defra
Department for Environment Food and Rural Affairs

Environment Agency

delivering benefits through evidence

Hydrological modelling using convective scale rainfall modelling – phase 1

Project: SC060087/R1

Flood and Coastal Erosion Risk Management Research and Development Programme

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 Evidence Directorate 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: 978-1-84911-179-9

Email: fcerm.evidence@environment-agency.gov.uk.

Further copies of this report are available from our publications catalogue: http://publications.environment-agency.gov.uk or our National Customer Contact Centre: T: 08708 506506
E: enquiries@environment-agency.gov.uk.

Dissemination Status:
Released to all regions
Publicly available

Keywords:
Flood forecasting, hydrological modelling, numerical weather prediction, probabilistic forecasting, uncertainty

Research Contractor:
WL Delft Hydraulics (now part of Deltares) and Centre for Ecology & Hydrology (CEH) Wallingford

Environment Agency's Project Manager:
Simon Hildon, Evidence Directorate

Theme manager:
(acting) Stefan Laeger, Incident Management and Community Engagement Theme

Project Number:
SC060087

Product Code:
SCHO0210BRYR-E-P

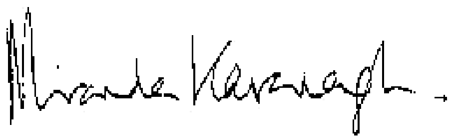
Evidence at the Environment Agency

Evidence underpins the work of the Environment Agency. It provides an up-to-date understanding of the world about us, helps us to develop tools and techniques to monitor and manage our environment as efficiently and effectively as possible. It also helps us to understand how the environment is changing and to identify what the future pressures may be.

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

The Research & Innovation programme focuses on four main areas of activity:

- **Setting the agenda**, by informing our evidence-based policies, advisory and regulatory roles;
- **Maintaining scientific credibility**, by ensuring that our programmes and projects are fit for purpose and executed according to international standards;
- **Carrying out 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.



Miranda Kavanagh

Director of Evidence

Executive summary

Hydrological models have the capability to provide useful river flow predictions and flood warnings. The aim of this project, 'Hydrological Modelling using Convective Scale Rainfall Modelling', is to investigate which models and associated computational methods would allow best use of the latest Met Office developments in numerical weather prediction (NWP). Two recent enhancements in particular offer interesting opportunities and open the door to the use of probabilistic flood forecasting. These two developments are:

- operation of the nowcasting system STEPS (Short Term Ensemble Prediction System) at 2 km resolution;
- a new system for longer term numerical weather prediction called MOGREPS (Met Office Global and Regional Ensemble Prediction System).

The three-phase project is concerned primarily with:

- how to use high resolution (convective scale) rainfall forecasts effectively for flood forecasting;
- how to make operational the use of ensembles of numerical weather prediction (MOGREPS) in flood forecasting and warning within the Environment Agency's National Flood Forecasting System (NFFS).

This report presents the results of Phase 1 (inventory and data collection) and forms the starting point for the rest of the project – Phase 2 (pilot case study).

Phase 1 involved:

- examining the suitability of three recent storm events as the pilot case study for Phase 2;
- obtaining detailed numerical weather predictions from the Met Office's Joint Centre for Mesoscale Meteorology (JCMM) in Reading;
- compiling an inventory of hydrological models (primarily rainfall-runoff models) suitable for predicting runoff generated by intensive storm events;
- configuring and calibrating four models (PRTF, PDM, G2G, REW) for transformation of high resolution rainfall predictions into accurate flood forecasts;
- carrying out an analysis of the use of high resolution NWP forecasts and the generation of 'pseudo' ensembles;
- determining the procedures and performance indicators for assessing model and ensemble performance (e.g. skill scores) when applied to the pilot case study;
- selecting two Environment Agency Regions (North East and Thames) to act as pilots for ensemble forecasting in NFFS.

The extreme weather event at Boscastle in north Cornwall on 16 August 2004 was selected as the pilot case study for Phase 2.

At the end of Phase 3 (analysis and verification), overall conclusions will be drawn on the general benefits of using high resolution NWP as input into a hydrological model for flood forecasting. In addition, a possible approach using the hydrological models – and calibration and computation methods – will be formulated.

Contents

List of tables and figures

1 Introduction

The Met Office continuously seeks to enhance its numerical weather prediction capability. For example:

- In the very near future, the nowcasting system STEPS[1] will become operational at 2 km resolution.
- For longer term numerical weather prediction, the Met Office has developed a new system called MOGREPS[2] which uses a coarser model resolution of 24 km.

Both systems will be run in ensemble mode.

These developments offer interesting opportunities for the Environment Agency and open the door to the use of probabilistic flood forecasting. However, operational research is required to realise the potential benefits of these developments for the Environment Agency's flood warning service, i.e. its National Flood Forecasting System (NFFS).

In addition, research is aiming to improve the prediction of convective events by using much finer grid sizes. The Storm Scale Numerical Modelling project[3] examined the ability of the new convective scale configuration of the Met Office's Numerical Weather Prediction (NWP) model to predict thunderstorm rainfall. It concluded that, if suitable post-processing is applied to the output, a substantial gain in capability will be achieved in changing from the current 12 km grid model to a 1 km grid. Even changing to a 4 km grid should give better results.

Hydrological models have the capability to provide useful river flow predictions and flood warnings provided the rainfall information with which they are supplied is sufficiently accurate. They have generally been used with raingauge data, radar analyses or extrapolated forecasts. More recently, longer term NWP model results have been used.

When introduced operationally, the rainfall prediction methods developed in the Storm Scale Numerical Modelling project will provide more accurate forecasts of intensive rainfall resulting from convective storms. With such data available as input to hydrological models, it should be possible to predict the risk of flooding more accurately and with longer lead times. However, the potential benefits for operational flood warning will only be fully realised if appropriate hydrological modelling concepts are applied.

The aim of this project ('Hydrological Modelling using Convective Scale Rainfall Modelling') was to investigate which hydrological model concepts and associated computational methods will allow the best use to be made of the latest Met Office developments in NWP. The project is primarily concerned with:

- making operational the use of ensemble data generated by the Met Office's regular weather models;
- considering the future potential of convective scale rainfall predictions.

The project has three phases:

- Phase 1 Inventory and data collection
- Phase 2 Pilot

---

[1] Short Term Ensemble Prediction System
[2] Met Office Global and Regional Ensemble Prediction System
[3] Joint Defra and Environment Agency Flood and Coastal Erosion Risk Management R&D Programme

- Phase 3 Verification and synthesis.

This report presents the results of Phase 1 and forms the starting point for the rest of the project. As such it provides the basis and direction for the pilot case study in Phase 2.

The project was undertaken for the Environment Agency by Deltares[4] based at Delft in the Netherlands in collaboration with the Centre for Ecology & Hydrology (CEH) based at Wallingford in the UK.

[4] Deltares (http://www.deltares.nl/xmlpages/page/deltares_en) was formed on 1 January 2008 from WL Delft Hydraulics, GeoDelft (subsurface and groundwater unit of TNO) and parts of Rijkswaterstaat to create a new independent Dutch institute for national and international delta issues. The contract for this project was awarded to WL Delft Hydraulics in March 2007.

2 Project approach

2.1 Project objectives

The project examined the following key issues:

- **How to use high resolution (convective scale) rainfall forecasts effectively for flood forecasting.** The project objectives in this respect are:
  - to identify the best methods of providing input to hydrological models from the output of convective scale NWP models;
  - to develop methods for improving the short-range prediction of flooding associated with thunderstorms by using post-processed output from high resolution NWP models as input into hydrological models to generate an ensemble of forecast scenarios in order to improve forecast warning.
- **How to make operational the use of ensembles of numerical weather prediction (MOGREPS) in flood forecasting and warning within the Environment Agency's National Flood Forecasting System (NFFS).** The project objectives in this respect are:
  - to identify an approach to probabilistic flood forecasting using ensembles of numerical weather predictions;
  - to make operational the use of ensembles of numerical weather predictions in a test environment running NFFS.

The main aim of the project was to develop a practical approach to make its results operational in the forecasting systems used by the Environment Agency (and potentially SEPA). A secondary aim was to contribute to the practical use of probabilistic flood forecasts in decision-making related to flood warning.

During the project the research team:

- looked at ways in which high resolution NWP model precipitation forecasts can be used as input to hydrological models for flood warning;
- examined the potential usefulness of such a system;
- made recommendations as to how improvements could be made and difficulties overcome.

The findings will be used by the Environment Agency to provide more accurate and reliable warnings of flood events.

2.2 Using convective scale rainfall forecasts in the National Flood Forecasting System

Progress in achieving this objective is discussed in Part 1 of this report.

A methodology for using convective scale rainfall predictions for flood forecasting was developed for testing in Phase 2 of the project for one pilot case study. This pilot featured a convective storm event over an area for which hydrological modelling is feasible and looked mainly at:

- how to model the response for such events;
- how to use the forecast information in flood warning.

2.2.1 High resolution numerical weather prediction

Detailed numerical weather predictions were obtained from the Met Office's Joint Centre for Mesoscale Meteorology (JCMM) in Reading, which is active in research on numerical modelling of convective scale events.

For the potential test cases (see Section 3) it was necessary to run the high resolution configuration of the Met Office Unified Model (UM). Model output data were used where available.

The decision on which model resolution to use was based on advice from JCMM. In principle it would be interesting to test a series of model resolutions, as previous studies have shown that the forecast ability of convective storms improves considerably with increasing NWP resolution.

In order to represent the positional uncertainty that comes with the high resolution rainfall predictions, 'pseudo' ensembles were created.

2.2.2 Hydrological modelling

A basic inventory of hydrological modelling concepts suitable for predicting runoff generated by intensive rain storms has been compiled. The inventory was prepared on the basis of available literature and concentrated on model algorithms available for operational use.

The inventory is mainly concerned with rainfall-runoff models. Routing and hydrodynamic models are considered less relevant within the framework of this research as they rely on accurate predictions of lateral inflows with rainfall-runoff models. The project compared the modelling concepts currently applied in NFFS with distributed hydrological models.

Modelling concepts currently applied in the areas of the potential pilot case studies are:

- transfer function model – Physically Realisable Transfer Function Model (PRTF);
- lumped conceptual hydrological model – 'standard' Probability Distribution Model (PDM), Thames Catchment Model (TCM), modified conductive rock matrix (MCRM) or North American Mesoscale Model (NAM).

Distributed modelling concepts are, by their nature, more suitable for computing the spatially distributed response to convective scale storm events. The following concepts were therefore tested in this study:

- the distributed conceptual hydrological model called Grid-to-Grid (G2G);
- the more physically based distributed hydrological model called Representative Elementary Watershed (REW).

Most of the analysis was carried out in the near operational environment of NFFS. More details of the currently applied models and distributed modelling concepts are given in Section 4.

All modelling concepts tested must be able to run in Delft-FEWS.[5] Delft-FEWS module adapters are available for all models currently used in NFFS. There is already an adapter for the REW model and a new module adapter for the G2G model will be developed during the early stages of Phase 2.

Appropriate geographical datasets were collected for the configuration of new hydrological models for the pilot catchment. The geographical datasets are not relevant where existing forecasting hydrological models are used, i.e. transfer functions or lumped hydrological models like PDM or TCM.

The model calibration is based on a continuous dataset with rainfall events; associated observed radar data (grids) and raingauge measurements were collected. Spatial observed radar data form the basis of the calibration once corrected with the help of raingauge data using available HyradK[6] functionality. In order to be able to run such ground truth corrections operationally in the future, a FEWS adapter was developed for use in Phase 2 of this project.

The model calibration strategy aims to properly represent flow generated under convective storm conditions. The calibration is carried out partly automatically and partly manually, using predefined criteria where possible.

Models of a conceptual or physics-based form have, by their very nature, strong parameter interdependence. Therefore a combination of manual estimation (supported by interactive visualisation tools) and automatic estimation of sub-sets of parameters has been found to work best. The calibration strategy may differ for different models but encompasses a set of agreed performance measures (including formal objective functions and visual hydrograph plots).

A number of performance measures for assessing deterministic and probabilistic forecasts are available. These were considered during the project along with any new ones developed.

How best to characterise uncertainty in model structure, initial states and parameter estimates was considered when developing and trialling probabilistic flood forecasting methodologies.

2.2.3 Analysis

It will be necessary to configure the processing of high resolution NWP data and the running of the hydrological models applied in the pilot case study into a test set-up of NFFS.

Flood forecasts will be produced using NFFS in order to stay as close as possible to the Environment Agency's regular forecasting environment. The required changes to the NFFS configuration will be added to the current regional configuration for the Environment Agency Region where the pilot case study is located. This test configuration will be installed on a system in Delft accessible to Environment Agency staff.

For the test cases outlined in Section 3, rainfall products will be generated containing multiple forecast scenarios from the high resolution NWP output. The rainfall products will be fed into the hydrological models to produce probabilistic forecasts within NFFS (following current forecasting procedures as much as possible). The forecasts will be produced and analysed only for the period covered by the pilot case study.

The 'raw' hydrological forecast data will be processed to form probabilistic forecasts and associated information. Existing presentation methods will be developed for presenting

[5] Deltares' Flood Early Warning System (see http://www.wldelft.nl/soft/fews/int/index.html)

[6] HYRAD radar hydrology kernel

spatial distributed forecasts and probabilistic forecasting data. These methods will be used to represent the results of the various methods applied to forecast convective storms on the basis of high resolution NWP data.

The performance of the hydrological predictions from the high resolution NWP output will be analysed with regard to:

- the impact of the applied hydrological model structure;
- the resolution of the NWP forecast data used.

The question of whether post-processing of NWP data has an impact on the flood forecasts will also be investigated. Objective functions (performance indicators) will be used to evaluate the forecast quality for the various combinations of factors.

2.2.4 Verification

The methods developed in Phase 2 will be applied in Phase 3. Data processing and analysis will be applied to a single 'verification' basin to test the general applicability of the approach. At present the scope is roughly defined and may be adjusted significantly. The applied methodology will be fine-tuned based on the outcome of Phase 2, i.e. steps that clearly do not seem to contribute to improved forecasts will be excluded. This refined approach will be applied to the selected verification basins. The project will run through the same sequence of steps as in Phase 2.

At the end of the verification phase, overall conclusions will be drawn on the general benefit of using high resolution NWP as input into a hydrological model for flood forecasting. In addition, a possible approach using the hydrological models – and calibration and computation methods – will be formulated.

2.3 Operational implementation of ensemble forecasting

Progress in achieving this objective is discussed in Part 2 of this report.

Making operational the use of ensembles generated by MOGREPS will be carried out initially for two Environment Agency Regions selected during Phase 1. Criteria for selection include a serious interest in probabilistic forecasting as the Region's forecasting team will be more intensively involved in the pilot than teams from other Regions. Preferably these will be other Regions than the one where pilot work on forecasting of floods generated by convective storms is carried out.

In a later stage of the project, the use of ensembles generated by STEPS may also be configured or made into a blended product.

Ensemble forecasting will be configured in NFFS for the selected Regions. The configuration will be based on the current configuration of NFFS. No distributed models will be run in this test case. The configuration will include:

- importing and processing of NWP ensembles (from MOGREPS);
- ensemble runs of forecasting models;
- data displays including statistical analyses;
- performance indicators focused on testing probabilistic forecasting skill.

A test environment was set up in Delft in which prototypes of the systems developed within the framework of this project could be run. A limited number of Environment Agency staff were given access to this live system to become acquainted with the project deliveries via a virtual private network (VPN). A data feed from the Met Office to Delft was set up with a duplicate feed to the Environment Agency's test HYRAD system at CEH Wallingford.

The potential and benefits of using NWP ensembles for flood forecasting will be assessed during Phase 2 at a workshop with the limited group of scientists and forecasters involved in the pilot application. Based on the outcome of this workshop, adjustments may be made to the configuration prior to presenting the results to a larger audience in a feedback workshop.

The workshop will complete Phase 2 of the project.

2.4 Project overview

The project overview prepared following Phases 2 and 3 will:

- examine how to improve flood forecasting on the basis of convective scale weather forecasts in the future;
- make recommendations about future steps and research;
- include a projection of how the project results could be used in the operational forecasting environment run by the Environment Agency;
- present an implementation plan for configuration and roll-out of ensemble forecasting in NFFS.

The actual operational implementation of ensemble forecasting is **not** part of this project.

Part 1: Using convective scale rainfall forecasts in NFFS

3 Selection of pilot test case

The first step was to select a storm event for the pilot case study planned for Phase 2 to test convective scale rainfall predictions.

Following discussions with Nigel Roberts of JCMM, the following convective storm events were identified as potentially interesting as a test case for Phase 2:

- south Wales, 10 May 2006
- west London, 3 August 2004
- Boscastle, 16 August 2004

The locations affected are highlighted in Figure 3.1.

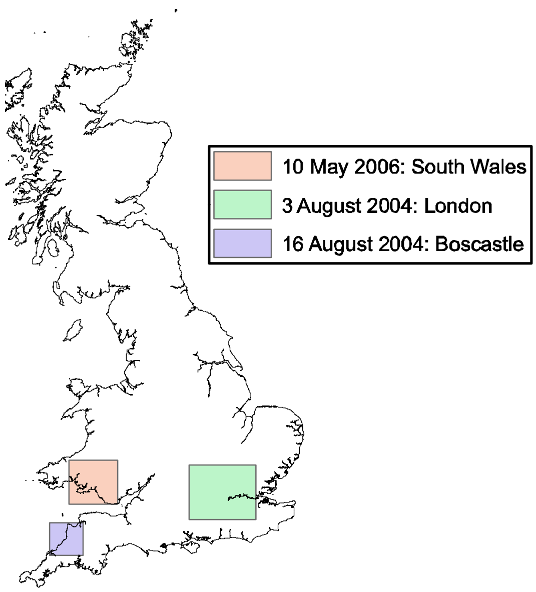


Figure 3.1 Locations of the three case study events.

The following analysis draws on:

- radar data available to CEH via its HYRAD backup discs;
- river flow data available to CEH/Delft;
- published sources.

Investigations on hydrological response and available data were carried out before making a final choice. One criterion is that the pilot area forms part of the current NFFS configuration.

3.2 Event of 10 May 2006 in south Wales

On 10 May 2006 a sequence of convective cells travelled westwards affecting the region between Reading and south Wales. In particular, the area around Pontardawe was badly affected with some flooding of homes reported.[7]

The sequence of storms lasted for around nine hours, starting at 15:00 (all times GMT). Figure 3.2 shows the accumulation of five-minute Nimrod composite radar data over this period. Gauged catchments in the vicinity are listed in Table 3.1.

The most significant catchment totals were for the rivers Ogmore at Bridgend, Dulais at Cilfrew and Tawe at Ynystanglws (gauging stations 58005, 58008 and 59001). A majority of the rainfall for each catchment fell within a two-hour period. Hyetographs for these catchments are presented in Figure 3.3 and show the eastern-most catchment (Ogmore) receiving the earliest rainfall.

Subsequent inspection of the river flow data through Delft-FEWS reveals that none of the water level triggers at the gauged sites were exceeded during the event. Despite the local flooding and meteorological significance, this case study therefore has a limited interest from a flood forecasting point of view.

