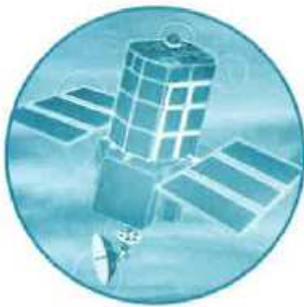


**Comparison of Rainfall-Runoff Models for Flood
Forecasting**
Part 2: Calibration and evaluation of models



Research and Development

**Technical Report
W242**



ENVIRONMENT AGENCY

Comparison of Rainfall-Runoff Models for Flood Forecasting

Part 2: Calibration and evaluation of models

R&D Technical Report W242

V A Bell, D S Carrington and R J Moore

Research Contractor:
Institute of Hydrology

Publishing Organisation

Environment Agency, Rio House, Waterside Drive, Aztec West, Almondsbury,
BRISTOL, BS32 4UD.

Tel: 01454 624400 Fax: 01454 624409
Website: www.environment-agency.gov.uk

© Environment Agency 2001

September 2001

ISBN: 1 85705 397 4

All rights reserved. No part of this document may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior permission of the Environment Agency.

The views expressed in this document are not necessarily those of the Environment Agency. Its officers, servants or agents accept no liability whatsoever for any loss or damage arising from the interpretation or use of the information, or reliance upon views contained herein.

Dissemination status

Internal: Released to Regions
External: Released to Public Domain

Statement of Use

This report provides users of hydrological models with an indication of the comparative performance of different rainfall-runoff models in varied catchment situations. It is intended to be used as a guide in the selection of appropriate models for flood forecasting purposes.

Research Contractor

This document was produced under R&D Contract W5-005 by:
Institute of Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford, Oxfordshire,
OX10 8BB

Tel: 01491 838800 Fax: 01491 692424

Environment Agency's Project Manager

The Environment Agency's Project Managers for Project W5-005 were:
Owen Wedgwood, North West Region
Nigel Outhwaite, Thames Region

ACKNOWLEDGEMENTS

Particular thanks are due to the following Environment Agency members of the Steering Committee:

Owen Wedgwood (Project Coordinator and Committee Chairman)
Tim Harrison
Mike Knowles
Nigel Outhwaite
Jennifer Soggee

The Flood Protection Commission of the Ministry of Agriculture, Fisheries and Food, through its strategic funding of flood forecasting research at the Institute of Hydrology, has provided the foundation for aspects of the work reported here.

CONTENTS

	Page
Acknowledgements	i
List of Tables, Figures and Plates	iv
Executive Summary	xi
Keywords	xiii
1. Introduction	1
1.1 Background	1
1.2 Models for intercomparison	2
1.3 Selection of study catchments	24
2. Strategy for model assessment	28
2.1 Introduction	28
2.2 Strategy for model calibration	28
2.3 Strategy for model evaluation	29
2.4 Model and data configuration tools	33
2.5 Preparation of models and data for assessment	40
2.6 Study catchments and data for model assessment	44
2.7 Data management	64
2.8 Selection of calibration and evaluation periods	71
3. Model calibration	75
3.1 Introduction	75
3.2 Model calibration using raingauge data	75
3.3 Model calibration using radar data	79
3.4 Overview	93
4. Model evaluation	94
4.1 Introduction	94
4.2 Results using raingauge data	94
4.3 Results using weather radar data	123
4.4 Overview	145
5. Choice of catchment model	147
5.1 Ease of model use	147
5.2 Data requirements of models	151
5.3 Guidelines for model choice	152
5.4 Summary of model choice	153
6. Summary, conclusions and recommendations	156
6.1 Summary	156
6.2 Main results	157

	Page	
6.3	General conclusions	159
6.4	Operational recommendations	160
6.5	Opportunities for further work	163
References		165
Appendix A: Radar calibration using multiquadrics		169
A.1	Introduction	169
A.2	HYRAD local calibration procedure	169
A.3	Hybrid calibration procedure	170
A.4	Application to the study catchments	171
Appendix B: Catchment maps		173
Appendix C: Model parameter sets		192
C.1	TCM models	192
C.2	NWS models	196
C.3	MCRM models	200
C.4	PDM models	204
C.5	IEM models	208
C.6	TF models	211
C.7	PRTF models	215
C.8	Grid models	218
Appendix D: Simulated and forecast flow hydrographs		221

LIST OF TABLES, FIGURES AND PLATES

	Page	
Table 1.2.1	Parameters in the Thames Catchment Model	4
Table 1.2.2	Parameters of the NWS Model	7
Table 1.2.3	Parameters in the Midlands Catchment Runoff Model	9
Table 1.2.4	Parameters of the PDM model	11
Table 1.2.5	Parameters of the Isolated Event Model	12
Table 1.2.8	Parameters of the Grid Model	21
Table 1.3.1	Catchments selected for model assessment	25
Table 2.3.1	Two-way contingency table for categorical evaluation	33
Table 2.4.1	Raingauge weights derived using different methods	36
Table 2.4.2	Spatial datasets used to delineate the TCM response zones	38
Table 2.4.3	WRAP (Winter Rain Acceptance Potential) data classes	38
Table 2.4.4	The proportion of each land type in each catchment	39
Table 2.6.1	Flow characteristics and Alarm Sensor levels for the study catchments	45
Table 2.6.2	Data availability for Trout Beck at Moor House	47
Table 2.6.3	Data availability for the Silk Stream at Colindeep Lane	49
Table 2.6.4	Data availability for the Dove at Izaak Walton	51
Table 2.6.5	Data availability for the Lavant at Graylingwell	53
Table 2.6.6	Data availability for the Rhondda at Trehafod	55
Table 2.6.7	Data availability for the Brue at Lovington	57
Table 2.6.8	Data availability for the Stour at Shipston	59
Table 2.6.9	Data availability for the Roch at Blackford Bridge	61
Table 2.6.10	Data availability for the Witham at Claypole Mill	63
Table 2.8.1	Periods used for model calibration and evaluation: raingauge data	72
Table 2.8.2	Menu of events for possible use with weather radar data	73
Table 2.8.3	Periods used for model calibration and evaluation: radar data	74
Table 3.2.1	Calibration performance (R^2 statistic) for different models, catchments and periods of record using raingauge data (best results in bold)	76
Table 3.2.2	Model order of ARMA(p,q) error predictors for different models and catchments	86
Table 3.2.3	Comparison of model-gain updating smoothing factors across catchments	86
Table 3.3.1	Calibration performance (R^2 statistic) for different models, catchments and periods of record using radar data (best results in bold)	90
Table 4.2.1	Evaluation performance for different models and catchments: simulation mode using raingauge estimates as input	95
Table 4.2.2	Evaluation performance for different models and catchments: forecast mode using raingauge data as input	108
Table 4.2.3	Best TF model updating procedure for each catchment as judged using the R^2 statistic and Threshold CSI criterion. State Updating: SU; Error prediction: ARMA	109

	Page	
Table 4.3.1	Evaluation performance for different models and catchments: simulation mode using raingauge, uncalibrated radar and calibrated radar estimates of rainfall as input	126
Table 4.3.2	Evaluation performance for different models and catchments using raingauge and calibrated radar estimates of rainfall as input	137
Table 5.4.1	Summary of advantages and disadvantages of the rainfall-runoff models	154
Table 6.4.1	Summary of current operational use of rainfall-runoff models by EA regions in the UK and operational recommendations for each	161
Table 6.4.2	Operational recommendations for rainfall-runoff models in terms of catchment size and dominant characteristics	162
Table A.1	Mean field bias of radar rainfall over catchments for different periods	172
Table C.1.1	Comparison of TCM parameters across catchments using raingauge data	193
Table C.1.2	Comparison of TCM parameters across catchments using uncalibrated radar data	194
Table C.1.3	Comparison of TCM parameters across catchments using calibrated radar data	195
Table C.2.1	Comparison of NWS model parameters across catchments using raingauge data	197
Table C.2.2	Comparison of NWS model parameters across catchments using uncalibrated radar data	198
Table C.2.3	Comparison of NWS model parameters across catchments using calibrated radar data	199
Table C.3.1	Comparison of MCRM parameters between catchments using raingauge data	201
Table C.3.2	Comparison of MCRM parameters between catchments using uncalibrated radar data	202
Table C.3.3	Comparison of MCRM parameters between catchments using calibrated radar data	203
Table C.4.1	Comparison of PDM parameters across catchments using raingauge data	205
Table C.4.2	Comparison of PDM parameters across catchments using uncalibrated radar data	206
Table C.4.3	Comparison of PDM parameters across catchments using calibrated radar data	207
Table C.5.1	Comparison of IEM parameters across catchments using raingauge data	209
Table C.5.2	Comparison of IEM parameters across catchments using radar data	210

	Page	
Table C.6.1	Comparison of simulation mode TF parameters across catchments using raingauge data	212
Table C.6.2	Comparison of simulation mode TF parameters across catchments using uncalibrated radar data	212
Table C.6.3	Comparison of simulation mode TF parameters across catchments using calibrated radar data	213
Table C.6.4	Comparison of forecast mode TF parameters across catchments using raingauge data	214
Table C.6.5	Comparison of forecast mode TF parameters across catchments using calibrated radar data	214
Table C.7.1	Comparison of simulation mode PRTF parameters across catchments using raingauge data	216
Table C.7.2	Comparison of simulation mode PRTF parameters across catchments using uncalibrated radar data	216
Table C.7.3	Comparison of simulation mode PRTF parameters across catchments using calibrated radar data	216
Table C.7.4	Comparison of forecast mode PRTF parameters across catchments using raingauge data	217
Table C.7.5	Comparison of forecast mode PRTF parameters across catchments using calibrated radar data	217
Table C.8.1	Comparison of Grid Model parameters across catchments using raingauge data	219
Table C.8.2	Comparison of Grid Model parameters across catchments using uncalibrated radar data	219
Table C.8.3	Comparison of Grid Model parameters across catchments using calibrated radar data	220
Figure 1.2.1	Representation of a hydrological response zone within the Thames Catchment Model.	3
Figure 1.2.2	The NWS Model.	5
Figure 1.2.3	The Midlands Catchment Runoff Model.	8
Figure 1.2.4	The PDM rainfall-runoff model.	10
Figure 1.2.8	The Grid Model.	20
Figure 1.3.1	Location of the study catchments.	26
Figure 2.3.1	Definition sketch for types of flow forecast.	30
Figure 2.3.2	Threshold crossing criteria.	32
Figure 2.4.1	Comparison of raingauge weights derived using the Thiessen and elevation schemes.	35
Figure 2.4.2	DTM-derived TCM response zones for the Roch.	39
Figure 2.6.2	Catchment map for Trout Beck at Moor House.	47
Figure 2.6.3	Catchment map for the Silk Stream at Colindeep Lane.	49
Figure 2.6.4	Catchment map for the Dove at Izaak Walton.	51
Figure 2.6.5	Catchment map for the Lavant at Graylingwell.	53
Figure 2.6.6	Catchment map for the Rhondda at Trehafod.	55
Figure 2.6.7	Catchment map of the Brue at Lovington.	57
Figure 2.6.8	Catchment map for the Stour at Shipston.	59

	Page
Figure 2.6.9	Catchment map for the Roch at Blackford Bridge. 61
Figure 2.6.10	Catchment map for the Witham at Claypole Mill. 63
Figure 2.7.1	Anomaly maps constructed from long-term radar rainfall totals: Hameldon Hill radar. (Roch catchment superimposed) 65
Figure 2.7.2	Anomaly maps constructed from long-term radar rainfall totals: Clee Hill radar. (Stour catchment superimposed) 66
Figure 2.7.3	Compositing of 2 and 5 km radar fields from Clee Hill. (Stour catchment superimposed) 67
Figure 2.7.4	Cumulative hyetographs of catchment average rainfall over the Stour estimated using uncalibrated and calibrated radar and raingauge data: 1 to 30 April 1998. 68
Figure 2.7.5	Cumulative hyetographs for the Stour showing uncalibrated and calibrated radar data and raingauge data. Clee Hill radar, 1 to 31 March 1998. 70
Figure 2.7.6	Radar rainfall field resulting from spurious clutter for the six day accumulation 13 to 19 March 1998 for Clee Hill. 71
Figure 3.2.1	Variation in model performance (median R^2 statistic) across catchments. 77
Figure 3.2.2	Model calibration: Flow hydrographs for the Silk Stream, 15/1/90 – 15/9/90. 80
Figure 3.2.3	Model calibration: Flow hydrographs for the Stour, 8/1/90 – 1/9/90. 82
Figure 3.2.4	Impulse response functions for the TF and PRTF models calibrated in simulation mode. 84
Figure 3.2.5	Impulse response functions of ARMA error-predictors. 86
Figure 3.2.6	Impulse response functions of TF and PRTF models calibrated in updating mode. 88
Figure 3.3.1	Simulation-mode hydrographs using (i) uncalibrated and (ii) calibrated radar data: Silk Stream, 12 to 28 December 1989. 91
Figure 3.3.2	Simulation-mode hydrographs using (i) uncalibrated and (ii) calibrated radar data: Stour, 8 to 28 January 1993. 92
Figure 4.2.1	Simulation-mode R^2 performance across catchments: evaluation period using raingauge data. 96
Figure 4.2.2	Scatter plots of observed and simulated peak flows: evaluation period using raingauge data. 99
Figure 4.2.3(a)	Simulated flow hydrographs from different models: Silk Stream, 21–25 September 1992. 101
Figure 4.2.3(b)	Simulated flow hydrographs from different models: The Stour, 6–18 April 1998 (Easter 1998 Flood). 102
Figure 4.2.3(c)	Simulated flow hydrographs from different models: The Roch, 20 January – 15 February 1995. 103
Figure 4.2.3(d)	Simulated flow hydrographs from different models: The Lavant, 8 December 1992 – 1 June 1994 (Chichester flood). 104
Figure 4.2.4	Variation in forecast performance (R^2 and Threshold CSI) with forecast lead time for different models. 110
Figure 4.2.5	Scatter plots of observed and peak flow, and peak timing error for different models at a specified forecast lead time. 113

	Page	
Figure 4.2.6	Examples of fixed lead time evaluation forecasts.	116
Figure 4.2.7	Evaluation mode fixed origin forecasts (a) Silk Stream: 22–23 September 1992.	117
Figure 4.2.7	(cont...) Evaluation mode fixed origin forecasts (b) The Stour: 9–10 April 1998 (Easter 1998 Flood).	118
Figure 4.2.8	Variation in peak forecast accuracy with lead time for a significant flood event in each catchment.	119
Figure 4.2.9	Variation in forecast performance (R^2 and Threshold CSI) with lead time: TF model using error prediction (bold line) and state updating (dotted line).	121
Figure 4.2.10	Variation in forecast performance (R^2 and Threshold CSI) with forecast lead time for different models: TF model with error prediction.	122
Figure 4.3.1	Simulation-mode model hydrographs using raingauge and calibrated radar data: (a) Silk Stream, 21 to 26 September 1989.	127
Figure 4.3.2(a)	Radar rainfall estimates at different scales from Chenies radar for Silk Stream: 21:30 22 September 1992.	130
Figure 4.3.2(b)	Radar rainfall estimates at different scales from Clee Hill radar for the Stour: 06:30 9 April 1998.	131
Figure 4.3.2(c)	Radar rainfall estimates at different scales from Hameldon Hill radar over the Roch: 16:15 8 December 1993.	132
Figure 4.3.3	Scatter plots of simulated and observed peak flow using (i) uncalibrated radar, (ii) calibrated radar and (iii) raingauge rainfall data for all models.	133
Figure 4.3.4	Variation of R^2 and Threshold CSI with forecast lead-time using raingauge and calibrated radar data: Silk Stream.	138
Figure 4.3.5	Variation of R^2 and Threshold CSI with forecast lead-time using raingauge and calibrated radar data: Stour.	139
Figure 4.3.6	Scatter plots of (i) observed and forecast peak flow and (ii) peak timing error, using calibrated radar and raingauge data: Silk Stream, September 1992.	140
Figure 4.3.7	Scatter plots of (i) observed and forecast peak flow and (ii) peak timing error, using calibrated radar and raingauge data: Stour.	141
Figure 4.3.8	Fixed lead-time hydrographs (1 to 6 hours ahead) using calibrated radar and raingauge data.	142
Figure 4.3.9	Evaluation mode fixed-origin forecasts using calibrated radar data: Silk Stream, 22 to 23 September 1992.	143
Figure 4.3.10	Evaluation mode fixed-origin forecasts using calibrated radar data: Stour: 9 to 10 April 1998.	144
Figure B.1.1	Map of catchment relief: Trout Beck.	174
Figure B.1.2	Map of DTM-derived TCM response zones: Trout Beck.	174
Figure B.1.3	Map showing distribution of SAAR across the catchment: Trout Beck.	175
Figure B.1.4	Areas used to define raingauge weights: Trout Beck.	175
Figure B.2.1	Map of catchment relief: Silk Stream.	176

	Page	
Figure B.2.2	Map of DTM-derived TCM response zones: Silk Stream	176
Figure B.2.3	Map showing distribution of SAAR across the catchment: Silk Stream.	177
Figure B.2.4	Areas used to define raingauge weights: Silk Stream.	177
Figure B.3.1	Map of catchment relief: Dove.	178
Figure B.3.2	Map of DTM-derived TCM response zones: Dove.	178
Figure B.3.3	Map showing distribution of SAAR across the catchment: Dove.	179
Figure B.3.4	Areas used to define raingauge weights: Dove.	179
Figure B.4.1	Map of catchment relief: Lavant.	180
Figure B.4.2	Map of DTM-derived TCM response zones: Lavant.	180
Figure B.4.3	Map showing distribution of SAAR across the catchment: Lavant.	181
Figure B.4.4	Areas used to define raingauge weights: Lavant.	181
Figure B.5.1	Map of catchment relief: Rhondda.	182
Figure B.5.2	Map of DTM-derived TCM response zones: Rhondda.	182
Figure B.5.3	Map showing distribution of SAAR across the catchment: Rhondda.	183
Figure B.5.4	Areas used to define raingauge weights: Rhondda.	183
Figure B.6.1	Map of catchment relief: Brue.	184
Figure B.6.2	Map of DTM-derived TCM response zones: Brue.	184
Figure B.6.3	Map showing distribution of SAAR across the catchment: Brue.	185
Figure B.7.1	Map of catchment relief: Stour.	186
Figure B.7.2	Map of DTM-derived TCM response zones: Stour.	186
Figure B.7.3	Map showing distribution of SAAR across the catchment: Stour.	187
Figure B.7.4	Areas used to define raingauge weights: Stour.	187
Figure B.8.1	Map of catchment relief: Roch.	188
Figure B.8.2	Map of DTM-derived TCM response zones: Roch.	188
Figure B.8.3	Map showing distribution of SAAR across the catchment: Roch.	189
Figure B.8.4	Areas used to define raingauge weights: Roch.	189
Figure B.9.1	Map of catchment relief: Witham.	190
Figure B.9.2	Map of DTM-derived TCM response zones: Witham.	190
Figure B.9.3	Map showing distribution of SAAR across the catchment: Witham.	191
Figure B.9.4	Areas used to define raingauge weights: Witham.	191
Figure D.1	Simulated flow hydrographs from different models: Trout Beck at Moor House, 16/11/95-17/11/95.	222
Figure D.2	Simulated flow hydrographs from different models: Silk Stream at Colindeep Lane, 21/6/92-22/6/92.	223
Figure D.3	Simulated flow hydrographs from different models: Dove at Izaak Walton, 1/3/98-15/3/98.	224
Figure D.4	Simulated flow hydrographs from different models: Rhondda at Trehafod, 15/11/92-15/12/92.	225
Figure D.5	Simulated flow hydrographs from different models: Brue at Lovington, 16/11/96-16/12/96.	226

	Page
Figure D.6	Simulated flow hydrographs from different models: Stour at Shipston, 27/12/93-26/1/94. 227
Figure D.7	Simulated flow hydrographs from different models: Roch at Blackford Bridge, 25/2/98-17/3/98. 228
Figure D.8	Simulated flow hydrographs from different models: Witham at Claypole Mill , 30/12/97-24/1/98. 229
Figure D.9	Fixed lead-time evaluation forecasts for different models: Trout Beck at Moor House, 16/11/95-17/11/95. 230
Figure D.10	Fixed lead-time evaluation forecasts for different models: Silk Stream at Colindeep Lane, 23/9/92-24/9/92. 231
Figure D.11	Fixed lead-time evaluation forecasts for different models: Dove at Izaak Walton, 7/3/98-9/3/98. 232
Figure D.12	Fixed lead-time evaluation forecasts for different models: Rhondda at Trehafod, 1/12/92-4/12/92. 233
Figure D.13	Fixed lead-time evaluation forecasts for different models: Brue at Lovington, 20/11/96-21/11/96. 234
Figure D.14	Fixed lead-time evaluation forecasts for different models: Stour at Shipston, 6/1/94-7/1/94. 235
Figure D.15	Fixed lead-time evaluation forecasts for different models: Roch at Blackford Bridge, 7/3/98-8/3/98. 236
Figure D.16	Fixed lead-time evaluation forecasts for different models: Witham at Claypole Mill, 3/1/98-6/1/98. 237
Figure D.17	Fixed lead-time evaluation forecasts for different models using calibrated radar data: Silk Stream at Colindeep Lane, 23/9/92-24/9/92. 238
Figure D.18	Fixed lead-time evaluation forecasts for different models using calibrated radar data: Stour at Shipston, 10/4/98-11/4/98. 239
Plate 2.6.2	Flooding at Croft on the Tees, March 1968 46
Plate 2.6.3	Flooding of the Silk Stream in Edgeware 48
Plate 2.6.4	Flooding in the Dove catchment at Hanging Bridge, 20 December 1991 50
Plate 2.6.5	Flooding of the River Lavant near Chichester 1994 52
Plate 2.6.6	Flooding of the River Rhondda at Trehafod, 1979 54
Plate 2.6.7	Flooding of the Brue at Bruton, July 1982 56
Plate 2.6.8	Flooding of the River Avon at Evesham, downstream of the confluence with the Stour, Easter 1998 58
Plate 2.6.9	Flooding of the Roch January 1995 60
Plate 2.6.10	Flooding of the River Witham at Lincoln, downstream of Bargate Weir 62

EXECUTIVE SUMMARY

The purpose of the project “Comparison of Rainfall-Runoff Models for Flood Forecasting” is to provide guidance to the Environment Agency on the choice of rainfall-runoff model for use in different catchments for flood forecasting purposes. A literature review of models presented in the Part 1 Report recognised that whilst there is a plethora of “brand-name” models there is much similarity between many of them. A rather small set of model functions is common to many models and they differ in the detail of their configuration. Eight models were selected for a more detailed assessment of performance using data from nine catchments of varied character and spread throughout the regions of the Agency. The results are reported in this Part 2 Report along with conclusions and recommendations.

