. Excluding these catchments would then leave the choice shown in Table 5.7.

Table 5.7 Potential case studies with only one general type per region.

Region	Type 1 Rapid Response Catchment	Type 2 Integrated Catchment Model	Type 3 Flow Routing Model(s)
Anglian			
Midlands		Upper Severn	Lower Trent
North East	Upper Calder		
North West	Douglas		Lower Eden
Southern		Sussex Ouse	
South West			
Thames		Ravensbourne	
EA Wales			

Again, applying the criterion that there should only be one general type of model per region, then the following considerations apply:

- Midlands Region – the Severn is perhaps of more general interest to other regions, compared to the relatively simple example of the Lower Trent.
- North West – the Lower Eden forecasting model has been extensively reviewed since the January 2005 floods in Carlisle, and is therefore perhaps of more interest than the Douglas.

If these criteria are applied, then Table 5.8 shows the remaining potential case studies.

Table 5.8 Potential case studies selecting only one potential case study per region.

Region	Type 1 Rapid Response Catchment	Type 2 Integrated Catchment Model	Type 3 Flow Routing Model(s)
Anglian			
Midlands		Upper Severn	
North East	Upper Calder		
North West			Lower Eden
Southern		Sussex Ouse	
South West			
Thames		Ravensbourne	
EA Wales			

Table 5.9 shows the key factors which would then be considered with this selection of case studies.

Table 5.9 Summary of catchments in Table 5.8

Region	Region	Model type	Category	Existing data assimilation
Upper Severn	Midlands	MCRM/DODO/ISIS	Integrated Catchment Model	ARMA
Upper Calder	North East	PDM	Rapid Response Catchment	ARMA/state updating
Lower Eden	North West	VPMC/ISIS	Flow Routing	ARMA
Sussex Ouse	Southern	PDM/ISIS	Integrated Catchment Model	ARMA
Ravensbourne	Thames	TCM/ISIS	Integrated Catchment Model	ARMA

This selection is close to the requirement for four catchments and spans five regions, three types of rainfall–runoff model, and three types of flow routing/hydrodynamic model. However, there is only scope to consider two integrated catchment models in Phase 2 of the project, so the choice needs to be narrowed down further (and this will be discussed with the Project Board).

The following comments (paraphrased) were made on these catchments during the consultation meetings:

- Lower Eden – the forecast relies primarily on flow routing. Ungauged inflows are a significant factor between Temple Sowerby and Great Corby. Note that PRTF models provide the hydrological inputs, and the lower boundary is not tidal. Floodplains are generally accounted for in the ISIS model build, although with some questions on the most appropriate cell representation to use (reservoir units or secondary channels). The region has also trialled the flood risk mapping module in NFFS, for the reach from Great Corby to Sheepmount on the Eden.

- Ravensbourne – this is the first hydrodynamic model in the region for a rapidly responding urban catchment with a distributed set of hydrological inputs. It is a TCM-ISIS model with radar rainfall inputs, and approximately 30–40% of the catchment area is gauged. It is not, however, certain that it will be fully configured on NFFS by the time that the case studies will be performed.

- Upper Severn – uncertainties arise regarding the extent to which flows go out of bank, and how to represent this storage in the model (e.g. in the Lower Severn between Tewkesbury and Gloucester, and for the Upper Severn around Shrewsbury in large flood events).
- Sussex Ouse – no particular comments noted.
- Upper Calder – could also provide examples of rapid response catchments (Todmorden, Walsden). Rainfall forecasts are required (ideally) to obtain sufficient lead time on some catchments (e.g. on the Upper Calder where catchment response times may only be 1.5 to 2 hours).

This process has therefore provided a possible basis for selection of the case studies. The next step is to consider how best to map the techniques that need to be tested to these case studies and this is discussed in the following section.

Also, given the timescale and resources available in Phase 2 of the project, it may be necessary to identify subcatchments within the larger models which will still be a valid test bed for the techniques to be tested, and this point is also discussed below.

5.3.4 Selection of techniques

Proposed techniques

The review studies described in Sections 2, 3 and 4 of this report have identified the range of techniques shown in Table 4.7. Table 5.10 summarises an assessment of the suitability of these techniques for use on this project.

Based on the review, and the experience in the project team, the items shown in bold font are seen as practicable to consider within Phase 2 of the project, for which the requirement is to implement draft NFFS test configurations for the case studies by November 2009, with final versions by March 2010.

Table 5.10 Summary of techniques which will be considered during the case studies (see text for description of bold font entries).

General technique	Method	Comment
Forward uncertainty propagation	**Ensemble techniques**	Ensemble nowcasts (STEPS) will be evaluated for two events
	Multi-model techniques	Not selected in the usual sense of considering multiple types of models at a site (which are generally not available in the Environment Agency), but multiple parameter sets will be considered for the same model (see simpler techniques)
	Monte Carlo sampling	Well-established technique and proposed for off-line use (although considered too computationally intensive at present for real-time implementation in integrated catchment models in NFFS)
	Analytical approaches	Not suitable for chains of nonlinear models and/or real-time evaluation of uncertainty
	Latin hypercube	Can be used in multivariate sampling where distributions and covariation of uncertain input variables are known; more efficient than Monte Carlo sampling but with some approximations
	Copula sampling	Can be used when marginal distributions of uncertain input variables are known but where interactions between variable are complex; more complex than Monte Carlo sampling
	Fuzzy sets	Can be used where uncertain input variables are poorly specified, for example where only a range is easily specified; technique under development for flood forecasting applications
	Emulators	Technique that holds great promise for forecasting, uncertainty emulation and other applications
	Simpler techniques	Off-line Monte Carlo sampling and sampling from a range will be explored as examples of quick and simple techniques to implement (see Table 5.11), and have been proven in operational studies on the Rhine FEWS-NL system, for example
Data assimilation	Error correction	Current Environment Agency approach
	State updating	Current Environment Agency approach
	Parameter updating	Technique generally not appropriate for conceptual and physically based models
	Particle filtering	Some advantages over Ensemble KF techniques but also some operational disadvantages (robustness, computational)
	Kalman Filter	Long-established technique with many applications
	Ensemble Kalman Filter	Technique established for a few years and trialled in several practical flood forecasting applications
	Adaptive Data Based Mechanistic	Models will be developed and implemented into NFFS for the Lower Eden and Upper Severn case studies, building on previous FRMRC and other work
	Machine learning techniques	A range of new techniques under development and demonstrated in research studies
	Variational methods	More appropriate for meteorological and coastal forecasting applications (2D/3D)
Conditioning approaches	Hydrological Model Output Statistics	Method being evaluated by the National Weather Service (NWS) in the USA which focuses on short range forecasts but is still under development
	Quantile regression	Simple method which can easily be tested to determine the potential in flood forecasting and is being used/evaluated in some applications
	Ensemble Pre-Processing	Method under development by the NWS and others
	Ensemble Post-Processing	Method used for many years to remove biases in long-term flow volume forecasts
	Bayesian Model Averaging	Experimental method that has already proved successful for both coastal and river applications and will soon be available within Delft-FEWS
	Bayesian Processor of Output	Experimental method and still many details unclear besides the large volume of historical data that is needed (although near operational studies on the Rhine showed potential of the method)

The entry for simpler forward uncertainty propagation techniques refers to the methods which are summarised in Section 4.7 of this report. Table 5.11 summarises these methods for the top three sources of uncertainty identified during the consultation meetings.

Table 5.11 Some potential simpler forward uncertainty propagation techniques for the top three sources of uncertainty identified in the consultation meetings.

Source of uncertainty	Potential methods
Catchment averaging of raingauge data	Off-line Monte Carlo sampling of catchment average rainfall, raingauge weights, or radar rainfall fields, assume a typical range
Validity of rating curves	Decomposition of error sources, addition of variances (ISO/British Standard procedures), off-line Monte Carlo sampling, evaluation of historical performance, assume a typical range
Model calibration (rainfall–runoff models)	Off-line Monte Carlo sampling of key model parameters (e.g. GLUE) and/or model initial conditions, multi-model comparisons, evaluation of historical performance, assume a typical range

Note that the word 'simpler' in this context is also relative, since implementation of some of these approaches could be a significant task.

The methods used in Phase 2 would be selected from this list, possibly with some equivalent methods substituted if initial exploratory modelling shows that to be more beneficial to the project. Perhaps the simplest technique of all is to specify a likely range for the parameter or variable of interest (e.g. ±10%), and to choose a few values which are representative of that range (e.g. max, mean, min). For the more complicated Monte Carlo based approaches, it is proposed that any simulations are done off-line, and that the results are implemented in the case studies using the following existing features of NFFS:

- Multiple data inputs – NFFS allows more than one data input (ensemble) to be evaluated per model at each time-step (as in the Hydrological Modelling with Convective Scale Rainfall project, for example). The Monte Carlo (or other) off-line simulations would be used to guide the selection of an appropriate number (ensemble) of rainfall inputs to represent the full distribution of catchment rainfall uncertainty and interpolation schemes.

- Multiple rating curves – NFFS allows more than one rating curve to be defined at a gauging station, and for a forecast to be derived for each curve at each model time-step, providing an ensemble of forecasts. The Monte Carlo (or other) off-line simulations would be used to guide the choice of an appropriate number of curves to represent the full distribution of rating curve uncertainty.

- Multiple parameter sets – NFFS allows more than one set of parameters per model to be evaluated at each time-step, allowing an ensemble of forecasts to be generated. The Monte Carlo (or other) off-line simulations would be used to guide the selection of the key parameters to consider, and their likely distributions. For simplicity, only a single key parameter would be considered at a time (rather than multiple parameters, and parameter interdependence).

For example, recent work by Deltares on the operational FEWS-NL model suggests that – for the Meuse catchment – as few as 5–6 ensemble members may be sufficient to capture the key parameter uncertainty.

For the simpler forward uncertainty propagation methods, this approach would avoid the need for further development in NFFS, allowing time and resources to be spent on evaluating the more advanced techniques instead. Deltares could also make the UATools uncertainty analysis software available to the project team and the Environment Agency to assist with the off-line studies.

A Bayesian Model Averaging approach might also be evaluated for one example to illustrate how this approach can be used to derive a single 'best' forecast from each of a set of ensemble members, although BMA techniques are more usually applied to multiple structures/types of models, rather than multiple sets of parameters with a single or multiple inputs for the same model.

Level of risk

In the high-level uncertainty framework, it is proposed that two levels of detail might be used linked to the level of risk:

- Method A – a simple approach suitable for a rapid first assessment of potential modelling solutions.
- Method B – the main model selection approach, which aims to arrive at a reasonable compromise between technical, cost, benefit and other considerations.

Table 5.12 presents a first attempt at classifying the methods selected in the previous section, based on a combination of the complexity of the method, and the past experience of the project team in implementing the approach.

Table 5.12 Classification of the selected techniques by level of risk.

Technique	Method A	Method B
Forward uncertainty propagation	Assume a typical range	Off-line Monte Carlo
Data assimilation	Kalman Filter/DBM	Ensemble Kalman Filter
Conditioning approaches	Quantile regression	Bayesian Model Averaging

Model run times

Model run times will also be a factor to consider during the case studies and, in the high-level uncertainty framework, the following four approaches are proposed for consideration:

- Computational improvements – e.g. parallel processing, faster processors.
- Model configuration changes – e.g. nested models, model simplification/rationalisation.
- Statistical approaches – e.g. sampling or grouping of ensembles.
- Model emulators – e.g. simpler models to emulate the behaviour of more complex models.

The issue of computational improvements and statistical approaches will be considered mainly through review, and possibly some exploratory testing/benchmarking, unrelated to the case studies.

It is proposed that the remaining two approaches – model reconfiguration and emulators – will also be illustrated in the case studies, using the following approaches:

- Model configuration – one case study showing the benefits of reconfiguration of a hydrodynamic model, based mainly on a previous study by the project team.
- Emulators – Data Based Mechanistic approach applied to two catchments.

Matching techniques to case studies

To maximise benefits from the project, the following process is proposed for the development of the case studies and uncertainty framework.

- Step 1 – trial simple and complex (Method A and B) forward uncertainty propagation methods for rainfall and rainfall–runoff model uncertainty on the Type 1 – Rapid Response case study.
- Step 2 – trial simple and complex (Method A and B) forward uncertainty propagation methods for rating curve uncertainty on the Type 3 – Flow Routing case study.
- Step 3 – trial simple and complex (Method A and B) conditioning approaches on the Type 1 – Rapid Response case study.
- Step 4 – trial simple and complex (Method A and B) data assimilation approaches on the Type 3 – Flow Routing case study.
- Step 5 – trial emulation and model reconfiguration approaches on the Type 3 – Flow Routing case study and compare approaches and run time reductions with Step 4.

This approach should help to reduce the project risk by testing techniques on the simpler models first, and in a phased manner (i.e. not applying all techniques at the same time).

At the end of this stage, all key methods and techniques would have been used on one or other of the two simple case studies. The uncertainty framework would also be developed further to include the lessons learned from these studies.

At this point we would propose to finalise the scope for the more complex integrated catchment model case studies based on what has been learned in the earlier case studies and before work starts on the more complex catchments. Figure 5.1 illustrates this process.

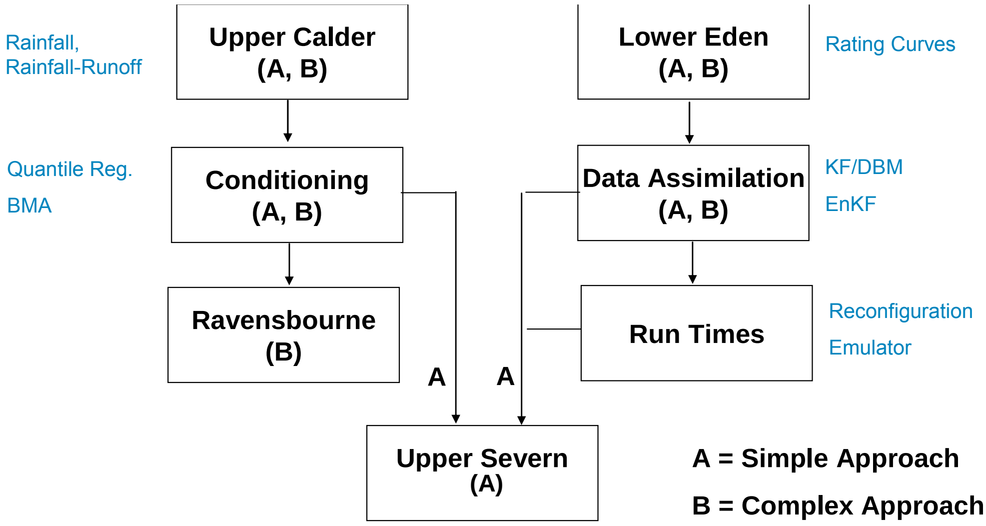


Figure 5.1 Illustration of the sequence in which the case studies will be performed.

The Ravensbourne will be a test bed for using the more complex hydrologically related techniques for an integrated catchment model, while the second integrated catchment model (the Upper Severn) will be used to combine both hydrological and flow routing related techniques; however, to keep the scope of this case study manageable, it will only consider simple techniques.

In this work, the focus of effort would be as follows for the two integrated catchment models to be considered:

- Type 2 – Integrated Catchment Model – Model 1 – focus on the hydrological aspects for forward uncertainty propagation, and conditioning of model outputs.
- Type 2 – Integrated Catchment Model – Model 2 – focus on the rating curve aspects for forward uncertainty propagation, data assimilation, model reconfiguration and emulation.

The Model 1 example would therefore be a step towards the more complicated Model 2 example, which would be the end point for the case studies. The lessons learned would be documented in factsheets outlining the approach, outputs, benefits and limitations. It may also be possible to draw preliminary conclusions from these studies about the performance of models in typical forecasting situations. The uncertainty framework might also be refined following completion of these studies.