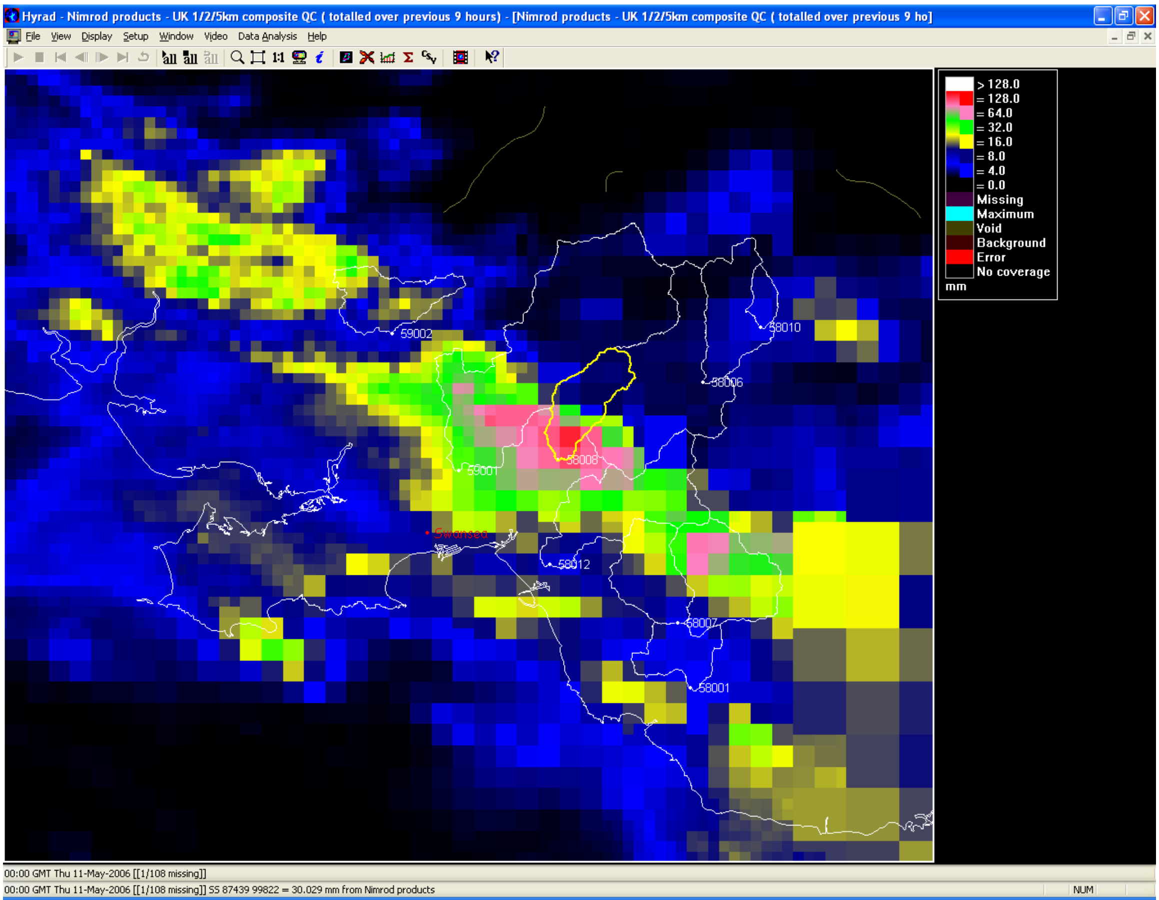


Figure 3.2 Accumulations of five-minute Nimrod radar data for the nine-hour period starting at 15:00 on 10 May 2006.

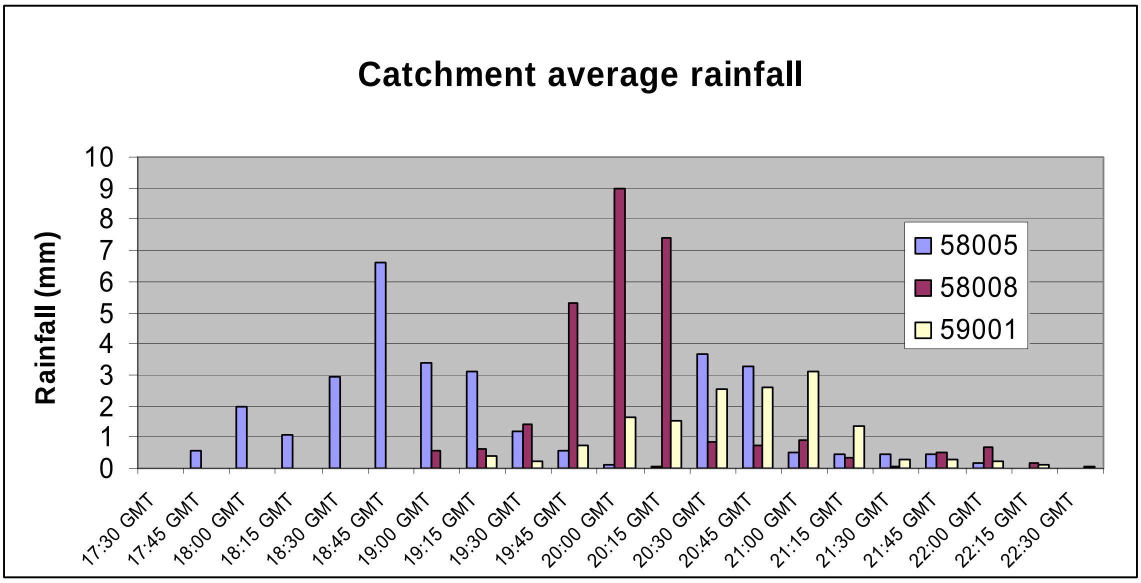
[7] 'Storms blamed for balloon crash', BBC News, 11 May 2006 (http://news.bbc.co.uk/1/hi/wales/4760693.stm)

Table 3.1 Details of gauging stations in the vicinity of the storm.

Number	Name	Location (NGR)	Area (km^2)
59001	Tawe at Ynystanglws	SS 685998	227.7
59002	Loughor at Tir-y-dail	SN 623127	46.4
58001	Ogmore at Bridgend	SS 904794	158.0
58002	Neath at Resolven	SN 815017	190.9
58005	Ogmore at Brynmenyn	SS 904844	74.3
58006	Mellte at Pontneddfechan	SN 915082	65.8
58007	Llynfi at Coytrahen	SS 891855	50.2
58008	Dulais at Cilfrew	SN 778008	43.0
58009	Ewenny at Keepers Lodge	SS 920782	62.8
58010	Hepste at Esgair Carnau	SN 969134	11.0
58012	Afan at Marcroft Weir	SS 771910	87.8

NGR = National Grid Reference

(a) 15 minute totals



(b) Cumulative rainfall

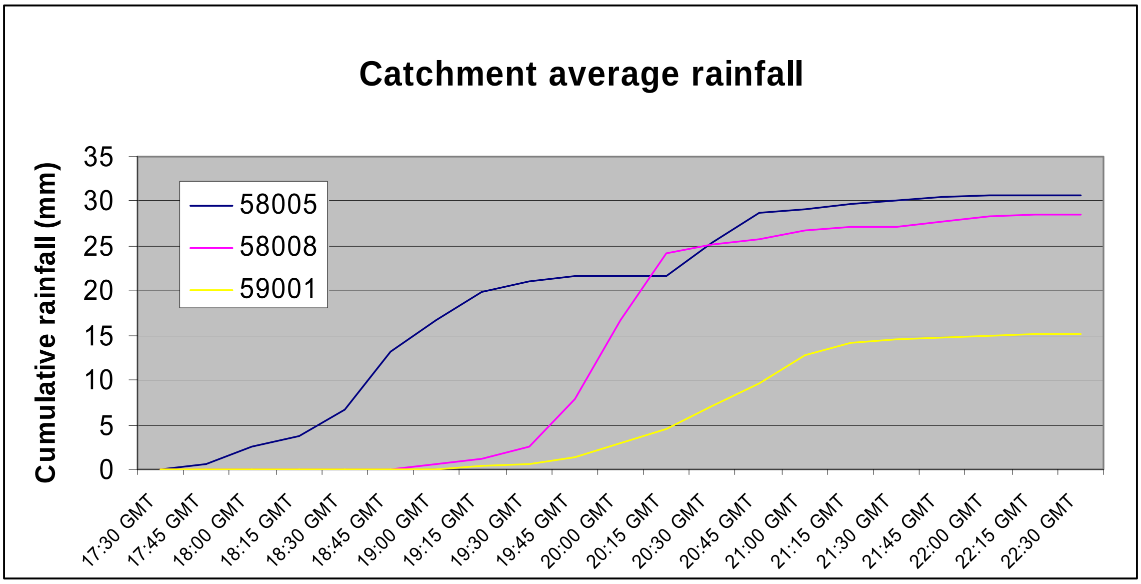


Figure 3.3 Hyetographs of catchment average rainfall using Nimrod composite radar data for selected catchments: Ogmore at Brynmenyn (58005), Dulais at Cilfrew (58008), Tawe at Ynystanglws (59001).

3.3 Event of 3 August 2004 in London

Table 3.2 shows the information obtained from Thames Region for the event in west London in August 2004.

Table 3.2 Details of gauging stations in the vicinity of the storm.

Name	File	LocationIdXML	LocationIdFews	Parameter
Wembley	Wembley.xml	3839TH	3839TH	Flow
Northolt (Yeading East)	Northolt.xml	3629TH	3629TH	Flow
Brent (Monks Park)	Monks_Park.xml	3850TH	3850TH	Flow
Hendon (Hendon Lane)	Hendon_Lane.xml	3809TH	3809TH	Flow
Gutteridge Wood	Gutteridge_Wood.xml	3625TH	3621TH	Flow
Cranford (Cranford Park)	Cranford_Park.xml	3660TH	3660TH	Flow
Greenford (Costons Lane)	Costons_Lane.xml	3870TH	3870TH	Flow
Colindale (Colindeep Lane)	Colindale.xml	3829TH	3829TH	Flow
Brent Cross	Brent_Cross.xml	3820TH	3820TH	Flow

The location of most of these gauges is shown in the map in Figure 3.4. Thames Region has developed TCM forecasting models and corresponding catchment average precipitation (using gauges and radar) for the gauging stations at Costons Lane, Monks Park, Wembley and Colindeep Lane. The maps in Figure 3.5 show the catchment areas defined for each of these stations.

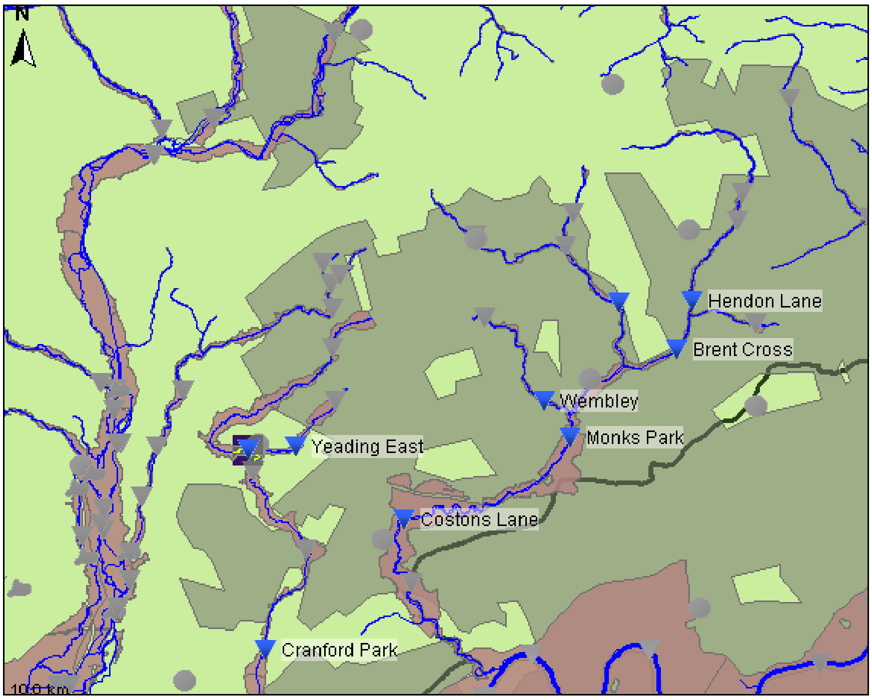


Figure 3.4 Location of gauging stations.

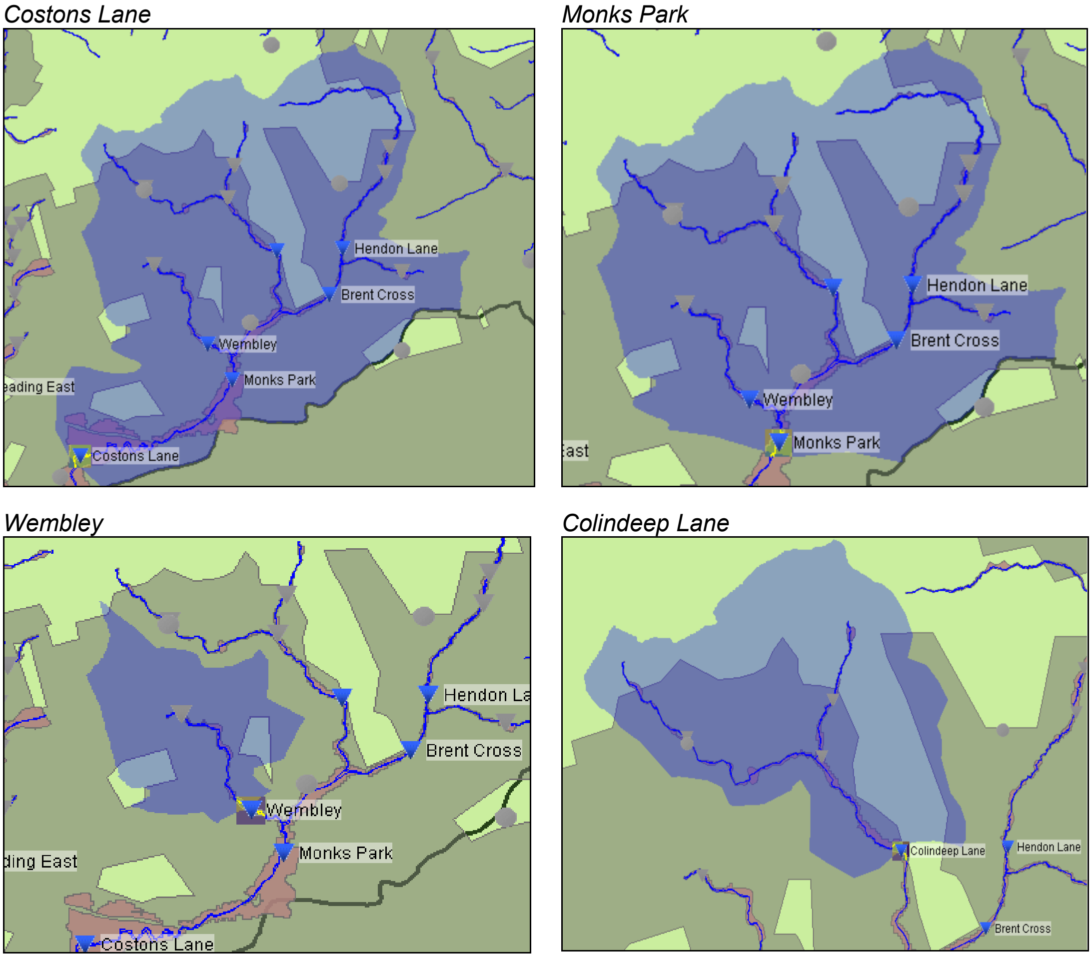
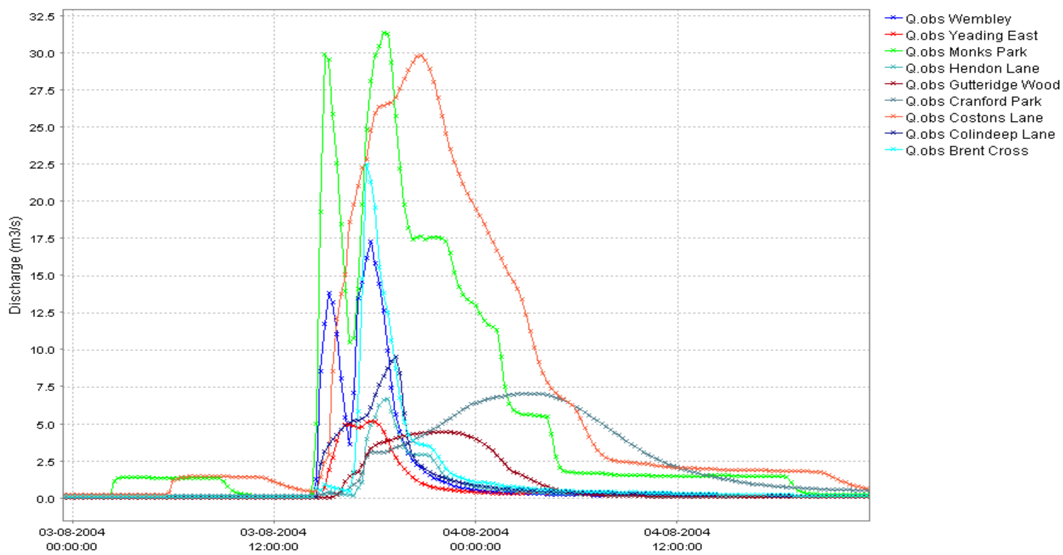


Figure 3.5 Catchment areas for Costons Lane, Monks Park, Wembley and Colindeep Lane gauging stations.

Flow data are available for the period 2–5 August 2004; hydrographs are shown in Figure 3.6.

All locations



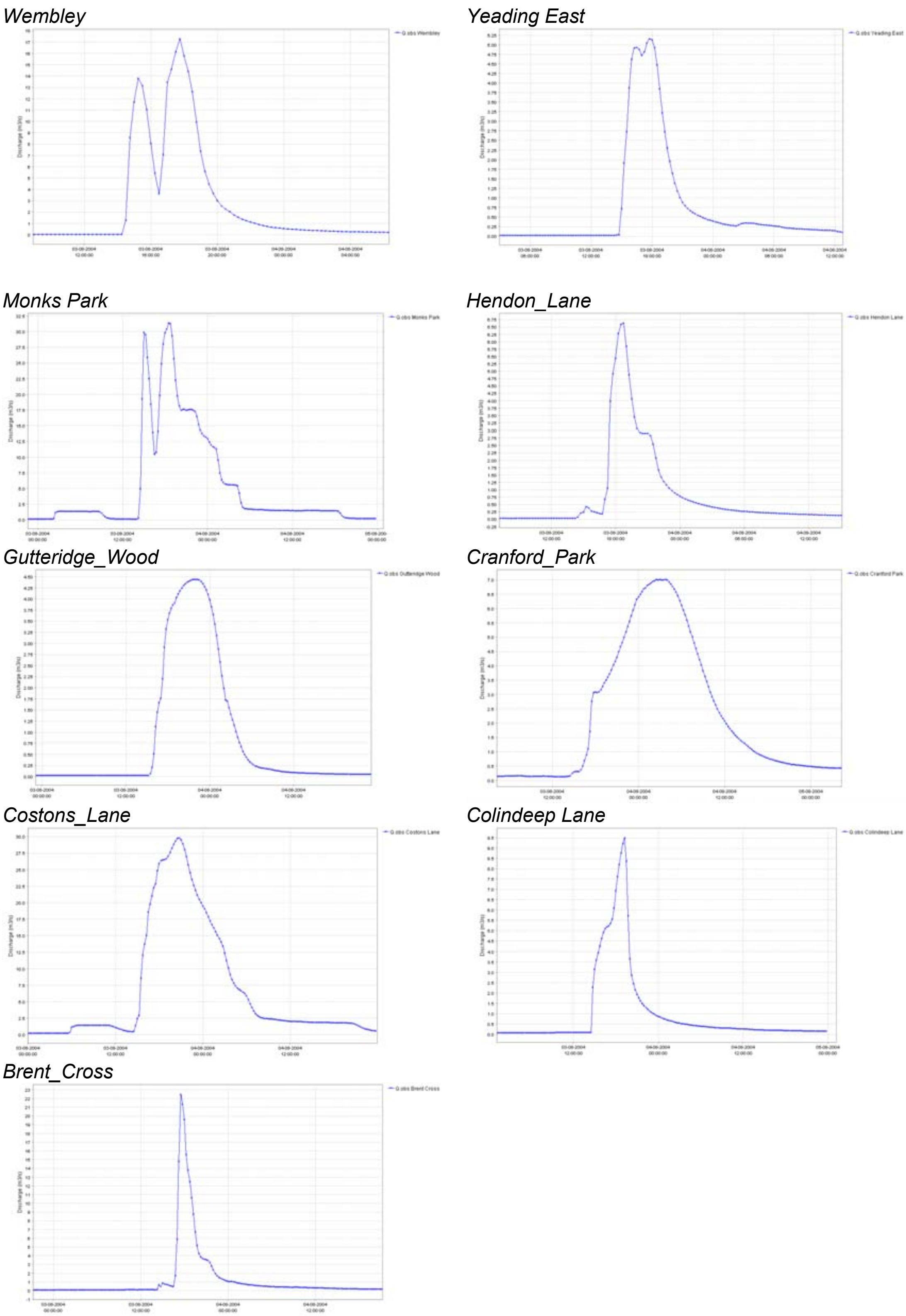


Figure 3.6 Hydrographs for all gauging stations in the vicinity of the storm.

Within Thames Region the event caused flooding of streets, houses and even subway stations. However, no thresholds were crossed for any of the measured flow files. The flooding was limited to local drains and these are not monitored. The amount of flow in the gauges shown in Figure 3.6 reached bank full for most stations – a significant flow.

3.4 Event of 16 August 2004 in Boscastle

The heavy rainfall that affected north Cornwall on 16 August 2004 fell predominantly between 12:00 and 16:00 GMT. It was produced by a sequence of convective storms that developed along a coastal convergence line caused by the change in friction between the land and sea. This effect was heightened by solar heating over land. The exact storm path of each heavy rain cell varied slightly, but the variation between the Camel Estuary and Bude was sufficiently small that the heaviest rain fell on the same catchments throughout the period. This is evident in Figure 3.7, which shows the rainfall accumulation using Nimrod composite radar data over the event.

