

delivering benefits through evidence



Hydrological modelling using convective
scale rainfall modelling – phase 2

Project: SC060087/R2

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: 978-1-84911-182-9

© Environment Agency – February, 2010

All rights reserved. This document may be reproduced with prior permission of the Environment Agency.

The views and statements expressed in this report are those of the author alone. The views or statements expressed in this publication do not necessarily represent the views of the Environment Agency and the Environment Agency cannot accept any responsibility for such views or statements.

This report is printed on Cyclus Print, a 100% recycled stock, which is 100% post consumer waste and is totally chlorine free. Water used is treated and in most cases returned to source in better condition than removed.

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:

SCHO0210BRYQ-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 2 (pilot case study). During Phase 1 (inventory and data collection), the storm event at Boscastle on 16 August 2004 was selected as the case study to test various hydrological modelling concepts for the transformation of high resolution rainfall predictions into accurate flood forecasts.

Three hydrological models (one lumped and two distributed) were applied to the north Cornish catchments affected by the Boscastle event:

- Probability Distributed Moisture (PDM) model;
- physical-conceptual Grid-to-Grid (G2G) model;
- physics-based Representative Elementary Watershed (REW) model.

The three models were configured and calibrated for the three gauged catchments (Ottery, Tamar and Camel) selected as the focus of the case study. Raingauge-adjusted radar rainfall data produced using HyradK were used as model input.

This phase also coupled the latest Met Office high resolution NWP products with the distributed hydrological model developed by the Centre for Hydrology & Ecology at Wallingford and considered the future potential of ensemble convective scale rainfall predictions. The term 'distributed forecasting' in this sense means the use of a spatially distributed (grid-based) hydrological model to forecast 'everywhere'. This contrasts with current hydrological model networks that comprise a connected set of (normally) lumped rainfall-runoff ('catchment') models feeding into hydrological and hydrodynamic river models which provide forecasts only at specific locations.

Ensemble forecasting was configured in a test NFFS system for two Environment Agency Regions (Thames and North East) set up at Deltares in the Netherlands as part of the Delft Flood Early Warning System (Delft-FEWS) and receiving MOGREPS forecasts from the Met Office. Particular attention was given to the effect on system performance as it is necessary, when running the models in ensemble mode, to repeat the forecast workflow 24 times.

Conclusions

Phase 2 demonstrated that a distributed hydrological model (set up using a digital terrain model) can be operated on the National Flood Forecasting System platform with short enough run times for use in real-time forecasting.

The PDM model gave excellent performance across catchments but was insensitive to the storm pattern. The G2G model gave good performance across catchments; ungauged performance was also good. The REW model gave good performance for winter periods but overall performance was only reasonable; ungauged performance was on a par with gauged performance.

Distributed models were considered to have a number of advantages including:

- sensitivity to spatio-temporal structure of storms;
- helpful in understanding storm and catchment shaping of flash floods;
- can identify locations vulnerable to flooding;
- help forecast floods shaped by 'unusual' storm and catchment conditions absent from the historical record;
- provide a complete spatial picture of flood hazard across a region;
- respond sensibly to ensemble rainfall forecasts that vary in position.

The performance of the G2G model was considered particularly promising with a number of its attributes being relevant to convective scale probabilistic flood forecasting.

Recommendations

- **Hydrological modelling concepts to be carried through to Phase 3.**
 - The distributed G2G model shows promise for Area-wide flood forecasting at gauged and ungauged locations.
 - An extended G2G formulation incorporating soil/geology datasets should be considered.
- **Case study selection for Phase 3.** A more focused 'regional assessment' should be undertaken based on an area within Midlands Region affected by the summer 2007 floods and utilising raingauge data in combination with radar data for improved rainfall estimation as model input. More detailed analyses (including use of high resolution NWP pseudo ensembles) should be performed for the selected case study area in the Midlands Region.
- **Continuation of the MOGREPS trial into Phase 3.** This will allow:
 - more events to be captured;
 - additional forecast stability tables to be added to the test configuration;
 - a wider range of Environment Agency staff to assess the results.

A more radical recommendation is to trial the G2G model countrywide across England and Wales in Phase 3.

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

1	Introduction	1
2	Project approach	3
2.1	Project objectives	3
2.2	Using convective scale rainfall forecasts in NFFS	3
2.3	Operational implementation of ensemble forecasting	6
2.4	Project overview	7
Part 1: Using convective scale rainfall forecasts in NFFS		8
3	Case study event	9
3.1	Meteorological synopsis	9
3.2	Flood damage	10
4	Hydrological models	11
4.1	Probability Distributed Moisture (PDM)	11
4.2	Grid-to-Grid model (G2G)	13
4.3	Representative Elementary Watershed (REW) model	16
5	Model configuration and calibration: application to the case study catchments	25
5.1	Catchment information	25
5.2	Configuration	28
5.3	REW model application	32
5.4	PDM application	51
5.5	G2G application	58
6	Analysis of use of high-resolution NWP forecasts	70
6.1	High resolution NWP forecasts	70
6.2	Generation of pseudo ensembles	71
6.3	Analysis of hydrological model forecasts using HyradK data	73
6.4	Analysis of hydrological model forecasts using deterministic high resolution NWP rainfalls	77
6.5	Analysis of hydrological model forecasts using pseudo ensembles of high resolution NWP rainfalls	82
Part 2: Operational implementation of ensemble forecasting in NFFS		87
7	Ensemble forecasting in NFFS using MOGREPS	88
7.1	Selection of pilot Environment Agency Regions	88
7.2	STEPS and MOGREPS ensembles	88
7.3	Configuring ensemble forecasting in NFFS	91

7.4	Test environment	102
7.5	Effects of (hydrological) model concept and size	102
7.6	System performance	115
8	Conclusions	120
8.1	Models	120
8.2	Implementation of MOGREPS ensembles	121
9	Recommendations	122
10	Proposed plan of work for Phase 3 (modelling)	123
10.1	Case study outline	123
10.2	Principal operational output	124
10.3	MOGREPS trial	124
	References	125
	List of abbreviations	128
	Appendix A: Report of the completion workshop	130
	Appendix B: Raingauge data quality control	136
	Appendix C: Instructions for using the test system	141
	Appendix D: Database optimisation using SQL queries	145
	Appendix E: Importing and processing pseudo ensembles with Delft-FEWS	152
	List of tables and figures	
Table 4.1	Parameters of the PDM model	12
Table 4.2	Hydrological fluxes within the REW model	24
Table 5.1	Gauging station details (main case study catchments are in bold)	26
Table 5.2	Raingauges used by HyradK	30
Table 5.3	Calibration and verification periods	31
Table 5.4	REW model parameters for Camel catchment	34
Table 5.5	REW model parameters for Tamar catchment	37
Table 5.6	Model performance for the calibration and verification periods, January 2006 to August 2007 and 2004–2005 respectively	39
Table 5.7	Event periods (start and end times are at 00:00)	46
Table 5.8	Settings used in module application	48
Table 5.9	R^2 model efficiency for all events ¹	50
Table 5.10	PDM model parameters for the Camel, Ottery and Tamar catchments	52
Table 5.11	Simulation mode performance of G2G and PDM models over calibration and verification periods	55
Table 5.12	G2G model parameters for the Camel and Tamar catchments	60
Table 6.1	Summary of different rainfall estimators accumulated over the five-hour period ending 17:00 16 August 2004 ¹	72
Table 6.2	Catchment average rainfall totals (mm) for the three case study catchments using different sources of rainfall ¹	79
Table 7.1	Files updated in North East Region NFFS configuration for import and display of MOGREPS data	92
Table 7.2	Files updated in North East Region NFFS configuration for processing of MOGREPS data	93
Table 7.3	Other files adjusted in North East Region NFFS configuration	94
Table 7.4	Other config files updated in North East Region NFFS	95
Table 7.5	Files updated in order to compute statistics	96
Table 7.6	Files updated in order to generate MOGREPS reports	98
Table 7.7	Files updated in Thames Region NFFS configuration in order to import and display MOGREPS data	99
Table 7.8	Thames Region South East area MOGREPS thresholds	101
Table 7.9	Thames Region North East area MOGREPS thresholds	101

Table 7.10	Thames Region West area MOGREPS thresholds	102
Table 7.11	System software	102
Table 7.12	Parameters in the Thames Catchment Model	105
Table 7.13	System specifications used in the test system	116
Table 7.14	Run times of the ensemble forecast in the Thames and North East Region test systems	117
Table 7.15	Estimated run times for the complete region forecast in ensemble mode ¹	117
Table 8.1	Summary of model performance	120
Table 8.2	Summary of general conclusions	121
Table B1	15-minute raingauge totals in excess of 70 mm (all checked as erroneous and set as missing within HyradK processing)	136
Table B2	15-minute raingauge totals in the range 20–70 mm (all checked and comments noted)	138
Table B3	A record of all the periods of raingauge record treated as missing (apart from 15-minute totals in excess of 70mm listed in Table B1)	139
Table D1	syncLevels – selected by convention	146
Table D2	Field in time series table	148
Table D3	Result for example query 2	149
Figure 3.1	Accumulations of five-minute Nimrod radar rainfall data for the seven-hour period starting at 10:00 16 August 2004	10
Figure 4.1	Overview of the PDM rainfall-runoff model	11
Figure 4.2	Grid-to-Grid distributed hydrological model	13
Figure 4.3	Binary structure of the channel network	17
Figure 4.4	A representative elementary watershed (REW) as a 3D spatial region	18
Figure 4.5	Overlap of a REW map with a soil map yields a smaller subdivision of the unsaturated zone within a REW into RECs	18
Figure 4.6	Conceptualisation of a REW in the model	20
Figure 4.7	Implementation of the Hardy-Cross method for the groundwater system	20
Figure 4.8	Calculation of the groundwater flow field for the Geer basin (Belgium)	21
Figure 4.9	Water table surface interpolated with the bi-cubic spline method	21
Figure 4.10	View of hydrological processes represented in the REW model	22
Figure 4.11	Cross section of the REW showing the river channel	23
Figure 5.1	Map of relief over the Boscastle region	26
Figure 5.2	Maps over the Boscastle region of (a) solid geology and (b) HOST	27
Figure 5.3	Modelling area	28
Figure 5.4	Daily evaporation profile used for REW modelling	29
Figure 5.5	Map of raingauges selected for use with HyradK	30
Figure 5.6	Camel watershed broken down into 51 REWs	34
Figure 5.7	Modelled and observed discharges at Denby, 1 January 2005 to 1 October 2007 (precipitation at REW 1)	36
Figure 5.8	Tamar watershed broken down into 81 REWs	36
Figure 5.9	Modelled and observed discharges at Gunnislake, 1 January 2005 to 1 October 2007 (precipitation at REW 1)	38
Figure 5.10	Modelled and observed discharges at Werrington, 1 January 2005 to 1 October 2007 (precipitation at REW 1)	38
Figure 5.11	Measured and modelled discharge for 18 October to 14 November 2005 for Gunnislake (top) and Werrington (bottom)	40
Figure 5.12	Measured and modelled discharge for 18 October to 14 November 2005 for Denby (top) and Crowford Bridge (bottom)	41
Figure 5.13	Results for a winter event for the internal sites Crowford Bridge (top), Lifton Park (middle) and Slaughterbridge (bottom)	42
Figure 5.14	Measured and modelled discharge for 18 October to 14 November 2005 for four internal sites	43
Figure 5.15	Results for the Boscastle event (in raw simulation mode) for, from top left to bottom right, Gunnislake, Werrington, Crowford Bridge (internal site) and Denby	44
Figure 5.16	Model efficiency (R^2) versus lead time for periods 2–5 for Gunnislake, Denby and Werrington	48
Figure 5.17	Model efficiency (R^2) versus lead time for periods 6–9 for Gunnislake, Denby and Werrington	49
Figure 5.18	Model efficiency (R^2) versus lead time for periods 10–12 for Gunnislake, Denby and Werrington	49
Figure 5.19	Modelled (PDM and G2G) and observed hydrographs at the Gunnislake, Werrington and Denby locations for the calibration period 9 January 2005 to 1 October 2007	53
Figure 5.20	Modelled (PDM and G2G) and observed hydrographs at the Gunnislake, Werrington and Denby locations for the evaluation period 1 January 2004 to 23 December 2005	55
Figure 5.21	Modelled (PDM and G2G) and observed hydrographs at the Gunnislake, Werrington and Denby locations for the evaluation winter period 18 October to 14 November 2005 and the summer period 19 to 27 May 2006	56
Figure 5.22	Modelled and observed hydrographs using the PDM and G2G models for the period 15 to 20 August 2004 covering the Boscastle event	57
Figure 5.23	Model efficiency (R^2) versus lead time for periods 1–12 and the calibration and evaluation period for Gunnislake, Denby and Werrington	58
Figure 5.24	Modelled and observed hydrographs using the PDM and Tamar G2G models for the period 19 November to 27 December 2006	61
Figure 5.25	Modelled and observed hydrographs using the PDM and Camel G2G models for the period 19 November to 27 December 2006	62
Figure 5.26	Model efficiency (R^2) for PDM and G2G over the calibration and evaluation periods	64
Figure 5.27	Modelled and observed hydrographs using G2G and extended G2G for the period 19 November to 27 December 2006	67
Figure 5.28	Model efficiency (R^2) for G2G and extended G2G over the calibration and evaluation periods	68
Figure 5.29	Modelled and observed hydrographs using the PDM, G2G and extended G2G models for the period 15 to 20 August 2004	69

Figure 6.1	Forecast rainfall data from the high resolution NWP model as displayed in Delft-FEWS	70
Figure 6.2	Rainfall accumulations (mm) for the five-hour period ending 17:00 16 August 2004 using different rainfall sources.	72
Figure 6.3	Rainfall accumulations (mm) for the five-hour period ending 17:00 16 August 2004 using different rainfall sources and accumulated over 12 km grid squares	73
Figure 6.4	Forecast data from the high resolution NWP model as displayed in Delft-FEWS	74
Figure 6.5	PDM simulation mode results using HyradK raingauge-adjusted rainfall data (left column) and fixed origin forecast results using state correction and HyradK 'perfect' rainfall forecasts (right column).	75
Figure 6.6	G2G fixed origin forecast results using state correction and HyradK raingauge-adjusted radar data as 'perfect' rainfall forecasts	76
Figure 6.7	REW fixed lead time forecasts using HyradK-adjusted precipitation for the Boscastle event for Denby	77
Figure 6.8	Fixed origin PDM model forecasts at Werrington (top row), Gunnislake (middle row) and Denby (bottom row) using 1, 4 and 12 km deterministic NWP and HyradK 'perfect' rainfall data.	79
Figure 6.9	Rainfall accumulations (mm) for the 5 hour period ending 17:00 16 August 2004 using 1 km (top left), 4 km (top right) and 12 km (bottom left) km NWP and HyradK raingauge-adjusted radar (bottom right)	80
Figure 6.10	Fixed-origin G2G model forecasts at Werrington using 1, 4 and 12 km deterministic NWP and HyradK 'perfect' rainfall data: (a) using the 00UTC forecast and (b) using the 03UTC forecast	81
Figure 6.11	Fixed origin G2G Model forecasts at Denby using 1, 4 and 12 km deterministic NWP and HyradK 'perfect' rainfall data: (a) using the 00UTC forecast and (b) using the 03UTC forecast	81
Figure 6.12	PDM model forecasts at Werrington, Gunnislake and Denby using the 00UTC 1km NWP pseudo ensemble	83
Figure 6.13	G2G Model forecasts at Werrington and Gunnislake using the 00UTC 1km NWP pseudo-ensemble	84
Figure 6.14	REW forecasts for Gunnislake and Denby made using the pseudo ensembles as input.	85
Figure 6.15	REW forecasts for Werrington and Slaughterbridge made using the pseudo ensembles as input	86
Figure 7.1	Model Coverage in MOGREPS	90
Figure 7.2	Display of single ensemble member in the T46 test system (at 18:00 20 March 2008)	91
Figure 7.3	Example of a pre-defined display	97
Figure 7.4	Example report of North East Region MOGREPS results	98
Figure 7.5	Representation of a hydrological response zone within the Thames Catchment Model	104
Figure 7.6	Ensemble member 0, 1, 2 and 3 of the MOGREPS forecast on 21 March used in the examples below	106
Figure 7.7	Albany Park catchment	107
Figure 7.8	Edmonton Green catchment	107
Figure 7.9	Wheatley catchment	108
Figure 7.10	Forecast for 11:00 20 March for Albany Park	109
Figure 7.11	Forecast for 23:00 20 March for Albany Park	109
Figure 7.12	Forecast for 11:00 20 March for Edmonton Green	110
Figure 7.13	Forecast for 23:00 20 March for Edmonton Green	110
Figure 7.14	Forecast for 11:00 20 March for Wheatley	111
Figure 7.15	Forecast for 23:00 20 March for Wheatley	112
Figure 7.16	15-minute rainfall accumulations at Farnley Hall, West Yorkshire	112
Figure 7.17	Example 1 – added value: Chesterfield T0 09:00 11 January 2008	113
Figure 7.18	Example 2 – Low probability extreme event: Denby Dale Road T0: 09:00 11 January 2008	114
Figure 7.19	Example 3 – hydrological model, initial conditions: Gypsy Lane T0 09:00 15 January 2008	115
Figure 7.20	Components of the FHSnet Delft-FEWS test system	116
Figure C1	Select the Ensemble Grid Data (Met Office) check box	141
Figure C2	Example of ensemble results for TCM/ARMA from the filters	142
Figure C3	Example of pre-defined display of ensemble results	143
Figure C4	Example ensemble results from for PDM/ARMA/KW?SNOW from the filters	144
Figure D1	Layout of the Delft-FEWS live system	145
Figure D2	Result for example query 4	150
Figure D3	Example output for example query 5	150
Figure E1	Converter program to help screen with command-line options	152

1 Introduction

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

- In the very near future, the nowcasting system STEPS¹ will become operational at 2 km resolution.
- For longer term numerical weather prediction, the Met Office has developed a new system called MOGREPS² 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 project³ 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 2. The storm/flood event at Boscastle on 16 August 2004 was selected during Phase 1 as the pilot case study. Boscastle is situated within the Environment Agency's South West Region. Two other Regions (North East and Thames) were therefore selected to act as pilots for ensemble forecasting in NFFS.

The project was undertaken for the Environment Agency by Deltares⁴ 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 NFFS

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 during Phase 1 for testing in Phase 2 on 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.

The use of convective scale rainfall forecasts in NFFS is discussed in Part 1 of this report.

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 pilot case studies,⁵ 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 was compiled during Phase 1. 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 of interest because they were currently applied in the areas of the potential pilot case studies were:

- 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 the Phase 1 report.

⁵ See Section 3 of the Phase 1 report.

All modelling concepts tested had to be able to run in Delft-FEWS.⁶ Delft-FEWS module adapters are available for all models currently used in NFFS; there was already an adapter for the REW model and a new module adapter for the G2G model was developed during the early stages of Phase 2.

During Phase 1, 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 HyradK⁷ 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.

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 was 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 were 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 were added to the current configuration for the Environment Agency Region where the pilot case study is located. This test configuration was installed on a system in Delft accessible to Environment Agency staff.

During Phase 1, rainfall products were generated containing multiple forecast scenarios from the high resolution NWP output for the potential pilot case study. The rainfall products were fed into the hydrological models to produce probabilistic forecasts within NFFS (following current forecasting procedures as much as possible). The forecasts were produced and analysed only for the period covered by the pilot case study.

The 'raw' hydrological forecast data were processed to form probabilistic forecasts and associated information. Existing presentation methods were developed for presenting spatial distributed forecasts and probabilistic forecasting data. These methods were used to

⁶ Deltares' Flood Early Warning System (see <http://www.wldelft.nl/soft/fews/int/index.html>)

⁷ HYRAD radar hydrology kernel

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 was 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 was also investigated. Objective functions (performance indicators) were 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 was carried out initially for the two Environment Agency Regions selected during Phase 1. Criteria for selection included a serious interest in probabilistic forecasting as the Region's forecasting team would be more intensively involved in the pilot than teams from other Regions. The selected Regions (North East and Thames) were not Environment Agency Regions 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 was configured in NFFS for the selected Regions. The configuration was based on the current configuration of NFFS. No distributed models were run in this test case. The configuration included:

- 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 were assessed during Phase 2 at a workshop with the limited group of scientists and forecasters involved in the pilot application (see Appendix A). The workshop completed Phase 2 of the project. Any adjustments to the configuration will be made prior to presenting the results to a larger audience in a feedback workshop.

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 Case study event

During Phase 1 of the project, the extreme convective event that affected Boscastle and the surrounding area on 16 August 2004 was chosen as the case study event for Phase 2.

Following the devastating floods associated with the storm, the Environment Agency commissioned a review of the meteorology, hydrology, hydraulics and impacts of the event (HR Wallingford 2005). The Met Office has also provided a detailed account of the meteorology (Golding 2005).

A brief meteorological synopsis and summary of the flood damage is given below. A detailed description of the catchments used in the case study is given in Section 5. The hydrological model results are presented in Sections 5 and 6.

3.1 Meteorological synopsis

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 3.1, which shows the rainfall accumulation using Nimrod composite radar data over the event.

The extreme rainfall event was captured by a network of tipping-bucket raingauges (triangles in Figure 3.1) and daily storage gauges. Three gauges were situated near the core of the storm and confirmed the presence of extreme rainfall totals.

In the 24-period to 09:00 GMT 17 August 2004, the daily storage gauge at Otterham (SX 169 916) recorded 200.4 mm, the daily storage gauge at Trevalec (SX 134 900) recorded 184.9 mm and the tipping bucket raingauge at Trevalec recorded 155.2 mm, of which 153.6 mm fell in a six-hour period. The discrepancy between this and the Trevalec daily raingauge totals is most likely to be due to known problems with tipping bucket raingauges during intense rainfall events.

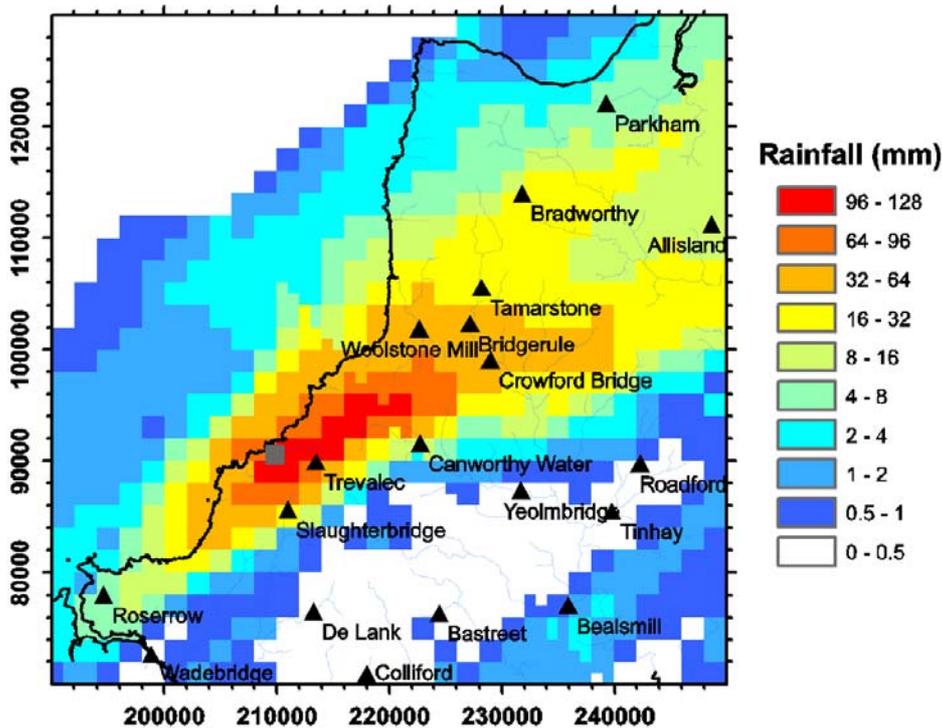


Figure 3.1 Accumulations of five-minute Nimrod radar rainfall data for the seven-hour period starting at 10:00 16 August 2004.

Notes: Solid triangles denote raingauge locations.
Northings and eastings are for National Grid Reference (NGR) co-ordinates in metres.

3.2 Flood damage

Several catchments across north Cornwall were affected by the resulting floods. The most severe flooding occurred on the Valency and Crackington Stream though the rivers Ottery and Neet also flooded.

A report commissioned by the Environment Agency (HR Wallingford 2005) contains a detailed account of the considerable damage caused to Boscastle and Crackington Haven. Flash flooding affected at least 100 homes and businesses with a total of six properties being destroyed. Roads, bridges and other infrastructure were badly damaged and 115 vehicles were swept away. Fortunately, due to the quick response of the emergency services, no lives were lost but around 100 people were rescued by helicopter. Other notable effects of the flash flood were the numerous trees swept away, causing trash dams and several new flow paths cut by the flows.

4 Hydrological models

The next sections introduce the modelling concepts applied in Phase 2. For a wider discussion of modelling concepts see Moore et al. (2006a).

As stated in the Phase 1 report, the intention had been to also run the PRTF model for the pilot case study. However, it was not realised initially that the PRTF model also needs MORECS soil moisture deficit (SMD) data to run. As these data were not available the PRTF runs were not performed.

4.1 Probability Distributed Moisture (PDM)

The Probability Distributed Moisture model is a fairly general conceptual rainfall-runoff model which transforms rainfall and evaporation data to flow at the catchment outlet (Moore 1985, Moore 1999, CEH Wallingford 2005a, Moore 2007). Figure 4.1 illustrates the general form of the model. The PDM model 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.

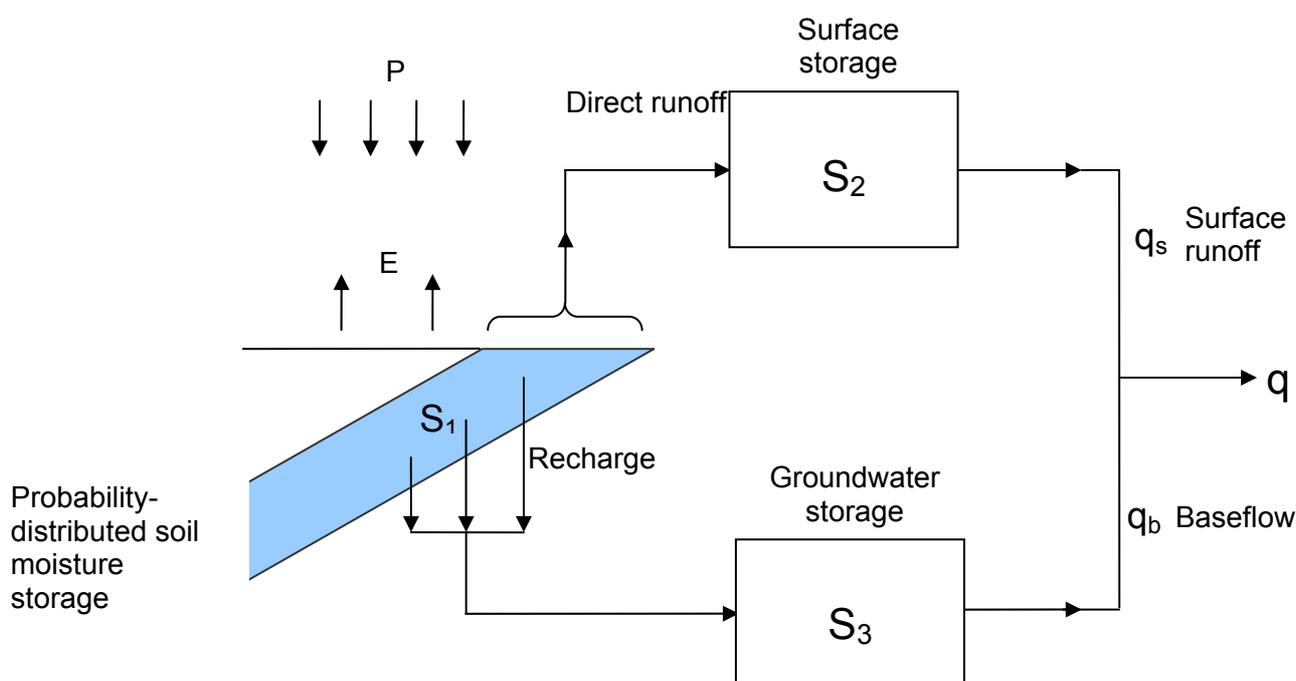


Figure 4.1 Overview of the PDM rainfall-runoff model.

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 considering 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.

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 here by a cascade of two linear reservoirs cast as an equivalent transfer function model (O'Connor 1982). The outflow from surface and subsurface storages, together with any fixed flow representing, say, compensation releases from reservoirs or constant abstractions, forms the model output. Table 4.1 summarises the parameters involved in the standard form of PDM.

Table 4.1 Parameters of the PDM model.

Parameter name	Unit	Description
fc	none	Rainfall factor
τ_d	h	Time delay
Probability-distributed store		
C_{min}	mm	Minimum store capacity
C_{max}	mm	Maximum store capacity
b	none	Exponent of Pareto distribution controlling spatial variability of store capacity
Evaporation function		
b_e	none	Exponent in actual evaporation function
Recharge function		
k_g	h mm ^{$b_g - 1$}	Groundwater recharge time constant
b_g	none	Exponent of recharge function
S_t	mm	Soil tension storage capacity
Surface routing		
k_1, k_2	h	Time constants of cascade of two linear reservoirs
Groundwater storage routing		
k_b	h mm ^{$m - 1$}	Baseflow time constant
m	none	Exponent of baseflow nonlinear storage
q_c	m ³ s ⁻¹	Constant flow representing returns/abstractions

For real-time application for flood forecasting, the PDM model is provided with forecast updating schemes. There are two basic types:

- error prediction – where the dependence in past model errors are used to predict future ones;
- state correction – where model errors are attributed to the model states going adrift and adjustments made to them to bring the model back on track.

The simple empirical state correction scheme is applied here in its 'super-proportional' adjustment form. Further details are provided in Moore (1999) and CEH Wallingford (2005a).

4.2 Grid-to-Grid model (G2G)

The Grid-to-Grid model is a grid-based runoff production and routing model (Moore et al. 2006a, Bell et al. 2007, Moore et al. 2007). It is a physical–conceptual distributed model configured on a grid for Area-wide flood forecasting and so can be used to forecast river flows at both gauged and ungauged sites. It is designed to be used with gridded rainfall estimates.

Its simple physical–conceptual formulation allows the model to be configured directly using spatial datasets on terrain and, where necessary, soil, geology and land cover properties. Here the simplest form of the G2G Model is used which requires only digital terrain data. Terrain slope is used to infer the capacity of the land to absorb water and to infer flow paths whose lengths control water translation through a catchment. This spatial dataset support leaves only a small number of regional model parameters to calibrate manually.

A schematic of the G2Grid model is shown in Figure 4.2. The model can be split into two distinct parts:

- the runoff production scheme, which acts in each grid square to generate fast ('surface') and slow ('subsurface') runoffs;
- the grid-to-grid flow routing scheme, which routes these runoffs across the domain.

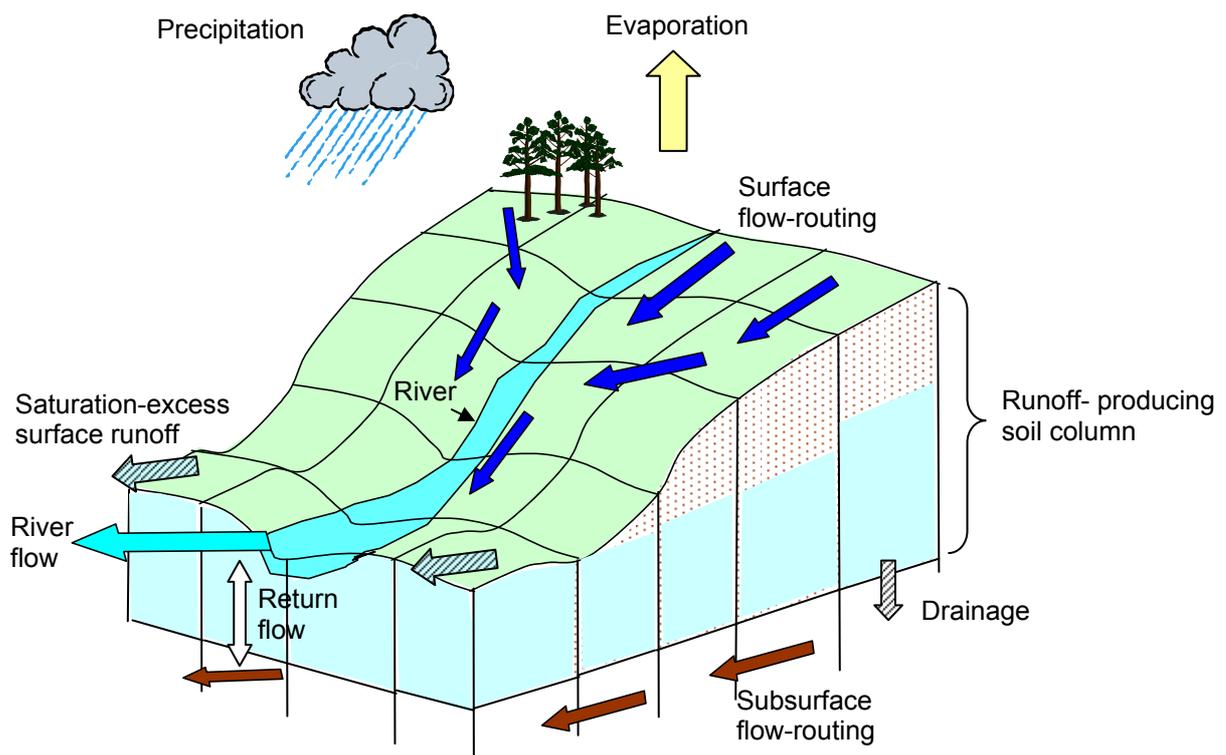


Figure 4.2 Grid-to-Grid distributed hydrological model.

4.2.1 Runoff production scheme

A topography-linked probability distributed runoff production scheme based on that employed by the Grid Model (Bell and Moore 1998a,b) is used. It generates surface and subsurface runoffs within each grid square which are then routed across the model domain using the routing scheme.

A simple empirical relation is assumed between topographic gradient, g , and moisture storage capacity, c , at a point

$$c = (1 - g/g_{\max})c_{\max} \quad (1)$$

where g_{\max} and c_{\max} are the maximum regional gradient and storage capacity values.

Terrain slope within a grid square is assumed to have the power distribution:

$$F(g) = \text{Prob}(\text{slope} \leq g) = \left(\frac{g}{g_{\max}} \right)^b \quad 0 \leq g \leq g_{\max} \quad (2)$$

where the exponent b is related to the mean gradient of the grid square,

$\bar{g} = \int_0^{g_{\max}} g f(g) dg$, by:

$$b = \frac{\bar{g}}{g_{\max} - \bar{g}}. \quad (3)$$

Based on these assumptions, the probability distribution function of storage capacity, c , within a grid square can be shown to have the Pareto form:

$$F(c) = 1 - \left(\frac{c_{\max} - c}{c_{\max} - c_{\min}} \right)^b \quad c_{\min} \leq c \leq c_{\max} \quad (4)$$

but with the minimum storage capacity $c_{\min} = 0$. The shape parameter, b , controls the form of variation between the minimum and maximum storage capacities.

Probability-distributed model theory presented by Moore (1985) can then be used to obtain the proportion of each grid square which is saturated and in turn, via analytical expressions (Moore 1999, 2007), calculate the volume of surface runoff generated and the grid-square water storage, $S(t)$, at time t .

Note that the maximum storage of the grid square, S_{\max} , is equal to the mean of the point storage capacities over this area, so that (for $c_{\min} = 0$):

$$S_{\max} = \bar{c} = \frac{c_{\max}}{b+1}. \quad (5)$$

The constraint $S_{\max} \geq \bar{c}_{\min}$ can be imposed to prevent any grid square having a zero maximum storage capacity; here \bar{c}_{\min} is the minimum mean store capacity of a grid square that is allowed and is treated as a regional parameter. For grid squares where this constraint applies, c_{\max} is recalculated using equation (4.3.5) with $S_{\max} = \bar{c}_{\min}$.

Losses from the grid square probability-distributed store via evaporation and drainage to groundwater vary as functions of its water storage, $S(t)$. Over the i^{th} time interval, $(t, t + \Delta t)$, water is lost as evaporation at a rate, E'_i , from the water in store as a function of the potential evaporation rate, E_i , and the soil moisture deficit, $S_{\max} - S(t)$, such that:

$$\frac{E'_i}{E_i} = 1 - \left\{ \frac{(S_{\max} - S(t))}{S_{\max}} \right\}^{b_e} \quad (6)$$

where the exponent b_e it is treated as a regional parameter (the same for all grid squares) and set to 2 here.

A power law function is used for the drainage, d_i , from the grid square probability-distributed store to groundwater storage:

$$d_i = k_g^{-1} (S(t) - S_t)^{b_g} \quad (7)$$

where k_g is a drainage time constant (here treated as a regional parameter), b_g is an exponent (here set to 3) and S_t is the threshold storage below which there is no drainage, water being held under soil tension. The tension threshold allows water to remain in soil storage and be made available to evaporation; this can be of particular importance for permeable catchments. It is treated as a regional parameter and if, for a particular grid square, $S_{\max} < S_t$, then drainage from that grid square can never occur.

The net rainfall rate, π_i , over the i^{th} time interval to the grid square is given by:

$$\pi_i = P_i - E'_i - d_i \quad (8)$$

where P_i is the grid square rainfall. Simple water accounting coupled to the probability-distributed analytical expressions for volume of runoff and water storage calculated for each grid square allow gridded surface and subsurface (drainage, d_i) runoffs to be generated for input to the G2G model routing scheme.

4.2.2 Grid-to-Grid flow routing scheme

The basis of the G2G flow routing scheme is a simple kinematic wave equation (Moore and Jones 1978), which relates channel flow, q , and lateral inflow per unit length of river, u . The equation is extended in the G2G model to include a return flow term, R , representing surface–subsurface water transfers per unit length of river.

In one-dimension (1D), the basic equation is of the form:

$$\frac{\partial q}{\partial t} + c \frac{\partial q}{\partial x} = c(u + R) \quad (9)$$

where c is the kinematic wave speed and x and t are distance along the reach and time respectively.

This equation is used to represent the movement of water from one grid cell to the next according to flow paths inferred from a digital terrain model. Equation (9) is applied separately to the surface and subsurface runoffs output from the runoff production scheme, thereby representing the simultaneous parallel water movement along fast (surface) and slow (subsurface) pathways. Different wave speeds over land and river (for surface and subsurface) pathways are accommodated. The return flow term allows transfer of water between subsurface and surface pathways, representing interactions on hill slopes and within river channels.

The finite difference representation of equation (9):

$$q_k^n = (1 - \theta)q_{k-1}^n + \theta(q_{k-1}^{n-1} + u_k^n + R_k^n) \quad (10)$$

is used where the dimensionless wave speed $\theta = c \Delta t / \Delta x$ ($0 < \theta < 1$), with Δx and Δt the time and space steps of the discretisation.

In this two-dimensional (2D) application, equation (10) provides a recursive formulation expressing flow out of the n^{th} grid cell at time k , q_k^n , as a linear-weighted combination of the flow out of the grid cell (at the previous time), inflow to the grid cell from adjacent grid cells (at the previous time) and the total lateral inflow (runoff production) plus return flow in the grid cell (at the same time).

The G2G routing scheme can be conceptualised as a network cascade of linear reservoirs (Moore et al. 2006a, Bell et al. 2007, Moore et al. 2007). The return flow to the surface routing pathway is given by a return flow fraction, r (between 0 and 1), of the water depth stored in the subsurface: this parameter can differ for land (denoted r_l) and river (denoted r_r) pathways. Note that to ensure numerical stability, the routing time step can be smaller than the model time step used in the runoff production scheme.

Methods have been developed for model initialisation and forecast updating of the G2G model for application in real-time flood forecasting. Initialising the states of a distributed model using river flow observations at gauged locations in the model domain is needed to avoid a long spin-up period for the model. Such initialisation will be needed when first installing the model within a forecast system and also in the event of a system or telemetry failure that precludes effective recovery from a previous set of stored model states. A simple initialisation scheme has been developed based on steady-state assumptions. Only an initial form of scheme is implemented at present. Test results have demonstrated the effective spatial transfer of information from a gauged site used for model initialisation to other locations within the model domain.

For forecast updating, a method of data assimilation is needed that incorporates flow measurements at gauged locations in the modelled region. The aim is to increase forecast accuracy by updating the states of the G2G distributed model in real-time using river flow observations sequentially at every time step up to the time the forecast is constructed and at every subsequent forecast time origin.

The sequential data assimilation scheme developed for the G2G model employs empirical state correction as a simple, pragmatic alternative to more complex procedures based on the Kalman filter. Only an initial form of scheme is implemented at present. Test results have shown that sequential data assimilation is more effective than a simple model re-initialisation at each time origin. Forecast hydrographs are generally improved as the forecast time origin approaches the flood peak. Overall forecast accuracy compared with model simulations is increased for lead times of interest at selected locations in the model domain assumed to be ungauged.

4.3 Representative Elementary Watershed (REW) model

4.3.1 Model capabilities

The Representative Elementary Watershed model is a complex hydrological simulation tool designed and developed for the simulation of a complete hydrological cycle system underlain by a regional aquifer which extends beyond the topographic boundaries.

The tool can be used for different water resources studies by looking at different components of the hydrological cycle and at processes that play a role at different time scales. It can for instance be used for event-based studies such as the response of a catchment to an extreme precipitation or the behaviour of the hydrological system under forcing conditions that are changing over longer time periods. Typical examples of possible applications and hydrological studies include:

- hydrological water balance;
- rainfall-runoff studies;
- groundwater recharge and development studies;
- impact of climate change on the hydrological cycle.

4.3.2 Spatial discretisation of the landscape into modelling units

In the REW model a catchment is partitioned into a series of discrete spatial units called representative elementary watersheds (REWs). REWs are defined by performing an analysis of the catchment topography and constitute a set of the interconnected elements that are organised around the tree-like structure of the stream channel network as shown in Figure 4.3.

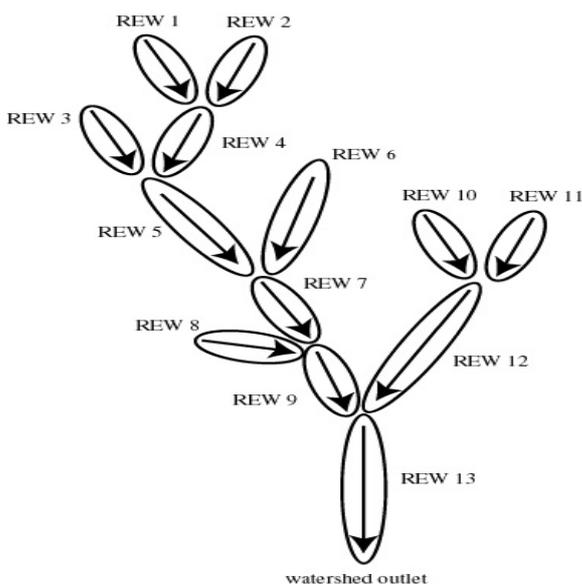


Figure 4.3 Binary structure of the channel network.

REWs constitute three-dimensional (3D) regions, with a vertical prismatic mantle surface defined by the REW boundaries. The REW boundaries coincide with topographic divides. They delineate portions of the land surface that capture precipitation. The contour of a REW mantle surface coincides with the perimeter of sub-basins. A schematic representation of a REW is depicted in Figure 4.4.

The REW is delimited by the atmosphere at the top and by an impermeable layer at the bottom. The impermeable layer can be either defined by a horizontal surface or can be given by interpolation of bedrock depth for a series of irregular points.

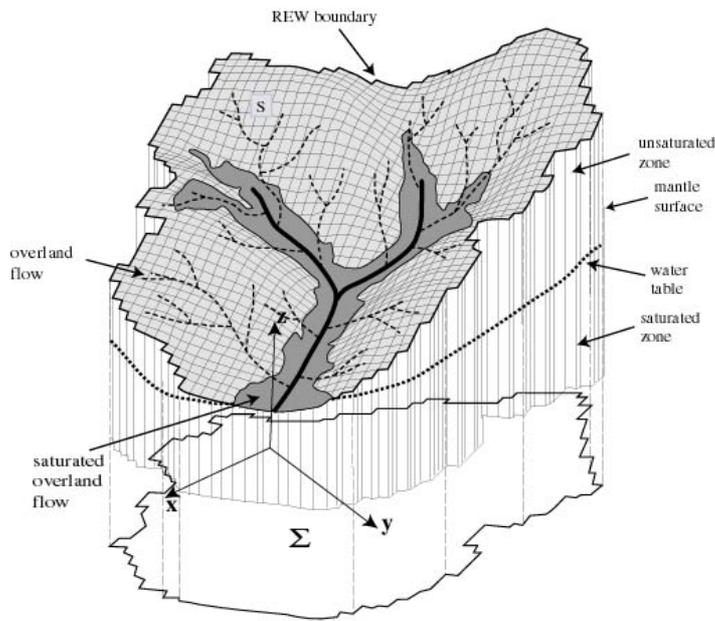


Figure 4.4 A representative elementary watershed (REW) as a 3D spatial region.

4.3.3 Sub-REW variability

To be able to account for hydrological variability within a REW with features at scales smaller than the REW due to land use pattern or soil properties, the unsaturated zone can be broken down further into smaller units called representative elementary columns (RECs). These RECs are defined on the basis of an overlapping series of geographic information system (GIS) maps such as land use and soil type. The procedure for breaking down the unsaturated zone allows different soil properties to be assigned to each unit. Figure 4.5 shows an example of a catchment broken down into RECs through combination of land use maps with REWs.

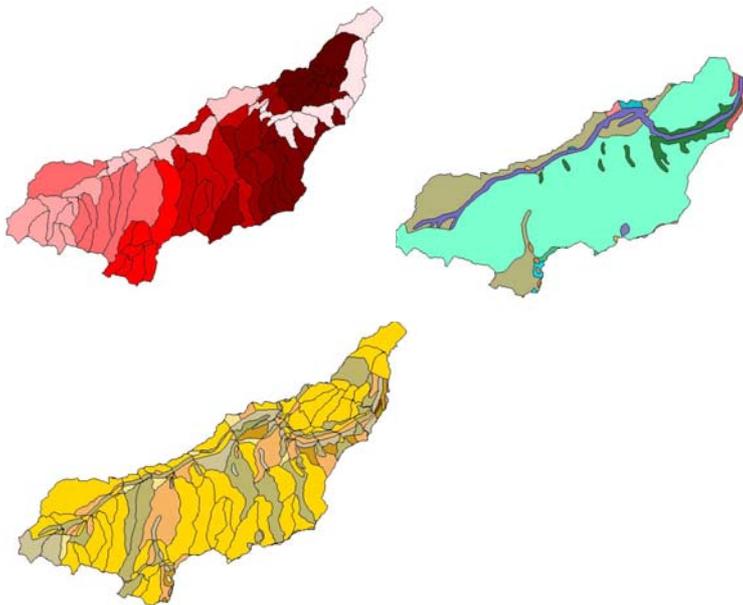


Figure 4.5 Overlap of a REW map with a soil map yields a smaller subdivision of the unsaturated zone within a REW into RECs.

4.3.4 Modelled processes

The volume occupied by a REW contains typical flow zones encountered in a catchment. The following zones can be modelled explicitly and for every REW:

- the unsaturated zone;
- the saturated zone;
- the subsurface storm-flow zone;
- the saturated overland flow area;
- the infiltration excess overland flow;
- the channel reach;
- a snow zone.

The flow within the various domains evolves over different temporal scales and encompasses phenomena such as unsaturated and saturated porous media flow (subsurface zones) as well as overland and channel flow (land surface zones). The modelling of the various flow processes is described separately in the following paragraphs.

Unsaturated zone (U-zone)

The unsaturated zone is modelled by means of a Richards' equation solver (Ross 2003). The chosen solver for the partial differential equation (PDE) governing flow in unsaturated soil has the property of linearising the mass flux between cells and allows a very fast solution of the equation, avoiding the need to search for iterative solutions. With respect to full non-linear solvers, the accuracy of the numerical solution is somewhat lower. But given the high uncertainty in the choice of the soil parameters, the errors of approximation made in the choice of the numerical method is considered of second order and thus negligible.

Saturated zone (S-zone)

The saturated zone is modelled as a 2D aquifer. The groundwater zone is recharged through recharge flux from the unsaturated zone. The groundwater is then distributed laterally via horizontal REW mantle fluxes based on piezometric head differences between REWs. The piezometric head is the average water table level calculated for a REW via the mass balance equation. The mass balance equation is an ordinary differential equation (ODE) solved analytically, given the recharge flux from:

- the unsaturated zone, e^{us} ;
- the lateral groundwater distribution fluxes between the REW and neighbouring REWs, e^m ;
- the seepage flux, e^{so} ;
- the exchange flux of groundwater with the river channel across the bed area, e^{sr} .

The seepage flux, e^{so} , feeds the overland flow zone shown in Figure 4.6.

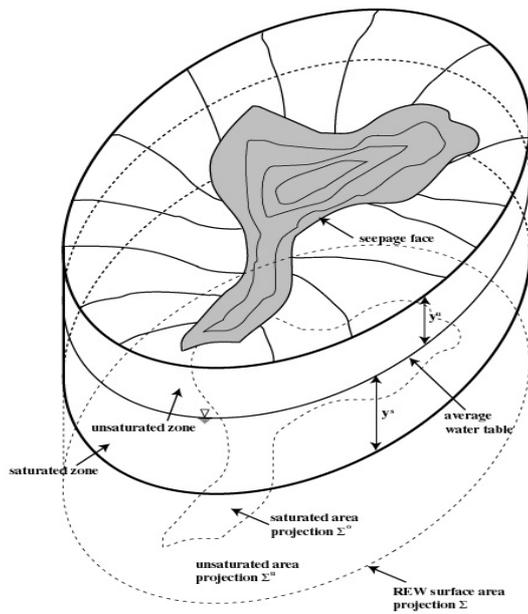


Figure 4.6 Conceptualisation of a REW in the model.

The length scale, Λ , over which piezometric head differences are dissipated between a REW and its neighbouring REWs is an unknown quantity which is recalculated at chosen time steps based on first principles. For this, the Hardy-Cross (Cross 1936) network balancing method is used (see Figure 4.7).

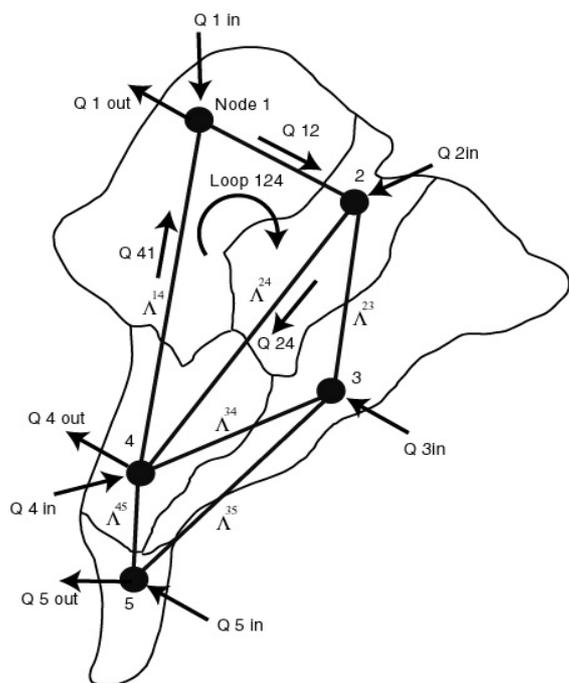


Figure 4.7 Implementation of the Hardy-Cross method for the groundwater system.

Given a piezometric head distribution calculated from the mass balance for the saturated zone of each REW at a given point in time and given known groundwater losses across the catchment boundaries, the dissipation length scales are calculated by successive approximation.

The procedure is parsimonious and based on a non-linear system of equations which preserve (i) mass at each network node and (ii) the head losses along a closed triangular loop (as shown in Figure 4.7). The horizontal aquifer flow field is subsequently calculated by resolving the momentum balance equation for the REW elements. An example of a vector of flow velocities for the Geer aquifer (Belgium) is shown in Figure 4.8.

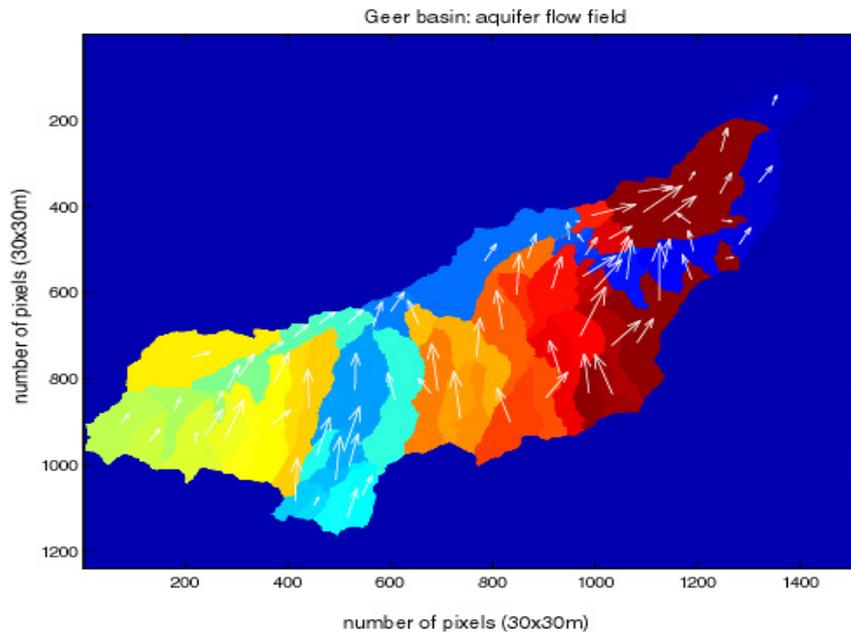


Figure 4.8 Calculation of the groundwater flow field for the Geer basin (Belgium).