The chosen models encompass those used operationally by the EA together with one overseas model and a simple distributed model previously developed for the Agency. Four of the models are lumped, conceptual models with continuous water accounting procedures: the Thames Catchment Model (TCM), the Midlands Catchment Runoff Model (MCRM), the Probability Distributed Moisture model (PDM), and the US National Weather Service Sacramento model (NWS). A fifth model, the Isolated Event Model (IEM), is an event model modified to operate continuously in real-time. Water balance principles are used for soil moisture accounting and water storage routing but an empirical function links the two components, controlling runoff production as a function of soil moisture. The sixth model is a simple Transfer Function (TF) model whilst the seventh is a constrained form of TF model, referred to as the Physically Realisable Transfer Function (PRTF). The TF types of model are black-box models which empirically relate rainfall and flow, can be related to unit hydrographs and can be subject to conceptual interpretation as forms of routing function. The last model, the Grid Model, is included as a simple form of distributed conceptual rainfall-runoff model suitable for use in flood forecasting and able to use weather radar estimates of rainfall in grid form. Each model is associated with an updating procedure whereby recent measurements of flow are incorporated into the model so as to improve forecast performance in real-time.

The strategy for assessment used is based on first calibrating all models in “simulation-mode”, where each model is used to transform rainfall (and potential evaporation) to runoff without using flow to update the model forecast. Each model is then evaluated using periods of data not used for calibration. This simulation-mode evaluation serves to focus on the process model capabilities of a given model. Subsequently, each model is evaluated in “forecast-mode” in which an updating scheme is used to incorporate measurements of flow up to the “forecast time-origin”. This emulates the forecast performance expected operationally at different forecast lead times. Perfect foreknowledge of rainfall is assumed so as not to confound the model assessment with uncertainties in rainfall forecasts. The statistics used for model assessment are R^2 and a Threshold CSI (Critical Success Index). The R^2 statistic, giving the proportion of the variability in the flow accounted for by the model forecasts, is used to provide a broad guide to model performance. The Threshold CSI statistic is used to judge the efficacy of a model to correctly forecast the exceedence of a set of flow thresholds, particularly relevant to the use of a forecast to trigger an alert level of a given severity. Forecasts are also judged more informally via hydrograph plots and scatter plots of observed and forecast flood peaks. Whilst the main assessment relates to the use of raingauge estimates of rainfall as input to the models, for three of the catchments the assessment extends to the use of weather radar, both in raw form and as raingauge-calibrated radar estimates of rainfall.

Whilst forecast accuracy is the focus of the model assessment, other issues are taken into consideration including ease of model configuration, initialisation and calibration.

The form of model assessment used, employing long continuous records at a fixed 15 minute time-step typically eight months in duration, has meant that it has been difficult to emulate the operational performance of TF and PRTF models. These models are used by the Agency in “event mode” and commonly operate on baseflow separated runoff where baseflow is taken as the flow at the start of the event. The opportunity exists to manually adjust the model parameters affecting the volume, shape and timing of the forecast as the flood develops. Also, the model time-step and model order are commonly chosen with regard to the response characteristics of the catchment. The results reported here relate to TF/PRTF models without baseflow separation, using an automated method of model parameter adjustment and using a fixed model time-step and model order. The approach most resembles that used on the River Medway to support the operation of the Leigh Flood Barrier and the method of automated model gain adjustment has also been used in Anglian Region. The results relating to TF/PRTF models should be interpreted against this background.

Overall, the results suggest that no one model consistently out-performs all others across all catchments. The TCM is one of the best performing models when judged using the R^2 statistic whilst the PDM is more successful according to the Threshold CSI criterion (relevant to the issuing of flood alerts) and in forecasting flood peaks. Whilst the TCM is the most complex model and can be a challenge to calibrate, the PDM is of intermediate complexity. For a simple model, the IEM is surprisingly successful, particularly in terms of Threshold CSI. The simplest models, the TF and PRTF, are easiest to calibrate and initialise and can provide acceptable forecasts for some catchments. For smaller catchments in particular, TF models compare favourably with other models when used with error prediction rather than state updating. The MCRM proved sensitive to initial conditions of soil moisture but can work well on small-to-medium sized catchments. The NWS model, despite its large number of parameters, proved easy to calibrate using automatic optimisation and provided reasonable performance. Use of radar data gave as good, and sometimes slightly better results than using raingauge data alone, provided the radar was functioning well, and raingauge-calibration generally helped. The Grid Model was the only distributed model assessed and can utilise radar data in grid form. For the three catchments on which it was evaluated in simulation-mode it consistently gave the second best model simulations in terms of R^2 but did less well according to the Threshold CSI criterion.

Operationally, the TCM, PDM and IEM models appear to be the most appropriate flood forecasting models to use, of those assessed, the choice depending on the complexity of catchment response whilst all models have value in the right situation. The advantage of model familiarity acquired through past use is employed to guide more specific recommendations for each EA region. It is recommended that more automated applications of TF/PRTF models be pursued which accommodate the effects of catchment wetness on runoff production through effective rainfall transformations and incorporate baseflow via parallel “fast” and “slow” transfer function routing components. Opportunities for further research on model formulation and configuration, updating schemes, and catchment-scale rainfall estimation are identified.

KEYWORDS

Rainfall-runoff model, Floods and Flooding, Flood forecasting, Flood warning, Modelling (Hydrological).

1. INTRODUCTION

1.1 Background

The purpose of the project “Comparison of Rainfall-Runoff Models for Flood Forecasting” is to provide guidance to the Environment Agency on the choice of rainfall-runoff model for use in different catchments for flood forecasting purposes. This has been approached in two phases, the first phase providing a literature review of models whilst the second phase assessed the performance of a selected set of models across a range of catchments. The first phase is reported in EA R&D Technical Report W242, referred to here as the Part 1 Report, whilst the second phase is the subject of this Part 2 Report.

The Part 1 Report recognised that whilst there is a plethora of “brand-name” models there is much similarity between many of them. A rather small set of model functions are common to many models and they differ in the detail of their configuration. Eight models were selected to carry through to the second phase of the project, involving assessment of performance using data from a variety of catchments. The chosen models encompass those used operationally by the EA together with one overseas model and a simple distributed model previously developed for the Agency. Four of the models are lumped, conceptual models with continuous water accounting procedures: the Thames Catchment Model (TCM), the Midlands Catchment Runoff Model (MCRM), the Probability Distributed Moisture model (PDM), and the US National Weather Service Sacramento model (NWS). A fifth model, the Isolated Event Model (IEM), is an event model modified to operate continuously in real-time. Water balance principles are used for soil moisture accounting and water storage routing but an empirical function links the two components, controlling runoff production as a function of soil moisture. The sixth model is a simple Transfer Function (TF) model whilst the seventh is a constrained form of TF model, referred to as the Physically Realisable Transfer Function (PRTF). The TF types of model are black-box models which empirically relate rainfall and flow (or baseflow separated flow), can be related to unit hydrographs and can be subject to conceptual interpretation as forms of routing function. The last model, the Grid Model, is included as a simple form of distributed conceptual rainfall-runoff model suitable for use in flood forecasting and able to use weather radar estimates of rainfall in grid form, as opposed to an average over the catchment.

Whilst a detailed description of these eight models is provided in the Part 1 Report, the next section provides an overview of each model by way of background. This includes information on the type of updating procedure used with each model type. The term “updating” refers to the technique used to incorporate recent measurements of flow so as to improve model performance in real-time. Section 1.3 outlines the selection of the nine catchments to be used in the detailed assessment of models. A more detailed description of the catchments and the supporting data for each is deferred to Section 2.

Section 2 outlines the strategy for assessment to be used. It is based on first calibrating all models in “simulation-mode”. In this mode each model is used to transform rainfall (and potential evaporation) to runoff without using flow to update the forecast from the model. Each model is then evaluated using periods of data not used for calibration. This simulation-mode evaluation serves to focus on the process model capabilities of a given model. Subsequently, each model is evaluated in “forecast-mode” in which an updating scheme is used to incorporate measurements of flow up to the “forecast time-origin”. This emulates the

forecast performance expected operationally at different forecast lead times. However, in order to focus on comparing models, perfect foreknowledge of rainfall is assumed. Section 2 also introduces the statistics used for model assessment. The R^2 statistic, giving the proportion of the variability in the flow accounted for by the model forecasts, is used to provide a broad guide to model performance. A Threshold CSI (Critical Success Index) statistic is used to judge the efficacy of a model to correctly forecast the exceedence of a set of flow thresholds, particularly relevant to the use of a forecast to trigger an alert level of a given severity. Forecasts are also judged more informally via hydrograph plots and scatter plots of observed and forecast flood peaks.

Sections 3 and 4 present the results of the model calibration and evaluation and form the core of the report. Whilst the main results relate to the use of raingauge estimates of rainfall as input to the models, for three of the catchments the assessment extends to the use of weather radar, both in raw form and as raingauge-calibrated radar estimates of rainfall. Section 5 discusses other issues relating to model choice beyond performance assessment, focusing on ease of model use. Finally, Section 6 provides a summary of the report and the main conclusions reached. A set of recommendations is presented relating to both general issues and more specific operational aspects. Opportunities for further work are identified.

1.2 Models for Intercomparison

The eight rainfall-runoff models selected for assessment across a variety of catchments are outlined here by way of background. Seven of the models are introduced in order of complexity starting with the Thames Catchment Model (TCM) as the most complex and ending with the Physically Realisable Transfer Function model (PRTF) as the simplest. The Grid Model is presented last as a special case, being a distributed flood forecasting model designed for use with weather radar data. This order is maintained in the model assessment that follows.

Thames Catchment Model (TCM)

The structure of the Thames Catchment Model, or TCM (Greenfield, 1984), 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 being 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. In this study the same, catchment-average, rainfall is used for all zones.

The conceptual representation of a hydrological response zone in the TCM is illustrated in Figure 1.2.1 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 (Penman, 1949) in which a near-surface storage, of depth related to the rooting depth of the associated

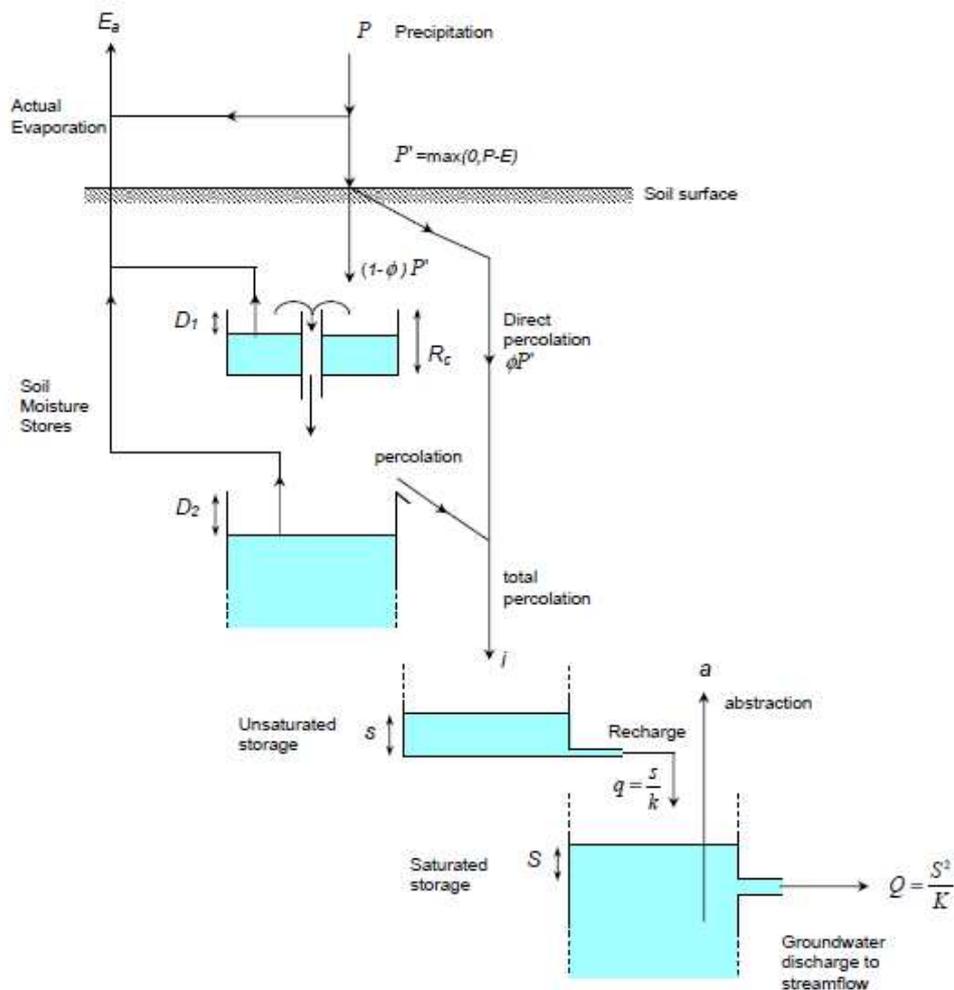


Figure 1.2.1 Representation of a hydrological response zone within the Thames Catchment Model.

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 whilst 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 1.2.1.

Table 1.2.1 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, $c\Delta t/\Delta x$

US National Weather Service Sacramento Model (NWS)

The US National Weather Service (NWS) rainfall-runoff model is also called the Sacramento Soil Moisture Accounting Model or simply the Sacramento Model. It was developed in the early 1970s at the NWS River Forecast Centre in Sacramento (California), principally by Bob Burnash and Larry Ferral, as a classic lumped, conceptual, soil moisture accounting model. The basic source document is the report by Burnash *et al.* (1973).

A schematic of the model is shown in Figure 1.2.2 which highlights that the model comprises three principal storages:

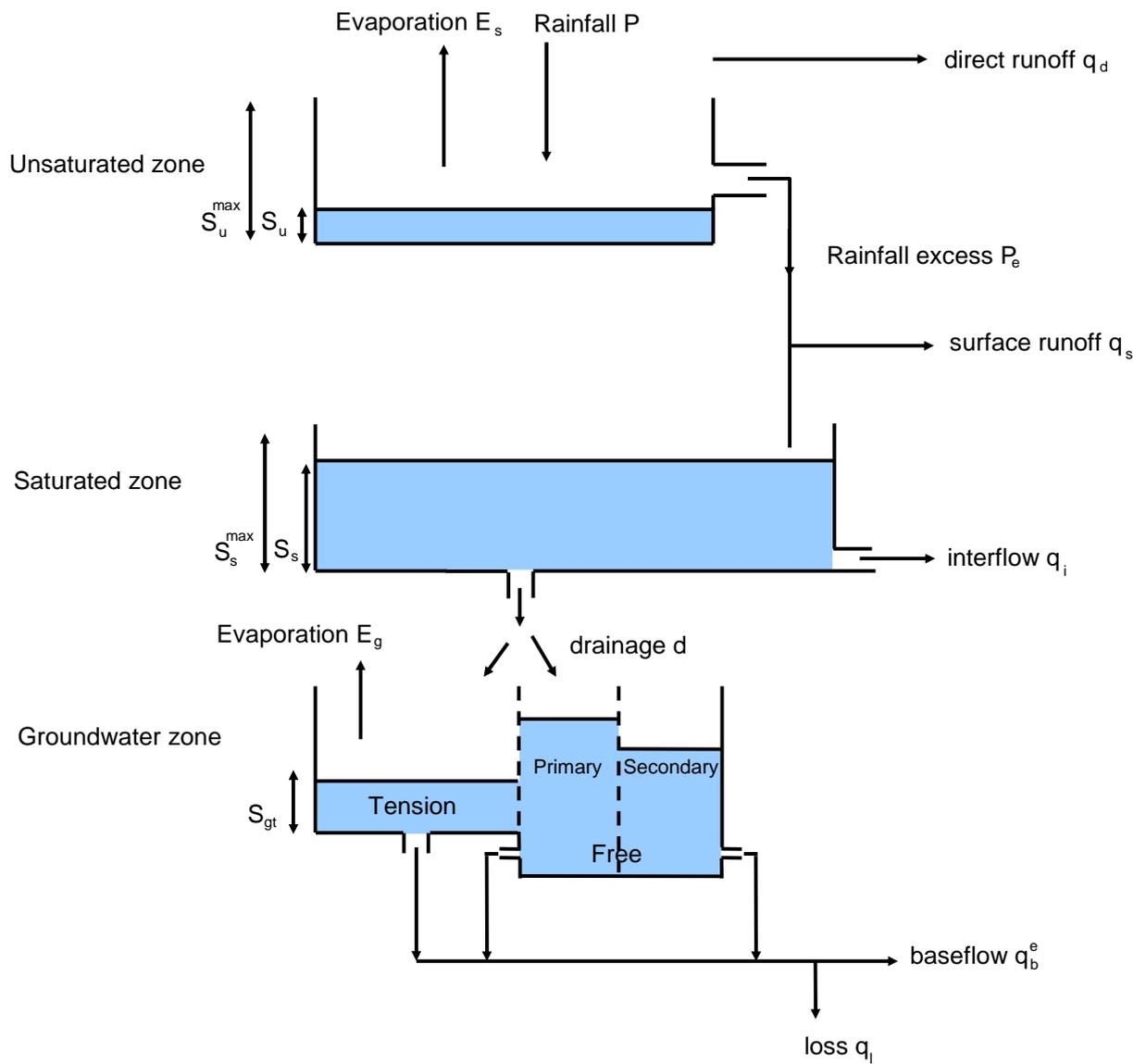


Figure 1.2.2 The NWS Model.

- (i) unsaturated zone store generating direct runoff to the basin outlet and rainfall excess feeding the saturated zone below after a proportion contributes to surface runoff;
- (ii) saturated zone store generating interflow and draining downwards as percolation to the groundwater zone; and
- (iii) groundwater zone store which is divided into water held under tension and water that is free to drain, both contributing to baseflow after losses have been taken into account.

A general outline of the NWS model formulation follows.

Evaporation from the unsaturated zone storage, E_s , reduces linearly with water in storage, S_u , from the potential rate, E , when the store is full to capacity, S_u^{\max} , to zero when empty. During periods of rain, direct runoff, q_d , is generated from the fraction of the catchment that is impervious, f . This fraction is made up of fixed impervious and replenished tension water fractions, f_i and f_w respectively. Rainfall excess is then calculated by continuity and passes downwards to the saturated zone store. Surface runoff, q_s , is generated as a fraction $(1-f)$ of the rainfall excess leaving a residual rainfall excess to enter the saturated zone store. Interflow, q_i , from the saturated zone store is proportional to the water in the saturated zone store, S_s , with an adjustment for the potential (maximum) impervious fraction, $f_{\max}=f_i+f_w^{\max}$. The lateral flows generated from the unsaturated and saturated zone storages are summed and routed using a classical unit hydrograph convolution. Channel evaporation is accounted for as a simple fraction, c , of the potential evaporation.

Drainage (percolation) from the saturated zone into the groundwater zone occurs as a function of the degree of saturation in the saturated zone and the deficit in the groundwater zone. Drainage to groundwater is split between tension water S_{gt} and free water S_{gf} . Tension water supplies evaporation loss from groundwater as a function of the potential evaporation still to be satisfied and the proportion of total groundwater storage that contains tension water. Free water is split between primary and secondary compartments, with maximum storage capacities S_{gp}^{\max} and S_{gs}^{\max} respectively, and generating separate primary and secondary baseflows. These are summed to give the total baseflow. Effective baseflow is calculated after taking into account losses, q_l . The total flow at the basin outlet is given by the sum of the effective baseflow and the routed lateral flows from the unsaturated and saturated zone storages. The parameters of the NWS Model are summarised in Table 1.2.2.

Midlands Catchment Runoff Model (MCRM)

The Midlands Catchment Runoff Model (MCRM) (Bailey and Dobson, 1981; Wallingford Water, 1994) comprises three main stores: an interception store, a soil moisture store and a groundwater store (Figure 1.2.3). Rapid runoff is generated from the soil moisture store, the proportion of the input to the store becoming runoff increasing exponentially with decreasing soil moisture deficit. “Percolation” to the groundwater store occurs when the soil is supersaturated, increasing as a linear function of the negative deficit. When supersaturation exceeds a critical value, “rapid drainage” also occurs as a power function of the negative deficit in excess of the critical value (the so-called excess water). This rapid drainage along with rapid runoff forms the soil store runoff. Evaporation occurs preferentially from the

Table 1.2.2 Parameters of the NWS Model

Parameter	Unit	Description
r_f	none	Rainfall factor
f_i	none	Fraction of the catchment that is impervious
f_w^{\max}	none	Maximum additional fraction of impervious area which develops as tension water requirements are met
c	none	Fraction of the catchment covered by streams, lakes and riparian vegetation
S_u^{\max}	mm	Capacity of unsaturated zone tension water store
S_s^{\max}	mm	Capacity of unsaturated zone free water store
k_i	day ⁻¹	Rate of interflow from saturated zone
γ	none	Proportional increase in percolation from saturated to dry conditions
δ	none	Exponent in equation for percolation rate
S_{gt}^{\max}	mm	Capacity of groundwater zone tension water store
S_{gs}^{\max}	mm	Capacity of groundwater zone secondary free water storage
k_{gs}	day ⁻¹	Lateral drainage rate from secondary groundwater zone
S_{gp}^{\max}	mm	Capacity of groundwater zone primary free water storage
k_{gp}	day ⁻¹	Lateral drainage rate from primary groundwater zone
p	none	Fraction of percolated water going directly to groundwater zone free water store in preference to tension water store
r_s	none	Fraction of groundwater zone free water not available for resupplying lower zone tension water store

interception store at a rate which is a fixed proportion of the catchment potential evaporation. A proportion of any residual evaporation demand is then met by water in the soil store, the proportion varying as a function of the soil moisture deficit. Drainage of the groundwater store to baseflow varies as a power function of water in storage, the exponent being fixed at 1.5. The total output, made up of baseflow and soil store runoff, is then lagged and spread evenly over a specified duration to represent the effect of translation of water from the ground to the catchment outlet. Finally, the flow is smoothed using two nonlinear storage functions, one for routing in-bank flow and the other out-of-bank flow, the two components being summed to give the catchment model outflow. A summary of the parameters involved in the MCRM are presented in Table 1.2.3.