- Black circles denote gauged catchments where rating equations and flow measurements are available; the most noteworthy response from these occurred for the Ottery at Werrington Park (47015) – see Figure 3.8.
- Other 'level-only' sites are denoted by purple circles; in particular the Ottery at Canworthy Water and the Neet at Woolstone Mill recorded notable floods.

The gauging station details are summarised in Table 3.3 and selected catchment hyetographs are presented in Figure 3.9.

Table 3.3 Gauging station details in the vicinity of the storm.

Number	Name	Location (NGR)	Area (km^2)
47010	Tamar at Crowford Bridge	SX 29050 99100	76.4
47005	Ottery at Werrington Park	SX 33740 86580	120.7
47019	Tamar at Polson Bridge	SX 35300 84900	470.3
N/A	Ottery at Canworthy Water	SX 22840 91680	50.4
N/A	Neet at Woolstone Mill	SS 22730 01810	37.2
N/A	Camel at Slaughterbridge	SX 10940 85720	9.04

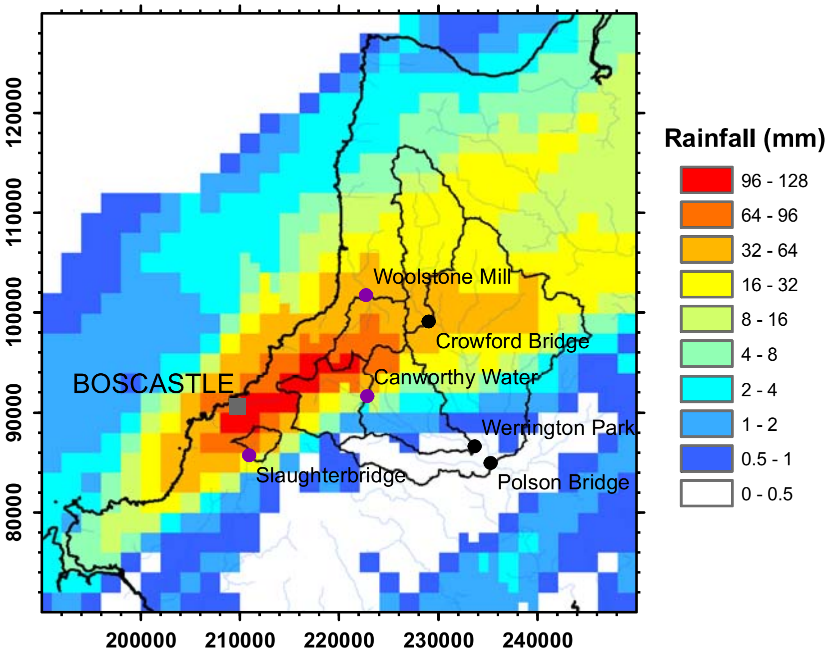


Figure 3.7 Accumulations of five-minute Nimrod radar rainfall data for the seven-hour period starting at 10:00 16 August 2004.

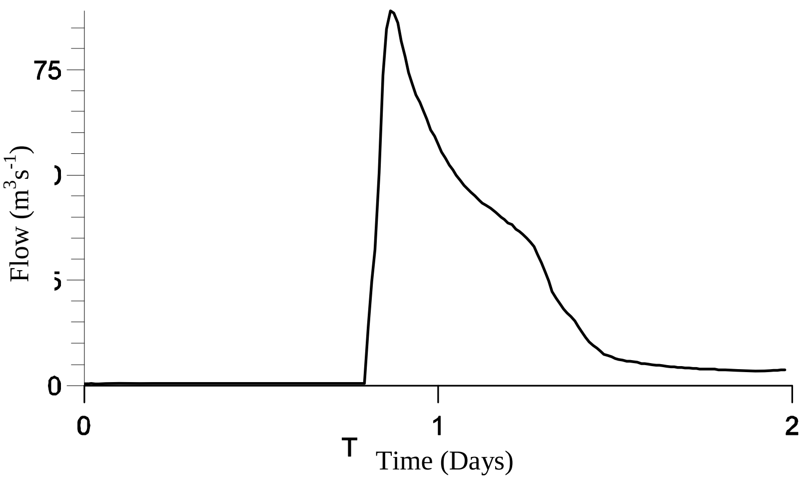
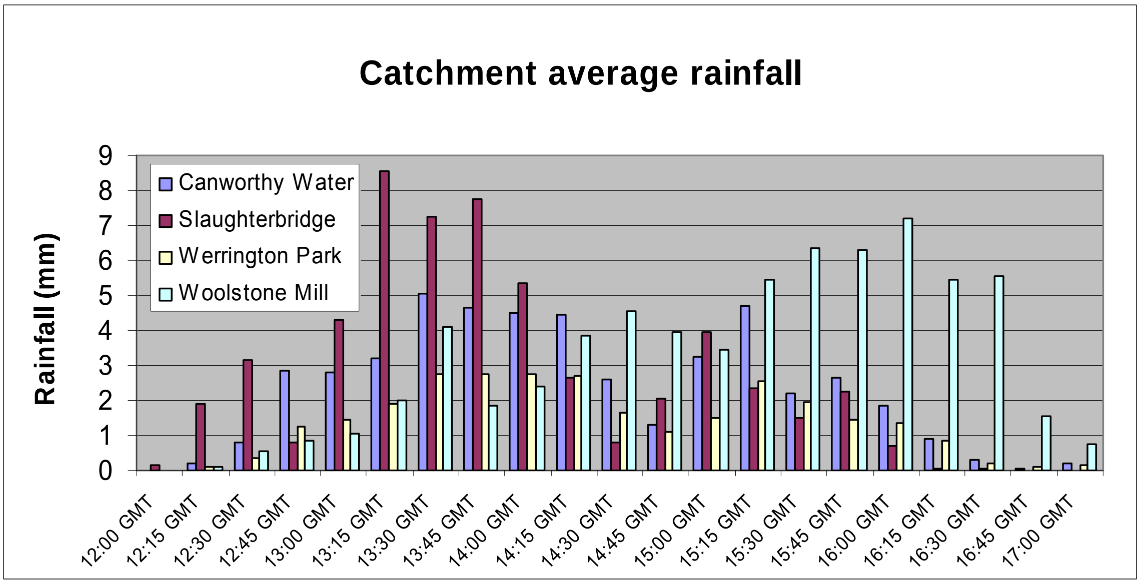


Figure 3.8 Flood hydrograph for the Ottery at Werrington Park, 16–17 August 2004.

(a) 15 minute totals



(b) Cumulative rainfall

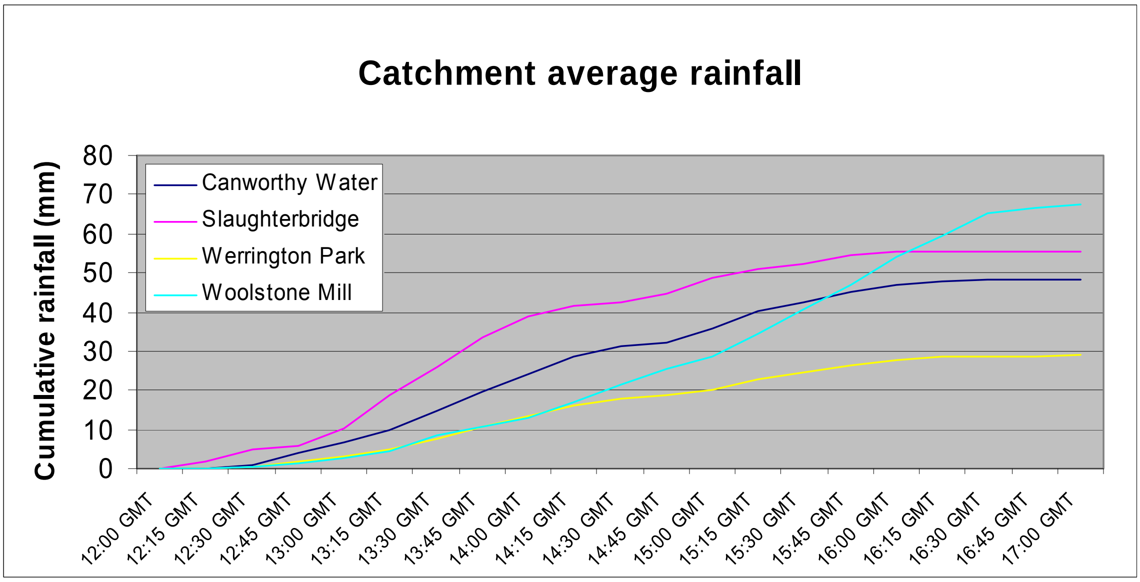


Figure 3.9 Hyetographs of catchment average rainfall using Nimrod composite radar data for selected catchments: Ottery at Canworthy Water, Camel at Slaughterbridge, Ottery at Werrington Park, Neet at Woolstone Mill.

3.5 Discussion

The event in south Wales was quickly ruled out because it had a limited interest from a flood forecasting point of view.

The west London case looked promising as a verification basin alongside any other candidates arising during the course of the project. This case had a number of advantages in that:

- several lumped operational models are available in the area;

- the gauges at the outlets of these models showed a reasonable (bank full) response to the event;
- there is a high density of raingauges in the area;
- the area has several level gauges that are not forecasting points which could be used for calibration and verification.

However, all in all, the Boscastle area seemed most suited for use as the pilot area. There are three main reasons for this choice:

- The close proximity of several fast responding headwater catchments would expose NWP forecast uncertainty in terms of ‘position’ errors and be a good case study from an ensemble point of view.
- The strong topographic influences (Figure 3.10) on flood response would mean that rainfall forecast errors are likely to be transformed in a relatively simple way to errors in modelled flood forecasts. In contrast, urban features of the west London catchments would be more likely to be affected by unique local features that may confound interpretation.
- Gauged catchments in the Boscastle vicinity recorded notable river flow responses (including some ‘level-only’ sites).

A disadvantage of the Boscastle area is the there are no current rainfall-runoff models for the area other than simple transfer function models, making the leap to distributed models rather large.

Figure 3.10 Three-dimensional view from the sea to the Boscastle catchment with a threefold vertical exaggeration applied.

4 Hydrological modelling

This section introduces the modelling concepts that will be applied in Phase 2:

- Physically Realisable Transfer Function Model (PRTF)
- Probability Distribution Model (PDM)
- Grid-to-Grid Model (G2G)
- Representative Elementary Watershed Model (REW).

For a wider discussion of modelling concepts see Moore et al. (2007).

4.1 Physically Realisable Transfer Function Model (PRTF)

A transfer function is a type of time-series model originally popularised by Box and Jenkins (1970). Flow at time t (Q_t) is related to past flow and rainfall through m flow parameters and n rainfall (R) parameters, giving a model with structure (m,n). A pure time delay (τ) can be incorporated into the model structure to lag the impact of rainfall on resultant flow. The form of a linear transfer function model is as follows:

$$Q_t = a_1Q_{t-1} + a_2Q_{t-2} + \ldots + a_mQ_{t-m} + b_1R_{t-1-\tau} + b_2R_{t-1-\tau} + \ldots + b_nR_{t-1-\tau}$$

where:

Q	=	flow
R	=	rainfall
a_1, a_2, a_m	=	flow parameters
$b_1, b_2,, b_n$	=	rainfall parameters
τ	=	pure time delay

The parameters of a transfer function rainfall-runoff model can sometimes be interpreted with respect to catchment dynamics (e.g. partitioning of flow between fast surface and subsurface runoff and slow subsurface runoff, percentage runoff, etc.). However, this is by inference rather than directly as with physically based or conceptual models (Young and Tomlin 2000).

Transfer functions of this form were developed by the Environment Agency's North West, South West, Southern and Anglian Regions. However, the models used an input of total rainfall and, in some cases, tended to perform poorly when applied to independent verification or test events – particularly in catchments with a significant base flow component or dry catchments. This resulted in a lack of confidence in the performance of the models which has limited their operational use.

Because of these problems, both South West and North West Regions use a variant of the simple linear transfer function model known as Physically Realisable Transfer Function (PRTF). The PRTF approach was developed by Han (1991) to address the problems of stability in model output by constraining the model structure and parameters to produce a positive and non-oscillatory (physically realisable) model output.

4.1.1 Model outline

The PRTF model has three additional parameters to adjust the shape, timing and magnitude of the impulse response. At present, these parameters cannot be updated automatically through the course of an event but may be set manually, either prior to the onset of an event or during an event. In addition, South West Region applies a two-stage calibration as a means of developing a look-up table for selecting the parameters based on catchment antecedent conditions. In real time forecasting, the parameters for different antecedent conditions are selected and used automatically by NFFS.

Transfer function models can use either forecast flows or observed flows when running in real time. Although updating the model state through the course of an event provides a powerful self-correction capability in the latter case, over-reliance on state (and parameter) updating to counteract the inadequacies of a linear model in simulating a non-linear process is regarded as unwise.

4.1.2 Application

Transfer function models are calibrated for individual flood warning locations and give therefore only general information about the regional spread of a flood event at ungauged sites. High resolution information about rainfall events is used by the transfer function models in a smoothed manner because the information needs to be translated into catchment average time series. Due to their limited relation to the physical characteristics of a catchment, regionalisation of transfer function models based on parameters derived for nearby locations is difficult.

Transfer function models (especially PRTF) have been successfully applied in South West and North West Regions. In this study, the application of the existing PRTF models for flood forecasting resulting from convective storm events will be tested against the use of more complicated rainfall-runoff models to investigate whether application of the latter actually provides better forecasts.

Table 4.1 lists the hydrological models available in the area around Boscastle; the catchment area covered by these models is shown in Figure 4.1. Beside PRTF models, some correlation models are mentioned in case their input parameters relate to a site for which a rainfall-runoff model (PRTF) is available.

Table 4.1 Available models in area around Boscastle.

River	Location	Model
Camel	Slaughterbridge	PRTF
	Denby	Correlation based on Slaughterbridge
Strat	Bush	PRTF
Neet	Woolstone Mill	PRTF
	Helebridge	PRTF; Correlation based on Woolstone Mill
Fowey	Restormel	PRTF
Lower Tamar	Gunnislake	PRTF
Lynher	Pillaton	PRTF
Walkham	Huckworthy	PRTF
	Horrabridge	PRTF; Correlation based on Huckworthy

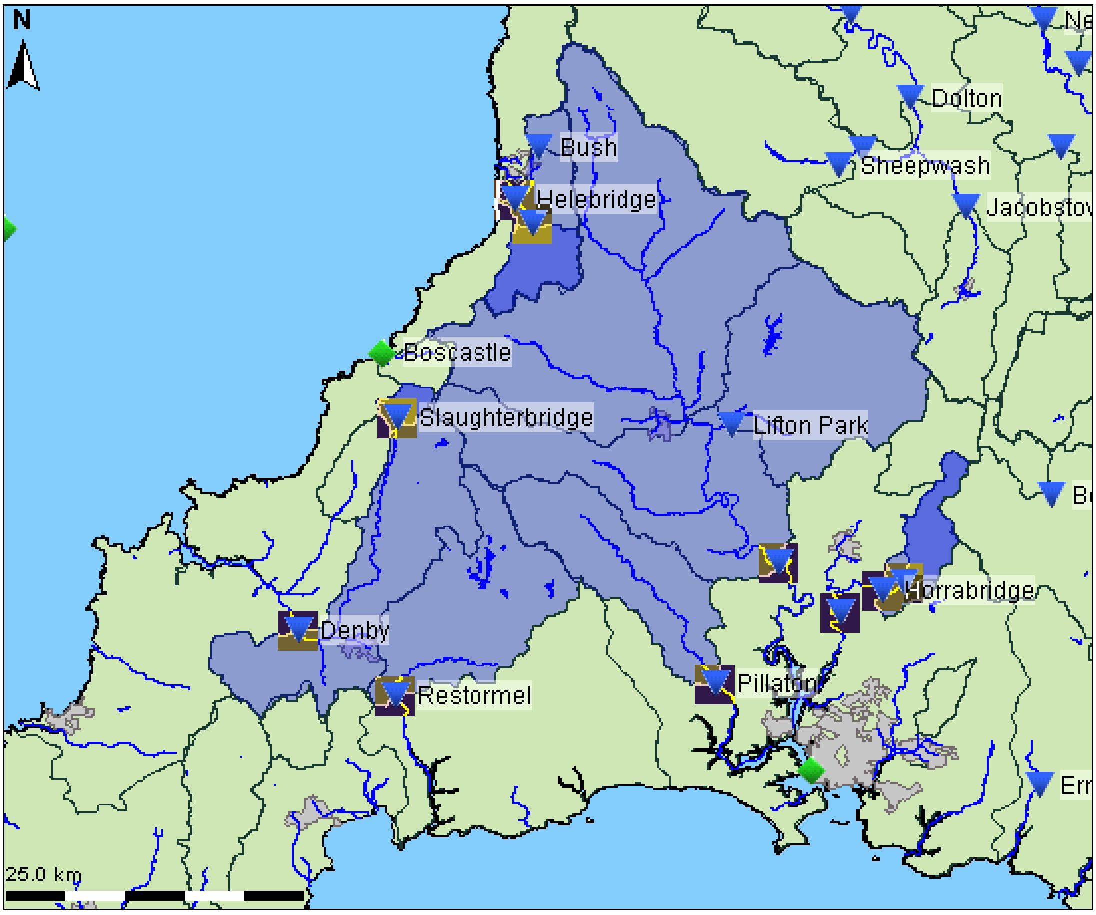


Figure 4.1 Catchment area of models available in the area around Boscastle.

4.2 Probability Distribution Model (PDM)

4.2.1 Model outline

The Probability Distributed Moisture model, or PDM, is a fairly general conceptual rainfall-runoff model which transforms rainfall and potential evaporation data to flow at the catchment outlet (Moore 1985, Moore 1999, CEH Wallingford 2005, Moore 2007). The PDM was designed more as a toolkit of model components than a fixed model construct. A number of options are available in the overall model formulation which allows a broad range of hydrological behaviours to be represented. Figure 4.2 illustrates the general form of the model.

Runoff production at a point in the catchment is controlled by the absorption capacity of the soil to take up water; this can be conceptualised as a simple store with a given storage capacity. By assuming that different points in a catchment have differing storage capacities and that the spatial variation of capacity can be described by a probability distribution, it is possible to formulate a simple runoff production model which integrates the point runoffs to yield the catchment surface runoff into surface storage.

The standard form of PDM employs a Pareto distribution of store capacities, with the shape parameter b controlling the form of variation between minimum and maximum values c_{min} and c_{max} respectively. Drainage from the probability-distributed moisture store passes into subsurface storage as recharge. The rate of drainage is in proportion to the water in store in excess of a tension water storage threshold.

The subsurface storage, representing translation along slow pathways to the basin outlet, is commonly chosen to be of cubic form, with outflow proportional to the cube of the water in store. An extended subsurface storage component (Moore and Bell 2002) can be used to represent pumped abstractions from groundwater; losses to underflow and external springs can also be accommodated.

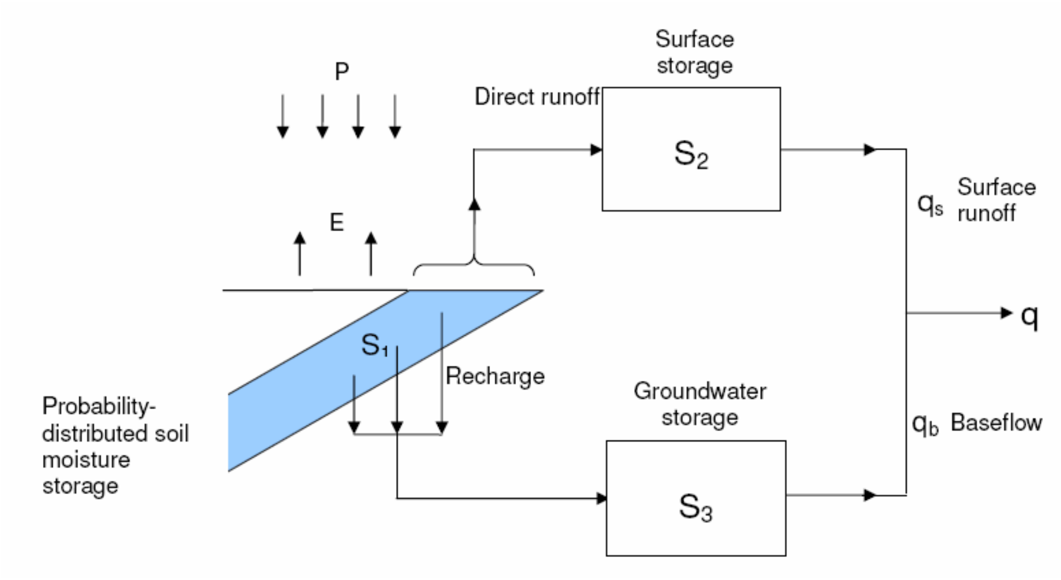


Figure 4.2 PDM rainfall-runoff model.