The REW-average groundwater levels are interpolated at selected time steps through bi-cubic spline functions (Inoue 1986), providing a smooth groundwater surface between REW-average groundwater points. The fitting of the smooth surface is based on the finite element method (FEM), which calculates the surface by minimising the elastic tension energy in the surface. The same procedure can be applied for the definition of the impermeable lower boundary of the catchment if sparse measurement points of the bedrock depth are available. Figure 4.9 shows an example of a fitted surface.

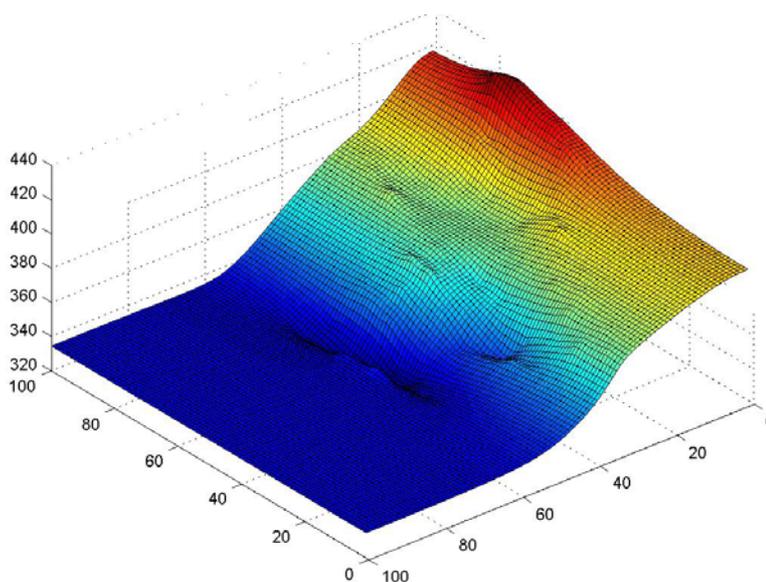


Figure 4.9 Water table surface interpolated with the bi-cubic spline method.

Saturation excess flow/Dunne-type flow (O-zone)

The saturation excess flow – also called Dunne-type flow in the literature – is caused by direct precipitation on top of saturated areas. In the REW model, the growth of the saturated areas is linked directly to the rise and fall of the REW-average groundwater level y^s (in Figure 4.11). By default, it is assumed that the relation between the groundwater levels and the growth of the saturated areas is linear. The saturated areas are alimeted through exfiltration from the saturated zone along conceptual seepage faces.

The model calculates the saturated REW area fraction, ω , as a dynamic variable. The runoff on the saturated areas is calculated by solution of the mass and momentum balance equations for overland flow (kinematic wave) analytically. The overland flow zone discharges laterally into the river channel, yielding a lateral channel inflow flux, e^{of} . The saturation excess zone is fed directly by precipitation and is exposed to potential evaporation during inter-storm periods. If infiltration excess flow is generated on the unsaturated part of the REW surface, the infiltration excess flow is discharged into the saturated overland flow zone through a flux, e^{oc} .

The hydrological processes represented in the REW model are illustrated in Figure 4.10.

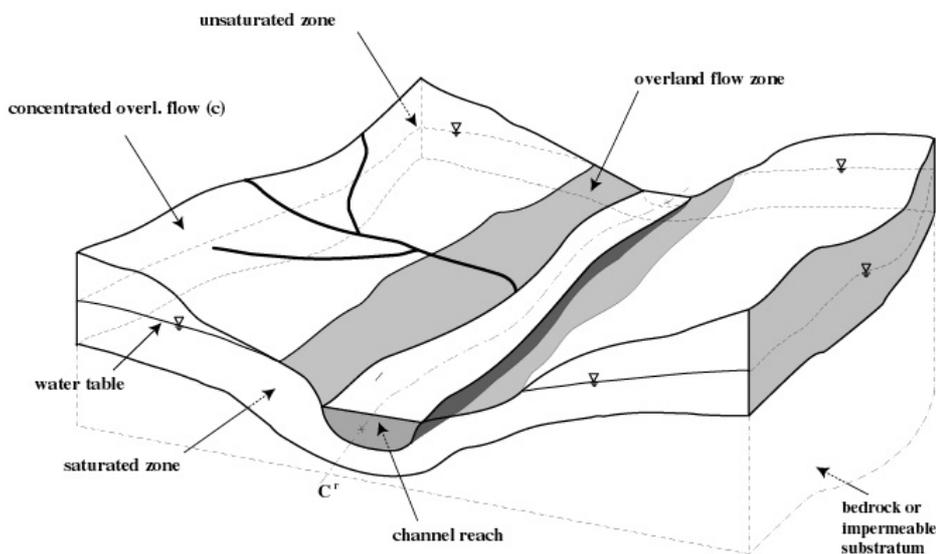


Figure 4.10 View of hydrological processes represented in the REW model.

Subsurface storm-flow (P-zone)

Subsurface storm-flow is generated in a shallow subsurface layer with high conductivity. For some catchments, the use of this zone is essential to capture certain rapid runoff phenomena. This zone can also be used to represent a perched aquifer system (thus the designation P-zone), which constitutes a shallow, suspended reservoir of groundwater. The subsurface storm-flow (or the perched system) is fed by direct infiltration of precipitation and discharges towards the channel as the flux term, e^{pr} . In case of saturation of the subsurface layer, the excess flow is discharged directly into the saturated overland flow zone as flux, e^{po} . The governing equations for the subsurface storm-flow are the mass and momentum balance equations for subsurface flow, which are combined into a kinematic wave equation and solved analytically.

Infiltration excess flow/Horton-type flow (C-zone)

The infiltration excess flow – also called Horton-type flow in the literature – is caused by precipitation that exceeds the infiltration capacity of the soil. As a result water builds up on the surface and runs off. In the REW model, the infiltration excess flow is modelled through analytical solution of the mass and momentum balance ordinary differential equations (ODEs). The runoff flux, e^{co} , is discharged directly into the saturated overland flow zone. The infiltration excess flow is fed by the precipitation rate during storm periods and by potential evaporation during inter-storm periods.

Channel flow (R-zone)

The channel flow zone is recharged by fluxes from upstream links, e^{rin} , the outflow to the downstream reach, e^{rout} , and lateral inflow fluxes, e^{or} , e^{sr} and e^{pr} , from the overland flow zone (O-zone), the aquifer (S-zone) and the subsurface storm-flow zone (or the perched zone, P-zone).

The lateral inflows due to overland flow and the shallow subsurface storm-flow zone are controlled by the governing equations for these respective zones. The exchange with the groundwater is dictated by the average head differences between the REW-average groundwater level and the river. The water between the two zones is exchanged through a river bed transition zone, for which a hydraulic conductivity and a thickness can be specified. For situations in which the average water level in the channel reach is higher than the water level in the surrounding aquifer, the flux, e^{sr} , causes the groundwater to be fed from the channel. If on the other hand the average water level in the aquifer increases with respect to the channel, the groundwater feeds the channel. This principle is explained schematically in Figure 4.11, which features the REW-average water level, the actual water level, the water table interpolated via the Inoue (1986) algorithm and the average water level in the channel.

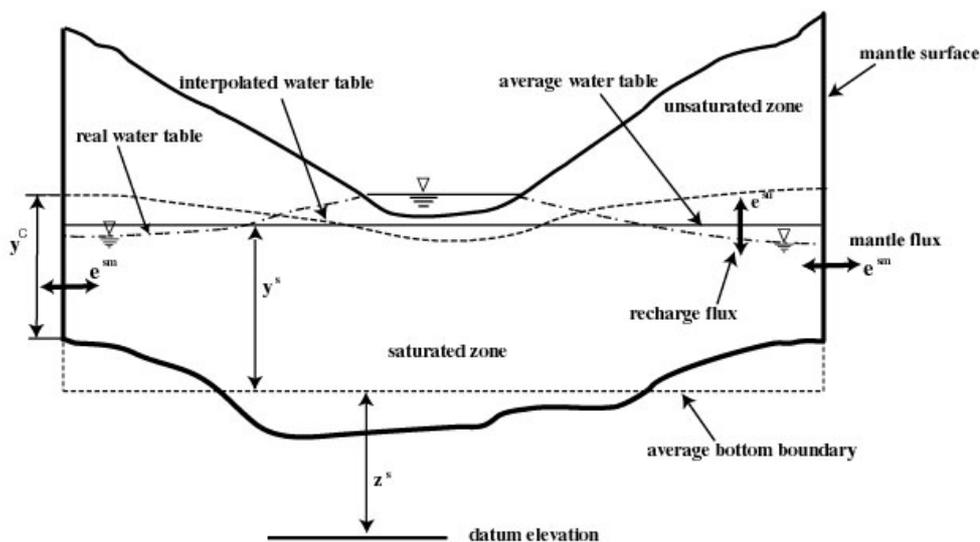


Figure 4.11 Cross section of the REW showing the river channel.

Summary of exchange fluxes in the REW model

The most relevant model internal and internal fluxes are shown in Table 4.2. This specifies which fluxes are within zones in a REW and which ones between a REW and either neighbouring REWs or the outside environment (i.e. across catchment boundaries).

Table 4.2 Hydrological fluxes within the REW model.

Flux description	Symbol	REW-internal flux	Inter-REW flux	External boundary flux
River-saturated zone	e^{sr}	yes	no	no
Water table flux (unsaturated zone-saturated zone)	e^{us}	yes	no	no
Infiltration	e^{cu}	yes	no	no
Inflow from infiltration excess flow zone to saturated overland flow zone	e^{co}	yes	no	no
Lateral channel inflow	e^{or}	yes	no	no
Inter-REW groundwater flow	e^m	no	yes	yes
Lateral channel inflow from subsurface storm-flow zone	e^{pr}	yes	no	no
Exfiltration from subsurface storm flow zone to saturated overland flow zone	e^{po}	yes	no	no
Exfiltration (seepage flow)	e^{so}	yes	no	no
Channel in and outflow	$e^{r out}$	no	yes	yes
	$e^{r out}$			

5 Model configuration and calibration: application to the case study catchments

This section discusses model configuration and calibration and presents the simulation and forecast results (using HyradK raingauge-adjusted radar data) of the three hydrological models – PDM, G2G and REW. Forecast performance of the hydrological models over the Boscastle event when using high resolution NWP is discussed in Section 6.

A summary of the case study catchments is given in Section 5.1 followed by discussion of the model configuration. The raingauge-adjusted radar rainfall data used as model input and produced by HyradK are detailed in Section 5.2.3. Sections 5.3 to 5.5 discuss the individual model calibration and assessment of their performance.

5.1 Catchment information

The areas worst affected by the case study storm were Boscastle and Crackington Haven. These are ungauged catchments and flow measurements during the flood are not available for these locations. Other stations in the region did register a significant response and the most noteworthy occurred for the Ottery at Werrington Park.

The network of gauging stations near Boscastle with flow measurements is listed in Table 5.1 and shown in Figure 5.1. Figure 5.1 also provides the 50 m resolution elevation data from the Integrated Hydrological Digital Terrain Model (IHDTM) (Morris and Flavin 1990); it shows clearly the location of Bodmin Moor to the south of Boscastle and the western edge of Dartmoor to the east of the map.

Maps of the solid geology and the dominant HOST (Hydrology Of Soil Types) class number (Boorman et al. 1995) at a 1 km resolution are shown in Figure 5.2.

Following analysis of the available data and consultation with the Environment Agency, three of these gauged catchments were selected for the main focus of the hydrological case study:

- Ottery at Werrington Park;
- Tamar at Gunnislake;
- Camel at Denby.

The remaining gauging stations will be used to assess the ungauged performance of the distributed models.

The selected catchments are discussed in more detail below based on the information given above along with station summaries and spatial catchment information from the UK National River Flow Archive (NRFA) (<http://www.ceh.ac.uk/data/nrfa/>).

Table 5.1 Gauging station details (main case study catchments are in bold).

Station	National Grid Reference	NRFA station number	Area (km ²)
Withey Brook at Bastreet	224400 076400	47013	15.74
Inny at Beals Mill	235900 077100	47020	104.99
Strat at Bush	223447 107996	N/A	10.75
St Neot at Craigs Hill	218400 066200	48009	22.89
Tamar at Crowford Bridge	229000 099100	47010	77.68
De Lank at De Lank	213300 076500	49003	21.74
Camel at Denby	201700 068200	49001	209.93
Tamar at Gunnislake	242600 072500	47001	920.11
Neet at Helebridge	221380 103830	N/A	76.40
Lyd at Lifton Park	238900 084200	47006	220.39
Lynher at Pillaton Mill	236900 062600	47004	135.27
Tamar at Polson Bridge	235300 084900	47019	471.74
Fowey at Restormel	209800 062400	48011	167.2
Camel at Slaughterbridge	210940 085720	N/A	9.04
Thrushel at Tinhay	239800 085600	47008	112.7
Fowey at Trekeivesteps	222700 069800	48001	36.80
Warleggan at Trengoffe	215900 067400	48004	25.26
Ottery at Werrington Park	233700 086600	47005	121.66
Neet at Woolstone Mill	222730 101810	N/A	37.15

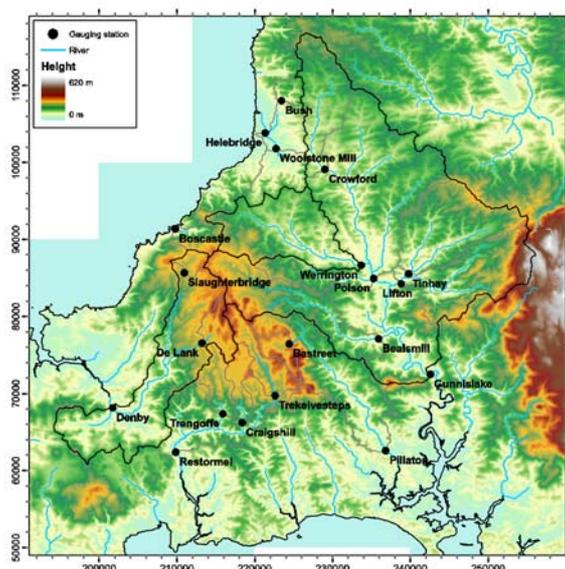


Figure 5.1 Map of relief over the Boscastle region.

Notes: Gauging stations and their catchment boundaries are also shown. Northings and Eastings are for NGR co-ordinates in metres.

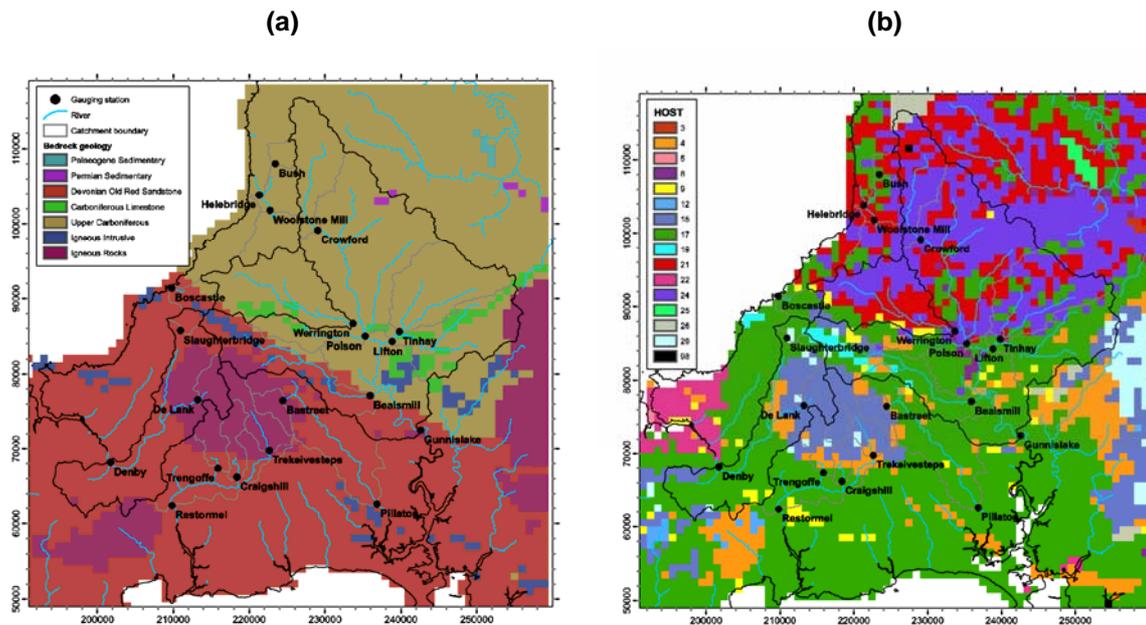


Figure 5.2 Maps over the Boscastle region of (a) solid geology and (b) HOST.

Notes: Northings and eastings are for NGR co-ordinates in metres.

5.1.1 Ottery at Werrington Park

This catchment is a responsive natural catchment with a small drainage area (121 km²) and moderate relief. It is an ideal candidate for PDM rainfall-runoff modelling. The geology of the catchment is mainly carboniferous culm measures which are classified to have very low permeability. There is little in the way of superficial deposits except for a swath of river terrace deposits and alluvium centred along the main river channel. The HOST classification is split between two main classes: class 24 to the north and 21 to the south of the catchment. These correspond to mineral soils overlying a slowly permeable substrate with the presence of a gleyed or impermeable layer within the first 100cm and no significant groundwater.

This gauging station is of particular interest as it recorded a significant flood response during the Boscastle storm and has a reasonable rating curve, although there is some out-of-bank flow and bypassing during large events. Used in combination with the downstream station at Gunnislake on the River Tamar, this pair of nested catchments will be very useful for calibrating the distributed hydrological models.

5.1.2 Tamar at Gunnislake

This is a fairly responsive rural catchment of moderate relief and the largest gauged catchment (920 km²) in the case study which, in combination with the interior gauge at Werrington, makes it important for calibrating the distributed models. Due to the localised nature of the Boscastle storm, the station only registered a very small flow response. It is not a natural choice for the PDM model as a network of models would normally be used, but the PDM model should still perform reasonably well.

The geology (Figure 5.2a) consists mainly of Carboniferous formations with some Devonian formations to the south-eastern edge. There are significant alluvial flats in the middle reaches. Apart from the small areas of Bodmin Moor and Dartmoor that cover the western and eastern tips of the catchment, the hydrogeology is classified as having very low

permeability. The HOST classification of the catchment is dominated by classes 21 and 24 to the north and 17 to the south. Class 17 corresponds to mineral soils overlying an impermeable (hard) substrate with no impermeable or gleyed layer within the first 100cm and no significant groundwater.

5.1.3 Camel at Denby

This is a small-to-medium sized catchment (209 km²) that should be suitable for both lumped and distributed models. The northern part of the catchment was affected by the Boscastle storm and the station registered a moderate response during the Boscastle event. It is believed to have a good rating curve. There is a small reservoir (Crowdy) in the north-east part of the catchment that affects runoff.

The geology of the catchment consists of igneous rocks of mixed permeability underlying Bodmin Moor with Devonian formations of very low permeability elsewhere. There are superficial deposits of peat over Bodmin Moor and small areas of alluvial deposits elsewhere. The HOST classification shows undrained peat soils with an unconsolidated substrate and the presence of groundwater within 2m (class 12) over a majority of Bodmin Moor. The remainder of the catchment is dominated by HOST class 17.

5.2 Configuration

The three modelling concepts discussed in Section 4 (PDM, G2G and REW) have been configured and calibrated for the three gauged catchments (Camel, Ottery and Tamar). The distributed models have also been configured (but not calibrated) for the Valency at Boscastle, as shown in Figure 5.3.

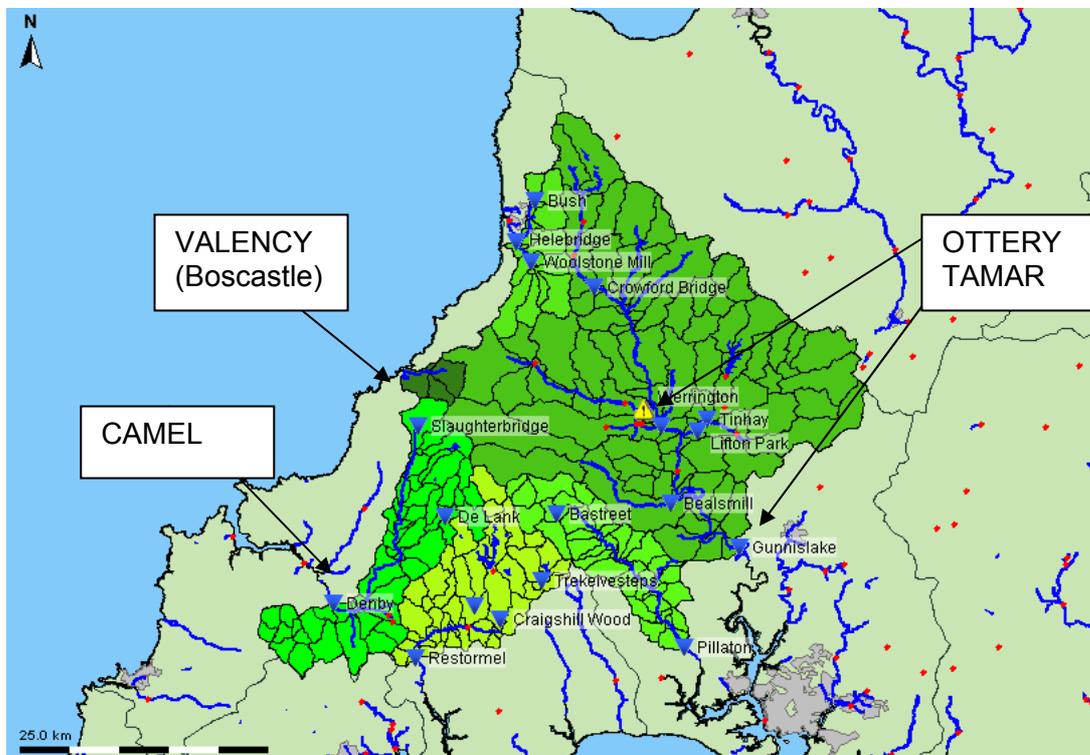


Figure 5.3 Modelling area.

The remainder of this section outlines the datasets used for model configuration and the calibration strategy. A detailed account of model calibration is provided for each of the three modelling concepts in Sections 5.3–5.5.

5.2.1 Digital Terrain Model (DTM)

In this case study, the simplest variants of the G2G and REW models were configured where the only spatial dataset employed is the 50m IHDTM. This dataset provides elevation and hydrologically consistent flow directions that are used within the runoff production and flow routing elements of the models. It also allows catchment boundaries to be delineated, which are used directly in the configuration of REW and PDM.

5.2.2 Potential evaporation

Monthly MORECS potential evaporation (PE) data (Hough et al. 1997) were used as input to all models. An evaporation profile derived from these data and used by the REW model is shown in Figure 5.4.

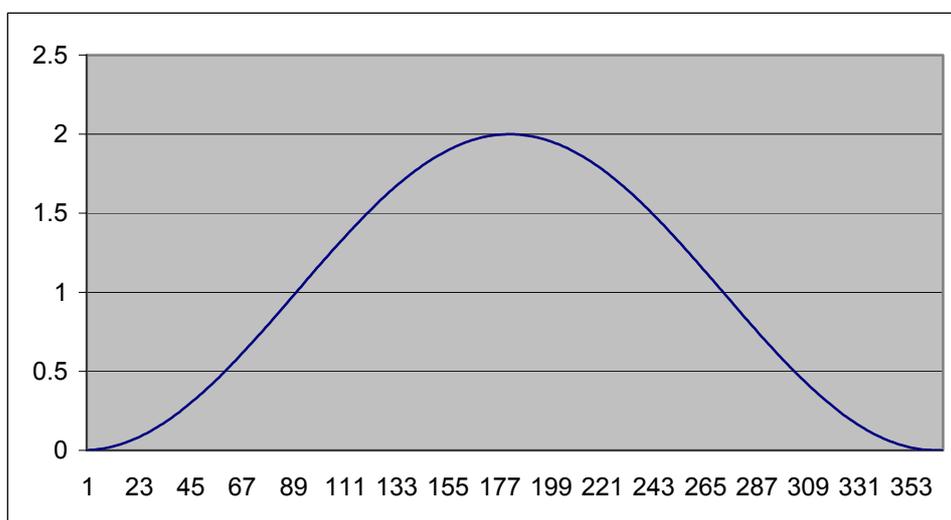


Figure 5.4 Daily evaporation profile used for REW modelling.

5.2.3 HyradK raingauge-adjusted radar data

The HyradK module adapter was used to generate raingauge-adjusted radar rainfall time series. Nimrod composite radar data and Environment Agency raingauge data were combined and the resulting HyradK rainfall grids used as input to all models. These gridded rainfall data are used directly by the G2G model and in the form of catchment averages for the REW and PDM models.

Raingauge data were provided for the area surrounding the case study catchments as well as within the catchments. This is important as the raingauges located outside catchments can still have a positive impact on the HyradK rainfall estimates inside a catchment.

Only a subset of the raingauges provided was included. The main criteria for selection were to maintain a consistent raingauge network over the 2002–2007 study period. Several new raingauges have been installed since 2005, but these were only included if they replaced an older raingauge that has ceased to be in operation. The raingauges selected are listed in Table 5.2 and their locations mapped in Figure 5.5.

The time series of 15-minute accumulations were then quality controlled using the following methods.

- **70mm filter** – 15 minute accumulations in excess of 70 mm were removed from the data and set to be missing. These were cross-referenced to radar data and Met Office weather summaries to confirm that they were indeed erroneous. Table B1 in Appendix B lists the values changed.
- **20–70mm filter** – 15 minute accumulations above 20mm were investigated further including use of radar data and Met Office weather summaries. Table B2 lists the values above 20mm and also indicates whether these were assumed valid or treated as missing.
- **Cumulative hyetographs** for groups of raingauges located close together were plotted and any anomalies (e.g. blocked raingauges, zero recorded rainfall) were investigated. Periods treated as suspect were replaced by missing values (-999.0); the periods are listed in Table B3.

5.2.4 Model calibration strategy

Case study hydrometric data were provided for the period 2002–2007. After studying the available data and realising that the Boscastle event (16 August 2004) needed to be used for model verification, a split sample methodology was used where distinct calibration and verification periods were identified (Table 5.3).

All models were calibrated using the calibration period data only. The robustness of the model calibrations was then tested independently over the verification period. The selection used encompassed a range of summer and winter events in both the calibration and verification periods.

Table 5.3 Calibration and verification periods.

Period	Start date	End date
Calibration period	00:00 09/01/2006	23:45 31/07/2007
Verification period	00:00 01/01/2004	23:45 23/12/2005

5.2.5 Generating hydrological model forecasts

All hydrological models employed the HyradK raingauge-adjusted radar data as a ‘perfect rainfall forecast’. Different approaches for generating the hydrological forecasts were used by the different models. For the PDM and G2G models, empirical state correction schemes were chosen which utilised gauged observation at the point of interest up to the start of the forecast. The REW model uses an autoregressive moving average (ARMA) error prediction model. More details about the approaches are given in the sections that describe the results obtained with the different models.

5.3 REW model application

5.3.1 Terrain analysis

Separate REW models are set up for the five catchments in the model domain covering parts of Cornwall and Devon. The selected catchments included the Tamar, Camel, Fowey, Lynher and Valency. Only the Tamar/Ottery and the Camel were calibrated and are discussed here.

The first step in the model set-up is the analysis of the 50 × 50 m digital terrain maps (DTMs) for the extraction of the stream channel network and the determination of the REWs. For the terrain analysis, the open source software TARDEM from the University of Utah was used having been extended with additional capabilities for the extraction of REWs.

The DTM analysis led to two separate catchment configurations:

- Tamar (with the Ottery as an internal catchment);
- Camel.

The stream channel network is extracted with the stream threshold area criterion. The threshold area is a minimal accumulated upstream area expressed as the number of pixels. Pixels with an accumulated area higher than the threshold area are defined as stream channel pixels. The network is extracted by assuming a cut-off Horton–Strahler threshold of order 1, meaning that first and larger order channels are all part of the network. If larger REWs are desired, the Horton–Strahler threshold order can be set equal to 2 or higher.

Once TARDEM has extracted the network and determined the sub-basin areas, the module REWANALYSIS is used to determine the 3D REW geometries and REW interconnections.

Figures 5.6 and 5.8 show the spatial discretisation of the Camel and the Tamar catchments into 51 and 81 REWs respectively.

5.3.2 REW model set-up

Preliminary considerations

Before describing the parameterisation and the results of the model for the two study catchments, it is important to emphasise the assumptions under which the catchments were modelled.

- The REW model does not consider the presence of a vegetation cover and the catchments are thus modelled as if they were bare soil. The net precipitation is given by the sum of precipitation minus potential evaporation. As a result, vegetation-related effects such as interception or a more sophisticated SVAT (surface–vegetation–atmosphere transfer) scheme with root extraction are neglected. Therefore, the non-linear effects of soil–water depletion during summer months are not fully accounted for.
- The parameters describing soil texture and structure are applied homogeneously across the catchment, the saturated zone (S-zone), the unsaturated zone (U-zone) and the subsurface storm flow zone (P-zone). In principle it would be possible to assign different properties to the various zones

and REWs but, in the absence of detailed soil information, uniform values were used.

- The raingauge-adjusted radar rainfall time series was been used to obtain areal averages over the REWs.
- The potential evaporation was estimated from the monthly average MORECS data. These data are supplied on a coarse national grid covering the entire UK. The MORECS potential evaporation estimates were taken relative to the grid cell covering the study area. The same evaporation time series were assigned to all REWs in the study catchments.
- It was assumed that there is no lateral groundwater exchange across the external catchment boundaries. Only the REW 1 (the REW in correspondence with the catchment outlet) was assumed to have a permeable outer boundary and thus be admitting a minimal groundwater flux across the external catchment boundary. In this context, at least one network node must be allowed to exchange water to preserve continuity of mass within the network for the groundwater distribution algorithm to converge (the algorithm is based on the Hardy-Cross discharge rebalancing over closed network loops).
- The calibration of the model parameters was performed over a 15-minute time series of precipitation and evaporation over a calibration period ranging from 9 January 2006 to 31 July 2007. The precipitation data were at 15-minutes intervals while the potential evaporation data were disaggregated from average monthly data to 15-minute data. The calibration was performed manually, whereby the hydrodynamic parameters (soil texture and structure, Manning coefficients) were kept constant while five parameters determining geometric characteristics were kept variable.

5.3.3 Camel

The Camel catchment has a surface area of 209 km². The catchment was broken down by the terrain analysis into a total of 51 modelling units or REWs (Figure 5.6) by assuming a surface threshold area of 100 pixels for channel heads.

To determine the REWs, a subsurface zone delimited by an infinite depth (3000m, parameter 4) bedrock layer was set. Uniform soil texture and structure parameters were assigned for all REWs as indicated in Table 5.4.

Calibration

The exchange between the river channel and aquifer (parameter 1) was chosen as very low (10^{-6}), effectively setting the river-groundwater exchange to zero. In this context the depth of the river bed transition zone (parameter 2) is set to 1.5 m, but remains irrelevant as a model parameter.

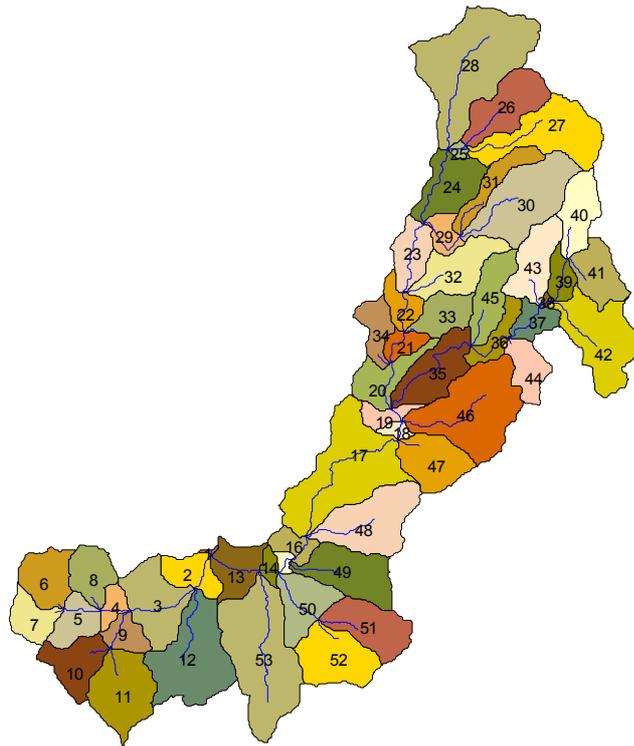


Figure 5.6 Camel watershed broken down into 51 REWs.

Table 5.4 REW model parameters for Camel catchment.

No	Parameter	Unit	Value	Calibrated
1	Hydraulic conductivity for channel bed	m/s	0.000001	N
2	River bed transition zone thickness	m	1.5	N
3	Exponent in power relationship (p=1 linear)		0.30	Y
4	Bedrock depth	m	3000	N
5	Soil porosity	–	0.5	N
6	Saturated hydraulic conductivity S-zone	m/s	0.00005	N
7	Saturated hydraulic conductivity U-zone	m/s	0.00005	N
8	Brooks–Corey soil parameter lambda	–	1.00	N
9	Brooks–Corey pressure scaling parameter	m	0.25	N
10	Water content at saturation	–	0.5	N
11	Saturated hydraulic conductivity P-zone	m/s	0.005	Y
12	Exponent on transmissivity law ($2 \leq g \leq 4$)		3.8	Y
13	Depth of saturated subsurface flow layer	m	0.5	Y
14	Exponent for surface precipitation partitioning		0.3	Y

An important calibration parameter is the exponent of the power law relationship which governs the expansion of the saturated areas as a function of water table position (parameter 3). This relationship was chosen with an exponent that is less than linear (0.3), causing larger increases of the saturated area fraction for water table levels close to average channel bed elevation of the REW, and with decreasing saturated area expansion for groundwater table levels above average channel bed elevation. This parameter was also assigned as constant for all REWs and should in principle be set as variable between REWs. The most common range for this parameter is 0–3.

The soil porosity (parameter 5) was set uniformly to 0.5 for the entire catchment and is not considered a calibration parameter. The hydraulic conductivity of the saturated zone (S-zone, parameter 6) and the unsaturated zone (U-zone, parameter 7) were both set to 5×10^{-5} m/s. This parameter was set as a constant value during the model setup. The Brooks–Corey parameters λ (parameter 8) and m (parameter 9) were set equal to 1 and 0.25 respectively and uniformly for all unsaturated zones of all REWs.

The residual water content in the saturated zone was assumed to be equal to 0 and the water content at saturation (parameter 10) to be equal to the soil porosity. The P-zone constitutes an important store for the present system. The P-zone is a subsurface storm-flow layer which is described as a subsurface kinematic wave equation and a transmissivity law controlled by an exponent (parameter 12). The transmissivity law exponent is a calibration parameter.

The net precipitation falling onto the soil surface is split into two parts:

- one part going directly into the unsaturated soil and therefore into the Richards equation column;
- one part going into the P-zone.

The splitting is governed by a power law relationship, in which the mean saturation of the top 10 cells of the Richards equation layers are raised to a power β (parameter 14). In this fashion, the wetter the top soil becomes, the more water is entering the P-zone.

The lateral flow in the P-zone occurs in a layer with a constant depth of 0.5 m (parameter 13) and joins the river channel (R-zone). Excess water, which cannot be transferred in this layer flows off as surface runoff, causing peaks in overland flow. Both parameters 13 and 14 are relevant calibration parameters controlling this process. This mechanism of runoff partitioning determines to a large part the reactions of the system and requires special attention in the calibration process.

Results

Figure 5.7 shows the results of the calibration for the gauging station on the Camel at Denby. The falling limb of the hydrograph can be seen to be too steep for low discharges. This shortcoming is essentially attributable to limitations in the representation of surface runoff as a sheet flow in terms of an analytically solved kinematic wave equation for low discharges. The issue does not constitute a problem in the case of the Tamar catchment (see Section 5.3.4) with higher surface runoff volumes.

Improved results can be achieved by either using Manning coefficients that are higher and thus lie outside the range of typical values reported in the literature, or by representing overland flow with equations other than kinematic surface wave equations (e.g. simple storage–discharge power law relationships). Section 5.3.5 discusses the calibration results in more detail.

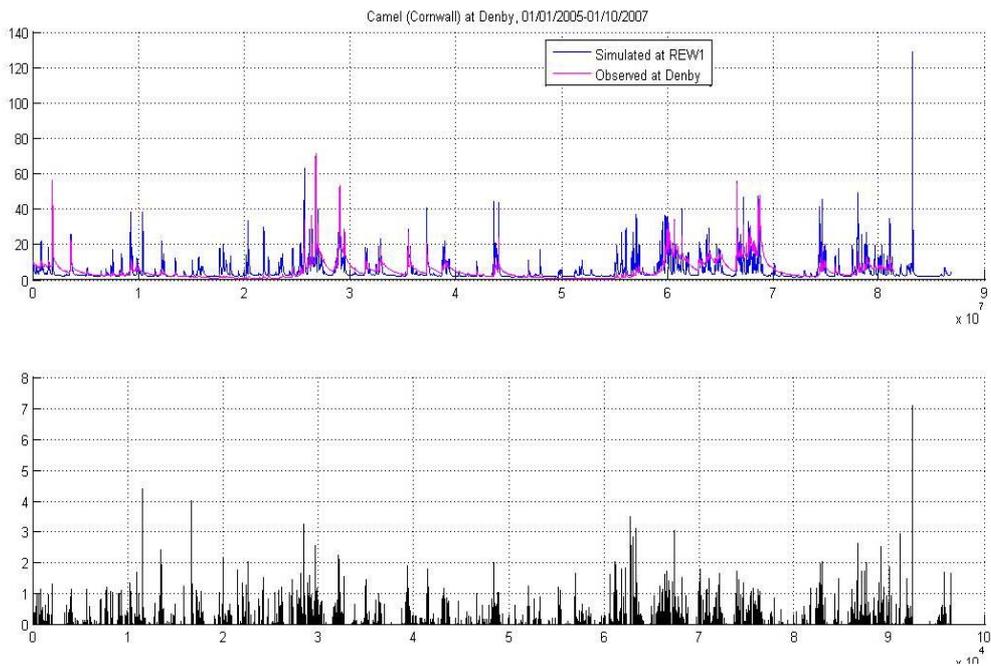


Figure 5.7 Modelled and observed discharges at Denby, 1 January 2005 to 1 October 2007 (precipitation at REW 1).

5.3.4 Tamar

The Tamar has a surface area of 920 km². The catchment was separated into a total of 81 REWs (Figure 5.8) by assuming a surface threshold area of 100 pixels for channel heads.

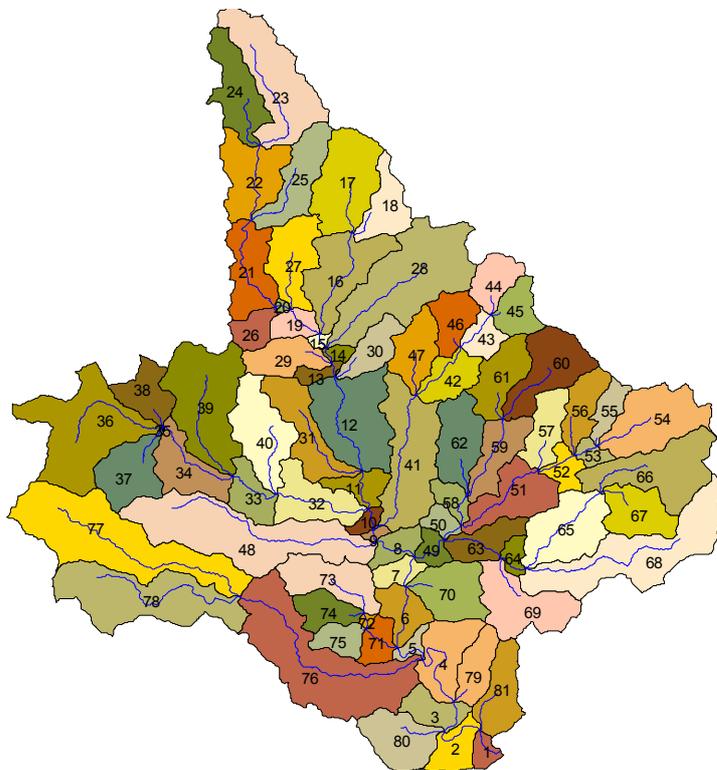


Figure 5.8 Tamar watershed broken down into 81 REWs.

Calibration

Regional uniformity between catchments was assumed and the same parameter sets were used for the calibration of the Tamar catchment as for the Camel. The respective parameter values are summarised in Table 5.5.

Table 5.5 REW model parameters for Tamar catchment.

No.	Parameter	Unit	Value	Calibrated
1	Hydraulic conductivity for channel bed	m/s	0.000001	N
2	River bed transition zone thickness	m	1.5	N
3	Exponent in power relationship ($p=1$ linear):		0.30	Y
4	Bedrock depth	m	3000	N
5	Soil porosity	–	0.5	N
6	Saturated hydraulic conductivity S-zone	m/s	0.00005	N
7	Saturated hydraulic conductivity U-zone	m/s	0.00005	N
8	Brooks–Corey soil parameter lambda	–	1.00	N
9	Brooks–Corey pressure scaling parameter	m	0.25	N
10	Water content at saturation	–	0.5	N
11	Saturated hydraulic conductivity P-zone	m/s	0.005	Y
12	Exponent on transmissivity law ($2 \leq g \leq 4$):		3.8	Y
13	Depth of saturated subsurface flow layer	m	0.5	Y
14	Exponent for surface precipitation partitioning		0.3	Y

Results

Figures 5.9 and 5.10 show the calibration results at the two gauging stations on the Tamar at Gunnislake (outlet) and Werrington Park further upstream on the Ottery.

Peak behaviour is well captured for both gauging stations and there is a clear presence of spurious peaks during low flow periods. These are caused by a consistent presence of water in the subsurface storm-flow layer, which leads to surface runoff also during rather small rainfall events.

Additional work is needed to improve the non-linear interaction between the P-zone and the Richards equations columns with the aim of reducing the water resident in the P-zone during low flow periods. Section 5.3.5 discusses the calibration results in more detail.

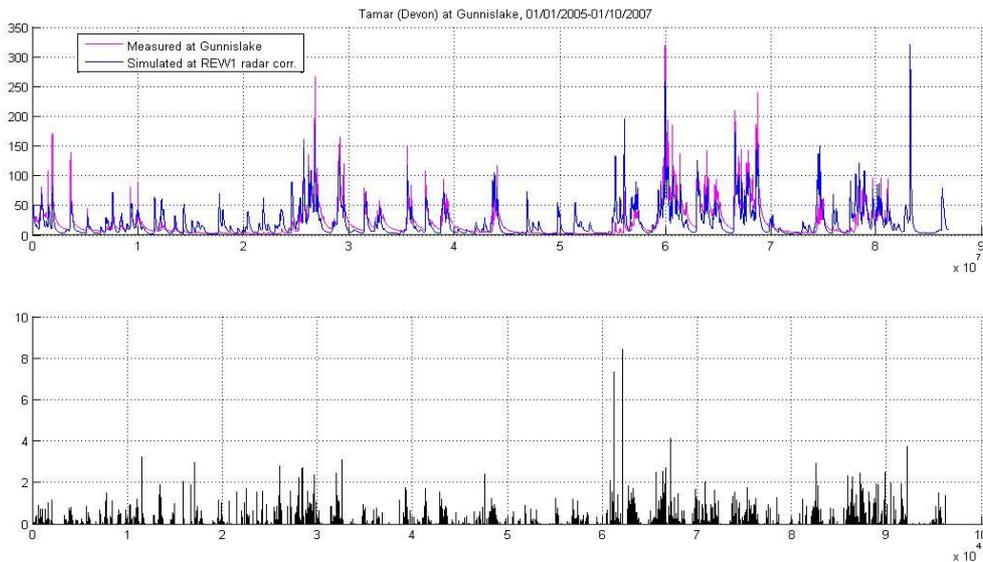


Figure 5.9 Modelled and observed discharges at Gunnislake, 1 January 2005 to 1 October 2007 (precipitation at REW 1).

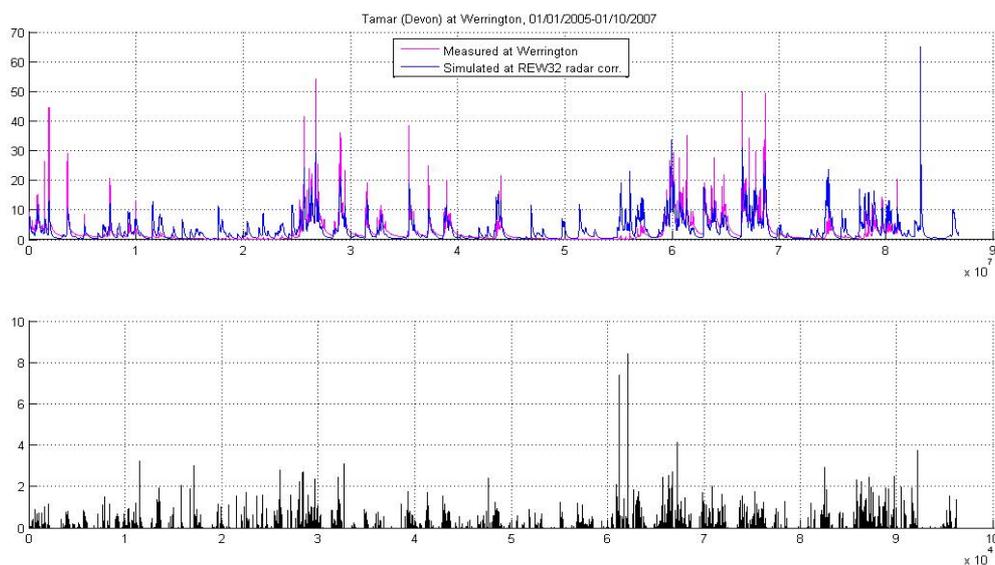


Figure 5.10 Modelled and observed discharges at Werrington , 1 January 2005 to 1 October 2007 (precipitation at REW 1).

5.3.5 Model performance (simulation mode)

As described above, the simulation results are acceptable for the Tamar gauging stations (Werrington Park and Gunnislake) but much less so for the Camel at Denby.

Table 5.6 lists selected performance measures for the simulated versus the observed discharge at the gauge locations. R^2 (Nash–Sutcliffe) efficiencies, such as those presented in Table 5.6, can range from $-\infty$ to 1. An efficiency of 1 corresponds to a perfect match of modelled discharge to the observed data while an efficiency of 0 indicates that the model predictions are as accurate as the mean of the observed data. Additionally, an efficiency of

less than zero occurs when the observed mean is a better predictor than the model. Clarke (2008) provides a recent review of this performance statistic.

Figures 5.11 and 5.12 show measured and modelled discharge for a winter period for the three calibration sites plus a site within the Tamar catchment (Crowford Bridge) which was treated as ungauged during calibration. More internal sites are displayed in Figures 5.13 and 5.14.

Although all model parameters were taken as uniform throughout each catchment, the results for the interior sites look promising and compare well with the sites used for calibration. One of the reasons for this may be attributed to the rather uniform geology within each of the two catchments. A closer look at the result for the Camel at Denby indicates that:

- the peaks are overestimated;
- the model reacts much too rapidly to rainfall.

On the other hand, the baseflow component is underestimated.

The results for summer events are less promising for both catchments. In particular, the results for the Boscastle event itself are not very good; this has implications on the result of the forecasts reported in the next section. In general the peaks in summer are overestimated but, for the Boscastle case, the peaks are underestimated in the Tamar at all stations and overestimated at Denby in the Camel (Figure 5.15).

Table 5.6 Model performance for the calibration and verification periods, January 2006 to August 2007 and 2004–2005 respectively.

	Location	R^2 efficiency	Mean absolute error	Bias
Calibration period	Gunnislake (Tamar)	0.5865	12.1972	0.7209
	Werrington (Ottery)	0.4693	1.9367	0.6792
	Denby (Camel)	-0.9380	4.4659	1.0399
Verification period	Gunnislake (Tamar)	0.5813	10.4516	2.0773
	Werrington (Ottery)	0.5258	1.8427	0.6856
	Denby (Camel)	-1.0693	4.2473	1.5415

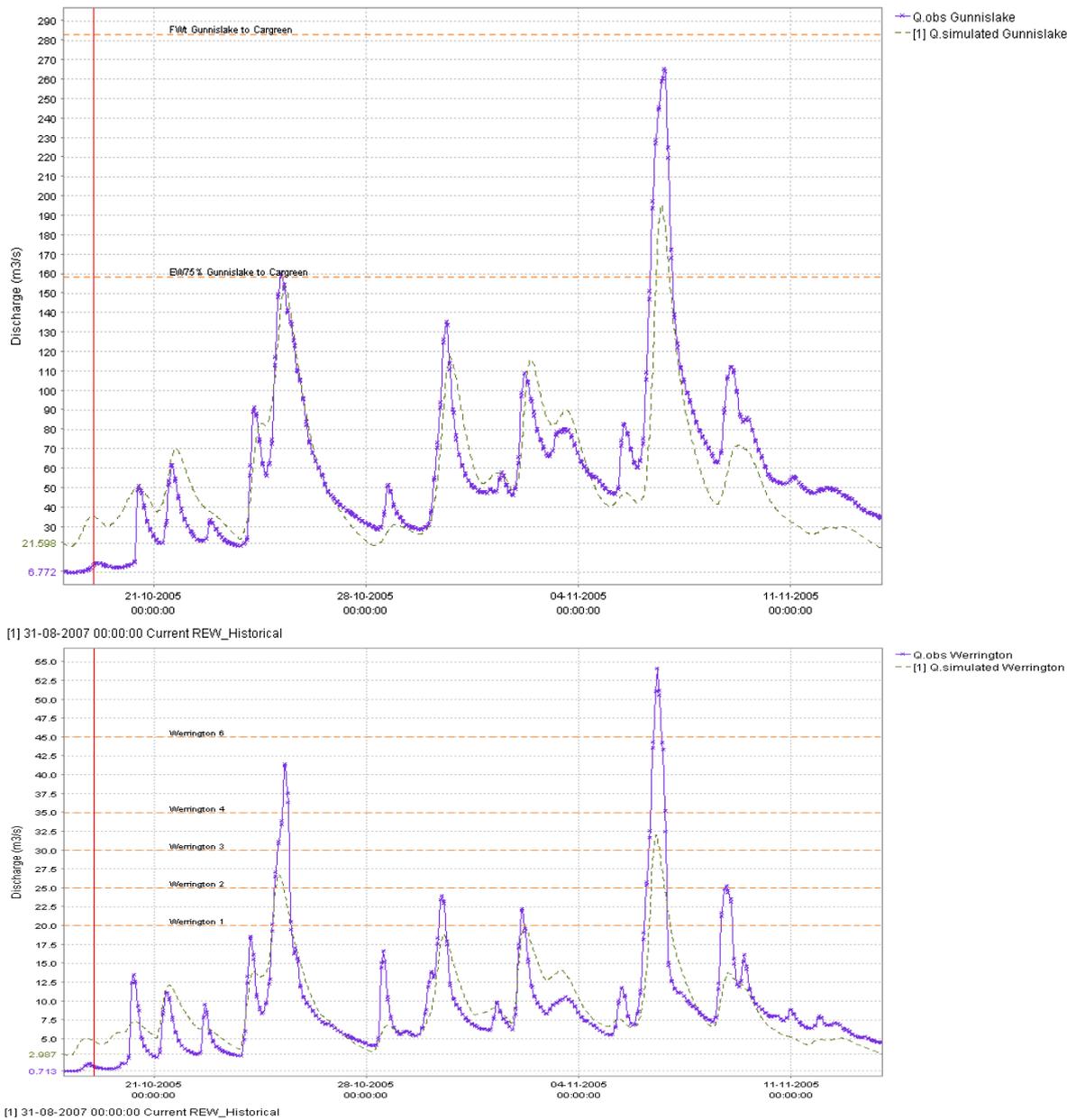


Figure 5.11 Measured and modelled discharge for 18 October to 14 November 2005 for Gunnislake (top) and Werrington (bottom).