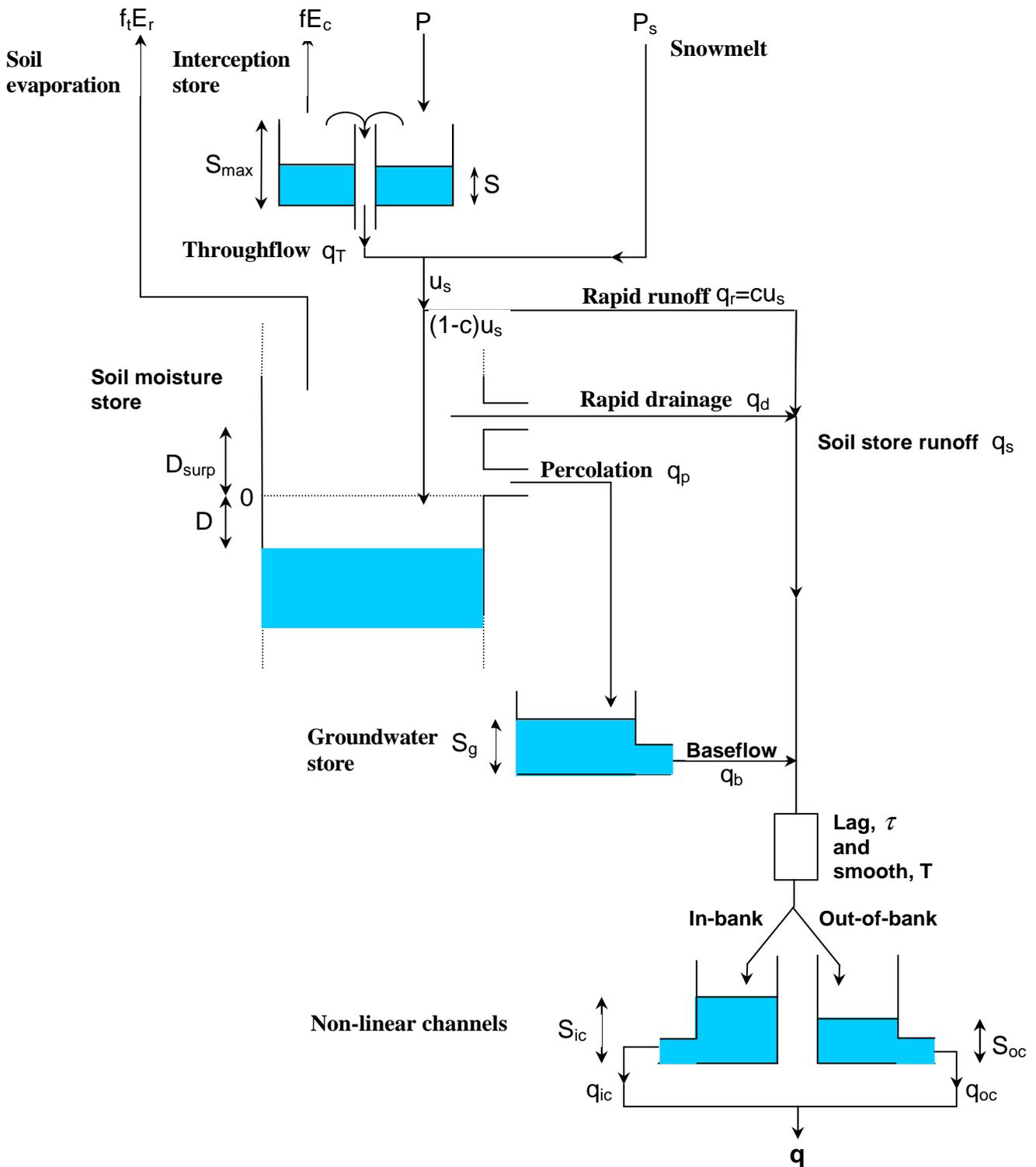


Figure 1.2.3 The Midlands Catchment Runoff Model.

Table 1.2.3 Parameters in the Midlands Catchment Runoff Model

Parameter	Unit	Description
f_c	none	Rainfall factor
S_{\max}	mm	Capacity of interception store
f	none	Fraction of catchment evaporation potentially met by interception storage
c_0	none	Minimum value of rapid runoff proportion
c_1	mm^{-1}	Parameter in rapid runoff proportion function
c_{\max}	none	Maximum value of rapid runoff proportion
q_p^{\max}	mm h^{-1}	Maximum percolation rate
D_{surp}	mm	Maximum soil store moisture surplus
γ_d	none	Soil function exponent controlling rapid drainage
k_d	h mm^{γ_d-1}	Soil function coefficient controlling rapid drainage
T_p	none	Potential transpiration factor
T_m	none	Minimum transpiration factor
E_{\max}^D	mm	Deficit below which potential transpiration factor applies
E_{\min}^D	mm	Deficit above which minimum transpiration factor applies
K_g	$\text{h mm}^{0.5}$	Time constant in baseflow storage function
τ	h	Time lag applied to total runoff
T	h	Duration of time spread applied to total runoff
S_{bf}	mm	Channel storage at bankfull
k_{cr}	$\text{h}^{-1} \text{mm}^{1-\gamma_{\text{cr}}}$	In-channel routing storage coefficient
γ_{cr}	none	In-channel routing storage exponent
k_{or}	$\text{h}^{-1} \text{mm}^{1-\gamma_{\text{or}}}$	Out-of-bank channel routing storage coefficient
γ_{or}	none	Out-of-bank channel routing storage exponent

Probability Distributed Model (PDM)

The Probability Distributed Moisture model, or PDM, is a fairly general conceptual rainfall-runoff model which transforms rainfall and evaporation data to flow at the catchment outlet. Figure 1.2.4 illustrates the general form of the model. The PDM has been 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. Here, only a standard form of the PDM has been used.

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 used here 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 restricted here to be of cubic form, with outflow proportional to the cube of the water in store. An extended subsurface storage component is 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 contribute 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. The parameters involved in the standard form of PDM model are summarised in Table 1.2.4.

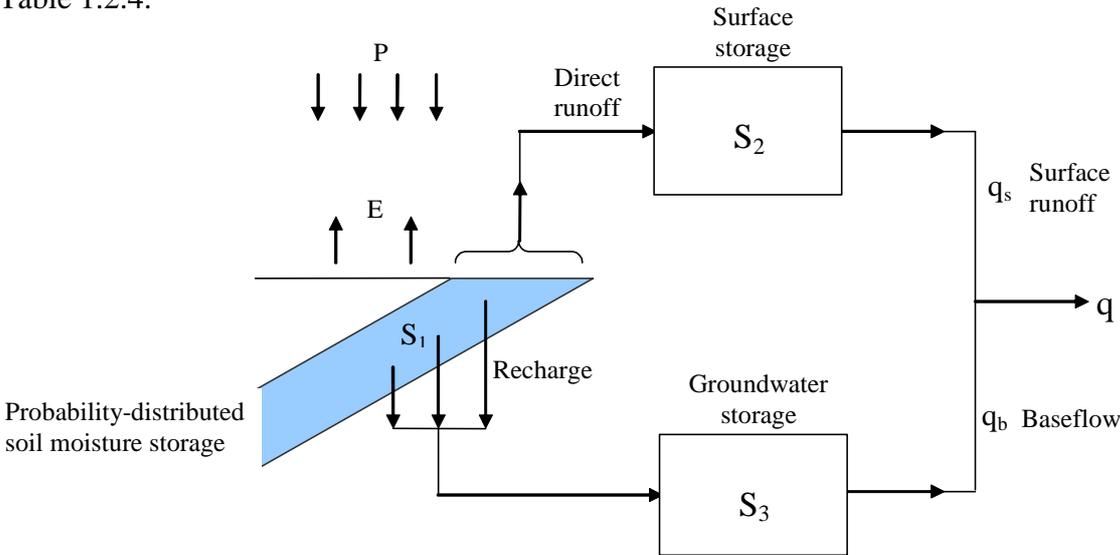


Figure 1.2.4 The PDM rainfall-runoff model.

Table 1.2.4 Parameters of the PDM model

Parameter name	Unit	Description
f_c	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_s	h	time constant of cascade of two equal linear reservoirs ($k_s=k_1=k_2$)
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

Isolated Event Model (IEM)

The Isolated Event Model, or IEM, was originally developed for design applications as part of the UK Flood Studies Project (NERC, 1975). In many respects it is very similar to the single zone representation of the Thames Catchment Model in using the Penman stores concept and a quadratic reservoir for routing. However, the use of the Penman stores concept is not done as part of an explicit soil moisture accounting procedure as is the case with the TCM. Rather, the soil moisture deficit it provides is used as an index of catchment wetness within an empirical equation which relates the proportion of rainfall that becomes runoff - the runoff coefficient - to the soil moisture deficit, D . Specifically the runoff coefficient is defined by the exponential function, $f = \alpha \exp(-\beta D)$, where β is a parameter with units (mm water)⁻¹ and α is a dimensionless parameter. Note that the IEM uses as standard a Penman upper store of depth 75 mm, the root constant for short grass, with no bypassing ($\phi=0$). Because the original formulation was event-based and for design, the runoff coefficient, f , was applied to the whole storm and D was the soil moisture deficit at the start of the storm. The parameter α can be interpreted as a “gauge representativeness factor” since, with zero deficit (saturated conditions), a proportion α of the rain becomes runoff. The storm rainfall time series is multiplied by the factor f to give an “effective rainfall” series. This is then subject to a time

delay before being used as input to the quadratic storage reservoir, the outflow from which forms the IEM model flow prediction.

For real-time flood forecasting applications the concept of an “event” is an awkward notion to work with. The IEM has been modified for real-time use by redefining the runoff coefficient, f , to be a time variant function of the deficit D . Thus, we have $f_t = \alpha \exp(-\beta D_t)$. The calculation of D_t is calculated continuously, within and between storm events, using the Penman stores water accounting procedure. No use is made by the IEM of the outflows from the Penman stores, only the deficit as an index of catchment wetness and its impact on the ensuing volume of flood runoff.

Further modifications of the classical IEM formulation resulted from trials undertaken in the context of the study by Moore *et al.* (1993). The first is to replace rainfall by net rainfall (rainfall less evaporation) prior to applying the factor f_t to yield effective rainfall. The second modification is to replace the simple time delay on the effective rainfall by a triangular time delay function. Thus the inflow to the quadratic storage is a weighted combination of delayed effective rainfalls up to the current time, with the weighting defined by a triangular function. The final modification is that a constant flow, q_c , can be added to the outflow from the quadratic storage to give the total basin outflow.

The similarity between the IEM and a single zone of the TCM has been exploited by implementing the IEM as a variant on the TCM, with the overall model code being referred to as the PSM (Penman Store Model). The IEM parameters are as for a TCM zone with $n_z=1$, A equal to the catchment area, $R_c=75$, $R_p=999$ and ϕ , k , a , τ_d and N set to zero. The remaining parameters, together with additional parameters specific to the IEM, are listed in Table 1.2.5.

Table 1.2.5 Parameters of the Isolated Event Model

Parameter name	Unit	Description
α	None	Coefficient in runoff proportion equation
β	None	Exponent in runoff proportion equation
K	mm h	Quadratic storage constant
τ_s	h	Delay to start of smoothing triangle
τ_p	h	Delay from start to peak of smoothing triangle
τ_e	h	Delay from start to end of smoothing triangle
q_c	$m^3 s^{-1}$	Constant flow

Transfer Function Model (TF)

Transfer Function or TF models are a class of time-series models popularised by Box and Jenkins (1970). They are linear models with which an output variable can be forecast as a linear weighted combination of past outputs and inputs. In a rainfall-runoff context the output is usually flow (or baseflow separated flow) and the input rainfall (or effective rainfall). Any residual model error can be represented through a noise model which is normally of autoregressive moving average (ARMA) form. The overall model is termed a Transfer Function Noise, or TFN, model.

A linear transfer function model relates an output at time t , y_t , to r previous values of the output and s previous values of an input with delay b , u_{t-b} , such that

$$y_t = -\delta_1 y_{t-1} - \delta_2 y_{t-2} - \dots - \delta_r y_{t-r} + \omega_0 u_{t-b} + \omega_1 u_{t-b-1} + \dots + \omega_{s-1} u_{t-b-s+1}$$

where $\{\delta_i\}$ are r autoregressive parameters and $\{\omega_i\}$ are s moving average parameters operating on the past outputs and inputs respectively. With y_t as basin runoff (or baseflow separated runoff) and u_t as rainfall (or effective rainfall) this TF model can be used as a simple rainfall-runoff model. The notation TF(r,s,b) is used to indicate the order of the model in terms of the number of parameters and the time delay.

The TF model is equivalent in form to the linear model

$$y_t = v_0 u_{t-b} + v_1 u_{t-b-1} + v_2 u_{t-b-2} + \dots$$

where the weights v_0, v_1, v_2, \dots define the model's impulse response function (equivalent to the unit hydrograph for effective rainfall as input and baseflow separated runoff as the output). In general the number of parameters $r+s$ in the transfer function representation is far fewer than in the impulse function representation: this is strictly infinite although in practice can be treated to correspond to a significant memory length. The transfer function model thus offers a parsimonious parameterisation of a linear system response.

The model output, y_t , can be related to the observed output, Y_t , through the relation

$$Y_t = y_t + \eta_t$$

where $\eta_t = Y_t - y_t$ is the simulation-mode model error. This model error may be represented by an ARMA error predictor (discussed later) to obtain real-time updated forecasts. In this form, the overall model is referred to as a Transfer Function Noise (TFN) model as popularised by Box and Jenkins (1970).

A special case of the TFN formulation, referred to as Autoregressive Moving Average on exogenous inputs or ARMAX, is given by

$$Y_t = -\delta_1 Y_{t-1} - \delta_2 Y_{t-2} - \dots - \delta_r Y_{t-r} + \omega_0 u_{t-b} + \omega_1 u_{t-b-1} + \dots + \omega_{s-1} u_{t-b-s+1} + \xi_t$$

where ξ_t also represents model error and can be represented by an ARMA noise model structure. If possible dependence in the model error is not explicitly represented then the above can be used to justify the TF prediction equation

$$y_t = -\delta_1 Y_{t-1} - \delta_2 Y_{t-2} - \dots - \delta_r Y_{t-r} + \omega_0 u_{t-b} + \omega_1 u_{t-b-1} + \dots + \omega_{s-1} u_{t-b-s+1}.$$

It is this TF predictor that is most commonly used as the basis of operational forecasts by the Environment Agency regions using TF models. The predictor simply operates to form a forecast as a weighted sum of present and past flows and lagged rainfall inputs. (The flows may be baseflow separated with baseflow taken as the flow at the start of an event. This possible distinction is assumed below without further comment.) Observed values of flow are used in the right hand side of the above equation but as the forecast lead time increases the latest forecast value replaces the not-yet-available observed flow at future times. This forecast formulation in which observed flow values are used directly can be referred to as “full state correction”.

With the input-output pair of a TF model being rainfall-runoff then the nonlinearity known to exist by hydrologists is clearly not represented explicitly. The state correction formulation is one way of reducing the effect of this weakness. Allowing the model parameters to be time-variant and tracking the variation using a recursive estimation scheme provides other opportunities for improvement. For example, Cluckie and Owens (1987) employ a TF model in such a way that a single model gain parameter, G_t , controlling the proportion of rainfall that becomes runoff, is recursively estimated. Specifically, they use the reparameterised TF model

$$y_t = -\delta_1 Y_{t-1} - \delta_2 Y_{t-2} - \dots - \delta_r Y_{t-r} + G_{t-1} (\omega_0 u_{t-b} + \omega_1 u_{t-b-1} + \dots + \omega_{s-1} u_{t-b-s+1})$$

for forecasting, with the time-varying model gain parameter calculated as

$$G_t = \mu G_{t-1} + (1 - \mu) \frac{Y_t + \delta_1 Y_{t-1} + \dots + \delta_r Y_{t-r}}{\omega_0 u_{t-b} + \omega_1 u_{t-b-1} + \dots + \omega_{s-1} u_{t-b-s+1}}.$$

Here, μ is a smoothing factor in the range (0,1) used to dampen out erratic fluctuations in G_t . This form of TF model with time-varying model gain is included here in the assessment of models using catchment data. It has been used operationally for flood forecasting in Anglian (Page, 1991), and Southern (Pollard, undated) regions of the Environment Agency. In Anglian region the output has been taken to be baseflow separated runoff. Also two sets of model parameters are sometimes used to cope with different responses under “fast response” and “average” conditions. In the assessment that follows the output is taken to be total flow and only a single set of model parameters has been used. This arises from the use of continuous long records in the assessment, typically of eight months duration, and where the concept of an event required to define baseflow has no place.

A related approach in focussing on real-time tracking of the model gain is used in the Nith flood forecasting system in Scotland (Lees *et al.*, 1993). In this case the model gain is tracked using recursive least squares, assuming a random walk process for the parameter variability. A drawback of such approaches involving recursive parameter updating is that the variation is merely “tracked” and not “anticipated”. Our understanding of hydrological science, for example, tells us that antecedent wetness can influence the gain or runoff proportion and that

soil moisture accounting model components can be used to anticipate this effect. This leads one to recognise that the role of the transfer function is primarily that of a linear routing operation and can be incorporated as such into a conceptual model as merely one component form. However, an important purpose of this study is to establish whether TF models used in practice provide acceptable model performance when compared to other models.

Another approach to accommodating nonlinear effects in a linear TF model is through the use of a nonlinear loss function to transform rainfall to “direct runoff” or “effective rainfall” and using this as the input variable u_t . Functionally, the transfer function serves as a simple linear routing function. Alternatively, a parallel system of two transfer function models can be envisaged together with a partitioning rule which directs rainfall to the two functions which operate as slow and fast translation pathways. A variety of nonlinear loss functions and parallel TF model functions were investigated in the UK for use in flood forecasting (Moore, 1980, 1982). Most recently, improved estimation schemes for this class of parallel TF model have been developed (see, for example, Young, 1992, Jakeman *et al.*, 1990) which overcome some of the problems encountered in this earlier work. These modified forms of TF model are not used by the EA operationally and are therefore not assessed here.

Physically Realisable Transfer Function (PRTF) Model

The basic idea in formulating the Physically Realisable Transfer Function, or PRTF, model (Han, 1991) is to choose a parameterisation which constrains the impulse response function to have a physically realistic form in a hydrological context. Principally, this means that it should be positive and not exhibit oscillatory behaviour (it is stable). The basic idea in the PRTF formulation is to replace the set of autoregressive parameters, $\delta_1, \delta_2, \delta_3, \dots$, by a single parameter, β , related to them by

$$\delta_i = (-\beta)^{-i} C_r^{r-i},$$

where the combinatorial has the standard definition

$$C_k^r = \frac{r!}{(r-k)!k!} = \frac{r(r-1)\dots(r-k+1)}{k!}.$$

This is referred to as an “equal root” parameterisation and gives a stable impulse response function for $\beta > 1$. An important feature of the equal root parameterisation is that it allows the r autoregressive parameters of the TF model to be reduced to one, the root β , through the use of the above relation. However, the form of TF model is restricted as a result.

It is of interest to note special cases of the above. For dependence on one past output ($r=1$) we have $\delta_1 = -1/\beta$ and for two past outputs ($r=2$) $\delta_1 = -2/\beta$ and $\delta_2 = 1/\beta^2$. The impulse response function for a single, unlagged input ($s=1, b=0$) is $v(t) = \beta^{-t}$ for $r=1$ and $v(t) = (1+t)\beta^{-t}$ for $r=2$. Han (1991) suggests that choosing r to be 2 or 3 provides sufficient flexibility of the impulse response function, provided the moving average parameters $\{\omega_i\}$ can take on negative values so as to lower the recession limb. To make the model more physically intuitive the equal root parameterisation β is substituted by the time-to-peak, t_{peak} , of the impulse response function. For $r=2$ when $v(t) = (1+t)\beta^{-t}$ we have for the reparameterisation

$$\beta = \exp\left\{\frac{1}{(1+t_{peak})}\right\}.$$

Note that t_{peak} is actually the time-to-peak of the impulse response function corresponding to the autoregressive part of the PRTF model, excluding moving average and pure time delay effects. Unless the moving average parameters are constrained to be positive then the parameter is better interpreted as indexing the rate at which the tail decays and fails to be indicative of the time-to-peak.

Han (1991) recognises that the TF model, with its fixed impulse response function, will not provide an adequate representation of the rainfall-runoff process which is both nonlinear and time variant. He chooses to address this problem by adjusting the form of the impulse response function to reflect each flood situation as it is encountered in real-time. To ease this task Han introduces three types of adjustment factor designed to alter the volume, shape and time response of the TF model. For volume adjustment the moving average parameters, $\{\omega_i\}$, are scaled using a factor α , the proportion of volume change, such that the adjusted parameters are given by

$$\omega_i^* = (1 + \alpha)\omega_i \quad i = 0, 1, \dots, s-1.$$

Note that the autoregressive parameters, $\{\delta_i\}$, are not affected by this adjustment.

The shape of the impulse response function is changed with reference to a shift in the position of the peak of the autoregressive part of the impulse response function. The shape adjustment factor, γ , is defined as

$$\gamma = t_{peak}^* - t_{peak}$$

where t_{peak}^* denotes the adjusted peak time. For $r=2$ this may be expressed in terms of the equal root parameterisation, β of the original model and the adjusted model β^* , to give

$$\beta^* = \exp\left\{\left(\gamma + \frac{1}{\ln \beta}\right)^{-1}\right\}.$$

It follows that the adjusted autoregressive parameters are obtained by substituting the above in

$$\delta_i^* = (-\beta^*)^{-i} C_{r-i}^r.$$

The third form of adjustment is to time shift the impulse response system. This simply involves a change to the pure time delay parameter, b , used to delay the rainfall inputs to the transfer model.

Wedgwood (1993) recognised the difficulty of implementing such simple adjustments, especially for fast responding catchments and where forecasts from many catchments may be

required. He explored knowledge-based procedures which employ logical rules, developed from an analysis of synthetic and historical storm data, to automate the adjustment of the PRTF model. A drawback of the approach is the initial acquisition of knowledge concerning the thresholds, linkages and relationships involved.

Presently, the PRTF model is used in North-west and South-west regions in a form that requires the user to manually adjust the three factors controlling the volume, shape and time response of the model as the flood develops to gain better agreement between past observed and forecast flows. The approach is not automatic or objective and could not be satisfactorily incorporated in the model assessment procedure used here. The PRTF model has been incorporated in the assessment as a restricted form of TF model and used in conjunction with two forms of updating: ARMA error prediction and state updating (full state correction with model gain updating).

To restrict the scope of assessment of both TF and PRTF models, the same order of model has been used throughout. This has dependence on two past flows and four lagged rainfalls such that the model takes the simple form

$$y_t = -\delta_1 y_{t-1} - \delta_2 y_{t-2} + \omega_0 u_{t-b} + \omega_1 u_{t-b-1} + \omega_2 u_{t-b-2} + \omega_3 u_{t-b-3} ,$$

with two autoregressive parameters δ_1 and δ_2 and four moving average parameters ω_0 , ω_1 , ω_2 and ω_3 . In PRTF form this reduces from six to five parameters through the relation of the equal root parameter β (reparameterised as the time-to-peak, t_{peak}) to δ_1 and δ_2 . This is thought to have sufficient flexibility in TF response behaviour to provide a reasonable indication of the performance of the TF approach across a range of catchments and is supported by comments made by Han (1991). It was also supported by exploratory studies aimed at identifying the optimal model order for use across the study catchments.