Runoff generated from the saturated probability-distributed moisture stores contributes to the surface storage, representing the fast pathways to the basin outlet. This is modelled by a cascade of two linear reservoirs cast as an equivalent transfer function model (O'Connor 1982). The model output is formed from the outflow from surface and subsurface storages, together with any fixed flow representing, say, compensation releases from reservoirs or constant abstractions.

4.2.2 Application

PDM is widely used for flood forecasting in the Environment Agency's North East, Wales, Anglian and Southern Regions. PDM parameters are often derived for a number of gauged catchments in a river basin. By regionalisation, PDM parameters for ungauged sites are then derived obtain a more complete coverage of the basin.

The physical–conceptual nature of the PDM and the model's level of intermediate complexity offer some hope of successful application to ungauged sites. Each of the model parameters has a clear physical meaning that invites attempts to establish physically based linkages with data on soil and geological properties, land cover, topography and stream network topology.

No PDM models are used operationally in the area around Boscastle. Within the framework of the project, PDM models will be derived for the gauged sites in the area around Boscastle roughly shown in Figure 4.1.

4.3 Grid-to-Grid Model (G2G)

4.3.1 Model outline

The Grid-to-Grid Model is a physical–conceptual hydrological model configured to:

- represent spatial variability in catchment response;
- make full use of the spatially distributed rainfall data derived from networks of radars and raingauges.

The model can be configured for use at (almost) any spatial resolution and temporal resolution. At present, it is typically run at a 1 km resolution and for a 15-minute time step. The model employs Digital Terrain Model (DTM) data to support its configuration and parameterisation. A modular formulation allows model revisions/extensions to be made.

Through adopting an area-wide formulation in contrast to a catchment-based one, it is well suited to support forecasting at any set of locations within a defined area. As a consequence, the model can be calibrated to groups of gauged locations within an area and forecasts extracted for any ungauged location within the same area. It also supports modelling of nested and parallel catchments.

The model structure is summarised in Figure 4.3. The simplest model variant employs a simple terrain-based runoff production scheme to derive surface and subsurface runoffs from gridded rainfall and potential evaporation inputs. Terrain slope is related to the capacity of the land to absorb water, which is allowed to vary within a grid-square as well as across grid-squares using the probability-distributed principle as invoked in the PDM. The grid-to-grid water routing component employs a kinematic wave formulation that is equivalent in conceptualisation to a network cascade of linear reservoirs. Surface and subsurface runoffs are routed via parallel fast and slow response pathways linked by a return flow component representing stream–soil–aquifer interactions. The terrain-following flow paths are configured using the DTM.

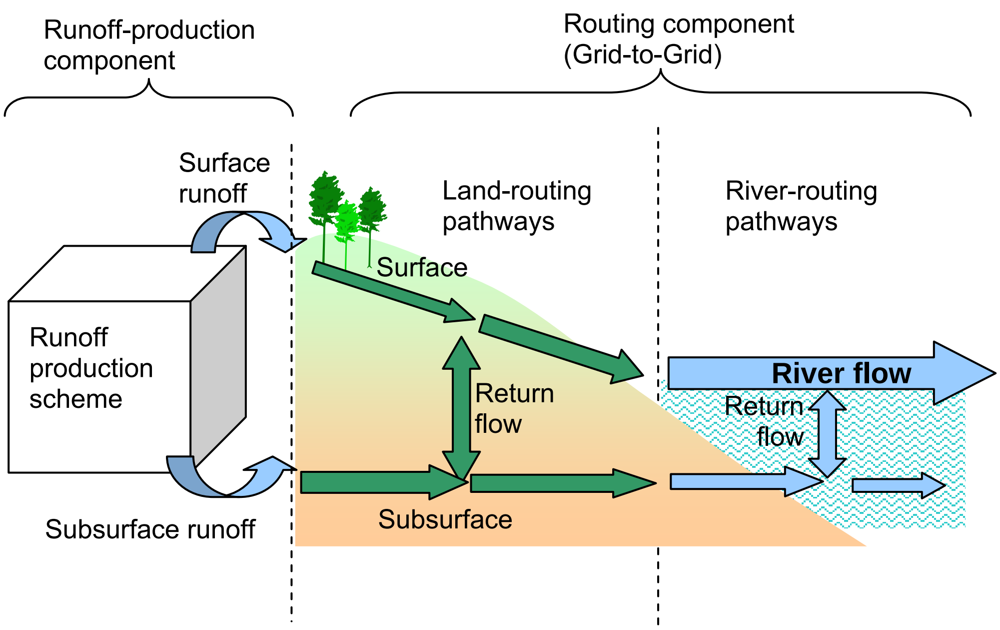


Figure 4.3 Schematic of the Grid-to-Grid model structure.

A more complex variant of the G2G model employs:

- runoff production functions controlled by soil and land cover properties;
- routing functions that may be nonlinear and linked to channel properties, in part inferred from geomorphologic relations.

A schematic is shown in Figure 4.4. Further details of the G2G model can be found in Moore et al. (2006, 2007) and Bell et al. (2007).

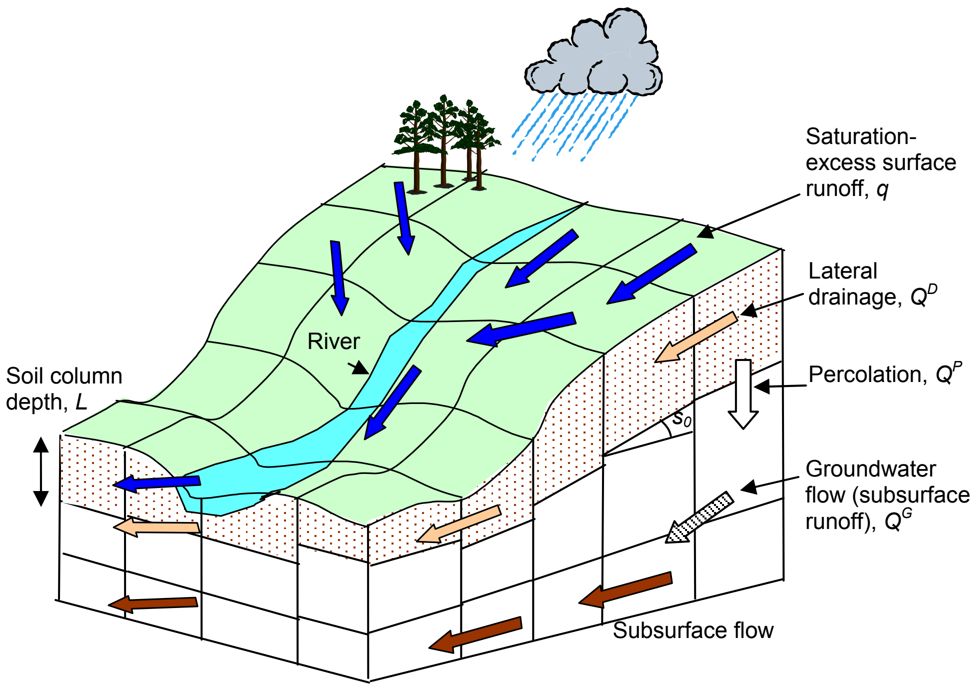


Figure 4.4 Important features of a G2G coupled runoff-production and routing scheme.

The G2G model is particularly suited for use with gridded rainfall time-series derived from:

- spatially interpolated raingauge network data;
- radar data;
- merged radar–raingauge data.

These derived gridded rainfall estimates can be realised operationally within the NFFS via CEH's HyradK module adapter.

For future times, rainfall forecasts in gridded form from STEPS and from NWP models can be used. Configured on a 1-km model domain, the G2G can fully exploit data from weather radar and high resolution NWP model sources. Ensemble forms of rainfall forecast products – from STEPS, high resolution NWP and MOGREPS – can also be used to obtain river flow, runoff and soil moisture estimates in probabilistic form.

An example output from the G2G model is shown in Figure 4.5 for a convective storm over the River Kent in Cumbria. River flow is measured at the Victoria Bridge gauging station in the town of Kendal for which there is a flood warning requirement. The station gauges water from a steep upland catchment responding quickly to rainfall. Five nested catchments are shown and the Victoria Bridge catchment is highlighted in yellow. The HYRAD display shows:

- gridded three-hour rainfall totals (top left);
- G2G model estimates of gridded soil moisture deficit (top right);
- surface/river flow;
- time-series of modelled river flow at Victoria Bridge.

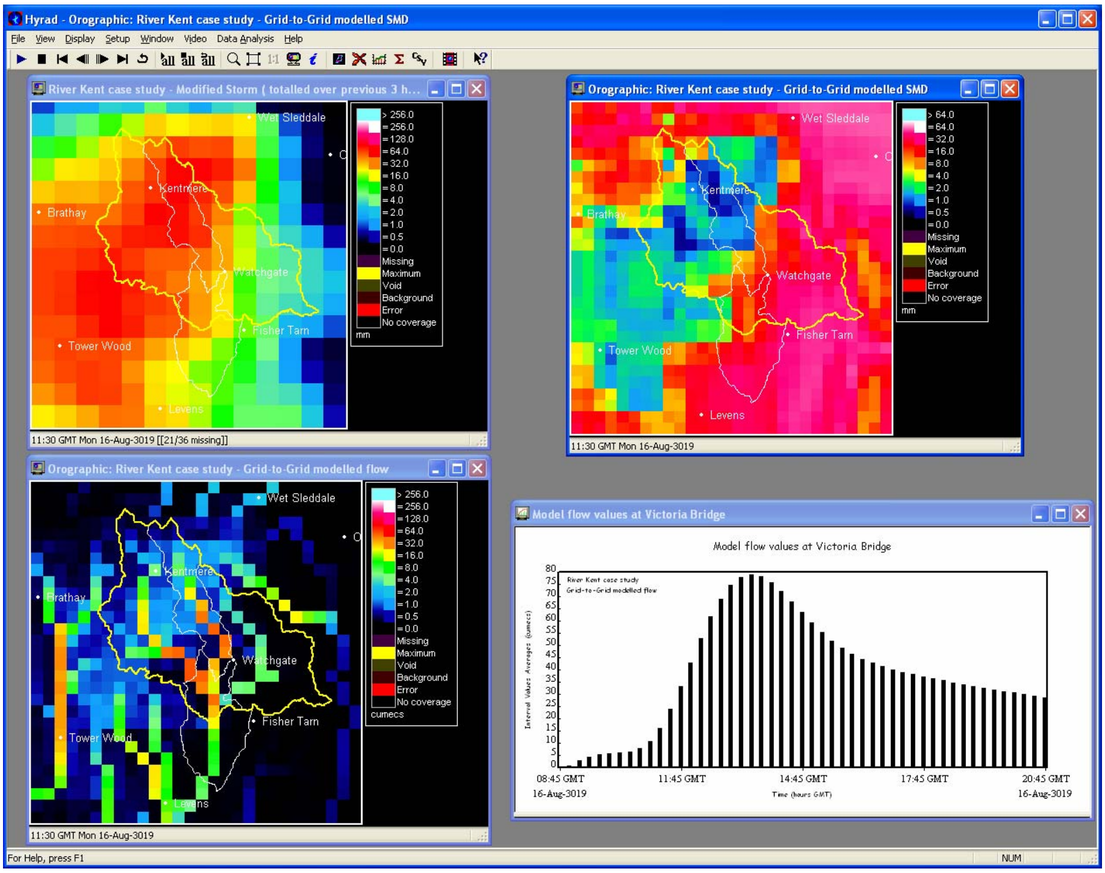


Figure 4.5 Example output from the G2G model visualised in HYRAD.

Notes: Top left: Rainfall accumulation (mm) for the three hours up to 11:30.
Top right: G2G model estimates of soil moisture deficit (mm) at 11:30.
Bottom left: G2G model estimates of surface/river flow (m^3s^{-1}) at 11:30.
Bottom right: Model hydrograph (m^3s^{-1}) for the grid square corresponding to Victoria Bridge (bars represent 15-minute intervals).

4.3.2 Application

The G2G model is being developed as a module adapter for use within the NFFS. Its distributed formulation means that the modelled flood response is sensitive to the pattern and path of storms at a detailed 'convective' scale. This will not be the case for a lumped model such as the PDM when applied to larger scale catchments. In particular, the G2G model will be sensitive to the correct positioning of rainfall forecasts, for example from STEPS, MOGREPS or high resolution NWP systems. When used as input to the G2G model, ensemble forms of these forecasts will be capable of yielding probabilistic forecasts that reflect this sensitivity to pattern and path.

Uncertainty in the position of convective storm cells is a significant source of error in rainfall forecasts. This feeds through to uncertainty in flood forecasts for a target location. The use of the distributed G2G model with forecast ensembles will provide a process-based methodology for characterising flood risk in probabilistic terms across the modelled domain. It will in turn provide a foundation from which to develop risk-based decision-support for flood warning.

4.4 Representative Elementary Watershed Model (REW)

Physically based distributed models have long been regarded as the only models that are able to capture the complexity of a watershed in a dynamic environment. All these models are based on the Freeze and Harlan (1969) blueprint which consists of sets of partial differential equations governing the flow through the soils, overland flow and open channel flow. This blueprint formed the basis of many of the distributed approaches described in the literature. Abbott et al. (1986a,b) presented the SHE[8] model that remains the most complete model to date. Other approaches include the IHDM[9], TOPOG (Vertessy et al. 1993) and HILLFLOW (Bronstert 1999) models.

The Freeze and Harlan blueprint received much criticism (Beven 1989), eventually leading to the bold statement by Beven (2002) that it should be abandoned. Although some still have a firm belief in the blueprint, many people have started working on alternative solutions to solve the notorious problems of scale that are apparent in hydrology. In addition, renewed appreciation of the ‘hydrological workhorses’ – conceptual models such as HBV[10] (Bergstrom 1995), Sacramento (Burnash et al. 1973), etc.) – can be seen in the literature.

In fact, most operational flood forecasting systems rely on technology and model concepts dating from the 1970s. Given that operational forecasters search continuously for better forecasting results, hydrological modellers must question why their latest models are not used for operational hydrology. Assuming that flood forecasters are not just conservative by nature, it can be concluded that recent developments in modelling have resulted in models that have limited use for operational hydrology due to their complexity and high demand for parameters and associated uncertainties.

This therefore leaves the following rather disappointing observations:

- Both the scientific development in hydrological modelling and the application of the limited amount of new concepts have slowed down considerably in the last 10 years.
- Microscale models that are able to evaluate the consequences of land use change cannot be used for larger watersheds without introducing huge computational demands and scaling problems. At the same time the data required to run these models are not usually available on larger scales or cannot be obtained (even if resources were unlimited).
- Conceptual models are the current ‘workhorses’ of hydrological modelling and are being used for operational forecasting. At the same time, there is a strong desire to ‘merge’ meteorological and hydrological forecasting and to use models for more that one purpose (e.g. for climate changes scenarios).

[8] Système Hydrologique Europèen
[9] Institute of Hydrology Distributed Model
[10] Hydrologiska Byråns Vattenbalansavdelning

There are a number of ways in which those involved in forecasting in hydrology can arrive at better predictions at catchment scale and solve some of these problems.

- Instead of providing a single model output that is regarded as the 'best' possible, the use of ensemble techniques now allows for an estimate of the combined model and parameter estimation errors (Yapo et al. 1998). However, models based on the Freeze and Harlan blueprint still demand too much computational resources to make ensemble modelling practical.
- Data assimilation can be used to reduce model uncertainty. Although simple output correction in operational forecasting systems gives good results in terms of a better short-term forecast at the basin outlet, it does nothing to improve the models predictive powers for other applications. Using data assimilation techniques that focus on model state variables both during model calibration and during prediction can improve results significantly. Others have reported success in reducing model uncertainty by incorporating chemical tracer data.
- Another improvement that is currently being investigated is the coupled hydrological and meteorological modelling at the mesoscale.

4.4.1 Model outline

A novel catchment modelling approach based on global balance laws for mass, momentum and energy has been presented by Reggiani et al. (1998, 1999, 2000). The purpose of their work was to integrate the microscale conservation equations for mass, momentum and energy over specially chosen integration regions which make up a representative elementary watershed (REW). Reggiani and Schellekens (2003) describe why the concept needs to be investigated further and deserves more practical application.

REWs are defined in such a way as to allow the definition to be globally applicable and scale independent; they are thus recognisable at various spatial scales reaching from small sub-catchments or patches of a few hectares to entire systems of many square kilometres. In principle, they can be seen as having been extracted with a pastry cutter from the landscape. An example of an REW is shown in Figure 4.6.

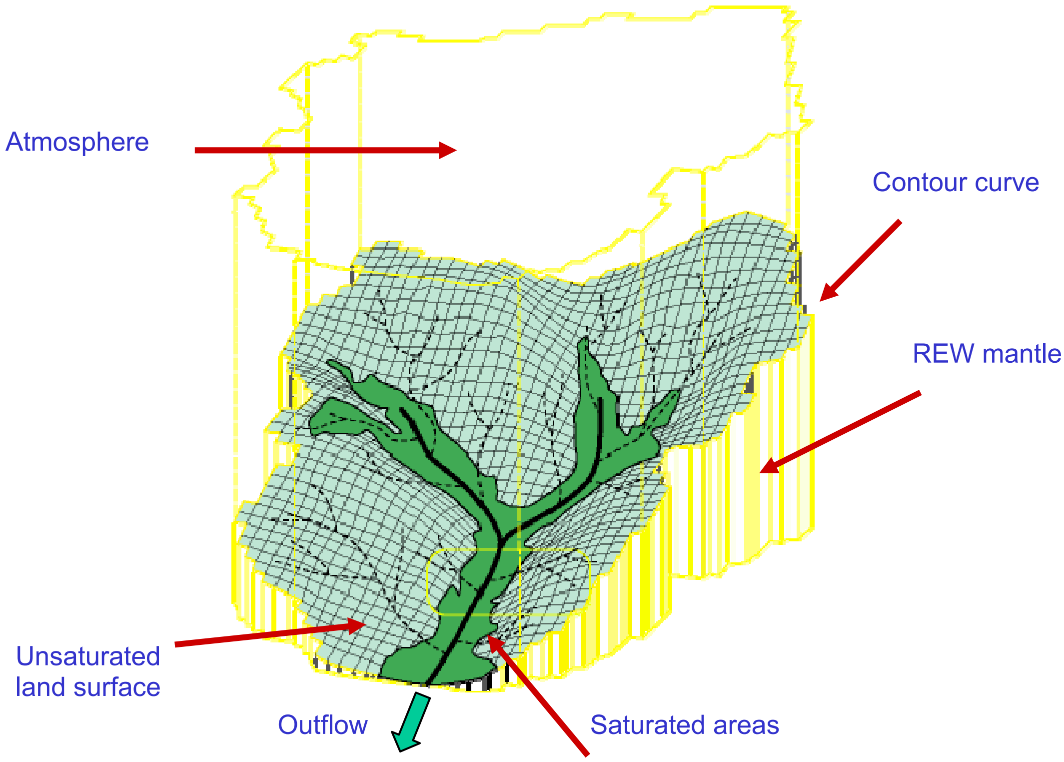


Figure 4.6 View of the volume constituting an REW.

The integration procedure yields – in contrast to microscale or macroscale formulations – so-called megascale balance laws[11] that are obtained without making any a priori assumptions on the importance of various terms. These laws constitute scale-independent ordinary differential equations (ODEs) which conserve physical properties for hydrologically representative zones within an REW in terms of spatially and temporally integrated variables.

The REW methodology is based on a formulation where a change in spatial scale from the microscale to the megascale transforms the spatial gradients into fluxes across the boundary of accurately chosen volumes. A successful application of this approach lies in representing these fluxes in the best possible way. Its implementation, in combination with uncertainty concepts such as 'landscape space' to 'model space' mapping for model hypothesis testing may give some hope for a shift in focus on simulating environmental systems in the near future (Beven 2002).

The REW concept is based on equations integrated from the microscale to the scale of a suitably chosen control volume (megascale), thus converting systematically internal spatial gradients into fluxes across volume boundaries. The resulting model structure is such that three-dimensional (3D) geometry can still be represented and relevant (physically observable) hydrological fluxes across the control volume boundaries are modelled explicitly. This is an important difference with respect to lumped models. In HBV or similar models, the reservoirs become zero-dimensional amorphous stores that have some physical meaning, but have lost their actual geometrical structure. This is not the case with REW; each single hydrological flux is representable, parameterisable and modellable.

[11] For a definition, see Gray et al. (1993).

4.4.2 Application

The REW model has been adjusted by Deltares so that it will run as part of a Delft-FEWS configuration. As such no development was necessary for this project in order to fully utilise the REW model. Like the G2G model, the REW model is sensitive to spatial patterns in the precipitation input and storm movement over the catchment. Clearly, this also depends on the chosen size of the REWs (Figure 4.7). The REW concept has already been used successfully for several sub-basins in the Rhine catchment in conjunction with ensemble forecasts produced by the European Centre for Medium-Range Weather Forecasts (ECMWF).