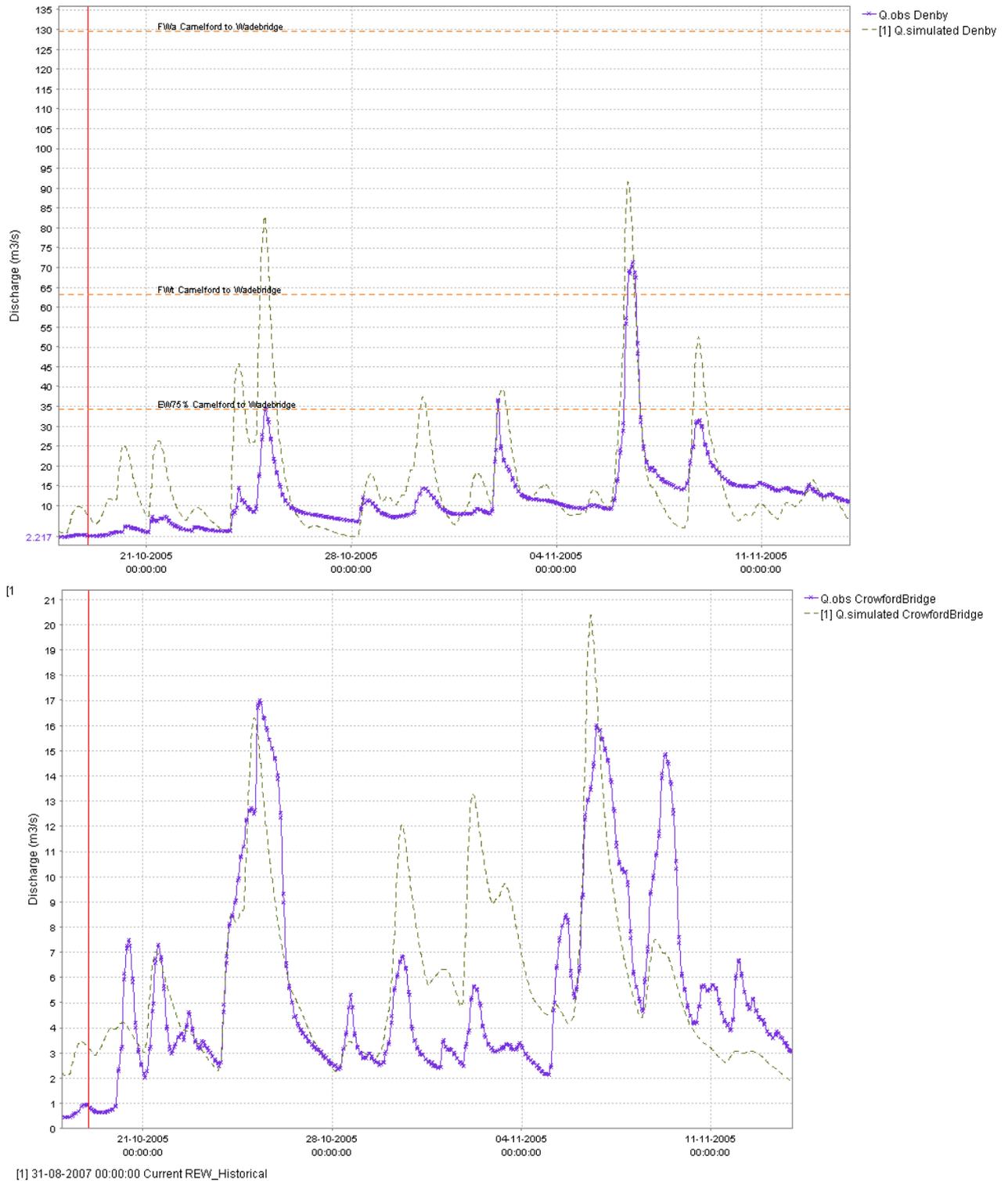


Figure 5.12 Measured and modelled discharge for 18 October to 14 November 2005 for Denby (top) and Crawford Bridge (bottom).

Notes: Crawford Bridge is an internal site treated as ungauged during calibration. This gives an indication of how distributed models can be used to estimate discharge (and flooding) for ungauged sites.

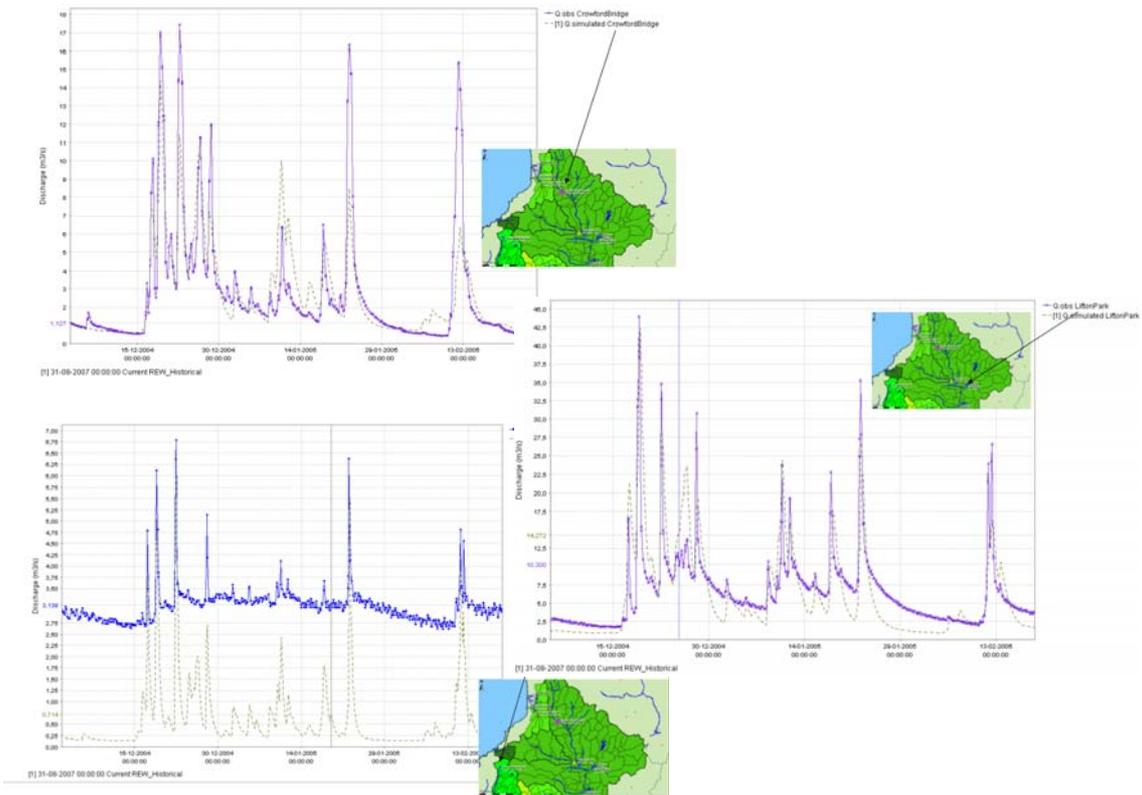


Figure 5.13 Results for a winter event for the internal sites Crawford Bridge (top), Lifton Park (middle) and Slaughterbridge (bottom).

Notes: Both Tamar sites perform rather well.
 For Slaughterbridge, a systematic error (offset) is visible. The latter will route down to Denby and influence those results as well.

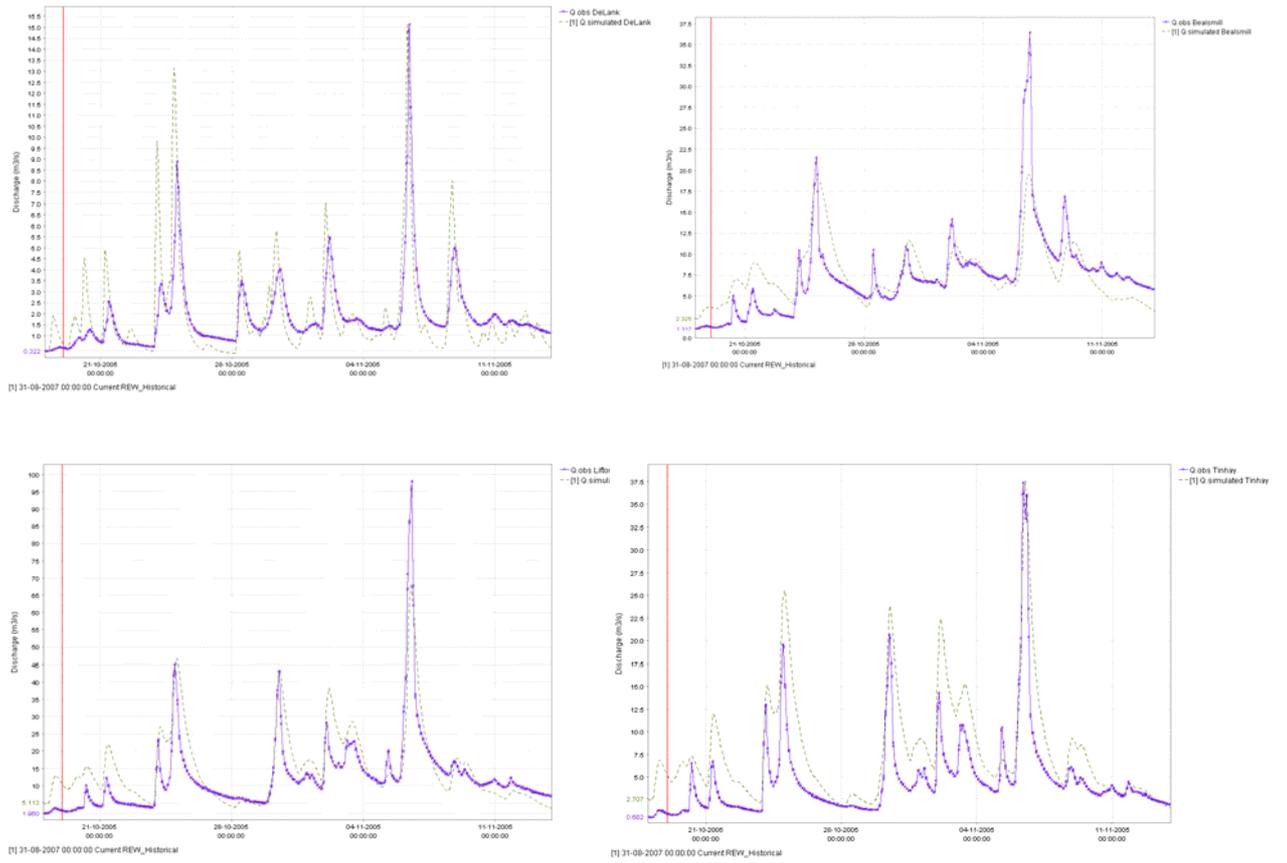


Figure 5.14 Measured and modelled discharge for 18 October to 14 November 2005 for four internal sites.

Notes: Top left to bottom right: De Lank, Beals Mill, Lifton Park and Tinhay

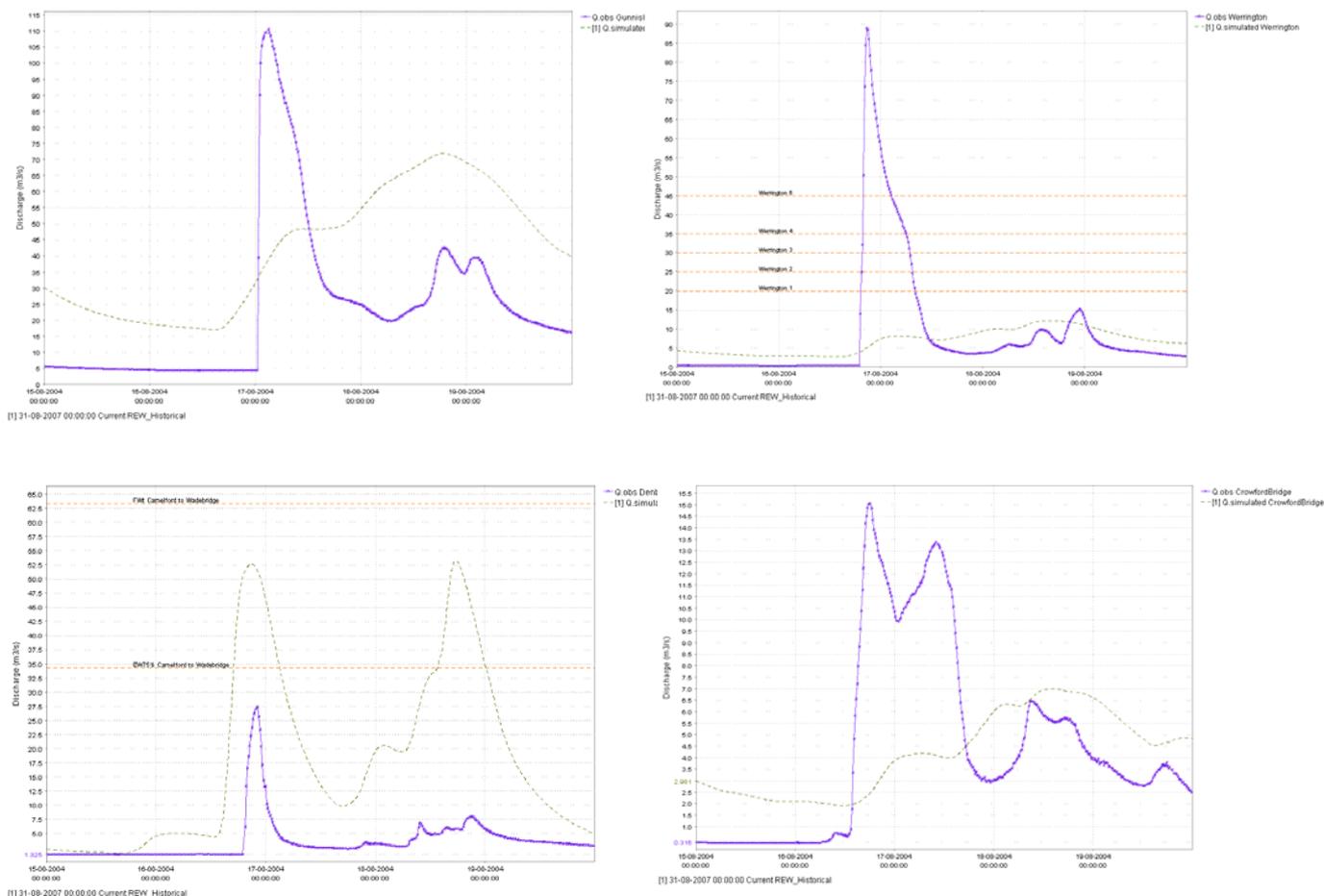


Figure 5.15 Results for the Boscastle event (in raw simulation mode) for, from top left to bottom right, Gunnislake, Werrington, Crowford Bridge (internal site) and Denby.

Although the type of performance described above is fairly typical for many simulation results, there are some features of the REW model that are lacking and require further examination. The two main shortcomings are both linked to the very simple SVAT scheme used in the present model:

- Transpiration losses are not taken from the soil component itself but are part of a bulk evapotranspiration (ET) component that is subtracted from the precipitation component. As a result the soil may not dry out enough in summertime (with high evapotranspiration losses and low precipitation). Therefore the model may start to generate quick runoff far too early.
- No interception component is modelled. In forested areas, in particular, this may lead to an overestimation of the net precipitation component.

Both limitations can be changed relatively easily in the current model structure and an experimental version is being worked on which will include a better interaction between vegetation and soil and between the different soil components.

If the REW model is used in the future, use of an adjusted version is suggested as the current limitations hamper its application to UK catchments and especially for those events to be forecast by the high resolution NWP, i.e. high intensity storms in summer.

However, the time available for calibration so far has been fairly limited. As such, only a limited number of parameters have been adjusted and no automatic calibration procedures to fine tune calibration could be used. It is assumed that more time for calibration would have allowed a better performance, regardless of the shortcomings mentioned above. This has been demonstrated in previous studies using the model.

For the Camel catchment, it is clear from the plots for Slaughterbridge that a groundwater component in the discharge is also being missed. This may be caused by an incorrect parameterisation of the groundwater component but could also indicate that the initial conditions (after running for 10 years) were not initialised properly.

Apart from the issues reported above, the Boscastle case shows extremely high rainfall intensities which were not present in the calibration period. The hydrological processes that may have been triggered in this event (e.g. widespread overland flow) are very rare in temperate regions and are poorly represented in most models. Although both infiltration excess overland flow (also called Hortonian overland flow, HOF) and saturation overland flow (SOF) are represented in all models, proper representation of these flow types usually requires a very high resolution model that includes these processes at the hillslope scale; the size of the current models does not allow for this. In addition, fast pathways to the stream may form during these high intensity events (e.g. by connecting zones of SOF and HOF areas that can form a continuous overland flow area to the stream channel that is normally not present), thereby invalidating the model set-up used for calibration because these flow processes did not occur in the calibration period.

5.3.6 Model performance (forecasting mode)

To evaluate the performance of the model in forecasting mode, workflows were set up in the standalone Delft-FEWS system to run the REW model in forecasting mode.

- For each forecast location with gauge data available, an ARMA model was set up using automatic parameter determination to represent the operational setting (see Section 5.3.6).
- For the entire period, historical runs were made saving initial conditions for each day (at 00:00). This was done in raw simulation mode without updating.
- Input to the model in forecast mode was made up of:
 - perfect rainfall forecast (the HyradK-adjusted radar) over the entire period;
 - measured discharge up to the time of forecast, T₀ (to the ARMA model).
- Forecasts were made for every 15-minute time step during the selected periods.
- Fixed lead time series were extracted from these forecasts for 1, 2, 4, 6, 8, 10 and 12 hour lead times

The results of this exercise gave an indication of the performance of the REW model set-up in forecasting mode itself, assuming that the HyradK-adjusted rainfall provides a good estimate of the actual rainfall over an REW.

Run time limitations meant it was not possible to run forecasts for each 15-minute interval for the entire period and a number of periods were selected to evaluate forecast performance using HyradK-adjusted rainfall (Table 5.7).

Table 5.7 Event periods (start and end times are at 00:00).

Period No.	Start date	End date	Over Gunnislake threshold	Summer / winter	Comments
1	25/07/2003	26/07/2003	N	S	Possible convective event Large response at Denby Relevant for Boscastle event NOT used for REW.
2	10/12/2003	10/02/2004	Y (three times)	W	
3	15/08/2004	20/08/2004	N	S	Boscastle event
4	02/10/2004	02/11/2004	Y (twice)	W	
5	14/12/2004	26/01/2005	Y (three times)	W	
6	10/02/2005	15/02/2005	N	W	Interesting double peak Isolated event Just under threshold
7	18/10/2005	14/11/2005	Y (twice)	W	
8	26/11/2005	10/12/2005	Y (once)	W	
9	19/05/2006	27/05/2006	N	S	Just under threshold (peak $\sim 125 \text{ m}^3 \text{ s}^{-1}$)
10	14/11/2006	14/12/2006	Y (four times)	W	Includes largest peak on record.
11	07/02/2007	12/03/2007	Y (four times)	W	
12	09/05/2007	19/05/2007	N	S	Under threshold (peak $\sim 110 \text{ m}^3 \text{ s}^{-1}$)

ARMA model and configuration

The ARMA model is applied to improve model time series predictions through combining modelled series and observed series. As input, it uses an output series from a forecasting module (typically discharge from a routing or rainfall-runoff module) and the observed series at the same location. An updated series for the module output is again returned by the module. Updating is applied through application of an error model to the residuals between module output and observed series. This error model is applied also to the forecast data from this module to allow correction of errors in the forecast.

The configuration of the error modelling module is used to determine its behaviour in establishing the statistical model of the error and how this is applied to derive the updated series.

Configuration items are as follows:

- Order_AR – (maximum) order of the AR component;

- Order_MA – 0;
- Order_Sel – option to determine if the orders are to be derived automatically (with the maxima as defined above) or as given;
- Transform – option to apply a transformation to residuals (may either be ‘none’, ‘mean’ or ‘boxcox’);
- Lambda – a required parameter for the ‘boxcox’ transformation option.

Time series definitions

Three types of time series models can be distinguished:

- autoregressive (AR);
- moving average (MA);
- combined ARMA type.

An ARMA(p,q) process for a variable x_n can be written as (Priestley 1981):

$$x_n + a_1x_{n-1} + \dots + a_px_{n-p} = \varepsilon_n + b_1\varepsilon_{n-1} + \dots + b_q\varepsilon_{n-q} \quad (11)$$

where variable, ε_n , derives from a purely random process giving a sequence of independent identically distributed stochastic variables with zero mean and variance, σ_ε^2 . The coefficients, a_i and b_i , are model parameters to be estimated. This process is purely AR for $q=0$ and MA for $p=0$.

AR estimation

Burg’s method – also denoted as maximum entropy (Burg 1967, Kay and Marple 1981) – is used for parameter estimation to ensure that the model will be stationary. Asymptotic AR order selection criteria can give wrong orders if candidate orders are higher than $0.1N$ (N is the signal length). The finite sample criterion, $CIC(p)$, is used for model selection (see Broersen 2000). The model with the smallest value of $CIC(p)$ is selected. CIC (combined information criterion) uses a compromise between the finite sample estimator for the Kullback–Leibler information (Broersen and Wensink 1998) and the optimal asymptotic penalty factor 3 (Broersen 2000, Broersen and Wensink 1996).

Box–Cox transformations

The Box–Cox transformation (Box and Cox 1964) can be applied in the order selection and estimation of the coefficients. The object in doing so is usually to make the residuals more homoscedastic and closer to a normal distribution. The transform for a variable y is defined as:

$$T(y) = (y^\lambda - 1) / \lambda \quad (12)$$

when the Box–Cox transform parameter, λ , is not equal to zero. When $\lambda=0$, then $T(y)=\log(y)$.

Application of the module

The implemented algorithm computes AR(p) models with $p=0,1,\dots,N/2$ and selects a single best AR model with CIC . This automatic mode was used in this study. The settings used are shown in Table 5.8.

Table 5.8 Settings used in module application.

Parameter	Setting
orderSelection	true
order_ar	3
order_ma	3
subtractMean	false
boxcoxTransformation	false
lambda	0

Results

Graphs showing lead time versus model efficiency for the periods defined in Table 5.7 are presented in Figures 5.16 to 5.18. A summary table is shown in Table 5.9.

For most periods the forecast model efficiency is better than the raw simulation (without the ARMA correction) for lead times up to about six hours and sometimes even up to 12 hours.

For some cases, especially for Werrington, the forecast efficiency drops below the raw simulation after four hours. Better performance for longer lead time for those cases might be obtained by changing the ARMA model configuration to always go back to the raw simulation after a number of hours.

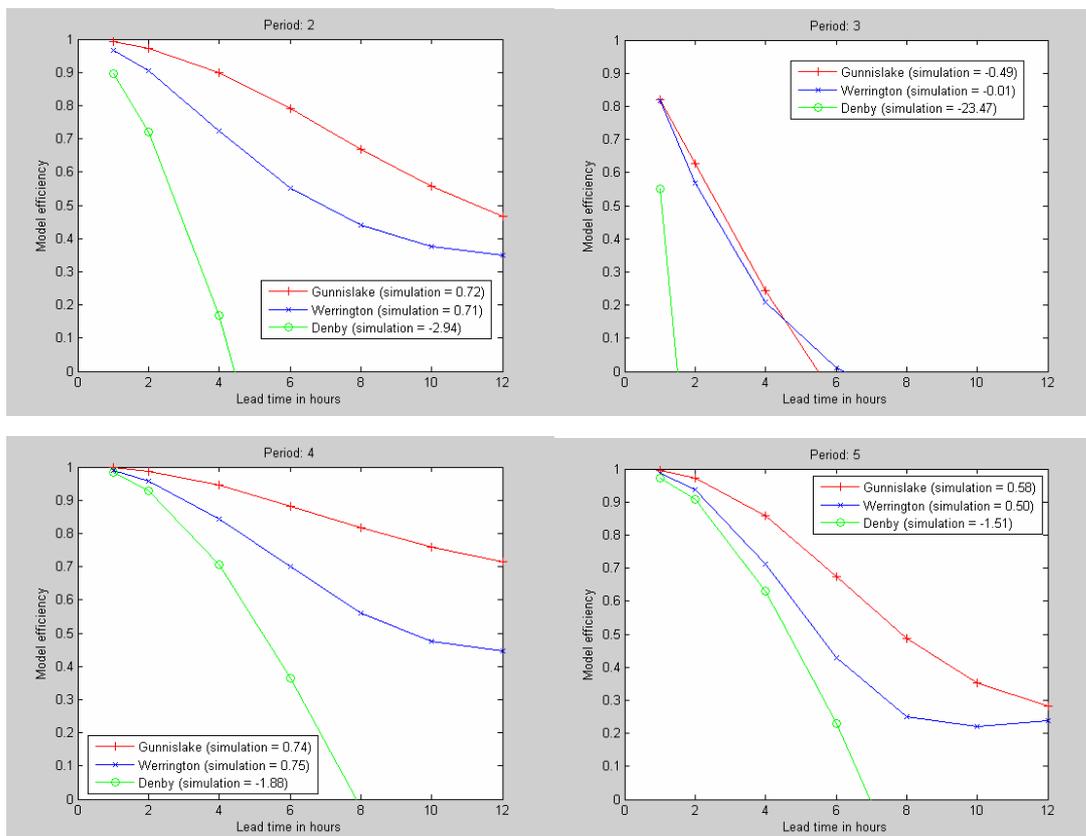


Figure 5.16 Model efficiency (R^2) versus lead time for periods 2–5 for Gunnislake, Denby and Werrington.

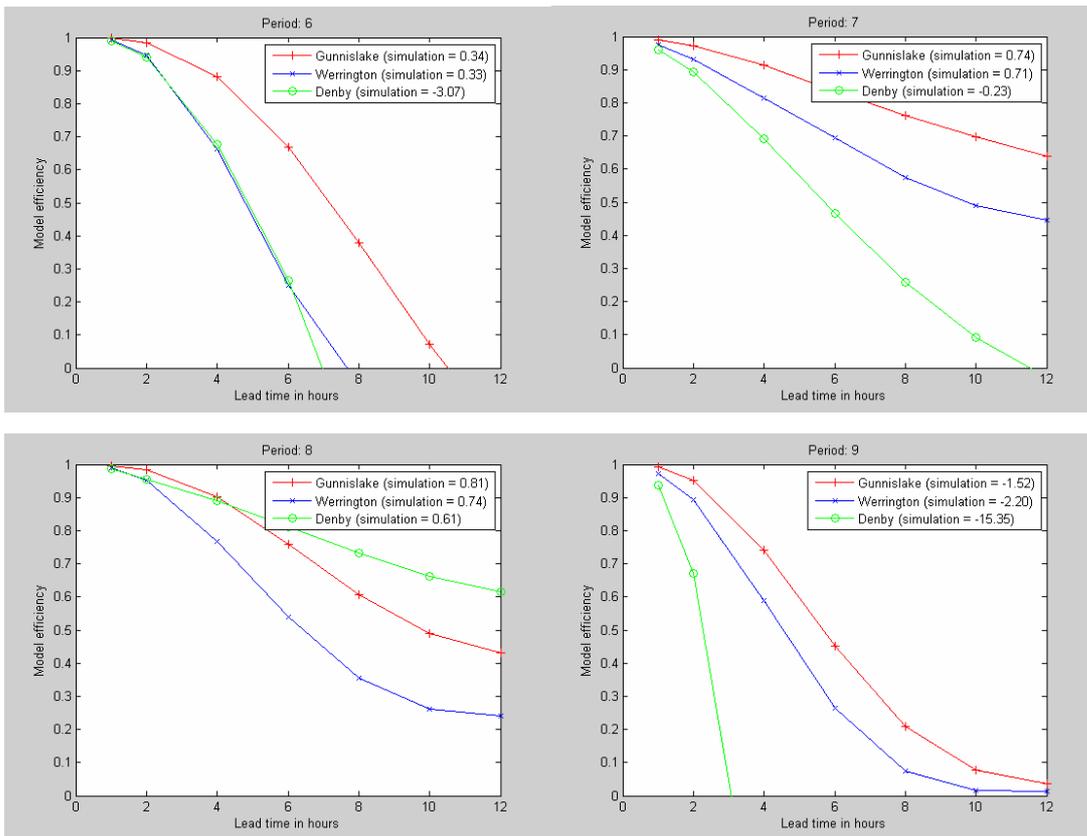


Figure 5.17 Model efficiency (R^2) versus lead time for periods 6–9 for Gunnislake, Denby and Werrington.

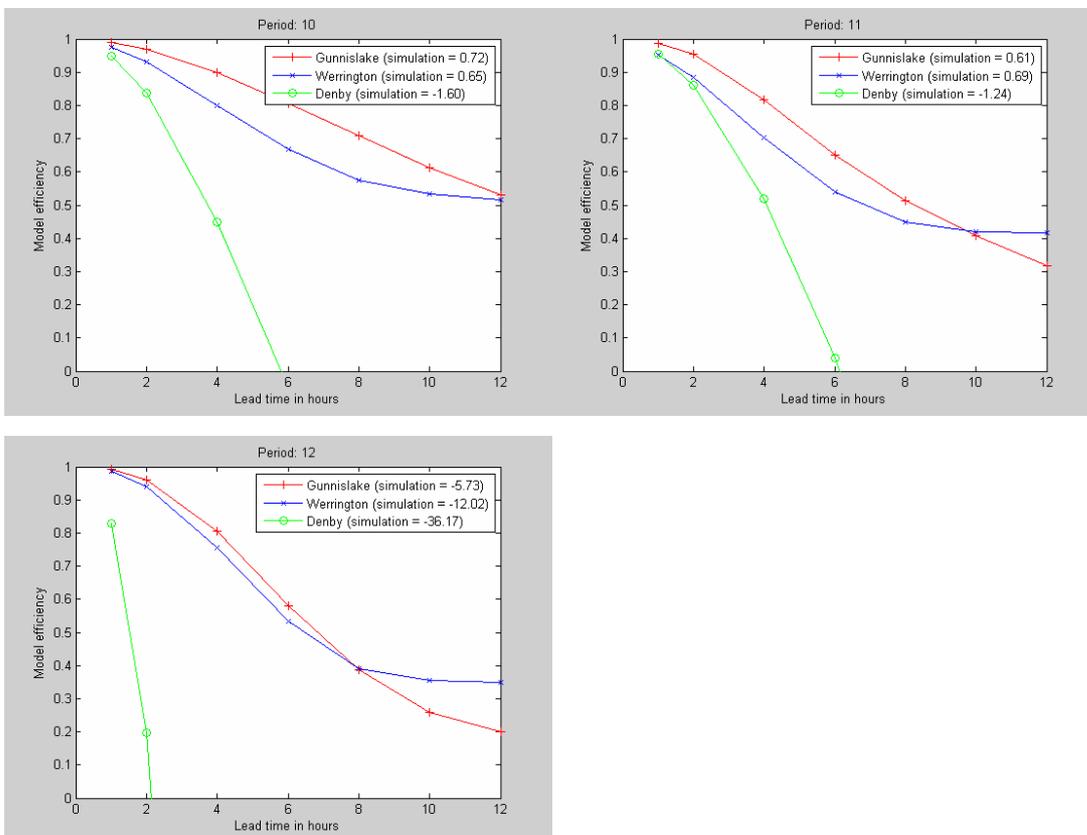


Figure 5.18 Model efficiency (R^2) versus lead time for periods 10–12 for Gunnislake, Denby and Werrington.

Table 5.9 R^2 model efficiency for all events ¹.

10-12-2003				10-02-2004				Period 2				15-08-2004				20-08-2004				Period 3				
Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	
1		0.99	0.97	0.9				1	0.82	0.82	0.55					1	0.99	0.99	0.97					
2		0.97	0.91	0.72				2	0.63	0.57	-0.55					2	0.99	0.97	0.94					
4		0.9	0.72	0.17				4	0.24	0.21	-3.47					4	0.86	0.71	0.63					
6		0.79	0.55	-0.64				6	-0.08	0.01	-6.53					6	0.67	0.43	0.23					
8		0.67	0.44	-1.58				8	-0.36	-0.08	-9.63					8	0.49	0.25	-0.24					
10		0.56	0.38	-2.41				10	-0.53	-0.12	-12.25					10	0.35	0.22	-0.71					
12		0.47	0.35	-3.06				12	-0.59	-0.15	-14.32					12	0.28	0.24	-1.11					
Simulation		0.72	0.71	-2.94				Simulation		-0.49	-0.01	-23.47				Simulation		0.58	0.5	-1.51				
02-10-2004				02-11-2004				Period 4				14-12-2004				26-01-2005				Period 5				
Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	
1		1	0.99	0.98				1	0.99	0.99	0.97					1	0.99	0.97	0.96					
2		0.99	0.96	0.93				2	0.97	0.94	0.91					2	0.97	0.93	0.89					
4		0.94	0.84	0.71				4	0.86	0.71	0.63					4	0.91	0.82	0.69					
6		0.88	0.7	0.36				6	0.67	0.43	0.23					6	0.84	0.69	0.47					
8		0.82	0.56	-0.03				8	0.49	0.25	-0.24					8	0.76	0.58	0.26					
10		0.76	0.48	-0.44				10	0.35	0.22	-0.71					10	0.7	0.49	0.09					
12		0.71	0.45	-0.81				12	0.28	0.24	-1.11					12	0.64	0.45	-0.03					
Simulation		0.74	0.75	-1.88				Simulation		0.58	0.5	-1.51				Simulation		0.74	0.71	-0.23				
10-02-2005				15-02-2005				Period 6				18-10-2005				14-11-2005				Period 7				
Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	
1		1	0.99	0.99				1	0.99	0.97	0.96					1	0.99	0.97	0.96					
2		0.98	0.95	0.94				2	0.97	0.93	0.89					2	0.97	0.93	0.89					
4		0.88	0.66	0.68				4	0.91	0.82	0.69					4	0.91	0.82	0.69					
6		0.67	0.25	0.26				6	0.84	0.69	0.47					6	0.84	0.69	0.47					
8		0.38	-0.05	-0.29				8	0.76	0.58	0.26					8	0.76	0.58	0.26					
10		0.07	-0.17	-1.03				10	0.7	0.49	0.09					10	0.7	0.49	0.09					
12		-0.2	-0.25	-1.81				12	0.64	0.45	-0.03					12	0.64	0.45	-0.03					
Simulation		0.34	0.33	-3.07				Simulation		0.74	0.71	-0.23				Simulation		0.74	0.71	-0.23				
26-11-2005				10-12-2005				Period 8				19-05-2006				27-05-2006				Period 9				
Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	
1		1	0.99	0.99				1	0.99	0.97	0.96					1	0.99	0.97	0.96					
2		0.98	0.95	0.96				2	0.95	0.89	0.67					2	0.95	0.89	0.67					
4		0.9	0.77	0.89				4	0.74	0.59	-0.58					4	0.74	0.59	-0.58					
6		0.76	0.54	0.81				6	0.45	0.26	-2.43					6	0.45	0.26	-2.43					
8		0.61	0.36	0.73				8	0.21	0.07	-4.62					8	0.21	0.07	-4.62					
10		0.49	0.26	0.66				10	0.08	0.02	-6.96					10	0.08	0.02	-6.96					
12		0.43	0.24	0.61				12	0.04	0.01	-9.2					12	0.04	0.01	-9.2					
Simulation		0.81	0.74	0.61				Simulation		-1.52	-2.2	-15.35				Simulation		-1.52	-2.2	-15.35				
14-11-2006				14-12-2006				Period 10				07-02-2007				12-03-2007				Period 11				
Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	
1		0.99	0.97	0.95				1	0.99	0.95	0.95					1	0.99	0.95	0.95					
2		0.97	0.93	0.84				2	0.95	0.88	0.86					2	0.95	0.88	0.86					
4		0.9	0.8	0.45				4	0.82	0.7	0.52					4	0.82	0.7	0.52					
6		0.81	0.67	-0.05				6	0.65	0.54	0.04					6	0.65	0.54	0.04					
8		0.71	0.57	-0.59				8	0.51	0.45	-0.48					8	0.51	0.45	-0.48					
10		0.61	0.53	-1.1				10	0.41	0.42	-0.93					10	0.41	0.42	-0.93					
12		0.53	0.52	-1.5				12	0.32	0.42	-1.25					12	0.32	0.42	-1.25					
Simulation		0.72	0.65	-1.6				Simulation		0.61	0.69	-1.24				Simulation		0.61	0.69	-1.24				
09-05-2007				19-05-2007				Period 12																
Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	Hour	Gunnislake	Werington	Denby	
1		0.99	0.99	0.83				1	0.99	0.99	0.83					1	0.99	0.99	0.83					
2		0.96	0.94	0.2				2	0.96	0.94	0.2					2	0.96	0.94	0.2					
4		0.81	0.76	-2.41				4	0.81	0.76	-2.41					4	0.81	0.76	-2.41					
6		0.58	0.53	-6.89				6	0.58	0.53	-6.89					6	0.58	0.53	-6.89					
8		0.39	0.39	-12.6				8	0.39	0.39	-12.6					8	0.39	0.39	-12.6					
10		0.26	0.36	-18.89				10	0.26	0.36	-18.89					10	0.26	0.36	-18.89					
12		0.2	0.35	-25.17				12	0.2	0.35	-25.17					12	0.2	0.35	-25.17					
Simulation		-5.73	-12.02	-36.17				Simulation		-5.73	-12.02	-36.17				Simulation		-5.73	-12.02	-36.17				

Notes ¹ Grey cells denote that the forecasting R^2 for that lead time is lower than that of the raw simulation indicating under/overshoot of the ARMA model.

5.4 PDM application

5.4.1 Model set-up

The standard form of the PDM model was initially used for all case studies with:

- a cubic baseflow storage;
- a cascade of two unequal reservoirs for the surface storage;
- a truncated Pareto distribution of soil/vegetation absorption capacity.

Where appropriate, the soil tension capacity, S_t , influencing drainage to groundwater and evaporation, was allowed to be non zero and modelling of catchment returns/abstractions was invoked through adding a constant flow, q_c . Section 4.1 provides a more comprehensive model description.

The calibration process was split into two parts. First, the process model parameters were calibrated in simulation mode where the model deterministically calculates simulated flow using only the input data (rainfall and potential evaporation), completely ignoring the observed flow (except for model initialisation). Secondly, when considering the real-time flood forecasting application, the model was run in forecast mode which aims to emulate real-time application in an offline environment and is used to calibrate the state updating parameters. In this case, the HyradK raingauge-adjusted rainfall data were used as 'perfect' foreknowledge of forecast rainfall. The calibrated PDM model parameters are presented in Table 5.10.

The 50-metre IHDTM was used to delineate the catchment boundaries for the case study catchments and to provide the catchment areas needed by the PDM model. The boundaries were used to calculate catchment average rainfall from the HyradK raingauge-adjusted radar rainfall estimates.

5.4.2 Tamar at Gunnislake

As stated above, flood forecasting of the Tamar catchment would normally be undertaken by a network of models so as to utilise the upstream gauging stations. This was beyond the scope of the project but the PDM model should still perform reasonably well. As this was the largest catchment and includes the Ottery sub-catchment, it was calibrated first.

The observations at Gunnislake reveal a fairly responsive catchment with a moderate baseflow contribution. The model parameters were calibrated manually (apart from the state updating parameter). The size of the catchment is reflected in the long time-delay parameter. A good model fit was achieved over the calibration period except for an overestimate of the model during the wetting up of the catchment after the very dry 2006 summer (this is evident across all catchments and models). This was alleviated somewhat by employing a minimum soil capacity storage which dampens the initial response of the catchment to rainfall.

Performance over the entire calibration period is presented in Figure 5.19, which shows that the PDM model captured the long-term slow response of the catchment very well, with excellent agreement on the recessions. The peaks also modelled well and overall the model calibration is very acceptable. Further analysis is given in Section 5.4.5.

Table 5.10 PDM model parameters for the Camel, Ottery and Tamar catchments.

Parameter name	Catchment		
	Camel at Denby	Ottery at Werrington Park	Tamar at Gunnislake
Rainfall factor f_c	1.0	1.05	1.0
Time Delay τ_d	0.75	2.5	3.25
Soil Moisture			
C_{min}	22.0	24.0	45.0
C_{max}	68.0	59.0	140.0
b	0.3	0.6	0.450
Evaporation function b_e	2.75	2.75	2.75
Recharge function			
k_g	40000	25000	40000
b_g	2.67	3.1	2.47
S_t	16.5	24.0	65.0
Surface routing <i>Cascade of 2 linear reservoirs</i>			
k_1	6.0	3.5	5.0
k_{21}	3.5	3.75	8.0
Baseflow storage (cubic) k_b	170.0	100.0	45.0
Returns/abstractions q_c	0.6	0.05	2.0
State updating			
$gain_s$	1.6607	1.7613	1.5870
$gain_b$	0.47636	1.3217	0.87134

5.4.3 Ottery at Werrington

Comparison of the observed flow hydrographs (Figure 5.19) at Werrington and the downstream station at Gunnislake immediately reveals differences in catchment behaviour. In particular, the Werrington catchment has a flashier response to rainfall. This is due to both the smaller catchment area and the different soil dominance within the catchment (see Section 5.1). In addition, analysis of the major peaks at Werrington reveals a strange behaviour of the hydrograph with the recession limb falling quicker than the rising limb – this makes calibration even trickier.

As Werrington is a sub-catchment of the Gunnislake catchment, the Gunnislake PDM parameters were used as a starting point for the Werrington calibration. These were then refined manually and required shallower soil stores, quicker surface routing parameters and a shorter time delay.

Interestingly, the Ottery catchment is the only site that required a rainfall factor other than 1; the setting of 1.05 suggests a slight underestimation of the catchment average rainfall by the raingauge-adjusted HyradK data.

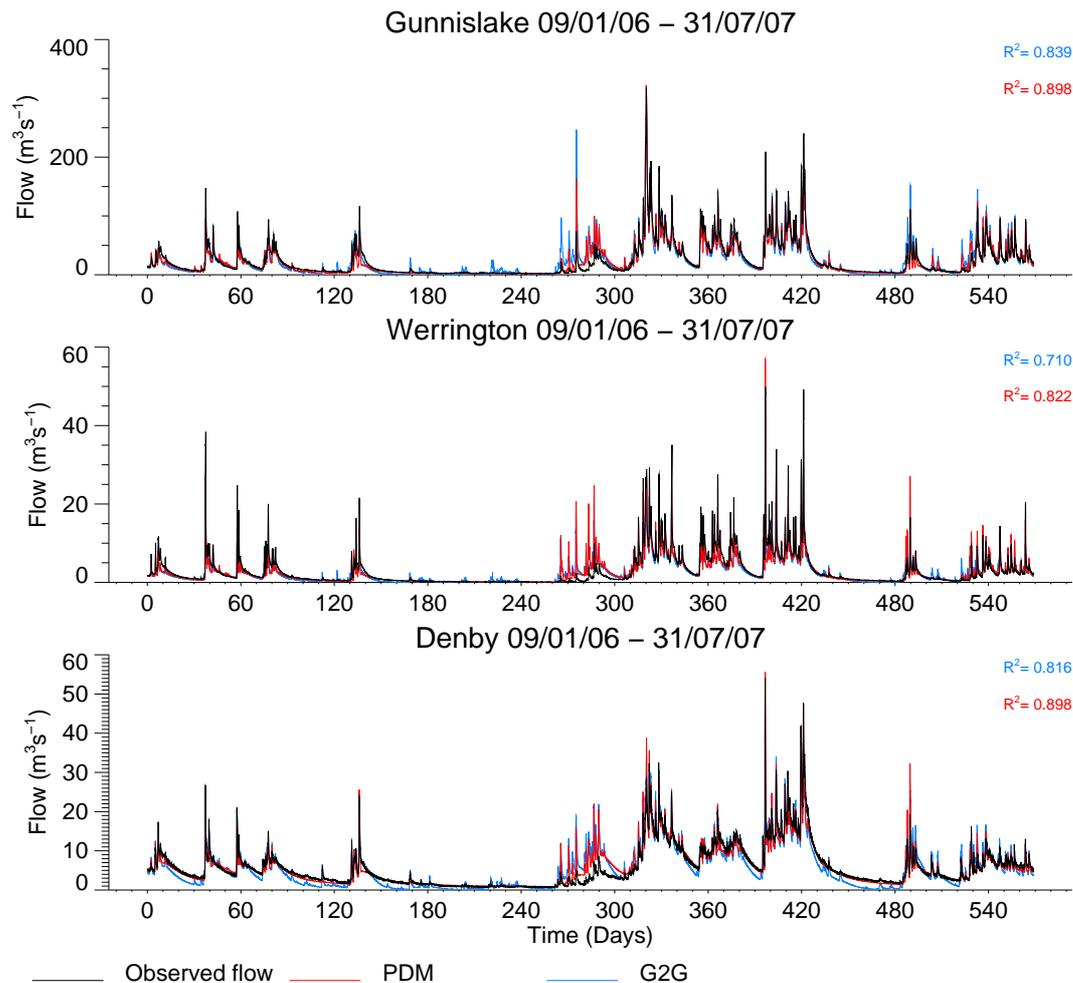


Figure 5.19 Modelled (PDM and G2G) and observed hydrographs at the Gunnislake, Werrington and Denby locations for the calibration period 9 January 2005 to 1 October 2007.

The performance over the entire calibration period is presented in Figure 5.19, which shows that the PDM model captured the long-term slow response of the catchment very well, with very good agreement on the recessions. Again problems over the wetting-up period after the dry 2006 summer are evident. There is a general trend for the observed peaks to be underestimated by the model. A satisfactory calibration that captured the peaks better could not be achieved. Overall the model calibration is acceptable. Further analysis is given in Section 5.4.5.

5.4.4 Camel at Denby

Comparison of the observed hydrograph at Denby with the Tamar catchment (Figure 5.19) highlights the different behaviour of the Camel catchment. In particular, there is a much larger baseflow component which has a prolonged seasonal effect but a flashy short-term response is still present. In addition, the peak flows are comparable to Werrington despite being a larger catchment (210 km² compared with 121 km² for Werrington). Again this can be attributed to the different soil and geology present in the Camel catchment relative to the

Tamar catchment (see Section 5.1). In particular the peat of Bodmin Moor and the deeper soils associated with the HOST classes give rise to the different baseflow characteristic.

The Denby PDM parameters were calibrated manually from scratch and show some differences from those calibrated for the Werrington and Gunnislake models. A slower release from the subsurface store (k_b) was needed to model the prolonged baseflow component of the observations. Keeping the correct balance between water moving to the subsurface store to maintain the baseflow component and the generation of sufficient surface runoff to model the short-term response requires careful selection of the soil store related components. In particular, the soil stores were not too deep (c_{max}) so as to maintain the possibility of saturation excess in the short term while the recharge function allowed for significant and sustained recharge to the subsurface store to occur.

Performance over the entire calibration period is presented in Figure 5.19, which shows that the PDM model captured the long-term slow response of the catchment very well, with very good agreement on the recessions. The problems over the wetting-up period after the dry 2006 summer are once again evident. The observed peaks are also well modelled and overall the calibration is very acceptable. Further analysis is given in Section 5.4.5.

5.4.5 Model performance (simulation mode)

The simulation mode performance of the PDM model is generally very good at all three locations (Gunnislake, Werrington and Denby). The performance statistics of the PDM model are presented in Table 5.11 and the simulated flows are compared with observed flows over the calibration period in Figure 5.19 and the evaluation period in Figure 5.20.

The hydrograph simulations clearly show that PDM successfully modelled the recession and seasonal behaviour of the catchments, only exception being the 'wetting up' period following the dry summer in 2006. The peak responses are also well modelled at Denby and Gunnislake. Modelling peak flows at Werrington was more challenging and marginally less successful, which is reflected in the slightly poorer performance statistics. To give an impression of the short-term behaviour of the model, example winter and summer events are presented in the left and right columns of Figure 5.21 respectively (note the relatively low peak flows for the summer).

The PDM simulation mode results over the Boscastle event are presented in Figure 5.22 and show mixed results. The Boscastle storm was an extreme convective event with very high and localised rainfall totals – an event very different to the 'typical' storms used in model calibration. This poses real difficulties for a conceptual lumped model. In particular, the catchment average rainfall time series used as model input will smooth out these peak intensities making it very difficult to generate the large surface runoffs that were observed. However, the PDM simulations were really rather good at Werrington in both magnitude and timing. This probably reflects the fact that Werrington had the largest percentage area coverage by the storm and so the lumped conceptualisation was still a good one. At Gunnislake and particularly Denby, the percentage coverage is much less and the observed flows were underestimated by the PDM model.

Table 5.11 Simulation mode performance of G2G and PDM models over calibration and verification periods.

	Location	R^2 Efficiency		Mean absolute error		Bias	
		PDM	G2G	PDM	G2G	PDM	G2G
Calibration period	Tamar at Gunnislake	0.898	0.839	5.076	6.818	-0.635	0.334
	Ottery at Werrington Park	0.822	0.710	0.864	1.094	-0.229	-0.377
	Camel at Denby	0.898	0.816	0.923	1.590	0.000	-0.228
Verification period	Tamar at Gunnislake	0.913	0.871	4.170	5.282	-0.454	0.752
	Ottery at Werrington Park	0.857	0.747	0.670	0.920	-0.046	-0.119
	Camel at Denby	0.922	0.839	0.722	1.285	0.103	0.023

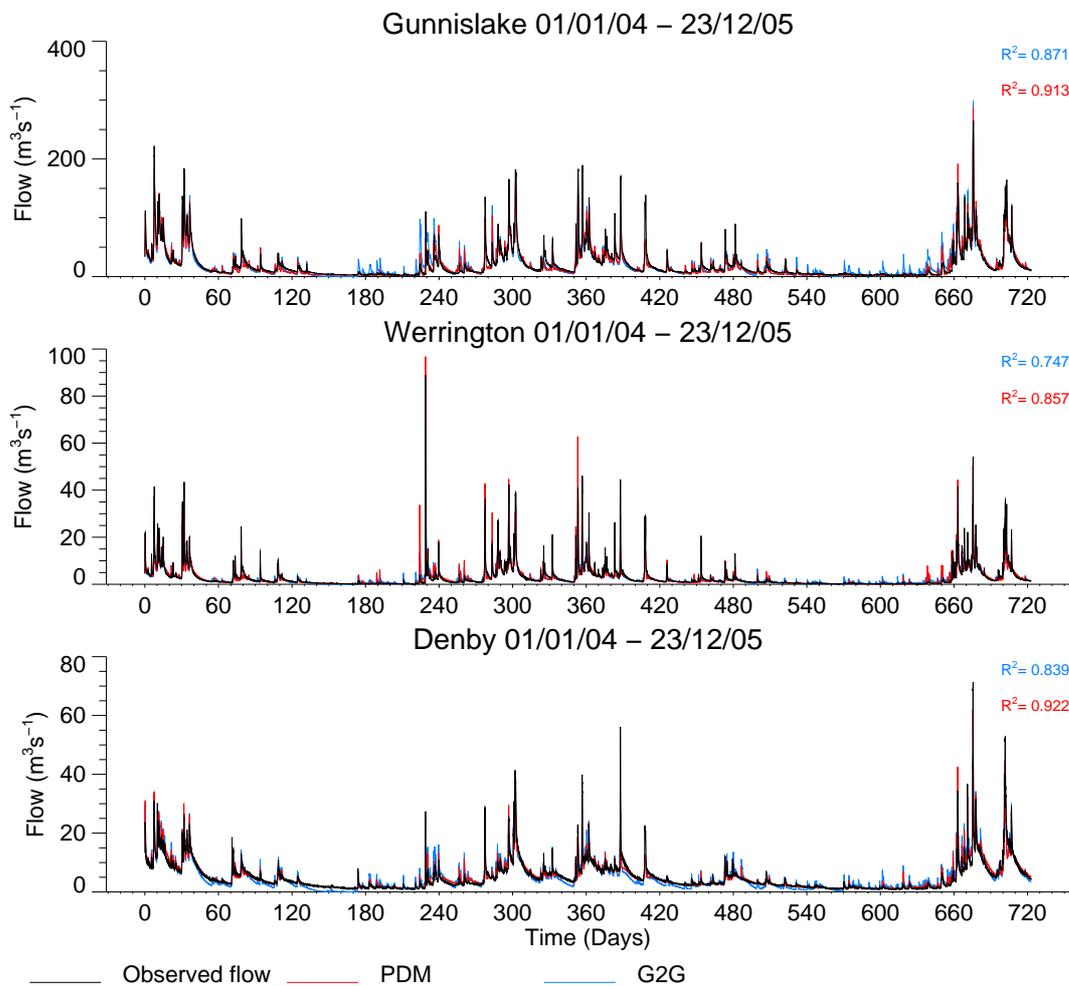


Figure 5.20 Modelled (PDM and G2G) and observed hydrographs at the Gunnislake, Werrington and Denby locations for the evaluation period 1 January 2004 to 23 December 2005.