Regarding the proposed choice of case studies in the previous section, the following catchment-specific issues are worth noting:

- Type 1 – Upper Calder – some limited rainfall forecast ensembles for 2 days in 2007 may also be available through the Hydrological Modelling with Convective Scale Rainfall project (although it is presently outside the scope of this project to consider these).
- Type 2 – Model 1 – Ravensbourne – the main issues relate to rainfall rainfall–runoff modelling and ungauged inflows.
- Type 2 – Model 2 – Upper Severn **OR** Sussex Ouse – the example chosen would represent a range of modelling uncertainty issues. For the Sussex Ouse, it is likely that it would not be feasible to consider the whole catchment, but just a major tributary – the Uck – which was one of four pilot studies nationally for NFFS. For the Severn, the issue of floodplain storage and backwater effects is an interesting dimension to the problem.
- Type 3 – Lower Eden – an interesting aspect of this case study would be to compare the model uncertainty (for the current operational model) with the ratings in operation before the January 2005 flood, and following the improvements made to high flow ratings after that event based on the new data recorded. Also, to examine the gains from data assimilation at the lower end of the model for two forms of the model; starting upstream of and immediately downstream of the major Eamont tributary.

For the Lower Eden, it would also be possible to document how reconfiguration of the model over the years has improved the run-time performance as suggested earlier under 'Model run times' since two members of the Project Team have previously worked on this aspect of the model (Atkins, Edenvale Young). This case study would also be an option for Task 2.4 (floodplain mapping) since it is one of the few examples for which this option has been implemented on NFFS so far.

The following figure (Figure 5.2) summarises these proposals for the case studies.

	Forward propagation						Data assimilation		Post-processing		Emulation	Reconfiguration
	Rainfall		RR model parameters		Rating							
	Simple	Complex	Simple	Complex	Simple	Complex	Simple	Complex	Simple	Complex	Simple	
Preparatory analysis (off-line)	Sampling from a range	Monte Carlo sampling	Sampling from a range	Monte Carlo sampling	Sampling from a range	Monte Carlo Sampling	N.A.	N.A.	Hindcast analysis	N.A.	Model identification	
Real-time application	Ensemble of rainfall inputs	Ensemble of rainfall inputs	Ensemble of model parameters	Ensemble of model parameters	Ensemble of rating curves	Ensemble of rating curves	Kalman Filter or Data Based Mechanistic	Ensemble Karman Filter (En KF)	Quantile regression	Bayesian Model Averaging (BMA)	Data Based Mechanistic	
Upper Calder	1	1	1	1					1	1		
Lower Eden					1	1	1	1			1	1
Ravensbourne		1		1						1		
Upper Severn or Sussex Ouse	2		2		2		1		1		2	

Key

(grey shading)	Minor contribution to model uncertainty
(diagonal line)	Project team chosen not to test technique (to keep scope of study manageable)
1	Assessment of these techniques will be the focus of case study
2	Case study will provide supporting information on the assessment of this technique

Figure 5.2 Summary of proposals.

6. References

ARONICA, G., HANKIN, B.G. AND BEVEN, K.J., 1998. Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. *Advances in Water Resources*, 22(4), 349–365.

BATES, P.D., HORRITT, M.S., ARONICA, G. AND BEVEN, K.J., 2004. Bayesian updating of flood inundation likelihoods conditioned on flood extent data. *Hydrological Processes*, 18, 3347–3370.

BECKERS, J.V.L., SPROKKEREEF, E. AND ROSCOE, K., 2008. Use of Bayesian model averaging to determine uncertainties in river discharges and water level. *4th International Symposium on Flood Defence: Managing Flood Risk, Reliability and Vulnerability*, Toronto, Ontario, Canada, 6–8 May 2008.

BELL, V.A. AND MOORE, R.J., 1999. An elevation dependent snowmelt model for upland Britain. *Hydrological Processes*, 13, 1887–1903.

BELL, V.A. AND MOORE, R.J., 2000. *Development and Assessment of Snowmelt Models for Flood Forecasting and Warning.* R&D Technical Report W227, Environment Agency, 60pp.

BELL, V.A., MOORE, R.J. AND BROWN, V., 2000. Snowmelt forecasting for flood warning in Upland Britain. In: M. Lees and P. Walsh (eds), *Flood Forecasting: What Does Current Research Offer the Practitioner?* pp. 11–25. Occasional Paper No. 12, British Hydrological Society.

BELL, V.A., CARRINGTON, D.S. AND MOORE, R.J., 2001. *Comparison of Rainfall–Runoff Models for Flood Forecasting. Part 2: Calibration and Evaluation of Models*. R&D Technical Report W242, Environment Agency, 239pp.

BEVEN, K.J., 2001. *Rainfall–Runoff Modelling: The Primer.* Wiley, Chichester.

BEVEN, K.J., 2002. Towards a coherent philosophy for environmental modelling. *Proceedings Royal Society of London*, 458, 2465–2484.

BEVEN, K.J., 2005. On the concept of model structural error. *Water Science and Technology,* 52(6), 165–175.

BEVEN K.J., 2006. A manifesto for the equifinality thesis. J.Hydrology, 320, 18-36

BEVEN, K.J., 2009*. Environmental Modelling: An Uncertain Future?* Routledge, London (see http://www.uncertain-future.org.uk/).

BEVEN, K.J. AND BINLEY, A.M., 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes,* 6, 279–298.

BEVEN, K.J. AND FREER, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems. *Journal of Hydrology*, 249, 11–29.

BEVEN, K. J., ROMANOWICZ, R., PAPPENBERGER, F., YOUNG, P.C. AND WERNER, M., 2005. The uncertainty cascade in flood forecasting. In: P. Balbanis, D. Lambroso and P. Samuels (eds), *Innovation, Advances and Implementation of Flood Forecasting Technology, Proceedings of the ACTIF Meeting*, Tromsø, abstract, p. 33, paper 9p on CD (ISBN 978–1-898485-12-4).

BEVEN, K.J., YOUNG, P., ROMANOWICZ, R., O'CONNELL, E., EWEN, J., O'DONNELL, G., HOMAN, I., POSTHUMUS, H., MORRIS, J., HOLLIS, J., ROSE, S., LAMB, R. AND ARCHER, D., 2008a. *Analysis of Historical Data Sets to Look for Impacts of Land Use and Management Change on Flood Generation*. FD2120 Project Report, Defra, London.

BEVEN, K.J., YOUNG, P.C., LEEDAL, D.T. AND ROMANOWICZ, R., 2008b. Computationally efficient flood water level prediction (with uncertainty). In: *Proceedings, European Conference on Flood Risk Management*, Oxford.

BEVEN, K J, SMITH, P J, AND FREER, J, 2008c, So just why would a modeller choose to be incoherent?. J. Hydrology, 354,15-32.

BOX, G.E.P. AND JENKINS, G.M., 1970. *Time series analysis forecasting and control*. Holden-Day, 553pp.

BROERSEN, P.M.T., 2007. Error correction of rainfall–runoff models with the ARMAsel program. *IEEE Transactions on Instrumentation and Measurement*, 56(6), 2212–2219.

BROERSEN, P.M.T, AND WEERTS, A.H., 2005. Automatic error correction of rainfall–runoff models in flood forecasting systems. In: *Proceedings IEEE/IMTC Conf*erence, Ottawa, Canada, 1531–1536.

BURGERS G, VAN LEEUWEN P J, EVENSEN G, 1998, Analysis Scheme in the Ensemble Kalman Filter. Monthly Weather Review 126, 1719-1724

BÜRGI, T 2002, Operational flood forecasting in mountainous areas - an interdisciplinary challenge. In International Conference in Flood Estimation. Spreafico, M. and Weingartner, R. (eds.). CHR Report II-17. Bern, CH, pp. 397–406

BUTTS, M.B., FALK, A.K., HARTNACK, J., MADSEN, H., KLINTING, A., VAN KALKEN, T., CADMAN, D. AND PRICE, D., 2005. Ensemble based methods for data assimilation and uncertainty estimation in the FLOODRELIEF project. *ACTIF International Conference on Innovation Advances and Implementation of Flood Forecasting Technology*, Tromso, Norway, 17–19 October 2005.

CEH WALLINGFORD, 2004. *CEH XML Module Adapter Guide to Configuration and Use, Version 1.0.006*. Centre for Ecology & Hydrology, Wallingford, 111pp.

CEH WALLINGFORD, 2005a. *PDM Rainfall–Runoff Model. Version 2.2, September 2005*. (Incorporates: Practical User Guide, Guide, Practical User Guide to the Calibration Shell, User Manual, Training Exercises).

CEH WALLINGFORD, 2005b. *PSM Rainfall–Runoff Model. (Penman Store Model Incorporating the TCM and the IEM.) Version 2.2, September 2005*. (Incorporates: Practical User Guide, Guide, Practical User Guide to the Calibration Shell, User Manual, Training Exercises).

CEH WALLINGFORD, 2005c. *KW Channel Flow Routing Model. Version 2.2.* (Incorporates: Practical User Guide, Guide, Practical User Guide to the Calibration Shell, User Manual, Training Exercises). Centre for Ecology & Hydrology, Wallingford.

CEH WALLINGFORD, 2005d. *Interpolation for ARMA Models*. Centre for Ecology & Hydrology, Wallingford, 2pp.

CEH WALLINGFORD, 2009. *A Quick Guide to Running the PDM in Ensemble Mode. Version 1.1*, Centre for Ecology & Hydrology, Wallingford, 7pp.

COLE, S.J. AND MOORE, R.J., 2008. Hydrological modelling using raingauge- and radar-based estimators of areal rainfall. *Journal of Hydrology*, 358, 159–181.

COLE, S.J. AND MOORE, R.J., 2009a. Distributed hydrological modelling using weather radar in gauged and ungauged basins. *Advances in Water Resources, 32, 1107-1120*.

COLE, S.J. AND MOORE, R.J., 2009b. *Hydrological Modelling for the Rivers Lavant and Ems: Interim Report*. Centre for Ecology & Hydrology, February 2009, 50pp.

COLE, S.J., ROBSON, A.J., BELL, V.A. AND MOORE, R.J., 2009. Model initialisation, data assimilation and probabilistic flood forecasting for distributed hydrological models. *Geophysical Research Abstracts,* 11, EGU2009-8048-3.

DAY, G.N., 1985. Extended streamflow forecasting using NWSRFS. *Journal of Water Resources Planning and Management*, *ASCE*, 111(2), 157–170.

DE ROO, A.P.J. et al, 2003. Development of a European flood forecasting system. *International Journal of River Basin Management*, 1(1), 49–59.

DOUGLAS, J.R. AND DOBSON, C., 1987. Real-time flood forecasting in diverse drainage basins. In: V.K. Collinge and C. Kirby (eds), *Weather Radar and Flood Forecasting*, Chapter 11, pp. 153–169, J. Wiley & Sons Ltd.

EL SERAFY G Y AND MYNETT A E, 2004, Comparison of EKF and EnKF in Sobek River: Case study Maxau-Ijssel. In: Liong S-Y, Phoon K-K and Baovic V. (Eds.) Proc. 6th International Conference on HydroInformatics, World Scientific, pp. 513-520

EL SERAFY G Y, GERRITSEN H, HUMMEL S, WEERTS A H, MYNETT A E, TANAKA M (2007) Application of data assimilation in portable operationalforecasting systems—the DATools assimilation environment. Ocean Dynamics.

ENVIRONMENT AGENCY, 2002 Real Time Modelling Guidelines. R&D project WSC013/5.

ENVIRONMENT AGENCY, 2005. *Performance Measures for Flood Forecasting: Review and Recommendations*. R&D Technical Report W5C-021/2b/TR.

ENVIRONMENT AGENCY, 2006. *Flood Warning Levels of Service*. AMS Work Instruction 137_05, Version 2, May 2006.

ENVIRONMENT AGENCY, 2007. *Probabilistic Flood Forecasting Scoping Study*. R&D Technical Report FD2901/TR.

ENVIRONMENT AGENCY, 2009. *Flood Warning Levels of Service*. AMS Work Instruction 137_05, Version 4, April 2009.

EVENSEN, G, 1994, Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res., 97(17), 905-924.

FAULKNER, H., PARKER, D., GREEN, C. AND BEVEN, K., 2007. Developing a translational discourse to communicate uncertainty in flood risk between science and the practitioner. *Ambio*, 16(7), 692–703.

GELB, A. (ed.), 1974. *Applied Optimal Estimation*. MIT Press, Cambridge, MA, 374pp.

GILL, P.E., MURRAY, W. AND WRIGHT, M.H., 1981. *Practical Optimisation*. Academic Press, London, 401pp.

GEORGAKAKOS K P (1986a, b) A Generalized Stochastic Hydrometeorological Model for Flood and Flash-Flood Forecasting, 1 Formulation and 2 Case Studies," Water Resources Research, 22(13), 2083-2106

HALL, J.W. AND SOLAMATINE, D., 2008. A framework for uncertainty analysis in flood risk management decisions. *Journal River Basin Management*, 6(2), 85–98.

HALQUIST J (2006) Use of HPC QPF Confidence Interval Forecasts to producxe a Hydrologic Ensemble of River Forecasts. American Meteorological Society 20th Conference on Hydrology, 28 January-2 February 2006

HAN, D., 1991. *Weather Radar Information Processing and Real-Time Flood Forecasting*. PhD Thesis, Water Resources Research Group, Department of Civil Engineering, University of Salford, 277pp.

HASHINO, T., BRADLEY, A.A. AND SCHWARTZ, S.S., 2007. Evaluation of bias-correction methods for ensemble streamflow volume forecasts. *Hydrology and Earth Systems Sciences*, 11, 939–950.

HOUGH, M., 2003. *An Historical Comparison Between the Met Office Surface Exchange Scheme-Probability Distributed Model (MOSES-PDM) and the Met Office Rainfall and Evaporation Calculation System (MORECS).* Final Report, Met Office.

JANSSEN, P.H.M., PETERSEN, A.C., VAN DER SLUIJS, J.P., RISBEY, J.S. AND RAVETZ, J.R., 2003. *RIVM/MNP Guidance for Uncertainty Assessment and Communication: Quickscan Hints and Action List, RIVM/MNP.* ISBN 90-6960-105-2.

JANSSEN, P.H.M., PETERSEN, A.C., VAN DER SLUIJS, J.P., RISBEY, J.S. AND RAVETZ, J.R., 2005. A guidance for assessing and communicating uncertainties. *Water Science and Technology*, 52, 125–131.

JAZWINSKI, A.H., 1970. *Stochastic Processes and Filtering Theory.* Academic Press, New York, 376pp.

JONES, A.E., JONES, D.A. AND MOORE, R.J., 2003. *Development of Rainfall Forecast Performance Monitoring Criteria. Phase 1: Development of Methodology and Algorithms*. Report to the Environment Agency and the Met Office, CEH Wallingford, 291pp.

JONES, A.E., JONES, D.A., MOORE, R.J. AND ROBSON, A.J., 2004. *Heavy Rainfall Warning Assessment Tool User Guide*. Report to the Environment Agency and the Met Office, CEH Wallingford, 65pp.

JONES, D.A. AND MOORE, R.J., 1980. A simple channel flow-routing model for real-time use. Hydrological forecasting, Proc. Oxford Symposium, April 1980, *IAHR-AISH Publ. No.129*, 397-408.

KACHROO, R.K., 1992. River flow forecasting. Part 1. A discussion of principles. *Journal of Hydrology*, 133, 1–15.

KALMAN, R, 1960. New approach to linear filtering and prediction problems, J. Basic Eng. 82D: 35-45.

KITANIDIS, P K AND BRAS, R L (1980), Adaptive Filtering through Detection of Isolated Transient Errors in Rainfall-Runoff Models, Water Resources Research, 16(4), 740-748

KLIR, G, 2006. Uncertainty and Information, Wiley: Chichester

KLIR, G AND FOLGER, T, 1988. Fuzzy Sets, Uncertainty and Information, Prentice Hall: Englewood Cliffs, NJ.

KOENKER, R AND BASSETT G, 1978. Regression quantiles. Econometrica, 46(1), 33-50.

KRZYSZTOFOWICZ, R., 1999. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resources Research*, 35(9), 2739–2750.

KRZYSZTOFOWICZ, R., 2001. The case for probabilistic forecasting in hydrology. *Journal of Hydrology,* 249, 2–9.