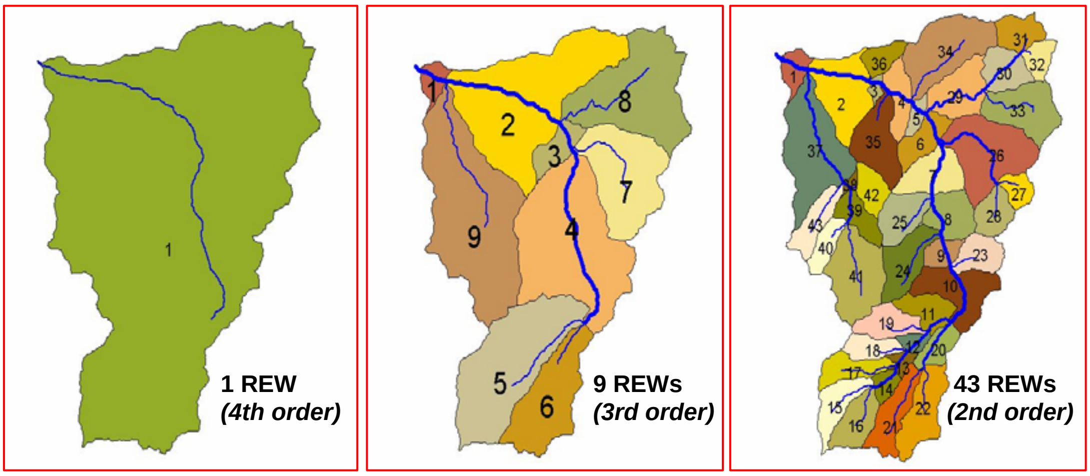


Figure 4.7 Example of different sizes of REWs.

5 Model configuration and calibration

5.1 Configuration

Four modelling concepts for transformation of high resolution rainfall predictions into flood forecasts were tested. These concepts were as discussed in Section 4:

- Physically Realisable Transfers Function Model (PRTF)
- Probability Distributed Model (PDM)
- Grid-to-Grid Model (G2G)
- Representative Elementary Watershed Model (REW).

The PRTF models were run in NFFS for the catchment around Boscastle. Application of this approach was compared to the application of the three other concepts to define whether application of more physically based and spatially distributed rainfall-runoff models offers a better forecasting skill. The three other models were schematised and calibrated for the catchments shown in Figure 5.1.

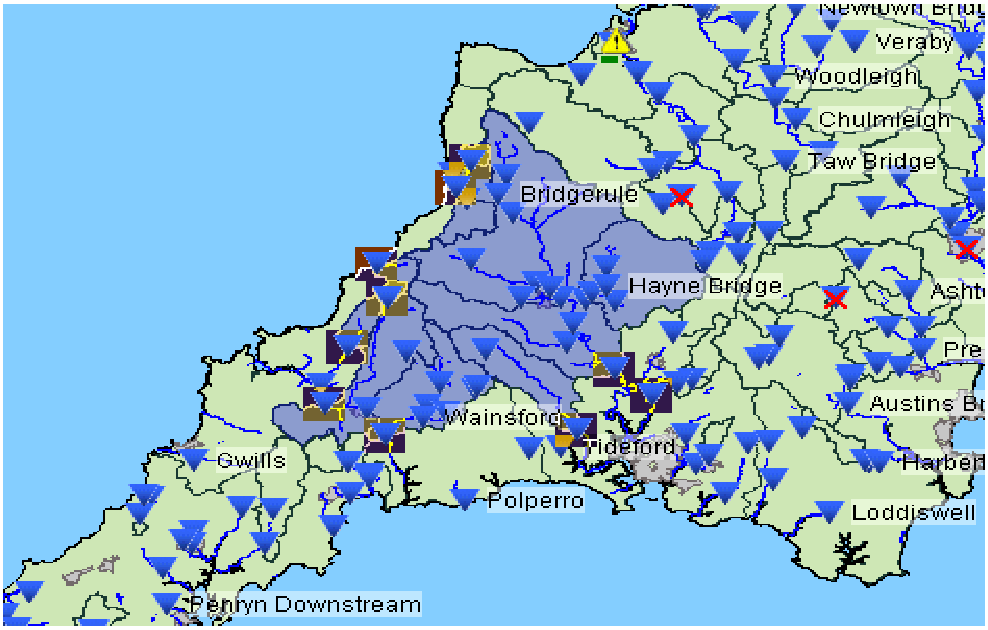


Figure 5.1 Modelling area.

The simplest variant of the G2G and REW model employs terrain data only from a DTM (e.g. 50-metre IHDTM[12]), for configuration of the runoff production and flow routing elements of the model.

A more complex variant of the G2G model – commonly needed for heterogeneous lowland river basins – employs soil and land cover property data to support the runoff production

[12] CEH Wallingford Integrated Hydrological Digital Terrain Model

component of the model; HOST (Hydrology of Soil Types[13]) and LCM2000 (Land Cover Map 2000[14]) spatial datasets are currently used. In addition geological data are required.

PDMs can also be set up based on these geographical data. No **new** PRTF models will be developed by the project.

5.2 Model calibration

5.2.1 Approach

Proper representation of flow generated under convective storm conditions is vital but, under operational conditions, models are generally used for a wider range of climatic conditions. The emphasis of calibration is on reproducing the peak flows accurately.

Calibration is carried out partly automatically and partly manually using predefined criteria where possible. Models of a conceptual or physics-based form have, by their very nature, strong parameter interdependence. Therefore a combination of manual estimation (supported by interactive visualisation tools) and automatic estimation of sub-sets of parameters has been found to work best.

It is anticipated that the calibration strategy may differ for individual modelling concepts, but that a set of performance measures can be agreed (to include formal objective functions and visual hydrograph plots).

A number of performance measures for assessing deterministic and probabilistic forecasts are available. These were considered during the project along with any new measures.

How best to characterise uncertainty in model structure, initial states and parameter estimates will be considered when developing and trialling probabilistic flood forecasting methodologies.

5.2.2 Temporal datasets

Model calibration was based on a continuous dataset with rainfall events. The associated observed radar data (grids) and raingauge measurements were collected. The spatial observed radar data were used as the basis after correction with the help of raingauge data using HyradK functionality. In order to be able to run such ground truth corrections operationally in the future, a Delft-FEWS adapter for HyradK was developed for use in Phase 2 (see Section 5.3).

Calibration required the following datasets:

- river gauge data at 15 minute time steps for 2002–2007;
- raingauge data at 15 minute time steps from 2002–2007;
- observed radar data at 15 minute time steps from 2002–2007 (grids).

The time-series of gauge data were retrieved from WISKI.[15] The radar data were already available at CEH.

[13] See http://www.macaulay.ac.uk/host/
[14] See http://www.ceh.ac.uk/sci_programmes/BioGeoChem/LandCoverMap2000.html
[15] Water Management Information System developed by Kisters AG.

5.2.3 PRTF

There are no plans in this project to develop and calibrate any new PRTF models.

5.2.4 PDM

Model calibration is performed using a combination of manual and automatic optimisation using TSCAL.[16] Performance measures such as root mean square error (RMSE) and R^2 efficiency are monitored along with percentage error and timing of flood peaks. Typically at least a year's worth of continuous river flow and rainfall data are employed for calibration.

5.2.5 G2G

The G2G model is a physical–conceptual model that employs spatial property datasets to support its configuration together with a small number of regional parameters that require calibration to the chosen model domain. The simplest variant employs DTM data to control runoff production variation over space and to configure the terrain-following paths for flow routing. A more complex variant employs runoff production functions controlled by terrain slope, soil and land cover properties and routing functions that may be non-linear and linked to channel properties in part inferred from geomorphologic relations.

Model calibration of the regional parameters is carried out manually using gauging station records within the model domain. The model is calibrated to groups of gauged locations within an area, while forecasts can be extracted for any gauged or ungauged set of locations within the model domain.

Interactive visualisation of the correspondence of observed and predicted flow hydrographs is the main calibration tool, although performance measures such as root mean square error (RMSE) and R^2 efficiency are monitored along with percentage error and timing of flood peaks. Typically at least a year's worth of continuous river flow and rainfall data are employed for calibration. For potential evaporation (PE), a simple sine curve over the year may be used. MOSES[17] (or MORECS) estimates are preferred, if available, or PE estimates from weather stations. Sometimes regional PE standard profiles are used.

Note that the area-wide grid-to-grid formulation of the G2G contrasts with the source-to-sink forms of grid-based model. The latter are calibrated to a chosen outlet point to obtain 'regional' parameters and re-applied to the ungauged location of interest (a nested, paired and/or similar catchment) using these regional parameters but configured using the topography (and any other supporting information) and rainfall for the ungauged catchment area.

The area-wide form of the G2G model allows a regional parameter set to be established providing reasonable performance across a number of gauged sites and, at the same time, allows forecasts for the ungauged sites within the delineated area.

5.2.6 REW

The REW model is a physically based hydrological model. The principal hydrological processes described in the model are:

- channel flow (kinematic wave);

[16] Time Series CALibration program
[17] Met Office Surface Exchange Scheme

- overland flow (kinematic wave);
- unsaturated zone flow (Richard's equation);
- saturated zone flow (Darcy's flow);
- snow model (energy and mass balance equation, Utah State University).

The strength of using a physically-based modelling system like the REW model has been highlighted by reporting a figure of various internal model fluxes and hydrological variables that can be computed by the model (Figure 5.2). All the processes are described by equations that take the physics of the underlying processes and gravity explicitly into account. The parameters have physical meaning and represent measurable quantities such as:

- Manning roughness coefficient;
- saturated hydraulic conductivity;
- soil texture and structure parameters;
- physical snow properties.

The presence of gravity is taken into account through use of a digital terrain model, which allows the slope and aspect of the land surface to be calculated. This is in contrast to purely conceptual models (e.g. HBV, Sacramento) in which the watershed is represented by conceptual stores. In conceptual models, mass is transferred between stores via power law relationships in which the parameters may have no explicit physical meaning and need to be determined by calibration.

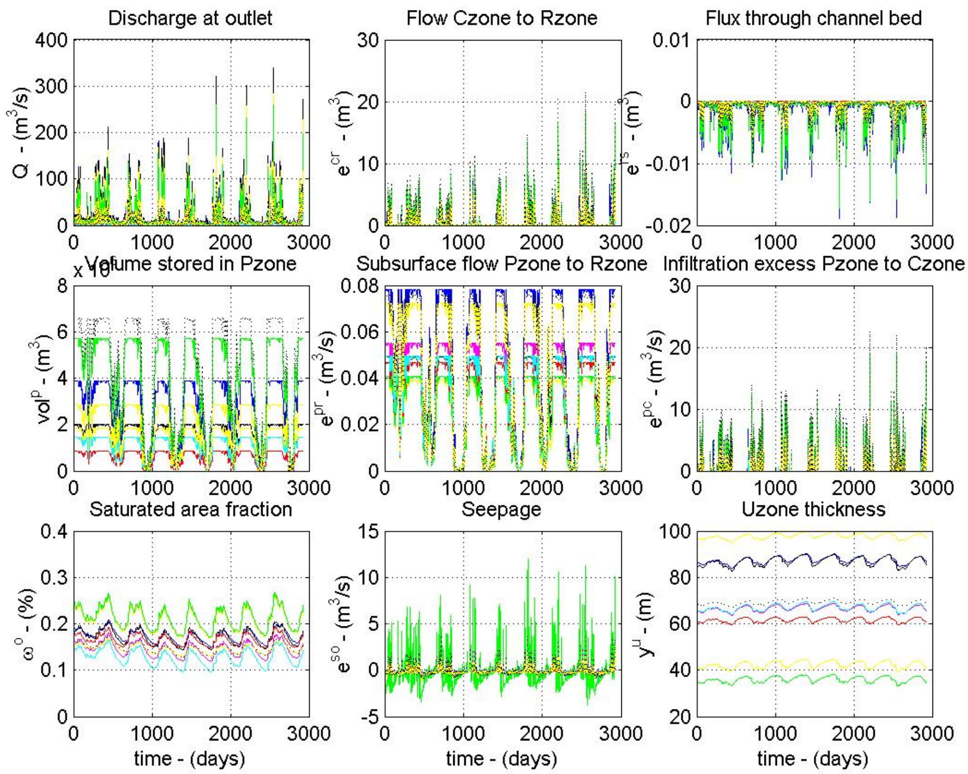


Figure 5.2 Model-internal hydrological fluxes for REWS.

In the REW model, the parameters are determined on the basis of:

- values of soil texture and structure measured in field experiments;
- surface and channel roughness properties observed in the catchments.

These observed parameter values usually allow already a reasonably good preliminary representation of the catchment response. Fine-tuning can be achieved by fixing all parameters apart from about five including channel roughness and soil hydraulic conductivity. These parameters can then be searched for in a relatively limited value range manually or by employing an automatic parameter search algorithm, e.g. the shuffled complex evolution (SCE) algorithm derived by Vrugt et al. (2006).

Performance measures such as root mean square error (RMSE) and R^2 efficiency are monitored along with percentage error and timing of flood peaks. Typically at least a year's worth of continuous river flow and rainfall data are employed for calibration.

The REW model was calibrated with the use of Delft-FEWS. Most of the calibration was performed manually or a link developed by Deltares between Delft-FEWS and the UATools library can be used. Steps in the calibration are summarised in Figure 5.3.

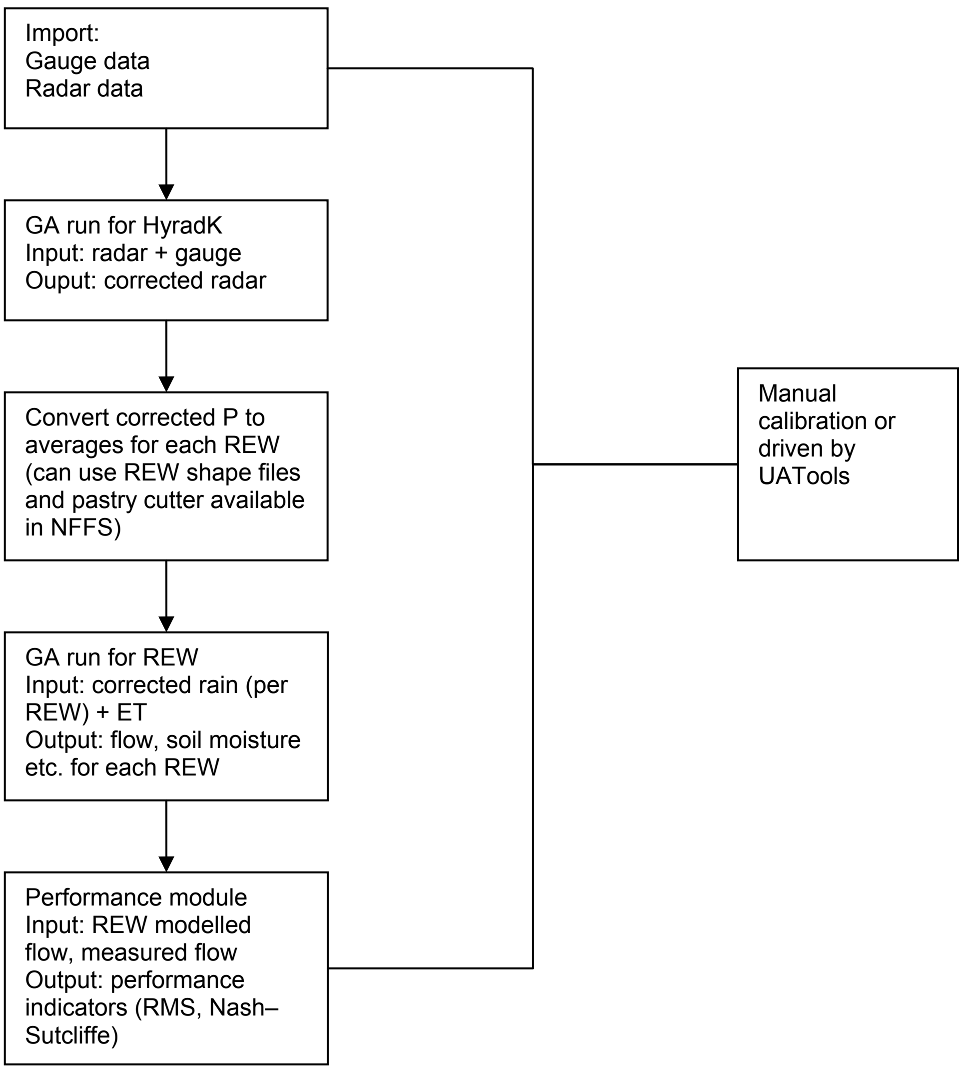


Figure 5.3 Steps taken in the calibration runs for the REW model.

The physical basis of the REW model makes it a platform that is particularly well-suited for applications in ungauged basins. This is because the physical properties of the system can be determined a priori by local surveys in the basins without requiring long discharge records for calibration as needed for purely conceptual models.

Thus far the REW model has been successfully calibrated for the 1,500 km^2 Ourthe Basin and the 1,000 km^2 Geer basin in Belgium, as well as the 30,000 km^2 Mosel basin in Germany/France/Luxembourg.

5.3 Use of HyradK

An NFFS module adapter for HyradK was developed to facilitate generation of gridded estimates of rainfall from raingauge-only data sources and merged raingauge and radar datasets suitable for use in lumped and distributed models.

The HyradK module adapter has two benefits for flood forecasting and warning.

- HyradK can provide improved quantitative estimates of rainfall using raingauge network data alone or merged radar-raingauge data.
- Distributed models such as G2G and REW can exploit the spatial information contained in the gridded HyradK rainfall products.

The benefits of the HyradK functionality and configuration are illustrated below.

5.3.1 Functionality and configuration

HyradK can be used to produce:

- gridded spatial rainfall fields from telemetry raingauge data;
- gridded gauge-adjusted radar products using telemetry raingauge data.

The raingauge-only and gauge-adjusted radar products are derived using a multiquadric surface fitting technique (Moore et al. 1991, Moore 1999). They can be viewed through the HYRAD Display Client or Delft-FEWS. Examples of these surfaces are compared to (unadjusted) radar data in Figure 5.4 for a localised convective event and in Figure 5.5 for a widespread orographic event.

Raingauge-only (HyradK) *Gauge-adjusted 1 km raw radar (HyradK)* *1 km raw radar*

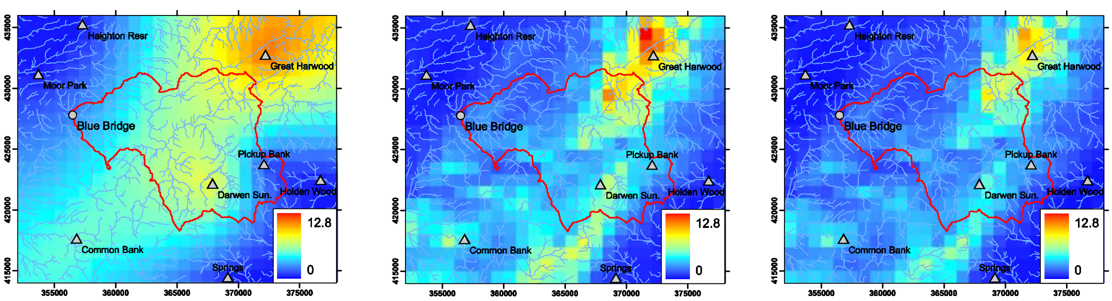


Figure 5.4 15-minute rainfall accumulations derived using different rainfall estimators for a convective storm affecting River Darwen up to 16:00 14 June 2002.

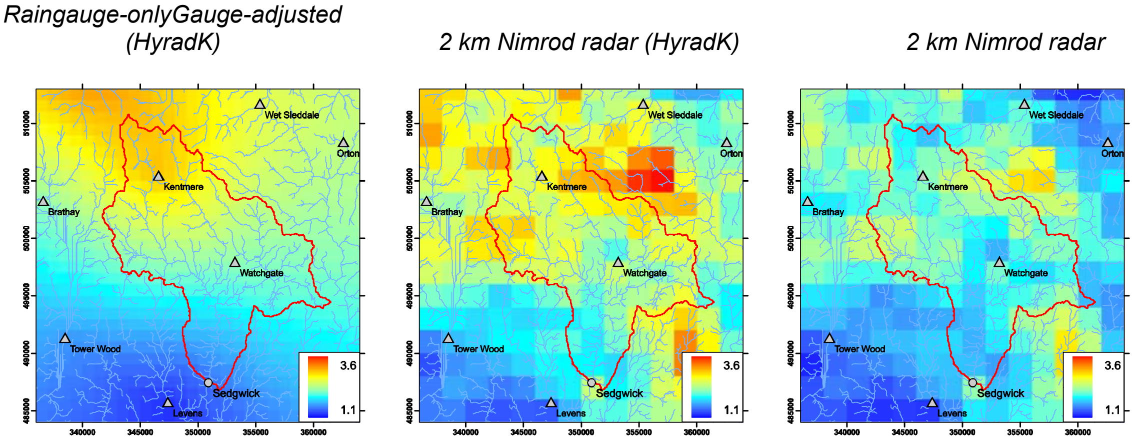


Figure 5.5 15-minute rainfall accumulations derived using different rainfall estimators for an orographic storm affecting River Kent up to 07:15 3 February 2004.