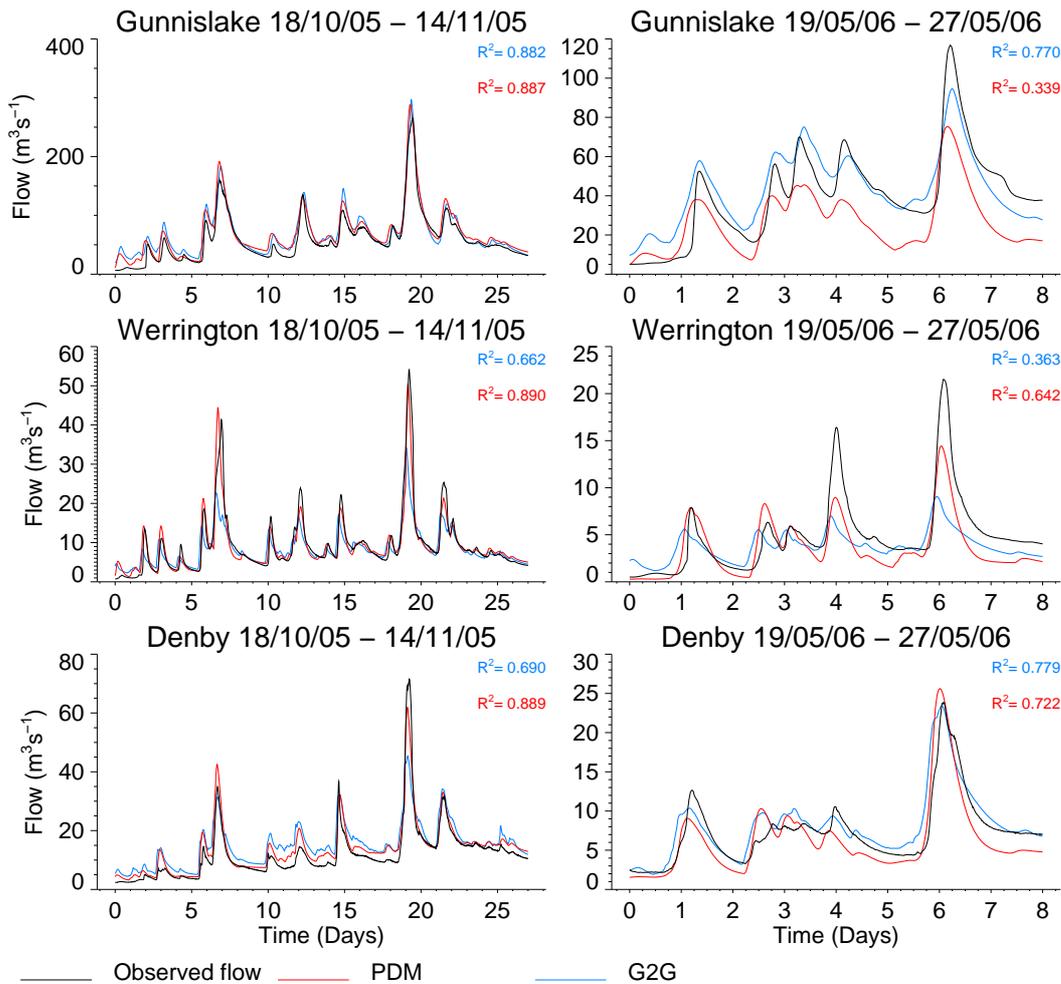


Figure 5.21 Modelled (PDM and G2G) and observed hydrographs at the Gunnislake, Werrington and Denby locations for the evaluation winter period 18 October to 14 November 2005 and the summer period 19 to 27 May 2006.

Encouragingly, the performance statistics and simulation hydrographs give consistent performance over both the calibration and evaluation periods, giving confidence that the model calibrations are robust. The impressive R^2 statistics are all in excess of 0.822 and show that the flexibility afforded by the lumped conceptual formulation of the PDM model can give rather good results and, in usual storm conditions, provide a tough benchmark to better.

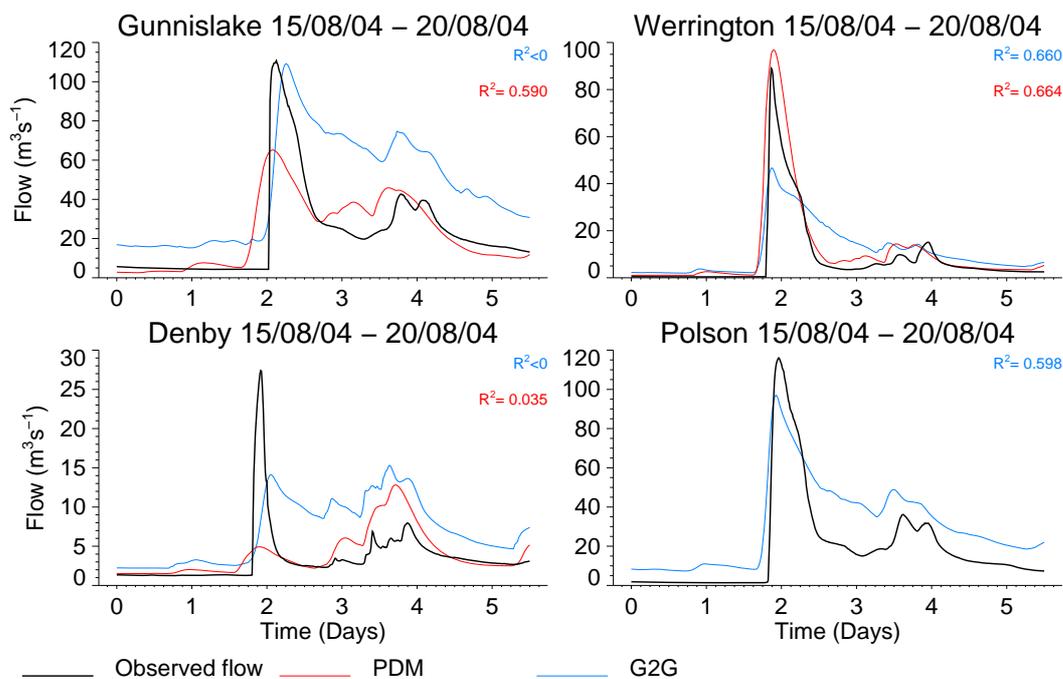


Figure 5.22 Modelled and observed hydrographs using the PDM and G2G models for the period 15 to 20 August 2004 covering the Boscastle event.

5.4.6 Model performance (forecasting mode)

The forecast mode parameters of the PDM model ($gain_s$ and $gain_b$) were calibrated by employing the HyradK raingauge-adjusted radar data as a ‘perfect rainfall forecast’. Forecasts were made for lead times out to 24 hours at every 15-minute time step within events. Empirical state correction, which utilises the observed flow data to correct the internal states of the PDM model, was applied up to each forecast origin.

Figure 5.23 presents model efficiency (R^2) versus lead time over the entire calibration and evaluation periods for Gunnislake, Denby and Werrington. Results are also shown for the 12 shorter periods listed in Table 5.7 and used within the REW analysis of in Section 5.3.6.

The forecast performance of all three PDM models is consistent over the calibration and evaluation periods. Figure 5.23 indicates the considerable performance improvement state correction offers for lead times out to around 12 hours, beyond which the forecast mode performance tails off to the simulation mode performance as expected. The analysis over the shorter periods (1–12) shows similar results to the calibration and evaluation periods for the longer winter time events (2, 4–8, 10, 11). Forecast performance over the summer events (1, 3, 9, 12) is much more variable.

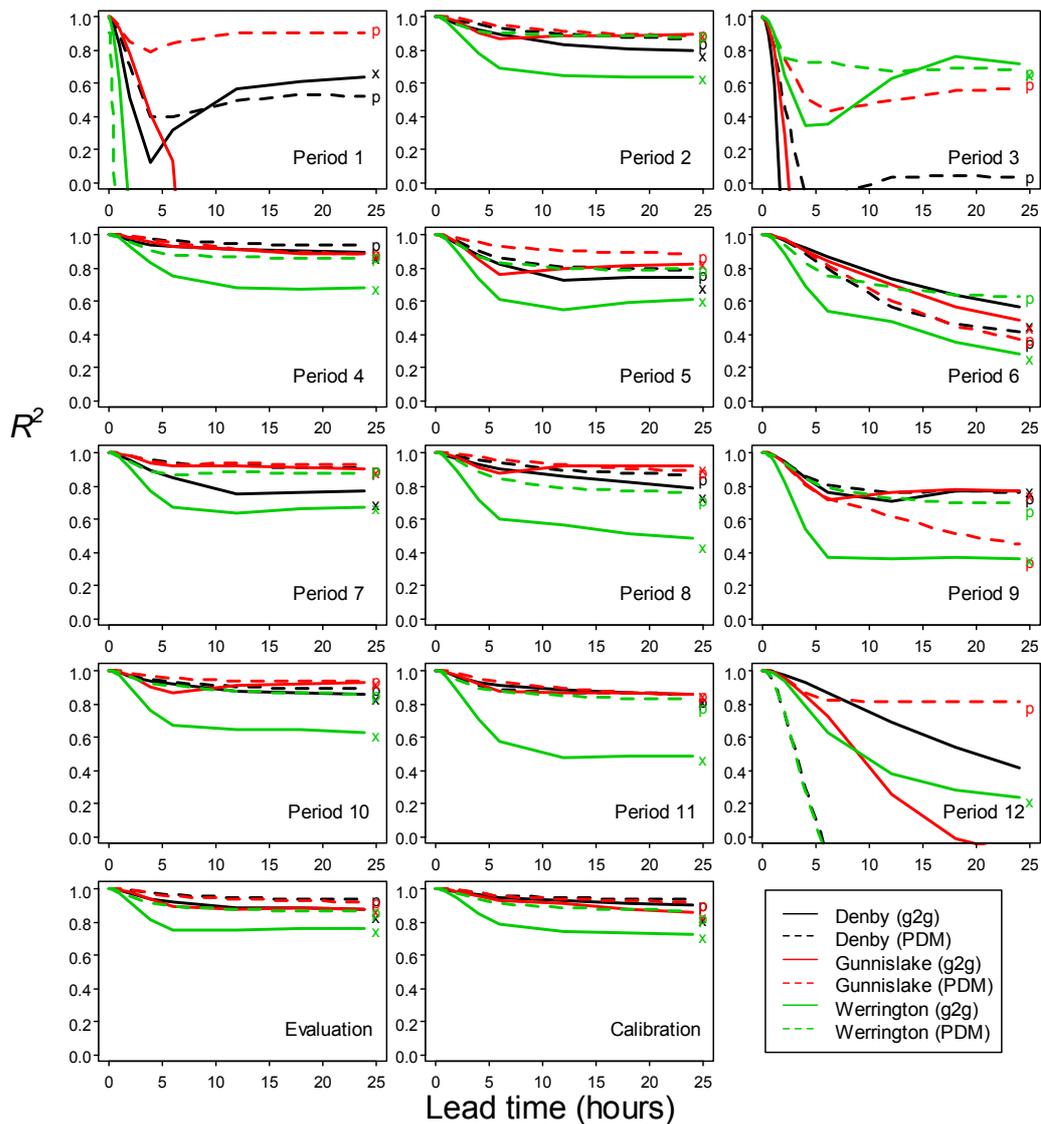


Figure 5.23 Model efficiency (R^2) versus lead time for periods 1–12 and the calibration and evaluation period for Gunnislake, Denby and Werrington.

Notes: Results for PDM and G2G models are shown. The symbols p and x refer to the simulation mode performance of the PDM and G2G models respectively.

5.5 G2G application

5.5.1 Model set-up

The only spatial dataset employed in configuring the G2G Model was the 50-metre IHDTM. This dataset was used in two ways:

- the average slope within each 1 km grid square was used within the runoff production scheme (see Section 4.2.1),
- flow paths were derived at a 1 km resolution using the COTAT+ method of Paz et al. (2006) and employed within the flow routing scheme (see Section 4.2.2).

Once the flow paths have been derived, cumulative upstream drainage areas can be calculated for each 1 km pixel and the appropriate grid square identified for each gauged location. Gridded inputs of HyradK raingauge-adjusted rainfall and MORECS potential evaporation (40 km resolution) were used directly by the G2G model.

At the outset, the aim was to have a single set of G2G model parameters that would be applied across the entire case study domain. However, initial analysis of the observed hydrographs in the region revealed a marked difference in behaviour between the faster responding catchments in the northeast of the case study domain (e.g. Tamar) and the more pronounced baseflow component of catchments in the southwest (e.g. Camel).

These differences in observed behaviour are not solely attributable to topographic controls and analysis of supporting datasets (see Section 5.1) revealed a north-east to south-west split in soil and geology characteristics. Thus, rather than use a single G2G model parameter set for the region, improved model simulations were obtained by splitting the region into two – respecting this soil/geology division. This division is discussed in more detail in Section 5.5.6.

The small number of regional model process parameters were calibrated manually in simulation mode for both the Tamar and Denby regions. Principally the main case study locations (Gunnislake, Werrington and Denby) were used in the model calibration, allowing the remaining gauged locations to be used to assess the ‘ungauged’ performance of the G2G model. The calibrated model parameters are given in Table 5.12. A standard form of empirical state correction was used for the forecast mode assessment.

5.5.2 Tamar catchment

The Tamar G2G model covered the north-east of the case study region and encompassed the following gauged locations:

- Bealsmill
- Bush
- Crowford Bridge
- Gunnislake
- Helebridge
- Lifton
- Polson Bridge
- Tinhay
- Werrington
- Woolstone.

Table 5.12 G2G model parameters for the Camel and Tamar catchments.

Parameter name	Catchment	
	Tamar (Werrington Park and Gunnislake)	Camel (Denby)
Wave speeds		
Surface land, c_l	0.05	0.04
Surface river, c_r	0.8	0.7
Subsurface land, c_{lb}	0.0015	0.001
Subsurface river, c_{rb}	0.004	0.003
Return flows		
Land, r_l	0.03	0.0012
River, r_r	0.05	0.002
Runoff generation		
c_{max} regional maximum	125	140
\bar{c}_{min} regional minimum	10	10
S_t	0	0
k_d	6.77×10^{-8}	8.3×10^{-8}
Land/river designation		
Accumulated area threshold, a_0	7	7
Routing time step (minutes)	5	5

During calibration, the main emphasis was on the Gunnislake and Werrington locations. Achieving a satisfactory manual calibration at both locations proved to be challenging, especially considering the strange behaviour of the observed flows noted at Werrington (see Section 5.4.3). However, using both gauging stations did help to calibrate the wave routing speeds of the model.

Performance over the entire calibration period is presented in Figure 5.19, which shows that the G2G model captured the long-term slow response of the catchment very well, with very good agreement on the recessions at Werrington and Gunnislake. Like the PDM model, the simulations reveal problems during the wetting up period after the dry 2006 summer. In general, the Tamar G2G model performs very well at Gunnislake but slightly worse at Werrington. The G2G model gives comparable results to the PDM model at Gunnislake but tends to be slightly too responsive to rainfall in the summer. At Werrington, the G2G model tends to underestimate the peaks – even more so than the PDM model. This is more apparent in Figure 5.24, which covers a shorter winter period.

Figure 5.24 also highlights the rather good G2G performance at other ‘ ungauged ’ sites in the region including sites outside but adjacent to the Tamar catchment (Woolstone and Helebridge). Unsurprisingly, the worst performance occurs at Bealsmill which is on the ‘ Denby ’ side of the soil/geology division discussed above and explains the higher baseflow component in the observations. Further model analysis is given in Section 5.5.4.

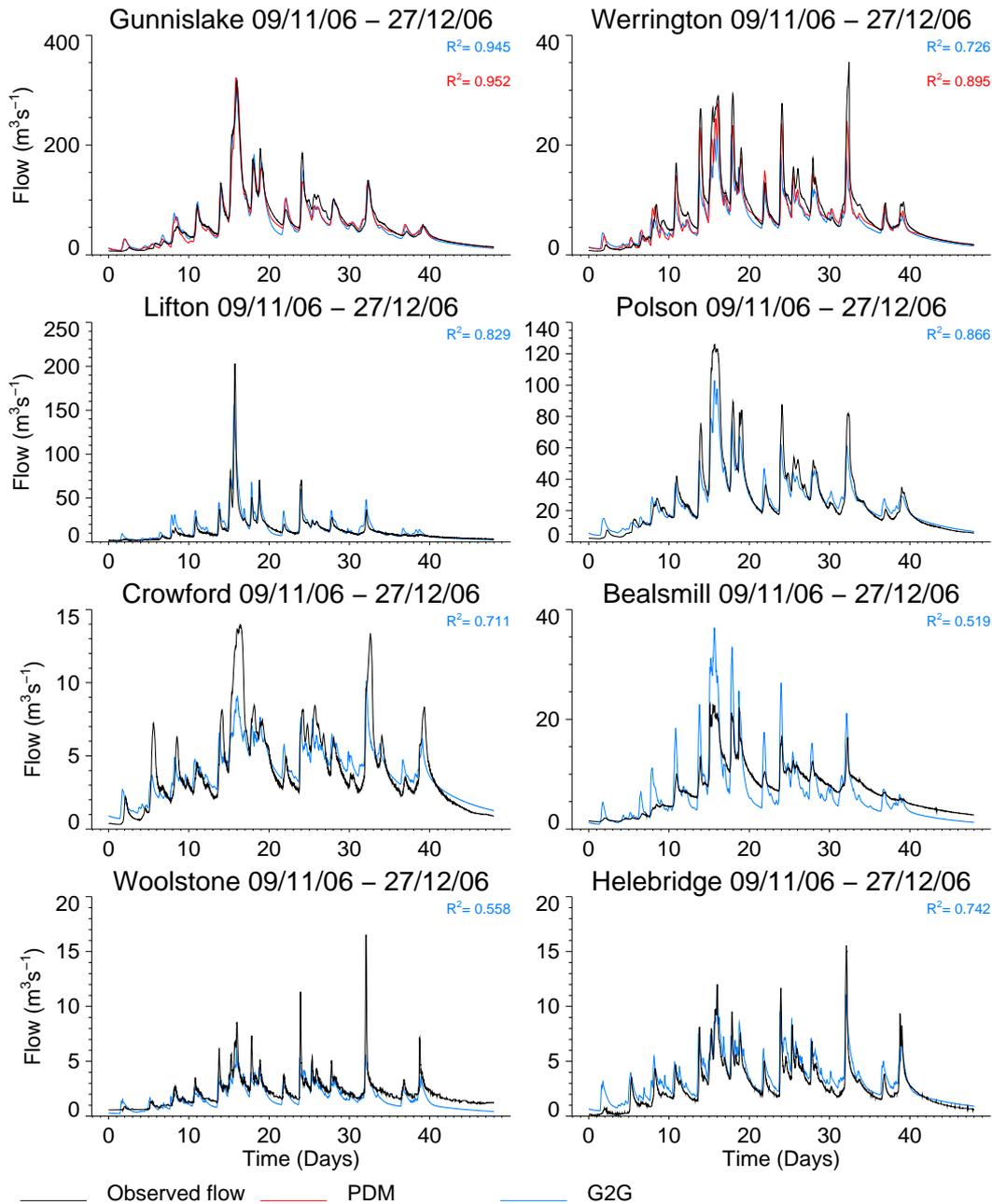


Figure 5.24 Modelled and observed hydrographs using the PDM and Tamar G2G models for the period 19 November to 27 December 2006.

5.5.3 Denby catchment

The Denby G2G model covered the south-west of the case study region and encompassed the following gauged locations:

- Bastreet
- Craigshill
- De Lank

- Denby
- Pillaton
- Restormel
- Slaughterbridge
- Trekeivesteps
- Trengoffe.

During calibration, the main emphasis was on the Denby location. Compared with the Tamar, the Denby G2G model required deeper soil storage, slower wave routing speeds, greater drainage to the subsurface stores and slower return flows.

Performance over the entire calibration period is presented in Figure 5.19, which shows that the G2G model captured the long-term slow response of the catchment reasonably well with just minor disagreement on the recession. Like the previous model results, the simulations reveal problems during the wetting up period after the dry 2006 summer. In general, the G2G model performs very well at Denby and, apart from the recession problems, is comparable with the PDM model results.

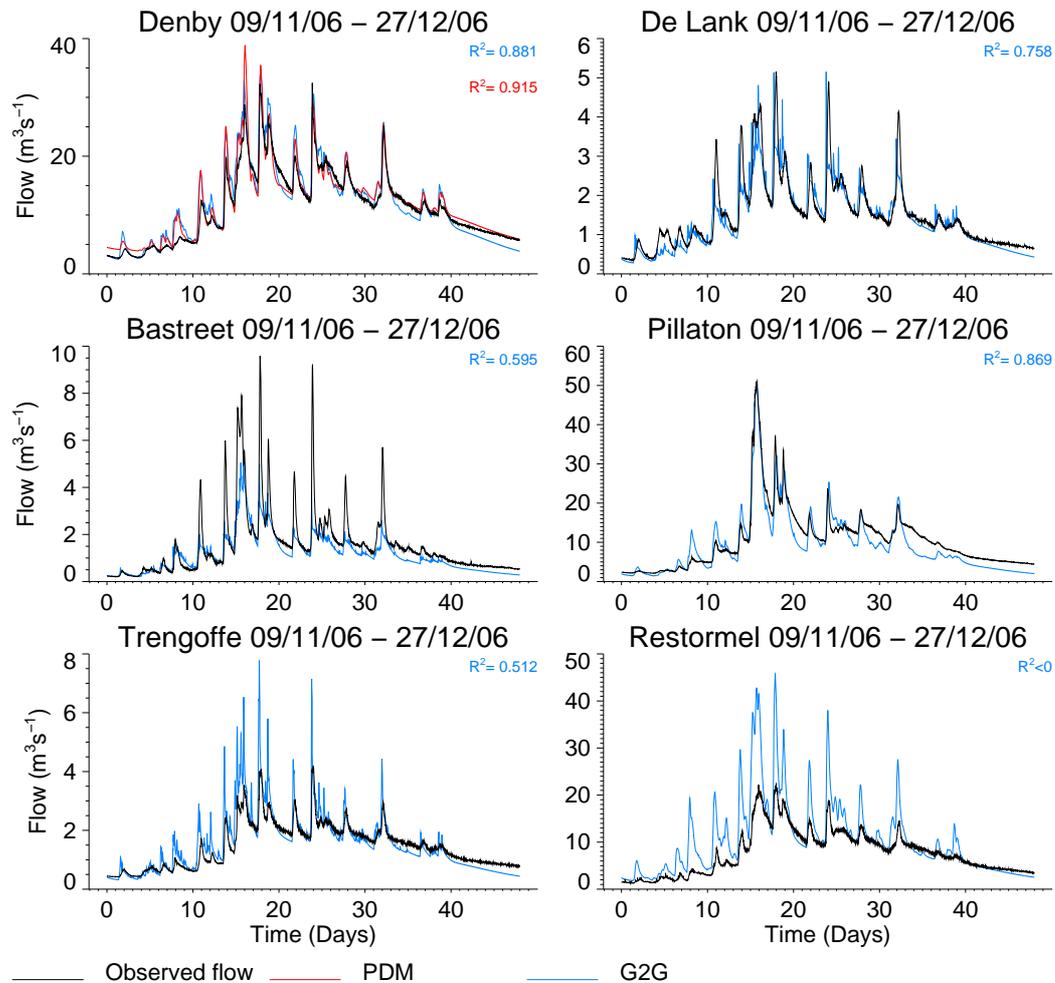


Figure 5.25 Modelled and observed hydrographs using the PDM and Camel G2G models for the period 19 November to 27 December 2006.

Figure 5.25 covers a shorter winter period and highlights the G2G performance at other ' ungauged ' sites in the region, including sites outside but adjacent to the Denby catchment. The performance at De Lank, an internal site, is good whereas the G2G tends to overestimate the peaks at Trengoffe and Restormel – though this is probably due to the reservoir influences which are not explicitly modelled here. Further model analysis is given in Section 5.5.4.

5.5.4 Model performance (simulation mode)

The simulation mode performance of the G2G model is generally good at all three of the main case study locations (Gunnislake, Werrington and Denby). The performance statistics of the G2G are presented in Table 5.5. The simulated flows are compared with observed flows over the calibration period in Figure 5.19 and the evaluation period in Figure 5.20.

The hydrograph simulations show that the G2G generally models the recession and seasonal behaviour of the catchments well, although there are minor disagreements on the recession for Denby. Otherwise the only exception is the 'wetting up' period following the dry summer in 2006, but this has been a problem for all models and catchments. The peak responses are also well modelled at Denby and Gunnislake. Modelling peak flows at Werrington was more challenging and marginally less successful, which is reflected in the slightly poorer performance statistics. The G2G model also tends to be marginally too sensitive to rain during summer periods.

The results provided by the PDM model proved difficult to better with the G2G model. However, this has to be balanced against the potential information at ' ungauged ' locations provided by the G2G model. Figure 5.26 shows graphically the model efficiency of G2G at all locations for which observed flow data were available. The PDM results are also given for locations where they are available (Gunnislake, Werrington and Denby).

The G2G model results show that it does provide some benefit at the ' ungauged ' locations. In particular, the results using the Tamar G2G show some real utility (e.g. Polson, Lifton, Tinhay and Bealsmill). Some of the poorer model performance can be attributed to uncertain flow measurements (e.g. Woolstone and Bush) or significant artificial influences such as reservoirs (e.g. Craigshill and Restormel). Encouragingly the G2G model performance is consistent across calibration and evaluation events.

The G2G simulation performance over the Boscastle event is presented in Figure 5.22. The plots show that the G2G model had difficulties simulating the peaks but, through the grid-based runoff production scheme, it appears to generate reasonably realistic amounts of localised surface runoff.

The main problem with the G2G results for the Boscastle event is that the simulated recession is too slow and prolonged. It is thought that this is due to the routing elements of the G2G model and, in particular, the slower wave speeds over land which are probably not suitable for this intense summer event.

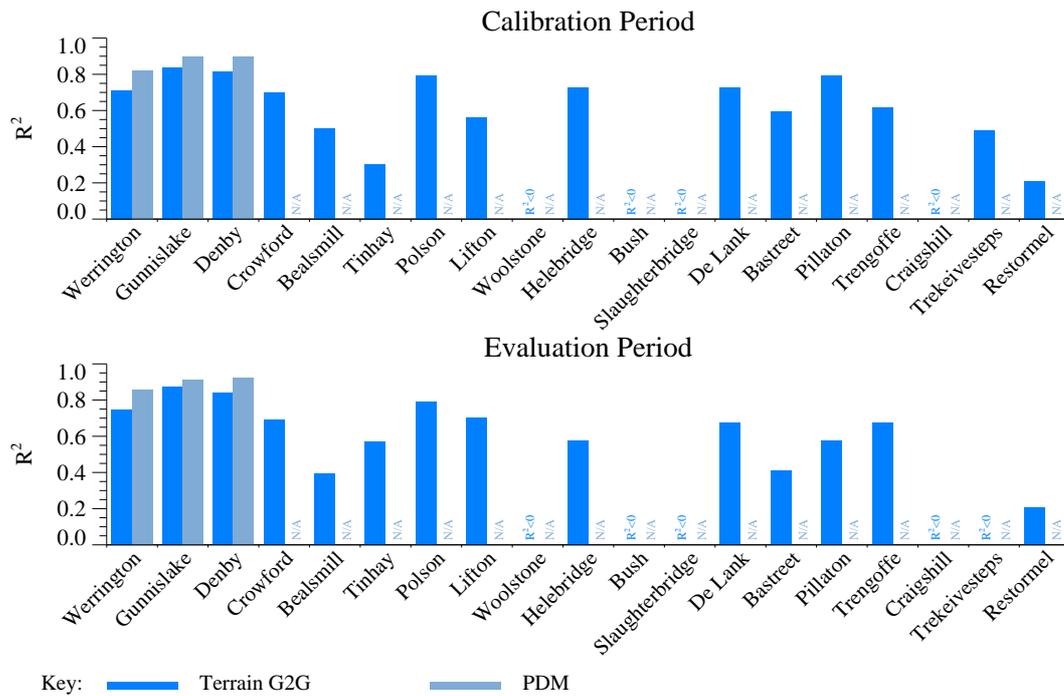


Figure 5.26 Model efficiency (R^2) for PDM and G2G over the calibration and evaluation periods.

Notes: Results for G2G at ‘ ungauged ’ locations are also given.

5.5.5 Model performance (forecasting mode)

The forecast mode parameters of the G2G model were set to standard values. HyradK raingauge-adjusted radar data were used as a ‘ perfect rainfall forecast ’. Forecasts were made for lead times out to 24 hours at every 15-minute time step within events. Empirical state correction, which utilises the observed flow data to correct the routing water of the G2G model, was applied up to each forecast origin.

Figure 5.23 presents model efficiency (R^2) versus lead time over the entire calibration and evaluation periods for the three main locations of Gunnislake, Denby and Werrington. Results are also shown for the 12 shorter periods listed in Table 5.7 and used within the REW analysis in Section 5.3.6.

The forecast performance of the G2G model was consistent at all three locations over both the calibration and evaluation periods, with the worst performance at Werrington. This is not surprising since the simulation mode results are also poorest at this location.

Figure 5.23 highlights the considerable performance improvement state correction which the G2G model offers for lead times out to around 10–12 hours for Denby and Gunnislake and 6–8 hours for Werrington. Beyond these lead times, the forecast mode performance tails off to the simulation mode performance as expected. The analysis over the shorter periods (1–12) shows similar results to the calibration and evaluation periods for the longer winter time events (2, 4–8, 10, 11). Forecast performance over the summer events (1, 3, 9, 12) is much more variable. Generally the PDM model forecasts perform best, reflecting (i) the better simulation provided by the PDM model at gauged locations and (ii) the fact that, due to time constraints, G2G forecast parameters were not optimised whereas PDM parameters were. Further research on state correction for the G2G model is ongoing.

5.5.6 Extended G2G model results

The geological map over the G2G modelled region shown in Figure 5.2(a) highlights the contrast between the Carboniferous rocks in the north-east half and the Devonian Old Red Sandstone with granitic igneous intrusions (e.g. Bodmin Moor) in the south-west half. This contrast is reflected in the HOST soil classes shown in Figure 5.2(b), with classes 24 (blue: shallow soils) and 21 (red: medium depth slowly permeable substrate) dominant in the north-east and 17 (green: deep soils over an impermeable layer) and 15 (blue: peat) in the south-west.

These contrasting soil/geology patterns exert a control on flood response to storm rainfall as seen through the G2G modelling across the region. Catchments to the south-west generally have a slower response on account of the deeper or peat soils. Rather than use a single G2G model parameter set for the region, improved model simulations have been obtained by splitting the region into two, respecting this soil/geology division. This provided a pragmatic approach.

A major reason for this need to subdivide the region relates to the inference of the soil's water holding capacity, which controls runoff production in the G2G model. In the simplest formulation used here, terrain slope is used as a surrogate for the capacity of a soil to absorb water through a linear relation. In simple terms, thin soils are associated with steep slopes and deeper soils with flatter areas. Such a relationship can break down across regions of contrasting soil types as evident in the case study region.

Splitting the region into the two contrasting areas and calibrating the G2G model for each has provided a simple way of improving modelled flows. However, use of a G2G model extended to incorporate soil/geology property information and requiring only one model parameter set is seen as a more strategic way forward. Such a model has been prototyped by CEH for the Environment Agency under the R&D project 'Rainfall-runoff and other modelling for ungauged/low-benefit locations' (Moore et al. 2007) and developed further since then. Although this variant of the G2G model is not yet available in module adapter form for real-time use in the NFFS, it was available for use in the project as a research tool to explore the potential advantages of the extended G2G model. The application of this model variant to the study region is discussed below.

HOST classes on a 1 km grid linked to soil properties through an association table currently provide the source of soil information. The soil properties are:

- soil depth;
- porosity;
- field capacity and residual values of water content,
- saturated hydraulic conductivity.

Soil depth and water content at field capacity are used to estimate the maximum storage capacity in each model grid square. This capacity is distributed within the grid square according to a Pareto distribution (like in the catchment PDM) but with the shape parameter, b , varying inversely with the square root of the maximum storage capacity. Downward percolation of water is controlled by available water in storage and the hydraulic conductivity of the soil. Lateral drainage is controlled in a similar way, but including the influence of terrain slope and using a conductivity appropriate to lateral movement. Groundwater accumulates by percolation from the soil and drains via a non-linear storage function with a rate constant parameter.

Within each grid square, surface runoff is generated via saturation excess flow and subsurface runoff derives from groundwater drainage. These runoffs are routed from grid-to-grid as with the simple G2G using the flow paths identified from the DTM. However, a

modified kinematic routing scheme is used for the surface runoff. This takes the form of a Horton–Izzard equation (non-linear storage routing) and can accommodate varying channel width (inferred via geomorphological relations) and roughness. A return flow is allowed between subsurface and surface pathways, as in the simple G2G model, to accommodate surface–groundwater interactions. Further details of the extended G2G model will be presented in the report for Phase 3.

Figure 5.27 compares the flow simulations from the G2G model with the extended model incorporating soil property information across eight catchments over 38 days of the two-year evaluation period.

In terms of R^2 efficiency, neither model performs better overall. A similar conclusion follows from an inspection of the hydrograph flood peaks and the models' abilities to reproduce them. This is encouraging given that the extended G2G model employs only a single parameter set for all catchments; the simple G2G model employs two to cope with the soil/geology heterogeneity across the modelled region. The recession behaviour for the Lynher at Pillaton Mill is improved using the extended G2G model but does not perform as well on the main peak. This contrasts with the De Lank in the Camel catchment where the recession behaviour is poorer whilst flood peak performance is similar. Flow simulations for other catchments can be compared leading to similarly contrasting, rather than consistent, conclusions.

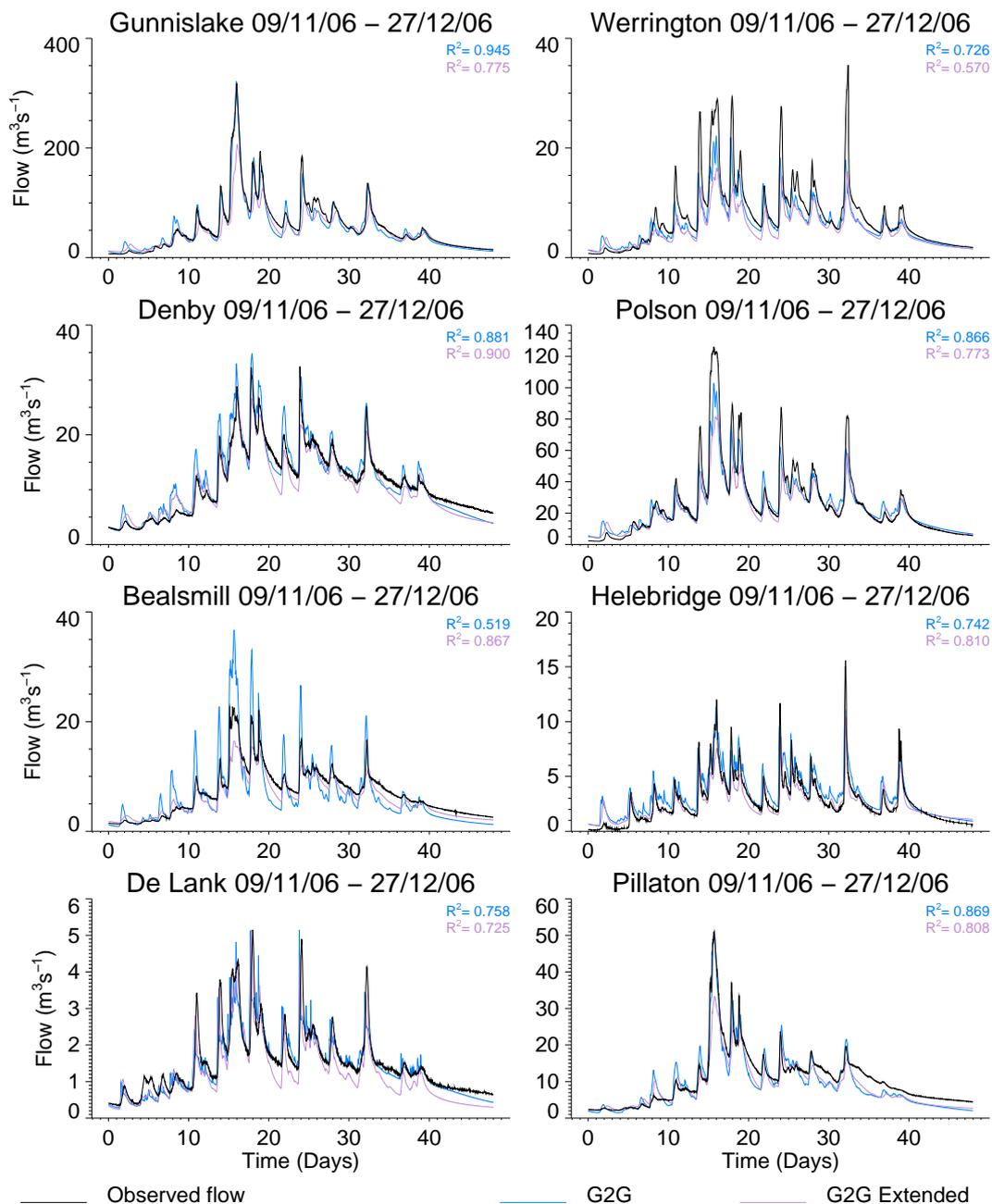


Figure 5.27 Modelled and observed hydrographs using G2G and extended G2G for the period 19 November to 27 December 2006.

Figure 5.28 compares the G2G and extended G2G model performance in terms of R^2 efficiency, plotted as bar charts, across all gauged catchments in the modelled region. The performance measure is plotted separately for calibration (~18 months) and evaluation (~24 months) periods. For a good majority of cases and for both periods, the extended G2G model is seen to perform best over these longer assessment periods.

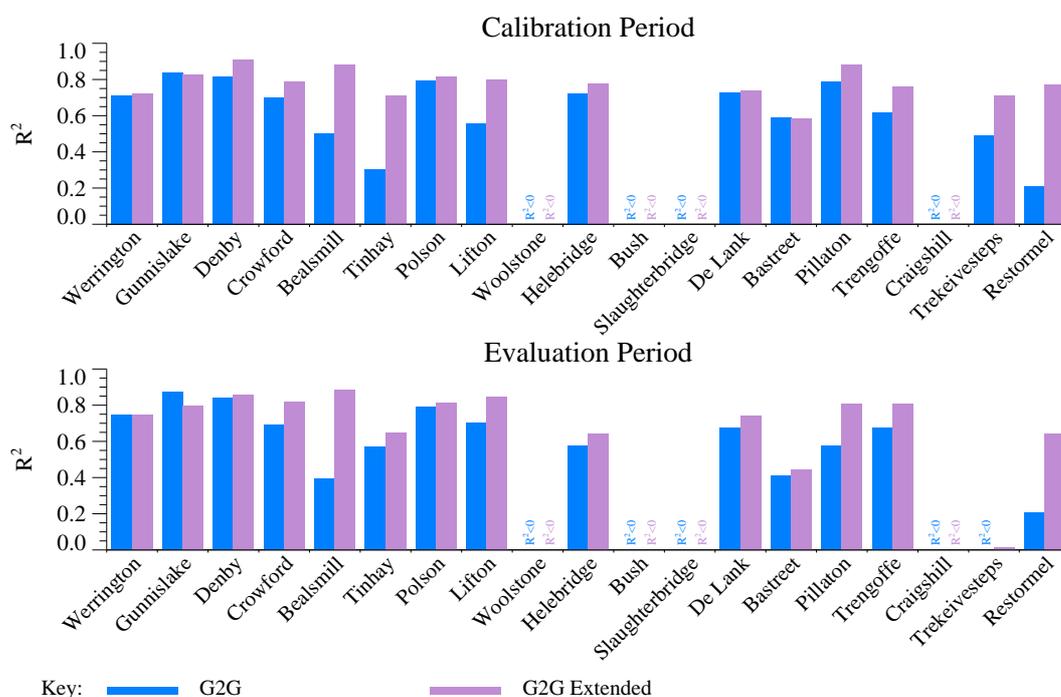


Figure 5.28 Model efficiency (R^2) for G2G and extended G2G over the calibration and evaluation periods.

Figure 5.29 compares the flood hydrographs relating to the Boscastle storm for three models – the lumped PDM and the two forms of distributed G2G model – for four of the catchments.

Excellent simulations are obtained for the Ottery at Werrington Park using the extended G2G and PDM models whilst the G2G model is much poorer. The extended G2G model also performs better for the Tamar at Polson Bridge (there is no PDM model for this catchment for comparison). All models perform consistently badly in simulating the sharp peak for the Camel at Denby. A rather mixed and generally unsatisfactory performance is obtained for the Tamar at Gunnislake from all models. It is likely that the delayed response seen in the extended G2G model simulations could be improved by using a shorter routing time step. The signature of this model response seems most amenable to obtaining a realistic simulation through further work.

Overall, the results obtained using the extended G2G – incorporating soil property information and a single model parameter set over the whole region – are very encouraging. They point to the great value of further work using the extended G2G in Phase 3, including the potential of its application throughout England and Wales.

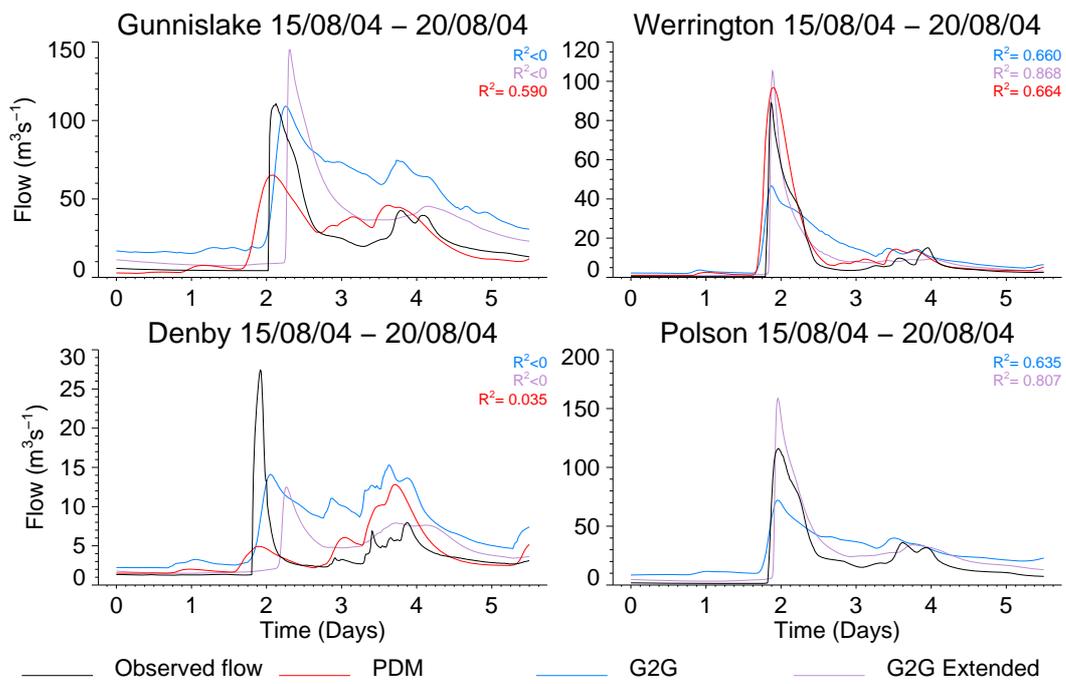


Figure 5.29 Modelled and observed hydrographs using the PDM, G2G and extended G2G models for the period 15 to 20 August 2004.

6 Analysis of use of high-resolution NWP forecasts

6.1 High resolution NWP forecasts

At the time of the Boscastle event (2004), the operational deterministic NWP had a resolution of 12 km. Since then, the operational NWP resolution has increased to 4 km with 1 km planned in the near future.

To assess the potential benefits of high resolution NWP rainfall for flood forecasting, 1, 4 and 12 km resolution NWP forecasts were provided for the Boscastle event by Nigel Roberts at Met Office JCMM. Forecasts from two origins were provided: 00UTC⁸ and 03UTC on 16 August 2004. Figure 6.1 gives example snapshots of the forecast rainfall data. Due to the method used to generate the high-resolution forecasts, the forecast rainfalls start at 01UTC and 04UTC respectively. More detail about the NWP forecasts is provided by Roberts (2006).

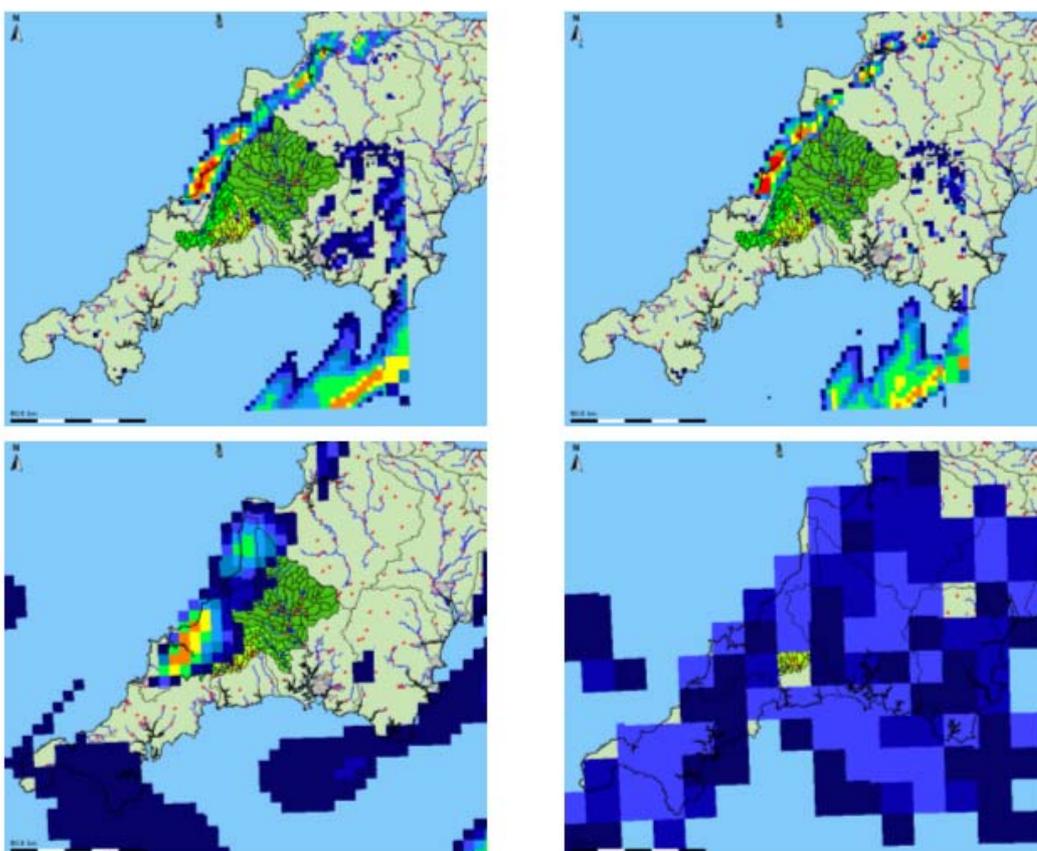


Figure 6.1 Forecast rainfall data from the high resolution NWP model as displayed in Delft-FEWS.

Notes: The panels (from top left to bottom right) show the raingauge-adjusted radar image, the radar image, the 4 km forecast and the 12 km forecast. The time for all panels is 12:00 16 August 2004 and the forecast origin for the forecast images is 01:00.

⁸ Co-ordinated Universal Time

6.2 Generation of pseudo ensembles

Ensembles of high-resolution rainfall forecasts are due to become available in the very near future. The nowcasting product STEPS available out to six hours is due to be made operational by the end of 2008 and ensembles of high resolution NWP are planned by 2010.

In preparation for the imminent availability of these high-resolution ensemble rainfall forecasts and as they were not available for the Boscastle case study, it was decided to generate 'pseudo' ensembles from the deterministic 1 km NWP forecast. The method developed for generating the pseudo ensembles closely involved Nigel Roberts from JCMM. To remain within the scope of the project, the intention was to derive a simple method that captures some of the spatial uncertainty associated with the deterministic NWP forecast. In its current form the method is **not** intended to be used immediately as a 'post-processor' of NWP rainfalls in an operational context.

The pseudo ensemble method developed consists of two stages:

1. Selecting a scaling factor to apply to each NWP forecast.
2. Generating ensemble members by randomly displacing the spatial origin within a given displacement radius.

The following sub-sections give more details of the method. The import and processing of pseudo ensembles with Delft-FEWS is outlined in Appendix E.

6.2.1 Selecting the scaling factor

Table 6.1 gives a summary of the 1 km NWP rainfall accumulations for the five-hour period ending 17:00 16 August 2004 from both the 00UTC and 03UTC forecasts. These are compared to 2 km single site Nimrod data from Cobbacombe, the Nimrod composite product (essentially 2 km data from the Predannack radar over the Boscastle catchment) and the raingauge-adjusted Nimrod composite data produced by HyradK (see Section 5.2.3). These results show that the 1 km NWP appears to underestimate the spatial peak rainfall accumulations and catchment average rainfall relative to the radar data. Furthermore, radar pixels coincident with the raingauge locations underestimated the raingauge totals and hence use of the HyradK raingauge-adjusted radar data has increased the Nimrod rain rates, giving even larger totals.

Table 6.1 Summary of different rainfall estimators accumulated over the five-hour period ending 17:00 16 August 2004 ¹.

Rainfall estimator	Peak pixel accumulation (mm)	Peak pixel location	Boscastle catchment average (mm)	Domain average (mm)
1 km NWP 00UTC	54.22 53.24	SX 005 835 SX 145 875	34.30	3.07
1 km NWP 03UTC	44.03	SX 125 985	16.77	2.54
Cobbacombe 2 km Nimrod	133.1	SX 150 890	92.66	N/A
Nimrod composite	115.68	SX 170 930	93.06	2.618
Raingauge-adjusted Nimrod composite	213.11	SX 125 905	170.0	3.996

Notes ¹ The domain used is (140000,000000) to (280000,140000).

At first glance, the figures in Table 6.1 would suggest a significant scaling factor (>2) needs to be applied to each NWP forecast to give rainfall accumulations (over Boscastle) commensurate with the radar and raingauge data.

However, analysis of the spatial accumulation maps presented in Figure 6.2 reveals that there are differences in the spatial distribution over the Boscastle region. The lower row shows the existence of three areas of high rainfall from the 00UTC NWP forecast compared with a single concentrated area from the radar-based accumulations. By stepping through the individual NWP images, this can be traced to individual convective cells having slightly different trajectories that meant the core of those cells just missed the Boscastle catchment. In comparison, the radar images reveal an almost constant trajectory of the cells over the Boscastle catchment and gives rise to the higher pixel rainfall accumulations.

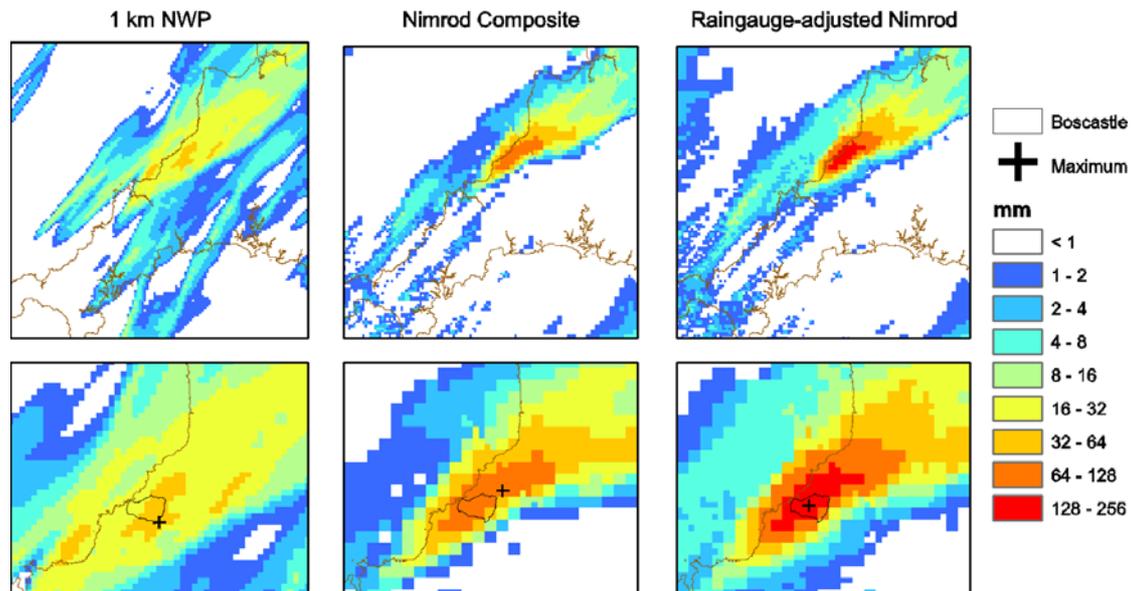


Figure 6.2 Rainfall accumulations (mm) for the five-hour period ending 17:00 16 August 2004 using different rainfall sources.

Notes: The 00UTC 1 km NWP forecast is used. The bottom row is a close up over Boscastle.

It is more appropriate to assess the 1 km NWP at a larger spatial scale. Figure 6.3 shows the different rainfall sources averaged over 12 km pixels using the 00UTC forecast. This shows that the general location of the heavy rain across the north-west Cornish coast is well predicted by NWP and in good agreement with the raingauge-adjusted Nimrod estimates apart for the amounts in the immediate vicinity of Boscastle, which are 2–4 times smaller. However, applying a blanket factor of two or more to the NWP data would distort the results away from Boscastle and give too high a domain average rainfall (see Table 6.1). As a compromise, a factor of 1.4 was used for the 00UTC forecast and 1.7 for the 03UTC forecast. The resulting rainfall accumulations for the scaled 00UTC forecast are presented in the right-hand element of Figure 6.3. Although this would not be possible in a real-time context, it provides a pragmatic way of creating a useful set of ensembles that reflects the high resolution NWP products which will be available in the near future.