KRZYSZTOFOWICZ, R. AND KELLY, K.S., 2000. Hydrologic uncertainty processor for probabilistic river stage forecasting. *Water Resources Research*, 36(11), 3265–3277.

KUCZERA, G., KAVETSKI, D., FRANKS, S. AND THYER, M., 2006. Towards a Bayesian total error analysis of conceptual rainfall–runoff models: characterising model error using storm dependent parameters. *Journal of Hydrology*, 331, 161–177.

LEEDAL, D.T., BEVEN, K.J., YOUNG, P.C. AND ROMANOWICZ, R., 2008. Data assimilation and adaptive real-time forecasting of water levels in the river Eden catchment, UK. In: *Proceedings FloodRisk2008 Conference*, HR Wallingford.

LEES, M., YOUNG P.C., FERGUSON S., BEVEN K.J. AND BURNS, J., 1994. An adaptive flood warning scheme for the River Nith at Dumfries. In: W.R. White and J. Watts (eds), *River Flood Hydraulics*, pp. 65–75, Wiley, Chichester.

LETTENMAIER, D P AND WOOD, E F 1993. Hydrologic forecasting. In Handbook of Hydrology, Maidment R.D. (Ed.) 26.1-26.30, McGraw-Hill

LI, Y., FREER, J., BEVEN, K.J. AND MATGEN, P., 2009. Towards a limits of acceptability aproach to the calibration of hydrological models: extending observation error. *Journal of Hydrology* (in press).

LIU Y, GUPTA, H V (2007) Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, Water Resour. Res., 43, W07401, doi:10.1029/2006WR005756.

MADSEN H., BUTTS M. B., KHU S. T. AND LIONG S.Y., 2000. Data assimilation in rainfall–runoff forecasting. *4th International Conference on Hydroinformatics*, July 2000, Cedar Rapids, Iowa, USA.

MANTOVAN P, TODINI E (2008) Hydrological Forecasting Uncertainty Assessment: Incoherence of the GLUE methodology. J Hydrology, 330, 368-381

MARSIGLI, C., BOCCANERA, F., MONTANI, A. AND PACCAGNELLA, T., 2005. The COSMO-LEPS Mesoscale Ensemble System: validation of the methodology and verification. *Non Linear Processes in Geophysics*, 12, 527–536.

MCCARTHY, S.S., TUNSTALL, S., PARKER, D., FAULKNER, H. AND HOWE, J., 2007. Risk communication in emergency response to extreme floods. *Proceedings of the Institute of Civil Engineers – Water Management*.

MOLTENI, F., BUIZZA R, PALMER T N, PETROLIAGIS T, 1996. The ECMWF ensemble prediction system: methodology and validation. Q.J.R.Meteorol.Soc., 122, 73-119.

MOORE, R.J., 1982. Transfer functions, noise predictors and the forecasting of flood events in real-time. In: V.P. Singh (ed.), *Statistical Analysis of Rainfall and Runoff*, pp. 229–250, Water Resources Publications.

MOORE, R.J., 1999. Real-time flood forecasting systems: perspectives and prospects. In: R. Casale and C. Margottini (eds), *Floods and Landslides: Integrated Risk Assessment*, Chapter 11, pp. 147–189, Springer.

MOORE, R.J., 2002. Aspects of uncertainty, reliability and risk in flood forecasting systems incorporating weather radar. In: J.J. Bogardi and Z.W. Kundzewicz (eds), *Risk, Reliability, Uncertainty and Robustness of Water Resources Systems*, pp. 30–40, Cambridge University Press.

MOORE, R.J., 2007. The PDM rainfall–runoff model. *Hydrology and Earth System Sciences*, 11(1), 483–499.

MOORE, R.J., AUSTIN, R.M. and CARRINGTON, D.S., 1993. *Evaluation of FRONTIERS and Local Radar Rainfall Forecasts for use in Flood Forecasting Models*. NRA R&D Note 225, National Rivers Authority, Bristol, 156pp.

MOORE, R.J. AND BELL, V.A., 2001. *Comparison of Rainfall–Runoff Models for Flood Forecasting. Part 1: Literature Review of Models*. Environment Agency R&D Technical Report W241, 94pp.

MOORE, R.J. AND BELL, V.A., 2002. Incorporation of groundwater losses and well level data in rainfall–runoff models illustrated using the PDM. *Hydrology and Earth System Sciences*, 6(1), 25–38.

MOORE, R.J., BELL, V.A., AUSTIN, R.M. AND HARDING, R.J., 1999. Methods for snowmelt forecasting in upland Britain. *Hydrology and Earth System Sciences*, 3(2), 233–246.

MOORE, R.J., BELL, V.A. AND CARRINGTON, D.S., 2000. Intercomparison of rainfall–runoff models for flood forecasting. In: M. Lees and P. Walsh (eds), *Flood Forecasting: What Does Current Research Offer the Practitioner?* pp. 69–76, Occasional Paper No. 12, British Hydrological Society.

MOORE, R.J., BELL, V.A., COLE, S.J. AND JONES, D.A., 2007. *Rainfall–Runoff and Other Modelling for Ungauged/Low-Benefit Locations*. Science Report – SC030227/SR1, Environment Agency, Bristol, 249pp.

MOORE, R.J., HARDING, R.J., AUSTIN, R.M., BELL, V.A. AND LEWIS, D.R., 1996. *Development of Improved Methods for Snowmelt Forecasting*. R&D Note 402, National Rivers Authority, 192pp.

MOORE, R.J. AND JONES, D.A., 1978. An adaptive finite-difference approach to real-time channel flow routing. In: Vansteenkiste, G.C. (ed.), *Modelling, Identification and Control in Environmental Systems*, North-Holland, Amsterdam, The Netherlands, 153-170.

MORADKHANI, H.S., SOROOSHIAN S, GUPTA H V AND HOUSER P 2005, Dual state-parameter estimation of hydrologic model using ensemble Kalman filter, Advances in Water Resources, 28(2), 135-147

NELDER, J.A. AND MEAD, R., 1965. A Simplex method for function minimisation. *Computer Journal*, 7, 308–313.

O'CONNELL, P.E., BEVEN, K.J., CARNEY, J.N., CLEMENTS, R.O., EWEN, J., FOWLER, H., HARRIS, G.L., HOLLIS, J., MORRIS, J., O'DONNELL, G.M., PACKMAN, J.C., PARKIN, A., QUINN, P.F., ROSE, S.C., SHEPHERD, M. AND TELLIER, S., 2005. *Review of Impacts of Rural Land Use and Management on Flood Generation: Impact Study Report*. Defra, Research and Development Technical Report FD2114/TR. Defra Flood Management Division, London, 142pp.

PAGE, T., BEVEN, K.J. AND FREER, J., 2007. Modelling the chloride signal at the Plynlimon Catchments, Wales, using a modified dynamic TOPMODEL. *Hydrological Processes*, 21, 292–307.

PAPPENBERGER, F. AND BEVEN, K.J., 2006. Ignorance is bliss: 7 reasons not to use uncertainty analysis. *Water Resources Research*, 42, W05302, doi:10.1029/2005WR004820.

PAPPENBERGER, F., BEVEN, K.J., HUNTER N., GOUWELEEUW, B., BATES, P., DE ROO, A. AND THIELEN, J., 2005a. Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall–runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). *Hydrology and Earth System Sciences*, 9(4), 381–393.

PAPPENBERGER, F., BEVEN, K., HORRITT, M. AND BLAZKOVA, S., 2005b. Uncertainty in the calibration of effective roughness parameters in HEC-RAS using inundation and downstream level observations. *Journal of Hydrology*, 302, 46–69.

PAPPENBERGER, F., HARVEY, H., BEVEN K., HALL, J. AND MEADOWCROFT, I., 2006a. Decision tree for choosing an uncertainty analysis methodology: a wiki experiment, http://www.floodrisknet.org.uk/methods http://www.floodrisk.net. *Hydrological Processes*, 20, 3793–3798.

PAPPENBERGER, F., MATGEN, P., BEVEN, K.J., HENRY, J-B., PFISTER, L. AND DE FRAIPONT, P., 2006b. Influence of uncertain boundary conditions and model structure on flood inundation predictions. *Advances in Water Resources*, 29(10), 1430–1449.

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

PAPPENBERGER, F., FRODSHAM, K., BEVEN, K.J., ROMANOVICZ, R. AND MATGEN, P., 2007b. Fuzzy set approach to calibrating distributed flood inundation models using remote sensing observations. *Hydrology and Earth System Sciences*, 11(2), 739–752.

PAPPENBERGER, F., BEVEN, K.J., RATTO, M. AND MATGEN, P., 2008. Multi-method global sensitivity analysis of flood inundation models. *Advances in Water Resources*, 31(1), 1–14.

QUIBY, J. AND DENHARD, M., 2003. SRNWP-DWD Poor Main Ensemble Prediction System: The PEPS Project. *Eumetnet News*, 8, http://www.eumetnet.eu.org

RAFTERY, A.E., BALABDAOUI, F., GNEITING, T. AND POLAKOWSKI, M., 2003. *Using Bayesian Model Averaging to Calibrate Forecast Ensembles*. Technical Report 440, Department of Statistics, University of Washington, Seattle.

RAFTERY, A.E., GNEITING, T., BALABDAOUI, F. AND POLAKOWSKI, M., 2005. Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*, 133, 1155–1174.

RAMSBOTTOM, D.M. AND WHITLOW, C.D., 2003. *Extension of Rating Curves at Gauging Stations. Best Practice Guidance Manual*. R&D Manual W6-061/M, Environment Agency, 254pp.

REED D W (1984) A Review of British Flood Forecasting Practice. Institute of Hydrology Report No. 90

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

REGGIANI, P. AND WEERTS, A.H., 2008a. Probabilistic quantitative precipitation forecast for flood prediction: an application. *Journal of Hydrometeorology*, 9, 76–95.

REGGIANI, P.M. AND WEERTS, A.H., 2008b. A Bayesian approach to decision-making under uncertainty: an application to real-time forecasting in the river Rhine. *Journal of Hydrology,* 356, 56–69.

REGGIANI, P., RENNER, M., WEERTS, A.H. AND VAN GELDER, P., 2009. Uncertainty assessment via Bayesian revision of ensemble streamflow predictions in the operational river Rhine forecasting system. *Water Resources Research* (in press).

ROBERTS, N., COLE, S.J., FORBES, R., MOORE, R.J. AND BOSWELL, D., 2009. Use of high-resolution NWP rainfall and river flow forecasts for advanced warning of the Carlisle flood. *Meteorological Applications*, 16(1), 23-34. doi: 10.1002/met.94..

ROMANOWICZ, R., BEVEN, K.J. AND TAWN, J., 1996. Bayesian calibration of flood inundation models. In: M.G. Anderson, D.E. Walling and P.D. Bates (eds), *Floodplain Processes*, pp. 333–360, Wiley, Chichester.

ROMANOWICZ, R. AND BEVEN, K. J., 1998. Dynamic real-time prediction of flood inundation probabilities. *Hydrological Sciences Journal*, 43(2), 181–196.

ROMANOWICZ, R., BEVEN, K.J. AND YOUNG, P.C., 2006a. Uncertainty propagation in a sequential model for flood forecasting. *IAHS Publication*, 303, 177–184.

ROMANOWICZ, R.J., YOUNG, P.C. AND BEVEN, K.J., 2006b. Data assimilation and adaptive forecasting of water levels in the River Severn catchment. *Water Resources Research*, 42(W06407):doi:10.1029/2005WR004373.

ROMANOWICZ, R.J., YOUNG, P.C., BEVEN, K.J. AND PAPPENBERGER, F., 2008. A data based mechanistic approach to nonlinear flood routing and adaptive flood level forecasting. *Advances in Water Resources*, 31, 1048–1056.

RUNGO, M., REFSGAARD, J.C. AND HAVNO, K., 1989. Improvement of the updating routine in the MIKE 11 modelling system for real-time flood forecasting. *Proceedings of HYDROCOMP'89 Conference*, Dubrovnik, Yugoslavia.

SCHAAKE, J., DEMARGNE, J., HARTMAN, R., MULLUSKY, M., WELLES, E., WU, L., HERR, H., FAN, X. AND SEO, D.J., 2007. Precipitation and temperature ensemble forecasts from single-value forecasts. *Hydrology and Earth System Sciences Discussions*, 4, 655–717.

SCHELLEKENS, J., 2008a. *AR Module Background Information*. Deltares Wiki page http://public.deltares.nl/display/FEWSDOC/AR+Module+Background+information

SCHELLEKENS, J., 2008b. *08 Error Correction Module (ARMA).* Deltares Wiki page http://public.deltares.nl/display/FEWSDOC/08+Error+Correction+Module+%28ARMA%29

SEO, D.J., HERR, H.D. AND SCHAAKE, J.C., 2006. A statistical post-processor for accounting of hydrologic uncertainty in short-range ensemble streamflow prediction. *Hydrology and Earth System Sciences Discussions*, 3, 1987–2035.

SHRESTHA, D.L. AND SOLOMATINE, D.P., 2006. Machine learning approaches for estimation of prediction interval for the model output. *Neural Networks,* 19, 225–235.

SMITH, P.J., BEVEN, K.J., TYCH, W., HUGHES, D., COULSON, G. AND BLAIR, G., 2008a. The provision of site specific flood warnings using wireless sensor networks. In: *Proceedings FloodRisk2008 Conference*, HR Wallingford.

SMITH, P.J., BEVEN, K.J. AND TAWN, J.A., 2008b. Detection of structural inadequacy in process-based hydrological models: a particle-filtering approach. *Water Resources Res*earch, 44(1) W01410, doi:10.1029/2006WR005205.

SMITH, P.J., TAWN, J. AND BEVEN, K.J., 2008c. Informal likelihood measures in model assessment: theoretic development and investigation. *Advances in Water Resources*, 31, 1087–1100.

SMITH, R.N.B., BLYTH, E.M., FINCH, J.W., GOODCHILD, S., HALL, R.L. AND MADRY, S., 2006. Soil state and surface hydrology diagnosis based on MOSES in the Met Office Nimrod nowcasting system. *Meteorological Applications,* 13, 89–109.

STEDINGER J. R., R. M. VOGEL, S. U. LEE, R. BATCHELDER (2008), Appraisal of the generalized likelihood uncertainty estimation (GLUE) method, Water Resour. Res., 44, W00B06, doi:10.1029/2008WR006822.

SZOLLOSI-NAGY, A., 1976. *Introductory Remarks on the State Space Modelling of Water Resource Systems.* International Institute for Applied Systems Analysis, RM-76–73, 81pp.

TAYLOR, C.J., PEDREGAL, D.J., YOUNG, P.C. AND TYCH, W., 2007. Environmental time series analysis and forecasting with the CAPTAIN toolbox. *Environmental Modelling & Software*, 22, 797–814.

THIELEN, J., DE ROO, A., NIEMEYER, S., GOUWELEEUW, B., VAN DER KNIJFF, J. AND FRANCHELLO, G., 2004. Development and testing of a Pan-European Flood Alert System (EFAS*). Abstracts of 2nd European Flood Alert System (EFAS) Workshop*, Ispra, 10–12 November 2004.

TODINI, E., 2004. Role and treatment of uncertainty in real-time flood forecasting. *Hydrological Processes*, 18, 2743–2746.

TODINI, E., 2007. Hydrological modeling: past, present and future. *Hydrology and Earth System Sciences*, 11(1), 468–482.

TODINI, E., 2008. A model conditional processor to assess predictive uncertainty in flood forecasting. *International Journal of River Basin Management*, 6(2), 123–137.

TRACTON, M S AND KALNAY E, 1993. Ensemble forecasting at NMC: Practical aspects. Weather and Forecasting, 8, 379-398

VAN DER SLUIJS, J., RISBEY, J.S., KLOPROGGE, P., RAVETZ, J.R., FUNTOWICZ, S.O., CORRAL QUINTANA, S., GUIMARAES PEREIRA, A., DE MARCHI, B., PETERSEN, A.C., JANSSEN, P.H.M., HOPPE, R. AND HUIJS, S.W.F., 2003. *RIVM/MNP Guidance for Uncertainty Assessment and Communication: Detailed Guidance.* Utrecht University and RIVM.