The selection of the optimal TF model structure has received much attention and can include consideration of both model time-step and model order together with the search procedure involved (Isermann, 1980; Powell and Cluckie, 1985; Owens, 1987). Isermann (1980) comments that model order determination is not critical for single input - single output systems and in many cases there is not one definite “best” order. Jakeman and Young (1984) comment that the orders are usually smaller than 5 and often only 1 or 2. The extensive datasets available here for model calibration means that there is little penalty in adopting a higher order than is needed for some catchments. Thus, the pragmatic approach adopted here of using a TF(2,4,b) model throughout is considered justified in providing sufficient flexibility of response behaviour across catchments. Using a fixed time-step of 15 minutes, even for slowly responding catchments, has not led to difficulties that can arise with the autoregressive parameters approaching the unstable region, although care has been taken to preserve the precision of the estimated parameters. Subtracting the lowest flow on record as one form of baseflow separation was considered but rejected on account of the very small values involved, when compared to flood flows.

Jakeman and Hornberger (1993) advise that the sampling interval should be selected to be preferably less than the time constant of the quickest identifiable response but caution that this can make identification of the slower components numerically difficult unless suitable estimation algorithms are used. Suitable algorithms, such as the simplified refined instrumental variable (SRIV) algorithm, have not been used here because these do not feature

in EA operational practice. This highlights a deficiency in operational practice where recursive least squares is used to identify TF models for catchments whose response is dominated by a very slow decay of smaller volume (baseflow or “slow flow”) superimposed on a higher volume and faster decaying component (storm runoff or “quick flow”). The practice of using a fixed value of baseflow corresponding to the flow at the start of the event and varying the model time-step according to catchment response is not considered the most appropriate way of dealing with this problem. The PDM toolkit, which includes forms of TF model in parallel as model options, provides another way of addressing this problem. A parallel structure of flow processes is assumed from the outset together with nonlinear estimation (a simplex search procedure). This releases the restriction of a linear model structure demanded for SRIV estimation to be used, and allows the parameters of the nonlinear loss accounting component to be estimated as well.

Grid Model

The Grid Model was developed by the Institute of Hydrology for the Environment Agency to exploit the distributed nature of radar data and new digital datasets on elevation, land use and soils (Moore *et al.*, 1992; Bell and Moore, 1998). It is configured to share the same grid as that used by the weather radar. Each radar grid square area is conceptualised in the catchment as a storage which receives water in the form of precipitation and loses water via overflow, evaporation and drainage. The storage used in the basic form of model is a simple store (tank or bucket) having a finite capacity S_{\max} . This capacity can be thought of as an absorption capacity characterising the area of the square grid encompassing surface detention, soil moisture storage, and the interception capacity of vegetation and other forms of land use. A fundamental idea used in the basic form of model is that absorption capacity is controlled by the average gradient, \bar{g} , of the topography in the grid square which can be calculated readily from a digital terrain model (DTM).

Specifically, for a given grid square, the following linkage function is used to relate the maximum storage capacity, S_{\max} , and the average gradient, \bar{g} , within a grid square:

$$S_{\max} = \left(1 - \frac{\bar{g}}{g_{\max}} \right) c_{\max} ,$$

for $\bar{g} \leq g_{\max}$. The parameters g_{\max} and c_{\max} are upper limits of gradient and storage capacity respectively and act as “regional parameters” for the basin model. A measurement of the mean gradient within each grid square of the river basin is obtained from the DTM. Values of S_{\max} for all grid squares are determined using only the two model parameters, g_{\max} and c_{\max} , together with measurements of \bar{g} for each square.

A grid storage loses water in three possible ways. If the storage is fully saturated from previous rainfall then any net addition of water spills over and contributes to the fast catchment response. Drainage from the base of the store is controlled by the volume of water in store and contributes to the slow catchment response. Thirdly, water is lost via evaporation to the atmosphere. Figure 1.2.8 (a) illustrates a typical grid storage and the components of the water balance involved. Evaporation loss occurs at the rate, E_a , which is related to the potential evaporation rate, E , and the store water deficit, D , through a simple linear decrease

from the potential rate as a threshold deficit, D^* , is exceeded. The value of D^* is common across grid squares. Drainage from the grid storage, which contributes to the slow catchment response, occurs at a rate controlled by a power function of the water in store. Finally, the updated water storage is given by continuity taking into account the initial storage and losses to evaporation and drainage. The direct runoff rate contributing to the fast basin response is calculated as the rainfall less loss to any storage left available.

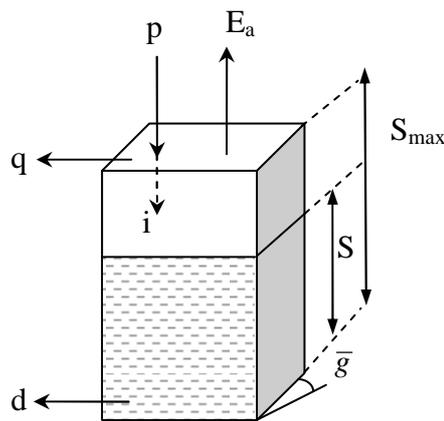
Water is routed from each grid square storage to the catchment outlet using DTM-derived isochrone pathways. The construction of isochrones – lines joining points of equal time of travel to the basin outlet – is achieved by assuming that water travels with only two velocities depending on whether the pathway involves a hillslope or a river channel. In this way it is relatively easy to construct isochrones by direct inference from the distance of a point to the basin outlet and the nature of the pathways involved. The catchment is subdivided into reaches according to these isochrones and water is routed along the reaches to the catchment outlet using a discrete kinematic wave routing procedure. This not only advects water between the reaches but also incorporates a diffusive component seen in observed hydrographs.

Figure 1.2.8 (b) shows an idealised catchment with isochrones overlaid onto the grid squares. Water storage accounting for any grid only partially inside the catchment is treated in the normal way and an adjustment made when accumulating the runoff/drainage across the catchment. The water input to each isochrone strip can be readily calculated as an area weighted summation of the outflow rates from the grid squares encompassed by the strip. The outflow can be the direct runoff rate, q_{tj} , or the drainage rate, d_{tj} , depending on whether the routing scheme relates to the fast or slow response pathway to the catchment outlet.

The n isochrone strips are represented by a cascade of n reaches, with each reach represented by a discrete kinematic wave equation with lateral inflow. The number of strips, n , together with a dimensionless wave speed parameter, θ , controls the lag and attenuation of water movement through the reaches (Moore and Jones, 1978). The lateral inflow, r_t^k , can be defined as direct runoff or drainage which are routed separately using two parallel discrete kinematic wave models, characterised by different wave speeds θ_s and θ_b respectively. A schematic depicting the overall structure of this basic form of Grid Model is shown in Figure 1.2.8 and Table 1.2.8 provides a summary of the model parameters.

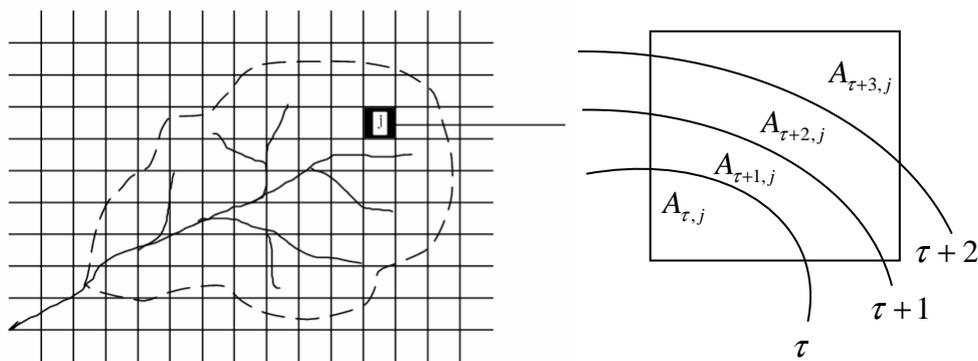
Note that only the basic form of Grid Model has been used in this study. Other variants involve using alternative mechanisms of runoff production which include: (i) the distribution of slope within a grid square to underpin a probability-distributed store formulation, (ii) a topographic index control, and (iii) an integrated air capacity control based on soil survey data. Also, land use classification of urban area can be used to delineate the fraction of each grid square that can be considered to have zero storage capacity. A variant of the translation component allows drainage from each grid square to travel to the basin outlet in a way governed by a separate set of isochrones, representing the slow response pathway, which is determined by the path length and a Darcy velocity of flow. This velocity is estimated from the local gradient of the terrain (calculated from the DTM) as an approximation to the hydraulic gradient.

(a) Water balance within a grid



- p rainfall
- E_a evaporation
- q direct runoff
- d drainage
- i infiltration
- \bar{g} average gradient
- S, S_{max} water storage and maximum storage capacity

(b) Catchment with superimposed weather radar grid and inset showing isochrone areas in grid square j



(c) Grid Model configuration

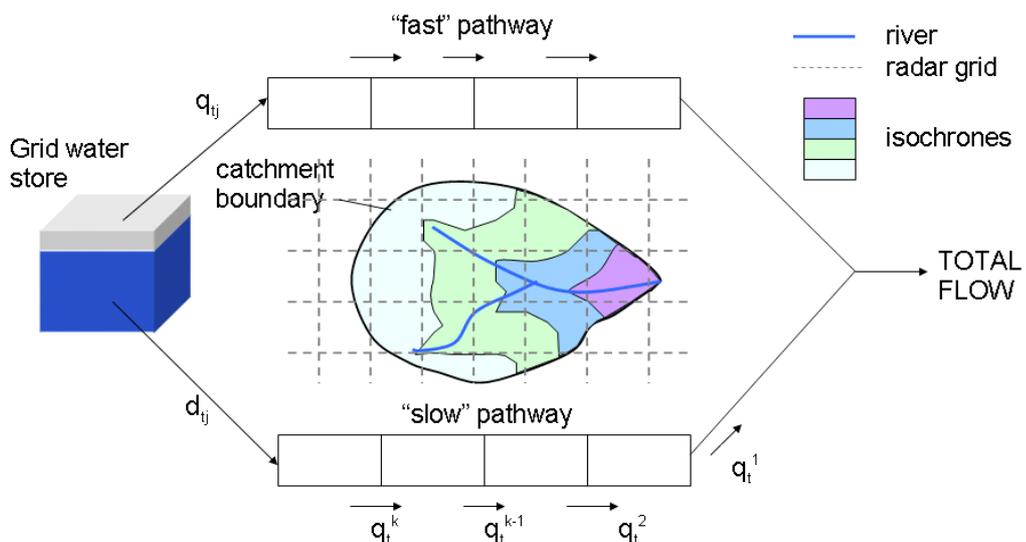


Figure 1.2.8 The Grid Model.

Table 1.2.8 Parameters of the Grid Model

Parameter	Description	Unit
f_r	Rainfall correction factor	-
D^*	Storage threshold deficit (or root constant) in evaporation function	mm
S_0	Proportion of total storage capacity initially full	-
g_{max}	Regional upper limit of gradient	-
c_{max}	Regional upper limit of storage capacity	mm
i_{max}	Maximum infiltration rate	mm h ⁻¹
k_d	Storage constant of (cubic) drainage function	h ⁻¹ mm ⁻²
θ_s	Wave speed parameter for routing direct runoff	-
θ_b	Wave speed parameter for routing drainage	-
v_L	Advection velocity of flow along land path	m s ⁻¹
v_R	Advection velocity of flow along river path	m s ⁻¹

Model updating methods

So far the set of rainfall-runoff models included here for assessment, with the exception of the TF and PRTF models, have been reviewed as simulation models transforming rainfall (and possibly evaporation) to runoff. Flows measured up to the time the forecast is made can be used for improving model performance using some form of *updating scheme*. In reviewing the TF and PRTF models we have seen that direct substitution of simulated flows by observed values in the forecast equation acts to re-initialise the model and is a form of *state correction*. Also, varying the model gain of the TF model with reference to recent observed flows provides a simple means of *parameter-adjustment*. The TF (and PRTF) model can also be used with a simple ARMA *error prediction* scheme, in which the dependence in present and past simulation-mode model errors are used to predict future ones. In this form the model is referred to as a TFN (Transfer Function Noise) model.

For the TCM, NWS and MCRM models a simple ARMA error prediction scheme has been used since this approach can be used with any model structure. The PDM model has its own state correction procedure which apportions model errors to adjust the water contents of different stores within the model. It can also be used with ARMA error prediction. Preliminary trials have been made in this study to determine which approach to use in different situations. The IEM is readily updated using state correction of the quadratic storage since its outflow is directly related to the observed flow. Whilst the Grid Model has its own form of state correction as one updating option, only the ARMA error prediction approach has been used here. A brief review of the ARMA error prediction scheme follows along with state correction schemes used by the PDM and IEM.

Error prediction

A feature of errors from a conceptual rainfall-runoff model is that there is a tendency for errors to persist so that sequences of positive errors (underestimation) or negative errors (overestimation) are common. This dependence structure in the error sequence may be exploited by developing error predictors which incorporate this structure and allow future errors to be predicted. Predictions of the error are added to the deterministic model prediction to obtain the updated model forecast of flows. The error prediction scheme is wholly external to the deterministic model operation and thus may be used in combination with any model. Error prediction is now a well established technique for forecast updating in real-time (Box and Jenkins, 1970; Moore, 1982). The basis of the technique is summarised below.

A real-time forecast of flow, $q_{t+P|t}$, made P time units ahead from a forecast origin at time t may be expressed as

$$q_{t+\ell|t} = q_{t+\ell} + \eta_{t+\ell|t}.$$

where q_{t+P} is the simulation-mode model forecast and $\eta_{t+P|t}$ denotes a prediction of the simulation-mode error, η_{t+P} , made P steps ahead from a forecast origin at time t . (The suffix notation $t+P|t$ should be read as $q_{t+P|t}$ being a forecast of the value at time $t+P$ given information up to time t .) Based on an ARMA model, the error prediction equation used is

$$\begin{aligned} \eta_{t+\ell|t} = & -\phi_1 \eta_{t+\ell-1|t} - \phi_2 \eta_{t+\ell-2|t} - \dots - \phi_p \eta_{t+\ell-p|t} + \theta_1 a_{t+\ell-1|t} + \theta_2 a_{t+\ell-2|t} \\ & + \dots + \theta_q a_{t+\ell-q|t}, \quad \ell = 1, 2, \dots \end{aligned}$$

where $\{\phi_i\}$ and $\{\theta_i\}$ are autoregressive and moving average parameters respectively,

$$a_{t+\ell-i|t} = \begin{cases} 0 & \ell - i > 0 \\ a_{t+\ell-i} & \text{otherwise} \end{cases}$$

and a_{t+P-i} is the one-step ahead prediction error

$$a_{t+\ell-i} = Q_{t+\ell-i} - q_{t+\ell-i|t},$$

where Q_{t+P-i} denotes the observed flow at time $t+P-i$ and the simulation-mode error is given by

$$\eta_{t+\ell-i|t} = \eta_{t+\ell-i} = Q_{t+\ell-i} - q_{t+\ell-i} \quad \text{for } \ell - i \leq 0.$$

The error prediction equation is used recursively to produce the error predictions $\eta_{t+1|t}$, $\eta_{t+2|t}$, ..., $\eta_{t+P|t}$, from the available values of a_t , a_{t-1} , ... and η_t , η_{t-1} , Using this error predictor methodology, the model simulation-mode forecast, q_{t+P} , is updated using the error prediction $\eta_{t+P|t}$ to calculate the required real-time forecast, $q_{t+P|t}$. Note that this real-time forecast incorporates information from the most recent observations of flow through the error predictor, and specifically through calculation of the one-step ahead forecast error, a_{t+P-i}

Alternative error predictor schemes may be devised by working with other definitions of the basic errors: for example by using proportional errors of errors from a logarithmic model. Only the standard additive form is used here. The Transfer Function Noise (TFN) Modelling Package has been used to identify the form of ARMA error predictor and to estimate its parameter values. Also a means exists within IH's TSCAL (Time Series CALibration) Model Calibration program to estimate the ARMA error predictor parameters for an assumed model structure. This has been used as a complement to the TFN package on occasions for parameter estimation and always to produce the final forecast results. Whilst error prediction provides a general technique which is easy to apply, its performance in providing improved forecasts will depend on the degree of persistence in the model errors. It is least successful at correcting the more variable errors that can occur in the vicinity of the rising limb and peak of the flood hydrograph.

State correction for the PDM and IEM

State correction techniques have been developed based on adjustment of the water content of conceptual storage elements in the belief that the main cause of the discrepancy between observed and modelled runoff will arise from errors in estimating basin average rainfall, which in turn accumulate as errors in water storage content. The main state variables in the PDM model are the water contents of the surface and groundwater stores, S_2 and S_3 , and of the probability-distributed soil storage, S_1 (using the notation of Figure 1.2.4). The flow rates out of the conceptual stores can also be regarded as state variables: examples are q_s , the flow out of the surface storage, and q_b , the flow out of the groundwater storage. When an error, $\varepsilon = Q - q = Q - (q_s + q_b)$, occurs between the model prediction, q , and the observed value of basin runoff, Q , it would seem sensible to "attribute the blame" to mis-specification of the state variables and attempt to "correct" the state values to achieve concordance between observed and model predicted flow.

A formal approach to state correction is provided by the Kalman filter algorithm (Jazwinski, 1970; Gelb, 1974; Moore and Weiss, 1980a,b). Whilst this provides an optimal correction for linear dynamic systems subject to random variations which may not necessarily be Gaussian in form, this is not the case for nonlinear dynamic models due to the linearisation approximations needed to be introduced. The implication of this is that simpler, intuitive adjustment schemes can be devised which potentially provide better adjustments than the more complex and formal extensions of the Kalman filter which accommodate nonlinear dynamics through approximations. Such schemes which make physically sensible adjustments will be referred to here as *empirical state adjustment schemes*. In the PDM a simple apportioning of the error, ε , between the surface and groundwater stores in proportion to their contribution to the total flow is used such that

$$q_b^* = q_b + \alpha g_b \varepsilon$$

$$q_s^* = q_s + (1 - \alpha) g_s \varepsilon$$

where

$$\alpha = q_b / (q_s + q_b)$$

and the superscript * indicates the value after adjustment. The *gain update* coefficients, g_b and g_s , when equal to unity yield the result that $q_b^* + q_s^*$ equals the observed flow, Q , thus achieving exact correction of the model flow to equal the observed value. Values of the coefficients other than unity allow for different adjustments to be made, and g_b and g_s can be regarded as model parameters whose values are established through optimisation to achieve the “best” fit between state-adjusted forecasts and observed flows. This scheme is referred to as the *proportional adjustment scheme*.

A *super-proportional adjustment scheme* is developed as a generalisation of the above with α defined as

$$\alpha = \frac{q_b}{\beta_1 q_s + \beta_2 q_b}$$

with the incidental parameters β_1 and β_2 chosen to weight the apportionment towards or away from one of the flow components; in practice β_1 and β_2 are assigned values of 10 and 0.1 to apportion more of the error adjustment to the surface store. Note that the adjustment is carried out at every time step and the time subscripts have been omitted for notational simplicity. Replacing α and $(1-\alpha)$ by unity yields the simplest *non-proportional adjustment scheme*. In this study only the super-proportional adjustment scheme has been included in the assessment.

The PDM also allows other state variables to be adjusted but these variants have not been assessed here. It should be noted that empirical state correction is of the same basic form as that employed by the Kalman filter in which an updated state estimate is formed from the sum of the current state value and the model error multiplied by a gain update coefficient. However, instead of defining the gain statistically, as the ratio of the uncertainty in the observation to that of the current state value, it is first related to a physical apportionment rule multiplied by a gain factor. This gain factor acts as a relaxation coefficient which is estimated through an off-line optimisation using past flood event data.

State correction of the IEM is particularly straightforward since only the quadratic storage is a candidate for correction. Adjustment of its outflow, q , is made using the standard form of adjustment with α equal to unity, such that $q^* = q + g\varepsilon$, where g is the gain update parameter and ε is the model error.

1.3 Selection of Study Catchments

The main aim of the Part 2 Report is to assess the performance of the eight rainfall-runoff models across a variety of catchments using datasets encompassing a wide range of flood events. In order to draw general guidelines on model choice in different situations, a careful selection of catchments for use in the assessment is required. The selection of catchments included in the model intercomparison is presented in Table 1.3.1 and their locations are mapped in Figure 1.3.1. At least one catchment has been chosen from each of the eight EA regions. The choice has been guided by operational importance as well as ensuring a mix of hydrological characteristics, such as small/medium/large, upland/lowland, ground/surface water dominant, and urban/rural. Very large catchments have been excluded as these would normally be represented by a network configuration of rainfall-runoff models and channel flow routing models, these being outside the scope of the present study.

Table 1.3.1 Catchments selected for model assessment

EA Region	Catchment	Area km²	Characteristics
1. North East	Trout Beck at Moorhouse	11.4	Small, upland rural.
2. Thames	Silk Stream at Colindeep Lane	29	Small urban, clay.
3. Midlands	Dove at Izaak Walton	83	Small/medium, high relief, rural, moorland mudstone/sandstone headwaters. Carboniferous limestone.
4. Southern	Lavant at Graylingwell	87.2	Small/medium, ephemeral stream in a permeable (Chalk) catchment, mainly rural, groundwater abstractions, well records.
5. Welsh	Rhondda at Trehaford	100.5	Medium, steep rural headwaters, urban/industrial development in valleys, 24% forest.
6. South West	Brue at Lovington	135.2	Medium, rural catchment, springfed headwaters. Clays, soils and oolites. Modest relief.
7. Midlands	Stour at Shipston	185.0	Medium, rural, Keuper Marl. Modest relief.
8. North West	Roch at Blackford Bridge	186	Medium, highly urbanised in lower half with moorland above.
9. Anglian	Witham at Claypole Mill	297.9	Large, mainly rural lowland catchment with clay (50%), limestone (40%) and gravel, low relief.