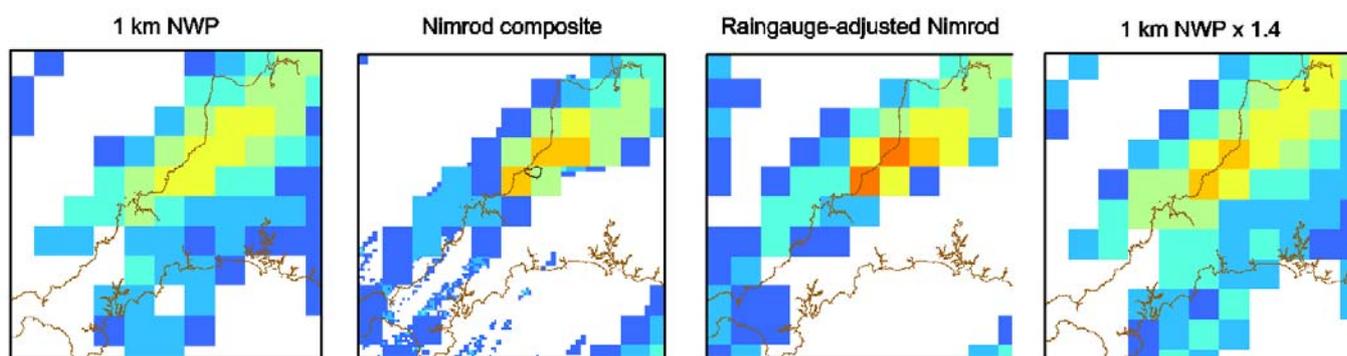


Figure 6.3 Rainfall accumulations (mm) for the five-hour period ending 17:00 16 August 2004 using different rainfall sources and accumulated over 12 km grid squares.

Notes: The 00UTC NWP forecast is used.

6.2.2 Ensemble generation

Following discussions with Nigel Roberts (JCMM), an appropriate way of generating the ensembles (within the constraints of the project) is to displace the spatial origin of the (scaled) 1 km forecast but maintain the temporal evolution. This will result in forecast rainfall accumulation maps with the same spatial distribution and totals, but shifted in space. For this particular storm, the appropriate maximum displacement is 20 km and is a representation of the perceived spatial accuracy of the forecast in this meteorological situation.

The argument is that any forecast displaced within the 20 km radius is equally likely to have occurred in reality. Any number of ensembles can be generated by randomly selecting the spatial displacement of the forecast within the 20 km radius. For simplicity, the selection was restricted to displacement units of 1 km in both northing and easting, and a maximum of 50 ensemble members. Figure 6.4 presents an example of three ensemble members along with the deterministic 1 km NWP forecast.

6.3 Analysis of hydrological model forecasts using HyradK data

Before assessing the hydrological model forecasts using both the deterministic and pseudo ensemble forms of high resolution NWP, it is informative to analyse the forecast mode

performance of the hydrological models using the HyradK raingauge-adjusted radar estimates as ‘perfect rainfall’ forecasts. While this is done at a broader level in Section 5, more detailed analysis of the Boscastle event reveals certain model behaviour that must be taken into account before considering NWP-based forecasts.

The lead time versus model efficiency plots using the HyradK data are presented in Section 5 (Figures 5.16–5.18 for REW and Figure 5.26 for PDM and G2G) and show more sensitivity in the summer periods (1, 3, 9 and 12). The Boscastle event is period 3. The performance of the state correction (PDM, G2G) or ARMA error prediction (REW) is heavily influenced by the simulation mode performance of the models (presented in Figure 5.15 for REW and Figure 5.22 for PDM and G2G).

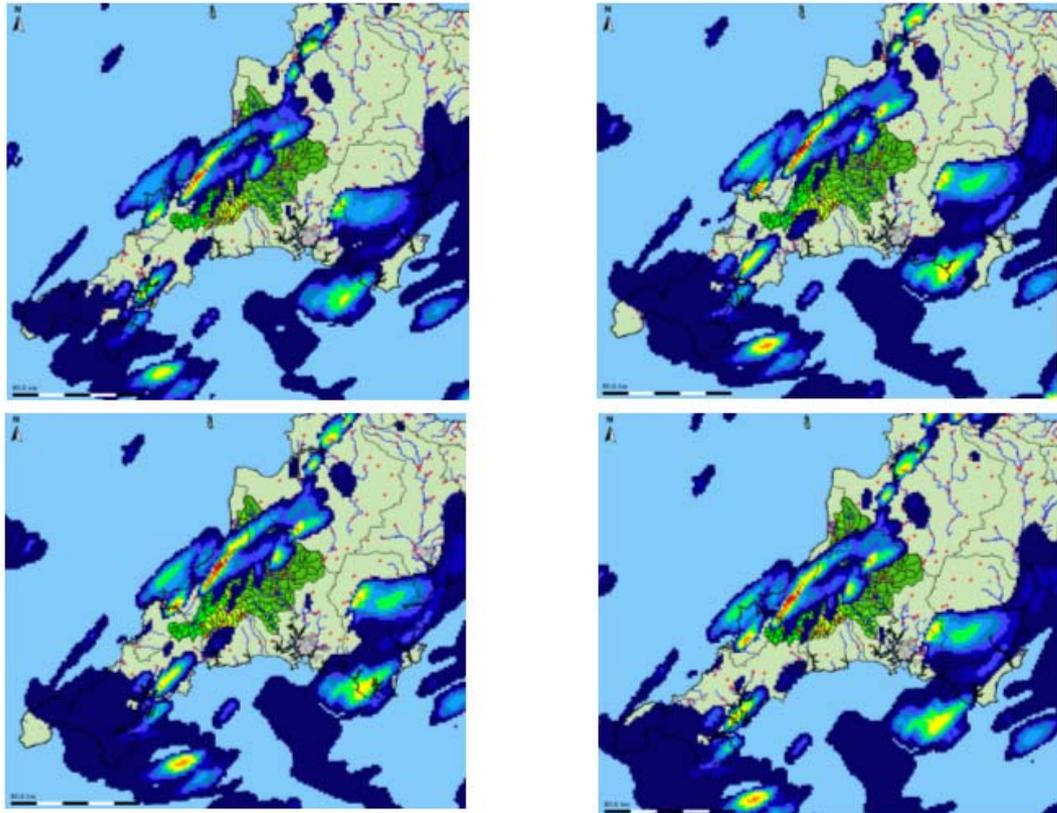


Figure 6.4 Forecast data from the high resolution NWP model as displayed in Delft-FEWS.

Notes The panels (from top left to bottom right) show the 1 km deterministic forecast and the ensemble members 01, 02 and 03. The time for all panels is 12:00 16 August 2008 and the forecast origin is 01:00.

To understand this further, Figure 6.5 presents the simulation mode results using the PDM model and also a sequence of fixed origin forecasts using state correction. The simulation results (left column) show a consistent trend with the PDM model providing a good simulation at the start of the period but then responding before the sharp rising limbs of the observed hydrographs. This means that the fixed origin forecasts made:

- *before* the PDM model responds (red line, right column) are basically very close to the simulation results;
- *after* the PDM model has responded but *before* the observations have started to rise are damped down through the state correction (e.g. blue line, right column).

The performance of the fixed origin forecasts made once the observations start to rise and after the peak vary from catchment to catchment.

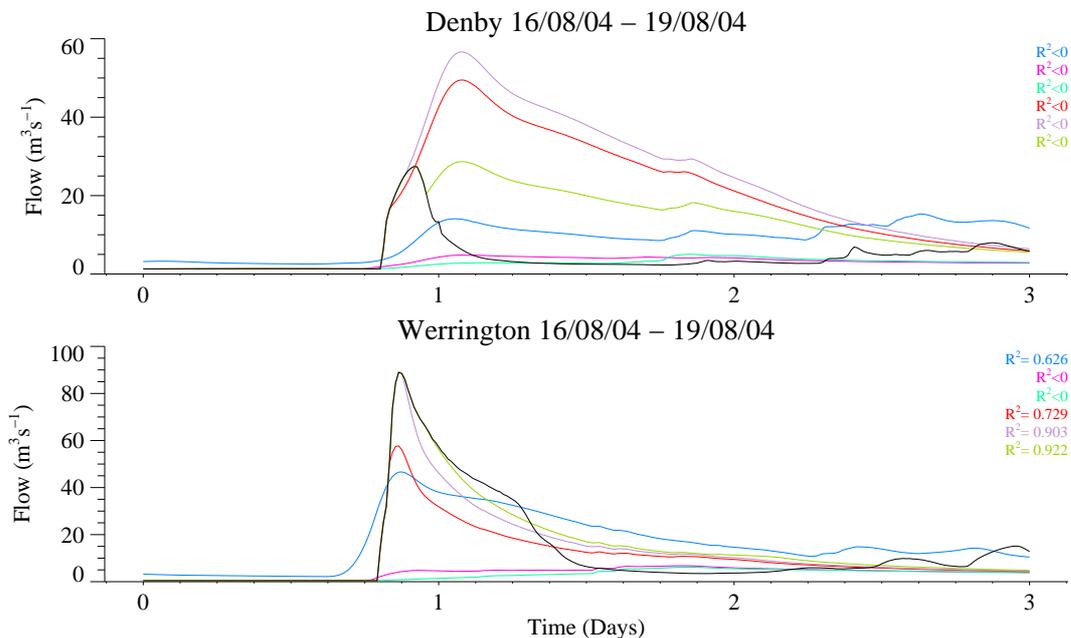


Figure 6.5 PDM simulation mode results using HyradK raingauge-adjusted rainfall data (left column) and fixed origin forecast results using state correction and HyradK ‘perfect’ rainfall forecasts (right column).

Fixed origin forecasts are presented for the G2G model in Figure 6.6 for Denby and Gunnislake, along with the simulation mode results (blue line). These show some similarities to the PDM results with the damping of the forecast made from origins *after* the G2G model has responded but *before* the observations have started to rise (pink and green lines). The fixed origin forecasts once the observations have started to rise vary from location to location. The results at Werrington are quite good; the poorer performance at Denby is attributable to the timing error in the simulation mode results.

For the REW-based forecasts, an ARMA error prediction model was used as described in Section 5.3.6. The raw REW simulation is poor and the sharp rise of the hydrograph is not captured by the model (see Figure 5.15). As such, the ARMA model can only improve the forecast once the hydrograph starts to rise. Fixed lead time forecasts are presented in Figure 6.7 for the REW at Denby and reveal oscillatory behaviour in the forecasts, particularly before and during the rising limb.

In summary, both state correction and ARMA error prediction approaches are very sensitive to any poor modelling of the rising limb of the Boscastle event – particularly if the model responds before the observations. This explains the sensitivity reported in the lead time versus model efficiency plots over the summer periods (see Section 5). These issues must be considered when assessing the high resolution NWP-based hydrological forecasts in the following sections.

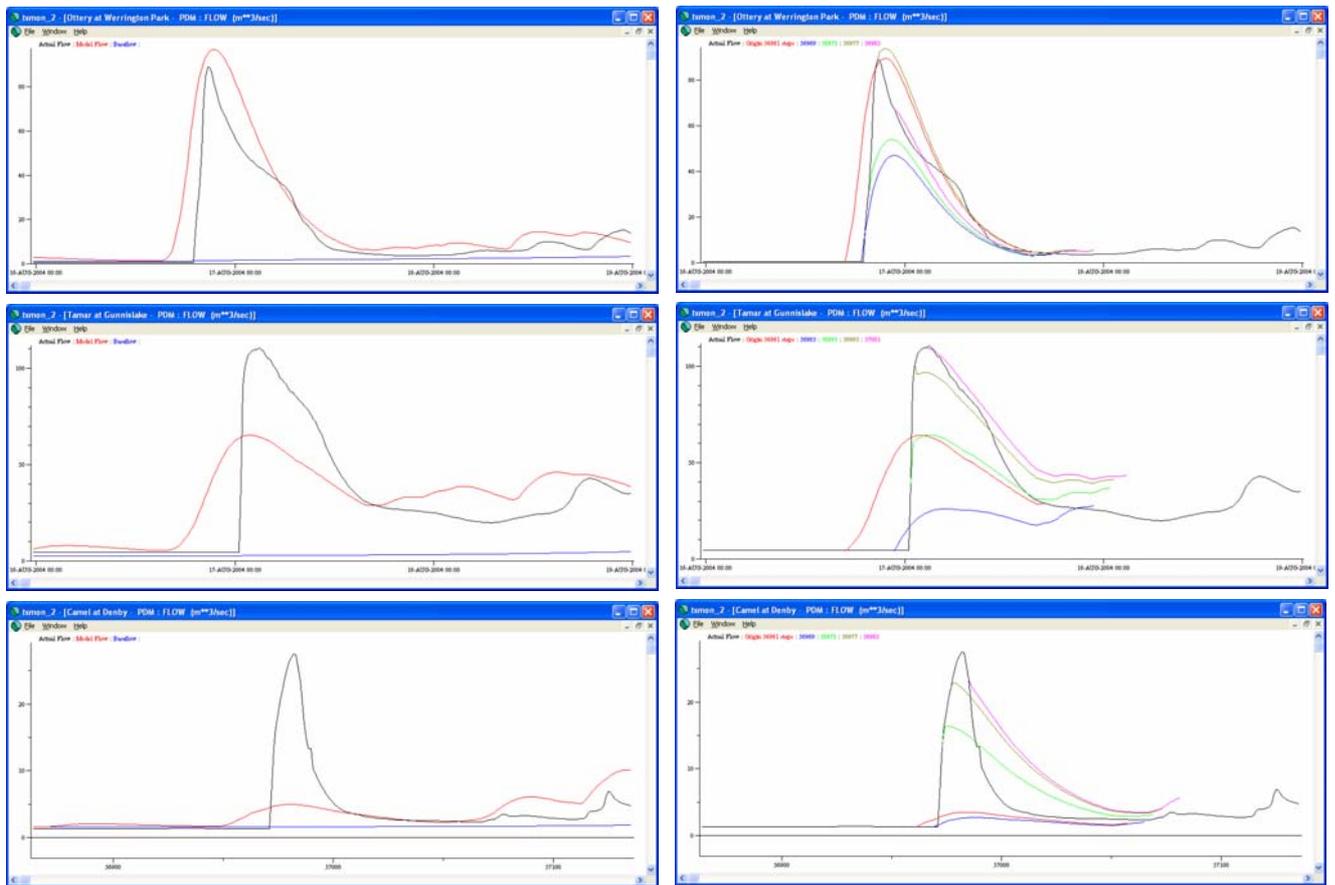


Figure 6.6 G2G fixed origin forecast results using state correction and HyradK raingauge-adjusted radar data as 'perfect' rainfall forecasts.

Notes The blue line is the simulation mode result.

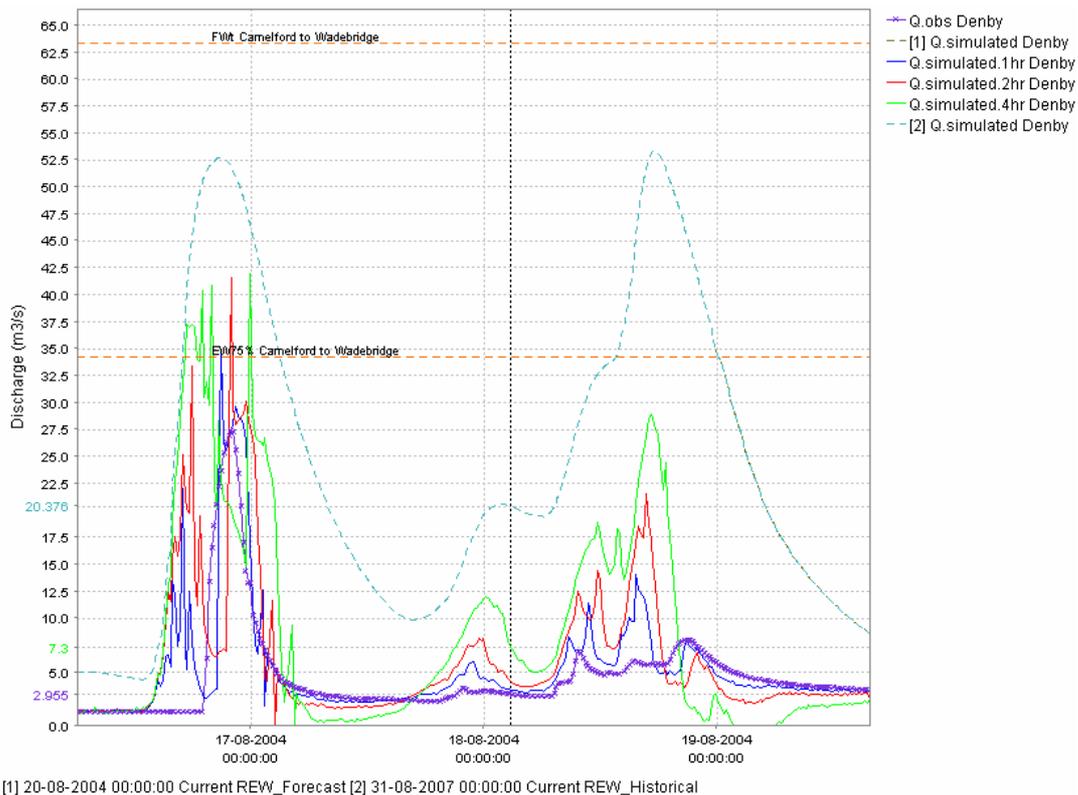


Figure 6.7 REW fixed lead time forecasts using HyradK-adjusted precipitation for the Boscastle event for Denby.

Notes: Blue = 1 hour, red = 2 hours and green = 4 hours.

6.4 Analysis of hydrological model forecasts using deterministic high resolution NWP rainfalls

NWP rainfall forecasts were provided at resolutions of 1, 4 and 12 km and for two forecast origins: 00UTC and 03UTC on 16 August 2008. Due to the methods used to generate the high-resolution (1 and 4 km) NWP model results, the rainfall forecasts start at 01 and 04UTC respectively. The high resolution NWP models were run specifically for the Boscastle storm so the end time of all the forecasts is 18UTC.

Since the NWP forecasts finish *before* the observed flows begin to rise, the NWP-based hydrological forecasts in the period *after* the model responds but *before* the observed flows rise will suffer the problems identified using the HyradK data in Section 6.3. This, combined with the fact that there are only two NWP forecast origins, means that plots of lead time performance against model efficiency are not very informative.

Bearing in mind the comments above, the most appropriate way to assess hydrological model performance using the different NWP resolutions is to run fixed origin forecasts using the different NWP resolutions and use the HyradK raingauge-adjusted radar as 'perfect' rainfall forecasts for comparison. The first three hours of each NWP forecast have been ignored as this represents the estimated time needed to generate and disseminate the forecasts in an operational context; this makes no real difference for the Boscastle case study as the rainfall did not start until around 12UTC.

Fixed origin forecasts using the PDM model and both the 00 and 03UTC NWP runs are presented in Figure 6.8 with hydrological forecast origins of 04 and 07UTC respectively (i.e. before the observations start to rise so as to avoid the ‘damping’ effect discussed above). This immediately reveals that:

- the hydrological forecasts based on high resolution NWP (1 or 4 km) generally perform better than ones based on 12 km NWP;
- the 00UTC NWP runs provide better hydrological forecasts than the 03UTC runs – this is in keeping with the analysis from a rainfall perspective by Roberts (2006).

The most conclusive evidence comes from the Werrington PDM model, which had the best simulation results and allowed a more direct analysis of the NWP-based hydrological forecasts without being confounded by shortcomings in hydrological model performance.

A more complete understanding of the hydrological forecasts comes from looking at the spatial maps of rainfall accumulations (Figure 6.9) and analysis of the catchment average rainfalls (Table 6.2) over the Boscastle storm. For example, Figure 6.8 shows that for 03UTC the 12 km NWP forecasts give the best NWP based hydrological forecasts at Denby; however, Figure 6.9 illustrates that 12 km hydrological results are best for the wrong reasons since the spatial distribution of the forecast rainfall is completely wrong.

Fixed origin forecasts using the G2G model are presented for Werrington and Denby in Figures 6.10 and 6.11 respectively. These show broad similarities with the PDM results, e.g. the high-resolution (1 or 4 km) NWP results generally perform best. However, there are some interesting differences. For example, the Werrington G2G results (Figure 6.10) show more sensitivity in terms of the timing of the forecast peak with the 00UTC NWP-based forecast having an earlier peak compared with the 03UTC one. This was not the case for the PDM model where the peak timings were similar.

The sensitivity of the distributed G2G model relates to the spatial distribution of the forecast rainfall; Figure 6.9 shows that, for Werrington, the 1 and 4 km 00UTC accumulations have too much rain in the lower parts of the catchment which, through the distributed runoff and routing formulation of the G2G, gives the earlier model response. This highlights the importance of accurate spatial distributions of NWP rainfall for producing accurate flood forecasts.

In summary, the high-resolution (1 or 4 km) NWP rainfall forecasts have been shown to provide real benefits for hydrological forecasting compared with the coarse 12 km NWP forecasts.

00UTC NWP forecasts

03UTC NWP forecasts

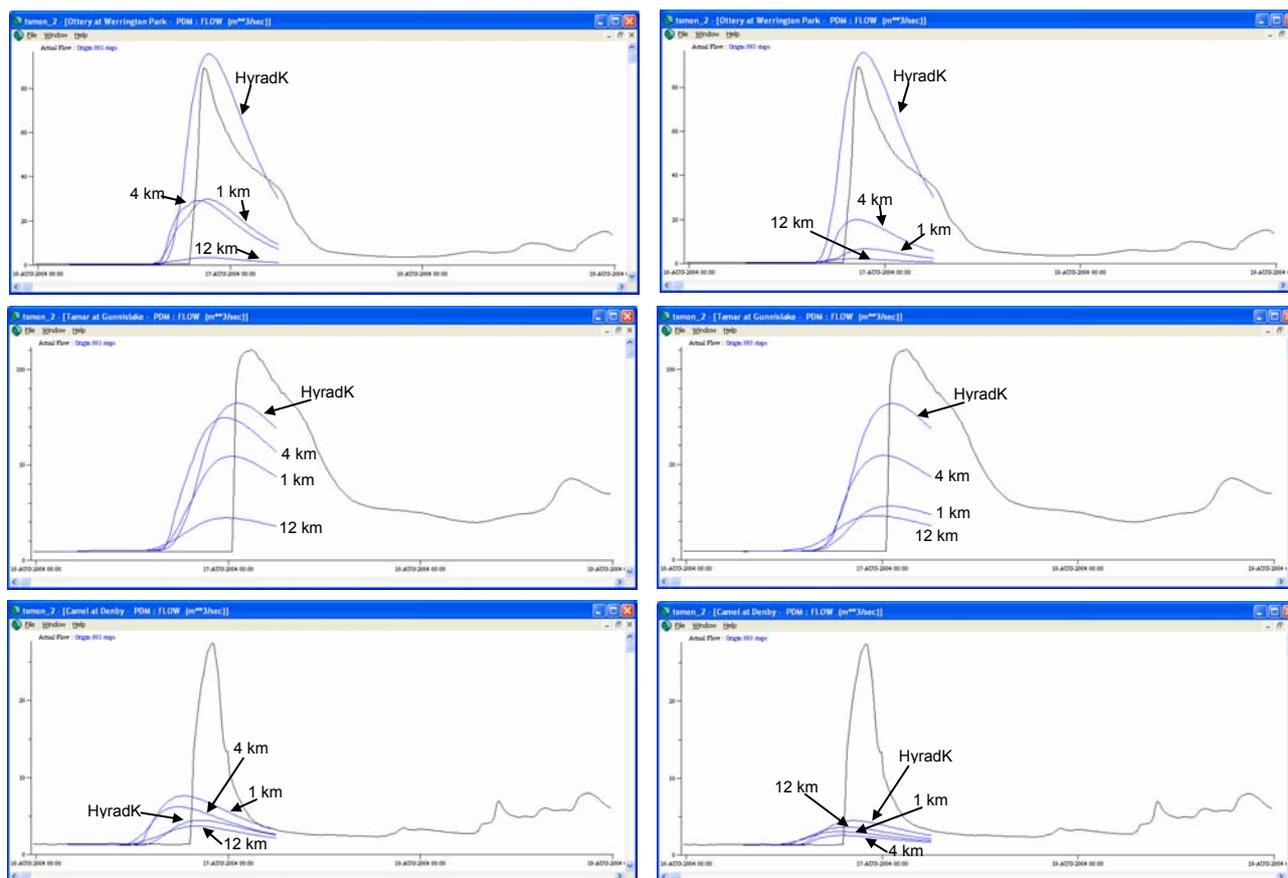


Figure 6.8 Fixed origin PDM model forecasts at Werrington (top row), Gunnislake (middle row) and Denby (bottom row) using 1, 4 and 12 km deterministic NWP and HyradK ‘perfect’ rainfall data.

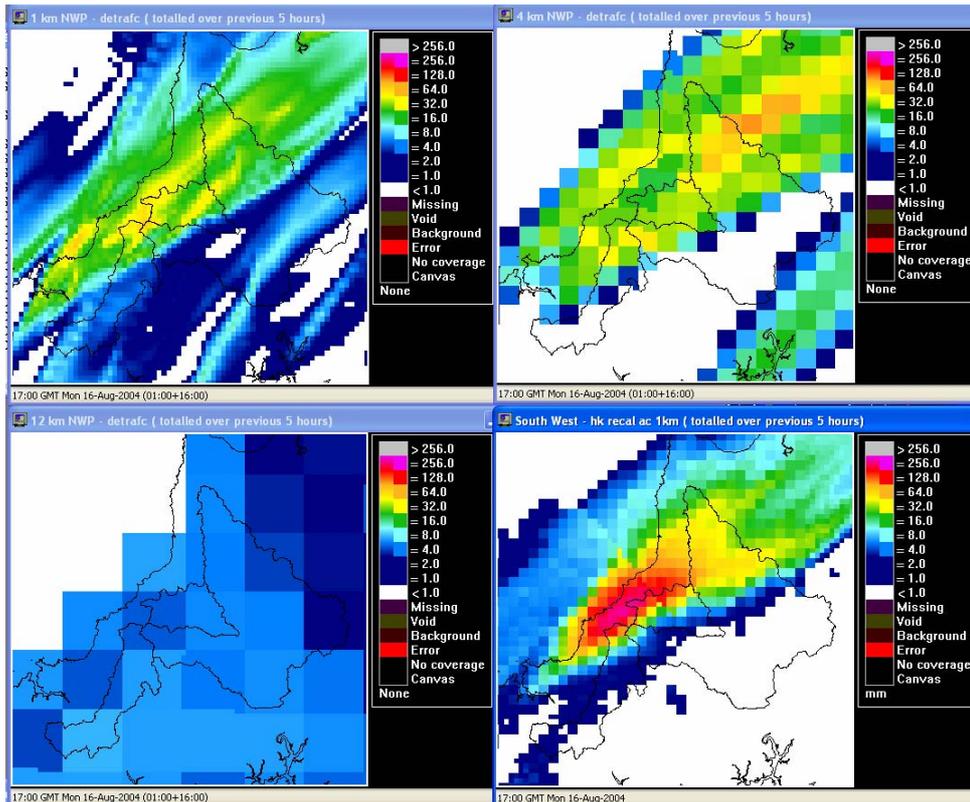
Notes: The 00UTC forecast is used in the left column and the 03UTC forecast is used in the right column.

Table 6.2 Catchment average rainfall totals (mm) for the three case study catchments using different sources of rainfall ¹.

Rainfall estimator	Ottery at Werrington	Tamar at Gunnislake	Camel at Denby
1 km NWP 00UTC	21.68	10.63	8.40
4 km NWP 00UTC	23.41	15.89	8.68
12 km NWP 00UTC	3.93	3.7	4.65
1 km NWP 03UTC	8.85	5.74	3.40
4 km NWP 03UTC	17.58	11.63	3.87
12 km NWP 03UTC	3.41	4.06	4.82
Nimrod composite	28.87	11.29	3.98
Raingauge-adjusted Nimrod composite	43.1	15.34	8.0

Notes ¹ Totals are for the five-hour period ending 17:00 16 August 2004.

00UTC NWP forecasts



03UTC NWP forecasts

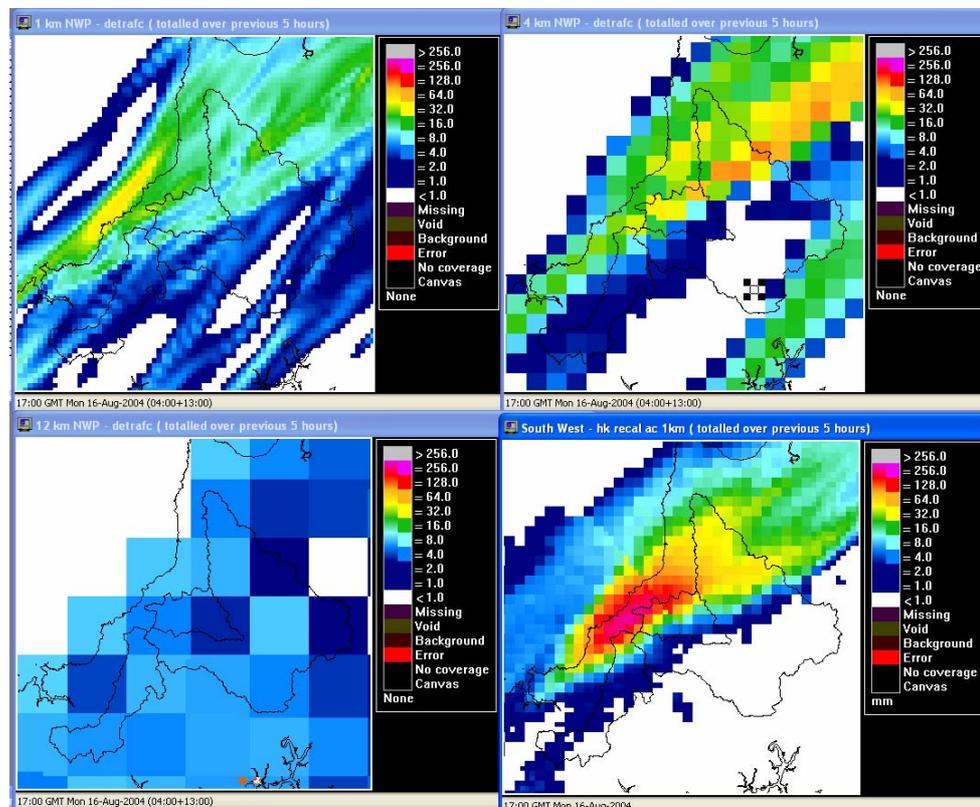


Figure 6.9 Rainfall accumulations (mm) for the 5 hour period ending 17:00 16 August 2004 using 1 km (top left), 4 km (top right) and 12 km (bottom left) km NWP and HyradK raingauge-adjusted radar (bottom right).

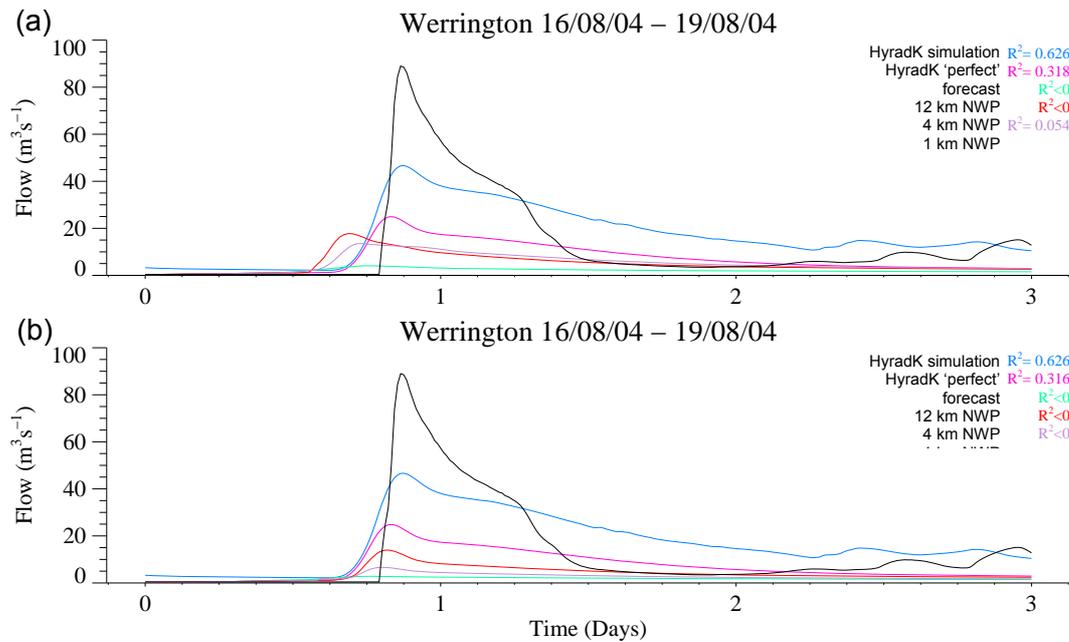


Figure 6.10 Fixed-origin G2G model forecasts at Werrington using 1, 4 and 12 km deterministic NWP and HyradK 'perfect' rainfall data: (a) using the 00UTC forecast and (b) using the 03UTC forecast.

Notes: The blue line gives the simulation mode result using HyradK rainfall data as a reference.

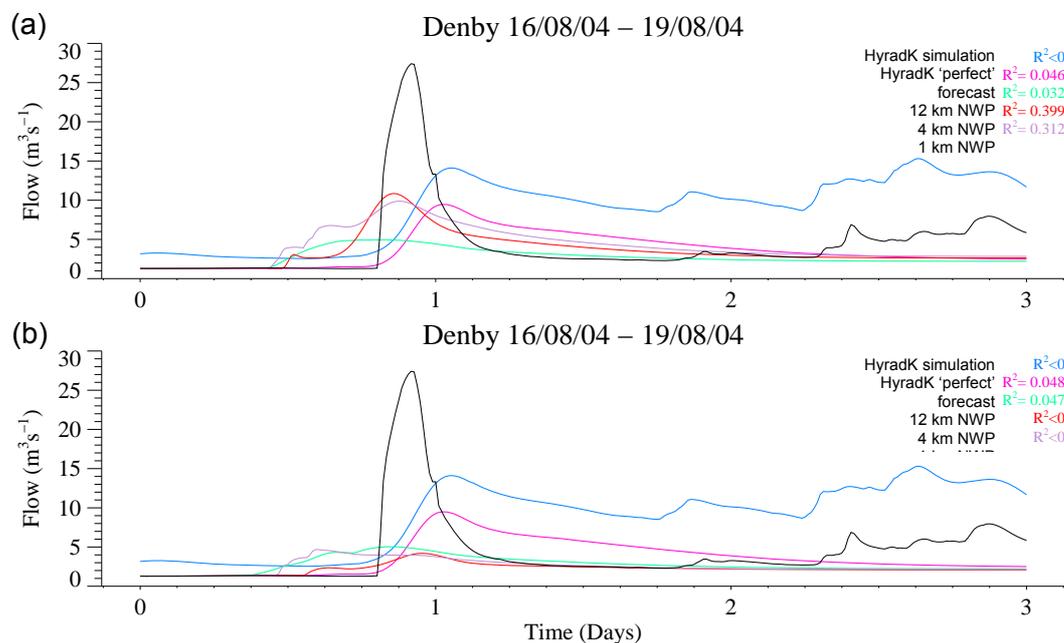


Figure 6.11 Fixed origin G2G Model forecasts at Denby using 1, 4 and 12 km deterministic NWP and HyradK 'perfect' rainfall data: (a) using the 00UTC forecast and (b) using the 03UTC forecast.

Notes: The blue line gives the simulation mode result using HyradK rainfall data as a reference.

6.5 Analysis of hydrological model forecasts using pseudo ensembles of high resolution NWP rainfalls

As in the analysis of the deterministic high resolution NWP rainfalls (Section 6.4), the most appropriate way to assess the hydrological model performance using the pseudo-ensembles of high-resolution 1 km NWP rainfalls is to run fixed origin forecasts.

Hydrological model forecasts using the 00UTC pseudo ensembles are presented for the PDM and G2G models in Figures 6.12 and 6.13 respectively using a forecast origin of 04UTC on 16 August 2004. For the REW model, forecasts have been made using the pseudo ensembles for forecast origins in hourly intervals from 04 to 20 UTC on 16 August 2004. Figure 6.14 presents the REW results for Gunnislake and Werrington, and Figure 6.15 presents the results for Denby and Slaughterbridge.

Analysis of the results shows that all models, whether lumped or distributed, are sensitive to the individual ensemble members. The spread of the hydrological ensembles looks encouraging and suggests that the simple method used to generate the pseudo ensembles is meaningful. At first, it may seem surprising that generating ensemble members through a small displacement (less than 20 km) could generate such sensitivity in the hydrological model outputs. However, as the Boscastle storm is small in spatial extent (see Figure 6.9) but large in rainfall magnitude, this relatively small displacement may cause large changes in total catchment rainfall. This means that both lumped and distributed models would be sensitive to the different ensemble members.

In addition, the distributed models (G2G and REW) are also sensitive to the placement of a storm within the catchment. For example, the Tamar catchment will show a much more pronounced reaction at Gunnislake if a storm falls at the bottom of the catchment than if the storm falls in the headwaters of the catchment. This is evident in the shape of the forecast hydrographs with the PDM model responding in a similar qualitative way to most members (see Figure 6.12) while the G2G and REW models (see Figures 6.13–6.15) show considerable variation in the hydrograph shape across the different members and catchments.

All this has implications for the use of high resolution NWP in an operational setting. The spatial displacement used in generating the pseudo ensembles has been taken so that it is in line with the perceived spatial accuracy of the NWP forecasts (in this particular meteorological scenario). As can be seen from the graphs, this can have an enormous effect on the flow from the investigated catchments. If only a deterministic forecast had been available in this case, the forecaster might have issued a warning for a single catchment only. However, the ensemble runs with a distributed model show that a serious event might have been possible in most of the catchments in the area. Although the simulation performance of the hydrological models for the Boscastle event leaves room for improvement, the combination of a distributed model with high-resolution ensemble rainfall forecasts gives a better indication of possible flood locations for such an extreme event.

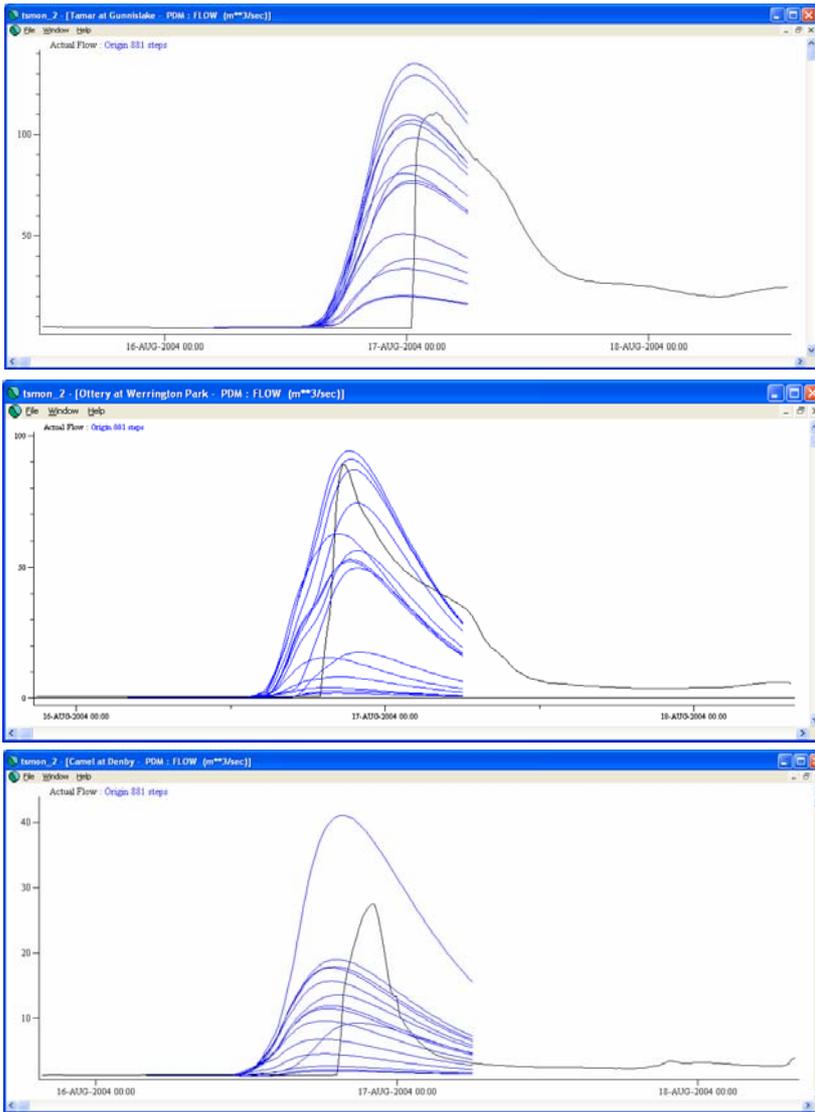


Figure 6.12 PDM model forecasts at Werrington, Gunnislake and Denby using the 00UTC 1km NWP pseudo ensemble.

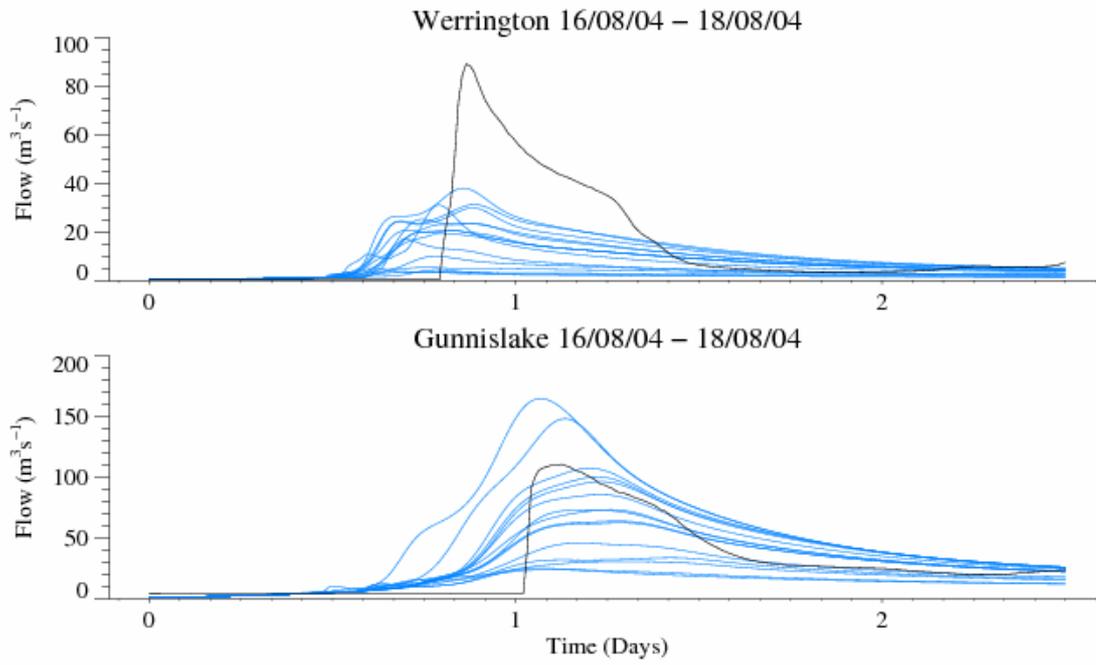


Figure 6.13 G2G Model forecasts at Werrington and Gunnislake using the 00UTC 1km NWP pseudo-ensemble.

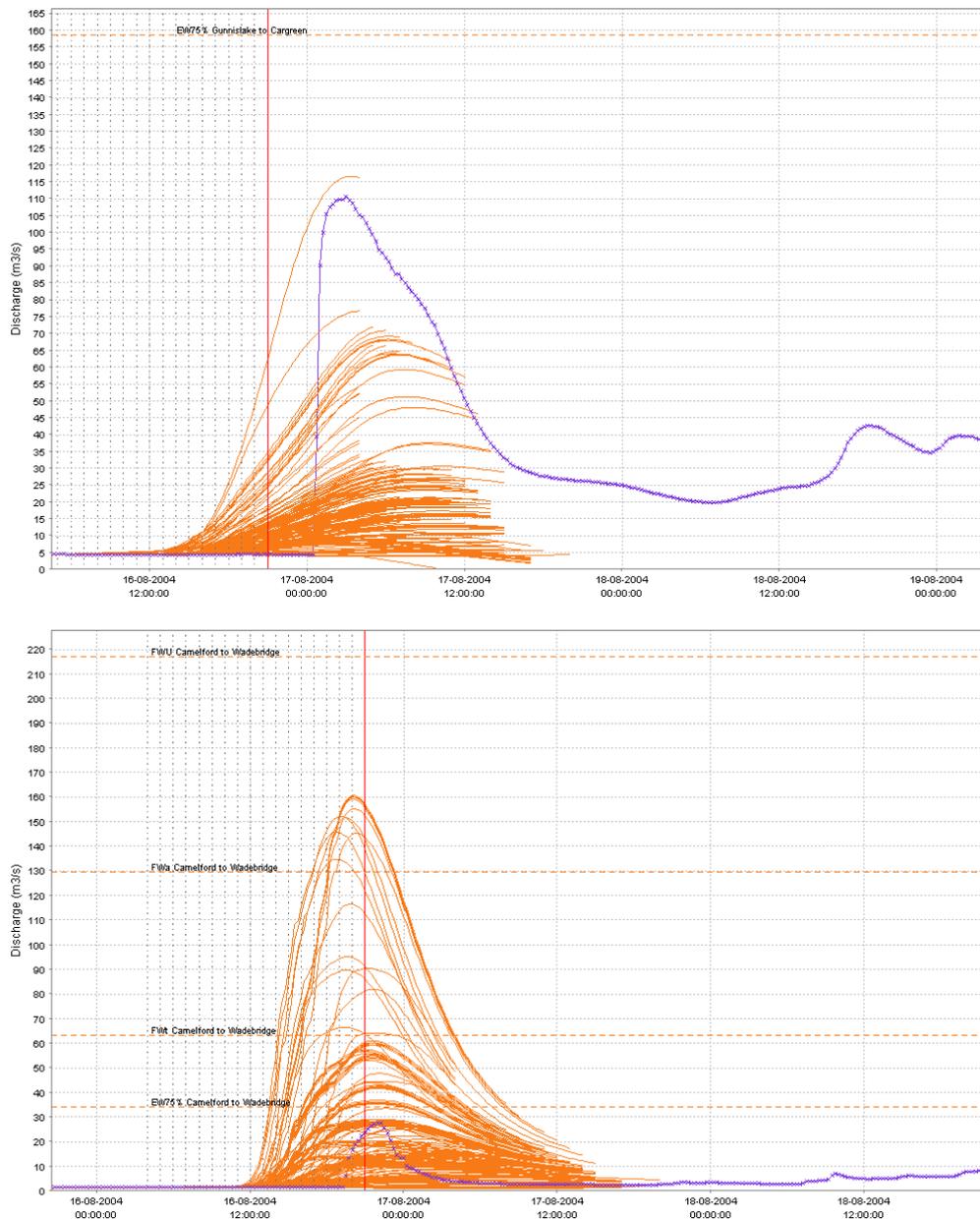


Figure 6.14 REW forecasts for Gunnislake and Denby made using the pseudo ensembles as input.

Notes Forecasts made at hourly intervals from 04:00 to 20:00 16 August 2004.

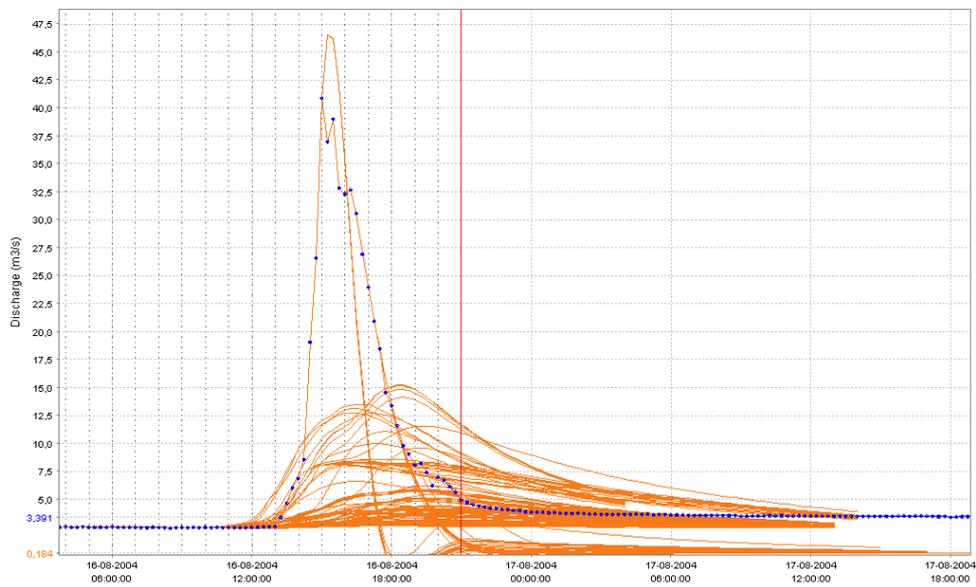
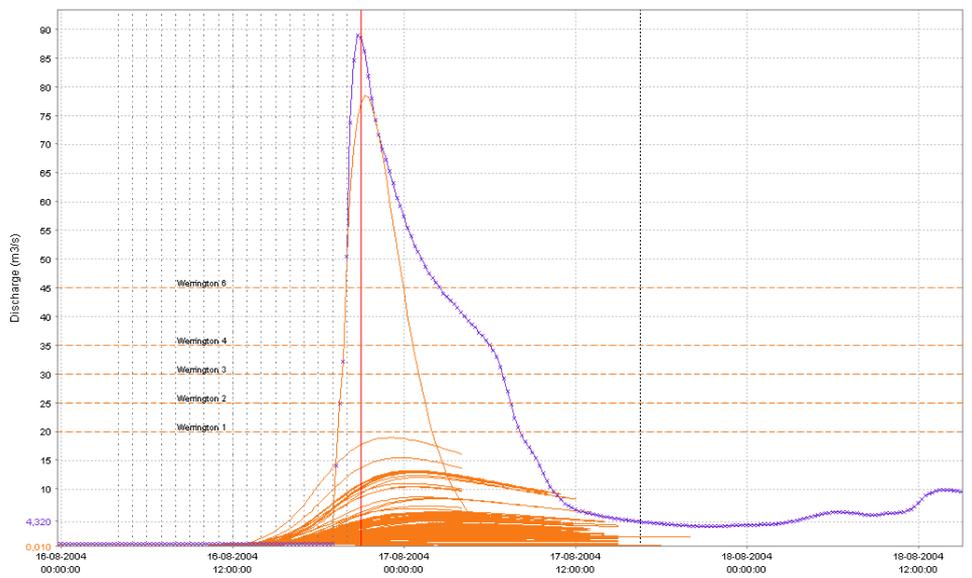


Figure 6.15 REW forecasts for Werrington and Slaughterbridge made using the pseudo ensembles as input.

Notes: Forecasts made at hourly intervals from 04:00 to 20:00 16 August 2004.

Part 2: Operational implementation of ensemble forecasting in NFFS

7 Ensemble forecasting in NFFS using MOGREPS

Ensemble forecasting was configured in the test NFFS system for two Environment Agency Regions – Thames and North East. The test system set up in Delft receives MOGREPS forecasts from the Met Office and processes these in an NFFS set-up. Selected Environment Agency staff were given remote access to the system (see Appendix C for instructions).

MOGREPS NWP rainfall ensembles were provided by the UK Met Office in experimental mode from the middle of January⁹ to the beginning of March 2008. Although this period of time was not long enough to determine conclusively the increase in skill through the use of ensembles, this section presents some sample outputs and discusses the preliminary operational outputs.

7.1 Selection of pilot Environment Agency Regions

Criteria for selection of pilot regions were:

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

These criteria led to the selection of Thames and North East Regions for the pilot application (see Phase 1 report).

- North East Region has short lead times to many of its upstream forecasting locations and is, to a large extent, covered with PDMs. The forecasting time step is 15 minutes and models are running fast.
- Thames Region has longer lead times to its most important forecasting locations but, on the other hand, has large, fast-responding urban areas within its forecasting responsibility. The Region is to a large extent covered with nested TCM models. The forecasting time step is 15 minutes and models are running fast. The nesting approach means that the larger currently available TCM models cover a long lead time, which makes them less beneficial when using nowcasting ensembles and, in some cases, for MOGREPS ensembles also.

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.

⁹ Data were supplied in real-time from 10 January 2008 to 1 March 2008 with data outages on 19 to 22 January, 26 to 27 January, and 26 to 27 February inclusive.

The ensemble rainfall forecasts should provide the input for the pilot on ensemble flood forecasting with NFFS. Background information from the Met Office about the ensemble prediction capability of both systems is presented below.

Please note that the STEPS ensembles will **not** be used. The description has been added for the sake of completeness.

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.

Nimrod¹⁰ and Gandolf¹¹ provided the Met Office's nowcasting capability until spring 2007 when Short Term Ensemble Prediction System (STEPS)¹² was introduced as a replacement.

STEPS provides ensemble prediction capability for nowcasting. This anticipates 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. 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.

7.2.2 MOGREPS

In 2005 the Met Office introduced a new ensemble system called MOGREPS (Met Office Global and Regional Ensemble Prediction System)¹³ 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 (days 0–3) for the UK and Ireland. It provides 24-member ensemble with a grid resolution of 24 km for a forecast length of 54 hours (36 hours are used in this research). 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. Model coverage is shown in Figure 7.2. Due to spin-up issues and the fact that only two forecasts are available per day, the first hours of the MOGREPS runs are generally not used.

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

¹⁰ <http://www.metoffice.gov.uk/water/nimrod.html>

¹¹ <http://www.metoffice.gov.uk/water/gandolf.html>

¹² <http://www.metoffice.gov.uk/science/creating/hoursahead/nowcasting.html>

¹³ <http://www.metoffice.gov.uk/science/creating/daysahead/ensembles/MOGREPS.html>

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 that is run at a 12 km resolution.