VAN DER SLUIJS, J., JANSSEN, P.H.M., PETERSEN, A.C., KLOPROGGE, P., RISBEY, J.S., TUINSTRA, W., VAN ASSELT, M.B.A. AND RAVETZ, J.R., 2005. *RIVM/MNP Guidance for Uncertainty Assessment and Communication: Tool Catalogue for Uncertainty Assessment*. Utrecht University and RIVM.

VERKADE, J., 2008. *On the Value of Flood Warning Systems*. MSc Thesis, TU Delft.

VILLARINI, G., MANDAPAKA, P.V., KRAJEWSKI, W.F. AND MOORE, R.J., 2008. Rainfall and sampling uncertainties: a rain gauge perspective. *Journal of Geophysical Research – Atmospheres*, 113, D11102, doi:10.1029/2007JD009214.

VRUGT, J.A. AND ROBINSON, B.A., 2007. Treatment of uncertainty using ensemble methods: comparison of sequential data assimilation and Bayesian model averaging. *Water Resources Research,* 43, W01411, doi:10.1029/2005WR004838.

WALLINGFORD WATER, 1994. *A Flood Forecasting and Warning System for the River Soar: Stage 2 Report.* Contract report to the National Rivers Authority Severn Trent Region (Version 2.0), January 1995, 99 pp, Wallingford Water, Wallingford (available from Centre for Ecology & Hydrology, Wallingford).

WEERTS, A.H., 2007. *DATools*. WL|Delft Hydraulics, December 2007.

WEERTS, A.H., 2008. Hindcast of water levels for the Rhine branches and the Meuse 2006 & 2007. *Deltares*, December 2008.

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., EL SERAFY, G., HUMMEL, S., DHONDIA, J. AND GERRITSEN, H., 2008a. Application of generic data assimilation tools (DATools) for operational flood forecasting: examples for the Rhine & Meuse. *CAHMDA-3 Workshop,* Melbourne, January 2008.

WEERTS, A.H., MEISNER, D. AND RADEMACHER, S., 2008b. *Input Data Rainfall–Runoff Model Operational Systems FEWS-NL/DE*. Deltares/BfG, December 2008.

WEERTS, A.H., EL SERAFY, G.Y.H., HUMMEL, S., DHONDIA, J. AND GERRITSEN, H., 2009. Application of generic data assimilation tools (DATools) for flood forecasting purposes. *Computers and Geoscience* (submitted).

WERNER, M., REGGIANI, P. AND WEERTS, A.H., 2009. Reducing and quantifying uncertainty in operational forecasting: examples from DELFT-FEWS forecasting platform/system. In: K. Beven and J. Hall (eds), *Applied Uncertainty Analysis for Flood Risk Management* (in press).

WMO, 1992. *Simulated Real Time Intercomparison of Hydrological Models*. WMO Operational Hydrology Report No. 38, WMO No. 779.

WOOD, S.J., JONES, D.A. AND MOORE, R.J., 2000a. Accuracy of rainfall measurement for scales of hydrological interest. *Hydrology and Earth System Sciences*, 4(4), 531–543.

WOOD, S.J., JONES, D.A. AND MOORE, R.J., 2000b. Static and dynamic calibration of radar data for hydrological use. *Hydrology and Earth System Sciences*, 4(4), 545–554.

WOOD, A.W. AND SCHAAKE, J.C., 2008. Correcting errors in streamflow forecast ensemble mean and spread. *Journal of Hydrometeorology*, 9, 132–148.

YANG, Z. AND HAN, D., 2006. Derivation of unit hydrograph using a transfer function approach. *Water Resources Research*, 42, W01501, doi:10.1029/2005WR004227.

YOUNG, P.C., 1993. Time variable and state dependent modelling of nonstationary and nonlinear time series. In: T. Subba Rao (ed.), *Developments in Time Series Analysis*, pp. 374–413, Chapman & Hall, London.

YOUNG, P.C., 2002. Advances in real time forecasting. *Philosophical Transactions of the Royal Society London*, A360, 1430–1450.

YOUNG, P.C., 2009. Real-time updating in flood forecasting and warning. In: I. Cluckie, G. Pender, C. Thorne and H. Faulkner (eds), *The Flood Management Handbook* (in press). Wiley-Blackwell (available from author).

YOUNG, P.C. AND BEVEN, K.J., 1994. Data-based mechanistic modelling and the rainfall–flow non-linearity. *Environmetrics*, 5, 335–363.

YOUNG, P.C. AND RATTO, M., 2008. A unified approach to environmental systems modeling. *Stochastic Environmental Research and Risk Assessment*, pp. (ISSN) 1436–3240 (Print), 1436–3259 (Online), DOI 10.1007/s00477–008–0271–1.

YOUNG, P.C., ROMANOWICZ, R.J., BEVEN, K.J., 2006. *Updating Algorithms in Flood Forecasting*. Technical Report Report UR5, http://www.floodrisk.org.uk, Flood Risk Management Research Consortium, UK.

YOUNG, P.C., LEEDAL, D. AND BEVEN, K.J., 2009. Reduced order emulation of distributed hydraulic models. In*: Proceedings 15th IFAC Symposium on System Identification SYSID09*, St Malo, France.

ZADEH, L. 1965, Fuzzy Sets, Control, 8:338-353

Appendix A – Note on Consultation and Review

1 Introduction

1.1 Background

Robust forecasts are vital in providing a comprehensive flood warning service to people and businesses at risk from flooding. For fluvial flood forecasting, rainfall–runoff, flow routing and hydraulic models are often combined into model cascades and are run automatically in the Environment Agency's National Flood Forecasting System (NFFS).

The outputs from these models are currently 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. 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 that the Environment Agency provides.

This R&D project will develop and test practical probabilistic methods to quantify and, where possible, reduce uncertainties around fluvial flood forecasts from sources other than predicted rainfall, which is already being addressed on the 'Hydrological Modelling with Convective Scale Rainfall' R&D project.[1] This will provide an overarching framework for assessing uncertainties in fluvial forecasting in a risk-based manner which, for completeness, will also provide the possibility to include rainfall forecasting uncertainty. Some of the specific project objectives are:

- To review current experience and consult key stakeholders to refine user needs.
- To recommend and test suitable techniques for the probabilistic treatment of the most important sources of uncertainty and combine them into an overarching uncertainty framework.
- To assess the possibilities and benefits of real-time/adaptive updating for probabilistic hydraulic/hydrological models.
- To demonstrate and validate the suggested techniques for linked forecasting models through case studies in NFFS.
- To recommend and investigate alternative ways of reducing run times for probabilistic flood forecasts.
- To provide updated guidance on probabilistic fluvial flood forecasting and develop an implementation plan.

The main output from this project will be up-to-date practical guidance on how to use probabilistic techniques in fluvial flood forecasting in order to interpret possible uncertainties around flood forecasts and ultimately allow for improved management of flood events. This will be supplemented by a number of practical case studies which will demonstrate how certain uncertainty techniques can add value to the forecasting process. The main focus of this project is

[1] It is worth noting that, in many of the larger catchments in the UK, satisfactory forecast lead times can be obtained from observations of rainfall and/or upstream river levels or flows alone

to demonstrate the practicality and benefits of applying probabilistic techniques to fluvial flood forecasting models which are sufficiently robust to be considered for use in an operational environment.

A national Flood Incident Management project is currently determining approaches and business change required to embed probabilistic flood forecasting and warning (PFFW) into current Environment Agency practices. R&D Project SC080030 is closely aligned with this project and will provide some of the required scientific evidence and methods.

1.2 Scope of Report

The project started in December 2008[2] and will complete in late 2010, and includes the following three main phases:

- Phase 1 – review, consultation and scoping (4–5 months).
- Phase 2 – development of suitable uncertainty framework and application to case studies (11–12 months).
- Phase 3 – best practice guidance (7–8 months).

One of the early tasks in Phase 1 of the project was to consult regional flood forecasting staff to determine gaps in available guidance, to assess end user requirements, to discuss ideas for possible case studies, and to allow for refinements in project scope. This report describes the main findings from those consultations and is the first main output from the project. The findings from the report will be discussed at the first project workshop, which will be held in Solihull on 16 March 2009.

The consultations were performed by telephone and meetings during the period 18 December 2008 to 16 January 2009 and, with the exception of the final meeting, at least one member of the Project Board was present at each discussion. Table 1.1 summarises the dates, locations and participants in each discussion.

A detailed agenda was issued in advance of each meeting and this report is structured around the following five main topics which appeared in the agenda:

- Views on the relative importance of different sources of uncertainty in different forecasting situations, and for different types of models and forecast lead time requirements.
- The selection criteria for case studies during Phase 2 of the project.
- Ongoing regional studies (if available) on sources of uncertainty in fluvial flood forecasting models.
- Experience with the performance of integrated catchment models combining rainfall–runoff, flow routing and/or hydrodynamic components.
- User requirements for the Real-Time Modelling Guidelines and other project outputs.

The majority of meetings lasted from 2 to 3 hours, with the shortest lasting just under 2 hours, and the longest lasting about 4 hours. Typically, the background to the project was presented during the first 10–20 minutes of each meeting, with approximately half of each meeting spent on a general summary of approaches to forecasting in each region, and a discussion of the first main topic on the list above (sources of uncertainty).

[2] The research contractors are Atkins (lead), Deltares, Lancaster University, CEH Wallingford and Edenvale Young

The report attempts to provide a summary of the full range of views which were expressed although, in many cases, comments have been abbreviated or paraphrased to provide more consistency throughout the report (e.g. in the length and format of each item). Also, at this stage, no interpretation or commentary on the comments is provided, and this will appear in the draft Phase 1 report which is to be issued in March 2009.

1.3 Layout of Report

The remainder of this report is presented as follows:

Section 2 – **Sources of Uncertainty** – describes views regarding sources of uncertainty in flood forecasting models, and the relative importance of each source.

Section 3 – **Case Studies** – summarises comments on the proposed criteria for selecting case studies, and suggestions from each region for possible case studies.

Section 4 – **Ongoing Studies** – describes ongoing regional and national studies which were mentioned during the consultations.

Section 5 – **Experience with integrated catchment models** – summarises the current types of integrated catchment models in each region, and experience gained in developing and using these types of model.

Section 6 – **User Requirements** – summarises initial views on the content, format and uses of the uncertainty framework and guidelines to be produced during Phase 3 of this project.

Section 7 – **Conclusions** – summarises the main conclusions from the consultation exercise.

Table 1.1 – Summary of Consultation Meetings and Phone Conferences (Project Board representatives shown in bold type).

Consultation	Location	Date	Participants
Southern Region	Worthing (by phone)	18 December 2008	**Mike Vaughan** Paul Swinburne Julia Farrell Kevin Sene (Atkins)
Midlands Region	Solihull	19 December 2008	**Richard Cross** **Karen Hudson** Ian Clayton Kevin Sene (Atkins)
North East Region	Leeds	22 December 2008	**Andy Lane** Dave Hill David Lindsay Susan Daffern *Richard Maxted (by email)* *Neil Ryan (by email)* Kevin Sene (Atkins)
Anglian Region	Peterborough (by phone)	6 January 2009	Steve Taylor David Price Deborah Cooper **Stefan Laeger** Kevin Sene (Atkins) Neil Breton (Atkins)
Thames Region	Reading (by phone)	8 January 2009	Nigel Outhwaite David Rylands Stuart Hyslop **Stefan Laeger** Kevin Sene (Atkins) Neil Breton (Atkins)
Thames Barrier	Thames Barrier (by phone)	13 January 2009	Colin Carron Sothi Sothiratnam **Stefan Laeger** Kevin Sene (Atkins) Neil Breton (Atkins)
EA Wales	Cardiff (by phone)	13 January 2009	Andy Lane Sam Taylor-Heard **Stefan Laeger** Kevin Sene (Atkins) Neil Breton (Atkins)
South West Region	Exeter (by phone)	14 January 2009	Oliver Pollard **Stefan Laeger** Kevin Sene (Atkins) Neil Breton (Atkins)
North West Region	Warrington	16 January 2009	Mark Franklin Peter Spencer Claire Wheeler Debbie Pinnington Kevin Sene (Atkins) Yiping Chen (Atkins)

2 Sources of Uncertainty

2.1 Introduction

In general terms, sources of spatial and temporal uncertainty can include (e.g. Beven *et al.* 2005; Butts *et al.* 2005):

- Random or systematic errors in model inputs (boundary or initial conditions).
- Random or systematic errors in observed data used to measure simulation accuracy.
- Uncertainties due to calibration of effective (model) parameter values.
- Uncertainties due to incomplete or biased model structures (i.e. model configuration).

In fluvial flood forecasting applications, sources of uncertainty can include rainfall observations and forecasts, antecedent conditions, high flow ratings, downstream boundary conditions (tidal levels, structure settings etc), and other sources. Real-time updating (data assimilation) techniques, such as error correction and state updating, can also help with reducing and quantifying the uncertainty in model outputs.

Model types which may need to be considered include rainfall–runoff, flow routing, and hydrodynamic models, and models for specific control structures (washlands, reservoirs, barrages). The magnitude of uncertainties can vary with lead time and the magnitude of the event. The influence of real-time updating or data assimilation also needs to be considered, together with other potential sources of uncertainty (e.g. channel blockage, human errors, defence breaches, infrastructure failure).

The discussions were structured around the following three main questions and the remainder of this section summarises the main responses to those questions:

- For river flood forecasting (rainfall–runoff and routing) models, what in your experience are the main sources of uncertainty in model forecasts, and how do these vary with type of model, location, type of rainfall event etc? (Section 2.1)
- What, for your region, would you say are the most important sources of uncertainty, and can you rank these in any particular order? (Section 2.2)
- For the suggestions from the Project Board, which specific aspects of these topics seem most important to you? Also, do you have any other suggestions for sources of uncertainty which might be considered on this project (with reasons)? (Section 2.3)

2.2 Main Sources of Uncertainty

An early suggestion from the Project Board was that the project could investigate the practicability and benefits of applying probabilistic techniques for the following sources of uncertainty:

- Catchment Averaging of Raingauge Data
- Validity of Rating Curves
- Model Calibration (Hydrodynamic Models)
- Model Calibration (Rainfall–Runoff Models)
- Representation of Floodplain Storage

- Representation of Antecedent Conditions
- Representation of Ungauged Inflows
- Influence of Structure Operations.

The following subsections describe the main discussions regarding each of these topics.

2.2.1 Catchment Averaging of Raingauge Data

The discussions suggested that uncertainties in raingauge-based inputs to catchment models can arise from a number of sources, relating to the raingauge network density, the type of rainfall event (frontal/convective etc), electrical and other instrumentation problems, and other factors (e.g. snowfall). The type of event is particularly an issue for convective type rainfall, when raingauges can miss the heaviest rainfall. All regions reported using raingauge data as inputs to some, or all, of the integrated catchment models used in the region, and Table 2.1 summarises the catchment averaging methods which were described.

Table 2.1 – Approaches to Catchment Averaging of Raingauge Data (values show number of regions).

Method	Most catchments	Some catchments
Thiessen Polygons	3	
Thiessen Polygons (adjusted subjectively for SAAR, topographic influences etc)	2	
Regional Surface Fitting Procedure	2	
Weights based on distance from catchment/topography/other factors	1	1

The number of raingauges used in catchment averaging procedures ranged from a small number (with either 1–2 or 5–6 quoted typically) through to the approach of using all of the gauges available in the region (some 90–100 gauges, in one case). Some advantages noted in using a small number of gauges were that the rainfall estimation procedure can be tailored to the individual catchment model (in particular, for topographic effects), while a regional approach is considered to be more robust to loss of data during an event, and provides greater consistency across catchments.