One catchment, the Stour to Shipston, was included during the course of the study in response to the “Easter 1998 Flood” and the need to have one catchment significantly affected by this major flood event. The conditions of super-saturation brought on by this event might bring into play different flood mechanisms that require to be modelled. The Lavant is included as representative of a catchment with a flood response superimposed on a dominant groundwater regime, and which has well level records and pumped abstractions. Only two of the models, the TCM and PDM, have the capability to accommodate pumped abstraction data and to represent ephemeral streamflow behaviour explicitly. As a consequence, only these models are applied to the Lavant catchment along with the TF model (using rainfall and flow data only), serving primarily as a point of reference. The Roch is included with the knowledge that

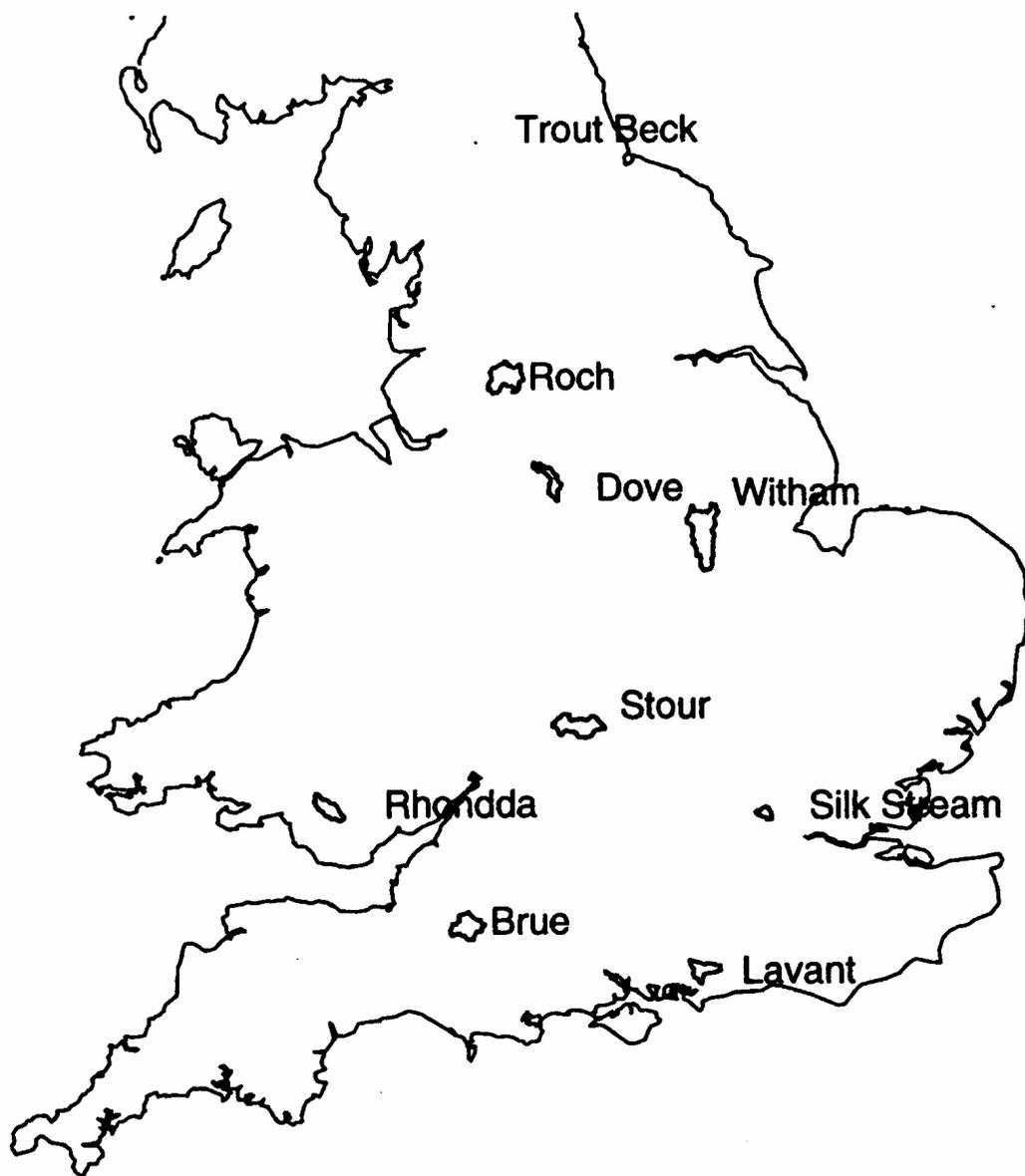


Figure 1.3.1 Location of the study catchments.

this has been extensively used in model development work in the EA Northwest, especially for TF and PRTF models. Use of the Brue reflects the availability of a very dense raingauge network installed by the EA under HYREX (Hydrological Radar EXperiment), a NERC Special Topic, along with supporting radar data from Wardon Hill in Dorset. Trout Beck, a typical small upland catchment, is the focus of a snowmelt forecasting project and an existing dataset at IH. The Dove at Izaak Walton was the first catchment in the Midlands to be chosen, because of its relevance to flood warning and that data had been collated in support of an extension to the snowmelt forecasting project. The Silk Stream is a small urban catchment in London with a good dataset of raingauge, flow and radar data brought together under the NRA Project “Evaluation of FRONTIERS and Local Radar Rainfall Forecasts for use in Flood Forecasting Models”. TCM, PDM and IEM models had already been developed for this catchment. The Rhondda to Trehafod provides a good example of a medium-sized catchment with steep slopes. This catchment has been the subject of study in the NRA Project “Development of Distributed Flood Forecasting Models using Weather Radar and Digital Terrain Data”, from which the Grid Model was developed and compared in performance to the PDM. The Witham was chosen by hydrologists in Anglian region as representative of a large, low-lying rural catchment in Eastern England and because of its importance to flooding of Lincoln. Further details of the catchments and their associated data are presented in Section 2.6.

Only three of the catchments – the Silk Stream, the Stour and the Roch – are assessed here using radar data whilst all catchments are assessed using raingauge data. This reflects the cost and time-consuming nature of handling radar data off-line at the present time. Data from the Clee Hill radar were used for the Stour, the Hameldon Hill radar for the Roch and the Chenies (London) radar for the Silk Stream. Single site (Type 2) 2 km radar data (or 5 km data outside the 75 km range) have been used. Nimrod corrected data (Golding, 1998) are not available over the period of a decade used for model assessment. Also, the Nimrod product has only been stable since circa 1998 and is available only for the coarser resolution of 5 km. Details of the periods of data used in the model assessment for both raingauge and radar data are presented in Section 2.8. The Grid Model has only been applied to the three largest catchments – the Stour, the Roch and the Witham – where the benefit of using a distributed model formulation is likely to be most beneficial.

2. STRATEGY FOR MODEL ASSESSMENT

2.1 Introduction

The basis of the model assessment is to compare eight rainfall-runoff models using datasets from nine catchments of varied character distributed throughout England and Wales. Sections 2.2 and 2.3 set out a split-sample strategy for model assessment based on *model calibration* using part of the datasets and *model evaluation* on datasets not used for calibration. The calibration phase focuses on the models used in simulation-mode, without correction using observed flows, although the parameters involved in updating are also estimated. Performance statistics and visual aids used for model assessment in both calibration and evaluation phases are introduced along with the specialised terminology used.

Section 2.4 presents some of the tools developed to support model and data configuration. This includes use of the IH Digital Terrain Model (DTM) and digital spatial datasets to support calculation of catchment average rainfall and to define response zones in the Thames Catchment Model. Also a new method to adjust weather radar data with reference to raingauge data is presented.

Section 2.5 outlines how the models have been prepared for assessment, using the RFFS Model Calibration Facility as a unified environment for model calibration and evaluation. Details relating to reconfiguring models to operate within this facility are given, including the setting of initial conditions and accommodating groundwater and abstraction data where needed.

Section 2.6 outlines the catchments used in the model assessment along with the supporting hydrometric records. Section 2.7 outlines issues relating to data management including data quality control and the creation of a Model Database. The selection of periods of data to be used for model calibration and model evaluation as part of a split-sample model assessment procedure is the subject of Section 2.8.

2.2 Strategy for Model Calibration

A rainfall-runoff model may be used to compute flows in **simulation mode**, based solely on rainfall and potential evaporation inputs, or in **updating mode**, when additional use is made of observed flow data. The updating procedures used here are ARMA (Autoregressive Moving Average) error-prediction and state correction. These procedures have been briefly reviewed in Section 1.2 whilst further details are given in Section 12 of the Part 1 Report. Model calibration involves estimation of the parameters of both the rainfall-runoff model and of the updating procedure before a forecast can be made. How these two tasks are accomplished is discussed below.

2.2.1 Calibration of rainfall-runoff model parameters

Calibration of a model is achieved by comparison of observed and modelled flows, both in the long term (comparing daily flows) and for sets of flood events (comparing 15-minute flows). It may involve a mixture of manual adjustment, requiring an understanding of the physical basis of

the models and visual inspection of the model hydrograph, and automatic optimization aimed at minimising an **objective function** which provides a measure of the difference between observed and computed flows. The objective function used in this study is the root mean square of the differences between observed and computed flows at each time step, with provision to exclude a warm-up period for each event and flows outside a given range. When calibrating the parameters of a rainfall-runoff model as a flow simulator, the computed flow used to form the objective function is the simulation-mode flow.

For a given model type, catchment, and source of rainfall data, the procedure is to adjust parameter values interactively and/or automatically, operating the model in simulation-mode to arrive at an optimum set. The automatic optimisation employs a robust and straightforward simplex (polytope) minimisation procedure (Nelder and Mead, 1965) modified to incorporate ideas suggested by Gill, Murray and Wright (1981). The slow components of a model may be roughly calibrated using daily data over a long period and then refined, along with the parameters of the fast components, using 15-minute data over periods encompassing many flood events.

The calibration procedure adopted for the TF and PRTF differs from that used for other models and further details are provided in Section 2.5.2.

2.2.2 Calibration of updating parameters

State-correction schemes are available for the PDM, IEM, TCM and Grid models. To calibrate the updating model parameters, the individual error at each time-step is taken as the difference between the observed flow and a computed flow based on updating up to the previous time-step (the **one-step-ahead error**).

Error-prediction parameters have been fitted using a Transfer Function Noise modelling package which estimates ARMA model parameters using an Approximate Maximum Likelihood Algorithm (Young, 1974). The Approximate Maximum Likelihood Algorithm estimates the p autoregressive and q moving average parameters of an ARMA(p,q) model of the model errors resulting from a simulation-mode model run. These ARMA model parameters can also be estimated or adjusted using automatic optimisation as required.

2.3 Strategy for Model Evaluation

2.3.1 Introduction

Model evaluation has been undertaken in both simulation-mode, which tests the rainfall-runoff model component alone, and in updating-mode, in which flow forecasts are computed making use of recent flow measurements. This section will explain the different forecast modes and types, clarify the terminology used and outline the different statistics used to compare model performance.

2.3.2 Forecast Construction and Terminology

A rainfall-runoff model may be used to compute flows in **simulation mode**, employing only rainfall and potential evaporation as inputs, or in **updating mode**, when additional use is made of observed flow data. A flow value computed from a given time, using information up to this time, for some future time is a **forecast value** with a **forecast origin** at the given time and a **lead time** equal to the difference between the times of the forecast origin and the forecast value. The time-series of forecast values computed from a single forecast origin is termed a **fixed-origin forecast**.

In this historical emulation of a real-time forecast, a fixed-origin forecast can be generated at every time-step using information on flows up to the forecast time-origin. To focus on the relative performance of different models it is assumed that the future rainfall inputs are known. Otherwise, uncertainty in the rainfall forecast would serve to confuse the assessment of different model types. A synthetic example is shown in Figure 2.3.1, in which the thick line shows an observed hydrograph and the thin lines show fixed-origin forecasts with origins at a selection of times and lead times up to τ hours ahead. The dotted line connects all the forecast values with lead times of τ hours and is called a **fixed lead-time forecast**. Such forecasts are very useful in characterising the average performance of a model for a given lead time, as discussed in the next section.

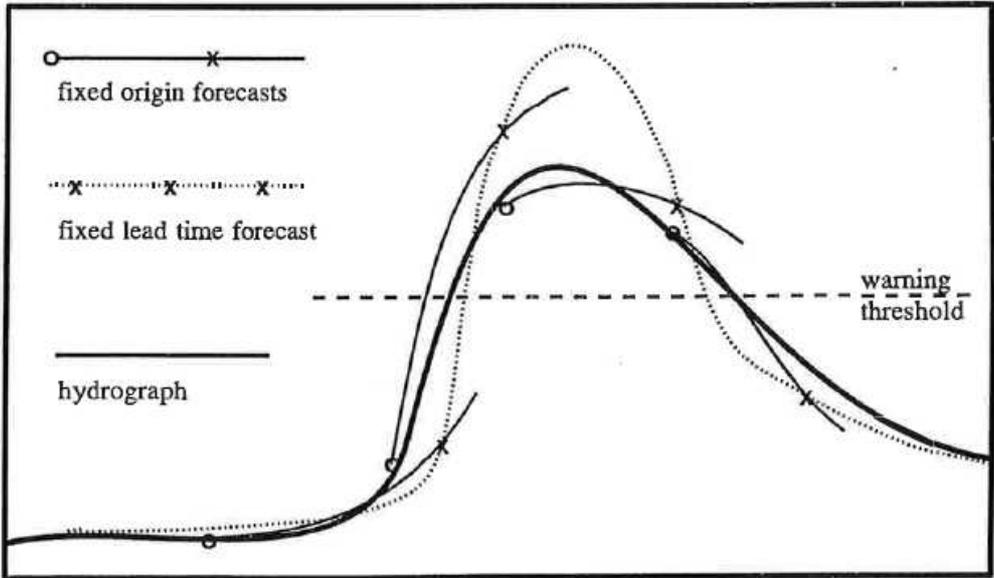


Figure 2.3.1 Definition sketch for types of flow forecast.

2.3.3 Measures of Forecast Performance

Various measures of the performance of a time-series of forecast flows, $\{q_t\}$, for times $t=1, 2, 3, \dots$, relative to observed flows, $\{Q_t\}$, are possible. These can be divided into quadratic measures, such as the root-mean-square error or *rmse*, and other categorical skill score measures, such as the Critical Success Index or CSI. Measures based on external factors, such as cost of flooding, are not considered here.

Quadratic measures

One of the simplest quadratic measures of performance is the root-mean-square error

$$rmse = \left(n^{-1} \sum e_t^2 \right)^{1/2} \quad (2.3.1)$$

where the error at time t is $e_t = Q_t - q_t$ and the summation is computed over n values. A related performance measure is the R^2 statistic:

$$R^2 = 1 - \frac{\sum e_t^2}{\sum (Q_t - \bar{Q})^2} \quad (2.3.2)$$

where \bar{Q} is the mean of the observed flows over the n values. This gives the proportion of variability in observed flows accounted for by the model forecast. Note that this can be negative if the forecasts are worse than that provided by the unknown mean flow.

For fixed lead-time forecasts these performance measures would be computed over the full period assessed, whilst for fixed-origin forecasts they would be computed over a period t_0 to $t_0 + \tau$, where t_0 is the forecast origin and τ is the lead time. However, they are most commonly applied to fixed lead-time forecasts because these give fewer measures per period assessed (one for each possible lead time rather than one for each possible origin) and relate to a quantity of interest (lead time) which has a common meaning across different assessment periods and catchments, so allowing model comparison. For this study the R^2 measure is preferred, since it is a relative measure of forecast accuracy which permits comparison across assessment periods and catchments of different size.

Categorical measures

The evaluation of forecast performance on the basis of categorical measures applied to fixed-origin forecasts is an attractive proposition, since it is closely related to the operational procedure whereby different categories of flood warning alert would be chosen based on the most recently available forecasts and measurements. However, care is required to design a measure that is simple, meaningful, does not rely on arbitrary decisions, and can be applied to any catchment. The procedure described below is an attempt to construct such a measure. The procedure is based around the notion of a threshold flow, which may or may not be exceeded during a given assessment period. The thresholds actually in use by the EA are not used here, firstly because they are essentially arbitrary, and secondly because there are very few exceedences of the higher

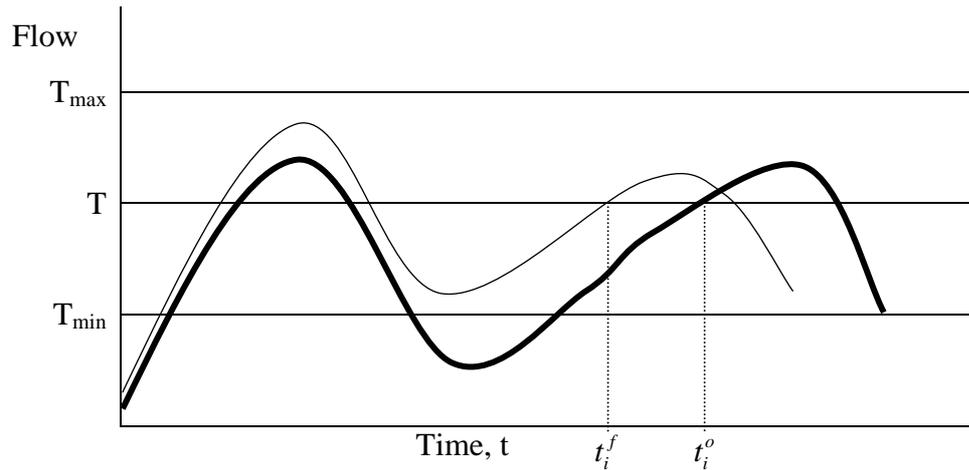


Figure 2.3.2 Threshold crossing criteria.

thresholds. Instead, the following procedure is repeated over a wide range of possible thresholds to give an overall skill score related to a particular lead time, τ .

Points are identified when a given threshold flow, T , is exceeded on rising limbs for both the observed and forecast hydrographs. If forecast and observed exceedences occur within a given time, Δt , of each other, they are matched up and denoted as t_i^f and t_i^o , with i identifying the occurrence. This is repeated for a range of thresholds from T_{\min} to T_{\max} in steps of magnitude ΔT . Skill scores are then calculated based on the ability of the model to forecast crossing of the range of thresholds within time Δt of the observed series also crossing the threshold. The threshold skill scores are defined as follows:

$$\text{Critical Success Index: } \quad CSI = n_{11} / (n_{11} + n_{12} + n_{21}) \quad (2.3.3)$$

$$\text{Probability of Detection: } \quad POD = n_{11} / (n_{11} + n_{12}) \quad (2.3.4)$$

$$\text{Correct Alarm Rate: } \quad CAR = n_{11} / (n_{11} + n_{21}) \quad (2.3.5)$$

where:

- n_{11} = number of times thresholds are crossed in both observed and forecast hydrographs within a time Δt of each other
- n_{12} = number of times thresholds are crossed in observed series, but not in forecast series within a time Δt of each other
- n_{21} = number of times thresholds are crossed in forecast series, but not in observed series within a time Δt of each other
- n_{22} = number of times thresholds are not crossed in both observed and forecast hydrographs within a time Δt of each other, and
- n = $n_{11} + n_{12} + n_{21} + n_{22}$, as shown in Table 2.3.1.

Table 2.3.1 Two-way contingency table for categorical evaluation

	Threshold exceeded during event	Threshold not exceeded during event
Threshold exceeded during forecast	n ₁₁	n ₂₁
Threshold not exceeded during forecast	n ₁₂	n ₂₂

These threshold skill scores may be plotted against lead-time to show the rate at which skill declines with increasing lead time. In each case, the skill index ranges from 0 to 1, with an index of 1 indicating a perfect forecast.

2.3.4 The forecasting process

The procedure for generating flow forecasts adopted here is straightforward and results in a set of flow forecasts for each of the evaluation periods in a particular catchment. Two sets have been produced.

- (i) For all catchments, except the slow responding groundwater catchment (the Lavant), 24 fixed lead-time forecasts have been produced for lead times of 0.25, 0.5, 0.75, 1.0, ...,5.75 and 6 hours. These are used to provide the forecast performance measures. For the groundwater-dominated catchment, lead times of 0.25, 0.5, 0.75, 1.0, ...,5.75 and 6 *days* have been used.
- (ii) Fixed-origin forecasts, six hours long, with between 4 and 6 time origins before a major flow peak. These forecasts are used to provide illustrative hydrographs.

2.4 Model and Data Configuration Tools

2.4.1 Methods of raingauge weighting for calculating catchment average rainfall

Rainfall data for catchments which have more than one raingauge located nearby need to be pooled to derive catchment average rainfall. Various methods are available for estimating areal rainfall from multiple raingauges, ranging from simple weighted averages to more complex surface fitting methods. In the present study, quality control of the rainfall data reveals that for eight of the nine catchments, no more than two raingauges per catchment can be relied upon to provide good quality data. As a result, simple weighted average schemes have been used in preference to more complex surface fitting methods.

Areal average rainfall is calculated as the linear weighted sum of n raingauges, such that

$$\bar{P} = \sum_{i=1}^n w_i P_i , \tag{2.4.1}$$

where p_i is the rainfall measured at the i 'th gauge and w_i is a raingauge weighting. The IH DTM (Digital Terrain Model) has been used to support derivation of these weights and the four schemes outlined below have been considered.

(i) Thiessen weights

Each raingauge is weighted according to the proportion of the catchment for which it is the closest gauge. The Thiessen weight (Thiessen, 1911) is therefore a distance weighting only, based on a “nearest neighbour” principle.

(ii) SAAR-adjusted Thiessen weights

For a given catchment where Thiessen weights have been calculated, the weight of each raingauge is adjusted according to the SAAR (Standard Average Annual Rainfall) in the region of the gauge. Specifically, an adjusted weight is defined as

$$w_T^{SAAR} = \frac{\bar{P}}{P} w_T \quad (2.4.2)$$

where w_T is the Thiessen weight, \bar{P} is the mean SAAR over the Thiessen area for that gauge, and P is the SAAR of the gauge. (The gauge suffix is omitted for notational simplicity). The effect of this adjustment is to reduce the weighting of a gauge that is in a region of higher rainfall relative to rainfall in the Thiessen area of that gauge.