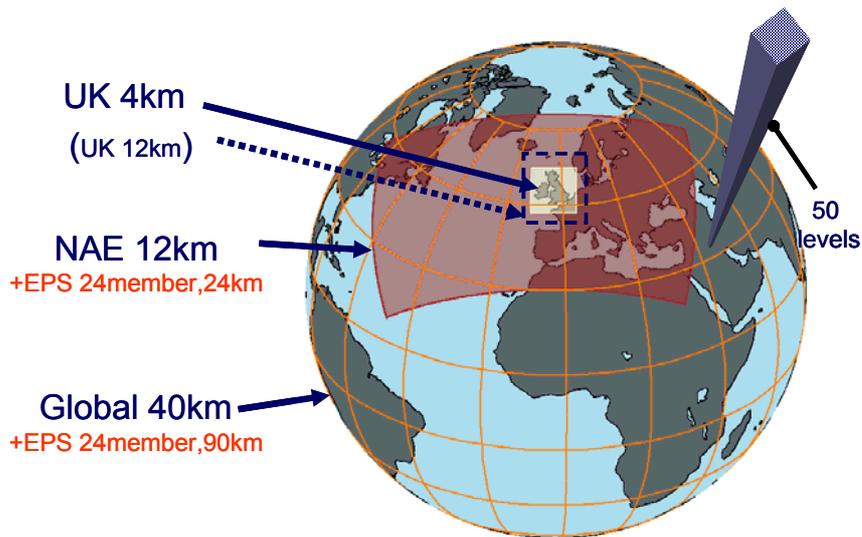


Figure 7.1 Model Coverage in MOGREPS.

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 should provide the quantitative basis for probabilistic flood forecasting.

7.2.3 Provision of MOGREPS Ensembles

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

The forecasts are sent in Nimrod file format. Delft-FEWS has been adjusted (catered for in its development budgets) to read these ensemble forecasts. The Met Office and the Environment Agency will have to decide about the file formats and how to make the feed to Delft operational.

7.3 Configuring ensemble forecasting in NFFS

The configuration is based on the current configuration of NFFS for North East and Thames Regions. No distributed models will be run.

The configuration changes include:

- importing and processing of NWP ensembles (MOGREPS);
- pre-processing of ensemble data to generate precipitation input;
- ensemble runs of forecasting models;
- data displays, including statistical analyses;
- reports for ensemble results;
- performance measures (implemented in code no results yet).

The NFFS configurations for North East and Thames Regions were extended to process the NWP ensembles and to display probabilistic forecast results.

For performance reasons, the gridded data for the individual ensemble members are not synchronised to the clients by default. However, they can be made visible when using a custom profile available in the test system (see Figure 7.2 and instructions in Appendix C).

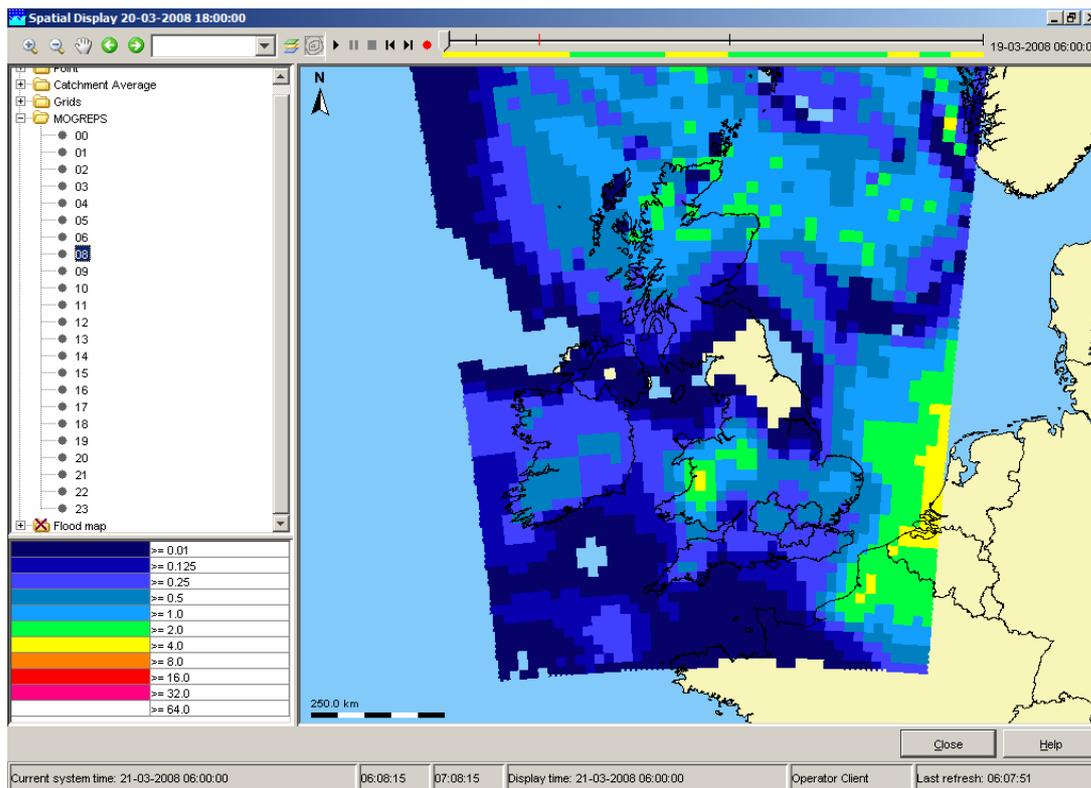


Figure 7.2 Display of single ensemble member in the T46 test system (at 18:00 20 March 2008).

7.3.1 Configuration changes in North East Region

Importing and displaying MOGREPS data

MOGREPS data are imported in the module instance, ImportMOGREPS 1.00 default.xml. All grids are stored with synclevel 7 so the data will not be sent to the clients automatically.

The data are read from 24 different directories (0–23) in which each directory contains an ensemble member. Because the Nimrod import is not ensemble-aware, each ensemble member is stored using ensembleId 0 and a different ensembleId (0–24). Later, the time series is made into a single ensemble (MOGREPS) in an interpolation module, MOGREPS_Spatial_Interpolation 1.00 default.xml.

Table 7.1 lists the files updated for importing and display of MOGREPS data.

Table 7.1 Files updated in North East Region NFFS configuration for import and display of MOGREPS data.

Workflow	Description
Import_workflow 2.33 default.xml	Import workflow
ImportMOGREPS 1.00 default.xml	MOGREPS Import Module Instance
IdImportMOGREPS 1.00 default.xml	MOGREPS ID mapping
IdMapDescriptors 2.24 default.xml	New ID mapping descriptor added (RegionConfig folder).
ModuleInstanceDescriptors 2.39 default.xml	New MI Descriptors added (RegionConfig folder).
Locations 2.39 default.xml	MOGREPS location added to regional locations file.
Grids 1.01 default.xml	MOGREPS grid properties added to regional grids file.
SpatialDisplay 2.28 default.xml	Spatial display of MOGREPS grids
sa_global.properties	MOGREPS import folder added to import folder tags.

Processing MOGREPS data and SNOWP models

The MOGREPS data are processed in a similar way to the non-ensemble Nimrod data that is part of the standard system. All processing is done in the Fluvial_FastResponse_Forecast_MOGREPS 1.00 default.xml workflow.

In all existing processing modules, the end time is set to 36 hours to match the length of the MOGREPS forecasts. Table 7.2 lists the files updated for processing MOGREPS data.

Table 7.2 Files updated in North East Region NFFS configuration for processing of MOGREPS data.

File	Description
Fluvial_FastResponse_Forecast_MOGREPS 1.00 default.xml	Fast responding catchments workflow with special MOGREPS modules and workflows included.
MOGREPS_Spatial_Interpolation 1.00 default.xml	Overlay MOGREPS grid with HYRAD polygons and SNOWP locations and compute catchment average
MOGREPS_CatchmentAveragePrecipitation 1.00 default.xml	Disaggregate from three hour to 15-minute intervals for catchments and SNOWP locations.
Fluvial_SNOWP_Forecast 1.00 default.xml	New workflow with all SNOWP models and input processing
SNOWP_Processing 2.22 default.xml	ensembleId=main added to all non rainfall series.
SNOWP_..... 2.21 default.xml (all models)	ensembleId=main added to all series with no ensemble input.
WorkflowDescriptors 2.25 default.xml	New workflow descriptors added (RegionConfig folder).

In the MOGREPS_Spatial_Interpolation file, the interpolation from MOGREPS grid to catchment average precipitation is carried out using three methods:

1. For the conversion of grids to all catchments that have polygons (locationset CatAvg_Spatial), the average of grid cells is used.
2. For the conversion of grids to all catchments that do not have polygons (one location in the Dales area and four locations in the Ridings area¹⁴), the value for the nearest cell centre is used.
3. For the conversion of grids to all SNOWP locations (locationset TemperatureSnowGenerated), the nearest cell centre is used.

After the extraction of catchment series (steps 1 and 2), the Catchment locationset can be used for all catchments in later operations.

Other files that have been adjusted while implementing the MOGREPS changes are listed in Table 7.3. New entries have been added to the filters to show the results of the MOGREPS data at the SNOWP locations. Only the main regional filter group has been updated; the area filter groups have not been updated.

¹⁴ There are three Areas in North East Region – Northumbria, Dales and Ridings.

Table 7.3 Other files adjusted in North East Region NFFS configuration.

File name	Adjustment
PREC_BACKUP_PROF 2.21 default.xml	syncLevel changed from 5 to 1, end time set to 36 hours.
Precip_CopyCatAvg 2.22 default.xml	End time set to 36 hours.
EVAP_..... 2.21 default.xml	End time set to 36 hours.
Filters 2.27 default.xml	Entries added for MOGREPS precipitation, merged precipitation and snow, SNOWP.

Running the MOGREPS ensembles in SNOW, PDM, KW and ARMA modules

The file, Fluvial_FastResponse_Forecast_MOGREPS 1.00 default.xml, includes three sub-workflows for the three areas in North East Region. These sub-workflows contain all the fast responding catchment modules as well as some input processing modules for precipitation and temperature. The flow-to-level modules that convert forecasted flow to levels are also included in these sub-workflows. The main changes to the modules are:

- Change the general adaptor (GA) config (forecast only) to include a main ensembleId in the non-ensemble series.
- Increase forecast length for all SNOW/PDM/KW/ARMA modules to 36 hours (as this is what MOGREPS provides).

Table 7.4 lists the config files that have been updated in addition to the GA module instances. New entries have been added to the filters to show the results of the SNOW/PDM/KW and ARMA modules.

Table 7.4 Other config files updated in North East Region NFFS.

File	Description
Fluvial_FastResponse_Forecast_MOGREPS 1.00 default.xml	Included fast responding Area workflows in ensemble mode.
Northumbria_Meteo_Processing 2.02 default.xml	ensembleld=main added and end time set to 36 hours.
Ridings_Meteo_Processing 2.01 default.xml	ensembleld=main added and end time set to 36 hours.
Dales_Meteo_Processing 2.03 default.xml	ensembleld=main added and end time set to 36 hours.
Snowconvertmm_Northumbria 2.01 default.xml	End time set to 36 hours.
Snowconvertmm_Aire 2.31 default.xml	End time set to 36 hours.
Snowconvertmm_Dales 2.01 default.xml	End time set to 36 hours.
Snowconvertmm_Ridings 2.01 default.xml	End time set to 36 hours.
TyneGenerate 2.21 default.xml	ensembleld=main added and end time set to 36 hours.
NiddGenerate 2.21 default.xml	ensembleld=main added and end time set to 36 hours.
TeesGenerate 2.21 default.xml	ensembleld=main added and end time set to 36 hours.
Gaunless_PDM_ErrorModel_MergelInputs 1.01 default.xml	End time set to 36 hours.
Gaunless_PDM_ErrorModel 1.01 default.xml	ensembleld=main added and end time set to 36 hours.
Swale_PDM_ErrorModel 1.01 default.xml	End time set to 36 hours.
Swale_PDM_ErrorModel_MergelInputs 1.01 default.xml	ensembleld=main added and end time set to 36 hours.
....FastFlowToLevel 2.21 default.xml	End time set to 36 hours.
Filters 2.27 default.xml	Time series of all models that use MOGREPS data added as well as merged precipitation.

Statistics

After running the catchment modules with the MOGREPS ensemble input, statistics are computed for the catchment rainfall and the output series of the PDM and ARMA modules. The following statistics for the ensemble series are computed: minimum, maximum, median, and 25, 33, 66, 75 percentiles. Table 7.5 lists the files that have been updated to compute the statistics.

Table 7.5 Files updated in order to compute statistics.

File	Description
Fluvial_FastResponse_Forecast_MOGREPS 1.00 default.xml	Statistics module added.
MOGREPS_PDM_Statistics 1.00 default.xml	Statistics for PDM, ARMA and rainfall catchment series added.
Parameters 1.90 default.xml	Statistics parameters for discharge and precipitation added.
LocationSets 2.36 default.xml	Added locationSets: HydroDischargeARMA_Fast .._Northumbria_Fast .._Dales_Fast ..ARMA_Ridings_Fast
LocationSets 2.36 default.xml	Added locationSets: HydroDischargeERRORModel_Fast .._Northumbria_Fast .._Dales_Fast

Other changes:

- location Dalton removed from HydroPDMDischargeUpdated_Dales locationSet
- location KIRBYW1 added to HydroPDMDischargeUpdated_Dales locationSet
- new entries added to the filters to show the results of the PDM and ARMA updated series statistics and the precipitation catchment statistics (i.e. statistics for PDM, ARMA and P.merged added to Filters 2.27 default.xml).

Pre-defined displays

In order to display the statistics as area graphs, the display groups have been updated with MOGREPS groups for all catchments. First five plot groups (RainfallMOGREPS, PDMSIMULATEDMOGREPS, PDMUPDATEDMOGREPS, ERRORMOGREPS and ARMAMOGRPS) are made in the display groups. Plot groups have been added to the file DisplayGroups 2.34 default.xml.

Figure 7.3 shows an example pre-defined display.

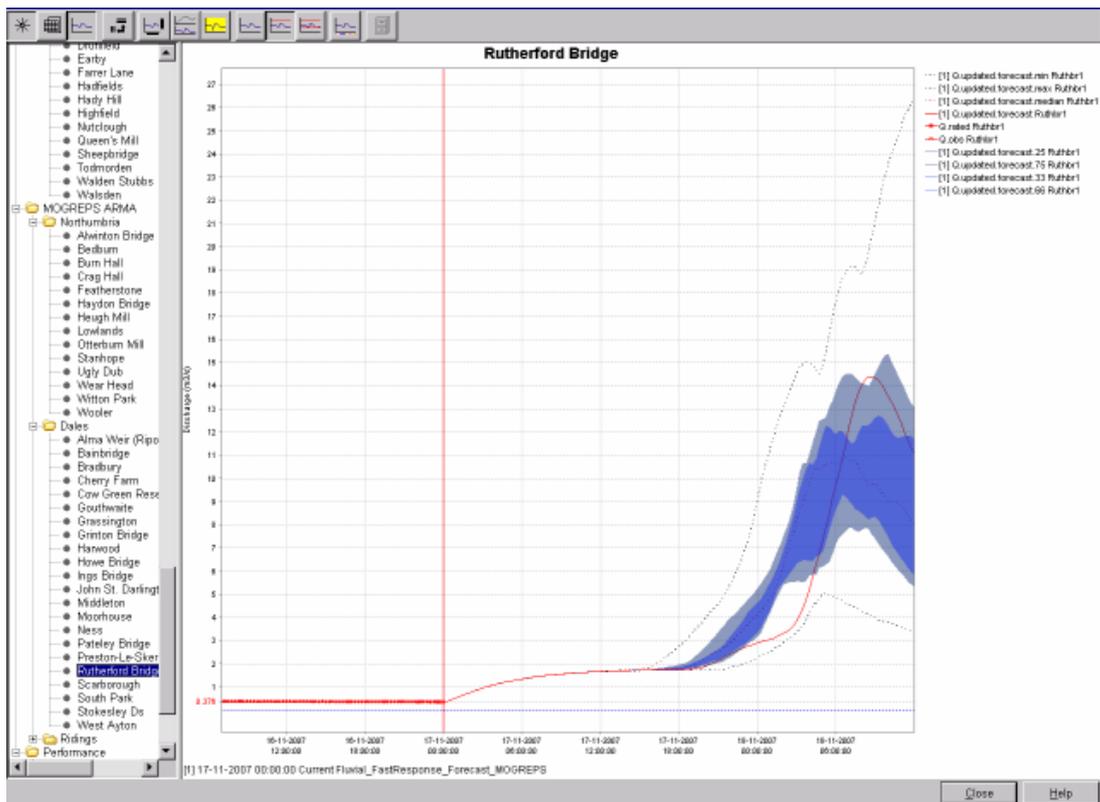


Figure 7.3 Example of a pre-defined display.

Reports

Reports of MOGREPS forecast output are generated in the new report module instance, Report_MOGREPS 1.00 default.xml. This report is part of the general Export_Current 2.35 default.xml workflow. Table 7.6 lists the files updated in order to generate MOGREPS reports.

Table 7.6 Files updated in order to generate MOGREPS reports.

File	Description
Export_Current 2.35 default.xml	Report workflow with new MOGREPS report module added.
Report_MOGREPS 1.00 default.xml	Module Instance that generates MOGREPS reports
fluvial_forecastlocation_template9 1.00 default.html	Template in reportTemplates folder copied from Thames Region
Report_Export_ZIPFile 2.22 default.xml	Export of MOGREPS report included.
Report_Export 2.23 default.xml	Export of MOGREPS report included.
Report_Export 2.49 default.zip	File northeast_navigation.js updated with MOGREPS links.
Report_Export_ZIPFile 2.49 default.zip	File northeast_navigation.js updated with MOGREPS links.

Figure 7.4 shows an example report.

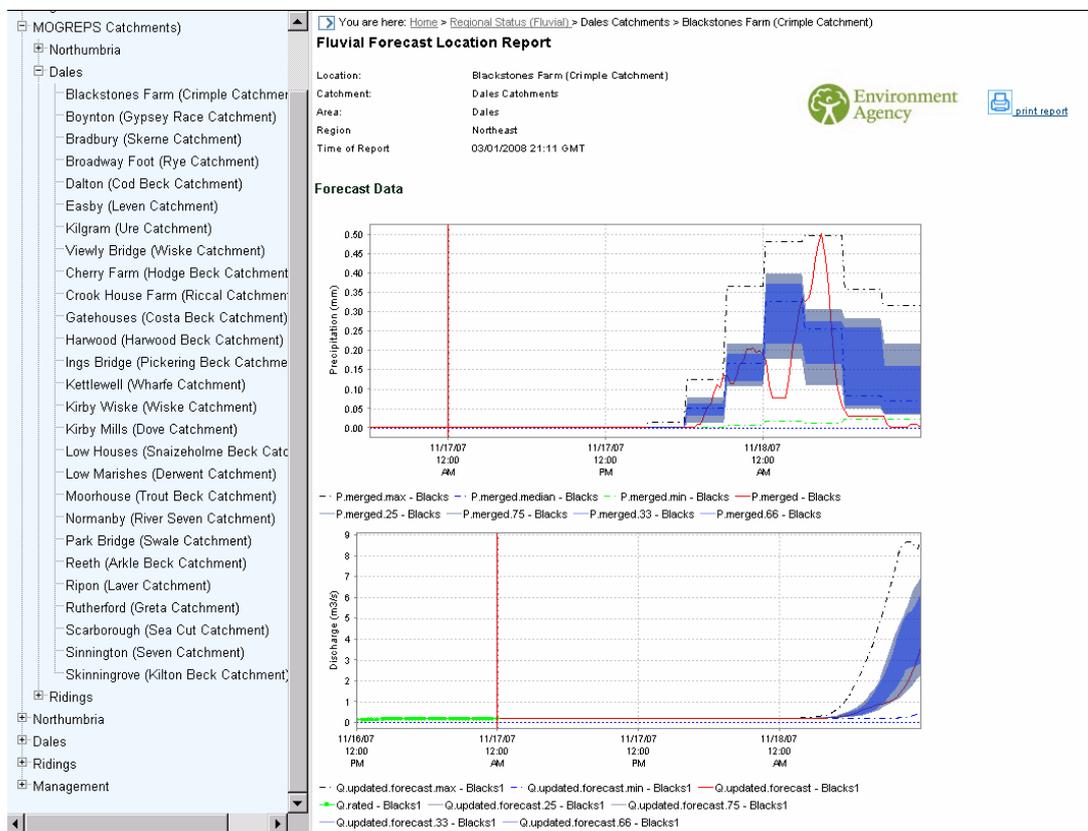


Figure 7.4 Example report of North East Region MOGREPS results.

7.3.2 Configuration changes in Thames Region

Importing and displaying MOGREPS data

MOGREPS data are imported in the module instance, ImportMOGREPS 1.00 default.xml. All grids are stored using synclevel 7. The data are read from 24 different directories (0–23) in which each directory contains an ensemble member. Because the Nimrod import is not ensemble-aware, each ensemble member is stored using ensembleId 0 and a different ensembleId (0–24). Subsequently the time series are made into a single ensemble (MOGREPS) in an interpolation module, MOGREPS_Spatial_Interpolation 1.00 default.xml. Table 7.7 lists the files updated in order to import and display MOGREPS data.

Table 7.7 Files updated in Thames Region NFFS configuration in order to import and display MOGREPS data.

WorkFlow	Description
ImportTelemetry 2.12 default.xml	Import workflow.
ImportMOGREPS 1.00 default.xml	MOGREPS import module instance
IdImportMOGREPS 1.00 default.xml	MOGREPS ID mapping
IdMapDescriptors 2.08 default.xml	New ID mapping descriptor added (RegionConfig folder).
ModuleInstanceDescriptors 2.83 default.xml	New MI descriptors added (RegionConfig folder).
Locations 2.14 default.xml	MOGREPS location added to regional locations file.
Grids 1.03 default.xml	MOGREPS grid properties added to regional grids file.
GridDisplay 2.11 default.xml	Spatial display of MOGREPS grids
sa_global.properties	MOGREPS import folder added to import folder tags.

Processing MOGREPS data

The MOGREPS data are processed in a similar way to the existing processing on non-ensemble Nimrod data. All processing is performed in the Fast_All_MOGREPS 1.00 default.xml workflow. In all existing processing modules, the end time has been set to 36 hours to match the length of the MOGREPS forecasts. New entries have been added to the filters to show the results of the MOGREPS data at the SNOWP locations. Only the main regional filter group has been updated; the area filter groups have not been updated.

Running the MOGREPS TCM and ARMA modules

The file, Fast_All_MOGREPS 1.00 default.xml, includes all the modules and sub-workflows (running the catchment models) for Thames Region. These sub-workflows contain all the catchment modules as well as some input processing modules for precipitation. The flow-to-level modules to convert forecasted flow to levels are also included in these sub-workflows.

The main changes in the modules are:

- change GA config (forecast only) to include a main ensemble in the non-ensemble series;
- increase forecast length for all TCM/ARMA modules to 36 hours (as this is what MOGREPS provides).

New entries have been made in the filters to show the results of MOGREPS forecasts.

Statistics

After running the catchment modules with the MOGREPS ensemble input, statistics are computed for the catchment rainfall and the output series of the TCM and ARMA modules. The following statistics for the ensemble series are computed: minimum, maximum, median and 25, 33, 66, 75 percentiles. New entries have been made to the filters to show the results of the TCM and ARMA updated series statistics and the precipitation catchment statistics.

Pre-defined displays

In order to display the statistics as area graphs, the display groups have been updated with MOGREPS groups for all catchments both for 'plain' TCM output and for the ARMA corrected discharge.

Reports

Reports of MOGREPS forecast output are generated in a new report module instance, Report_MOGREPS 1.01 default.xml. This report is part of the general Export_Current 2.01 default.xml *workflow*.

Thresholds

Thames Region has not yet set thresholds for TCM/ARMA results. For this project, the existing level thresholds were converted to flow for all TCM model locations with a threshold. This was set for the deterministic forecast and for the 66% of the ensemble forecast.

The MOGREPS thresholds for the three Areas in Thames Region (South East, North East and West) are given in Tables 7.8–7.10.

Table 7.8 Thames Region South East area MOGREPS thresholds.

Done?	LocationId	threshold?	rating?	name/Rating ID
X	2989TH	2989TH	2989TH	Addlestone
X	2620TH	2620TH	2620TH	Binfield
X	4370TH	4370TH	4370TH	Catford Hill
–	4180TH	#N/A	#N/A	#N/A
–	3270TH	3270TH	#N/A	#N/A
–	3290TH	3290TH	#N/A	#N/A
–	2427TH	2427TH	#N/A	#N/A
X	3061TH	3061TH	3061TH	Flash Bridge
X	3229TH	3229TH	3229TH	Gatwick Link
–	3210TH	3210TH	#N/A	#N/A
–	3080TH	3080TH	#N/A	#N/A
X	2936TH	2936TH	2936TH	Guildford Street
X	4310TH	4310TH	4310TH	Hayes Lane
X	3230TH	3230TH	3230TH	Horley
X	3369TH	3369TH	3369TH	3369TH
X	3390TH	3390TH	3390TH	Kingston Hogsmill
X	3240TH	3240TH	3240TH	Kinnersley Manor
–	2442TH	2442TH	#N/A	#N/A
X	4389TH	4389TH	4389TH	Manor House Gardens
–	2420TH	#N/A	#N/A	#N/A
–	2469TH	#N/A	#N/A	#N/A
X	3040TH	3040TH	3040TH	Tilford
X	2927TH	2927TH	2927TH	Trumps Green
–	2490TH	2490TH	#N/A	#N/A
–	3090TH	3090TH	#N/A	#N/A
–	2700TH	2700TH	#N/A	#N/A
–	3350TH	3350TH	#N/A	#N/A

Table 7.9 Thames Region North East area MOGREPS thresholds.

Done?	LocationId	threshold?	rating?	name/Rating ID
–	5427TH	5427TH	#N/A	#N/A
X	3829TH	3829TH	3829TH	Colindeep Lane
–	3870TH	3870TH	#N/A	#N/A
X	2870TH	2870TH	2870TH	Denham Colne
–	2879TH	#N/A	#N/A	#N/A
X	3826TH	3826TH	3826TH	Edgware Hospital
X	5357TH	5357TH	5357TH	Edmonton Green
X	5189TH	5189TH	5189TH	Elizabeth Way
X	5420TH	5420TH	5420TH	High Ongar
X	5470TH	5470TH	5470TH	Loughton
–	5080TH	5080TH	#N/A	#N/A
X	3680TH	3680TH	3680TH	Marsh Farm
X	3850TH	3850TH	3850TH	Monks Park
X	5480TH	5480TH	5480TH	Redbridge
–	5169TH	5169TH	#N/A	#N/A
X	2829TH	2829TH	2829TH	Uxbridge PSTN Level/Flow
X	2810TH	2810TH	2810TH	Warren Gate Road
–	4690TH	4690TH	#N/A	#N/A
X	3839TH	3839TH	3839TH	Wembley
–	4827TH	4827TH	#N/A	#N/A
X	3824TH	3824TH	3824TH	Wolverton Road
–	5349TH	5349TH	#N/A	#N/A
–	5369TH	5369TH	#N/A	#N/A
–	5129TH	5129TH	#N/A	#N/A

Table 7.10 Thames Region West area MOGREPS thresholds.

Done?	LocationId	threshold?	rating?	name/Rating ID
–	0260TH	0260TH	#N/A	#N/A
–	0660TH	0660TH	#N/A	#N/A
–	0790TH	0790TH	#N/A	#N/A
–	1020TH	1020TH	#N/A	#N/A
–	1080TH	1080TH	#N/A	#N/A
–	1090TH	#N/A	#N/A	#N/A
–	1290TH	#N/A	#N/A	#N/A
–	1290_w1TH	1290_w1TH	#N/A	#N/A
–	1290_w2TH	#N/A	1290_w2TH	Cassington
–	1420TH	1420TH	#N/A	#N/A
–	1460TH	#N/A	#N/A	#N/A
–	1790TH	1790TH	#N/A	#N/A
–	1925TH	1925TH	#N/A	#N/A
–	1980TH	1980TH	#N/A	#N/A
X	2210TH	2210TH	2210TH	Marlborough
–	2250TH	2250TH	#N/A	#N/A
X	2290TH	2290TH	2290TH	Theale HMFF set to PSTN
–	2590TH	2590TH	#N/A	#N/A

7.4 Test environment

The complete NFFS configurations for the two Regions have been set up on the test environment in Delft with the exception of some of the coastal forecasting and the ISIS model runs.

The ISIS models were not included at this stage because that would have required additional licences. In order to run ISIS in ensemble mode, hardware would be needed to bring down the longer run times associated with hydrodynamic modelling.

The test environment was set up with low cost system software detailed in Table 7.11. A further description can be found in Appendix C and Section 7.5.2.

Table 7.11 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 Effects of (hydrological) model concept and size

A short (and incomplete) analysis of the results of the pilot in a hydrological sense is presented below. A more complete analysis will be presented in the final project report when it is hoped that the number of events since the start of the MOGREPS data feed will have increased enabling a better analysis.

7.5.1 Catchment models used in the pilot

Thames Catchment Model

The structure of the Thames Catchment Model (TCM) (Greenfield 1984, Wilby *et al.* 1994, Moore and Bell 2001) is based on subdivision of a basin into different response zones representing, for example, runoff from aquifer, clay, riparian and paved areas and sewage effluent sources. Within each zone, the same vertical conceptualisation of water movement is used; the different characteristic responses from the zonal areas are achieved through an appropriate choice of parameter set, some negating the effect of a particular component used in the vertical conceptualisation. The zonal flows are combined, passed through a simple routing model (optional) and go to make up the basin runoff.

A conceptual representation of a hydrological response zone in the TCM is presented in Figure 7.5 using nomenclature appropriate to an aquifer zone. This zone structure is used for all types of response zone but with differing nomenclature; for example, percolation is better described as rainfall excess for zones other than aquifer. Within a given zone, water movement in the soil is controlled by the classical Penman storage configuration in which a near-surface storage – of a depth related to the rooting depth of the associated vegetation and to the soil moisture retention characteristics of the soil (the root constant depth) – drains only when full into a lower storage of notional infinite depth. Evaporation occurs at the Penman potential rate while the upper store contains water and at a lower rate when only water from the lower store is available. The Penman stores are replenished by rainfall but a fraction, ϕ (typically 0.15, and usually only relevant to aquifer zones), is bypassed to contribute directly as percolation to a lower ‘unsaturated storage’. Percolation occurs from the Penman stores only when the total soil moisture deficit has been made up.

The total percolation forms the input to the unsaturated storage. This behaves as a linear reservoir, releasing water in proportion to the water stored at a rate controlled by the reservoir time constant, k . This outflow represents ‘recharge’ to a further storage representing storage of water below the phreatic surface in an aquifer. Withdrawals are allowed from this storage to allow pumped groundwater abstractions to be represented. A quadratic storage representation is used, with outflow proportional to the square of the water in store and controlled by the nonlinear storage constant, K .

Total basin runoff derives from the sum of the flows from the quadratic store of each zonal component of the model delayed by a time, τ_d . Provision is also made to include a constant contribution from an effluent zone if required.

A more recent extension of the model passes the combined flows through an additional channel flow routing component if required. This component of the model derives from the channel flow routing model developed by the Institute of Hydrology (Moore and Jones 1978, Jones and Moore 1980) which, in its basic form, takes the kinematic wave speed as fixed. The model employs a finite difference approximation to the kinematic wave model with lateral inflow. The delay and attenuation of the flood wave is controlled by the spatial discretisation used and a dimensionless wave speed parameter, θ . The parameters of the TCM are summarised in Table 7.12.

The TCM features – along with the Isolated Event Model (IEM) – within the PSM (Penman Store Model) software, where further details can be found (CEH Wallingford 2005b). This includes details of the error prediction and state corrections methods available with the TCM for real-time forecast updating using river flow measurements.

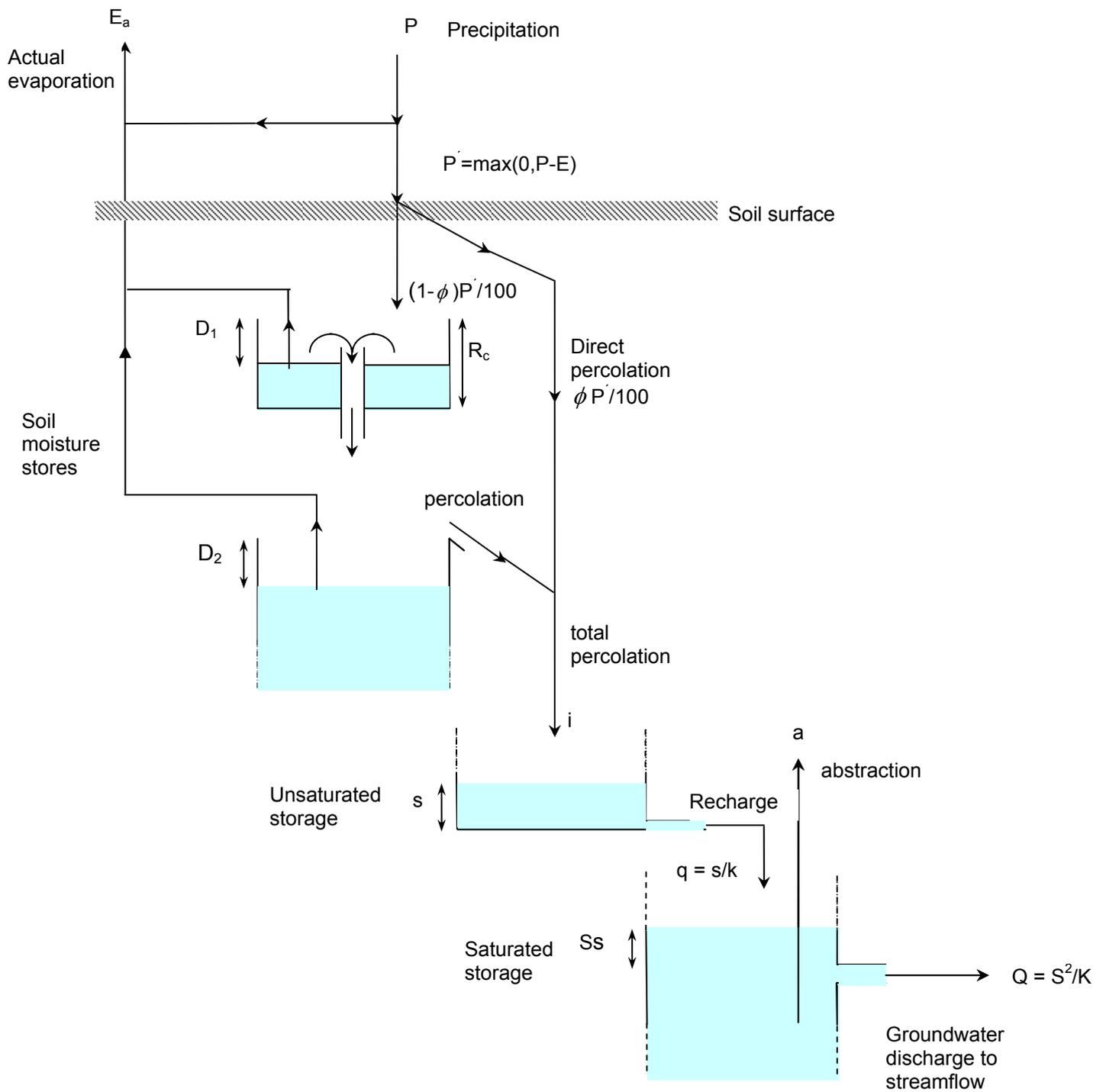


Figure 7.5 Representation of a hydrological response zone within the Thames Catchment Model.

Table 7.12 Parameters in the Thames Catchment Model.

Parameter name	Unit	Description
<i>Zone parameters</i>		
A	km ²	Area of hydrological response zone
γ	none	Drying rate in lower soil zone (usually $\gamma = 0.3$)
R_c	mm	Depth of upper soil zone (drying or root constant)
R_p	mm	Depth of lower soil zone (notionally infinite)
ϕ	none	Direct percolation factor (proportion of rainfall bypassing soil storage)
k	h	Linear reservoir time constant
K	mm h	Quadratic reservoir time constant
a	m ³ s ⁻¹	Abstraction rate from quadratic reservoir
<i>Other parameters</i>		
n_z	none	Number of zones
q_c	m ³ s ⁻¹	Constant flow (effluent or river abstraction)
τ_d	h	Time delay
N	none	Number of channel sub-reaches
θ	none	Dimensionless wave speed, $c\Delta t / \Delta x$

In the Thames system, application of the time-delay parameter is used for larger catchments, effectively shifting the precipitation input forward in time. For the larger catchments (e.g. Sutton Courtney), a delay of up to 30 hours is used. As a result, the forecasted precipitation is hardly used in the modelling. Therefore, these catchments will show no (or very little) spread in resulting discharge as a result of the different forecasted MOGREPS rainfall inputs.

Another factor that plays an important role is the size of the catchment itself compared to the resolution of the MOGREPS (and deterministic) grid size. An example is shown in Figure 7.19. In this case the amount of rain forecasted by the deterministic forecast is very similar to the mean of the ensemble. However, the intensity is very different. In this case (the catchment is rather small), the grid size of the deterministic is very similar to the size of the catchment while the grid of MOGREPS is larger than the actual catchment. The latter will always result in a lower intensity of the forecasted rainfall.

On the other end of the spectrum are the large TCMs in the West area of Thames Region. Even without the precipitation delay, the models cover an area that is so large that a lot of the variation between the ensemble members is smeared out as the spatial variation may very well occur *within* the catchment area. In that case, the resulting catchment average precipitation input may have very little variation.

Probability Distribution Moisture model

A description of the PDM model is given in Section 4.1. Within North East Region, two types of PDM configurations can be found. Some PDMs have been configured with state updating; others do not include state updating but a separate CEH ARMA module has been configured to create an updated forecast time series.

7.5.2 Two forecasts on 21 March for Thames Region

This section presents a few examples taken from the test system. Although there was no major event, the difference in reaction of different models can nonetheless be demonstrated.

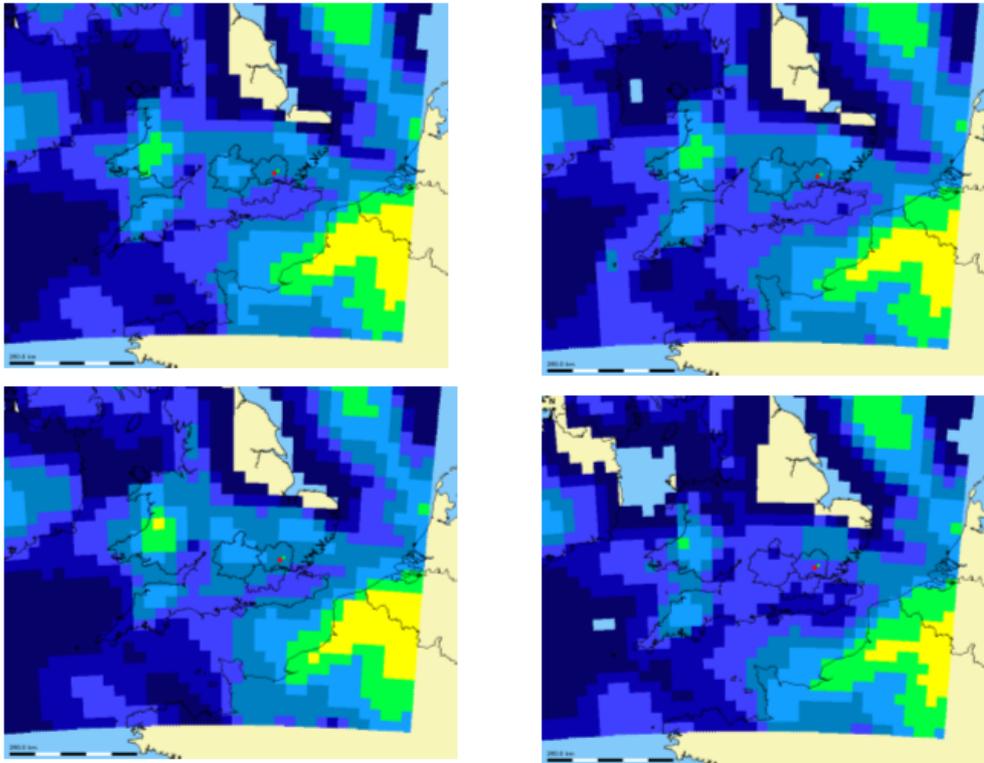


Figure 7.6 Ensemble member 0, 1, 2 and 3 of the MOGREPS forecast on 21 March used in the examples below.

Figure 7.6 shows four MOGREPS ensemble members (for the same time) of the forecast used in the examples of hydrological output shown below. The following catchments in Thames Region were used in the analysis:

- Albany Park in north London;
- Edmonton Green in north London;
- Wheatley near Oxford.

Both Albany Park and Edmonton Green are relatively small and fast-responding catchments in the North East Area of Thames Region, while Wheatley is a large catchment in its West Area. Figures 7.7–7.9 show the size and location of these catchments.

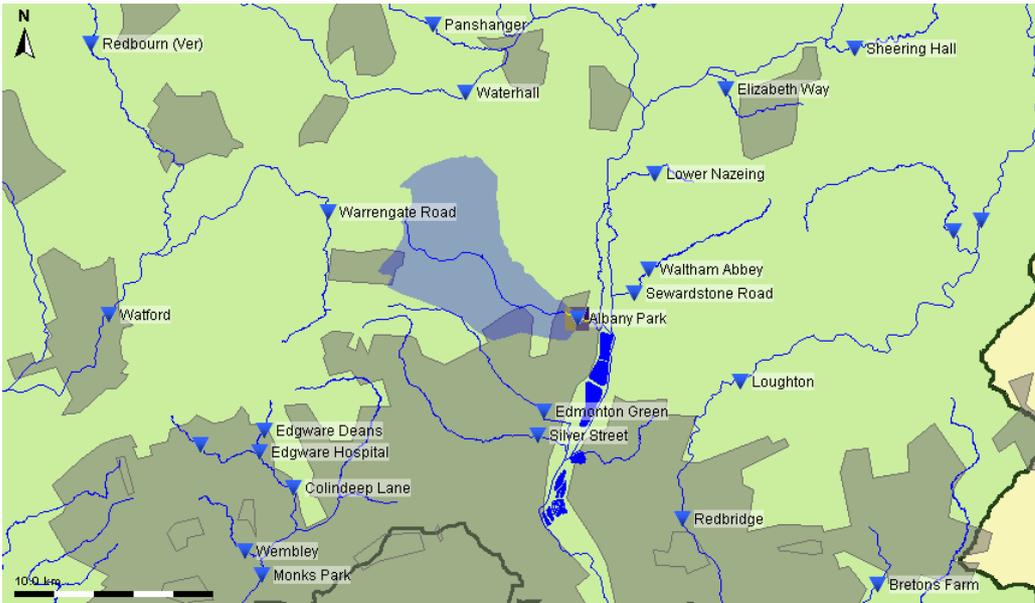


Figure 7.7 Albany Park catchment.

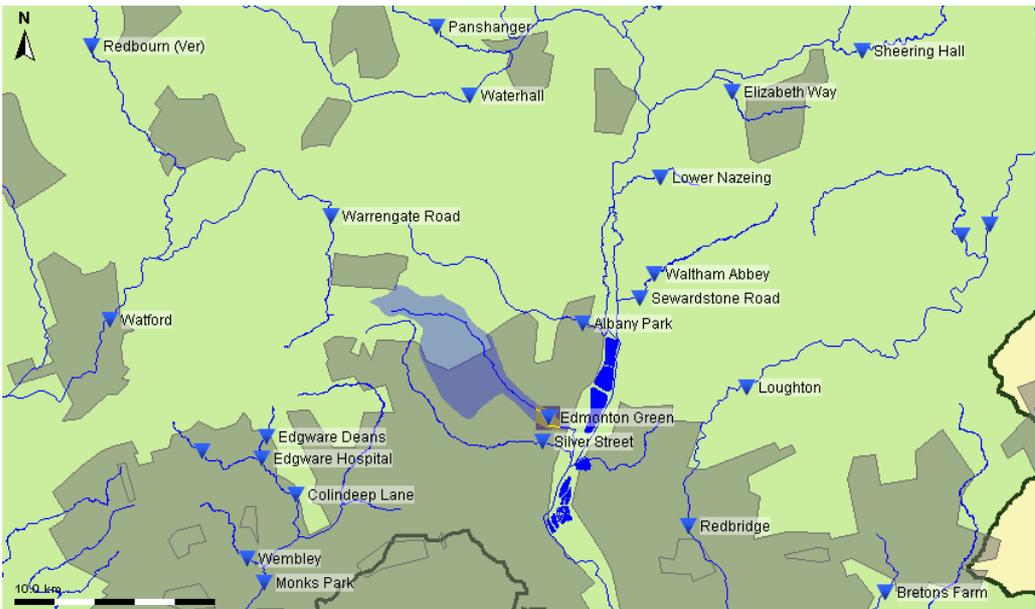


Figure 7.8 Edmorton Green catchment.

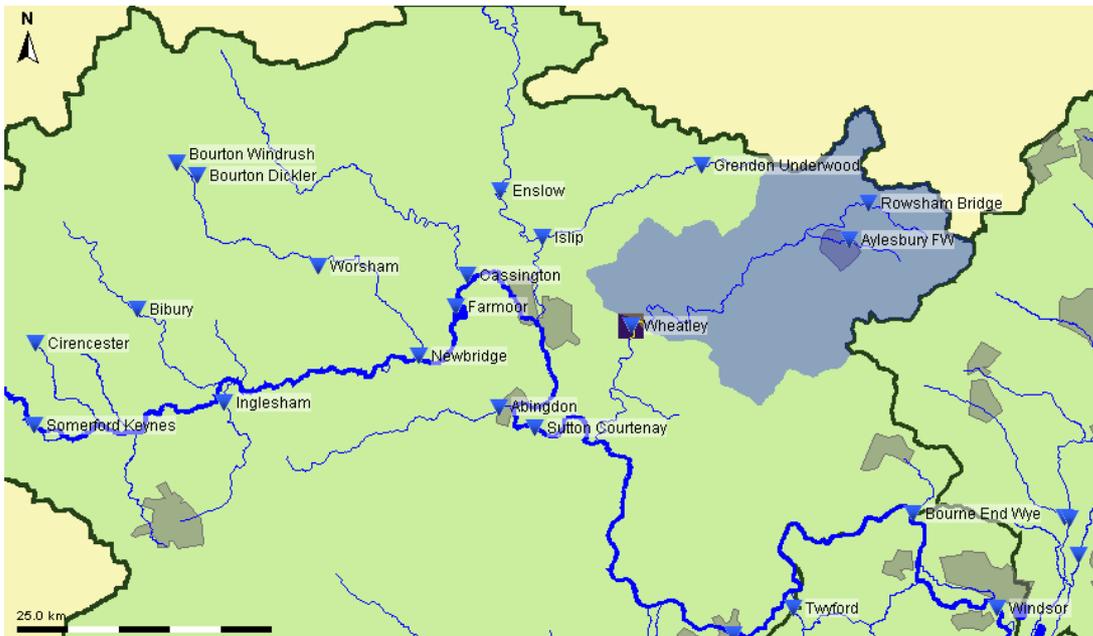


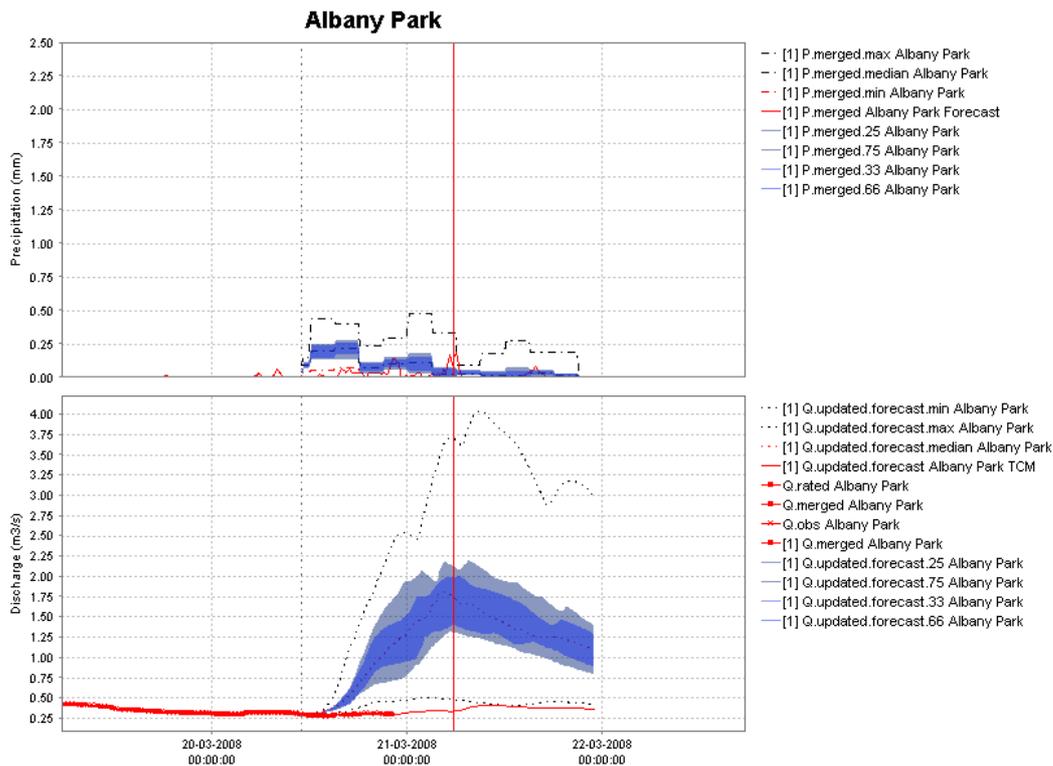
Figure 7.9 Wheatley catchment.

Figures 7.10 and Figure 7.11 show two consecutive forecasts for Albany Park.

In the first forecast, the MOGREPS forecast assumes more precipitation than the deterministic forecast for this area. As a result, the deterministic forecast lies outside the entire spread of the MOGREPS ensemble. Contrastingly, the next forecast shows a more 'normal' picture with the deterministic result within the ensemble spread. In addition, the total amount of forecasted precipitation is much less than in the previous forecast.

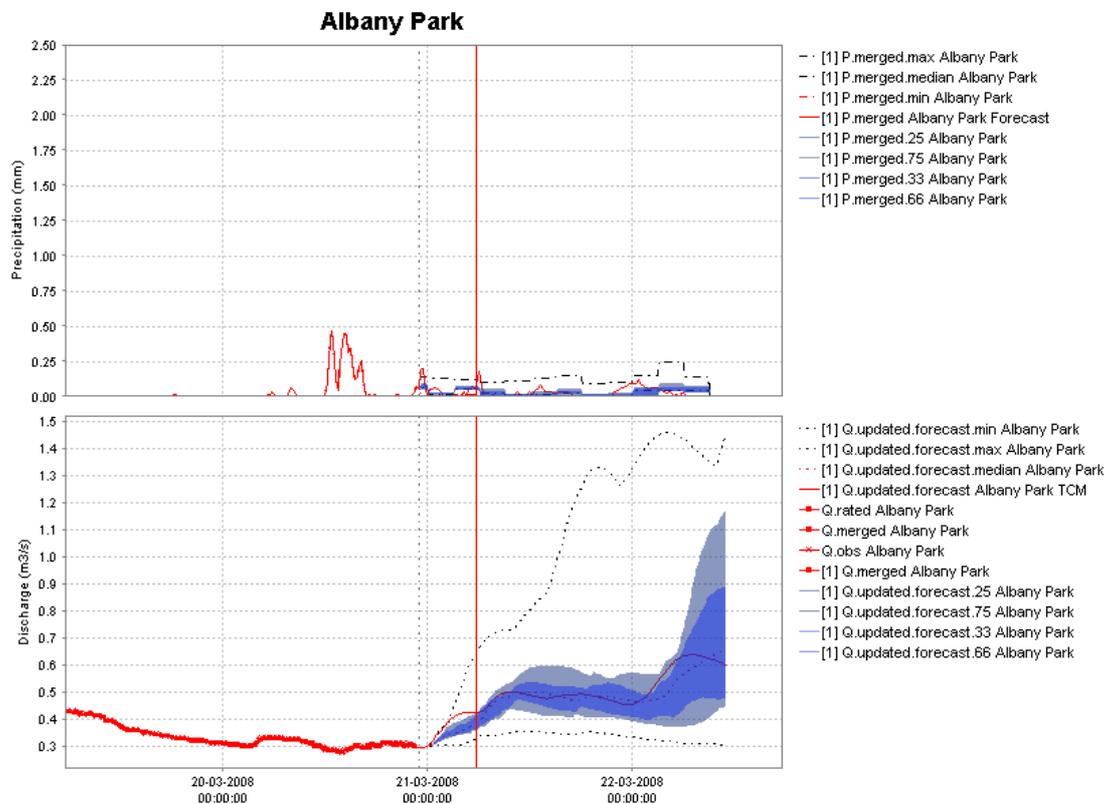
In this case (not a significant event), there seems to be little added value from the ensemble spread apart from that fact that it reminds us that two consecutive forecasts can produce very different results.

Results for this event using the similar sized Edmonton Green catchment show the same pattern (Figure 7.12 and Figure 7.13).



[1] 20-03-2008 11:00:00 Fast_All_MOGREPS

Figure 7.10 Forecast for 11:00 20 March for Albany Park.