The use of weather radar based rainfall actuals was also discussed. All regions reported that the images are used to monitor the progression and development of rainfall events, and that rainfall accumulation values are used as a backup in case of failure of the raingauge inputs. Two regions also reported the quantitative use of radar rainfall data as a primary input to rainfall–runoff models for catchments where intercomparison studies had shown that the catchment rainfall estimates outperformed the raingauge based estimates. In the past, one region had also previously explored real-time adjustment of radar data with raingauge data, and the new hourly Met Office bias correction approach was also noted (although this is based on a less dense network of raingauges than that operated by the Environment Agency).

Several regions noted that they had observed considerable differences between raingauge and radar rainfall based estimates of catchment rainfall, and that sometimes these were not systematic, in the sense of being consistently higher or lower during rainfall events. Some regions see the accuracy of radar rainfall actuals as comparable to that of raingauge data, with the best approach to use depending on the location of the catchment relative to the nearest radar, topography, the raingauge network density, and other factors, while others doubt the ability of radar rainfall estimates to capture orographic effects. However, both types of measurement are subject to uncertainties, and it is often difficult to say which is more representative of the actual values. The need to calibrate rainfall–runoff models to the same rainfall inputs as used in real-time was therefore emphasised on several occasions.

In most meetings, the performance of precipitation nowcasts (0–6 hours ahead) and Numerical Weather Prediction model outputs (6–36 hours ahead) was also discussed, in addition to that of raingauge-based and radar rainfall actuals.[3] At present, within the Environment Agency, it is a time consuming task to compare these types of data and forecasts, with doubts about the appropriate performance measures to use. However, it was noted that the ongoing national Skill Scores project, which is due to complete in 2009, will for the first time provide tools in NFFS to allow a consistent approach to the intercomparison of rainfall forecasts, radar rainfall actuals, and raingauge data at a catchment scale across England and Wales. Also, several regions noted that improvements should be seen once the new Thurnham, Old Buckenham and Northumbrian radar become operational.

Regarding the use of rainfall forecasts in catchment models, in some regions, many catchments are large enough to provide sufficient lead time by waiting for rainfall to 'hit the ground', so that the use of rainfall forecasts is primarily for early warning for mobilisation, and a general 'heads up', and sometimes this is in a negative sense; for example, that no flooding is likely over the next weekend. In other regions, rainfall forecasts provide the only practical way of providing sufficient lead time on fast response catchments, and at least two regions are currently making limited use of rainfall–runoff model outputs with rainfall forecast inputs as a key tool to help with issuing flood warnings for a few fast response catchments, using thresholds for the onset of flooding (i.e. 'Result Thresholds').

2.2.2 Validity of Rating Curves

A number of issues can lead to uncertainty in rating curves, including seasonal effects, backwater influences, artificial influences, tidal influences, a lack of spot gauging data (for natural river reaches), and issues with gauge datums. However, for flood forecasting applications, it is the uncertainty at the high flow end of the rating which is of most importance, with problems including the difficulty of gauging at high flows, and the rarity of high flow events. Also, additional flow mechanisms may become apparent at high flows, such as floodplain flows, non-modular flows, and bypassing of gauges.

Many points were raised during the discussion of this topic, and these can be briefly summarised as follows:

- High Flow Issues – the problems with high flow ratings are well known and to some extent can be mitigated by an appropriate model design (e.g. omitting gauges which are known to be less accurate at high flows), and additional modelling studies (e.g. extending ratings using 1D or 2D hydraulic modelling, and additional spot gauging campaigns). The issue of rating curve limits was also raised, and whether it is better to truncate the rating (because it is uncertain) or to derive an estimate for the full hydrograph (which may be considerably in error) to provide inflows to locations further downstream.
- Seasonal Influences – seasonal influences were reported to significantly affect the accuracy of rating curves in 3 regions (and these were typically regions in lower lying parts of England and Wales). These effects cannot be easily represented in NFFS at present although it is understood that this functionality is being developed, together with the option to switch ratings (e.g. from a rated section to an ultrasonic gauge when flows go out of bank). Seasonal effects can also to some extent 'balance out' over a number of events. In one example which was quoted, weed growth could be sufficient

[3] These discussions were useful although it was noted in each meeting that uncertainty in rainfall forecasts is outside the scope of the present project (other than in developing the uncertainty framework) since this is being considered in the 'Hydrological Modelling with Convective Rainfall' project

to change the flow regime, leading to water bypassing the station. In another, seasonal influences had been found to change the behaviour of error correction updating routines over the course of a year.

- Mobile River Beds – the issue of changes in bed profiles due to tidal effects, and following major flood events, was raised as an issue affecting rating curves and model calibration for certain locations in 2 regions.
- Development of Rating Curves – all regions reported that rating curves are normally developed by Field Monitoring and Data, but are sometimes improved as part of flood forecasting model development projects. Also, that curves are usually developed independently of integrated catchment models, and the rating curve parameters are not tuned to improve the model performance. Rainfall-level calibration options (e.g. as in the PDM model) are generally not used. Rating curves are normally based on spot gaugings (if available), survey and backwater calculations, or 1D or 2D hydrodynamic modelling, although the derivation of effective ratings, based on upstream and downstream flows, and low flow spot gaugings, was also described as a temporary solution.
- Maintenance of Rating Curves – although outside the scope of this project, another issue which was discussed was that of updating models following improvements by FMD to rating curves due to new spot gauging data, to account for changes to the river or gauge characteristics (e.g. flood defences, bridge structures, gauge datums etc), or to account for seasonal effects.
- Use of Rating Curves in Models – one point which was noted was that typically models are calibrated to flows which are derived from observed levels using a rating curve, and the forecast flows are then converted back to levels using the same curve. This to some extent mitigates errors in the rating, but only partially since it can be difficult to calibrate a model if flows are significantly in error (e.g. runoff may exceed rainfall), and incorrect flows at an upstream location can lead to errors in forecasts further downstream.
- Other Approaches to Measuring Flows – it was also noted that rating curves are usually not an issue where flows are measured directly (e.g. at ultrasonic gauges) or using a theoretical rating in which there is reasonable confidence (e.g. for a Crump Weir which fully contains high flows).

2.2.3 Model Calibration (Hydrodynamic Models)

All regions reported using hydrodynamic models as part of integrated catchment models, although in some cases this was only a recent development (in the past 4–5 years), and experience is still being gained with operation of this type of model. Some points which were raised during the discussions included:

- General – models are often developed from existing floodplain mapping and/or flood defence scheme models, rather than being built as part of flood forecasting model development projects. This means that, provided that the base model has satisfactory performance, the hydrodynamic component of the forecasting model should also perform in a similar way (ignoring, for the moment, uncertainties in rainfall–runoff model inputs and ungauged inflows). The issue of model uncertainty has therefore, to some extent, already been considered in the original model development (although improvements can always be made). The point was also made that, compared to hydrological models, the errors are often more systematic, and easier to constrain and control.

- Roughness Coefficient – this was generally reported as a secondary issue, related mainly to fine tuning of models, rather than a significant source of uncertainty. However, one region noted that values can be affected by seasonal influences, and (as for rating curves) that weed growth can sometimes be sufficient to change the flow regime (e.g. flows bypassing a gauge). Another region noted that the effective roughness can often be flow dependent, with higher values for in-channel flows (depending on type of channel, and floodplain characteristics), and with difficulties in estimating appropriate values for floodplain flows. The potential uses of the Conveyance Estimation System were also discussed to guide estimates for the roughness coefficient.

- Tidal Boundaries (general) – tidal boundaries were noted as a source of uncertainty in some cases, although related mainly to volumes on the floodplain, rather than the estimates or forecasts for levels at the downstream boundary. Two regions noted that, where tidal influences extend a significant way inland, some key issues can include how well models can propagate the tidal effects upstream, and the influences of floodplain storage, tide-locking, structure overtopping, and gate operations. For regions which use STFS forecasts of tidal levels, uncertainties in this component were generally reported not to be a significant factor for integrated catchment models, compared to the hydrological and hydraulic sources of uncertainty. Where tidal forecasts were important, this tended to be for specific locations or estuaries; for example, to the operation of the Thames Barrier, where ongoing work to develop a finer grid (3.5 km) version of the CS3X model was described for the Thames Estuary and Bristol Channel, and a number of studies into other sources of uncertainty (see Table 3.1).

- Real-time Updating – there was a range of views on how to use updating techniques with real-time hydrodynamic models, with a general move away from ‘monolithic’ models towards a number of models linking the main river flow gauges in a catchment. Error correction techniques are then easy to apply at the gauge locations, although require the model to be split at updating locations, meaning that downstream influences are not necessarily correctly propagated upstream. Internal model node updating is used in one region.

- Datums/Bed Levels – one region reported that these factors can cause major uncertainties in calibrating and using hydrodynamic models, with bed movements of 4–6 metres having been observed at one location. For this same region, a region-wide programme of GPS surveying is addressing issues with gauge datums.

- Full Flow Models – one region reported use of full flow models for a range of purposes other than flood forecasting, including navigation and water resources. In this case, the uncertainties arising from abstractions and discharges are more significant than for a model developed primarily for flood forecasting.

At this point, it was also convenient to discuss hydrological flow routing models, with two regions noting that representation of the wave-speed–discharge curve can be a major source of uncertainty, as can representation of the influence of flows to and from the floodplain. There was also some discussion of simpler approaches to flow routing, including the use of single or multiple level to level correlations, and how often they are typically developed on peak values, but applied across the full flow range (an approach which can work well in some situations). One region also noted a significant issue with matching of model performance at the boundaries between flow routing and hydrodynamic models (i.e. achieving consistency in forecasts for both flows and levels).

2.2.4 Model Calibration (Rainfall–Runoff Models)

Lumped rainfall–runoff models are widely used to estimate the inflows in integrated catchment models. Sources of uncertainty can include catchment rainfall estimates (see Section 2.2.1), rating curves (see Section 2.2.2) and the representation of antecedent conditions (see Section 2.2.6) and ungauged inflows (see Section 2.2.7).

Considering the rainfall–runoff modelling component alone, then the sources of uncertainty can include the parameter values which are used, and the model structure (where there is a choice). Also, the perceived model performance can depend on the performance measures which are chosen to calibrate and evaluate the model (for example, peak flow, timing of peak flow, timing of threshold crossing, flow volumes, root mean square error), and may vary with forecast lead time and whether real-time updating is used. Model performance for high flows, outside the range of calibration, is also an issue.

This topic generated considerable discussion during the consultations and the items which were discussed included:

- Model Structural Issues – several regions highlighted the distinction between model structural errors, and model calibration errors. Structural errors relate to the choice of model, how the catchment is conceptualised (e.g. the number and size of sub-catchments) and other factors, such as the use of alternative options in the modelling package (if there is a choice). For integrated catchment models, these issues were considered less amenable to uncertainty estimation techniques, unless the route is taken of developing a number of alternative models, and model configurations, and running them in parallel on NFFS to compare the model outputs.

- Model Calibration – for the types of models which are currently used within the Environment Agency, with the exception of event based and MCRM models, all regions now calibrate models to at least a few months of data, allowing a range of high flow and low to medium flow events to be included in the calibration. Nevertheless, given the performance measures which are often used (peak flow, timing of peak), and the fact that there are historically more flood events in winter, this tends to bias models towards performing better for widespread frontal events, rather than for the convective events which have been a feature of several major flood events in recent years. This issue also relates to the representation of catchment antecedent conditions (see Section 2.2.6). There can sometimes also be a case for using different calibration criteria for different locations around a catchment; for example, optimising a model for peak levels on tributaries, and for the overall shape/volume of the hydrograph for the main river channel.

- Extreme Events – it was noted several times that models may also not perform well for extreme events, due to a lack of data for calibration, and possibly due to additional runoff processes occurring for very high flows/high intensity rainfall (for example, soil compaction, and development of additional flow routes).

- Catchment Size – there were some brief discussions of what a reasonable lower limit is for the size of subcatchments in integrated catchment models, and how this relates to the scale/resolution of data inputs (e.g. weather radar grid scale, raingauge network density). Values in the range 2–60 km^2 were quoted as the smallest in current use.

- Result Thresholds – from the consultations, the use of rainfall–runoff model forecasts as the main basis for issuing flood warnings appears to be increasing, although is still not widespread. Of the two regions which discussed this point, one uses Result Thresholds at nearly all Forecasting Points, and one uses Result Thresholds in a few fast response catchments (but has plans to migrate all existing thresholds at Forecasting Points in the near future).

- Model Parameters – of the four main types of conceptual rainfall–runoff model used within the Environment Agency (see Section 5.2), the maximum number of parameters available for model calibration is typically in the range 10–20. However, for all types of model, typically only 3–5 parameters have a major influence on flood flow estimates (depending on the modelling package).

2.2.5 Representation of Floodplain Storage

Floodplain storage can affect both the timing and magnitude of flows further downstream in a river network, and so can be an important consideration in the design and operation of the flow routing and/or hydrodynamic components of integrated catchment models. Floodplain effects may also cause some of the difficulties which are sometimes experienced with the calibration of rainfall–runoff models, such as overestimation of peaks and timing errors.

When hydrodynamic models are used, the floodplain is usually included (where survey data is available) and hence the model performance implicitly allows for floodplain storage to some extent. However, uncertainties can arise from inaccuracies in survey data and – as noted earlier – the estimation of roughness coefficients. Also, model performance can depend on how the floodplain is represented, and how robust the methods are to wetting and drying of the floodplain (reservoir units, extended sections etc). In general though, these uncertainties were only considered to be significant on the lower and middle reaches of major rivers/areas, such as the Severn, Dee, Wye, Ouse, Somerset Levels, and Derwent. However, one issue which can arise is in multiple events, when cumulative errors in the filling and emptying of floodplain can become significant. Similar issues can also arise with controlled storage (e.g. washlands) and these are discussed in Section 2.2.8. Where floodplains are extensive, there can sometimes also be issues with understanding the main flow paths, and real-time 2D models might be an option for the future.

Where simpler modelling approaches are used (e.g. hydrological flow routing, or correlations) the effects of floodplain storage tend to be represented only if sufficient high flow data is available to parameterise the influence on downstream flows. However, in one region, floodplain storage levels in routing models can be updated during an event; for example based on observations taken by people on site, or from telemetered floodplain gauges.

2.2.6 Representation of Antecedent Conditions

The representation of antecedent conditions can have a major influence on the performance of the rainfall–runoff modelling component of an integrated catchment model. Primarily, this relates to soil moisture conditions, but other factors such as soil compaction (related to farming practices), snowmelt, frozen ground and groundwater levels were also mentioned. Other issues, such as floodplain storage, and reservoir levels, might also be considered under this heading, but are discussed in Sections 2.2.6 and 2.2.8 respectively.

During the consultations, this topic generated considerable discussion, and the key points which were raised included:

- Soil Moisture – for conceptual rainfall–runoff models (e.g. PDM, NAM), the soil store contents are modified at each time step based on observed (or forecast) rainfall and indicative evaporation. Values are unconstrained except in very wet or dry periods, when stores may fill or empty. In some models, there is the facility to update state values in real-time (e.g. PDM) or using manually entered parameters (e.g. MCRM) while, for HYADES models, one of four pre-defined runoff factors can be chosen to represent current conditions. For PRTF models, a Catchment Wetness Index (CWI) estimation procedure is used to select the most appropriate model parameters to use, and this is based on the MORECS approach, with actual evaporation values based on hourly MOSES estimates, and a PRTF-specific approach to estimating Antecedent Precipitation Index. These various approaches generally work well except that several regions reported that there are sometimes issues during the transition from dry summer conditions to wetter winter conditions, and with performance being better for intense rainfall, rather than for more moderate events (although real-time updating can help to reduce these effects). For most methods, without state updating, soil moisture values are unconstrained (not ‘ground truthed’), although it was also noted that any long-term drift in estimates tends to be halted as model stores fill in wet periods and empty in dry periods. It was suggested that alternative approaches, for example using MOSES or the IHACRES methods, might in the future lead to improved real-time estimates for soil moisture.

- Evaporation – regarding evaporation values, a variety of approaches was reported and these are summarised in Table 2.2. Views differed on whether using values which can vary over time is beneficial, with one region reporting significant improvements in model performance when using historical (MORECS) data in model calibration. Weekly MORECS values, or hourly MOSES values, were suggested as sources of ‘real-time’ evaporation data, although at present are only used operationally in one region as part of the procedure to estimate soil moisture values (see above).