(iii) Elevation-dependent weights

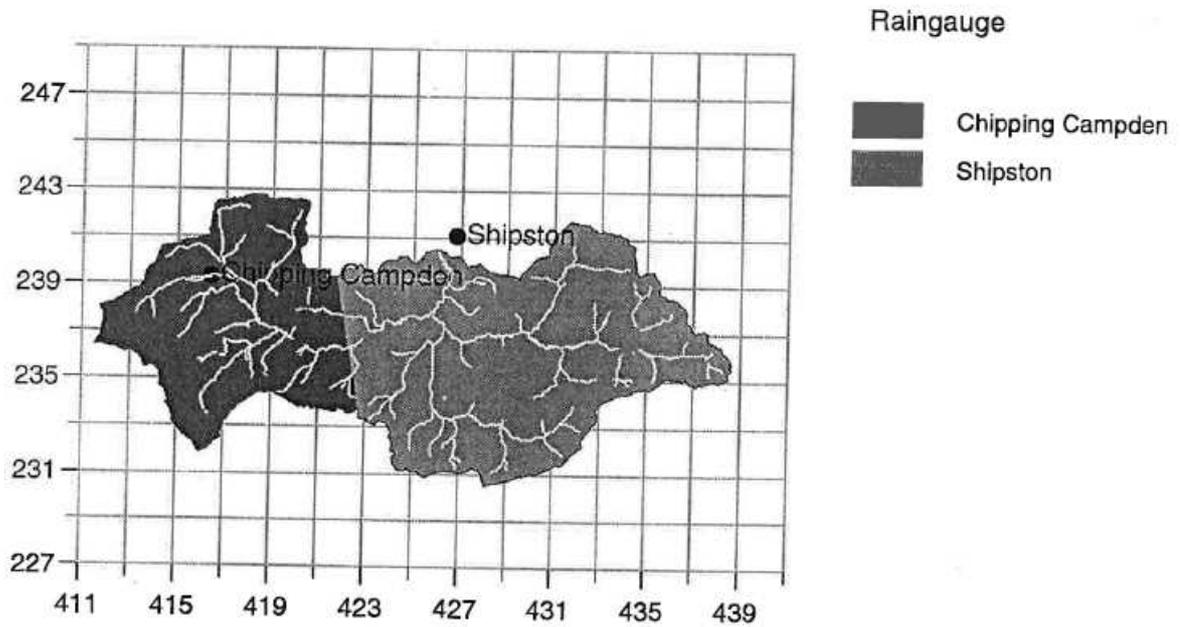
For regions where orographic effects are important and precipitation is dependent on elevation, a new scheme combining the Thiessen distance weighting with a DTM-derived elevation weighting has been developed. Thiessen distance weighting is used first to select the nearest gauges to the catchment. Gauges that would not be included in a Thiessen scheme are similarly excluded from the elevation-dependent scheme. The raingauge weights are then recalculated according to the proportion of the catchment that is closest in *elevation* to each gauge, as illustrated in Figure 2.4.1. This scheme has been of particular use in the Stour catchment, for which two raingauges are available: Chipping Campden and Shipston at elevations of 123 and 65 m respectively. Figure 2.4.1 compares the Thiessen raingauge areas with the elevation-derived area weights for these two gauges. The Thiessen method gives greater weight to the Shipston raingauge because much of the catchment lies closer to that gauge. In contrast the elevation method gives greater weight to the Chipping Campden gauge, which may be more representative of rainfall in the higher relief areas in the catchment.

(iv) SAAR-dependent weights

This fourth method is similar to that used to derive elevation-dependent weights, but uses SAAR instead of elevation. More specifically, after the Thiessen distance scheme has been used to select the nearest gauges to the catchment, raingauge weights are recalculated according to the proportion of the catchment that is closest in *SAAR* to each gauge.

Table 2.4.1 compares the raingauge weights derived using each of the four schemes above, together with the weight used in this study and the comparable weights used by the EA (if known). The EA weights for the Stour, which were derived empirically by hydrologists in the

(a) Thiessen scheme



(b) Elevation-dependent scheme

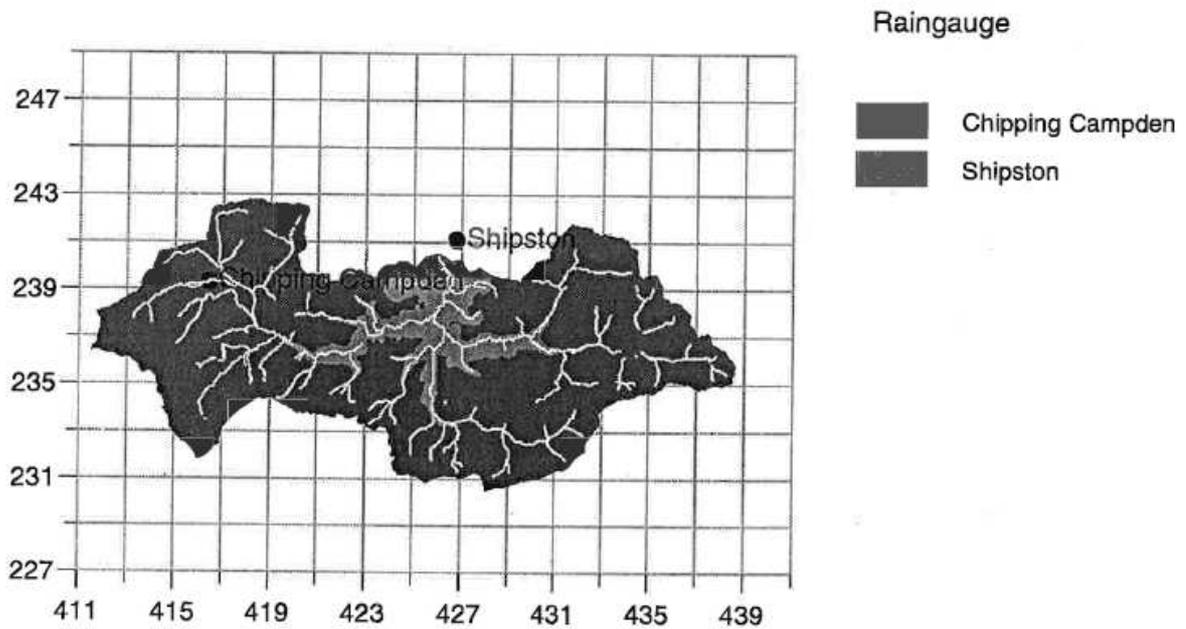


Figure 2.4.1 Comparison of raingauge weights derived using the Thiessen and elevation schemes.

Table 2.4.1 Raingauge weights derived using different methods

Catchment	Raingauge(s)	Thiessen	Thiessen/SAAR	Elevation	SAAR	Weights used	EA Weights
Trout Beck	AWS 31	1.	1.	1.	1.	1.	1.
Silk Stream	Mill Hill	0.962	0.961	0.584	0.997	0.600	-
	Radlett	0.038	0.040	0.416	0.003	0.400	
Dove	Hollinsclough	0.608	0.597	0.474	0.459	0.500	0.666
	Longcliffe	0.392	0.402	0.526	0.541	0.500	0.333
Lavant	Chichester	1.	1.	1.	1.	1.	1.
Rhondda	Tyn-y-waun	1.	1.	1.	1.	1.	1.
Stour	Chipping Campden	0.377	0.375	0.878	0.903	0.878	0.666
	Shipston	0.623	0.711	0.122	0.097	0.122	0.333
Roch	Spring Mill	0.645	0.575	0.569	0.179	0.569	-
	Kitcliffe	0.355	0.335	0.431	0.821	0.431	
Witham	Saltersford	0.761	0.733	0.460	0.552	0.460	-
	South Witham	0.239	0.241	0.540	0.448	0.540	

region, are closest to the elevation-dependent weights. The final choice of weight used in this study depended on the circumstances. In most catchments the elevation weight has been used in preference to the Thiessen distance weight. For catchments such as the Dove where there is little difference in the weights, the raingauge weights have been pooled to give a 50-50 weighting between the two gauges. In the case of the Silk Stream, the weights used are in line with those used in the Frontiers/Local Radar Rainfall Forecasting study (Moore *et al.*, 1993) and are similar to those derived using the elevation weighting scheme. The Brue catchment is excluded from the table because the catchment contains a dense network of 49 raingauges as part of the HYREX (Hydrological Radar Experiment) project. Data from the raingauge network is quality controlled and averaged to form high quality, 15 minute rainfall accumulations for the Brue catchment.

2.4.2 Raingauge adjustment of weather radar data

An improved estimator for the rainfall field can be obtained by combining information from a radar with data from a raingauge network. Moore *et al.* (1989, 1991, 1994a,b) present a method which fits a multiquadric surface to “calibration factors” formed by taking an adjusted ratio of the gauge to the coincident pixel radar value at each raingauge site. This calibration factor surface is applied to the radar values to obtain “calibrated radar” estimates. (The term “calibrated” is used here to mean adjusted with reference to raingauge data, and not in relation to the electronic calibration of the radar.) This approach has been adopted to obtain calibrated radar fields for the Silk Stream catchment. For this catchment data from an extensive raingauge network exists and, as part of the Frontiers and Local Rainfall Forecasting Study, time-series of catchment average rainfall have already been obtained by this method. However, for the remaining catchments, typically data from only two raingauges are available in the Model Database. This has led to a new calibration scheme being developed which is more appropriate when only data from a few raingauges are available. The scheme first adjusts for any bias in the mean of the radar rainfall field. The multiquadric surface fitting method is then applied to the calibration factors obtained using the mean field bias corrected radar values instead of the raw values. A different form of multiquadric surface is used so that the calibration factor surface tends towards a fixed value of 1 with increasing distance from the raingauges. This reflects the requirement for the calibration factor field to tend to 1 since the radar field has been adjusted to remove the mean field bias so the ratio of radar to gauge value should be unity, on average. The approach may be conceptualised as one in which the dynamic gauge adjustment dominates at short distances whilst at longer distances the rainfall estimate tapers towards the climatological mean field value of rainfall. Detail of the theory of the new method and its application to the study catchments is presented in Appendix A. The calibrated grid square radar estimates of rainfall are aggregated to form a catchment average estimate for each 15 minute time interval, using the proportion of each square within the catchment as a weighting.

2.4.3 Derivation of TCM response zones using digital datasets

The structure of the Thames Catchment Model, or TCM, is based on subdivision of a basin into different response zones representing, for example, runoff from aquifer, clay, riparian and paved areas. Delineation of these zones has been achieved using the IH DTM for the UK in

conjunction with a number of digital spatial datasets. Table 2.4.2 summarises the spatial datasets used in defining differing response zones within a given catchment.

Table 2.4.2 Spatial datasets used to delineate the TCM response zones

Dataset	TCM response zone
IH urban areas	Urban
IH 100 year flood risk map	Riparian
WRAP (Winter Rain Acceptance Potential)	Clay, aquifer, baseflow, ...

Each of the spatial datasets has been combined with the IH DTM catchment delineation procedures to calculate the proportion of each study catchment that is urban and riparian and as a guide to other response zones such as clay, aquifer and slow response (“baseflow”). The WRAP class dataset divides the UK into five soil types at a 1 km resolution, according to the Winter Rain Acceptance Potential of the region (NERC, 1975). The five soil classes are summarised in Table 2.4.3. The IH urban areas dataset has been produced by combining the ITE (Institute of Terrestrial Ecology) Land Cover Map of Great Britain with settlement polygons from the Ordnance Survey. These data include land classified as urban and suburban; although available at a 25 m resolution, they have been used here at a 50 m resolution to correspond to the resolution of the IH DTM. Riparian areas have been delineated using the 100-year flood risk map of England and Wales. These data have also been used at a 50 m resolution.

Table 2.4.3 WRAP (Winter Rain Acceptance Potential) data classes

WRAP class	Description
1	Well drained permeable, sandy or loamy soils
2	Permeable soils
3	Mixed permeable/impermeable soils
4	Clayey, loam over clay soils or impervious layer at shallow depths
5	Wet upland soils, peaty, shallow permeable, or rocky soils on steep slopes

Table 2.4.4 summarises the properties of the nine study catchments characterised by the proportion of differing land types. By way of example, the TCM response zones for the Roch catchment are shown in Figure 2.4.2. A complete set of catchment maps is provided in Appendix B. In calculating the proportions of each land type, a dominance hierarchy is assumed with urban areas wholly incorporated, then riparian excluding urban, and the remainder apportioned using the WRAP class proportions as a guide. In the Thames Catchment Model these proportions are multiplied by the area of the catchment to give an initial value for the size of each response zone. Final values for the area of each zone may differ from these initial values, as zone size is a parameter in the TCM and can be adjusted to

give optimum model performance. In addition, the zone size can act as a multiplicative rainfall factor, adjusting for the representativeness of the raingauges used. Therefore the total area of the zones may differ from the size of the catchment after calibration and also their relative proportions.

Table 2.4.4 The proportion of each land type in each catchment

Catchment	Urban	Riparian	WRAP Class				
			1	2	3	4	5
Trout Beck	0.	0.002	0.	0.	0.	0.	0.998
Silk Stream	0.543	0.003	0.	0.	0.	0.454	0.
Dove	0.0023	0.018	0.678	0.	0.	0.016	0.285
Lavant	0.011	0.013	0.844	0.	0.	0.132	0.
Rhondda	0.126	0.015	0.	0.	0.416	0.	0.443
Brue	0.012	0.032	0.023	0.001	0.656	0.276	0.
Stour	0.029	0.019	0.141	0.	0.308	0.503	0.
Roch	0.173	0.021	0.078	0.	0.	0.307	0.421
Witham	0.033	0.060	0.354	0.033	0.	0.520	0.

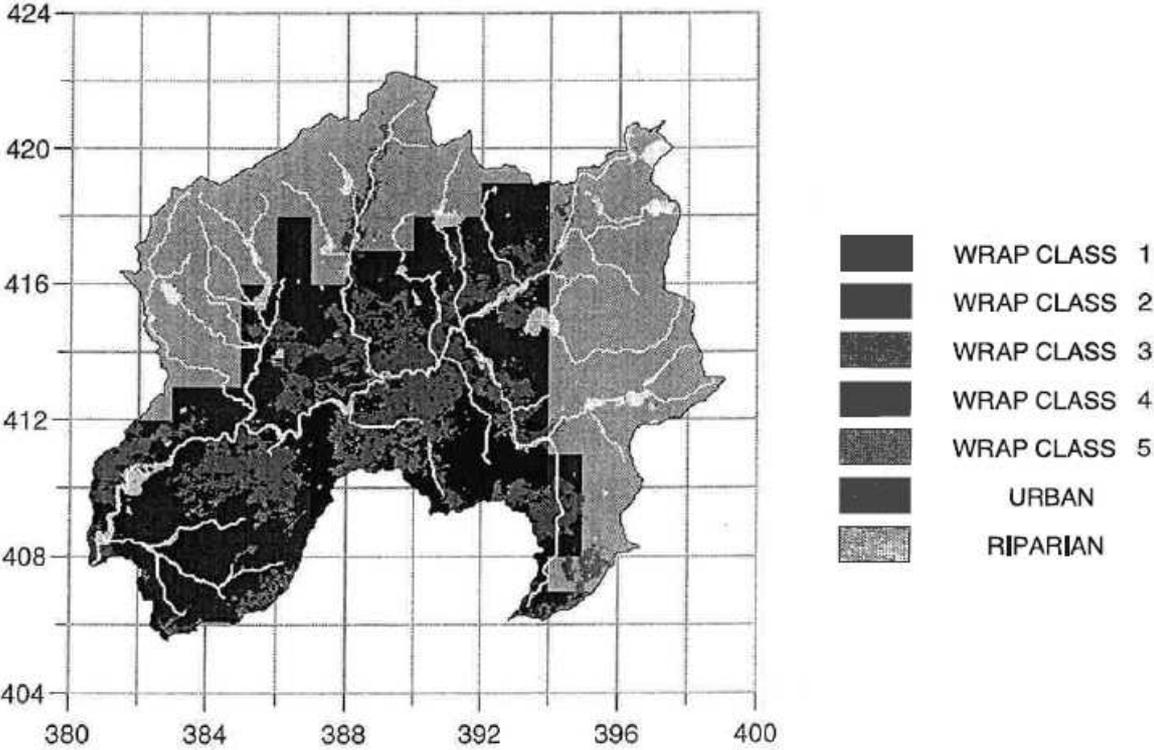


Figure 2.4.2 DTM-derived TCM response zones for the Roch.

2.5 Preparation of Models and Data for Assessment

2.5.1 Introduction

In order to assess the eight rainfall-runoff models in an impartial way, the IH River Flood Forecasting System (RFFS) Model Calibration Facility has been used as the single platform for model intercomparison. This provides a consistent environment for model calibration and evaluation across the models chosen for intercomparison. The eight models selected for assessment are:

TCM	Thames Catchment Model
MCRM	Midlands Catchment Runoff Model
PDM	Probability Distributed Moisture model
IEM	Isolated Event Model
NWS	National Weather Service model
TF	Transfer Function model
PRTF	Physically Realisable Transfer Function model
Grid Model	Simple distributed Grid-based Model.

At the start of the project, four of these models were already configured for use within the RFFS Model Calibration Facility: the TCM, PDM, IEM and the Grid Model. In addition, the PDM model included a TF(2,2) model as a constituent part. The remaining models - the NWS, MCRM and TF/PRTF - required reconfiguration in order to fit into the RFFS Model Calibration Facility. The NWS already existed in a form suitable for use within an evaluation framework, but had only previously been used at a daily time-step, while the MCRM and TF/PRTF required substantial coding in order to fit within the RFFS framework. The following sections outline the main changes that were made to these models, and discusses model initialisation and data requirements.

2.5.2 Model configuration and initial conditions

National Weather Service Model (NWS)

The main modifications made to the NWS model were related to the change from the original daily time-step to the 15 minute time-step used throughout the model intercomparison. In its original formulation the NWS was run on a daily time-step and flow was routed via an externally derived unit hydrograph coupled with a layered routing component described as “Muskingum” routing. While the use of a prescribed unit hydrograph is feasible at a daily time step (typically only 3 to 5 hydrograph ordinates may be required), at a 15 minute time interval this routing method becomes onerous as hundreds of hydrograph ordinates will be needed. Instead the routing component, which can be described more accurately as layers of linear reservoirs rather than “Muskingum” routing, has been used to translate flow across the catchment. A single rather than multi-layered linear reservoir was thought to be sufficient, and for a 15 minute time-step, this was found to be the case. When a daily time-step was used for soil moisture accounting, the unit hydrograph option could be invoked as originally coded.

Early trials indicated that for some catchments the performance of the model could be improved through the use of a time delay in the model’s response to rainfall, and this has

therefore been added to the model formulation and invoked as appropriate. Use of a time delay is consistent with the other rainfall-runoff models assessed here.

Midlands Catchment Runoff Model (MCRM)

Reconfiguring the MCRM for use within the RFFS Model Calibration Facility was found to be relatively straightforward, and few changes were made to the model structure. The model time-step was generalised from one hour to any time-step to allow operation at the 15 minute and daily time-steps used here. However, finding a robust algorithm for initialisation was a challenge as the model is extremely sensitive to the initial model states. When the MCRM is used operationally, the soil moisture is adjusted each week with reference to MORECS data. This was thought to be impractical in a model intercomparison, and instead the model has been initialised using a “warm-up” period where necessary; this is discussed in more detail in Section 2.4.3. Use of a different, and external, form of soil moisture model (MORECS) to adjust the MCRM’s soil moisture model is also seen as inconsistent, albeit useful in practice.

Grid Model

The Grid Model is designed to exploit the distributed nature of radar data and has been configured so as to share the same grid as that used by the weather radar data. The Grid Model has been calibrated on the three largest catchments, which were thought most likely to exhibit a distributed catchment response to rainfall. Unlike other models, which were calibrated using a mixture of manual adjustment and automatic optimisation of model parameters, the Grid Model was calibrated by manual adjustment alone. For a large catchment and a model resolution of 2 km, the Grid Model is extremely slow to calibrate using automatic methods and manual adjustment was preferred.

Transfer Function Models (TF/PRTF)

The Transfer Function (TF) model is a class of time-series model with which an output variable can be forecast as a linear weighted combination of past outputs and inputs. Specifically, for a TF(r,s,b) model, the output at time t, y_t , is related to r previous values of output and s previous values of input with delay b, u_{t-b} , such that

$$y_t = -\delta_1 y_{t-1} - \delta_2 y_{t-2} - \dots - \delta_r y_{t-r} + \omega_0 u_{t-b} + \omega_1 u_{t-b-1} + \dots + \omega_{s-1} u_{t-b-s+1} \quad (2.5.1)$$

where $\{\delta_i\}$ are the r autoregressive parameters and $\{\omega_i\}$ are s moving average parameters. The number of autoregressive and moving average parameters, r and s determine the model order, which must be as small as possible, but large enough to provide a physically realistic impulse response function. A TF(2,4,b) model order has been adopted here to provide sufficient flexibility in the impulse response function to accommodate most river flow situations, and has been used here for both the TF and the PRTF forms of transfer function model. Formal identification of model order for each catchment was trialled but the results were not clear-cut. It was not critical to obtain a minimum order model for each catchment on account of the extensive dataset available for calibration. This might not be the case if single flood events were used in isolation for model calibration, as is the usual practice in the EA where TF/PRTF models are used.

Calibration of a conceptual rainfall-runoff model, such as the TCM or the NWS, results in an optimal parameter set which can be used in both simulation and forecast mode. In simulation mode, rainfall data alone are used to calculate flow, while in forecast mode, ARMA error prediction or state-updating is used to correct the model output with reference to observed flow. In order to achieve comparable simulation and forecast results from the TF/PRTF models, the models have been calibrated twice. Once in simulation mode, where the calculated flow is based purely on past rainfall input, and once in updating mode, for which the forecast is determined with reference to past observed flow and rainfall values.

For the simulation-mode calibration, an Instrumental Variable (IV) algorithm (Young, 1974) has been used to estimate the parameters of a TF(2,4,b) model for each catchment. Experience suggests that the Instrumental Variable algorithm is quicker to run, and more likely to arrive at a physically appropriate set of parameters, than the simplex-based automatic parameter optimisation used with the conceptually-based rainfall-runoff models. However, the simplex scheme was used to obtain an initial estimate for the pure time delay, b, for use with the IV algorithm. The resulting set of parameters was generally found to provide a robust TF model for smaller responsive catchments, though for larger catchments with a significant groundwater component, some further calibration was required to achieve an acceptable model simulation. The special form of the PRTF meant that its parameters could not be estimated using the IV algorithm. Instead, the TF(2,4,b) parameter set was used to provide an initial set of parameter values from which to start a simplex optimisation.

In forecast mode, the parameters of the state-corrected form of the TF model (in which observed flows are used to forecast future flows) were estimated using a recursive least squares algorithm. For the PRTF, only the moving average parameters can be determined in this way, given a t_{peak} value, which determines the autoregressive parameters. This is achieved by embedding recursive least squares estimation of the moving average parameters within a simplex optimisation of t_{peak} . Whilst not identical in detail, this approach is consistent with that used in the MATH code used in practice by the EA for PRTF model calibration. Possible autocorrelation in the forecast errors is ignored with these model formulations.

Both error-prediction and gain-updating (Owens, 1986) can be used with the TF/PRTF models. The time-varying model gain updating scheme outlined in the Part 1 Report has been used in preference to ARMA error-prediction because this scheme is used in practice by the EA. However, trials suggest that a classical Transfer Function Noise model (a TF model with ARMA error-prediction) can give as good or better results than gain-updating. Owens (1986) recommends a number of updating constraints that should be applied to gain-updating to ensure that large, erratic variations in model gain do not occur. Most of these constraints have been applied in the TF/PRTF model formulation and appear to result in a robust updating scheme.

2.5.3 Model initialisation

In operational flood-forecasting systems, rainfall-runoff models are generally initialised using the model states stored from a previous model run. One way of emulating this in an off-line model intercomparison study is to automatically set initial model states to reflect saturated conditions, such as would be expected during winter periods, and to allow the models a period of “warm-up” if the chosen period begins during dry conditions. During model calibration

some experimentation with initial conditions was carried out to determine which models were sensitive to initialisation, and to ensure that model calibration was not unduly affected by the setting of initial states. Once calibration was complete, the scheme used for model evaluation was as follows. If the evaluation period started between 1 April and 1 November, the model was allowed a warm-up period beginning 1 April. If the period started after 1 November but before 1 April, conditions were assumed to be wet, and the automatic model initialisation was assumed to be adequate. Initial trials suggested that three models – the TCM, IEM and MCRM – should be initialised in this way, while the NWS, PDM and Grid Model were less sensitive to initial conditions and were initialised assuming predominantly wet conditions. The TF and PRTF are readily initialised using a few past observed flow and rainfall values, dictated by the model order.