[1] 20-03-2008 23:00:00 CurrentFast_All_MOGREPS

Figure 7.11 Forecast for 23:00 20 March for Albany Park.

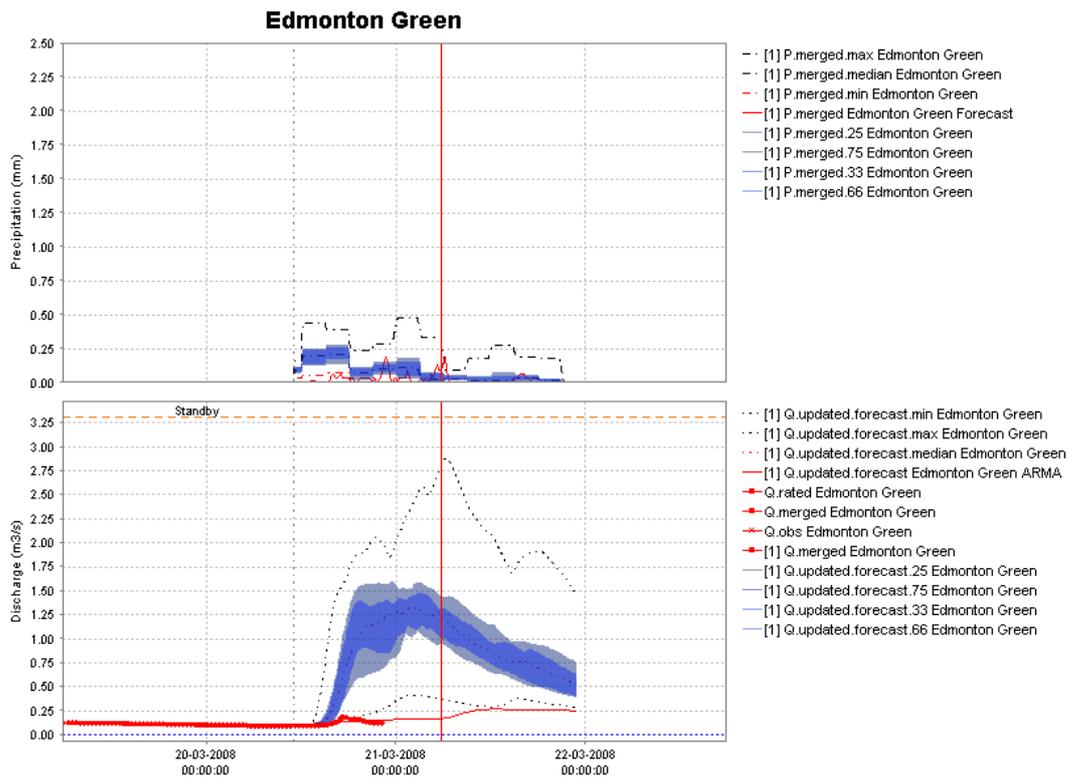


Figure 7.12 Forecast for 11:00 20 March for Edmonton Green.

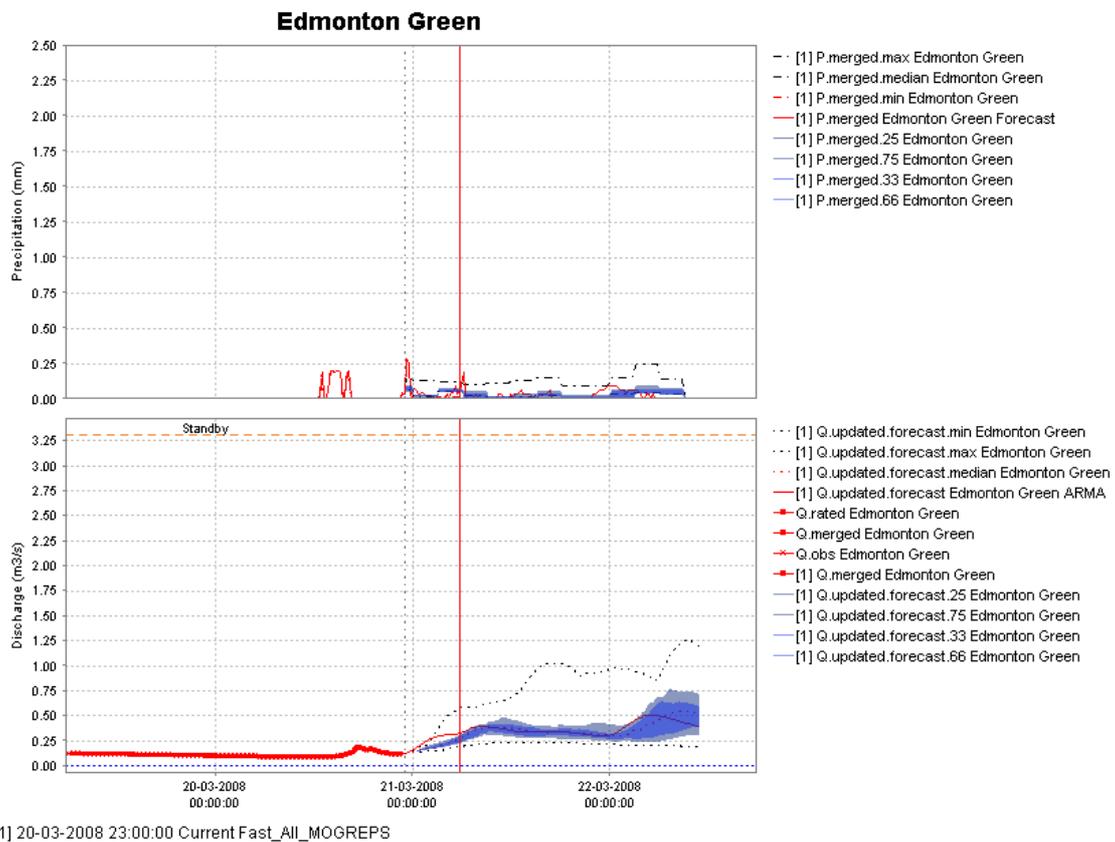


Figure 7.13 Forecast for 23:00 20 March for Edmonton Green.

The Wheatley catchment is a rather large catchment in the West Area of Thames Region. As can be seen from Figure 7.14 and Figure 7.15, the large delay in the catchment ensures that very little of the variation in precipitation input for this (small) event ends up in the updated discharge output. In this case, any forecasted precipitation has very little influence on the forecasted discharge. This is due to the delay factor used in the TCM configuration for this catchment.

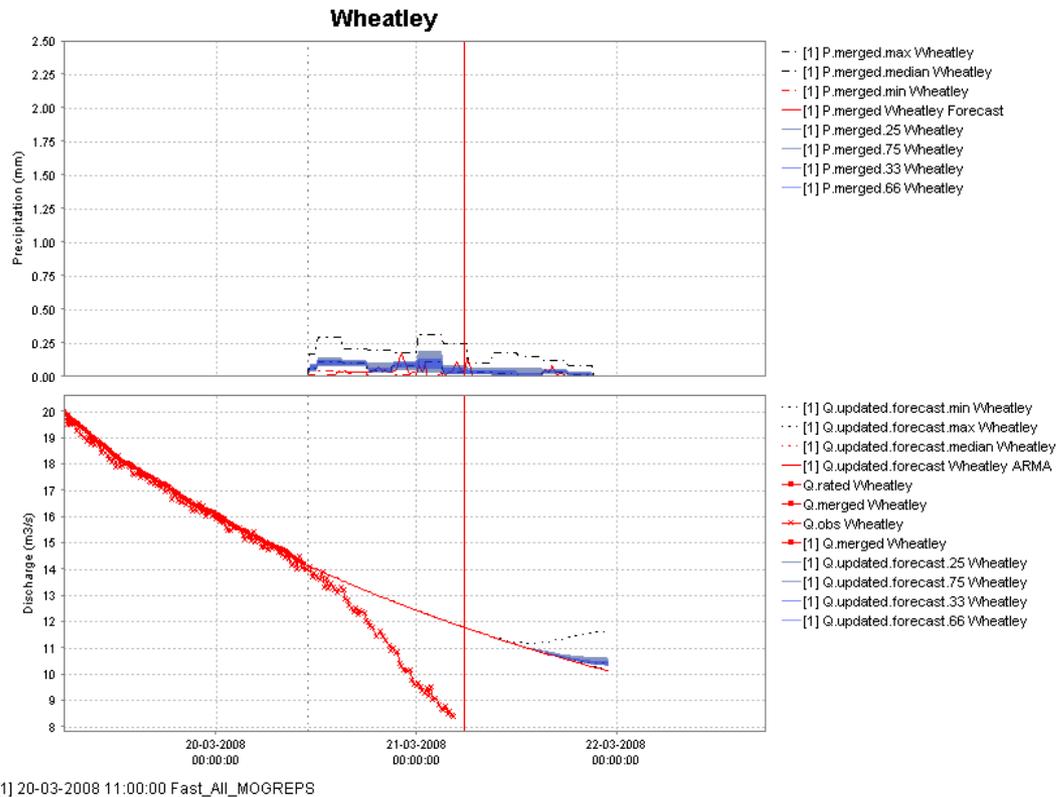
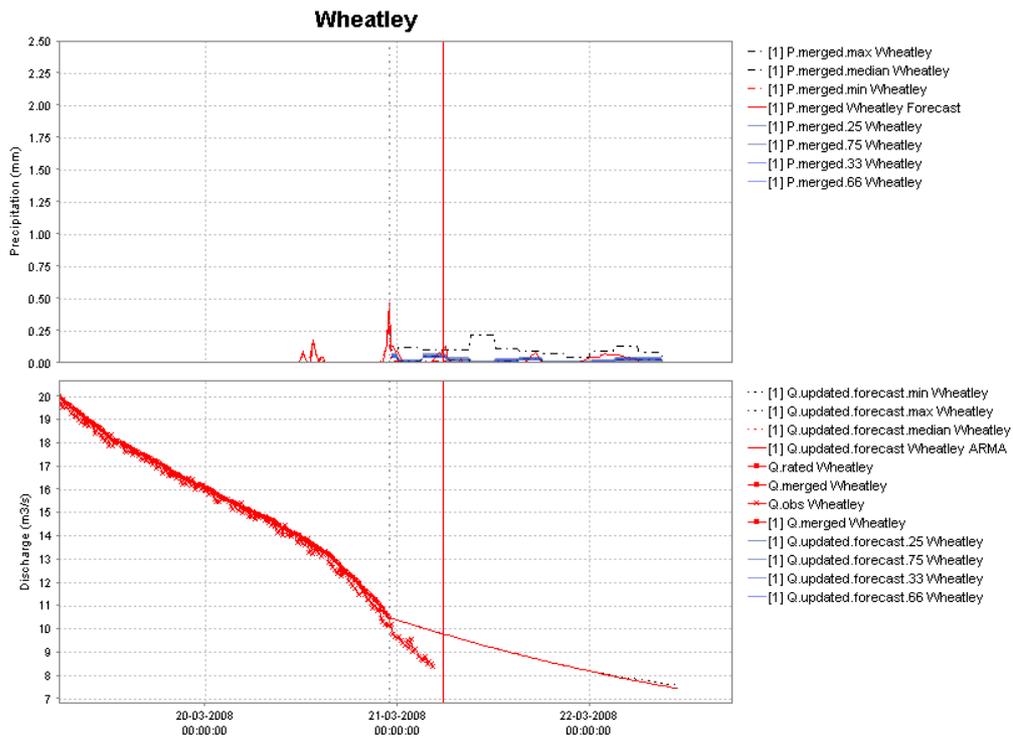


Figure 7.14 Forecast for 11:00 20 March for Wheatley.



[1] 20-03-2008 23:00:00 Current Fast_All_MOGREPS

Figure 7.15 Forecast for 23:00 20 March for Wheatley.

7.5.3 Brief analysis of an event in January 2008

Hindcasts were recreated by manually re-running the system for the period 11–16 January during which the UK experienced a number of minor precipitation events (Figure 7.16). Example outputs from this period are taken from North East and Thames Regions.

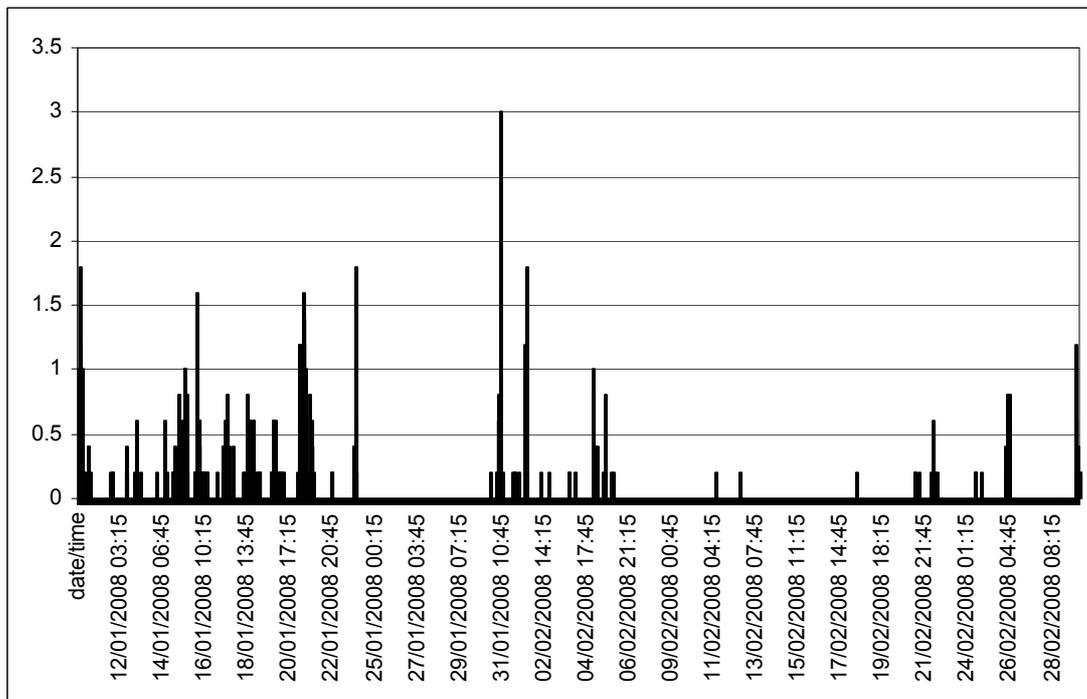


Figure 7.16 15-minute rainfall accumulations at Farnley Hall, West Yorkshire.

In an operational context, the use of ensembles can potentially provide a number of advantages. First it can provide an early warning of an extreme event, i.e. a low probability high impact event. By analysing the impact of the 'worst case' precipitation scenario, forecasters can feel more confident of the likelihood of a major flood. However, not all uncertainties are represented, e.g. uncertainties inherent in hydrological and hydrodynamic modelling. Thus, knowledge of the quality of the actual model may also be needed to interpret the results.

Figure 7.17 shows an example where the use of ensembles adds value for the forecaster. In this case, the deterministic forecast shows no catchment response while the range derived from the ensemble members indicates a much larger range of possibilities. In this case, the realised flow (parameter Q.rated) was indeed predicted more accurately by the 50th percentile flow than the deterministic (shown as parameter Q.updated.forecast).

It is important to note that the statistical functions applied to the time series (blue bands) are per time step and not per ensemble. This may have the effect of exaggerating the extreme events as the 'max' forecast does not constitute the forecast from one ensemble member.

The propagation of such 'statistical' time series (to, for example, a hydrodynamic model) can also be considered in the context of the physical meaning of the modelled process since calibration is not possible for all possible outcomes.

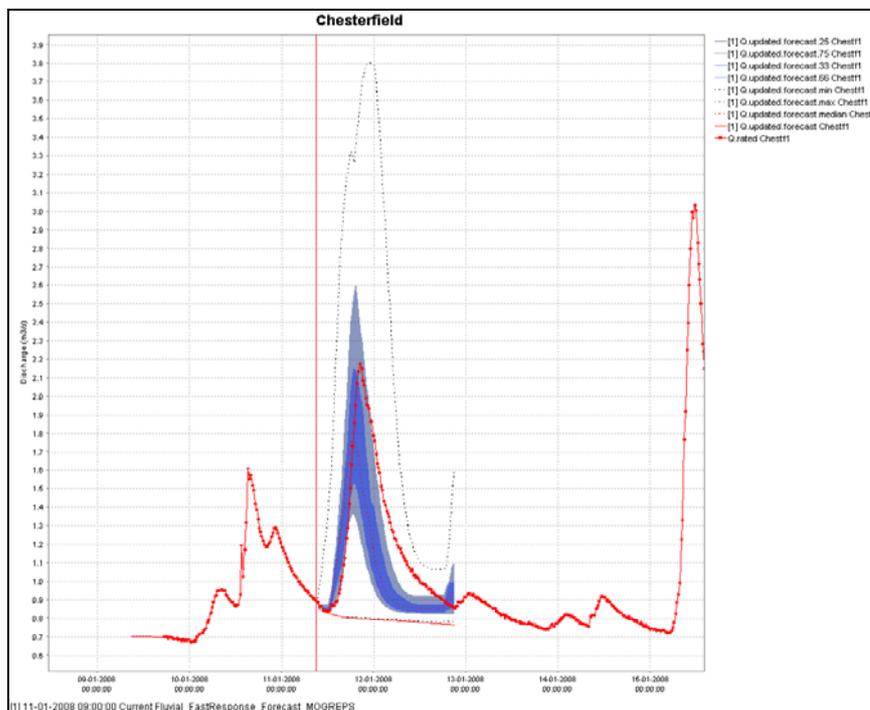


Figure 7.17 Example 1 – added value: Chesterfield T0 09:00 11 January 2008.

The example shown in Figure 7.18 highlights the sometimes large variation between the flow in the 75th percentile and the maximum value per time step. This range and behaviour depends on the sensitivity of the hydrological models to precipitation and the size of the catchment, and will vary per forecasting location. The inference is that, if warnings are given on a 'worst case' scenario' basis, a large number of false warnings is likely.

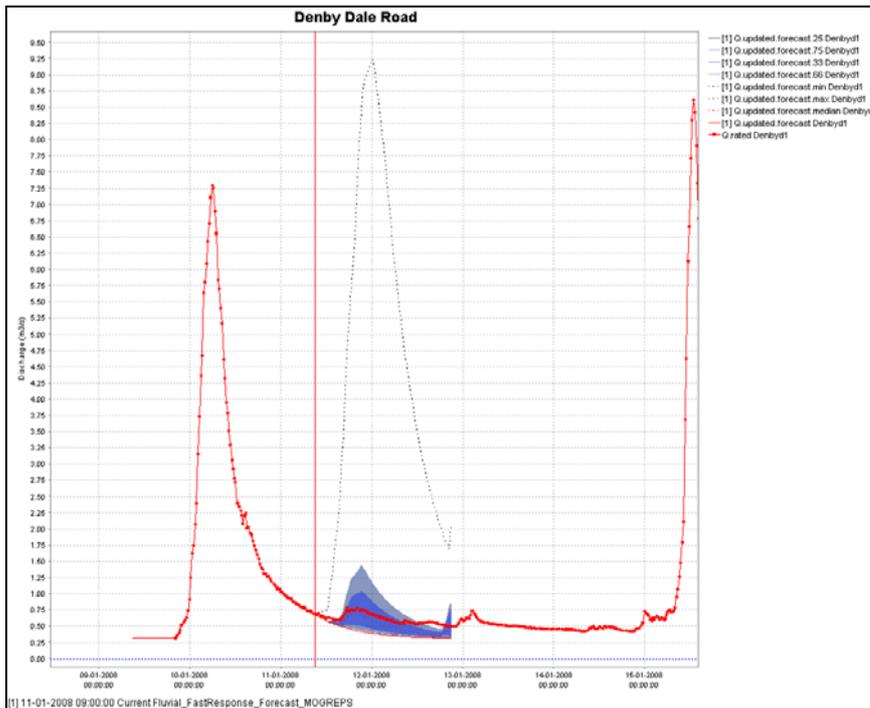


Figure 7.18 Example 2 – Low probability extreme event: Denby Dale Road T0: 09:00 11 January 2008.

In order to create a good forecast, the model must be well calibrated and assimilated with sufficient data. In the case of a hydrological model, aspects such as temperature and soil moisture must be representative of actual antecedent conditions. The example shown in Figure 7.19 highlights the fact that, although a range of rainfall possibilities was included, the model seems to be more sensitive to rainfall than in reality. In this case, the initial conditions at the start of the run are too wet, resulting in an overestimation of the baseflow component and a large sensitivity to precipitation input.

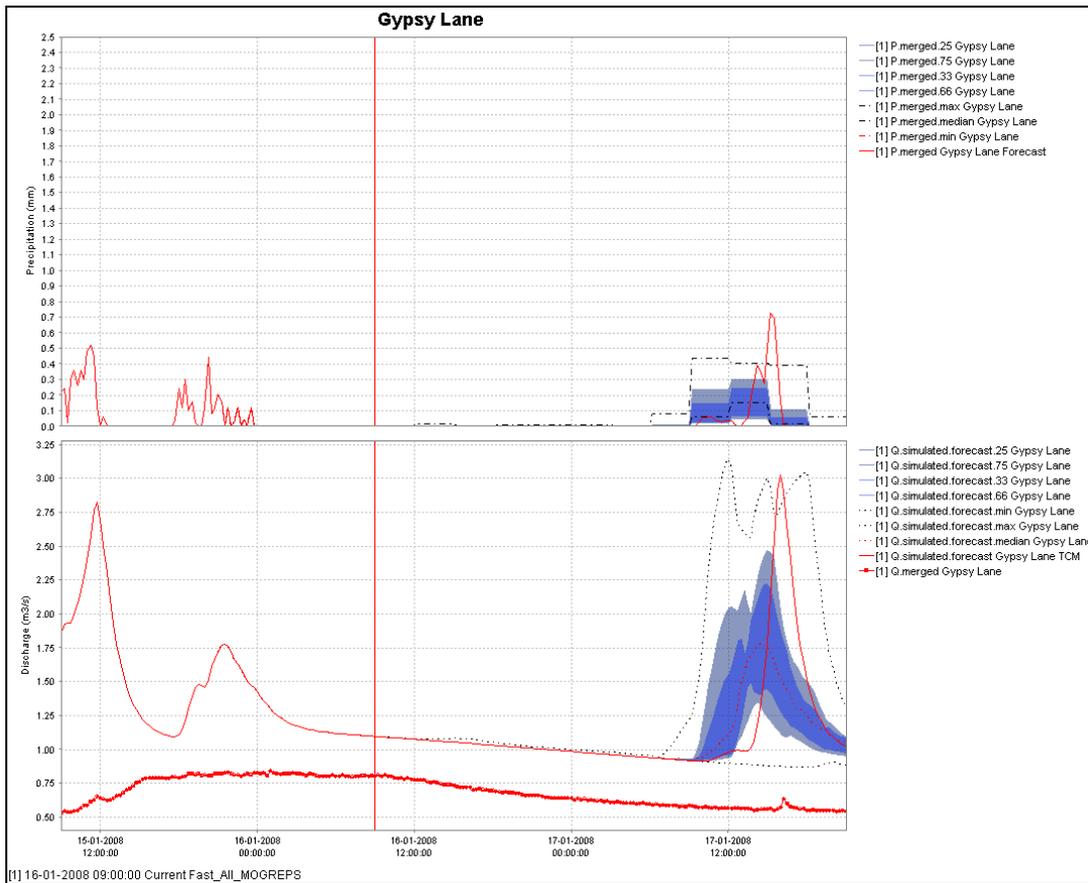


Figure 7.19 Example 3 – hydrological model, initial conditions: Gypsy Lane T0 09:00 15 January 2008.

7.6 System performance

Running the models in ensemble mode for the MOGREPS ensemble members will significantly impact the forecast run times because the forecast workflow has to be repeated 24 times. This section discusses how ensemble forecasts can be used practically.

The total performance of the system is governed by the following factors:

- system hardware components;
- forecast run times (run times of internal and external modules);
- database performance (and size);
- amount of data synchronised and network performance.

7.6.1 System specifications and set-up

Table 7.13 lists the specifications of the test system while Figure 7.20 provides an overview on all components in the test system (also known as FHSnet).

Table 7.13 System specifications used in the test system.

Component	Hardware
FSS ¹	AMD Dual core 2.19 Ghz, 3Gb RAM
DATABASE server	AMD Dual core 2.19 Ghz, 3Gb RAM
MC server ²	AMD Dual core 2.19 Ghz, 3Gb RAM

Notes: ¹ forecasting shell server
² master controller

All internal network connections are 100Mb copper.

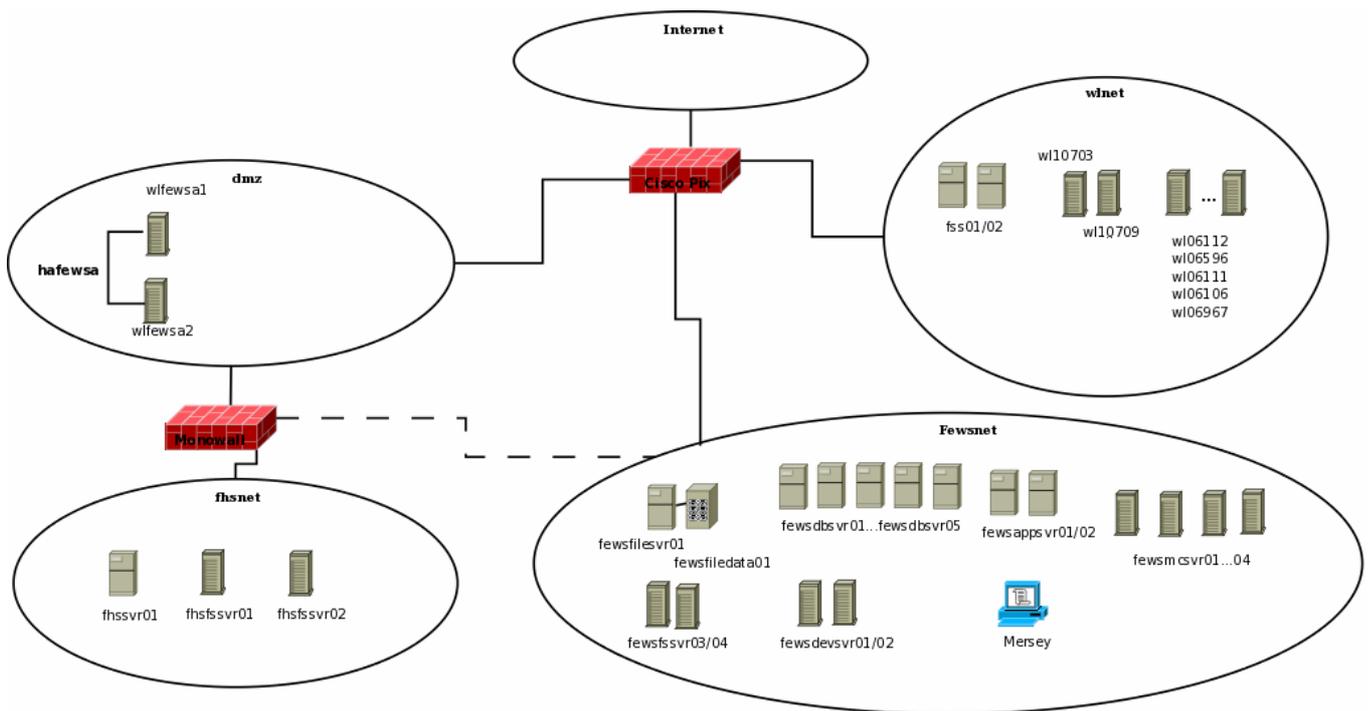


Figure 7.20 Components of the FHSnet Delft-FEWS test system.

7.6.2 Forecast run times (run times of internal and external modules)

In a distributed system such as Delft-FEWS, the total forecast run time is made up of separate components. This section deals only with the time it takes the forecast to run on the FSS after initial synchronisation has finished up to where outgoing synchronisation begins.

The times given in this section are based on the hardware used on the test system and actual numbers will be different with the use of other hardware. Average run times for the Thames and North East Region test systems are shown in Table 7.14.

Table 7.14 Run times of the ensemble forecast in the Thames and North East Region test systems.

	Run time
Thames	
Fast_All_MOGREPS	7 minutes
Fast_All	30 seconds
NorthEast	
Fluvial_FastResponse_Forecast_MOGREPS	16 minutes
Fluvial_FastResponse_Forecast	1 minute

MOGREPS forecasts are currently produced twice a day so the run times shown in Table 7.14 are acceptable; although longer than the normal runs, forecasts can be delivered in a timely manner. However, the picture will change if the ensemble runs are not limited to fast-responding catchments and, for example, also include the ISIS models. The estimated run times for this situation are shown in Table 7.15.

Table 7.15 Estimated run times for the complete region forecast in ensemble mode ¹.

Thames		North East	
Fluvial forecast	00:06:04	Fluvial_FastResponse_Forecast	00:03:41
ISIS_TThames_Forecast	00:00:41	RiverFlow_Forecast	00:08:09
Estimated ensemble run-time (Worst case= * 24)			
Fluvial forecast	02:24:00	Fluvial_FastResponse_Forecast	01:28:00
ISIS_TThames_Forecast	00:27:00	RiverFlow_Forecast	03:26:00

Notes: ¹ Normal run times are taken from the Environment Agency's online system.

It is clear from Table 7.15 that some speeding up of the forecasts would be required if the full forecast was run in ensemble mode on the current online Environment Agency system. There would also need a speed increase if the ensemble runs of just the fast-responding workflows are to be run more frequently.

There are a number of options for speeding up the ensemble runs.

- Optimise external models.
- Optimise data exchange.
- Install faster central processing units (CPUs) (cores) on each FSS.
- Run ensemble members in parallel:
 - Split up workflows (e.g. members 1–12 on FSS01 and 13–24 on FSS02).
 - Run the actual models at the GA level on a grid engine.
 - Improve the Delft-FEWS core to split (parts of) a workflow in several threads on multi-CPU/core FSSs.

Optimise external models

There is not usually much room for speed improvement in hydrological models as their equations can usually be solved analytically at great speed. A hydrodynamic model that was originally developed for a study that required high accuracy – and was later used for forecasting – could be adjusted to sacrifice some accuracy for speed. Alternatively, a hydrodynamic model could be replaced by a simple hydraulic routing model in the ensemble runs for which less accuracy may be acceptable.

Optimise data exchange

For fast running external models that require a significant amount of data, the exchange via PI XML only could be replaced by the binary version of the PI XML (XML header with binary payload) to speed up the file reading and writing. This may improve performance significantly but could require an update to the adapters involved. At present, only a selected number of adapters support this feature.

Install faster CPUs (cores) on FSSs

An easy gain (in terms of work needed for configuration, etc.) might be obtained by upgrading the CPUs on the FSSs. In general, a new generation of CPUs provides twice the performance (at the same price) every two years.¹⁵ Replacing a two-year-old FSS with a new one could obtain a theoretical speed increase of 100 per cent. This is for CPU speed only; the speed of hard drives increases at a slower rate.

Run ensemble members in parallel

- **Split up workflows (e.g. members 1–12 on FSS01 and 13–24 on FSS02).** Assuming each FSS has its own dedicated CPU, this would speed run times considerably. However, this would be a significant change in configuration and very hard to maintain.
- **Run the actual models at the GA level on a grid engine.** If most of the run time is within the actual models this option might be very efficient. It does require the set up (and maintenance) of a grid engine but this could be shared between Regions. Both the Sun grid engine (<http://gridengine.sunsource.net/>) and Condor high throughput computing (<http://www.cs.wisc.edu/condor/>) have been tested successfully with FEWS.
- **Improve Delft-FEWS core to split (parts of) a workflow in several threads on multi-CPU/core FSSs.** Most new machines today come with dual core CPUs. While these can be used, for example, to run two FSSs on a single box (one on each core), it does not speed up the individual runs. By improving the Delft-FEWS code, separate ensembles could be started in separate threads and run in parallel within a single FSS instance. This option has several advantages:
 - it does not require any configuration (or very little);
 - it can be used in combination with all the other options.

¹⁵ See http://en.wikipedia.org/wiki/Moore%27s_law for more information.

7.6.3 Database size

Because the MOGREPS data feed was not available at the time of this analysis, it was not possible to determine the effect of the size of the operational database. This will be undertaken in Phase 3.

A number of SQL queries have been developed to determine the size of the time series table of the Delft-FEWS central database and these will be used once the MOGREPS data feed has been restarted. The proposed procedure is presented in Appendix D.

7.6.4 Synchronisation times and network load

Because the MOGREPS data feed was not available at the time of this analysis, it was not possible to determine the sync time in an operational set-up. This will be undertaken in Phase 3.

8 Conclusions

Phase 2 of the project involved extensive modelling work across a large region of south-west England using a lumped model (the PDM model) and two distributed models – the physical–conceptual G2G model and the physics-based REW model.

The brief set of conclusions from this modelling investigation set out below serves to highlight the main outcomes. Recommendations and proposed work under Phase 3 are set out in Section 9.

8.1 Models

Model performance is summarised in Table 8.1.

Table 8.1 Summary of model performance.

Model	Performance
PDM	<ul style="list-style-type: none">• Excellent performance across catchments – R^2 efficiency 0.82 to 0.92 for both calibration and evaluation periods.• Simple lumped model and effective state correction, but insensitive to storm pattern.
G2G	<ul style="list-style-type: none">• Good performance across catchments – R^2 efficiency 0.64 [0.71] to 0.87 for both calibration and evaluation periods.• Ungauged performance good (0.64–0.75).• Camel improved by ‘at site’ calibration (to 0.82).
REW	<ul style="list-style-type: none">• Good performance for winter periods but overall performance only reasonable with R^2 around 0.58 for both calibration and evaluation periods.• Ungauged performance on a par with gauged performance. This shows some of the possible strengths of a distributed model for operational applications. Even with uniform parameters throughout the catchments, the results for internal (ungauged) sites may be useful to a forecaster.• Camel performance poor with current calibration.

The attributes of the G2G model relevant to convective scale probabilistic flood forecasting include:

- Area-wide model formulation well suited to ungauged problem;
- reasonably easy and quick to calibrate;
- simple physical–conceptual form is fast to run and thus well suited to ensemble application:
 - takes less than 10 minutes for England and Wales for two-day forecast;
 - distributed form sensitive to storm cell location;
- prospect of more stable calibration across Environment Agency Regions using extended soil/geology formulation (suggested by Camel results);
- simple, fast and effective state correction.

More general conclusions and practical implications relating to the potential use of distributed models such as the G2G and REW for operational flood forecasting and warning are summarised in Table 8.2.

Table 8.2 Summary of general conclusions.

Aspect	Comments
Advantages of distributed models	<ul style="list-style-type: none"> • Sensitive to spatio-temporal structure of storms. • Helpful in understanding storm and catchment shaping of flash floods, e.g. relative orientation of storm and river network.
Benefits of distributed modelling	<ul style="list-style-type: none"> • Identify locations vulnerable to flooding. • Help forecast floods shaped by 'unusual' storm and catchment conditions absent from the historical record. • Provide a complete spatial picture of flood hazard across a region. • Respond sensibly to ensemble rainfall forecasts that vary in position.
Practical implications	Lumped modelling for a given location: <ul style="list-style-type: none"> • can provide good flood forecasts in typical conditions; • is hard to better when calibrated to gauged catchments.

8.2 Implementation of MOGREPS ensembles

Ensemble runs of catchment models within Thames and North East Regions can be configured within NFFS with relatively limited effort. Run times are acceptable with current hardware if the ISIS model runs are not run in ensemble mode.

The display of a bandwidth of possible forecast outcomes together with the deterministic forecast provides useful extra information to the forecaster.

The number of events within the period the system ran in Phase 2 was too small to give any indication on the relative performance of the MOGREPS ensemble forecasts.

9 Recommendations

The recommendations following on from the results and conclusions of Phase 2 need to underpin a plan for work in Phase 3, leading finally to improvements in operational practice.

The recommendations concern three main issues:

- **Hydrological modelling concepts to be carried through to Phase 3:**
 - distributed G2G model shows promise for Area-wide flood forecasting at gauged and ungauged locations;
 - extended G2G formulation incorporating soil/geology datasets should be considered.
- **Case study selection for Phase 3:**
 - summer 2007 floods, with a focus on Midlands Region;
 - hydrometric data collation;
 - high resolution NWP data.
- **Continuation of the MOGREPS trial.** As discussed during the Phase 2 completion workshop, the MOGREPS trial will be continued in Phase 3. This will allow:
 - more events to be captured;
 - additional forecast stability tables to be added to the test configuration;
 - a wider range of Environment Agency staff to assess the results.

A more radical recommendation – stemming from the success of the G2G model for Area-wide forecasting in Phase 2 – is to trial the G2G model countrywide across England and Wales in Phase 3.

More detailed analyses – including use of high resolution NWP pseudo ensembles – should be performed for the selected case study area in the Midlands Region.

Section 10 provides a more detailed proposed plan for Phase 3 work.

10 Proposed plan of work for Phase 3 (modelling)

Phase 2 demonstrated that a distributed hydrological model (set up using a digital terrain model) can be operated on the Environment Agency's National Flood Forecasting System platform with short enough run times for use in real-time forecasting.

This phase has also coupled the latest Met Office high resolution NWP products with the CEH distributed hydrological model and considered the future potential of ensemble convective scale rainfall predictions. The term 'distributed forecasting' in this sense means the use of a spatially distributed (grid-based) hydrological model to forecast 'everywhere'. This contrasts with current hydrological model networks that comprise a connected set of (normally) lumped rainfall-runoff ('catchment') models feeding into hydrological and hydrodynamic river models, which provide forecasts only at specific locations.

Recent work at CEH outside this project, has:

- configured the G2G model across England and Wales as 'proof-of-concept';
- extended the terrain-based formulation to include soil/geology and land cover datasets (as research code) and a preliminary calibration obtained across Britain.

Work under Phase 3 should focus on the trial implementation and assessment of the G2G model across England and Wales, including its use for probability forecasting. The model will be used to provide risk-based forecasts of future flooding in a countrywide spatial context.

Against the background of work under Phase 2 and recent work at CEH, it is proposed that Phase 3 focuses on a trial implementation and assessment of the G2G model across England and Wales. Specifically it is proposed to:

- implement the existing terrain-based G2G module adapter within the NFFS (configured across England and Wales) as a trial system and assess performance;
- extend the G2G module adapter code to incorporate the research code formulation that uses soil/geology/land cover datasets;
- trial and assess performance;
- make recommendations for stakeholder trials in 2009 including capability:
 - to provide forecasts for ungauged and rapid response catchments;
 - for input to (hydrological and hydrodynamic) flow routing models of downstream river reaches.

10.1 Case study outline

The case study trial will employ weather radar and hydrometric data from 2007 with a focus on the summer 2007 floods. The performance of the G2G model will be assessed at a selection of river gauging station sites across the Environment Agency Regions.

This 'countrywide assessment' will initially employ weather radar estimates of rainfall only and assume 'perfect foreknowledge' of rainfall in the spirit of a 'proof-of-concept' and test of

NFFS infrastructure. Subsequently, the feasibility of using an improved merged raingauge/radar estimate of rainfall (using HyradK) would be investigated along with introduction of the extended G2G model.

It is envisaged that eight pairs of river gauging sites will be used to support the countrywide assessment of the G2G model to forecast at gauged and 'ungauged' locations.

A more focused 'regional assessment' will take an area within the Midlands Region affected by the summer 2007 floods and utilise raingauge data in combination with radar data for improved rainfall estimation as model input.

The summer 2007 storms over the Midlands Region exhibited significant spatial variability in rainfall due to embedded convective activity and will provide an important test of the G2G distributed model structure. As part of this regional assessment, trials will be made of the use of pseudo ensemble high resolution NWP forecast rainfalls as input to the G2G model to produce probabilistic flood forecasts.

The assessments will make use of conventional forecasts from lumped rainfall-runoff models as a benchmark reference.

10.2 Principal operational output

The G2G model will be configured to operate within the NFFS (Delft-FEWS) environment and to provide forecast coverage on a 1 km grid across England and Wales. Subject to the recommendations on operational use made at the end of Phase 3, this configuration will be available for Environment Agency and stakeholder trials in 2009. Further details of the proposed Phase 3 work, including a task breakdown, are contained in the revised Project Plan.

10.3 MOGREPS trial

The current test system, including the remote access to the system, will be maintained for the duration of the project. The configuration will be extended with extra tables indicating the stability of the ensemble forecasts.

An extra master controller will be set up in the system to run the countrywide model.

References

- Bell V A and Moore R J, 1998a A grid-based distributed flood forecasting model for use with weather radar data: Part 1. Formulation. *Hydrology and Earth System Sciences*, 2(2-3), 265-281.
- Bell V A and Moore R J, 1998b A grid-based distributed flood forecasting model for use with weather radar data: Part 2. Case studies. *Hydrology and Earth System Sciences*, 2(2-3), 283-298.
- 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.
- Boorman D B, Hollis J M and Lilly A, 1995 *Hydrology of soil types: a hydrologically based classification of the soils of the United Kingdom*. IH Report No. 126. Wallingford: Institute of Hydrology.
- Box G E P and Cox D R, 1964 An analysis of transformations. *Journal of the Royal Statistical Society (Series B)*, 26, 211-252.
- Broersen P M T, 2000 Finite sample criteria for autoregressive order selection. *IEEE Transactions on Signal Processing*, 48, 3550-3558.
- Broersen P M T and Wensink H E, 1996 On the penalty factor for autoregressive order selection in finite samples, *IEEE Transactions on Signal Processing*, 44, 748-752.
- Broersen P M T and Wensink H E, 1998 Autoregressive model order selection by a finite sample estimator for the Kullback-Leibler discrepancy. *IEEE Transactions on Signal Processing*, 46, 2058-2061.
- Burg J P, 1967 *Maximum entropy spectral analysis*. In Proceedings of the 37th Meeting of the Society for Exploration Geophysicists (Oklahoma City, 1967), 1-6.
- CEH Wallingford, 2005a *PDM Rainfall-Runoff Model. Version 2.2*. Wallingford: Centre for Ecology & Hydrology [Includes guide, practical user guides, user manual and training exercises].
- CEH Wallingford, 2005b *PSM Rainfall-Runoff Model. Version 2.2*. Wallingford: Centre for Ecology & Hydrology [Includes guide, practical user guides, user manual and training exercises].
- Clarke R T, 2008 A critique of present procedures used to compare performance of rainfall-runoff models. *Journal of Hydrology*, 352, 379-387.
- Cross H, 1936 *Analysis of flow in networks of conduits or conductors*. Bulletin 286. Urbana, IL: University of Illinois. Available from: <http://www.ideals.uiuc.edu/bitstream/handle/2142/4433/engineeringexperv00000i00286.pdf?sequence=3> [Accessed 7 October 2009].
- Golding B, 2005 Editor *Boscastle and North Cornwall post flood event study – meteorological analysis of the conditions leading to flooding on 16 August 2004*. Forecasting Research Technical Report No. 459. Bracknell: Met Office.
- Greenfield B J, 1984 *The Thames Water catchment model*. Internal report. Thames Water Technology and Development Division.
- Hough M, Palmer S, Weir A, Lee M and Barrie I, 1997 *The Meteorological Office Rainfall and Evaporation Calculation System: MORECS version 2.0 (1995)*. An update to Hydrological Memorandum 45. Bracknell: Met Office.

- HR Wallingford, 2005 *Flooding in Boscastle and North Cornwall, August 2004. Phase 2 studies report*. Wallingford: HR Wallingford.
- Inoue H, 1986. A least-squares smooth fitting for irregularly spaced data: finite-element approach using the cubic B-spline basis. *Geophysics*, 51(11), 2051–2066.
- Jones D A and Moore R J, 1980 *A simple channel flow routing model for real-time use*. In Hydrological Forecasting, Proceedings of the Oxford Symposium, April 1980, 397–408. IAHS Publication No. 129. Wallingford: IAHS Press. Available from: <http://iahs.info/redbooks/129.htm> [Accessed 7 October 2009].
- Kay S M and Marple S L, 1981 Spectrum analysis—A modern perspective. *Proceedings of the IEEE*, 69, 1380–1419.
- 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 Jones D A, 1978 *An adaptive finite-difference approach to real-time channel flow routing*. In *Modelling, Identification and Control in Environmental Systems* (ed. G C Vansteenkiste), pp. 153–170. Amsterdam: North-Holland.
- Moore R J and Bell V A, 2001 *Comparison of rainfall-runoff models for flood forecasting. Part 1: literature review of models*. Environment Agency R&D Technical Report W241, Bristol: Environment Agency.
- Moore R J and Bell V A, 2002 Incorporation of groundwater losses and well level data in rainfall-runoff models illustrated using the PDM. *Hydrology and Earth System Sciences*, 6(1), 25–38.
- Moore R J, Bell V A, 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 and 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 and Jones D A, 2007 *Rainfall-runoff and other modelling for ungauged/low-benefit locations*. Science Report SC030227/SR1. Bristol: Environment Agency. Available from: <http://publications.environment-agency.gov.uk/pdf/SCHO0307BMER-e-e.pdf> [Accessed 16 September 2009].
- Morris D G and Flavin R W, 1990 *A digital terrain model for hydrology*. In *Proceedings of the Fourth International Symposium on Spatial Data Handling (Zürich, 1990)*, Volume 1, 250–262. Dordrecht, The Netherlands: Springer.
- 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.
- Paz A R, Collischonn W and Lopes da Silveira A L, 2006 Improvements in large-scale drainage networks derived from digital elevation models. *Water Resources Research*, 42(8), W08502, doi:10.1029/2005WR004544.

Priestley M B, 1981 *Spectral Analysis and Time Series*. New York: Academic Press.

Roberts N, 2006 *Simulations of extreme rainfall events using the Unified Model with a grid spacing of 12, 4 and 1 km*. Met Office Forecasting Research Technical Report No. 486. Bracknell: Met Office.

Ross P J, 2003 Modelling soil water and solute transport—fast, simplified numerical solutions, *Agronomy Journal*, 95, 1352–1361.

Wilby R, Greenfield B and Glenny C, 1994 A coupled synoptic hydrological model for climate change impact assessment. *Journal of Hydrology*, 153, 265–290.

List of abbreviations

1D	one-dimension
2D	two-dimensional
3D	three-dimensional
AR	autoregressive
ARMA	autoregressive moving average
CEH	Centre for Ecology & Hydrology
CIC	combined information criterion
CPU	central processing unit
DTM	Digital Terrain Map/Model
FEM	finite element method
FEWS	Flood Early Warning System [Deltares]
FSS	forecasting shell server
G2G	Grid-to-Grid [model]
GA	general adapter [Delft-FEWS]
GIS	geographic information system
GMT	Greenwich Mean Time
HOF	Hortonian overland flow
HOST	Hydrology of Soil Types
IHDTM	Integrated Hydrological Digital Terrain Model
ISE	Isolated Event Model
JCMM	Joint Centre for Mesoscale Meteorology [Met Office]
MA	moving average
MC	master controller [server; Delft-FEWS]
MCRM	modified conductive rock matrix
MOGREPS	Met Office Global and Regional Ensemble Prediction System
NAM	North American Mesoscale Model
NFFS	National Flood Forecasting System
NGR	National Grid Reference
NRFA	National River Flow Archive
NWP	Numerical Weather Prediction
OC	operating client
ODE	ordinary differential equation

PDE	partial differential equation
PDM	Probability Distribution Model
PE	potential evaporation
PRTF	Physically Realisable Transfer Function Model
PSM	Penman Store Model
REC	representative elementary column
REW	representative elementary watershed
SEPA	Scottish Environment Protection Agency
SDM	soil moisture deficit
SOF	saturation overland flow
STEPS	Short Term Ensemble Prediction System
SVAT	surface–vegetation–atmosphere transfer)
TCM	total catchment management
UN	Unified Model [Met Office]
UTC	Co-ordinated Universal Time
VPN	virtual private network

Appendix A: Report of the completion workshop

Day 1 – Phase 2 completion workshop

Doug Whitfield – Introduction

The meeting started with an overview of where this activity sits within the national science R&D programme and what we hoped to achieve during the workshop. Some progress towards understanding and managing the impact of low probability, high impact events (e.g. Boscastle) was seen as one possible useful outcome of the workshop.

Karel Heynert – Objectives of the workshop

Karel introduced the topics for presentation and discussion. How to make use of NWP ensembles and how to make use of the high-resolution forecasts were the key research questions which this study (and workshop) aims to answer.

Bob Moore – Convective scale rainfall

Bob presented the current products available from the Met Office 4 km NWP to T+36 hours; in 2009, 1.5 km resolution (to T+48 hours) is expected.

Higher resolution weather prediction was shown to have significant benefits for predicting convective storms over the current 12 km resolution with better representation of synoptic (large scale) and local effects (e.g. orographic effects) as well as improved predictions of the evolution of storms (e.g. storm tracks).

The fundamental unpredictability was highlighted with variable nature in terms of exact location and timing for convective storms. The aim is to produce a forecast which can approximate the sizes, intensities and direction of the storm. The predictability depends on the area of interest. For forecasting, the timing and accumulations over a geographical area are the most important.

The uncertainty can be dealt with through the use of ensembles; however for high resolutions NWP (1.5 km), this is currently too computationally expensive. The use of 'pseudo ensembles', which perturb high-resolution forecasts in space, may offer a way to quantify the uncertainty. The methodology used for perturbation is currently based on expert knowledge and should vary per event and meteorological phenomena.

Bob Moore, Steve Cole and Jaap Schellekens – Model concepts, calibration and evaluation

An overview of the catchments studied was given, highlighting their location and size as well as differences in geology and soils.

Steve gave an overview of the methodology used to adjust the radar observations using 'observed' precipitation from raingauges. Monthly MORECS potential evaporation data from the Met Office were used in the calibration of the hydrological models.

Bob gave a technical overview of the lumped conceptual model PDM and the physical-conceptual distributed G2G model.

G2G in its simplest form employs a relationship between terrain slope and the capacity to absorb rainfall, which is probability-distributed, to represent runoff production within each grid cell. (An extended form can employ soil properties in situations where this slope-based

representation does not suffice.) The model uses gridded rainfall estimates as input to each grid cell. Water is routed from grid cell to grid cell over the modelled domain, thus providing an Area-wide approach to flood forecasting. This means that it can predict various quantities (e.g. river flow, soil moisture and potentially flood risk) in all grid squares, whether gauged or ungauged. This contrasts with source-to-sink distributed models that route flows directly from each grid square to the catchment outlet of interest.

Hillslope or river grid cells are differentiated using a threshold drainage area approach and are assigned different velocities of travel. Flow propagation along surface and subsurface pathways from grid cell to grid cell employs a kinematic wave routing approach. The results of the calibration show good model simulation of flows for both calibration and verification periods (R^2 efficiency 0.64–0.87). The results are promising for flood forecasting in ungauged catchments. The model is quick and easy to calibrate, and includes effective state correction routines for forecast updating.

The PDM model is well-established in the UK and produces excellent results for both calibration and verification periods. An overview was given of the methodology and some minor ‘wetting up’ deviations explained. R^2 efficiency measures were in the range 0.82–0.92. The PDM model also has robust and effective state correction. However the model is not sensitive to storm pattern due to its lumped nature.

Lastly, Jaap Schellekens presented the calibration of the semi-distributed, physically based REW model. This model is based on sound physically based formulations such as Richard’s equation and includes a groundwater model. However, the calibration revealed some problems with recession in the Camel catchment. Further effort could improve the calibration further.

Bob Moore, Steve Cole and Jaap Schellekens – Forecasting the Boscastle event

All models produced significant underestimation of peak flows when hindcasting the Boscastle event of August 2004 using high-resolution NWP. Using pseudo-ensembles, PDM was the only model to encompass the peak flow generated during this event.

The discussion raised a number of interesting points about how the pseudo-ensembles were generated and whether this could be formalised to create a useful product. Also discussed was whether more point data could be assimilated into NWP predictions.

It was noted that the rate of rise experienced during the Boscastle event was extraordinary and none of the hydrological models was able to reproduce this behaviour sufficiently. It was speculated that this could be due to sheeting effects of flow with extraordinarily large precipitation events or due to soil compaction in the area.

The distributed models were not proven to perform better than lumped in predicting this event in simulation mode. However, the distributed models (both REW and G2G) showed a much wider response in discharge when the pseudo-ensembles were used compared with the PDM model. This indicated that the spatial variation in precipitation was better captured by these models.

Jaap Schellekens – Introduction to Phase 3 and further case studies

Further verification events for Phase 3 are needed to determine whether additional benefit can be gained through the use of distributed models using high-resolution NWP. The REW model will not be used further in Phase 3. The additional development needed to improve the calibration is thought to be beyond the scope of this project.

Several events were mentioned including:

- June and July 2007 events in the Midlands;
- Albrighton (Shropshire) in 2006;

- July 2006 in Todmorden.

The events of summer 2007 were seen to be an interesting (and high profile) case study during which a series of convective and frontal storms were generated. Further investigation will be carried out in the early part of Phase 3 to assess their suitability as a test location. Some concerns were raised that radar data might not be available for this period due to an outage.

It was concluded that the most important condition is the availability of high-resolution NWP forecasts for the event. This should be verified with the Met Office soon.

Jaap Schellekens – Results of ensemble test system

Jaap presented some interesting outputs from the ensemble test system (located in Delft). He highlighted the fact that the plume plots presented only represent some of the uncertainty in rainfall and do not allow for other sources of uncertainty. No hydrodynamic models were run due to the computational restrictions (runs would take 2–3 hours); however, he highlighted this could be relatively easily implemented given further computational capacity.