Table 2.2 – Summary of Approaches to Estimating Sub-Daily Evaporation.

Method	Most catchments	Some catchments
Fixed sinusoidal profile of daily values	2	
Fixed monthly profile, with weekly updates based on MORECS outputs	1	
Disaggregated daily MORECS values (for calibration only); fixed profile in real-time	1	
Choice of 4 fixed profiles		1
Single regional profile	3	
Other	1	1

- Snow Cover – snowmelt was reported by all regions not to be a significant issue at present, while noting the potential for flooding problems in the future, and some major historical events (e.g. 1946/47, 1962/63, 1999). Both Midlands and North East Regions include a snowmelt component in rainfall–runoff models, and in the past used to receive data from an extensive network of snow observers, although this has declined in recent years. Where snowmelt models are used, it is important to keep the snowcover/depth stores updated since otherwise flows can increase significantly (and unrealistically) when air temperatures rise. It was noted that, even where a snowmelt model is not available, if raingauges are heated, the influence of snowfall can sometimes to some extent be inferred from the telemetered data.

- Groundwater – groundwater related issues (e.g. chalk, limestone, karstic formations, sink holes) were reported for some catchments within 3 regions. However, at present the general view was that the current rainfall–runoff models which are used operationally do not always adequately represent groundwater influences, although a new module is currently being developed in PDM, and the TCM model was originally developed for groundwater/water resources modelling. The MCRM model also represents groundwater, and groundwater level conditions can be updated manually if data are available. One region also reported use of a simple cumulative rainfall based index to estimating the likelihood of groundwater flooding.

2.2.7 Representation of Ungauged Inflows

A common problem in developing integrated catchment models is often that significant proportions of the catchment are ungauged. The inflows to the river network from these catchments must therefore be modelled in some way. A variety of approaches was reported for estimating ungauged inflows and these included the following techniques:

- Transfer of model parameters from a gauged catchment with similar characteristics (analogue catchments).
- As above, but then fine tuning parameter values to achieve a good calibration in the routing component of the model.
- Interpolation of parameters from gauged catchments upstream and downstream.
- Estimation of an impulse function based on Flood Estimation Handbook values (PRTF models only).
- Scaling flows on gauged flows in a nearby catchment (e.g. based on catchment area), possibly including a time difference also.

The use of Flood Estimation Handbook (and Flood Studies Report) catchment characteristics was noted as one way to help with identifying appropriate analogue catchments. The use of parameter transfer techniques seems to be becoming more widespread, with several regions reporting recent adoption of this approach (with lag and scaling methods used mainly in the past). The main advantage of transferring parameters is that this allows the overall model to better represent the spatial variations in runoff from variations in rainfall around a catchment.

In general, the importance of the ungauged component depends on the location and timing of inflows relative to the main channel flows and Forecasting Points, and the fraction of the catchment which is ungauged. Also, if the ungauged proportion of a catchment is high, there are questions about whether a model should be developed, and the risks due to the high uncertainty, and that at the very least this uncertainty should be quantified as part of the model development process, particularly for high risk locations.

2.2.8 Influence of Structure Operations

The operation of control structures such as barrages, flow regulators, lock gates, sluices and reservoir outfalls can have a significant influence on flows further downstream, and was reported as a source of uncertainty in several regions. Off-line storage areas such as washlands are also a factor in North East and Anglian regions, while Anglian Region also has several examples of pumped catchments, as does North West Region (e.g. the Alt Crossens catchment). Demountable defences were also noted, although the influence on flows downstream is probably small in most cases.

Even where real-time information is available, a key problem can be that there are no well-defined operating rules for sluices, gates etc, and structures are not always owned by the Environment Agency. Also, even if the rules are known, operators may vary these for good reasons, and the effectiveness of these variations can depend on the experience and skill of individual staff. Gates may also fail during an event.

By contrast, two regions noted that, in practice, structures may be opened or uncontrolled during flood events. For example, for reservoirs, although outflows may be controlled at low flows for compensation and scour releases, spillway flows are often unconstrained while, for rivers with canals, lock gates may normally be opened to allow the flood to pass. In these cases, there may be little need to explicitly represent these structures (assuming of course that, for gates, these are always opened in advance of flooding).

Flood defence breaches were also mentioned as another source of uncertainty although it is difficult to predict the likely locations of breaches and dam breaks or the extent of the resulting flooding (and this is probably outside the scope of this project). Similar points were also made about the difficulties (and uncertainties arising) in forecasting the effects of blockages at bridges, culverts and control structures, with instrumentation (e.g. at trash screens, CCTV) being the only realistic solution at present to detect problems.

Perhaps the most significant issue for several regions was the representation of reservoirs, and of the initial conditions (levels) at the start of a simulation. It was noted that even small errors in levels can have a large influence on outflows, particularly when the reservoir is spilling. To obtain a better understanding of reservoir influences, two regions noted that models had been developed with and without a reservoir included to assess the sensitivity to reservoir influences in real-time (for one case) and for post event analysis (for the second case). The MCRM model also includes an in-built reservoir balance model. Two regions also noted that a new reservoir module for NFFS is being developed and, once implemented, may significantly reduce the uncertainty in reservoir outflows, provided that real-time information is available on reservoir levels. This module, although not hydrodynamic, will use a simple reservoir routing approach, based on observed reservoir levels, and a representation of the spillway/outflow control rules, and should be available later in 2009.

2.3 Most Important Sources of Uncertainty

Following the general discussions on sources of uncertainty, each region was asked, if possible, to rank these sources in order of importance. Table 2.4 summarises the replies which were given, with a score of 1 being the highest, and 8 the lowest.

Table 2.4 – Most Important Sources of Uncertainty (1 = highest, 8 = lowest, c = depends on the catchment, forecasting issues, locations of forecasting points, lead time requirements etc).

Source of Uncertainty	Anglian	Midlands	North East	North West	Southern	South West	Thames	EA Wales
Catchment Averaging of Raingauge Data	1	1	6	1	2	c	3	1
Validity of Rating Curves	4	4c	7	3	c	2	2	2
Model Calibration (Hydrodynamic Models)	2/3	5c	5	5	c	c	c	2/3
Model Calibration (Rainfall–Runoff Models)	2/3	2	1	4	1	3	1	2/3
Representation of Floodplain Storage	c	3	4	c	c	c	c	c
Representation of Antecedent Conditions	c	8c	3	c	c	1	c	4
Representation of Ungauged Inflows	5	6c	2	2	3	c	c	c
Influence of Structure Operations	c	7c	8	c	c	c	c	c

For convenience, scores of 1, 2 or 3 are shown shaded. Several regions were reasonably confident in providing the first three or four entries in the list, while noting that the remaining entries all depended on catchment issues, lead time requirements, forecasting issues, the types of model, and other factors, and could merit a score of 1, 2 or 3 in some situations. One point which was noted several times was how the uncertainty varies depending on the locations of Forecasting Points in the catchment, and the catchment response time, with the flow routing/hydrodynamic components becoming more important (for a given lead time) in the lower reaches of a catchment.

Entries which could not be placed in any particular order, or whose score might vary, are shown by the letter c. In particular, structure operations and antecedent conditions were both noted as entries which could be considerably higher for some specific catchments. Also, the hydrological topics (catchment averaging, rainfall–runoff models, antecedent conditions, and ungauged inflows) are all interconnected to some extent, as are the remaining hydrodynamic topics (with rating curves providing a link between these two general categories of uncertainty).

Specific forecasting situations may also suggest a very different order; for example, during the Thames Barrier consultation meeting, the following prioritisation was provided: coastal surge forecasts, astronomical tide estimates, inflows at the western boundary of the Lower Thames model, roughness coefficients/operation of structures, and observed tidal levels at the downstream boundary of the estuary model.

The answers to this question will provide the Project Board with an indication of the most important sources of uncertainty, while recognising that this is a difficult question to answer, and may in some cases depend on regional skills and expertise, and the modelling tools which are used. For example, where reservoirs are an important influence on flood flows, the uncertainty was ranked as high but, once the new reservoir module becomes available in NFFS, models which adopt this module may be significantly more accurate, and so this source of uncertainty will drop towards the bottom of the list. A key step in developing plans for the main phase of the project will be to consider which sources of uncertainty are technically feasible to consider, within the time available during Phase 2 of the project, and the relative importance of each source.

It is also worth noting that, for all regions, uncertainties in forecast rainfall were stated to be at or near the top of the list, while acknowledging (see earlier) that the development of techniques to generate ensemble rainfall forecasts falls outside the scope of this project.

2.4 Specific Aspects and Other Sources

2.4.1 Specific Aspects

For some of the entries in Table 2.4, the following additional clarifications were provided on specific aspects that were most important:

- Antecedent Conditions – while all regions noted that soil moisture was probably the most important factor, one region noted that snowcover/depth is also an important regional issue, while another noted groundwater as important, and another felt that estimating evaporation was a critical factor.
- Validity of Rating Curves – one region noted that errors in datum values were probably the most important factor to consider, and another that rating curves at the boundaries between different types of model are probably the most important consideration.

2.4.2 Other Sources of Uncertainty

The main additional source of uncertainty which was noted was in the performance of real-time updating routines (although this is not normally considered as a source of uncertainty). For example, one region gave updating a score of 3 in the ranking exercise described in the previous section.

Although it was generally agreed that real-time updating usually provides significant improvements to forecast accuracy, and is recommended as best practice, problems can arise with poor data quality, and with the particular characteristics of individual updating algorithms. Some examples which were noted included:

- Erroneous results with error correction routines due to minor timing errors in hydrographs.
- Step changes in river flows due to minimum flow conditions in hydrodynamic models, or drift in the flow estimates from rainfall–runoff models during low flow periods.
- Time delays and a 'feedback loop' arising from state updating of a particular rainfall–runoff model store.
- Phantom peaks being inserted during relatively minor events by an error correction routine.
- The overall shape/volume of the hydrograph not being preserved/realistic after updating.

However, it was generally agreed that these problems are linked to specific approaches to updating, and the ongoing national initiative on error correction (see later) should lead to some specific recommendation on the best approaches to model updating. One region also reported that the use of 1st order error correctors helps to constrain the magnitude of adjustments, although with the disadvantage of the adjustments decaying rapidly with increasing lead time.

Another issue which was noted by two regions was that model performance is increasingly evaluated in terms of contingency measures (POD, FAR etc) and the lead time provided to properties which were flooded, which requires a knowledge of how gauge thresholds (Result Thresholds) relate to property flooding thresholds. The accuracy of this relationship (gauge to property thresholds), and of the surveyed elevations for the property thresholds, are therefore also factors in model performance, and are additional sources of uncertainty (although this topic is outside the scope of this project).

The topic of model structural/conceptualisation errors was also discussed during all meetings and, although important, was generally felt to be less easily amenable to an uncertainty analysis, i.e. for this project, the starting point for the assessment of uncertainty should be a completed model, rather than a model design, with the focus on techniques which can be applied in real-time (or simulated real-time).

3 Case Studies

3.1 Introduction

During Phase 2 of the project, a number of approaches to assessing uncertainty will be tested on a small number of case studies. We intend to have 4 case studies for practical forecasting problems based on forecasting techniques using continuous simulation approaches which we anticipate will still be used in 5 years time. These will consist of two complex linked integrated catchment models, combining rainfall–runoff, flow routing and/or hydrodynamic models (e.g. based on PDM/KW/ISIS, MRCM/DODO/ISIS) and two simpler examples to illustrate key concepts: for example:

- a rapid response catchment with rainfall and other sources of uncertainty;
- a well-defined river reach including structures, rating curve uncertainty and other factors.

At least one of these examples will include real-time updating and the influence on model performance with increasing lead time. Multi-model techniques might also be considered, such as comparisons of model outputs in real-time for different inputs (e.g. raingauge, radar rainfall), or different sets of model parameters.

During the consultations, it was noted that the choice should reflect the breadth of Environment Agency forecasting situations, and that the Project Board is open to other suggestions (while noting that rainfall forecasts fall outside the scope of this project, since they are being considered within the 'Hydrological Modelling with Convective Scale Rainfall' project). The discussions were structured around the following two questions and the remainder of this section summarises the main responses to those questions:

- Do you have any comments on the initial suggestions from the Project Board for case studies? (Section 3.1)
- Do you have any suggestions for specific catchments for consideration as case studies (with reasons)? (Section 3.2)

It was also noted as part of the introduction to the topic that, with eight regions, and the Thames Barrier, and only four case studies planned, then obviously it will not be possible to have case studies in all regions. However, the case studies will be selected so that the experience gained will be of benefit to most, if not all, other regions. One other proviso placed on the selection of case studies was that models should already be built and configured onto NFFS (since no new model development will be performed during this project).

3.2 Comments on Initial Suggestions

The general consensus from the consultations seemed to be that the choice of types of case study seemed sensible, and would provide examples which will be of interest and relevant to all regions. There was also a general view that the case studies should be chosen on technical grounds (i.e. specific types of models, and forecasting situations), rather than other criteria such as risk, or current model performance regarding level of service. One region also made the point that probabilistic techniques, although valuable, could sometimes require more time to make decisions when used operationally, so the initial applications are likely to be on catchments (or at Forecasting Points) with longer lead times, and this could be a factor in the choice of case studies.

In abbreviated form, the following more detailed points were also made:

- Rapid Response case study – some regions were interested to know the lead times covered by the term 'rapid' and it was noted that perhaps the term 'fast response' would have been more appropriate; the aim being to examine a model which relies primarily on a rainfall–runoff model to provide a forecast. The actual lead time is unimportant and could be several hours. However, it was noted that the project is not exploring the use of probabilistic techniques to extend the forecasting service to new types of forecasting issue; for example, where catchment response times are less than the level of service target.
- Integrated Catchment Models – three regions noted that their integrated catchment models do not usually include a hydrological flow routing component i.e. that rainfall–runoff models join directly to hydrodynamic models. The hydrological flow routing aspect is therefore of less interest, although should not be ruled out if it is relevant in other regions. More generally, one region noted that, within the constraints of technical feasibility, the more sources of uncertainty which can be combined, the better; for example, ungauged catchments, confluence flooding issues, structures, and uncertainty arising from rating curves and the effects of updating.
- Other Suggestions – confluence flooding was suggested as another possible type of simple case study, and a basic component of many larger integrated catchment models. An example could be provided for a confluence with a caravan park at risk, for example.

3.3 Suggestions for Specific Catchments

Suggestions for possible case studies were received from all 8 regions, and these are summarised in Table 3.1, and Figure 3.1 (although note that PRTF models are outside the scope of the present study).

Table 3.1 – Regional suggestions for possible case studies.