Determination of a robust initialisation scheme for all models is a challenging problem, and provides scope for further research. The scheme outlined here is thought to be an acceptable compromise for a difficult problem, and is thought to be as fair as possible to all models. The long (approximately 6-8 month) periods used for model calibration and evaluation will also ensure that any poor model performance arising from mis-specified initial conditions will have a negligible effect on the overall model performance.

2.5.4 Incorporation of groundwater abstraction and well level data

The River Lavant has been included in the model intercomparison as an example of a catchment where a regime of natural recharge and artificial abstraction exerts a primary control on flood generation. Abstractions are made at two sites whilst well level measurements are made at three locations. Extensions to the models are required to accommodate such data.

Two models have been reconfigured for use with daily abstraction data: the TCM, which already included the option of a constant abstraction rate, and the PDM. Within these models, daily abstraction data from the two sites are combined, converted to a 15 minute time-step and divided by the catchment area to give a value for the depth of water extracted (in mm per 15 minute time-step). The well level data have not yet been fully incorporated into the rainfall-runoff models. There is scope for further work on the use of these data in calibration, state-updating and for assessing model forecasts of groundwater level.

Reconfiguration of the TCM and PDM for use with groundwater abstraction data has resulted in the introduction of four model parameters:

- K_u Underflow time constant (hours)
- D_{max} Maximum depth of groundwater available for underflow
- f_s Proportion of the baseflow lost to the catchment as external springs, (set here to 0.)
- f_a Abstraction factor (usually set to 1).

Only two of the new parameters have been used in practice, K_u , the underflow time constant, and D_{max} , the groundwater depth above which water leaves the catchment as underflow. The spring factor f_s which determines the proportion of the baseflow lost to springs outside the catchment has not been invoked. The abstraction factor, f_a , applied to the abstraction data to account for unmeasured abstractions has been set to unity.

2.6 Study Catchments and Data for Model Assessment

2.6.1 Introduction

This section provides a brief description of the catchments used in the study gleaned from the IH/BGS publication “Hydrological data UK: Hydrometric Register and Statistics 1991-95” and documents held in the IH Surface Water Archive. Table 2.6.1(a) provides an overall summary of the flow characteristics. Table 2.6.2(b) gives the current alarm sensor settings used for operational flood warning. These serve as a point of reference to indicate the relative severity of the events used in this study.

For each catchment the following information is provided: (i) a brief description of the catchment and its gauging station, (ii) a photograph of flooding, either within the catchment or downstream, serving to highlight the relevance to flood warning, (iii) a table summarising the hydrometric stations and data availability, and (iv) a map of the catchment and its hydrometric stations. In reading the catchment descriptions that follow, reference should be made to these Tables and Figures to obtain an overall picture of the study catchments and their associated hydrometric data.

Table 2.6.1 Flow characteristics and Alarm Sensor levels for the study catchments

(a) Flow characteristics

River	Gauging Station	Area Km ²	Mean flow m ³ s ⁻¹	Peak flow m ³ s ⁻¹	Baseflow index	Loss mm	Runoff coefficient
Trout Beck	Moor House	11.4	0.55	45.5	0.15	402	0.79
Silk Stream	Colindeep Lane	29.0	0.25	28.7e	0.28	417	0.39
Dove	Izaak Walton	83.0	1.95	20.7	0.79	380	0.66
Lavant	Graylingwell	87.2	0.28	8.1	0.84	841	0.11
Rhondda	Trehafod	100.5	5.56	206.4	0.42	501	0.78
Brue	Lovington	135.2	1.85	95.5	0.47	457	0.49
Stour	Shipston	185.0	1.47	91.4	0.45	425	0.63
Roch	Blackford Bridge	186.0	4.94	282.9	0.50	413	0.67
Witham	Claypole Mill	297.9	1.75	37.5	0.67	432	0.30

(b) Alarm sensor levels

River	Gauging Station	Alarm level				
Trout Beck	Moor House	0.80 m 9.98 m ³ s ⁻¹ Standby	0.90 m 13.02 m ³ s ⁻¹ Area Alert	No flood warnings for this station		1.548 m 41.5 m ³ s ⁻¹ Bankfull
Silk Stream	Colindeep Lane	1.00 m Standby	1.35m 6.2 m ³ s ⁻¹ Low	1.55 m 9.2 m ³ s ⁻¹ Medium	1.80 m 13.1 m ³ s ⁻¹ High	14.3 m ³ s ⁻¹ Bankfull
Dove	Izaak Walton		10 m ³ s ⁻¹ Yellow	20 m ³ s ⁻¹ Amber		1.607 m 27.98 m ³ s ⁻¹ Highest
Lavant	Graylingwell		0.46 m 1.995 m ³ s ⁻¹ Yellow	0.60 m 3.406 m ³ s ⁻¹ Amber	0.70 m 4.555 m ³ s ⁻¹ Red	
Rhondda	Trehafod		1.60 m 58.7 m ³ s ⁻¹ Yellow	2.90 m 113.3 m ³ s ⁻¹ Amber	3.40 m 134.2 m ³ s ⁻¹ Red	
Brue	Lovington		1.9 m 24.9 m ³ s ⁻¹ Yellow	2.4 m 36.0 m ³ s ⁻¹ Amber	3.7 m 71.33 m ³ s ⁻¹ Red	1.9 m 24.9 m ³ s ⁻¹ Bankfull
Stour	Shipston	1.5 m 9.38 m ³ s ⁻¹ Standby	2.1 m 15.5 m ³ s ⁻¹ Yellow	2.5 m 20.0 m ³ s ⁻¹ Amber	3.6 m 57.2 m ³ s ⁻¹ Red	3.94 m 91.4 m ³ s ⁻¹ Highest
Roch	Blackford Bridge	0.75 m Standby				27.8 m ³ s ⁻¹ Bankfull
Witham	Claypole Mill	0.69 m 6.1 m ³ s ⁻¹ Control room open	0.77 m 20 m ³ s ⁻¹ Operational	0.85 m 25 m ³ s ⁻¹ Operational (2)	0.986 m 35 m ³ s ⁻¹ Red (Trigger 2)	1.086 m 43 m ³ s ⁻¹ Red (Trigger 3)

2.6.2 Trout Beck at Moor House

The Trout Beck catchment is the smallest of the study catchments with a drainage area of 11.4 km² to the gauging station at Moor House. This relatively remote upland catchment reaches an altitude of 847 m and supports peaty moorland developed mainly on Carboniferous Limestone. The flow regime is responsive and natural. A compound Crump weir is used for flow measurement and was designed to gauge flows in excess of 70 m³ s⁻¹, although the highest recorded flow is 45.5 m³ s⁻¹. Upstream shoaling and low winter temperatures can affect the precision of flow measurements.



Plate 2.6.2 Flooding at Croft on the Tees, March 1968

Table 2.6.2 Data availability for Trout Beck at Moor House

Data Type	Station Name	NGR Coords.	Station Number	Data Available
River Level	Trout Beck at Moor House	NY759336	25003	18/6/91-1/6/98
Flow	Trout Beck at Moor House	NY759336	25003	18/6/91-1/6/98
Rain	Cow Green Reservoir	NY817291	26644	5/6/91-1/6/98
Rain (hourly)	Moor House AWS	NY758328	AWS31	28/5/91-1/1/99

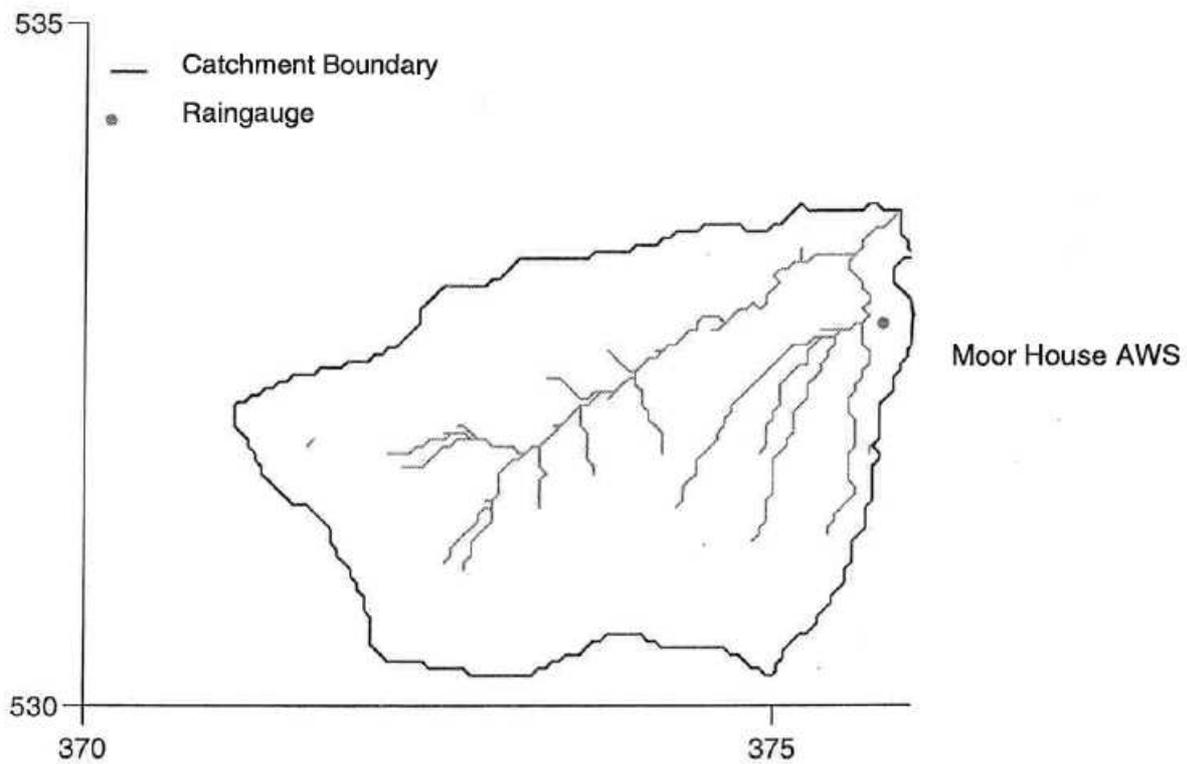


Figure 2.6.2 Catchment map for Trout Beck at Moor House.

2.6.3 Silk Stream at Colindeep Lane

The Silk Stream at Colindeep Lane extends over a 29 km² area of north-west London which is heavily urbanised (Stanmore and Edgware) except for its rural/suburban headwaters with woods on the highest ground. This together with London Clay being the exclusive lithology makes it a very responsive catchment. Altitude ranges from 40m at the gauging station to 153 m. With an average annual rainfall and runoff of 688 and 271 mm, giving a loss of 417 mm and a runoff coefficient of only .39, it is clearly significantly affected by artificial influences. Whilst artificial influences are evident at low flows, the net effect of abstractions and returns are uncertain. The flat-V weir used for gauging appears good except for some bypassing at flood flows. A bridge downstream may cause the weir to drown out at high flows. The mean annual flood is 15.2 m³ s⁻¹.



Plate 2.6.3 Flooding of the Silk Stream in Edgware

Table 2.6.3 Data availability for the Silk Stream at Colindeep Lane

Data Type	Station Name	NGR Coords.	Station Number	Data available
River Level	Colindeep Lane	TQ217895	39049	16/9/85-1/8/93
Flow	Colindeep Lane	TQ217895	39049	16/9/85-1/8/93
Rain	Mill Hill	TQ241920	246627	15/9/85-1/8/93
	Radlett	TL148002	277406	1/1/87-1/7/94
PE (15 min)	Lower Lee	-	-	1/9/85-1/9/93

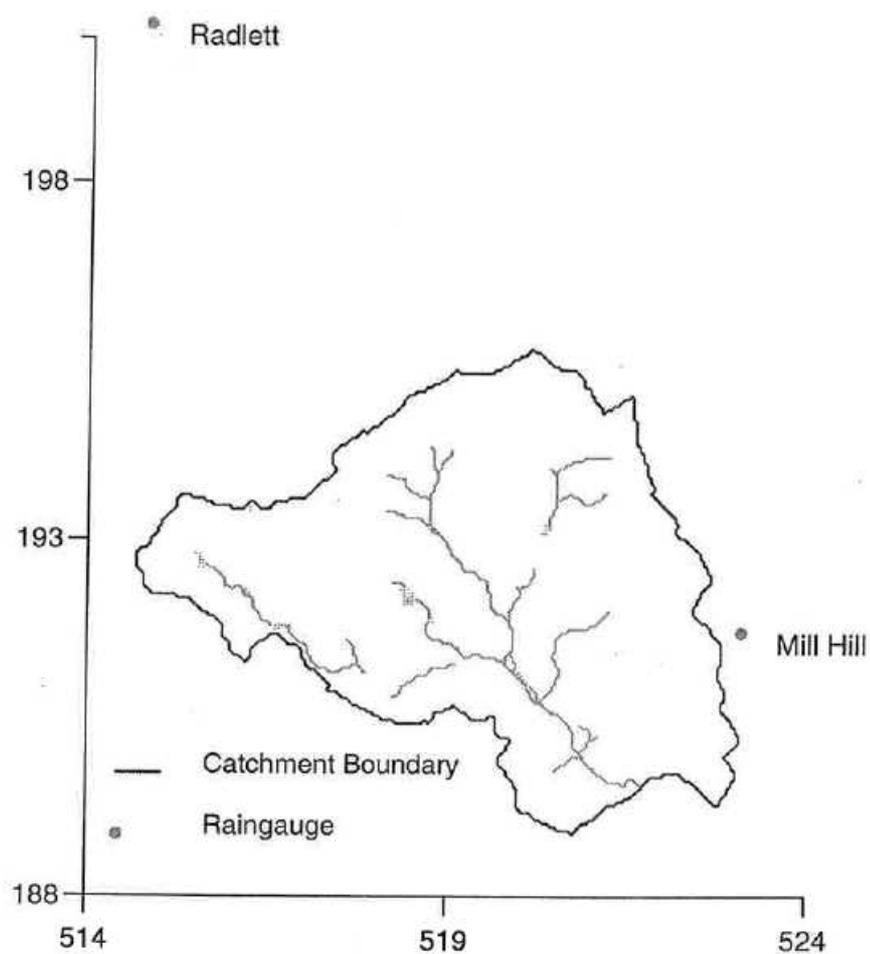


Figure 2.6.3 Catchment map for the Silk Stream at Colindeep Lane.

2.6.4 Dove at Izaak Walton

The Dove catchment in the Midlands region is a long narrow rural catchment draining an area of 83 km² to its gauging station at Izaak Walton. The upper catchment is moorland with Millstone Grit mudstone and sandstone; a steep Carboniferous Limestone ridge forms the western boundary. The river passes through deep limestone gorges, such as Wolfscote Dale and Dove Dale. The gauging structure is a Crump profile Flat-V weir which is modular up to bankfull. Some bypassing on the left bank may occur at high flows. Table A.2 presents the yellow and amber “forecast level thresholds” used for the Dove based on the Izaak Walton station. Forecast flows of 80 and 90 m³ s⁻¹ on the Manifold at Ilam (an adjacent branch of the Dove) are also used to issue yellow and amber alerts whilst the red alert is triggered by forecast levels exceeding 2 m at Rocester Weir downstream.



Plate 2.6.4 Flooding in the Dove catchment at Hanging Bridge, 20 December 1991

Table 2.6.4 Data availability for the Dove at Izaak Walton

Data Type	Station Name	NGR Coords.	Station Number	Data Available
River Level	Dove at Izaak Walton	SK146509	28046	1/1/90-17/12/97
Flow	Dove at Izaak Walton	SK146509	28046	1/1/90- 20/7/98
Rain	Ashbourne St. Oswalds	SK173465	3322	1/1/98-14/7/98
Rain	Carsington Dam	SK242503	3580	1/1/98-14/7/98
Rain	Cauldon Low	SK058480	3570	1/1/90-14/7/98
Rain	Chapel Reservoir	SK069795	3572	1/1/90-1/1/98
Rain	Hollinsclough	SK066665	3307	1/1/90-15/7/98
Rain	Longcliffe	SK228553	3576	1/1/90-14/7/98
Rain	Stanley Reservoir	SJ929519	3573	1/1/90-1/1//98
Rain	Tideswell	SK155746	3578	1/1/90-1/1/98

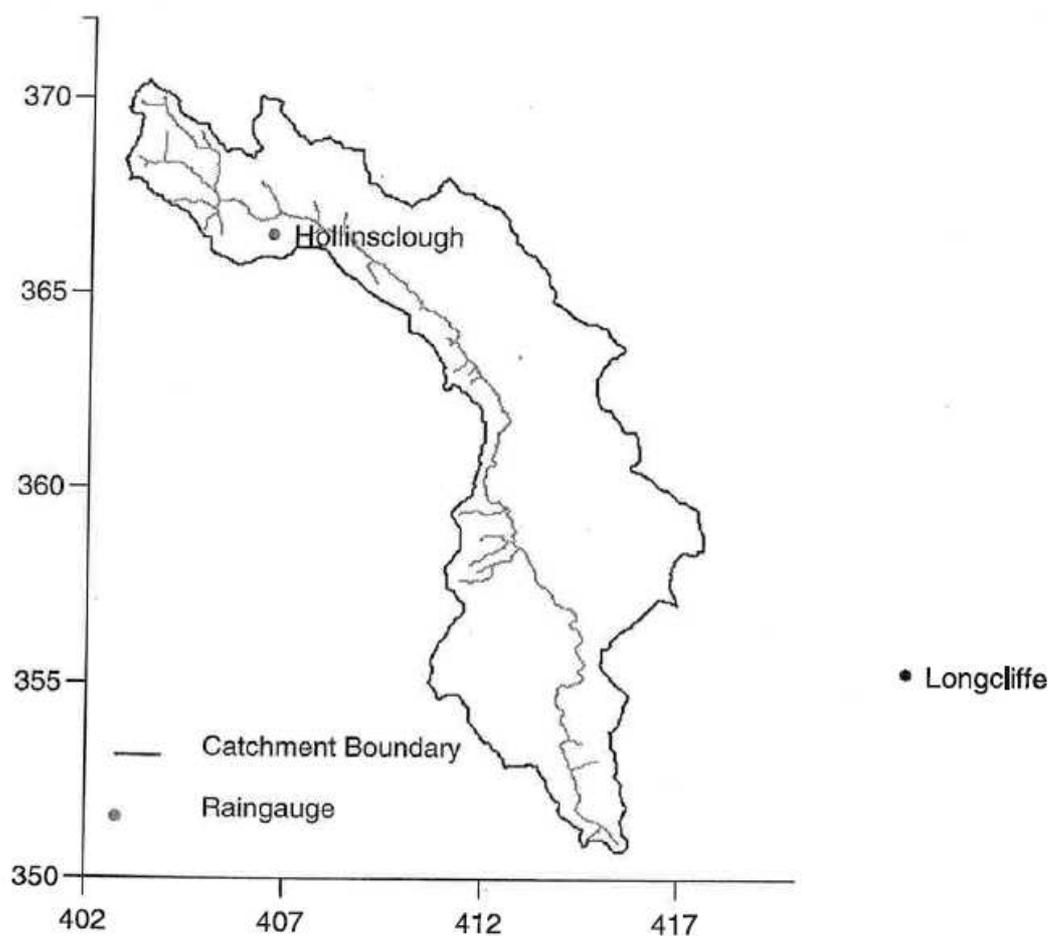


Figure 2.6.4 Catchment map for the Dove at Izaak Walton.

2.6.5 Lavant to Graylingwell

The groundwater-dominated catchment of the River Lavant in southern England drains an area of 87 km² to its gauging station at Graylingwell. It is an ephemeral stream on the dip-slope of the South Downs. This rural Chalk catchment is highly permeable. Land use is agricultural with significant woodland and only a little urban development close to Graylingwell. Significant groundwater abstractions from wells at Brick Kiln and Lavant reduce river flows. The gauging structure is a flat-V weir with a weir capacity of 6 m³ s⁻¹. Bypassing occurs during extreme events, such as the January 1994 flood peak estimated at 8.1 m³ s⁻¹.



Plate 2.6.5 Flooding of the River Lavant near Chichester 1994

Table 2.6.5 Data availability for the Lavant at Graylingwell

Data Type	Station Name	NGR Coords.	Station Number	Data Available
River Level	Lavant at Graylingwell	SU871064	41023	1/1/90-31/3/98
Flow	Lavant at Graylingwell	SU871064	41023	1/1/90-31/3/98
Rain	Chichester	SU879052	265313501	8/10/90-28/4/98
Daily Rain	Chilgrove	SU360144	320994	1/1/90-30/6/98
Daily Rain	S. Mundham	SU880003	320401	1/1/90-30/6/98
Daily Rain	Walderton	SU788105	321551	1/1/90-31/5/98
Well Level	Chilgrove	SU836144	81001	21/6/93-10/7/98
Well Level	Portfield Depot	SU878052	245313030	1/1/90-8/8/97
Well Level	West Dean Colworth	SU852151	245221016	7/1/82-29/9/97
Well Level	West Dean Nursery	SU863131	245222008	21/1/76-1/8/97
Daily abstraction	Brick Kiln	SU835124	1051	1990-1998
Daily abstraction	Lavant	SU856097	2051	1990-1998

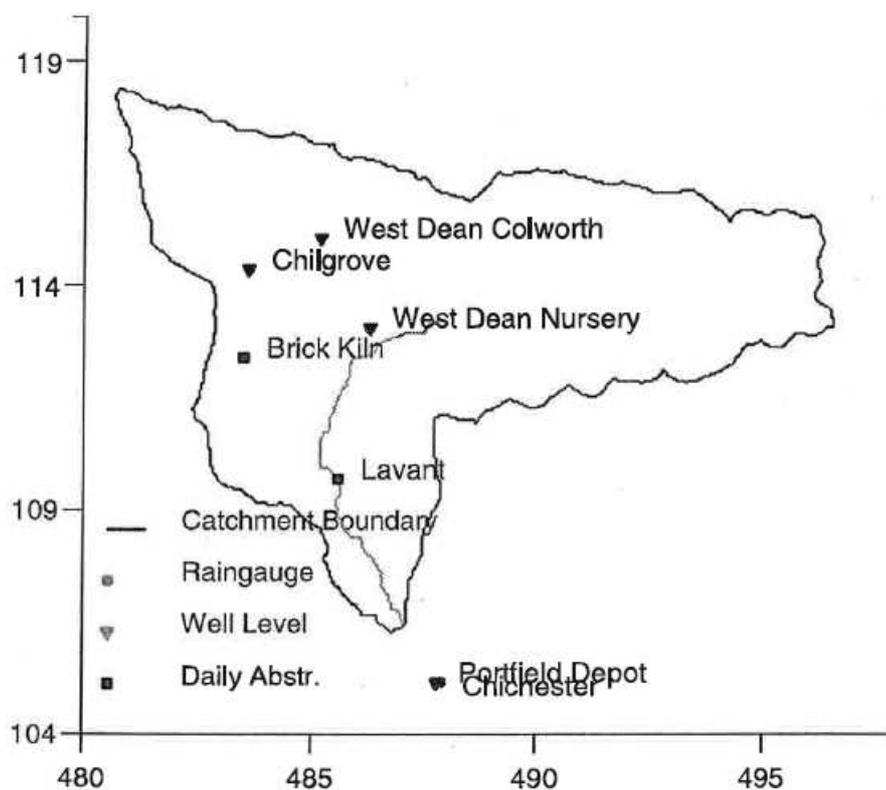


Figure 2.6.5 Catchment map for the Lavant at Graylingwell.