It was noted that the control run should be plotted separately on the time series display since the deterministic output is based on higher resolution data, i.e. it is not a like-for-like comparison. The computational burden of the nowcasting product from the Met Office – STEPS – was not tested in the scope of this project.

Day 2 – Probabilistic forecasting workshop

Day 2 brought together a wider audience of professionals from within the Environment Agency. Representatives from operational managers, policy-makers and forecasters were present to discuss broader issues of probabilistic forecasting.

Doug Whitfield gave an introduction to the day, again highlighting where this piece of science sits within the national R&D programme.

Marc Huband – Common sources of uncertainty

Marc introduced the common sources of uncertainty in flood forecasting highlighting that the uncertainty in rainfall prediction is only one source of uncertainty (albeit an important one) in a process which includes many uncertainties (e.g. in model parameters, high flow rating curves).

He presented examples of how uncertainty varies within a catchment or along a coastal reach, and made the point that the overall model development process should include steps to reduce, correct (via updating) and understand the main sources of uncertainty

Kevin Sene – Recent international developments

Kevin presented a brief introduction to the historical background to probabilistic flood forecasting and its development since the 1980s in the USA. Examples included the European Flood Alert System (EFAS) and international collaborations within the HEPEX and COST-731 programmes. In the UK and Europe, there are currently three large academic research programmes – FRMRC (Flood Risk Management Research Consortium), FREE (Flood Risk for Extreme Events) and FloodSite – all of which include some research components related to probabilistic flood forecasting.

Marc Huband – Some international examples

Marc gave an overview of some current examples of probabilistic flood forecasting techniques used worldwide or under development drawing on examples for river forecasting

from the National Weather Service in the USA and SYKE in Finland and for coastal surge forecasting from the National Hurricane Centre in the USA.

Karel Heynert – Probabilistic Flood Forecasting with NFFS

Karel presented some examples from the current ensemble test system in Delft highlighting the current potential for probabilistic forecasting within the Environment Agency. He also gave a general introduction to some reasons for adopting a probabilistic approach and the issues to consider.

Nigel Outhwaite – Pilot study in Thames Region – first impressions

Nigel presented outputs from the pilot conducted within Thames Region. His initial impressions were that the percentile plots were useful and that it was good to see the deterministic forecast as a comparison. The spaghetti plots were also seen to be useful since they give an idea of how many forecasts might cross a particular threshold.

There was some confusion over the interpretation of the percentile plots and the difference in NWP prediction grid scale between MOGREPS and the deterministic NWP forecast, which produced forecasts that did not necessarily match.

He questioned the usefulness over short lead times (i.e. until STEPS is available) but thought that the longer term forecasts currently available might be a useful planning tool for staff resources and mobilisation. Extensions of the system to include hydrodynamic models and of the MOGREPS lead time (to say T+48 hours) were seen to be key areas for improvement.

The blending project to merge STEPS and MOGREPS will allow more accurate forecasting at short lead times, although this is not due to complete until 2010.

Kevin Sene – Pilot study in North East Region – first impressions

On behalf of Andy Lane from the North East regional forecasting team, Kevin presented the current methodologies used in the region to quantify uncertainty. These include 'what if' scenarios for rainfall conducted in consultation with Met Office forecasters, worst-case scenario modelling, comparison with historical events, comparisons with level-to-level correlations, and scenarios for washland operations. The use of summer/winter calibration comparisons has also been considered (although is not used at present).

Kevin also discussed some possible applications of a probabilistic approach suggested by Area and Regional staff during the consultations performed during the Probabilistic Flood Forecasting Scoping Study. Some specific examples for North East Region were also discussed, including examples from an event in January 2008. The general examples included:

- early warning for operational staff;
- operation of demountable defences and washlands;
- potential improvements for forecasts in urban, fast response and ungauged catchments.

Another key potential benefit was better understanding of how models were performing for input to future model development and data improvement programmes. Several examples were also provided of potential uses for professional partners and in complex real-time control examples.

Discussion (focus on forecasting)

Some comments made in a wide ranging discussion included:

- Several ways of interpreting and post-processing probabilistic outputs were discussed including the visualisation tools being developed under the coastal flood forecasting R&D project.
- In some Regions, there are high expectations at senior level of the potential of probabilistic forecasts and a probability-based warning system. There is a need to keep key decision-makers informed about the development of probabilistic forecasting so that decisions about the future of the Environment Agency's warning service are based on sound and realistic information.
- Some concerns were raised about the level of time and skill required by duty officers to interpret forecasts and over the potential additional workload that implementation of a probabilistic-based warning process would bring; however, others thought this was not likely to pose a serious burden. The need for training and other guidance was discussed.
- Some forecasters were surprised that the Met Office considers each ensemble to have equal statistical value and perhaps had not appreciated the spread of inherent uncertainty in meteorological forecasting at longer lead times.
- The discussion also related to the new products which will become available from the Met Office (e.g. STEPS, outputs from the blending ensembles project) and that the Environment Agency should be ready to receive these new products as they become available through research and adaptation.
- Sources of uncertainty other than rainfall were also discussed. It was acknowledged that other uncertainties do exist, but that rainfall generates by far the greatest uncertainty for forecasting. The common view was to focus on the biggest uncertainty first.

Marc Huband – Practical exercise

The practical session involved the determination of a simple cost–loss function based on a scenario provided by the facilitators and was intended to generate discussion on how warnings could potentially be tailored to the needs of individual recipients according to their risk tolerance.

The exercise was based on a simple evacuation scenario, though it could equally have applied to the situation of installing a demountable barrier. Participants were divided up into five teams, with each team provided with a different probability threshold to consider.

The results of the exercise were then presented, followed by a general discussion on the assumptions in the analysis. This generated a good discussion on the cost associated with false alarms and hence the Environment Agency's reputation, the changing costs over time, and the associated responsibilities for estimating risk, etc.

Kevin Sene – Requirements for optimal use of probabilistic forecasts

Kevin presented an introduction to a range of approaches to interpretation of probabilistic forecasts including qualitative interpretation, threshold based measures, decision theory and decision support frameworks. The simpler approaches include visual interpretations of clustering and persistence, and have been shown to work well in some studies. Various map-based, graphical and tabulated formats were presented.

Drawing on techniques developed in meteorology and other fields (e.g. reservoir design), Kevin then discussed how cost–loss theory and utility functions can be used to provide a more objective approach to decision-making, based on a user's risk profile and the economic value of forecasts, and other factors such as tolerance to false alarms. These techniques, already well established in other fields, show potential for application to flood

warning situations, although this requires further research, particularly for extreme (rare) events.

Marc Huband – Emergency response/early mobilisation

Marc presented some ideas on the topic of whether being warned earlier but more frequently would assist in preparing for a possible flood. The presentation considered four key groups:

- Environment Agency flood forecasting staff;
- Environment Agency flood warning staff;
- professional partners;
- general public.

Examples included:

- planning of staff rotas;
- installation of demountable defences;
- operational response and emergency works;
- widespread/major events;
- installation of flood resilience measures.

Common themes requiring additional research and study were highlighted for each example including:

- communication of information;
- threshold setting;
- managing the perception of false alarms.

Kevin Sene – Real-time control

Kevin presented some potential applications of probabilistic forecasts in improving the real-time operation of structures within the Environment Agency for flood control. Examples included washlands, tidal barriers, reservoirs, and river regulators. The optimisation problems were outlined in each case including the need to consider multiple objectives. Three international examples of applying these techniques were then discussed from Italy, Taiwan and the USA.

Doug Whitfield – Discussion (the wider vision)

The day concluded with a wider discussion of developments in probabilistic flood forecasting within the Environment Agency and future plans. Topics covered included:

- policy;
- NFFS developments;
- STEPS;
- training;
- The Pitt Review;
- communication of uncertainty.

Appendix B: Raingauge data quality control

Table B1 15-minute raingauge totals in excess of 70 mm (all checked as erroneous and set as missing within HyradK processing).

Raingauge	Date	Time	15-minute accumulation	Raingauge	Date	Time	15-minute accumulation
Allet	30/04/2003	06:15:00	374.2	Luxulyan	02/03/2003	09:15:00	125
Allet	12/07/2003	06:15:00	557.6	Luxulyan	02/08/2003	11:00:00	1180
Allet	13/09/2003	06:15:00	674.2	Luxulyan	02/08/2003	11:30:00	103
Allet	04/11/2003	06:15:00	791.4	Luxulyan	02/08/2003	19:15:00	141
Bealsmill	22/06/2004	03:30:00	6142	Luxulyan	02/08/2003	20:45:00	446
Boscadjack	29/03/2003	06:15:00	1905.4	Luxulyan	03/08/2003	08:30:00	786
Boscadjack	30/04/2003	06:15:00	1960.8	Luxulyan	03/08/2003	11:15:00	264
Boscadjack	12/07/2003	06:15:00	2189.8	Luxulyan	04/08/2003	10:15:00	429
Boscadjack	09/09/2003	14:15:00	2313.6	Luxulyan	29/01/2004	11:30:00	5552
Boscadjack	29/11/2003	06:15:00	2517	Luxulyan	29/01/2004	12:15:00	279
Boscadjack	10/02/2004	13:15:00	2902.6	Luxulyan	29/01/2004	13:45:00	266
Boscadjack	05/05/2004	06:15:00	3107.2	Luxulyan	29/01/2004	14:00:00	343
Boscadjack	23/07/2004	07:15:00	3255.6	Luxulyan	29/01/2004	22:30:00	5534
Boscadjack	24/07/2004	07:15:00	3255.6	Luxulyan	29/01/2004	22:45:00	586
Boscadjack	25/07/2004	07:15:00	3255.6	Luxulyan	30/01/2004	06:15:00	422
Boscadjack	27/07/2004	07:15:00	3255.6	Luxulyan	30/01/2004	14:30:00	215
Boscadjack	28/07/2004	07:15:00	3255.6	Luxulyan	30/01/2004	15:15:00	96
Boscadjack	29/07/2004	07:15:00	3255.6	Luxulyan	30/01/2004	16:30:00	2236
Boscadjack	30/07/2004	07:15:00	3261.4	Luxulyan	30/01/2004	16:45:00	768
Boscadjack	01/08/2004	07:15:00	3261.4	Luxulyan	30/01/2004	19:45:00	1048
Boscadjack	02/08/2004	07:15:00	3262.2	Luxulyan	30/01/2004	20:30:00	237
Boscadjack	02/08/2004	11:30:00	3262.2	Luxulyan	31/01/2004	02:00:00	107
Bradworthy	25/11/2006	11:30:00	7361.6	Luxulyan	31/01/2004	02:45:00	133
Bradworthy	25/11/2006	13:45:00	*****	Luxulyan	31/01/2004	04:00:00	181
Bradworthy	25/11/2006	14:45:00	*****	Luxulyan	31/01/2004	04:45:00	108
Bridgerule	03/03/2004	07:00:00	466.5	Luxulyan	31/01/2004	05:15:00	98
Canworthy	11/02/2004	07:00:00	3307.5	Luxulyan	31/01/2004	13:00:00	362
Canworthy	13/06/2004	07:15:00	3545	Luxulyan	01/02/2004	04:15:00	122
Canworthy	14/06/2004	07:15:00	3545.5	Luxulyan	02/02/2004	02:30:00	80
Canworthy	15/06/2004	07:15:00	3545.5	Luxulyan	02/02/2004	13:00:00	161
Canworthy	16/06/2004	07:15:00	3545.5	Luxulyan	02/02/2004	13:30:00	295
Canworthy	17/06/2004	07:15:00	3545.5	Luxulyan	02/02/2004	13:45:00	470
Canworthy	18/06/2004	07:15:00	3545.5	Luxulyan	30/04/2006	21:00:00	7999.6
Canworthy	19/06/2004	07:15:00	3545.5	Mary Tavy	27/01/2004	09:00:00	846.6
Canworthy	20/06/2004	07:15:00	3551	Pillaton	23/04/2004	06:00:00	12274.2
Canworthy	21/06/2004	07:15:00	3556	Pillaton	02/02/2005	14:00:00	73.5
Canworthy	22/06/2004	07:15:00	3556	Slaughterbridge	13/09/2003	06:45:00	4724
Canworthy	23/06/2004	07:15:00	3593	Slaughterbridge	01/08/2006	08:15:00	8000.4
Canworthy	24/06/2004	07:15:00	3595.5	Sticklepath	26/11/2002	10:00:00	77
Canworthy	25/06/2004	07:15:00	3596	Sticklepath	15/11/2004	12:45:00	8100.8
Canworthy	30/06/2004	07:15:00	3605.5	Sticklepath	15/11/2004	15:45:00	379.6
Canworthy	01/07/2004	07:15:00	3607.5	Sticklepath	15/11/2004	16:00:00	1518.6
Canworthy	02/07/2004	07:15:00	3614	Sticklepath	20/12/2004	14:00:00	7999.6
Canworthy	03/07/2004	07:15:00	3623.5	Sticklepath	16/11/2006	11:15:00	9933.4
Canworthy	04/07/2004	07:15:00	3627	Sticklepath	16/11/2006	14:00:00	9998.4
Canworthy	05/07/2004	07:15:00	3627.5	Tamarstone	29/03/2003	06:45:00	2934.5
Canworthy	06/07/2004	07:15:00	3627.5	Tamarstone	12/07/2003	06:45:00	3127
Canworthy	07/07/2004	07:15:00	3627.5	Tamarstone	04/11/2003	06:45:00	3366.5

Canworthy	21/12/2004	09:15:00	7271.4	Tamarstone	01/01/2004	19:45:00	3551.5
Canworthy	20/02/2007	06:45:00	5847	Tamarstone	31/03/2004	06:45:00	3779.5
Canworthy	20/02/2007	07:00:00	4152	Tamarstone	14/04/2004	06:45:00	3797
Canworthy	20/02/2007	07:15:00	5847	Tamarstone	05/05/2004	06:45:00	3847
Canworthy	20/02/2007	09:00:00	4152	Tamarstone	15/07/2004	06:45:00	3959
Canworthy	20/02/2007	09:15:00	5847	Tamarstone	01/09/2004	07:15:00	4168
Canworthy	20/02/2007	09:45:00	4152	Tideford	17/03/2005	14:30:00	107.75
Canworthy	20/02/2007	10:00:00	5847	Problems at Tideford until (too many to list):			
Canworthy	20/02/2007	14:45:00	4152	Tideford	11/04/2005	08:30:00	139.62
Crowford	30/12/2004	13:15:00	19161.5	Tinhay	25/11/2003	11:15:00	3789
Crowford	27/07/2007	14:00:00	9979	Trebrownbridge	19/05/2003	16:30:00	1252.5
Gwills	30/04/2003	06:15:00	3121	Trebrownbridge	27/05/2003	07:45:00	1265
Gwills	12/07/2003	06:15:00	3280	Trebrownbridge	29/06/2003	07:45:00	1340.5
Gwills	13/09/2003	06:15:00	3413	Trebrownbridge	04/11/2003	07:00:00	1590
Gwills	02/11/2003	00:15:00	3492.5	Trebrownbridge	20/12/2003	10:30:00	1722
Gwills	04/11/2003	06:15:00	3506.5	Trebrownbridge	17/03/2004	11:45:00	2086.5
Gwills	06/11/2003	06:00:00	3508	Trebrownbridge	17/03/2004	12:45:00	2086.5
Gwills	30/01/2004	00:30:00	3754.5	Trebrownbridge	18/11/2004	16:30:00	2764
Gwills	11/02/2004	06:30:00	3801.5	Trebrownbridge	17/04/2005	20:15:00	16851.5
Gwills	23/02/2004	17:45:00	3803	Trebrownbridge	17/04/2005	22:00:00	3149
Gwills	23/02/2004	18:15:00	3803	Trengwainton	30/04/2003	06:15:00	2141.4
Gwills	24/02/2004	00:00:00	3803	Trengwainton	12/07/2003	06:15:00	2346.8
Gwills	24/02/2004	02:15:00	3803	Trengwainton	13/09/2003	06:15:00	2564
Gwills	24/02/2004	04:00:00	3803	Trengwainton	04/11/2003	06:15:00	2699.6
Gwills	24/02/2004	04:30:00	3803	Trengwainton	29/11/2003	06:15:00	2766.6
Gwills	03/03/2004	05:15:00	3815.5	Trengwainton	05/05/2004	06:15:00	3325.2
Gwills	28/05/2004	05:45:00	799.8	Trengwainton	28/01/2005	14:30:00	3942.2
Gwills	03/07/2004	04:30:00	892	Trengwainton	28/01/2005	15:00:00	16058.2
Gwills	25/09/2004	00:00:00	1656.3	Woolstone	17/05/2003	06:45:00	579.2
Gwills	03/12/2004	09:15:00	1307.8	Woolstone	03/06/2003	06:45:00	598.6
Gwills	17/12/2004	00:00:00	1972.6	Woolstone	12/07/2003	06:45:00	661.4
Gwills	17/01/2005	07:15:00	6511.5	Woolstone	19/12/2003	06:45:00	986
Gwills	17/01/2005	11:45:00	3684.5	Woolstone	20/12/2003	06:45:00	990.6
Gwills	20/01/2005	09:15:00	7775.6	Woolstone	03/03/2004	06:45:00	1222
Gwills	12/02/2005	01:30:00	8480.4	Woolstone	26/04/2004	06:45:00	1346.6
Gwills	12/02/2005	02:15:00	1518.6	Woolstone	05/05/2004	06:45:00	1364.4
Gwills	02/03/2005	14:30:00	8464.4	Woolstone	27/05/2004	06:45:00	1367.8
Gwills	02/03/2005	14:45:00	1534.6	Woolstone	30/05/2004	06:45:00	1368.6
Gwills	02/03/2005	15:00:00	8464.4	Woolstone	26/06/2004	06:45:00	1412.4
Gwills	02/03/2005	16:45:00	1534.6	Woolstone	22/07/2004	06:45:00	1471
Gwills	15/08/2005	14:45:00	8246	Woolstone	12/10/2004	07:00:00	456.2
Gwills	15/08/2005	15:00:00	1753	Woolstone	02/10/2006	00:45:00	8002.4
Gwills	15/08/2005	15:15:00	8246	Yeolmbridge	13/10/2003	06:15:00	3289
Gwills	15/08/2005	17:00:00	1753	Yeolmbridge	20/05/2004	09:15:00	3986
Gwills	15/08/2005	17:30:00	8246	Yeolmbridge	06/07/2004	09:45:00	4067
Gwills	15/08/2005	19:00:00	1753	Yeolmbridge	19/11/2004	05:15:00	1000
Gwills	15/08/2005	19:30:00	8246	Data recorded as thousands at Yeolmbridge until:			
Gwills	15/08/2005	22:15:00	1753	Yeolmbridge	07/02/2005	00:00:00	1000
Gwills	25/11/2005	20:45:00	8001.8	Yeolmbridge	01/10/2005	07:15:00	5691
Huckworthy	13/06/2007	10:00:00	2060.6	Yeolmbridge	30/01/2006	07:15:00	8524.2
Luxulyan	01/03/2003	00:30:00	464	Yeolmbridge	30/01/2006	07:30:00	1474.8
Luxulyan	01/03/2003	01:00:00	369	Yeolmbridge	19/03/2006	07:15:00	8455.4
Luxulyan	01/03/2003	01:15:00	318	Yeolmbridge	19/03/2006	07:30:00	1543.6
Luxulyan	01/03/2003	02:45:00	291	Yeolmbridge	24/11/2006	10:30:00	7999.4
Luxulyan	02/03/2003	09:00:00	395				

Table B2 15-minute raingauge totals in the range 20–70 mm (all checked and comments noted).

Raingauge	Date	Time	15-minute accumulation	Comment
Bastreet	01/09/2004	10:30	48	Treated as missing.
Bastreet	23/01/2005	09:45	21	Treated as missing.
Bealsmill	02/07/2004	13:30	27.5	Treated as missing (raingauge just comes back online).
Bridgerule	11/07/2007	08:30	22	Treated as missing, near end of record.
Cornwood	03/07/2001	18:15	37	Before 2002
Cornwood	05/07/2001	06:15	23	Before 2002
Cornwood	06/07/2001	03:45	29	Before 2002
Crowford	16/08/2004	13:30	22	OK, Boscastle event
Holsworthy	19/07/2007	12:30	20.8	Raingauge not used.
Huckworthy	24/01/2002	20:15	31	Treated as missing.
Huckworthy	23/09/2002	17:00	30	Treated as missing.
Huckworthy	10/10/2006	23:15	24	OK – heavy thunderstorms reported during 10/11th (MO website)
Lanreath	11/08/2004	21:45	22.4	OK – localised showers during 11/8, could be real.
Luxulyan	05/01/2003	08:15	25	Treated as missing – part of suspect period.
Luxulyan	28/02/2003	14:15	33	Treated as missing – part of suspect period.
Luxulyan	01/03/2003	00:45	33	Treated as missing – part of suspect period.
Luxulyan	01/03/2003	02:30	68	Treated as missing – part of suspect period.
Luxulyan	01/03/2003	03:00	54	Treated as missing – part of suspect period.
Luxulyan	01/03/2003	03:15	39	Treated as missing – part of suspect period.
Luxulyan	01/03/2003	04:00	57	Treated as missing – part of suspect period.
Luxulyan	01/03/2003	05:00	48	Treated as missing – part of suspect period.
Luxulyan	02/08/2003	10:45	30	Treated as missing – part of suspect period.
Luxulyan	02/08/2003	22:30	22	Treated as missing – part of suspect period.
Luxulyan	04/08/2003	14:00	43	Treated as missing – part of suspect period.
Luxulyan	04/08/2003	14:15	22	Treated as missing – part of suspect period.
Luxulyan	30/01/2004	14:45	25	Treated as missing – part of suspect period.
Luxulyan	30/01/2004	15:45	32	Treated as missing – part of suspect period.
Luxulyan	30/01/2004	16:00	26	Treated as missing – part of suspect period.
Luxulyan	30/01/2004	16:15	63	Treated as missing – part of suspect period.
Luxulyan	30/01/2004	17:15	56	Treated as missing – part of suspect period.
Luxulyan	30/01/2004	23:45	43	Treated as missing – part of suspect period.
Luxulyan	31/01/2004	01:45	26	Treated as missing – part of suspect period.
Luxulyan	31/01/2004	02:15	22	Treated as missing – part of suspect period.
Luxulyan	31/01/2004	02:30	33	Treated as missing – part of suspect period.
Luxulyan	31/01/2004	05:45	70	Treated as missing – part of suspect period.
Luxulyan	31/01/2004	06:45	44	Treated as missing – part of suspect period.
Luxulyan	31/01/2004	07:45	35	Treated as missing – part of suspect period.
Luxulyan	31/01/2004	12:45	25	Treated as missing – part of suspect period.
Luxulyan	01/02/2004	04:30	46	Treated as missing – part of suspect period.
Luxulyan	01/02/2004	05:45	33	Treated as missing – part of suspect period.
Luxulyan	02/02/2004	07:30	28	Treated as missing – part of suspect period.
Luxulyan	29/03/2004	11:45	28	Treated as missing – part of suspect period.
Luxulyan	29/10/2006	11:15	29.8	Treated as missing – part of suspect period.
Luxulyan	31/10/2006	15:15	23.8	Treated as missing – part of suspect period.
Slaughterbridge	26/02/2002	11:45	40	Treated as missing.
Slaughterbridge	16/08/2004	13:15	21.5	OK, during Boscastle event
Sticklepath	23/10/2002	12:00	47.4	Treated as missing.
Trebrownbridge	22/02/2006	17:15	63.5	Treated as missing.
Trengwainton	02/03/2003	17:00	28.8	Treated as missing – part of suspect period.
Trengwainton	09/03/2003	13:00	49	Treated as missing – part of suspect period.
Trengwainton	09/03/2003	13:15	25.2	Treated as missing – part of suspect period.
Trengwainton	10/03/2003	03:15	20.6	Treated as missing – part of suspect period.
Trengwainton	10/03/2003	03:30	23	Treated as missing – part of suspect period.
Trevalec	16/08/2004	15:45	24.2	OK, Boscastle storm
Woolstone	10/01/2007	13:15	25.2	Treated as missing – part of suspect period.

Table B3 A record of all the periods of rain gauge record treated as missing (apart from 15-minute totals in excess of 70mm listed in Table B1).

Raingauge	Period	Comment
Luxulyan	24/11/2002 15:15:00 to 10/01/2003 04:30:00	Looks like a blockage.
Luxulyan	28/02/2003 14:15:00 to 12/03/2003 16:00:00	Suspect behaviour, looks blocked.
Luxulyan	02/08/2003 10:00:00 to 02/08/2003 11:15:00	
Luxulyan	02/08/2003 19:00:00 to 03/08/2003 00:30:00	
Luxulyan	04/08/2003 14:00	43 mm
Luxulyan	04/08/2003 14:15	22 mm
Luxulyan	07/10/2003 11:00	7 mm
Luxulyan	30/01/2004 14:00:00 to 2/02/2004 13:45:00	Suspect – period is quiet and ends with 470 mm.
Luxulyan	29/03/2004 11:30	16 mm
Luxulyan	29/03/2004 11:45	28 mm
Luxulyan	29/03/2004 12:00	3 mm
Luxulyan	08/04/2004 10:45	4 mm
Luxulyan	08/04/2004 11:00	20 mm
Luxulyan	08/04/2004 11:15	1 mm
Luxulyan	10/04/2004 23:00:00 to 16/04/2004 06:00:00	Looks like gauge was down.
Luxulyan	15/07/2004 03:45:00 to 10/08/2004 00:00:00	Looks like gauge was down.
Luxulyan	18/10/2006 21:16:14 to 16/11/2006 09:32:53	Suspect
Roserrow	25/06/2002 07:22:30 to 13/07/2002 17:31:20	Looks like a blocked gauge.
Roserrow	31/07/2003 12:54:50 to 09/10/2003 10:27:00	Suspect
Roserrow	30/06/2007 05:53:50 to 11/07/2007 15:12:30	At end of record – looks like a blocked gauge.
Wadebridge	01/12/2002 01:25:20 to 29/12/2002 06:13:30	Appears to be significantly under recording rain.
Trevalect	08/08/2003 06:44:30 to 22/09/2003 01:11:50	No tips between these dates which looks suspect.
Trevalect	12/01/2006 13:44:09 to 21/01/2006 19:35:20	Appears blocked.
Trevalect	06/03/2007 20:20:20 to 22/03/2007 05:12:20	Appears to be under recording.
Bastreet	01/09/2004 10:30	48 mm
Bastreet	23/01/2005 09:40	21 mm
De Lank	06/01/2005 22:15:00 to 11/01/2005 09:00:00	Appears to be under recording.
De Lank	03/11/2005 09:28:49 to 11/11/2005 09:28:32	Appears to be under recording.
De Lank	09/03/2006 23:30:43 to 30/04/2006 18:47:22	Appears to be under recording.
De Lank	27/05/2006 12:32:36 to 25/06/2006 23:01:55	Appears to be under recording.
Woolstone	08/05/2002 05:15:01 to 20/05/2002 08:15:01	Suspect readings
Woolstone	26/10/2002 16:15:01 to 11/11/2002 14:15:01	Appears to be under recording.
Woolstone	15/12/2006 00:00:00 to 11/01/2007 12:30:00	Looks like a blockage.
Bridgerule	16/08/2004 12:45:00 to 10/09/2004 09:00:00	Raingauge appears to be blocked.
Bridgerule	11/07/2007 08:15:00 to 12/07/2007 09:00:01	Suspect to end of record – includes 15-minute value of 22.00 at 11/07/2007 08:30:00
Yeolmbridge	19/11/2004 05:15:00 to 07/02/2005 00:00:00	Values appear to have been multiplied by 1,000.
Yeolmbridge	15/09/2005 11:00:00 to 08/10/2005 09:30:00	Suspect under recording (changed from 1 to 0.2 mm resolution during this period).
Yeolmbridge	25/02/2006 08:00:00 to 23/03/2006 18:00:00	Suspect under recording.
Yeolmbridge	06/10/2006 09:00:00 to 25/10/2006 02:00:00	Suspect under recording.
Crowford	17/07/2007 04:30:00 to 28/07/2007 15:45:00	Looks like a blockage.
Cornwood	01/08/2002 03:15:00 to 08/08/2002 14:00:01	Suspect
Cornwood	15/05/2003 00:00:00 to 11/06/2003 14:45:00	Suspect under recording.
Huckworthy	24/01/2002 20:15	31 mm
Huckworthy	23/09/2002 15:15:00 to 23/09/2002 17:00:00	Suspect– ends with 30 mm
Huckworthy	05/08/2004 13:00:00 to 09/08/2004 10:45:00	Suspect values – daily accumulation >470 mm
Huckworthy	26/05/2006 10:00:00 to 13/06/2006 09:30:00	Looks like a blockage.
Pillaton	01/05/2003 06:00:00 to 10/05/2003 06:00:00	Suspect under recording.
Lee Moor	21/12/2003 18:30:00 to 01/01/2004 03:00:00	Suspect period at start of record.

Raingauge	Period	Comment
Sticklepath	15/11/2002 10:45:00 to 15/01/2003 05:30:00	Suspect under recording.
Sticklepath	23/10/2002 12:00	47.4 mm
Sticklepath	06/12/2006 02:45:00 to 18/12/2006 16:30:00	Suspect under recording.
Parkham	14/09/2004 07:00:00 to 25/09/2004 15:15:00	Suspect gauge was down/under recording.
Allisland	18/11/2004 08:00:00 to 16/12/2004 00:15:00	Suspect gauge was down/under recording.
Tamarstone	30/10/2006 09:00:00 to 17/11/2006 00:45:00	Suspect gauge was down/under recording.
Penryn	20/02/2002 06:59:50 to 19/03/2002 12:19:30	Suspect gauge was down/under recording.
Penryn	24/08/2005 06:15:00 to 30/08/2005 09:45:00	Erroneous tips in period
Trengwainton	26/02/2003 00:00:00 to 11/03/2003 13:00:00	Suspect behaviour, several large values >20 mm.
Allet	26/02/2004 19:00:00 to 15/03/2004 18:30:00	Suspect gauge was down/under recording.
Allet	13/04/2005 04:30:00 to 19/04/2005 12:30:00	Suspect gauge was down/under recording.
Bissoe	14/04/2005 08:36:21 to 13/05/2005 11:04:46	Looks like a blockage.
Trebrownbridge	22/02/2006 17:15	63.5 mm
Bealsmill	02/07/2004 13:28	Suspect value of 27.5 mm – raingauge just come back on line.
Slaughterbridge	26/02/2002 11:45	40 mm

Appendix C: Instructions for using the test system

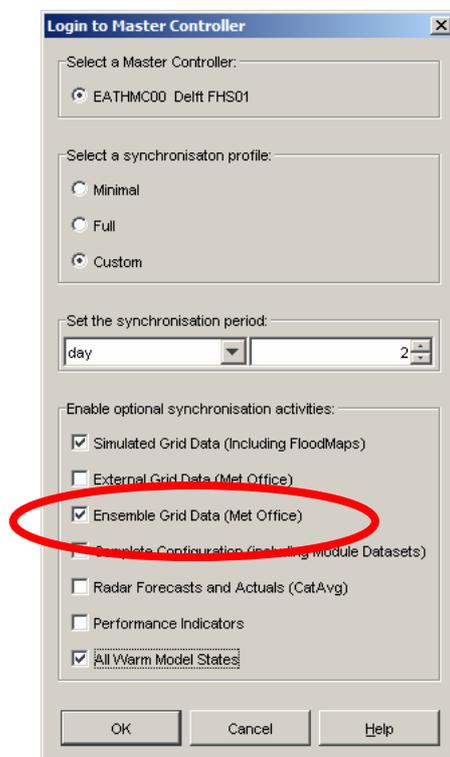
NOTE:

- If you run the system for the first time, you must restart after the initial synchronisation to be able to download the MOGREPS grids.

Start-up instructions

- Install by unzipping the T46_Client.zip file in your drive of choice. Be sure to unzip the full paths. This will create a T46_FEWS directory (with subdirectories).
- Start the client using the **Thames_T46OC** or **Northeast_T46OC** executable located in the bin directory.
- If new (root) configuration has been uploaded by Delft, you may to restart the client after the initial synchronisation.
- If you want to look at the MOGREPS input grids themselves (not needed to see the forecast results), you must choose a custom profile and select the Ensemble Grid Data (Met Office) check box (the ensemble results of the models are always synchronised) (Figure C1).

Figure C1 Select the Ensemble Grid Data (Met Office) check box.



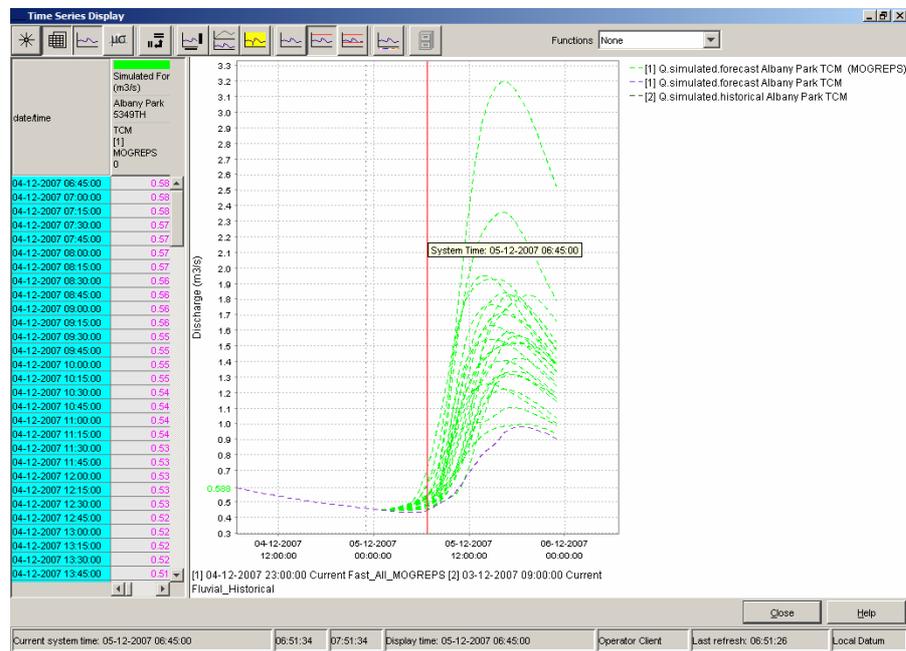
Thames

Ensemble results for TCM/ARMA are available from the filters (as spaghetti plots).
Example:

- Select filter Thames Region → Forecast Points → Hydrological.
- Select location Albany Park, parameters Simulated Historical Discharge and Simulated Forecast discharge and click on the time series display icon.

You should see something like the screen shown in Figure C2. In this case, each of the green lines is the result of an individual MOGREPS ensemble run. The blue line is the normal forecast based on the NWP input.

Figure C2 Example of ensemble results for TCM/ARMA from the filters.

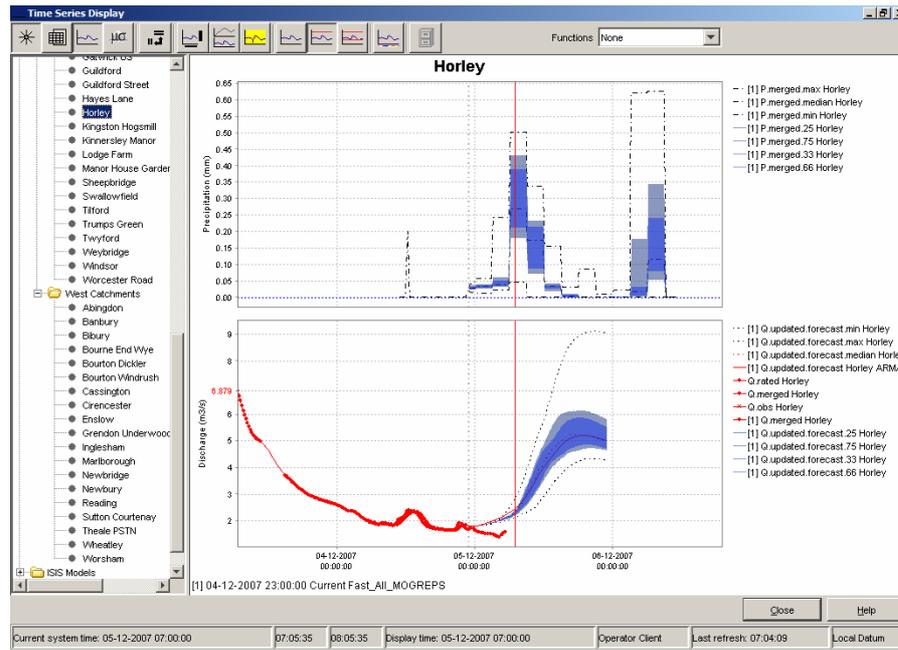


Ensemble results are also available in the pre-defined displays (Area graphs). For example:

- Open the pre-define display. Here you will find two new entries – MOGREPS Catchments TCM and MOGREPS Catchments ARMA showing the ‘raw’ TCM and error corrected ARMA results respectively.
- The plots are divided in two panels – the top one shows the rainfall input, the bottom the resulting discharge. The maximum, minimum and the area between the 25–75 percentiles and the area between the 33–66 percentiles are shown for each graph (Figure C3).

To see the individual MOGREPS grids (only possible when the custom sync has been chosen), open the spatial display and navigate to the MOGREPS filter. Here, each of the 24 individual members may be selected.

Figure C3 Example of pre-defined display of ensemble results.



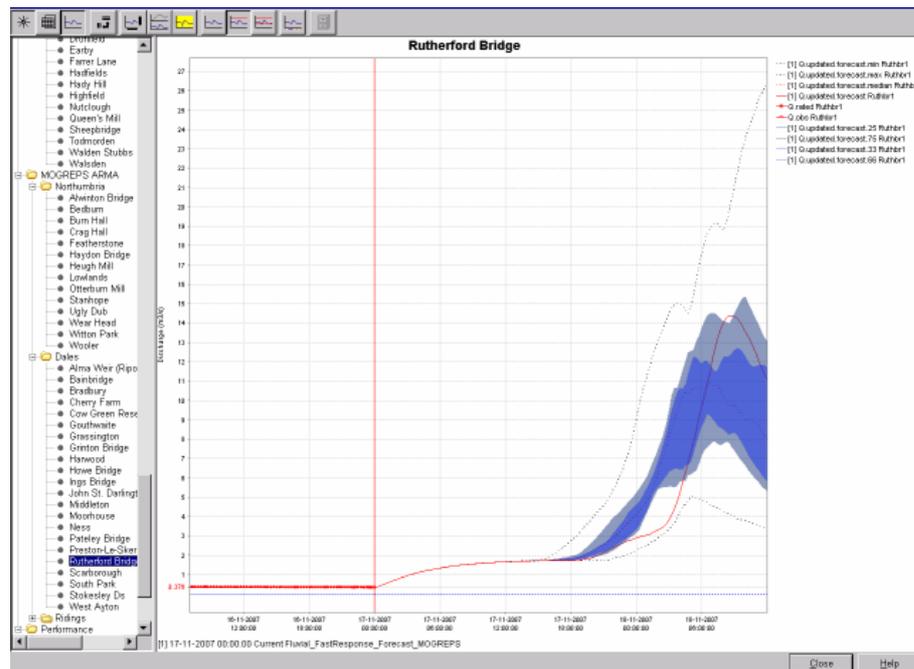
North East Region

Ensemble results for PDM/ARMA/KW?SNOW are available from the filters (as spaghetti plots). For example:

- Select filter Dales → Forecast Points → Hydrological Forecast Points.
- Select location Rutherford Bridge, parameters Simulated Historical Discharge and Simulated Forecast discharge and click on the time series display icon.

For display of the statistics as area graphs, the display groups have been updated with MOGREPS groups for all catchments. The first five plot groups are made (RainfallMOGREPS, PDMSIMULATEDMOGREPS, PDMUPDATEDMOGREPS, ERRORMOGREPS and ARMAMOGREPS) using the display groups (Figure C4).

Figure C4 Example ensemble results from for PDM/ARMA/KW?SNOW from the filters.



Appendix D: Database optimisation using SQL queries

FEWS stores all time series in the database as blobs (binary large objects). The contents of the blobs cannot be read directly in this manner (to do so use the jdbc server¹⁶) but the rest of the table can be read directly.

The example queries below focus on determining the size of certain time series in the database as the time series table is usually by far the largest table in the database. This can help in optimising the configuration of the system allowing you to focus on the largest chunks of data.

⚠ These examples have been made using a post gresql database. When using these on another database, some SQL statements may need to be modified to cater for small language differences.

Database sizes and system performance

Delft-FEWS uses a database to store and retrieve forecast data. Although there is a central database, each client – operator client (OC) and forecasting shell server (FSS), the calculation node – uses its own local datastore, which is a synchronised mirror of the central database. Synchronisation profiles and synchronisation levels are used to determine which data are synchronised between the clients and the central database. To achieve optimal performance, it is vital to only synchronise data that need to be synchronised. Figure D1 illustrates the layout of Delft-FEWS.

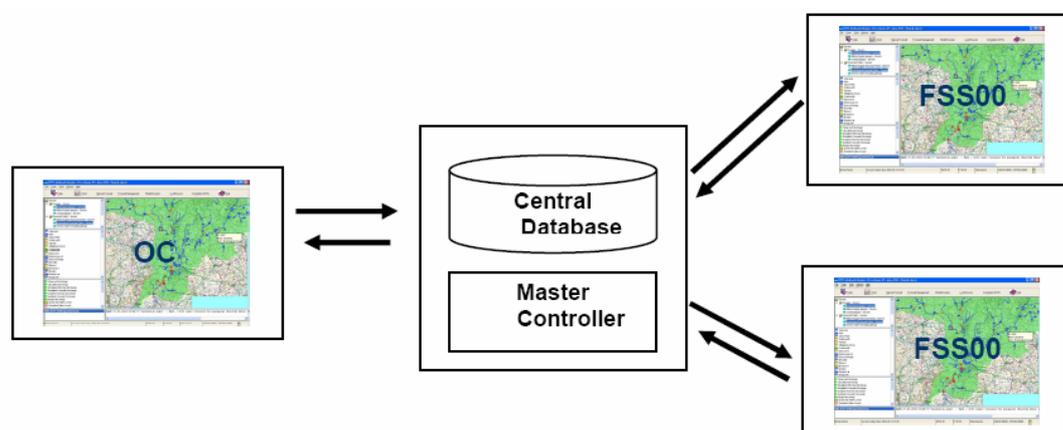


Figure D1 Layout of the Delft-FEWS live system.

Tuning the data streams between the different components can be done on a global level (using synchronisation profiles – not discussed here) and on a per time series level by setting a synchronisation level to a specific time series or by compressing time series.

This appendix focuses on the amount of data sent by the FSS to the master controller (for later retrieval) and the amount of data that are needed to visualise results on the OC.

Apart from synchronisation, Delft-FEWS can also use other techniques to minimise the amount of data. These are:

¹⁶ See <http://public.deltares.nl/display/FEWSDOC/Home> for documentation on how to do this.

- compression of data by specifying a precision;
- marking bits of data as temporary.

Table D1 lists the syncLevels currently used in Delft-FEWS:

Table D1 syncLevels – selected by convention.

Level	Description
0	(Default) All data from a forecast run
1	Scalar time series imported from telemetry
2	All grid data from a forecast run (e.g. flood kaps/Lisflood results)
3	CatAvg data (forecasts, actuals and NWP)
4	Astronomical
5	Data edited on OC
6	Grid data imported from external forecast (synchronised to OC)
7	Grid data imported from external forecast (FSS and MC only)
9	Temporary time series (not synchronised)

The general steps to follow to minimise the amount of data being transferred are as follows:

1. Determine what the bulk of the data in your systems are. Are they forecast data or imported grids?
2. Only synchronise data to the OC which are used in the OC (e.g. to make graphs). For example, ensemble grids are almost never interesting to look at in the OC, so they should be marked with synclevel 7.
3. Make sure temporary data (intermediate steps in calculations) get synclevel 9 and a short expiry time.
4. Set the valueResolution parameter in Parameters.xml for the parameters that take up a lot of space. By sacrificing some accuracy, you can achieve very high compression ratios for certain parameters. This can be very effective for large grids with a different value for each cell such a temperature fields. Precipitation generally already compresses very well without sacrificing accuracy because a precipitation field usually contain lots of repetitive values.

Calculating database size

The size of the database can be estimate using the following calculations:

$$\text{Forecastdata} = \text{FS} * \text{FF} * \text{RB}$$

$$\text{Historic/Staterun} = \text{SS} * \text{SF} * \text{RB}$$

$$\text{ModelState} = \text{MS} * \text{SF} * \text{RB}$$

$$\text{Historical date} = \text{RB} * \text{IS}$$

where:

FS = Size of a single forecast

SS = Size of a single Historic run

IS = Size of one day of import data

FS = forecast frequency (per day)
SF = Staterun/Historic frequency (per day)
RB = Rolling barrel length in days
MS = Size of the model state for one Historical run

In addition, several other (usually small) components will need to be added such as configuration data, model states, etc. However, the above make up the bulk of the data.

Fields in the time series table

Table D2 lists the fields in the time series table in the FEWS database. These fields can be used to query the database.

Please note that it will be difficult to use the locationId as the system optimises the database by storing data (for the same period) for several locations in one record. In this process, the location field will be changed to indicate a list of locations (in a run length encoded format).

Table D2 Field in time series table.

Field	Description
localIntId	Only filled local datastore, not in the MC database. Used to determine when a blob is created.
creatorTaskRunId	PM
blobId	ID of this blob
modifierId	PM
moduleInstanceId	ModuleInstance that created this series
parameterId	ParameterId for this time series
locationId	location ID (or IDs). A blob can hold data for multiple locations for optimisation purposes.
beginTime	Date/time of first value in the blob
endTime	Date/time of last value in the blob
timeSeriesType	Type of the time series according to the following enumeration: 0 external historical 1 external forecasting 2 simulated historical 3 simulated forecasting
taskRunId	Holds the taskrun that created this series. This field is null in the case of external data.
ensembleMemberId	EnsemblememberId. Default is main\$0 for non-ensemble date.
blob	The actual blob
blobSize	Size of the blob in bytes
creationTime	Time the blob was created.
synchLevel	Level used to determine when and how to synchronise the data.
expiryTime	Expiry time of this record. After this date/time the record will be removed.
localAvailableTime	Time this record was first available in the system
valueType	Type of the actual times series data: 0 scalar 1 grid 2 longitudinal profile 3 polygon 4 sample
timeStepId	Time step of this time series
externalForecastingStartTime	Forecast time in case of an external forecasting time series
constantFlag	
maxValue	Maximum value in the blob

Example queries

Example 1

The query below sums all the blobs in the time series table to determine the total size in megabytes (MB):

```
SELECT Sum(Length(TimeSeries.blob)/1024.0/1024.0)
FROM Timeseries
where moduleinstanceid LIKE '%'
```

Example 2

The query below produces a table (Table D3) showing the size of all the import data in the database (assuming all the import modules start with Import):

```
SELECT moduleinstanceid, parameterid, timeseriestype, synchlevel,
count(synchLevel),
Sum(Length(TimeSeries.blob))
FROM Timeseries
where moduleinstanceid like 'Import%'
GROUP BY moduleInstanceId,timeSeriesType, synchLevel, parameterid;
```

Table D3 Result for example query 2.

moduleinstanceid	parameterid	timeseriestype	synchlevel	count	sum 
ImportGfs	LH.simulated	1	6	898	324061976
ImportGfs	Wind.v	1	6	855	232308255
ImportGfs	Wind.u	1	6	858	231092616
ImportGfs	T.simulated	1	6	835	207716596
ImportGfs	RH.simulated	1	6	835	187757956
ImportGfs	Sol.shortwave.incoming	1	6	888	101572969
ImportGfs	LH.pot.simulated	1	6	884	97282656
ImportGfs	P.nwp.forecast	1	6	884	81574898
ImportGfs	Sol.shortwave.incoming	1	1	404	82295
ImportExternal	Eff.inverter	0	1	458	61615
ImportExternal	W.panel	0	1	989	61238
ImportExternal	V.panel	0	1	458	59994
ImportExternal	A.panel	0	1	458	59166
ImportExternal	Time.inverter	0	1	458	57829
ImportExternal	Time.inverter.today	0	1	458	57613
ImportExternal	Hz.grid	0	1	458	54990
ImportExternal	V.grid	0	1	458	53520
ImportExternal	T.inverter	0	1	458	51621
ImportExternal	kWh.total	0	1	458	49917
ImportExternal	kWh	0	1	413	38140
ImportExternal	kWh.today	0	1	123	12065

Example 3

To determine the size of a single forecast, one must find all the records belonging to a specific taskrunId:

```
SELECT TimeSeries.creatorTaskRunId, Timeseries.parameterid,
TimeSeries.moduleInstanceId,
TimeSeries.synchLevel,
Count(TimeSeries.synchLevel) AS NumBlobs,
Sum(Length(TimeSeries.blob))
FROM TimeSeries
WHERE (((TimeSeries.creatorTaskRunId)='EFASMC00:000024672'))
```

```
GROUP BY TimeSeries.creatorTaskRunId, TimeSeries.moduleInstanceId,
timeseries.parameterid,
TimeSeries.synchLevel;
```

Example 4

The amount of import data for one day can be determined using the following query:

```
SELECT
Sum(Length(TimeSeries.blob)/1024.0/1024.0)
FROM Timeseries
WHERE creationtime BETWEEN to_date('02-02-2008','MM-DD-YYYY') AND
to_date('02-03-
2008','MM-DD-YYYY')
and moduleinstanceid like 'Import%'
```

Figure D1 shows an example output for this query.

Figure D2 Result for example query 4.

sum
77.132251739501953125000000

Example 5

Alternatively we can group this to see which part of the import data take up most space:

```
SELECT moduleinstanceid, parameterid, timeseriestype, synchlevel,
count(synchLevel),
Sum(Length(TimeSeries.blob)/1024.0/1024.0)
FROM Timeseries
WHERE creationtime BETWEEN to_date('02-02-2008','MM-DD-YYYY') AND
to_date('02-03-
2008','MM-DD-YYYY')
and moduleinstanceid like 'Import%'
GROUP BY moduleInstanceId,timeSeriesType, synchLevel, parameterid;
```

Figure D3 Example output for example query 5

moduleinstanceid	parameterid	timeseriestype	synchlevel	count	sum
ImportGfs	LH.simulated	1	6	60	20.21313095092773437500
ImportGfs	T.simulated	1	6	61	14.92627239227294921875
ImportGfs	RH.simulated	1	6	61	13.12361526489257812500
ImportGfs	Sol.shortwave.incoming	1	6	60	6.96641540527343750000
ImportGfs	LH.pot.simulated	1	6	59	6.20453834533691406250
ImportGfs	Wind.u	1	6	22	5.32510375976562500000
ImportGfs	P.nwp.forecast	1	6	59	5.20136451721191406250
ImportGfs	Wind.v	1	6	20	5.16388320922851562500
ImportExternal	A.panel	0	1	2	0.00090408325195312500
ImportExternal	W.panel	0	1	7	0.000809669494628906250000
ImportExternal	V.panel	0	1	2	0.00080776214599609375
ImportExternal	Eff.inverter	0	1	2	0.00079441070556640625
ImportExternal	H.z.grid	0	1	2	0.00067996978759765625
ImportExternal	Time.inverter	0	1	2	0.00063323974609375000
ImportExternal	Time.inverter.today	0	1	2	0.00063323974609375000
ImportGfs	Sol.shortwave.incoming	1	1	3	0.00056648254394531250
ImportExternal	kWh	0	1	4	0.000535964965820312500000
ImportExternal	kWh.total	0	1	2	0.00053405761718750000
ImportExternal	T.inverter	0	1	2	0.00042438507080078125

Example 6

The amount of data for the module states and logfiles can be determined using:

```
SELECT moduleInstanceId, count(  
moduleinstanceid),  
Sum(Length(warmstates.blob)) FROM  
WarmStates GROUP BY moduleInstanceId;
```

and

```
SELECT eventcode, count(  
eventcode),  
Sum(Length(logmessage)) FROM  
logentries GROUP BY eventcode;
```

Appendix E: Importing and processing pseudo ensembles with Delft-FEWS

A simple C program has been constructed on the basis of the CEH sidb read library to convert the sidb data to ArcInfo ASCII files that can be imported into the system.

Figure E1 Converter program to help screen with command-line options.

```
Usage: sidbconv [options] (*default, +required)
-h Print this help
-F format Set output format: 1, 2 or 3
  1 = ArcView ascii for use with Delft-Fews import *
  2 = ArcView ascii for use with Delft-Fews general adapter
  3 = bil ascii for use with Delft-Fews general adapter (NOT IMPLEMENTED)
-i interval interval of data to read/write in minutes +
-n steps number of steps to process +
-s sidbpath full path to sidb +
-o outputfilename prefix for naming the outputfiles +
  This will be the external locationid when using the Delft-Fews import
-m month month to start reading data from +
-y year year to start reading data from +
-d day day of month to start reading data from +
-t time time in minutes in the day to start reading data from +
-b write bil format (instead of arc ascii) NOT IMPLEMENTED
-T sourcetype sidb sourcetype (default=5)
-I sourceid sidb sourceid (default=555)
-D datatype sidb datatype (default=10)
-S datasubtype sidb datasubtype (default=4613)
-E datatypeindexex sidb time since start of forecast (default=0)
-X x_size nr of columns (default=140)
-Y y_size nr of rows (default=140)
```

After conversion and importing, the forecasts are available within the FEWS database for processing and display. The grids are processed to calculate catchment average precipitation for the PDM model and to create average precipitation for each REW.

**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.