Region	Catchment	Category	Model Types	Description
Anglian	Cam	Integrated Catchment Model	PDM/ISIS	The model includes four tributaries upstream of a key risk area, where there is a structure. One of the tributaries is ungauged, one is well calibrated, and two are candidates for real-time updating, with approximately 70% of the flow gauged
Midlands	Severn	Integrated Catchment Model	MCRM/DODO/ISIS	Cascades of models, and types of models, in the middle and lower reaches of the catchment
	Lower Trent	Flow Routing	MCRM/DODO/ISIS	A well defined reach for which most of the major inflows are gauged, and ratings have been reviewed following the Summer 2007 floods. Various structures and a tidal downstream boundary
	Soar	Flow Routing	MCRM/DODO/ISIS	Significant ungauged inflows and complex structure operations; forecasting model currently under development
	Warwickshire Avon	Integrated Catchment Model	MCRM/DODO/ISIS	Major tributary of the river Severn at Tewkesbury. Several gauged fast response inflows, and no major issues with structures during flood events
North East	South Tyne	Integrated Catchment Model	PDM/ISIS	3 gauged PDM models, ungauged catchments, and an ISIS reach. The catchment has been used for pilot tests of performance measures such as POD and FAR, and variations in performance with lead time, and examining performance with raingauge data, no rain, and radar rainfall data
	Wansbeck	Integrated Catchment Model	PDM/KW	3 PDM models and a KW routing reach. The main flood risk area is Morpeth for which severe flooding occurred in 2008
	Upper Calder	Rapid Response	PDM	Fast response catchments to Todmorden and Walsden Water. Forecasts for the exceedance of Result Thresholds used to guide the issuing of flood warnings
	Derwent	Flow Routing	PDM/KW/ISIS	A catchment which could provide an example of storage issues, and which has PDM models with 4–5 routing reaches, and an ISIS model
North West	Upper Ribble	Rapid Response	PRTF	Model for the Upper Ribble to Ewood
	Douglas	Rapid Response	PDM	Model for the catchment to Wigan, with a reservoir in the catchment
	Greta	Rapid Response	PDM/VPMC	Model for the Greta at Keswick consisting of PDM and routing models, with reservoir influences
	Lower Eden	Flow Routing	VPMC/ISIS	Flow routing and hydrodynamic reaches to Carlisle, with a fluvial lower boundary. There is an adequate lead time from flow routing, although PRTF rainfall–runoff models have also been developed
	Lower Ribble	Flow Routing	PRTF/ISIS	Floodplains and tributary inflows in the lower catchment to Preston, with a tidal downstream boundary
	Derwent	Integrated Catchment Model	PDM/VPMC/ISIS	An integrated catchment model combining rainfall–runoff, reservoir, routing and hydrodynamic components

Table 3.1 (continued)

Region	Catchment	Category	Model Types	Description
Southern	Medway	Integrated Catchment Model	PDM/ISIS	A complex model with many rainfall–runoff inputs and structures (gates, locks/navigable reaches), and the Leigh Barrier
	Sussex Ouse	Integrated Catchment Model	PDM/ISIS	A complex model for the catchment to Lewes, with a reservoir, tidal downstream boundary, and four major subcatchments
	Adur	Integrated Catchment Model	PDM/ISIS	A complex model with an extensive area of tidally influenced floodplain, with flap gates controlling flows to and from the floodplain, and issues with stop logs in summer months
South West	Bude/Neet	Integrated Catchment Model	PRTF/ISIS	A fast response catchment with a tidal downstream boundary in the town of Bude, with two rainfall–runoff models and a hydrodynamic model reach
	Bristol Avon	Integrated Catchment Model	PRTF/ISIS	A complex model to the city of Bristol with several linked ISIS models and rainfall–runoff model inputs
Thames	Ravensbourne	Integrated Catchment Model	TCM/ISIS	A rapidly responding urban catchment with a distributed set of hydrological inputs. The integrated catchment model is a TCM-ISIS model with radar rainfall inputs, and approximately 30–40% of the catchment area is gauged
	Thames Barrier	Flow Routing	ISIS	The Thames Barrier is a possibility for a case study, and an investigation into how to combine ensemble surge forecasts with probabilistic fluvial forecasts would be of particular interest (although this is outside the current scope of the project). The current long-term programme of continuing improvements is currently focusing on uncertainties (where relevant) in surge estimates, wind shear effects, astronomical tide estimates, tidal levels at the estuary mouth, and in further improving the representation of tidal flows over Teddington weir, minor discharges into tributaries, and other factors
EA Wales	Tawe	Integrated Catchment Model	PDM/ISIS	A rapid response catchment to Swansea with 3 PDM models and one ISIS routing and a hydrodynamic model. The downstream boundary is at a structure
	Taff	Integrated Catchment Model	PDM/ISIS	A complex model for the Rhondda, Cynon, Upper Taff and Lower Taff reaches down to Cardiff, with 2 reservoir models in the upper reaches
	Solva	Integrated Catchment Model	PDM/KW/ISIS	A small fast response catchment to the town of Solva in Pembrokeshire, with an ISIS model for a flood detention reservoir in the middle reaches

Figure 3.1 – Locations of the catchments suggested for possible case studies.

4 Ongoing Studies

4.1 Introduction

The aim of this section of the consultation meetings was to briefly review the current approaches to assessing uncertainty which are used in each region, and to compile a list of area, regional, national and other studies which regional teams consider might be useful to the present study.

The discussions were structured around the following three main questions and the remainder of this section summarises the main responses to those questions:

- Are any other methods used in your region to assess uncertainty in flood forecasts in real-time? (Section 4.2)
- Have any regional or area studies been performed recently, or in the past, on the main sources of uncertainty in fluvial flood forecasting models, and what approaches are used? Are there any reports available for the project team to review? (Section 4.3)
- Are you aware of any other studies which might be useful to this project (within, or outside, the Environment Agency)? (Section 4.4)

4.2 Real-time Methods to Assess Uncertainty

During the Probabilistic Flood Forecasting Scoping Study (2006–07), and subsequent discussions during 2008 linked to the Hydrological Modelling with Convective Rainfall Study, the regional consultations suggested that the following methods are currently used for assessing uncertainty in flood forecasts:

- Use of the probability estimates in Heavy Rainfall Warnings
- Scenario modelling of gate settings, closures and blockages
- Sensitivity studies for surge estimates (e.g. +200 mm)
- Running of fluvial models with alternative rainfall inputs and scenarios (raingauge, radar actuals, forecasts etc)
- Look-up tables on sensitivity of levels/flows to rainfall estimates
- Running of alternative surge models (CS3, local, manual calculations)
- Evaluation of outputs from alternative approaches, e.g. rainfall–runoff, rainfall/catchment state assessment, threshold-based approaches
- Running of rainfall–runoff models for alternative catchment states (dry/wet etc)
- Attaching a likely range (in metres) to forecasts of peak levels.

The present consultations suggested that the following additional methods have also recently been adopted, or are being evaluated:

- Time lagged ensembles – an NFFS option to display successive forecasts on the same graph, giving an indication of the consistency in forecasts between model runs (see Figure 4.1).

- Peak level forecast range plots – plots which show how successive model forecasts have changed as the peak approaches, for comparison with the observed flows (to time now), and bands showing the best estimates at each lead time for the range of the forecast.

- 'Poor man's ensembles' – comparisons of flow forecasts for different models and rainfall inputs.

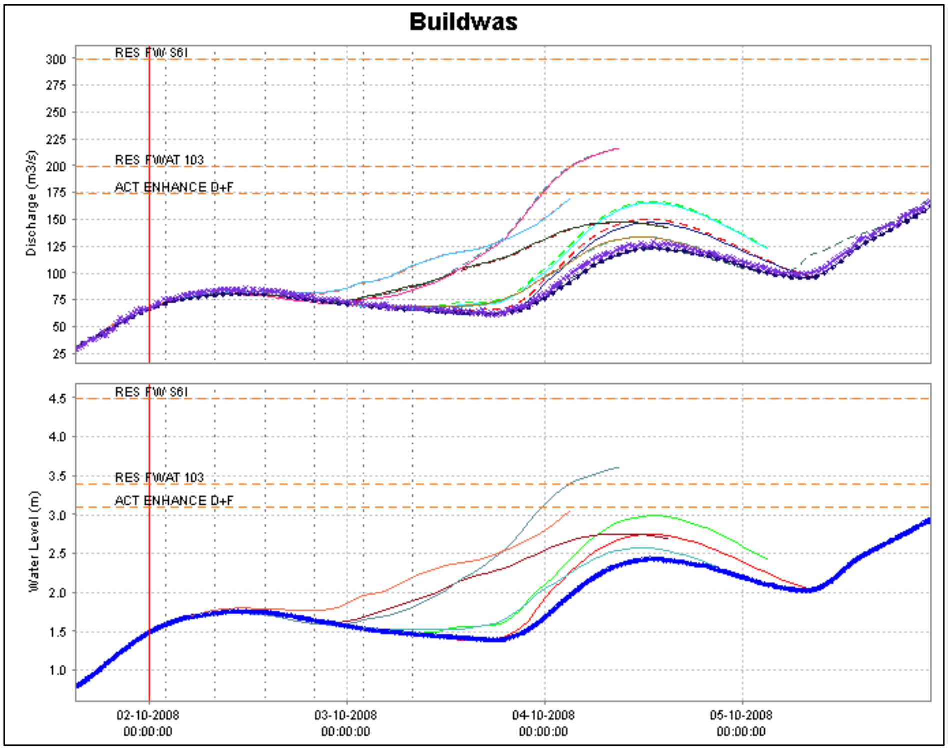


Figure 4.1 – Example of a time lagged ensemble plot provided by Midlands Region.

4.3 Regional and Area Studies

The following studies were noted as being possibly useful to the present study, and copies of the relevant reports were provided (when available):

- Rainfall–Runoff Model Sensitivity Tests – a study into the performance of the PDM components of an integrated catchment model, using stochastically generated rainfall, and sampling for different parameter sets. Runs have also been performed with and without rainfall to investigate the likely limits on forecast lead time.

- Rainfall Actuals Sensitivity Tests – regional studies into the performance of integrated catchment models using catchment average rainfall estimates based on raingauge data and radar rainfall data.

- Performance Measures/Level of Service studies – pilot tests for a catchment of performance measures such as POD and FAR, including variations in performance with lead time, and also examining performance with raingauge data, radar rainfall data, and assuming no future rainfall (with a region-wide study under development).

- Reservoir Influences – studies for a catchment which was affected during the Summer 2007 floods into the variations in forecast flows when including and excluding one or more reservoirs in the upper reaches of the catchment.
- More generally, it was noted that there may also be useful information within the model development reports for the four case studies which will be chosen for development during Phase 2 of the project (to be determined).

4.4 Other Relevant Studies

The following ongoing and recent national projects were noted as being potentially useful to this project:[4]

- FEWS User Day 16–17 October 2008 – a two day workshop to outline current and future developments in FEWS.
- Error Correction Workshop – an NFFS workshop held during October 2008, and related ongoing national study into the performance of error correction routines.
- Communication and Dissemination of Uncertainty – an R&D project which is close to completion and is considering the issues surrounding communication and dissemination of uncertainty in flood forecasts.
- Fast Response Catchments – a study into potential approaches to flood forecasting for fast response catchments, which was commissioned by an area office during 2008.
- Data Assimilation – ongoing research in the Netherlands on Kalman Filter and other approaches to updating of coastal forecasts. Also, ongoing studies by Proudman Oceanographic Laboratory.
- Coastal Flood Forecasting project – an ongoing R&D project to develop techniques for generating ensemble surge forecasts and probabilistic wave and wave overtopping forecasts.
- Skill Scores project – an ongoing project which seeks to recommend which performance measures should be routinely computed in NFFS (with appropriate enhancements to NFFS). Provisionally the list includes:
    - rainfall forecasts and radar rainfall actuals (0–3 hours, 3–6 hours, then 6 hourly blocks from 6 to 36 hours ahead);
    - catchment average rainfall from raingauges and radar rainfall;
    - forecast flows and observed flows (2, 6, 12, 24 and – where possible – 36 hours ahead).

Objective functions which will be used include bias, R^2, RMSE and RMSF, and threshold-based measures. A range of procedural and computational issues are currently under investigation (for example, the approach to event matching).

[4] Note that the Phase 1 report will refer to additional studies which we are aware of

5 Integrated Catchment Models

5.1 Introduction

Most regions now operate a number of integrated catchment models, and the aim of this section of the consultations was to discuss experience with developing and using these models, and any key sources of uncertainty which have been identified, and their relative importance in the accuracy of flood forecasts provided. Experience with real-time updating (data assimilation) was also of particular interest, including error correction and state updating.

The discussions were structured around the following three main questions and the remainder of this section summarises the main responses to those questions:

- What experience has been gained with the performance of integrated catchment models combining rainfall–runoff, flow routing and/or hydrodynamic components? (Section 5.2)
- What do you find are the most difficult/challenging aspects of model calibration (catchment rainfall estimation, permeable catchments, high flow ratings, representation of control structures etc)? (Section 5.3)
- Can you provide any descriptions or assessments of model performance which might be useful to this project? (Section 5.4)

5.2 Regional Experience

All regions reported experience with developing and using integrated catchment models. Table 5.1 summarises the types of models which were described as being currently in use or being tested with a view to implementation operationally.

Table 5.1 – Indicative summary of current uses of integrated catchment models for catchments where models are available.

Method	All Catchments	Many/Several Catchments	Some Catchments
PDM/ISIS	1	2	
PDM/KW/ISIS			2
PDM/KW		1	1
MCRM/DODO		1	
MCRM/DODO/ISIS			1
NAM/MIKE11		1	
PRTF/ISIS		1	1
PRTF/VPMC/ISIS			1
TCM/ISIS	1		

Several regions also described the use of rainfall–runoff models alone for fast response catchments and/or lower risk catchments, and the use of level to level correlations. The types of rainfall–runoff model currently in use included PDM, HYADES, MCRM, TCM, NAM and PRTF models.

The current approaches to real-time updating in each region were also discussed, and Table 5.2 summarises the methods which were reported:

Table 5.2 – Indicative summary of current uses of approaches to data assimilation for catchments where integrated catchment models are available.

Method	All Catchments	Most/Several Catchments	Some/A Few Catchments
FEWS ARMA	1	2	4
ARMA (various forms)		2	2
PDM state updating		2	1
Manual – snow			1
Manual – groundwater			1
Manual – soil moisture deficit		1	
Manual – gate settings			1
ISIS GAUGE			1
MIKE11 error correction			
MIKE11 state updating			1
PRTF updating		2	

5.3 Most Difficult/Challenging Aspects

During most meetings, this topic was largely covered within the initial discussions on sources of uncertainty (Section 2.1). The most difficult and challenging aspects which were noted might therefore be summarised as follows:

- Catchment Averaging of Raingauge Data – number of raingauges to include, weighting factors to use, relative performance of raingauge and radar rainfall based actuals.

- Validity of Rating Curves – seasonal influences, mobile bed levels, high flow/out of bank ratings, maintenance of ratings in models.

- Model Calibration (Hydrodynamic Models) – understanding/improving the performance of 'monolithic' models, deciding where/how/whether to implement real-time updating, seasonal influences, mobile bed levels, propagation of tidal influences upstream, representation of abstractions/discharges (if relevant), problems at model boundaries.

- Model Calibration (Rainfall–Runoff Models) – overall model conceptualisation/smallest catchment size, model performance for convective events and in moderate rather than intense rainfall events, most appropriate performance measures to use.

- Representation of Floodplain Storage – accuracy of survey data, understanding/representing flow paths on floodplains, roughness coefficient for in-bank/floodplain flows, real-time updating for floodplain storage, draining of the floodplain (in multiple events).

- Representation of Antecedent Conditions – representation of soil moisture during the transition from summer to winter conditions; real-time information/updating for snowmelt modelling; representation of groundwater influences, evaporation modelling.

- Representation of Ungauged Inflows – catchment conceptualisation (using a few large subcatchments or many smaller catchments), the general approach to estimating ungauged flows (scaling, parameter transfer, FEH-based etc).

- Influence of Structure Operations – representation of reservoir storage and operations, lack of real-time information on structure operations, differences between actual and design operating rules, operational problems (gate failures, blockages, flood defence breaches etc).
- Real-time Updating – allowing for ‘drift’ in baseflow conditions, updating in tidal conditions, updating when there are timing errors, the influence of poor data quality.
- Flood Mapping Module – run times, setting up and maintaining survey/defence heights etc.

As already noted in previous sections, a number of general issues were also discussed relating to data quality, data availability, catchment changes over time (ratings, channel characteristics, flood defences etc), model performance measures etc.

5.4 Model Performance Assessments

This topic was largely covered in the response to the question about recent regional and area studies (see Section 4.3), for which the main studies which were noted involved sensitivity tests to different types of rainfall inputs, and for different rainfall forecast lead times, investigations of the most appropriate performance measures to use (and regional application of those measures), and studies into the influence of reservoirs.

6 User Requirements

6.1 Introduction

The anticipated benefits of the project are as follows:

- Improved understanding of major sources of uncertainties in fluvial flood forecasting.
- Consistent and tested unified uncertainty framework with robust methods to quantify and reduce (where possible) uncertainty for integrated catchment models in fluvial flood forecasting from ‘end to end’.
- Lessons learned from applying and validating selected uncertainty techniques to case studies.
- Practical guidance on how to apply probabilistic fluvial flood forecasting operationally.