Raingauge-only surface

A form of raingauge-only surface is available which does not need any incidental surface parameters to be set. Therefore this can be applied immediately to any desired region. Usually a resolution of 1 km is used as in Figures 5.4 and 5.5.

Gauge-adjusted radar

Incidental parameters are needed for the gauge-adjusted radar surface fitting scheme. These are usually derived for each radar and resolution by optimisation using historical data (e.g. Moore et al. 2006a).

5.3.2 Application for hydrological modelling and forecasting

For flood forecasting purposes, the practical interest is in the quality of the river flows simulated by the hydrological model when these rainfall estimates are used as input. The River Kent at Sedgwick is used here as a case study to assess the quality of simulations from the PDM (a lumped conceptual model) and the G2G model (a distributed model). For clarity the models are run in ‘simulation-mode’, i.e. no forecast updating via state correction or error prediction is invoked and forecast rainfalls are not used.

Figure 5.6 presents model hydrographs for the River Kent at Sedgwick over an evaluation event using the different rainfall estimators as model input. The evaluation event is an extreme flood during February 2004, thus recreating the common operational scenario where calibrated models are exposed to events more extreme than those contained in the historical record.

- The raingauge-only results using HyradK spatially-interpolated raingauge data are consistently good for both the lumped and distributed models.
- There is a clear demonstration of the added value that the HyradK gauge-adjustment of radar can have over using unadjusted radar data as hydrological model input.
- Model simulations using the gauge-adjusted radar data bring the erroneous under-/over-predictions associated with the unadjusted radar simulations much closer to the observed peaks.

- The gauge-adjusted radar simulations can provide model performance that is as good as, or better than, raingauge-only simulations.

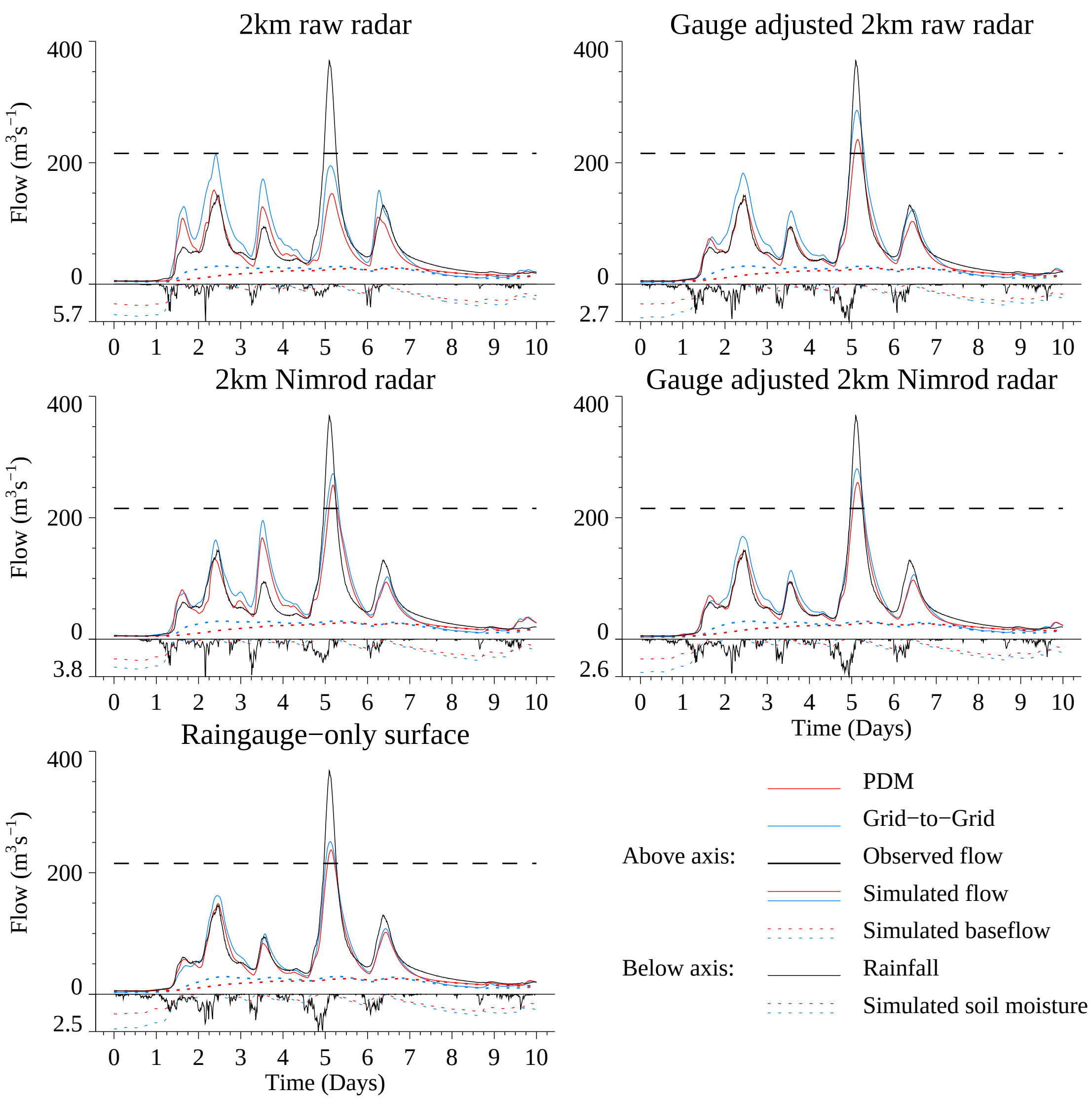


Figure 5.6 Hydrographs for the Kent catchment using different rainfall estimators and hydrological models for the evaluation period 29 January to 8 February 2004.

Notes: The horizontal dashed line indicates the maximum flow used in deriving the rating equation for the river gauging station.
The figure below the axis is the maximum 15 catchment average rainfall.

6 Analysis of the use of high resolution NWP forecasts

6.1 Overview

A number of steps will be taken during Phase 2 to evaluate:

- model performance;
- the added value of using the high resolution NWP forecasts for the pilot catchments.

This evaluation will be carried out for each of the models under investigation for the pilot case study at Boscastle:

- transfer function model (PRTF);
- lumped PDM model;
- G2G model;
- REW model.

Each model type will be calibrated for the pilot catchment. Gridded radar data corrected with HyradK on the basis of raingauge data (see Section 5.3) will be used as rainfall input.

For the PRTF and PDM models, catchment average precipitation time-series will be derived from the corrected grid stacks using the pastry cutter facility in Delft-FEWS.

The input for the REW model will be derived in a similar way but, in this case, an average precipitation will be derived for each REW.

If the radar grid resolution matches the 1 km grid of the G2G model, the radar data may be used directly. Otherwise the spatial interpolation utility in Delft-FEWS will be used to convert the radar grid to the grid required by the G2G model.

The models will be configured in NFFS so as to test them in a forecasting environment. Workflows will be set up that run the module datasets derived for each model algorithm separately. The existing forecast workflow in NFFS, which runs PRTF, will be used as a reference.

The added value of the four modelling concepts when using high resolution rainfall data will be determined by carrying out the following performance analyses for the complete calibration period (2003–2007):

- **Hydrological model simulation performance.** This defines how well a hydrological model can reproduce observed flows on the basis of observed rainfall input (corrected radar data).
- **Flood forecast performance using perfect rainfall forecasts.** This defines how well threshold crossings and flood levels can be predicted running the four modelling concepts in NFFS using a perfect rainfall forecast (i.e. observed rainfall) as input.

These analyses should provide an insight in the performance of the four modelling concepts and whether more physically based – and more complicated – hydrological models have an advantage over the current PRTF approach.

The next step will be to investigate the added value of using convective scale rainfall forecasts. As well as using deterministic NWP forecasts, attention will also be paid to using ensemble forecasts representing the inherent uncertainty of rainfall forecasts.

The potential for improving flood forecasts using convective scale rainfall forecasts will determined by carrying out the following performance analyses for the selected convective event:

- **Rainfall forecast performance***.* This defines how well NWP forecasts at various resolutions can predict rainfall patterns and quantities during convective events.
- **Flood forecast performance using NWP forecasts.** This defines how well threshold crossings and flood levels can be predicted running the four modelling concepts in NFFS using deterministic and ensemble NWP forecasts at various resolutions.

These analyses are discussed in more detail below.

6.2 Rainfall forecasts

A priori knowledge of the observed rainfall would provide the 'perfect' rainfall forecast. The best possible flood forecasts are provided by hindcasting while using observed rainfall as forecast rainfall. All inaccuracies present in such a forecast are mainly the result of model errors.

Deterministic numerical weather predictions are generated without applying perturbations to the initial conditions. In general, the deterministic runs are made with higher grid resolutions and ensemble NWP runs.

Numerical weather prediction is uncertain by nature due the chaotic behaviour of the modelled phenomena. This uncertainty can be represented by perturbing the initial conditions of NWP runs. As such ensembles of NWP runs are generated in which the individual ensemble member is assumed to have a more of less equal probability. Creating NWP ensembles requires high computational capacities. Therefore in practice the generation of NWP ensembles remains limited to coarser resolution weather models or, to a certain extent, limited area models.

Creating ensembles for convective scale rainfall forecasts is to date not an option. To circumvent this limitation and still to somehow be able to represent the uncertainty in high resolution, NWP pseudo ensembles can be created. This technique will be applied in this study and is described below.

6.2.1 Generating 'pseudo' ensembles

One of the major concerns often raised about a storm-resolving NWP model is that it attempts to resolve features that are likely to be inherently unpredictable within the time period over which forecasts will be run. In other words, a model with a grid spacing of around 1 km will be capable of resolving quite small showers but the exact location of any individual shower is not predictable beyond some short period of time (perhaps less than an hour).

The problem is that, for the smallest most unpredictable scales, forecast errors can grow so fast that eventually the error on these scales becomes the same for all forecasts regardless of the initial state, and can be regarded as random noise. This means we should not believe the fine-scale detail at some moment in time in a high resolution precipitation forecast. This statement does not mean that a high resolution forecast system has no skill rather that:

- the focus should be on the most predictable aspects such as rainfall accumulations over larger areas (e.g. river catchments);
- it is necessary to somehow take account of the spatial uncertainty associated with the detail in the forecast.

A simple but effective way of addressing this uncertainty is to assume that the forecast has the correct representation of the weather (e.g. showers, thunderstorms, frontal rain, etc.), but that the location and timing of the rain are never going to be exactly correct.

It is then possible to examine a particular river catchment and ask how much rain will fall on that catchment over some period (e.g. one hour), followed by how much rain would fall on the catchment if the forecast was relocated somewhere else nearby. If this was done for all possible locations within some distance from the original forecast location, we would end up with a collection (or 'ensemble') of equally likely catchment-average accumulations.

The appropriate temporal scaling (accumulation period) and relocation distance would need to be determined in advance at this stage and will generally increase with forecast lead-time. For the purposes of this project, these will be identified on a case-by-case basis.

The effectiveness of the technique is based on the assumption that:

- the rainfall rates produced by the model are reliable;
- the larger scale meteorological pattern is correct;
- the area over which the forecasts are relocated is close to the expected finer scale uncertainty.

More sophisticated methods might try to somehow factor in or reduce the impact of these assumptions (e.g. make use of coarser resolution model ensembles). But for now, the simple approach proposed should be a quick way of obtaining indicative high resolution NWP rainfall ensembles and allow investigation of their use for flood warning.

When this method is repeated for subsequent forecast periods, the number of possible interactions between the different collections of equally likely rainfall scenarios (one collection for each forecast period) soon becomes unmanageable. A method using the probability distribution of the forecast rainfall will therefore be developed for selecting a suitable subset (or 'ensemble') of the numerous possible rainfall scenarios.

6.2.2 Rainfall forecast performance

Although this research concerns the use of the high resolution NWP and not the performance of rainfall forecasts themselves, a meaningful analysis cannot be performed without a good idea about the performance of these forecasts.

The performance of the deterministic NWP forecast will be assessed on the basis of the subcatchment and aggregated grid cell average values of rainfall totals over different forecast horizons. This analysis will be performed for different NWP resolutions from the regular grid size of 12 km down to the convective scale (1–2 km). The analysis will be carried out only for the time of the convective event (2–3 days).

Rainfall forecast performance will be quantified using the indicators currently configured in NFFS for routine comparison of rainfall estimations with observed data:

- bias
- root mean square error (RMSE)
- root mean square factor (RMSF).

The temporal and spatial aggregation level applied to compute the indicators will be defined when the observed and forecast data are available.

The similarity between actual and forecasted spatial rainfall patterns will also be determined. For the latter analysis, it is hoped to build on existing work by JCMM in this field for the Boscastle event.

Testing of the skill of the pseudo ensembles is not anticipated.

6.3 NFFS configuration

The NFFS configuration for the Environment Agency's South West Region will be extended to support the analyses and to make it possible to define the forecast performance. The main changes to the configuration for the purpose of this study are discussed below.

6.3.1 Import and pre-processing

The NFFS already imports the current Met Office's deterministic NWP forecast, receiving the data in Nimrod format. High resolution NWP forecasts are also assumed to be available in Nimrod format and import of higher resolution deterministic NWP forecasts will be configured within the framework of this study.

A module adapter will be developed to allow running of HyradK from NFFS. HyradK will be configured to correct radar actual grid stacks on the basis of raingauge data.

The preference of the research team is to generate the pseudo ensembles on the basis of deterministic rainfall forecasts to be configured in NFFS or though this could potentially be done offline. A decision will be taken in a later stage. Delft-FEWS can be configured to create a pseudo ensemble by repositioning grids. Alternatively, an external module could be used for this purpose. Should pseudo ensembles be generated offline, an import of the pseudo ensemble will be configured in NFFS.

6.3.2 Data processing

Data processing involves all the steps currently being performed to prepare for the current models. In addition, it will be necessary to prepare data for the G2G and REW models.

Input series will be prepared:

- in grid format for the G2G model;
- in time series per REW for the REW model;
- per HYRAD catchment for PDM.

6.3.3 Forecast workflows

Separate forecast workflows will be configured running PDM, G2G and REW using alternative rainfall forecasts. State runs will also be set up for these models. For PRTF, workflows will be added running alternative rainfall forecast inputs.

The PDM and G2G models will be run by the Delft-FEWS general adapter (GA) through a module-specific adapter. A FEWS adapter for PDM is already available and a new adapter for G2G will be developed as part of this project. The REW and PRTF models can read XML processing instruction (PI) files directly and can interface with the GA without the need for an adapter (this is part of REW functionality).

The generation of ensemble forecasts will be configured for all four model concepts. The general adapter will be configured in such a way that the models will be executed once for each ensemble member. The model configurations themselves do not need to be adjusted for the ensemble runs as this will be handled by the general adapter.

6.3.4 Post-processing

The Delft-FEWS time-series display and spatial display will be updated to show new input and output.

Performance indicators will cover deterministic and ensemble forecasts. For the latter, the Delft-FEWS performance indicator module has recently been extended to compute continuous Brier scores and Brier skill scores.

The processing of the NWP forecast in NFFS is illustrated in Figure 6.1. Four main steps can be identified:

- Import the (non-ensemble) forecasts.
- Generate the ensembles.
- Run the forecast in ensemble mode, i.e. loop over each ensemble and run the 'normal' forecast for the configured models (lumped and distributed).
- Run a forecast statistics module to evaluate forecast performance.

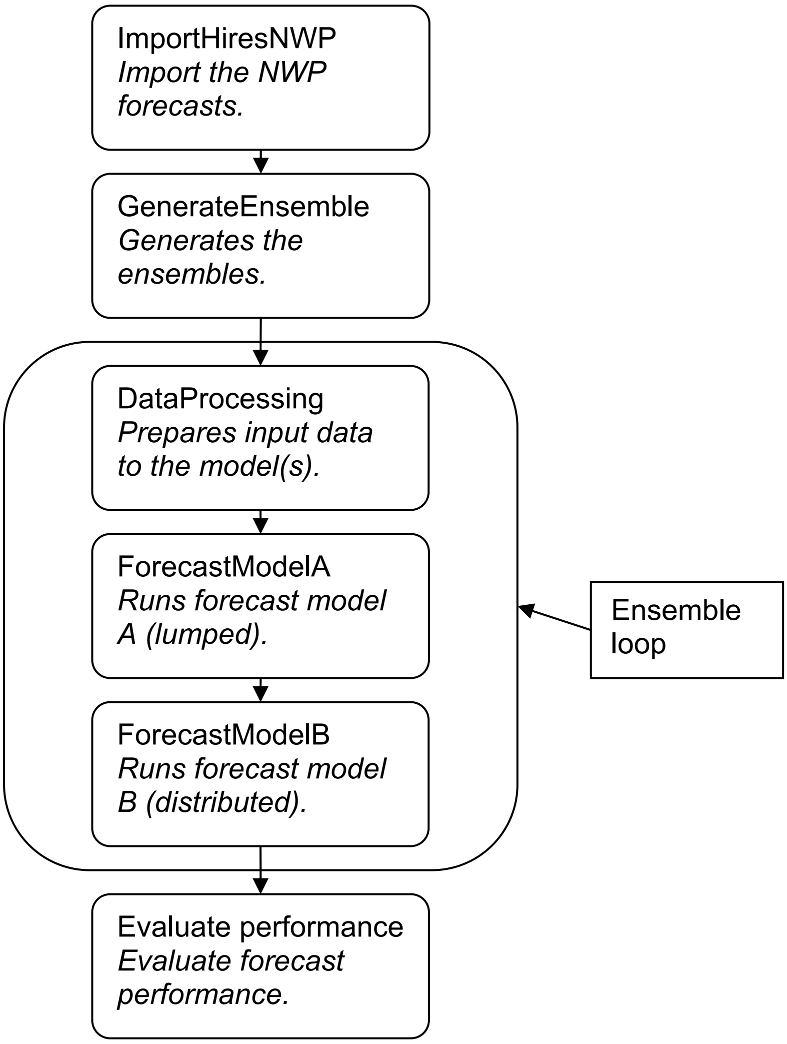


Figure 6.1 Schematic representation of a forecast workflow for the high resolution NWP forecasts.

6.4 Hydrological model performance

The model performance in the calibration period was analysed with respect to the difference in spatial representation between the lumped and distributed models. For example, a single storm with the same intensity that falls either in the upper or lower part of a catchment will give an identical response in a lumped model but may generate a different response in a distributed model (see Figure 6.2).

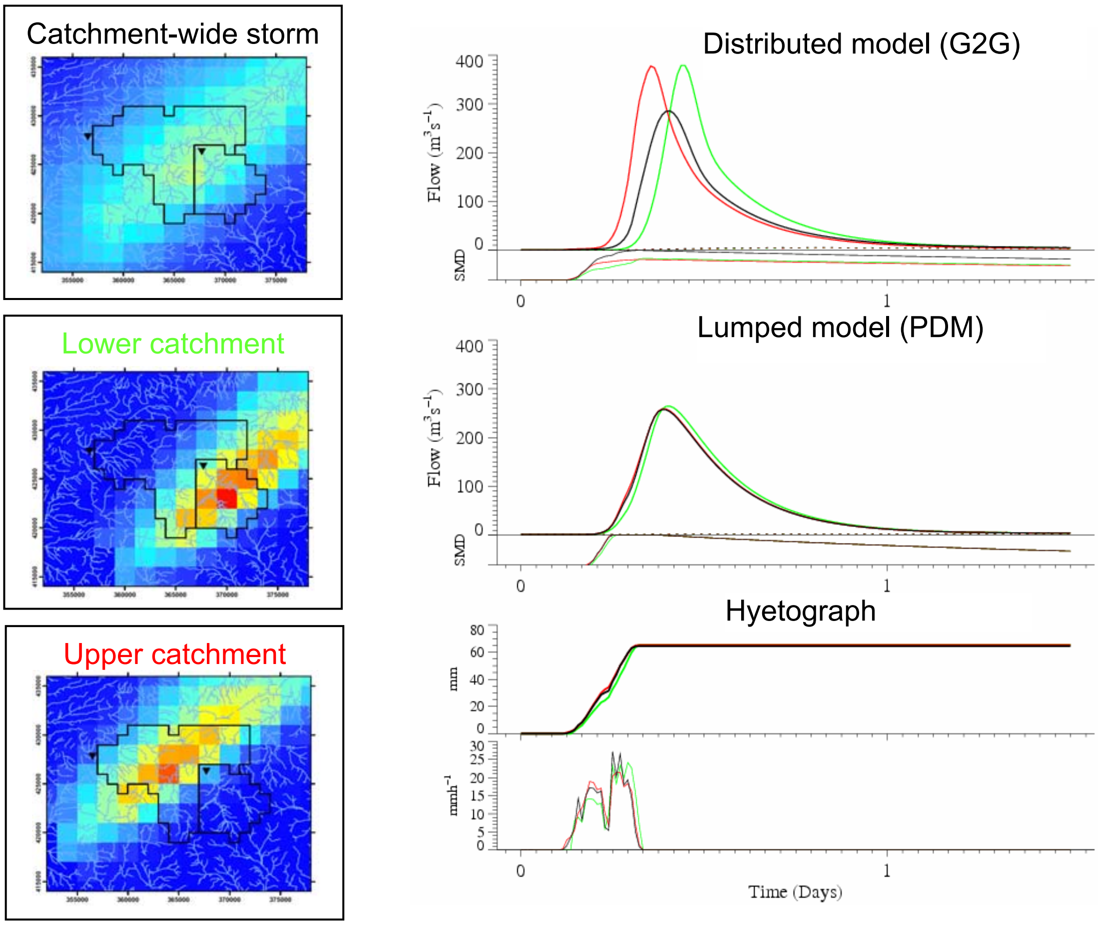


Figure 6.2 Differences in response to a storm in a distributed and lumped model.