The main project outputs will include the case studies, the unified framework, guidelines on applying probabilistic techniques, and the project technical and summary reports. The format of the guidelines will be discussed as the project progresses, but some possible contents could include:

- Risk-based approaches for selection of appropriate techniques, accounting for catchment characteristics and forecast lead time.
- Short case studies of best practice in the form of Factsheets.
- Flowcharts and diagrams to guide users through the selection process.

The framework could also consider good practice issues such as the need to evaluate model performance after flood events (and recommended approaches), continuing model maintenance, and appropriate performance measures for assessing different aspects of model performance, such as skill scores and contingency measures (e.g. POD, FAR), skill of the mean forecast, the spread-skill relationship of the probabilistic forecast, the Brier score, ranked probability score, and vulnerability based measures. However, the focus of the review of performance measures will be on a small number which could be used to assess the probabilistic outputs on the case studies (and which would be practical to implement operationally). As with the Real-Time Modelling guidelines, the framework will provide clear guidance on the data requirements, benefits and limitations of each approach.

The discussions were structured around the following four main questions and the remainder of this section summarises the main responses to those questions:

- Of the suggestions above, which would you find useful in a guideline document on probabilistic fluvial flood forecasting? (Section 6.2)
- What formats do you find most useful in the currently available guidance on approaches to model development (flowcharts, case studies, itemised instructions, tips and advice, tables of strengths/limitations of approaches etc)? (Section 6.3)
- Do you have any experience of using the Real-time Modelling guidelines from 2002, and subsequent documents which have incorporated ideas from

those guidelines (AMS Work Instructions, Tips and Guidance notes etc)? (Section 6.4)

- What do you see as the main risk-based criteria to consider in selection of an appropriate modelling approach? (Section 6.5)

As part of the introduction to the topic, it was noted that preparation of the guidelines is one of the final tasks on the project (in 2010) and the topics which will be considered will depend to some extent on the findings from Phase 2 of the project. However, it is useful to have any early ideas on content and format as this may guide the development of the framework and guidelines in later stages of the project.

6.2 Content of the Guidelines

The discussions on this topic suggested that there was a general view that the guideline document should set out a national approach to how the selected probabilistic techniques should be implemented, with a general feeling (with one exception) that it should not be too prescriptive ('not a cookbook', 'guide not tell').

Several consultees also said that, ideally, the document should be linked to, or supplemented by, the NFFS case studies, so that a practical implementation can be viewed, and that this should be combined with generic 'lessons learned'. Also, that a balance needs to be kept between providing long lists of choices and options, and providing only a small number of approaches (or a single approach) which is not generally applicable to other catchments or regions. The various options and choices could also be linked to risk (see later). At some meetings, there was also some discussion about whether the guidelines would only be used a few times (for example, when learning new techniques), or referred to frequently for reference during model design and development.

In abbreviated form, some other points which were made included the following observations:

- Factsheets could be useful, and tables summarising the strengths and limitations of different approaches.
- Information on the staff resources to implement solutions would be useful, together with descriptive information on the likely benefits, and possible 'quick wins'.
- The document should not be too lengthy and should be easily accessible for quick reference to specific items.
- A checklist approach could be useful, highlighting issues to consider, with short guidance and pointers.
- More detailed technical background on specific techniques may be useful.
- Documents should be accessible but not too long, and may need to be aimed at two levels; model users and model developers.
- Advice on appropriate performance measures would be very useful.
- The likely system requirements should be noted, with 'health warnings' regarding model run times and post-processing requirements (if required).
- Sharing of best practice between regions is always useful.

- The document should set out a national approach to how probabilistic techniques should be implemented ('a framework for how to do it').
- The case studies and examples will be useful, maybe with full descriptions/factsheets available for further reading in more depth (a document to 'dip in to').
- The development of a 'tool kit' approach would be useful.
- In addition to using guidelines during the initial model design, there could also be advantages in identifying potential issues during the model build phase, and in planning long-term programmes for model development.
- The guidelines should be flexible enough to adapt to new approaches in future (and not a list of instructions).
- Being able to see the thought process in setting up the 4 case studies would be particularly useful.
- The guidance could also include protocols and checklists containing information on probabilistic issues and uncertainty.
- The guidance should be shared with modelling partners that develop the models.
- The guidelines could suggest a tiered approach, starting with a simple method, then working up to more complex methods.
- Guidance on combining ensemble flow forecasts with ensemble surge forecasts would be useful where this is relevant (although this is outside the scope of this project).
- The framework/guidelines should also consider how to combine ensemble flow forecasts from the Hydrological Modelling with Convective Rainfall (T46) study with the probabilistic outputs from this project.
- Descriptions of assessments of model performance for historical events would be very useful (sources of uncertainty, how to treat probabilistically etc).
- Although outside the scope of the present project, a new NFFS training module could be created for probabilistic forecasting, allowing for ongoing training (not just at the project workshop in Phase 3).
- Guidance on assessing the critical areas of uncertainty in models, and in targeting work where the maximum benefit is obtained, would be useful; for example, the choices between improving ratings versus improving radar forecasts versus more rainfall–runoff work on tributaries (written reply).
- The introduction of national 'what-if' scenarios should be avoided as there are always local variations, and the most appropriate scenarios would depend on local factors and the characteristics of the event (written reply).
- Whatever techniques are recommended, they should be able to generate probabilistic forecasts in a similar time to current approaches, for multiple rivers/forecasting points (written reply).

It was also noted that there may be regional differences in what is useful; for example, between regions with an extensive coverage of integrated catchment models, where the current focus is on improving the performance of existing models, and regions with only a few models, where the focus is on extending the coverage of models.

A number of other, more general, comments were made on the guidance and other information currently available for the Project Board to consider (but which were outside the scope of this project and so are not reported here). The issues of providing guidance on the generation of rainfall forecasts, and the communication and dissemination of information on uncertainty, were also raised at some meetings (and in a written submission), but are being considered on other projects, and are outside the scope of the present study (although the uncertainty framework will consider how uncertainty in rainfall forecasts might be combined with other sources of uncertainty).

6.3 Currently Available Guidance

This topic generally did not generate much discussion and, in abbreviated form, the main points and observations which were made were:

- An entry for key words (a 'prompt list'), with links to definitions, as found in some current documents, would be useful.
- The comments on flowcharts and checklists ranged from the view that these are useful through to flowcharts not being favoured and checklists/protocols not widely liked.
- Tips and Guidance documents are useful for providing additional information.
- More detail on the technical aspects would be useful to supplement current guidance which seems weighted towards process/policy aspects.

6.4 Real-time Modelling Guidelines

The Real-time Modelling Guidelines were produced during 2001 and 2002 and provided a risk-based approach to the selection of modelling techniques for fluvial flood forecasting problems. In addition to the main document, some aspects have subsequently been included in work instructions and other documents, and in model development specifications. In abbreviated form, the points which were made on the guidelines included:

- The mixture of flow charts and case studies in the guidelines has been found to work well in the past. It was also felt that this document benefits from being brief and non-technical.
- The guidelines contained guidance on the staff time inputs required, and it would be useful to adopt a similar approach in the new document.
- The guidelines were useful and need a relaunch, and could be repackaged. This would provide the opportunity to make sure that everyone has the same information available to them. It was felt that there is too little technical content in the other currently available guidance.
- The guidelines were produced at a time when each region operated its own models and forecasting systems, and pre-dated the more consistent approach which has become possible with the introduction of the NFFS. The guideline document was therefore 'of its time' and requires updating.
- The economic (cost-benefit) aspects, although seen as important when the guidelines were produced, are perhaps less important nowadays, and therefore any new document should focus mainly on technical issues.

6.5 Risk Based Criteria

One option for the guidelines and uncertainty framework is that the choice of method could be tailored to the level of risk, in addition to accounting for catchment characteristics and forecast lead time requirements. The distinction between computationally complex methods and scientifically complex approaches was also discussed (with the former possibly being easy to implement, but taking a lot of computer time). In abbreviated form, the points which were made on a risk-based approach included:

- The methods could vary in depth and extent according to risk, perhaps as estimated from the 'Anglian Matrix' (HHH etc), and/or linked to numbers of properties and risk to life (these often go together).
- It is assumed that the options presented will link into simple measures of risk e.g. a risk matrix (HHH) approach.
- There was some doubt about the value of a risk based approach unless starting out model development work from scratch (with no model for the catchment(s) of interest). Alternatives to the risk matrix values (HHH etc) could include properties at risk, and cost-benefit. Technical feasibility is also important to consider (e.g. on fast response catchments).
- A risk based (HHH) approach seems sensible. Catchment and data issues are important as well, e.g. the location of Forecasting Points in the catchment, the locations of structures, lead time requirements, data availability and quality etc.
- Possibilities for a risk based approach could include the number of properties at risk and Risk Matrix values (HHH etc); also, the presence of a Severe Flood Warning. Technical feasibility is also important to consider e.g. on fast response catchments.
- One more general point was the suggestion that risk and uncertainty in model development should be identified/quantified at the feasibility/inception stage, and should ideally be part of the criteria for identifying whether the model development project is worthwhile, highlighting the various trade-offs between model complexity, data availability, and model uncertainty. This would help to mitigate against the development of a model with high uncertainty being developed in a high risk area, as this could have unfortunate consequences.

7 Conclusions

The consultation meetings and phone conferences allowed discussions to be held with regional flood forecasting teams early in the project to determine gaps in available guidance, to assess end user requirements, to discuss ideas for possible case studies, and to allow for refinements in project scope. The following topics were considered:

- Sources of uncertainty in Integrated Catchment Models
- The selection criteria for case studies during Phase 2 of the project
- Ongoing regional studies (if available) on sources of uncertainty
- Experience with the performance of integrated catchment models
- User requirements for the Real-Time Modelling Guidelines and other project outputs.

The consultations were performed by telephone and meetings during the period 18 December 2008 to 16 January 2009. Approximately 25 Regional Flood Forecasting and Area Flood Warning Duty Officers participated in the consultations, with written contributions received from an additional two people. With the exception of the final meeting, a Project Board member participated in each meeting. The majority of meetings lasted from 2 to 3 hours, with the shortest lasting just under 2 hours, and the longest lasting about 4 hours.

The consultations have provided a good picture of the current usage of integrated catchment models within the Environment Agency, and the sources of uncertainty in models of this type, and the related benefits. Key issues which have been highlighted include uncertainties in both the hydrological and hydrodynamic components of models, with more than half of the issues raised concerning the hydrological aspects (and rating curves). Some useful indications have also been provided of what would be useful in the guidelines and uncertainty framework which will be produced later in the project. Approximately 24 catchments were also suggested as potential case studies.

The findings from the consultations will inform planning for Phase 2 of the project, and selection of the four case studies to be considered. This report will also be discussed at the Phase 1 Project Workshop which is planned to be held in Solihull on 16 March 2009.

Glossary

Term	Description
Agency Management System	A series of Work Instructions and related documents which clarify procedures, targets etc for Environment Agency staff
Baseflow	The stream flow component arising from water moving through the aquifer to the stream channel
Catchment model	A model (or models) using observations of rainfall and/or upstream flows and/or levels to forecast flows and/or levels at a point within a Flood Warning Area (typically a gauging station)
Damage reduction	The amount of pre-flooding action that can be taken to reduce the cost of the flooding event expressed as a percentage factor, taking into consideration the lead time of the warning (i.e. the length of time between when a warning was issued and when flooding occurs) that allows the pre-flooding action to be carried out
DODO	Flow routing model (Midlands Region)
Flood Risk Area	An area at risk from flooding which may or may not have an existing warning service
Flood Warning Area	A proportion of the floodplain containing a community at risk from flooding which is provided with an appropriate flood warning service as per the Risk Decision Box
HYADES	Unit hydrograph model (rainfall–runoff)
HYRAD	A weather radar processing and display system produced by CEH Wallingford
ISIS	1D hydrodynamic modelling package (HR Wallingford)
Lead time	The maximum time ahead which a model or rainfall forecast can predict flows or rainfall. Also flood warning lead time
MIKE11	1D hydrodynamic model (DHI)
MNP	Netherlands Environmental Assessment Agency
MOGREPS	An ensemble Numerical Weather Prediction modelling approach (currently 0–36 hours) developed by the Met Office
NAM	Rainfall–runoff model (DHI)
Nimrod	Deterministic nowcasting technique now replaced by STEPS (Met Office)
Rainfall actuals	Observations of rainfall occurring at present using raingauges or radar
Rainfall–runoff model	A model which converts observed or forecast rainfall into estimated river flows
Routing model	A model which translates flows from the upstream to the downstream end of a river reach allowing for floodplain effects, tributary inflows etc
Service effectiveness	The proportion of flood serviced properties which were sent a timely, accurate and reliable flood warning
STEPS	Nowcasting technique that provides deterministic forecasts (0 to 6 hours) operationally now with ensembles planned for release in 2009 (Met Office)
Trigger	A river level above which a flood warning is issued (or considered); now called Threshold
Ungauged catchment	A catchment with no river level recorder
Updating	The use of observed river levels or flows to attempt to improve a forecast by correcting the forecast to better match the observed values

List of abbreviations

Term	Meaning/definition
AMS	Agency Management System
ARMA	AutoRegressive Moving Average
BMA	Bayesian Model Averaging
BPO	Bayesian Processor of Output
CAPTAIN	Computer Aided Program for Time-series Analysis and Identification of Noisy systems
CEH	Centre for Ecology & Hydrology (Wallingford)
CHPS	Community Hydrological Prediction System
CWI	Catchment Wetness Index
DBM	Data Based Mechanistic
D-PHASE	Demonstration of Probabilistic Hydrological and Atmospheric Simulation of flood Events
ECMWF	European Centre for Medium-range Weather Forecasts
EFAS	European Flood Alert System
EFFS	European Flood Forecasting System
EKF	Extended Kalman Filter
EnKF	Ensemble Kalman Filter
EPS	Ensemble Prediction System
FAR	False Alarm Rate
FEWS	Flood Early Warning System
FOEN	Federal Office for the ENvironment (Switzerland)
FREE	Flood Risk from Extreme Events
FRMMS	Flood Risk Management Modelling Strategy
FRMRC	Flood Risk Management Research Consortium
G2G	Grid-to-Grid Model (CEH Wallingford)
GLUE	Generalised Likelihood Uncertainty Estimation
HEPEX	Hydrologic Ensemble Prediction EXperiment
KF	Kalman Filter
KW	Kinematic Wave flow routing model (CEH Wallingford)
MAP	Mesoscale Alpine Programme
MCRM	Midlands Catchment Runoff Model (rainfall–runoff)
MOGREPS	Met Office Global and Regional Ensemble Prediction System
MORECS	Met Office Rainfall and Evaporation Calculation System
MOSES	Met Office Surface Exchange Scheme
NFFS	National Flood Forecasting System
NWP	Numerical Weather Prediction
NWS	National Weather Service (USA)
PDM	Probability Distributed Model (CEH Wallingford)
PFFW	probabilistic flood forecasting and warning
POD	Probability of Detection
PRTF	Physically Realisable Transfer Function model (rainfall–runoff)
PSM	Penman Store Model (incorporatesTCM and IEM)
QPF	Quantitative Precipitation Forecast
RMSE	root mean squared error
STFS	Storm Tide Forecasting Service (Met Office)
TCM	Thames Catchment Model (rainfall–runoff) (provided in PSM along with IEM)

Term	Meaning/definition
TF	transfer function
UNEEC	UNcertainty Estimation based on local Errors and Clustering
VPMC	Variable Parameter Muskingum-Cunge flow routing model
WMO	World Meteorological Organization
XEFS	eXperimental Ensemble Forecasting System

Would you like to find out more about us, or about your environment?

Then call us on

08708 506 506* (Mon-Fri 8-6)

email

enquiries@environment-agency.gov.uk

or visit our website

www.environment-agency.gov.uk

incident hotline 0800 80 70 60 (24hrs)
floodline 0845 988 1188

* Approximate call costs: 8p plus 6p per minute (standard landline).
Please note charges will vary across telephone providers

Environment first: This publication is printed on recycled paper.