The three storms depicted in Figure 6.2 are similar in terms of the catchment hyetograph but different in location (upper and lower storms) and coverage (stretched to give a catchment wide storm). The lumped model gives nearly identical hydrographs for all storms. The distributed model has a different response to each storm. When a storm is closer to the outlet of a catchment, the distributed model has a faster response to the rainfall; for a storm in a further part of the catchment, the response is slower (the lumped model has the same response in both cases).

The hydrographs presented in Figure 6.2 suggest that the spread of the ensemble flow forecasts will probably be larger for the distributed models compared with the lumped models. This is because the catchment averaging procedure required for the lumped models will spread the rain over the entire catchment whereas forecasts made with the distributed model will retain more of the spatial information contained in the original input signal. How important this difference is depends on the size of the catchment relative to the grid size of the NWP forecast. The spatial extent of the actual rainfall event is also crucial.

The track of a storm over a catchment also causes different responses from distributed and lumped models. For example, a storm moving upstream to downstream would be expected to generate a sharper river response at the outlet compared with a storm moving from downstream to upstream. A distributed model has the potential to represent this behaviour while a lumped model does not.

Distributed models are, in principle, able to simulate a more realistic response to rainfall events. In this project, the performance of the model concepts will be quantified with performance indicators to investigate whether the theoretical advantage translates itself into a better simulation performance.

6.4.1 Procedure

Hydrological model forecast performance will be determined by comparing model output with actual hydrological gauge readings for the complete calibration period (2003–2007). The comparison will be made for the module datasets calibrated for the various model concepts at all hydrological gauge stations.

The performance will be determined for two cases:

1. Model simulation (without updating or error correction).
2. Hydrological forecasting in NFFS using 'perfect' rainfall forecasts (observed rainfall).

Case 1 will provide an insight in the general behaviour performance when used without updating or error correction. The models will be run for the complete period 2003–2007 with corrected radar actuals (i.e. observed rainfall).

Case 2 simulates the use of the models in an actual forecasting environment. The models will be run as part of regular forecast workflows and all regular tools for improving forecast performance (e.g. updating and error correction) will be applied. This means that forecasts will be produced at regular intervals (two hours) during the whole period 2003–2007. Initial conditions for the forecast runs will be provided by daily state runs. This hindcast analysis will be made by running the NFFS standalone system in batch mode.

Case 2 will provide an upper estimate of the forecast performance achievable on the basis of rainfall forecasts. The remaining uncertainty will be mainly due to errors in the models themselves and the rating curves.

6.4.2 Performance

The hydrological model simulation performance will be quantified using the performance indicators presently configured in NFFS. These are:

- Lead time accuracy using mean square error (MSE), bias and mean absolute error (MAE) for selected lead times (see Figure 6.3). In NFFS, MSE and BIAS are computed for lead times of 2, 6, 12, 18 and 24 hours. MAE will be added as an indicator.
- Forecast performance skill scores for predicting threshold crossings. The following skill scores will be computed:
  - false alarm rate (FAR)
  - probability of detection (POD)
  - critical success index (CSI)
  - first forecast of threshold (FFOT)
  - bias.

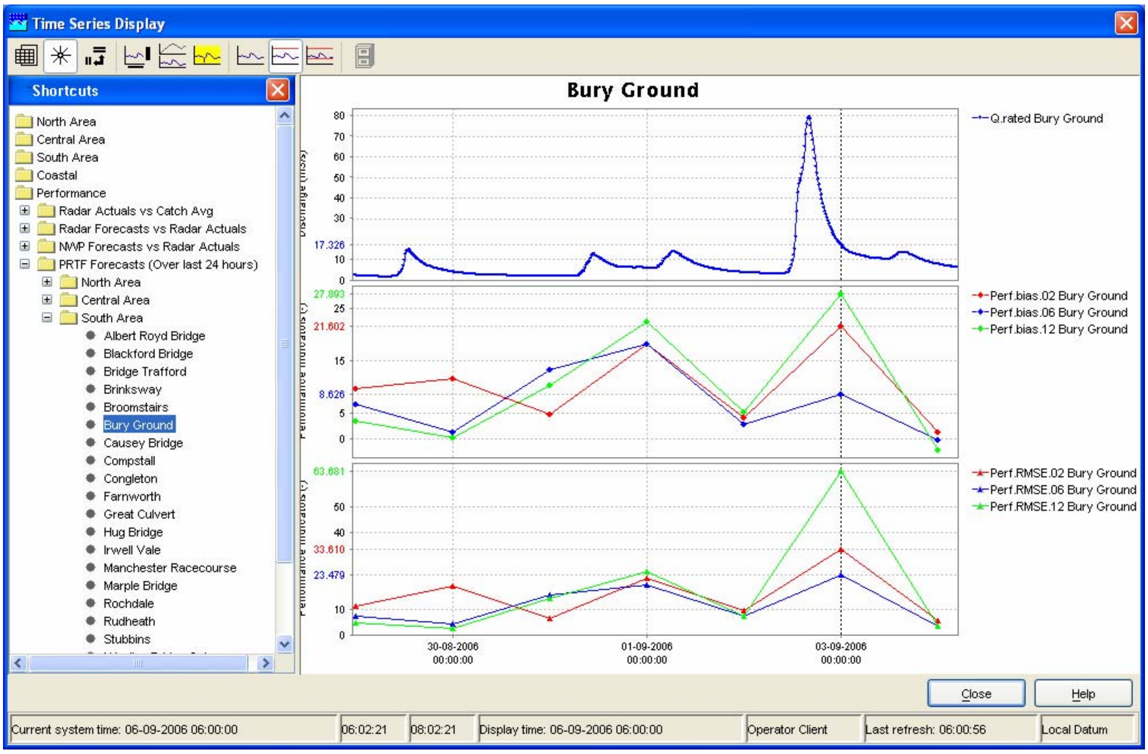


Figure 6.3 Example graph showing lead time accuracy using bias and RMSE for 2, 6 and 12 hour lead times.

6.4.3 Lead time performance

Performance of forecast is assessed on the basis of lead time accuracy. This is done by comparing the forecast lead time value against the observed value at the same. For each lead time, this value is assessed over a given number of forecasts.

Lead time accuracy is evaluated using the mean square error (MSE), BIAS and mean absolute error (MAE) as follows:

Mean square error in lead time accuracy (LEAD_MSE)

$$LEAD\_MSE^{\tau} = \frac{1}{J}\sum_{j=1}^{J}(\hat{x}_j^{\tau} - x_j^{\tau})^2$$

Bias in lead time accuracy (LEAD_BIAS)

$$LEAD\_BIAS^{\tau} = \frac{1}{J}\sum_{j=1}^{J}|\,\hat{x}_j^{\tau} - x_j^{\tau}\,|^2$$

Mean absolute error in lead time accuracy (LEAD_MAE)

$$LEAD\_MAE^{\tau} = \frac{1}{J}\sum_{j=1}^{J}(\hat{x}_j^{\tau} - x_j^{\tau})$$

where:

$LEAD\_\#\#\#^{\tau}$	=	lead time accuracy at time τ
J	=	number of forecasts considered
x_j^{τ}	=	reference value at time τ
$\hat{x}_j^{\tau}$	=	estimated value at time τ .

Note that reference values x_j^{τ} may not yet be available; if this is the case, the number of forecasts considered (*J*) is reduced accordingly.

6.4.4 Skill scores

By matching observed and forecast thresholds, various skill scores can be established. The first step is to set up a contingency table like the one shown in Table 6.1. Once criteria have been set for the time intervals to consider when matching thresholds, the values of *a*, *b*, *c* and *d* can be filled in (where, for example, *a* is the number of matched observed and forecast thresholds).

Table 6.1 Example contingency table.

Forecast threshold crossing	Observed threshold crossing	
	Yes	**No**
Yes	*a*	*b*
No	*c*	*d*

Four criteria are used to match the threshold crossings.

The first two criteria (min/max T_0 difference) determine how long before the actual observed threshold (the T_0 of the forecast) will be of use. Alternatively, if a threshold is predicted too soon before it happens, the statistic may contain spurious results.

The last two criteria (maximum time threshold is too early/late) determine how close the forecast threshold crossing should be to the observed threshold crossing in order for it to be considered the same threshold event.

Once the contingency table has been filled, different skill scores (FAR, POD, CSI) can be established:

$$FAR = \frac{b}{a+b}$$

$$POD = \frac{a}{a+c}$$

$$CSI = \frac{a}{a+b+c}$$

FFOT is determined as the average time between the T_0 of the forecast run in which a threshold was detected and the time of the threshold (i.e. the average lead time of the category *a* thresholds in the contingency table).

The bias of paired thresholds is the average time between paired observed and forecast events.

6.5 Hydrological forecast performance

The discussion on model performance assumes perfect rainfall forecasts. In reality, hydrological forecasts are often at least partly based on rainfall forecasts. The rest of this section examines:

- the performance of hydrological modelling concepts when dealing with deterministic and ensemble rainfall forecasts;
- the added value of convective scale rainfall forecasts and how to make effective use of them in flood forecasting.

6.5.1 Procedure

The analyses described below will be carried out only for the Boscastle event on 16 August 2004. For practical reasons, the availability of high resolution NWP runs is limited to the period around this event.

Within NFFS, hydrological forecasts will be generated throughout the pilot event at 30-minute intervals. Forecasts will be made for the following three cases:

1 ‘Perfect’ rainfall forecasts (see also Section 6.4);

2 Deterministic NWP forecasts with grid resolutions ranging from 12 km down to the convective scale (1–2 km);

3 Pseudo ensembles of NWP forecasts.

A MOGREPS hindcast could be used for reference purposes if this proves practical.

6.5.2 Performance

The hydrological forecast performance for cases 1 and 2 will be quantified with the same performance indicators as applied in Section 6.4. However, their use will be limited because this part of the research will focus on only a single event.

The performance indicators computed for cases 1 and 2 will be as follows:

- Lead time accuracy using MSE, bias and MAE for selected lead times (see Figure 6.3). In NFFS, MSE and BIAS are computed for lead times of 2, 6, 12, 18 and 24 hours. MAE will be added as an indicator.
- Forecast performance skill scores for predicting threshold crossings (FAR, POD, CSI, FFOT and bias).

These performance indicators are not applicable to probabilistic forecasts. Only those performance indicators suitable for determining the skill of ensemble forecasts are discussed in Section 6.5.3.

6.5.3 Ensemble performance

Brier skill scores will be used to define the skill of the ensemble forecasts generated for the Boscastle event. But because the analysis will be based on a single event, the computed scores will have a limited value. Nevertheless it is important to investigate the procedure of measuring the skill of ensemble forecasts by applying the methods described below.

The **Brier score** (BS) measures the mean square error of probability forecasts and is defined as the total of squared probability errors divided by the number of forecasts. For example, if an event was forecast with a probability of 60 per cent and the event occurred, the probability error is –40 per cent (i.e. 60 – 100). The Brier Score (BS) is defined as:

$$BS = \frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2$$

where *p* denotes the forecast and *o* the observation.

The **Brier skill score** (BSS) is defined as:

$$BSS = \frac{BS - BS_{reference}}{0 - BS_{reference}}$$

Figure 6.4 gives the typical size of the forecast error in probability terms.

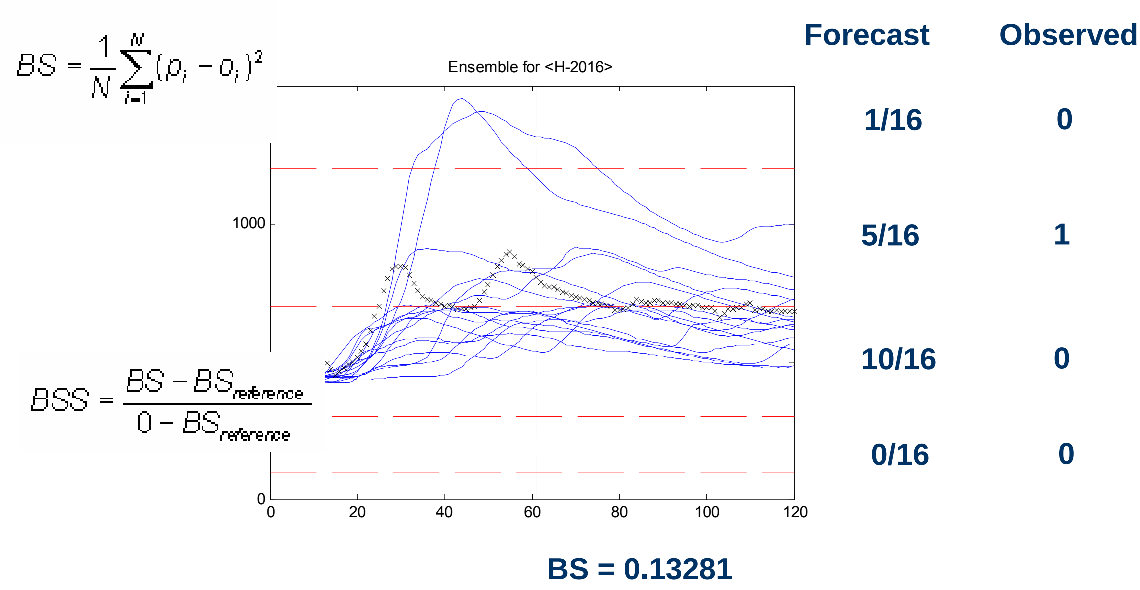


Figure 6.4 Example of a Brier skill score.

Use of the **continuous Brier skill score** will also be considered. This gives a balanced measure of location and spread of a forecast relative to outcome and is defined as the integrated mean square probability error:

$$CBS = n^{-1}\sum\int(O_i(x) - P_i(x))^2 dx$$

where:

$O_i(x)$ = indicator of event $o_i \leq x$ in the observed sample; equal to 1 if event $o_i \leq x$ does occur, 0 if not

$P_i(x)$ = probability of event $o_i \leq x$ occurring as stated in the probability forecast, value in the range 0 to 1.

o_i = observed value of sample *i* (*i* = 1, 2, ... *n*)

x = a variable threshold value covering all possible values of rainfall amount or rate.

Part 2: Operational implementation of ensemble forecasting in NFFS

7 Ensemble forecasting in NFFS

The aim is to configure ensemble forecasting in NFFS for two Environment Agency Regions – preferably not the South West Region where the pilot case study at Boscastle for forecasting of floods generated by convective storms is located. The configuration will be set up to run operationally on the test environment in Delft to which those Environment Agency staff involved will have access.

The potential and benefits of using NWP ensembles for flood forecasting were assessed in a workshop with a limited group of scientists and forecasters involved in the pilot application. Based on the outcome of this workshop, adjustments may be made to the configuration prior to presenting the results to a larger audience in the Phase 2 feedback workshop.

7.1 Selection of pilot regions

The plan is to make the use of ensembles generated by MOGREPS operational initially in two Environment Agency Regions. Criteria for selection of the pilot Region included:

- widespread application of conceptual rainfall-runoff models;
- fast running models;
- forecasting team interested in participating in the study.

Midlands, Thames and North East Regions were all considered suitable candidates.

- Midlands Region was fully covered with rainfall-runoff models (complemented with routing models). It has a varied topography and many of the models are gauged. In addition, the configuration in NFFS is such that the run-time of the models is relatively short so that the ensemble processing would not be hindered by a lack of computational resources. The downside is that Midlands applied an hour time step for forecasting instead of 15 minutes as in all other Regions.
- North East Region has short lead times to many of its upstream forecasting locations and is, to a large extent, covered by PDM models. The forecasting time step is 15 minutes and models run fast.
- Thames Region has longer lead times to its most important forecasting locations but has large urban areas within its forecasting responsibility. The Region is, to a large extent, covered with nested versions of its Thames Catchment Model (TCM). The forecasting time step is 15 minutes and models run fast. The nesting approach means that the larger currently models cover a long lead time, making them less beneficial when using nowcasting ensembles.

During the workshop, North East Region and Thames Region were selected as the pilot regions for testing the operational implementation of probabilistic forecasting. Midlands Region may be included if resources allow and there is a project requirement to do so.

7.2 STEPS and MOGREPS ensembles

Since spring 2007 the Met Office has used two systems to generate ensemble forecasts:

- STEPS – for short-term nowcasting of smaller scale short-lived weather features;
- MOGREPS – for short and medium range weather forecasting.

The ensemble forecasts for precipitation should provide the input to the testing of ensemble flood forecasting within NFFS in North East Region and Thames Region.

Background information from the Met Office on the ensemble prediction capability of STEPS and MOGREPS is presented below. The description of STEPS is added for sake of completeness; at this stage of the project, there are no plans to use STEPS ensembles.

7.2.1 STEPS

Nowcasting bridges the gap between telemetry and radar observations on the one hand, and numerical weather prediction on the other. For the first hours into the future, NWP is relatively unreliable. Nowcasting is therefore aimed at the prediction of the weather conditions for several hours ahead (up to six hours). It is run at much higher spatial and temporal resolutions in order to capture the smaller scale weather features.

Nimrod[18] and Gandolf[19] provided the Met Office's nowcasting capability until spring 2007 when Short Term Ensemble Prediction System (STEPS)[20] was introduced as a replacement.

STEPS provides ensemble prediction capability for nowcasting. This anticipates to the fact that the smaller scale weather features – like convective storms generating intensive flooding – are shorter lived and less predictable. With an ensemble prediction approach, the uncertainty of the nowcasts of weather condition can, to a certain extent, be quantified.

STEPS blends extrapolation of radar observations, noise and NWP on a hierarchy of scales. Output from STEPS includes ensemble rain rate and accumulations. An example domain is shown in Figure 7.1. Nowcasts are generated up to six hours ahead for a 2 km grid with a five minute time step.

The system produces a 50-member ensemble. Apart from the deterministic run, the individual members are currently not blended into the MOGREPS forecasts. A research project is underway to develop a methodology for this purpose.

[18] http://www.metoffice.gov.uk/water/nimrod.html
[19] http://www.metoffice.gov.uk/water/gandolf.html
[20] http://www.metoffice.gov.uk/science/creating/hoursahead/nowcasting.html

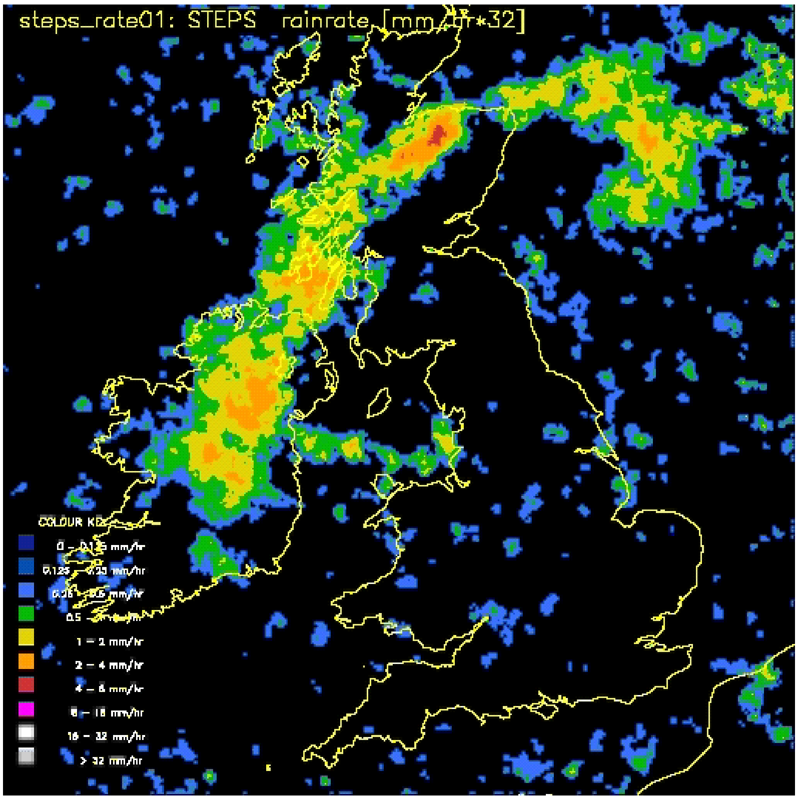


Figure 7.1 Example of STEPS output.

7.2.2 MOGREPS

In 2005 the Met Office introduced a new ensemble system called MOGREPS (Met Office Global and Regional Ensemble Prediction System)[21] which includes a 24 km resolution regional ensemble for the Atlantic and Europe. Ensemble forecasting is based on the principle of adding small perturbations to the best guess of the initial state of the atmosphere. The model is then run forward from the perturbed starting conditions to generate an ensemble of different forecasts.

The regional model (MOGREPS-R) is designed to provide ensemble forecasts for the short-range for the UK and Ireland (days 0–3). It provides a 24-member ensemble with a grid resolution of 24 km for a forecast length of 36 hours. Boundary conditions for the regional model are provided by a global model (MOGREPS-G) with a 90 km grid and a forecast time of 72 hours producing a 24-member ensemble. Both models are run twice daily at 0 and 12 UTC.[22] Model coverage is shown in Figure 7.2.

The ensembles consist of one control run and 23 additional members. The control forecast is run at the same resolution as the other ensemble members, but does not contain any perturbations to account for initial condition or model uncertainties; as such it runs from the best analysis of the initial state of the atmosphere. The control run can be compared with the standard deterministic weather forecast run at a 12 km resolution.

[21] http://www.metoffice.gov.uk/science/creating/daysahead/ensembles/MOGREPS.html
[22] Co-ordinated Universal Time

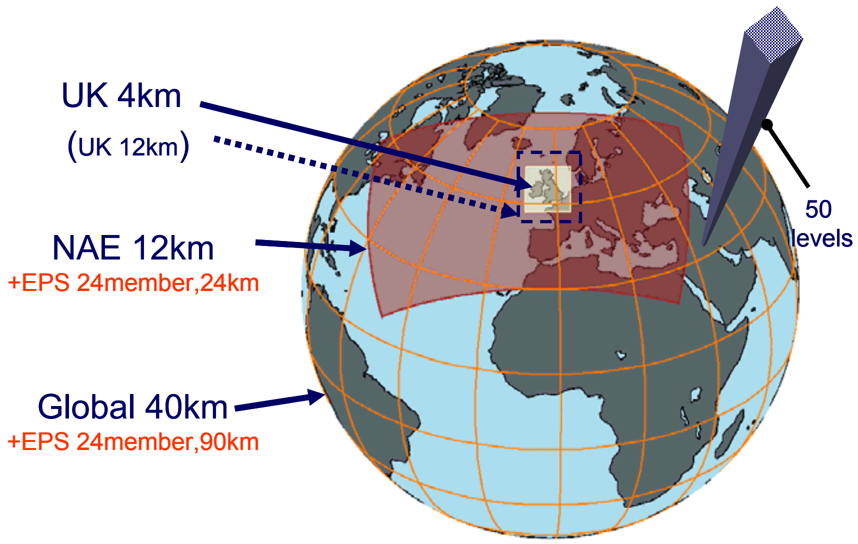


Figure 7.2 Model coverage in MOGREPS.

Notes: Source: Met Office

The 24 different predictions produced by the ensemble show a range of possible forecasts, allowing forecasters to quantify the uncertainty in an objective manner. If all 24 forecasts give similar solutions, this suggests a high confidence; when confidence is lower, the ensembles can help the forecaster to identify the most likely outcome, and also assess the risks of alternative solutions including more severe weather.

Meteorologists now believe that the ensemble prediction systems provide a method of quantitatively assessing the uncertainty associated with numerical weather prediction forecasts.

To provide a basis for probabilistic forecasting, meteorologists assume that the generated ensemble members have an equal probability. The latter is an important notion for when ensemble forecasting provides the quantitative basis for probabilistic flood forecasting.

7.2.3 Provision of MOGREPS ensembles

The MOGREPS ensembles will be provided by Met Office on a real-time basis to Deltares and CEH Wallingford. This will allow simulation of the use of these data for real-time forecasting on the test environment in Delft. Following this research and a test period, routine usage of the MOGREPS ensembles within the Environment Agency may be introduced.

The ensembles will be sent in Nimrod format. It is important that all time-steps are sent in a single file.

The minor developments in Delft-FEWS needed to accommodate both formats (STEPS and MOGREPS) are catered for within its development budgets. The Met Office and the Environment Agency will need to decide about the file formats and how to make the data feed to Delft operational. In the meantime, samples for the STEPS and MOGREPS ensembles can be made available in Nimrod format to allow for implementation of the required NFFS configuration changes and initial testing.

7.3 Configuring ensemble forecasting in NFFS

The configuration will be based on the current configuration of NFFS. No distributed models will be run.

The configuration changes will include:

- importing and processing of NWP ensembles (STEPS and MOGREPS);
- ensemble runs of forecasting models;
- data displays, including statistical analyses;
- performance indicators focused on testing probabilistic forecasting skill.

The NFFS configurations for North East, Thames and potentially Midlands Regions will be extended to process the NWP ensembles and display probabilistic forecast results.

7.4 Test environment

A test environment will be set up in Delft on which the complete NFFS configurations for North East and Thames Regions will be set up. Ensemble processing will be configured in the current (offline) NFFS configurations.

A limited number of Environment Agency staff will be given access to the test system via VPN to facilitate testing in a semi live environment.

The test systems will be fed with live data feeds provided by the Environment Agency and the Met Office.

The systems will run continuously and will, in principle, be permanently accessible. However no service level will be guaranteed.

The test environment will be set up with low cost system software detailed in Table 7.1.

Table 7.1 System software.

Item	Software	Existing live system at the Environment Agency
Operating system	Linux RedHat	HP-UX
Application server	Jboss	WebLogic
Database	PostgreSQL 8	Oracle 9i

7.5 Forecast performance

Evaluation of the system will be based on several parameters including:

- forecast accuracy (using a number of skill scores/performance indicators);
- forecast run times.

The skill scores and performance indicators used are listed below.

A factor that affects the quality of EPS streamflow forecasts is model bias (see Hashino et al. 2007). Within NFFS most forecasting locations have error correction base on observed flows applied. Therefore, we will use the error corrected ensemble forecast to evaluate forecast performance.

More information on this topic is given in the Phase 2 report.

7.5.1 Simple ensemble statistics

For each forecast location that is part of the pre-defined display, the following ensemble parameters will be determined for each time step:

- ensemble max – the maximum value in the ensemble;
- ensemble min – the minimum value in the ensemble;
- 25, 75, 33 and 66 percentiles in the ensemble;
- ensemble mean – the mean value of the ensemble.

The main purpose of these parameters is to set up forecast displays such as shown in Figure 7.3. They can also be used to set up contingency tables.

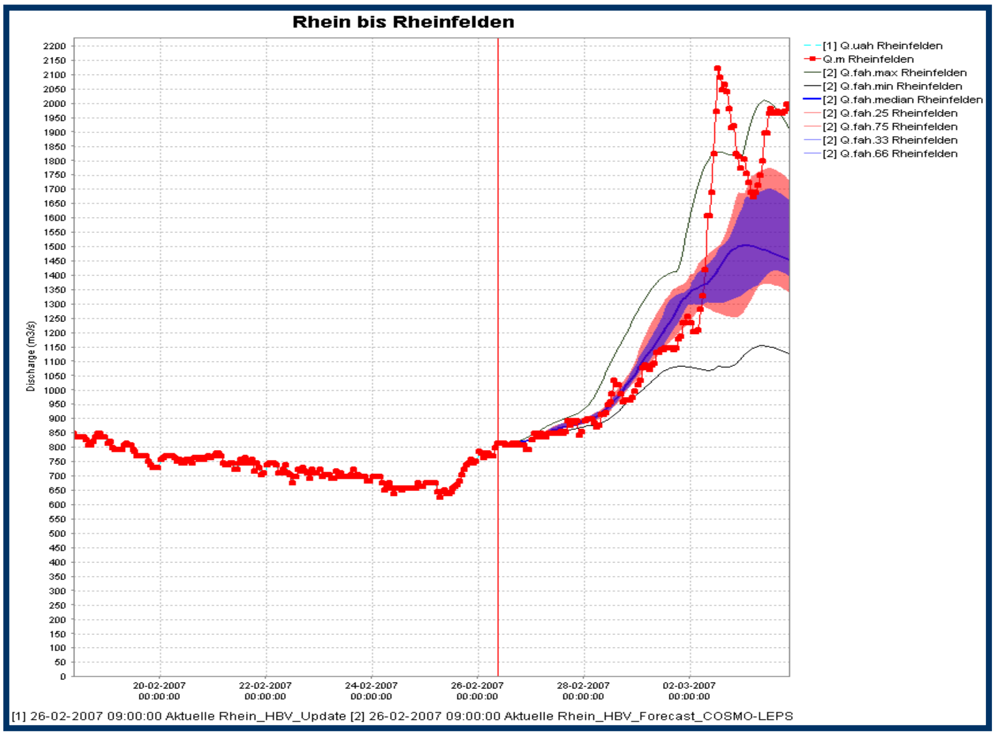


Figure 7.3 Example of an ensemble forecast showing the 25, 33, 66, 75 percentile, the ensemble maximum and ensemble minimum compared to the deterministic forecast and observed data.

7.5.2 Probability of streamflow exceeding a certain level

For locations with defined thresholds, a probability diagram can be produced showing the change of streamflow exceeding a threshold. In the example shown in Figure 7.4, each row is a separate forecast.

Such diagrams could be used to supplement the catchment threshold status tables used in current NFFS reports. This type of diagram also provides an easy visual check of forecast stability.

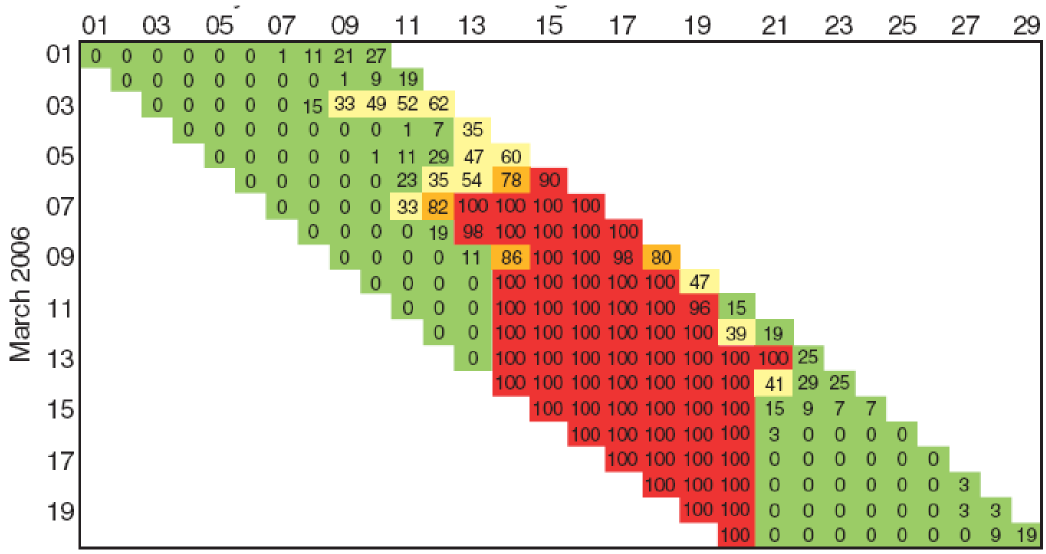


Figure 7.4 Probability diagram for streamflow exceeding a certain threshold.

Notes Source: ECMWF Newsletter No. 111, Spring 2007.

7.5.3 Brier score and Brier skill score

The Brier score is used to measure the performance of ensemble forecasts. The skill score provides a measure to compare the skill of the probabilistic forecast with a deterministic forecast. Computation of both will be configured. For further discussion of the Brier scores and Skill Scores, see Section 6.5.3.

7.6 System performance

System performance will also be analysed. Running NFFS in ensemble mode will have a significant impact on the forecast run times because the forecast workflow has to be repeated 30 times. The practical use of ensemble forecasts will therefore be assessed at this stage.

The performance of the following aspects will be quantified and analysed for each of the configurations:

- forecast run times;
- database size;
- synchronisation times and network load.

Possible areas for improvement of system performance will be identified. Recommendations will then be formulated how to optimise system performance.

References

Abbott M B, Bathurst J C, Cunge J A, O'Connell P E and Rasmussen J, 1986a An introduction to the European Hydrologic System–Système Hydrologique Européen, 'SHE' 1: history and philosophy of a physically based, distributed modelling system. *Journal of Hydrology*, 87, 45-59.

Abbott M B, Bathurst J C, Cunge J A and O'Connell P E, 1986b An introduction to the European Hydrologic System–Système Hydrologique Européen, 'SHE' 2: structure of a physically based, distributed modelling system. *Journal of Hydrology*, 87, 61-77.

Bergström A, 1995 *The HBV model*. In Computer Models of Watershed Hydrology. (ed. V P Singh), pp. 443-476. Highlands Ranch, CO: Water Resources Publications.

Beven K J, 1989 Changing ideas in hydrology – the case of physically-based models. *Journal of Hydrology*, 105, 157-172.

Beven K J, 2002 Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system. *Hydrological Processes*, 16, 189-206.

Bell V A, Kay A L, Jones R G and Moore R J, 2007 Development of a high resolution grid-based river flow model for use with regional climate model output. *Hydrology and Earth System Sciences,* 11(1), 532-549.

Box G E P and Jenkins G M, 1970 *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

Bronstert A, 1999 Capabilities and limitations of detailed hillslope hydrological modelling. *Hydrological Processes*, 13: 21-48.

Burnash R J C, Ferral R L and Mcguire R A, 1973 *A generalized streamflow simulation system. Conceptual modelling for digital computers.* Technical report. Silver Spring, MD: National Weather Service.

CEH Wallingford, 2005 PDM Rainfall-Runoff Model. Version 2.2 (includes guide, practical user guides, user manual and training exercises). Wallingford: Centre for Ecology & Hydrology.

Freeze R A and Harlan R L, 1969 Blueprint for a physically-based, digitally-simulated hydrologic response model. *Journal of Hydrology*, 9: 237-258.

Gray W G, Leijnse A, Kolar R L and Blain C A, 1993 *Mathematical Tools for Changing Spatial Scales in the Analysis of Physical Systems*. Boca Raton, FL: CRC Press.

Han D, 1991 *Weather radar information processing and real-time flood forecasting*. PhD thesis, University of Salford.

Hashino T, Bradley A A and Schwartz S S, 2007 Evaluation of bias-correction methods for ensemble streamflow volume forecasts. *Hydrology and Earth System Sciences*, 11, 939-950.

Moore R J, 1985 The probability-distributed principle and runoff production at point and basin scales. *Hydrological Sciences Journal*, 30(2), 273-297.

Moore R J, 1999 *Real-time flood forecasting systems: perspectives and prospects*. In Floods and Landslides: Integrated Risk Assessment (ed. R Casale and C Margottini), pp. 147-189. Berlin: Springer.

Moore R J, 2007 The PDM rainfall-runoff model. *Hydrology and Earth System Sciences*, 11(1), 483-499.

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, Watson B C, Jones D A and Black K B, 1991 *Local recalibration of weather radar*. In Hydrological Applications of Weather Radar (ed. I D Cluckie and C G Collier), pp. 65-73. Chichester: Ellis Horwood.

Moore R J, Bell V A, Cole S J and Jones D A, 2006a *Spatio-temporal rainfall datasets and their use in evaluating the extreme event performance of hydrological models.* R&D Project Report FD2208. Report to the Environment Agency and Defra. London: Defra.

Moore, R.J., Cole, S.J., Bell, V.A., Jones, D.A. 2006b. *Issues in flood forecasting: ungauged basins, extreme floods and uncertainty.* In Frontiers in Flood Research, Proceedings of 8th Kovacs Colloquium, UNESCO (Paris 2006) edited by I Tchiguirinskaia, K N N Thein and P Hubert, IAHS Publication 305, 103-122. Wallingford: IAHS Press.

Moore, R.J., Bell, V.A., Cole, S.J., Jones, D.A. 2007. *Rainfall-runoff and other modelling for ungauged/low-benefit locations*. R&D Technical SC030227/SR1. Bristol: Environment Agency. Available from: http://publications.environment-agency.gov.uk/pdf/SCHO0307BMER-e-e.pdf [Accessed 16 September 2009].

O'Connor K M, 1982 Derivation of discretely coincident forms of continuous linear time-invariant models using the transfer function approach. *Journal of Hydrology*, 59, 1-48.

Reggiani P and Schellekens J, 2003 Modelling of hydrological responses: the representative elementary watershed approach as an alternative blueprint for watershed modelling. *Hydrological Processes*, 17, 3785-3789.

Paniconi C, Marrocu M, Putti M and Verbunt M, 2003 Newtonian nudging for a Richards equation-based distributed hydrological model. *Advances in Water Resources*, 26(2), 161–178.

Reggiani P, Sivapalan M and Hassanizadeh S M, 1998 A unifying framework of watershed thermodynamics: 1. Balance equations for mass, momentum, energy and entropy and the second law of thermodynamics. *Advances in Water Resources*, 22(4), 367-398.

Reggiani P, Sivapalan M, Hassanizadeh S M and Gray W G, 1999 A unifying framework of watershed thermodynamics: 2. Constitutive relationships. *Advances in Water Resources,* 23(1), 15-39.

Reggiani P, Sivapalan M and Hassanizadeh S M, 2000 Conservation equations governing hillslope responses. *Water Resources Research*, 38(7), 1845-1863.

Vertessy R A, Hatton T J, O'Shaughnessy P J and Jayasuriya M D A, 1993 Predicting water yield from a mountain ash forest catchment using a terrain analysis based catchment model. *Journal of Hydrology*, 150, 665-700.

Vrugt J A, Nuallain B O, Robinson B A, Bouten W, Dekker S C and Sloot P M A, 2006 Application of parallel computing to stochastic parameter estimation in environmental models, *Computers & Geosciences*, 32(8), 1139-1155.

Yapo P O, Gupta H V and Sorooshian S, 1998 Multi-objective global optimization for hydrologic models, *Journal of Hydrology*, 204, 83-97.

Young P C and Tomlin C M, 2000 *Data-based mechanistic modelling and adaptive flow forecasting.* In Flood Forecasting: What does Current Research offer the Practitioner? (ed. M Lees and P Walsh). BHS Occasional Paper No. 12. British Hydrological Society

List of abbreviations

BS	Briar score
BSS	Briar skill score
CEH	Centre for Ecology & Hyrology
CSI	critical success index
DTM	Digital Terrain Model
ECMWF	European Centre for Medium-Range Weather Forecasts
ET	evapotranspiration
FAR	false alarm rate
FEWS	Flood Early Warning System [Deltares]
FFOT	first forecast of threshold
G2G	Grid-to-Grid [model]
GA	general adapter [Delft-FEWS]
HOST	Hydrology of Soil Types
IHDM	Institute of Hydrology Distributed Model
JCMM	Joint Centre for Mesoscale Meteorology [Met Office]
LCM2000	Land Cover Map 2000
MAE	mean absolute error
MCRM	modified conductive rock matrix
MOGREPS	Met Office Global and Regional Ensemble Prediction System
MOSES	Met Office Surface Exchange Scheme
MSE	mean square error
NAM	North American Mesoscale Model
NFFS	National Flood Forecasting System
NGR	National Grid Reference
NWP	Numerical Weather Prediction
ODE	ordinary differential equation
P	probability
PDM	PDM Probability Distributed Moisture [model]
PE	potential evaporation
PI	processing instruction
POD	probability of detection
PRTF	Physically Realisable Transfer Function Model

REW	Representative Elementary Watershed [model]
RMS	root mean square
RMSE	root mean square error
RMSF	root mean square factor
SEPA	Scottish Environment Protection Agency
STEPS	Short Term Ensemble Prediction System
TCM	Thames Catchment Model
UN	Unified Model [Met Office]
UTC	Co-ordinated Universal Time
VPN	virtual private network

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.