2.6.6 Rhondda at Trehafod

The River Rhondda in south Wales drains an area of 100 km² to its gauging station at Trehafod. The hilly upland areas are rural with livestock farming whilst there is industrial and urban development in the valleys. Geology is Coal Measures with valley alluvium. About 24% of the catchment is forested. There is an impounding reservoir in the upper catchment and flows are also affected by mine-water discharges. The gauging station is of velocity-area type in a formalised trapezoidal channel able to measure over the full flow range.



Plate 2.6.6 Flooding of the River Rhondda at Trehafod, 1979

Table 2.6.6 Data availability for the Rhondda at Trehafod

Data Type	Station Name	NGR Coords	Station Number	Data available
River Level	Rhondda at Trehafod	ST054909	57006	6/10/88-1/7/93
Flow	Rhondda at Trehafod	ST054909	57006	6/10/88-1/7/93
Rain	Tyn-y-Waun	SS933992	490291	21/9/89-1/7/93

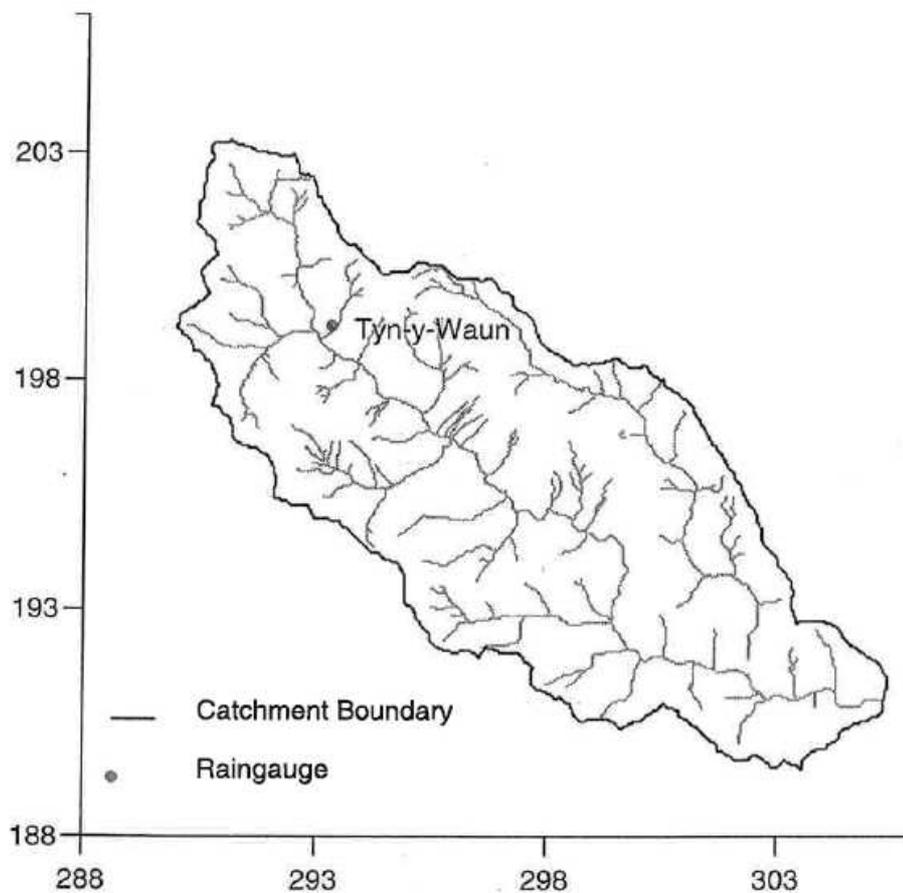


Figure 2.6.6 Catchment map for the Rhondda at Trehafod.

2.6.7 Brue at Lovington

The River Brue in South-west England drains a predominantly rural catchment area of 135 km² to its gauging station at Lovington. Headwater streams are fed by springs from the Mendips and Salisbury Plain. The responsive catchment is largely impermeable with Oxford Clay and Great Oolite in the upper part and Yeovil Sands and Inferior Oolite in the lower part. Low flows are gauged with a Crump weir whilst for higher flows (above 2.2 m³ s⁻¹) the station is used as a velocity-area type with a reliable rating up to bankfull (1.9 m); summer weed growth affects the stability of the rating. Table 2.6.1(b) indicates the flood warning alarm levels for this station. Above a level of 2.16 m, the Lower Brue at North Drain and Westhay overtops and the B3151 road floods. At 2.6 m (41 m³ s⁻¹) the road at Cow Bridge starts to flood. Above 3.1 m, (54 m³ s⁻¹), minor roads at West Lidford flood.

The Brue catchment was used as the focus of HYREX (Hydrological Radar EXperiment), a NERC Special Topic which ran from May 1993 to April 1997 (Moore, 1999). The Environment Agency funded the installation of a dense raingauge network over the catchment with most 2 km radar squares coincident with the catchment having at least one raingauge. Data management, quality control and archiving was the responsibility of the Institute of Hydrology with MAFF funding. Data from this network has been used to provide an unusually accurate estimate of catchment average rainfall.



Plate 2.6.7 Flooding of the Brue at Bruton, July 1982

Table 2.6.7 Data availability for the Brue at Lovington

Data Type	Station Name	NGR Coords	Station Number	Data available
River Level	Brue at Lovington	ST590318	52010	1/1/92-1/7/98
Flow	Brue at Lovington	ST590318	52010	1/1/92-1/7/98
Rain	Brue catchment average	-	-	19/9/93-1/10/98
Rain (daily)	Brue catchment average	-	-	19/9/93-1/10/98
PE (daily)	Bridge Farm AWS	ST638351	IHIS/029	2/9/93-26/8/96

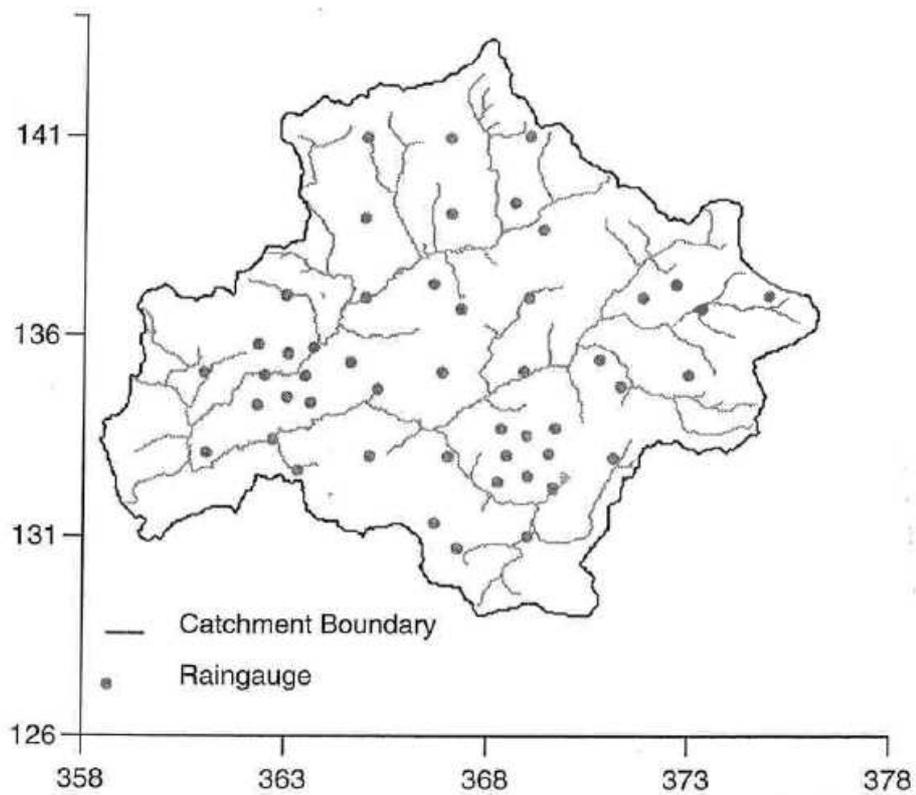


Figure 2.6.7 Catchment map of the Brue at Lovington.

2.6.8 Stour at Shipston

The Stour in the Midlands region is a major tributary of the Worcester Avon draining an area of 349 km² to the Avon confluence and 185 km² to the gauging station at Shipston used as the study catchment. The primarily northern flowing river has a relatively narrow floodplain and steep gradient. Soils originate predominantly from Keuper Marl. The village of Shipston lies 15 km south-east of Stratford-on-Avon on the left bank of the Stour. It has experienced a number of floods in recent years, notably in 1947 and 1968 in addition to the flood of Easter 1998. The EA Midlands Region in its Lower Severn Area provide a flood warning service supported by flood wardens and telephoned warnings to some individual properties.

River levels can fall below the intake of the gauging station stilling well: this has caused flows to be overestimated at low flows for the periods August-September 1995, July-October 1996 and June-October 1997.

The Easter 1998 flooding was associated with an almost saturated catchment prior to the onset of prolonged heavy rain starting on 9 April and lasting for two days. Rainfall totalling 66 mm over the 15 hour period from 04:00 9 April was recorded by the Shipston raingauge. The hydrograph peaked at 16:45 9 April at 63.21 mAOD, higher than the previous level of 62.85 mAOD recorded in 1968, and estimated to have a return period of 40-80 years. A total of 20 properties were flooded. Flood warnings were issued at 10:13 (yellow), 11:40 (amber) and 14:05 (red) on 9 April.



Plate 2.6.8 Flooding of the River Avon at Evesham, downstream of the confluence with the Stour, Easter 1998

Table 2.6.8 Data availability for the Stour at Shipston

Data Type	Station Name	NGR Coords.	Station Number	Data available
River Level	Stour at Shipston	SP260405	2092	1/1/90-13/9/98
Flow	Stour at Shipston	SP260405	2092	1/1/90-13/9/98
Rain	Chipping Campden	SP164393	1761	1/1/90-1/9/98
Rain	Langley	SP005282	1005	1/1/90-1/9/98
Rain	Stratford, Milcote	SP182529	1086	1/1/90-1/9/98
Rain	Shipston	SP268411	1087	1/1/90-1/9/98
Rain	Wellesbourne	SP271565	1165	1/1/90-1/9/98

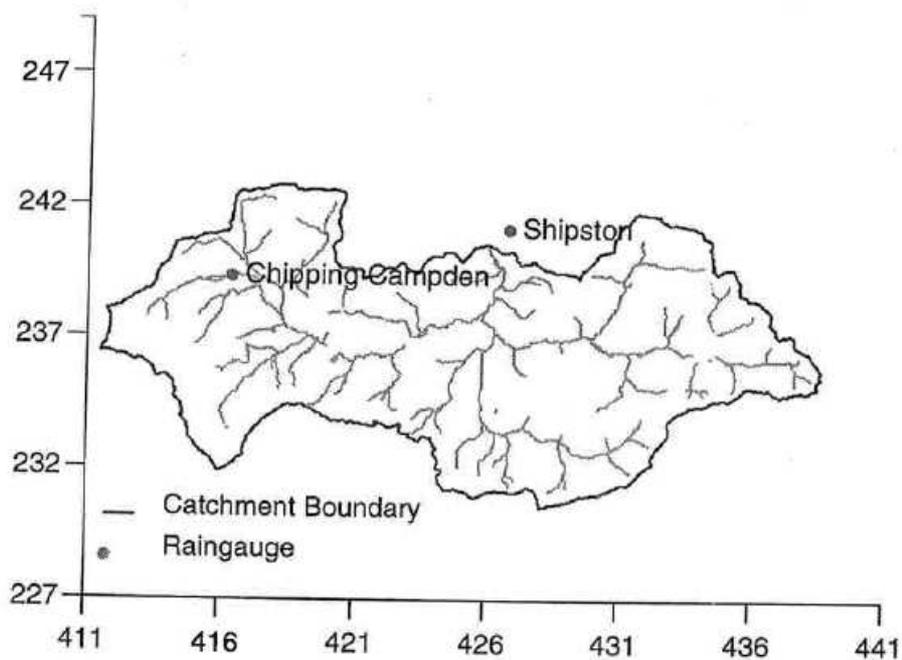


Figure 2.6.8 Catchment map for the Stour at Shipston.

2.6.9 Roch at Blackford Bridge

The catchment of the River Roch in north-west England drains an area of 186 km² to the gauging station at Blackford Bridge. It is highly urbanised in its lower half, encompassing the town of Rochdale, whilst there are peat moorlands in the upper reaches and mostly Coal Measures with Millstone Grit to the east. There are several water supply reservoirs in the headwaters. The gauging structure is a broad-crested mill-type curved (in plan) weir with a damaged crest and affected by debris. A rating derived from current metering is used with high flows measured from a road bridge 0.5 km upstream.



Plate 2.6.9 Flooding of the Roch, January 1995

Table 2.6.9 Data availability for the Roch at Blackford Bridge

Data Type	Station Name	NGR coords.	Station Number	Data available
River Level	Roch at Rochdale ETW	SD882127	690203	26/2/93-3/4/98
River Level	Roch at Blackford Bridge	SD807077	600205	1/1/90-3/4/98
Rain	Kitcliffe	SD960124	561468	1/2/90-10/3/98
Rain	Naden H'er Hill	SD837184	561777	1/1/90-10/3/98
Rain	Spring Mill Reservoir	SD875168	561660	1/3/90-12/3/98

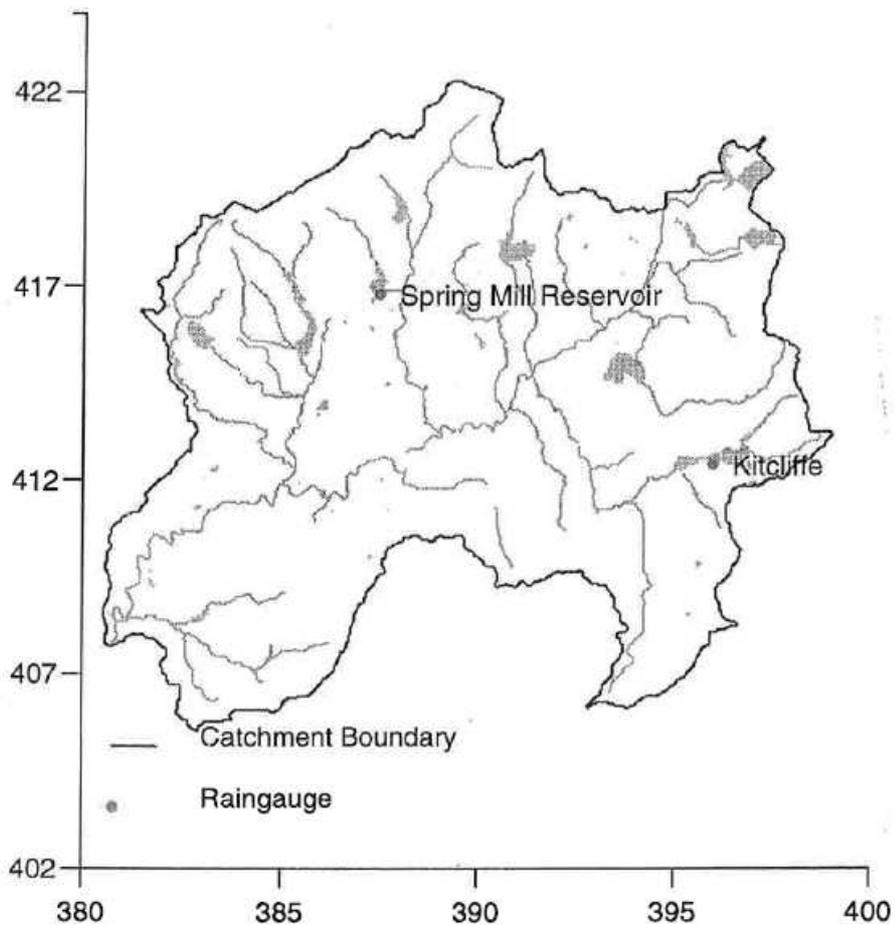


Figure 2.6.9 Catchment map for the Roch at Blackford Bridge.

2.6.10 Witham at Claypole Mill

The largely rural catchment of the River Witham to Claypole Mill in East Anglia drains an area of 298 km² with a lithology of clay (50%), limestone (40%) and gravel. The gauging structure is a Lea-designed broad-crested weir converted from an old weir at three levels. The structure is rated theoretically and there is no bypassing or drowning. There is an abstraction for public water supply at Saltersford.



Plate 2.6.10 Flooding of the River Witham at Lincoln, downstream of Bargate Weir

Table 2.6.10 Data availability for the Witham at Claypole Mill

Data Type	Station Name	NGR Coords.	Station Number	Data available
River Level	Witham at Claypole Mill	SK842480	30001	1/1/90-21/5/98
Rain	Brant Broughton	SK927546	S15	1/1/90-16/4/98
Rain	Ropsley	TF001336	U17	1/1/90-16/4/98
Rain	Saltersford	SK926335	S12	1/1/90-1/4/98
Rain	South Witham	SK929198	R05	1/1/90-16/4/98

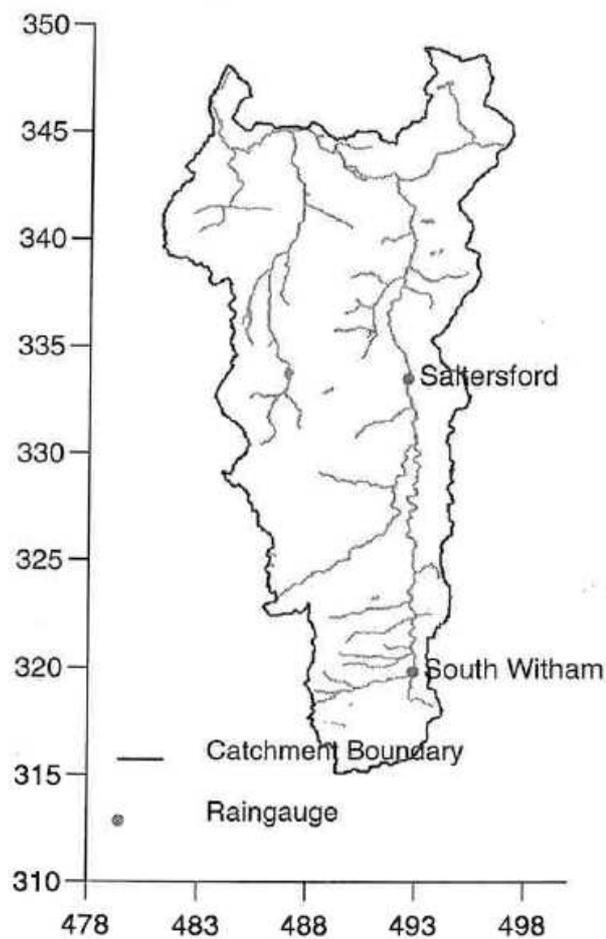


Figure 2.6.10 Catchment map for the Witham at Claypole Mill.

2.7 Data Management

2.7.1 Data quality control

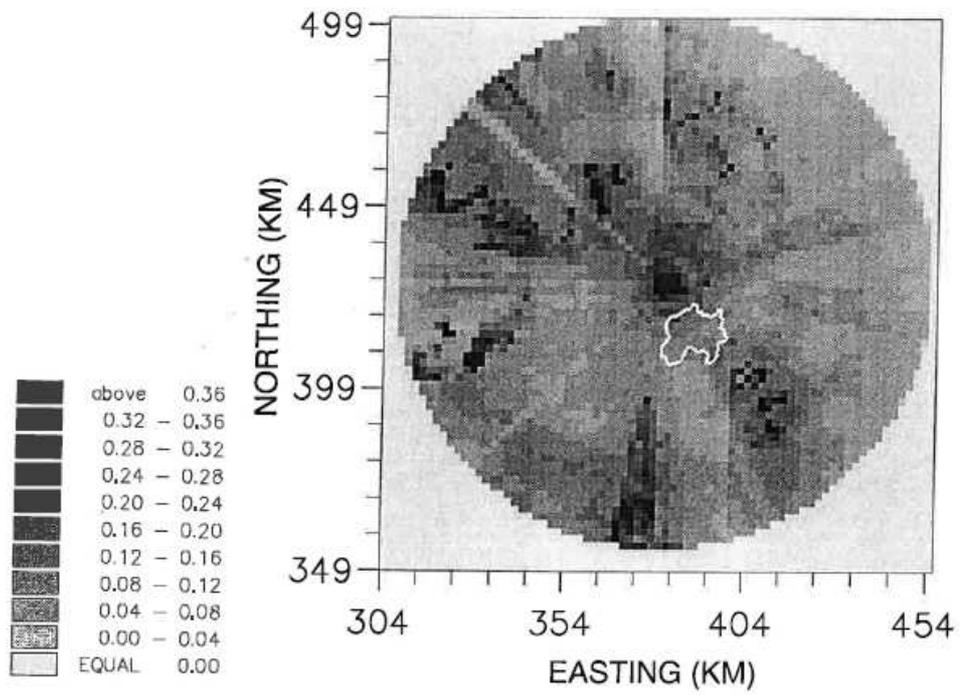
Raingauge data were quality controlled by comparing cumulative hyetographs of several neighbouring raingauge stations with each other. This served to expose periods when certain raingauges were malfunctioning or where data were absent. When assessing model performance a consistent set of raingauges was required for all periods, both for calibration and evaluation. In some cases raingauges were excluded from the study altogether because they malfunctioned at frequent intervals and so could not be used in a consistent way over all periods. Raingauges available for the Witham catchment were particularly affected by intermittent missing data. This, combined with only a small number of significant high flow periods, restricted the period of record available for model assessment for the Witham.

For Trout Beck, hourly data at raingauge AWS 31 was preferred to the fifteen minute data available for Cow Green since more good quality data were available and model performance was generally improved. This contrasts with results obtained within the Snowmelt Forecasting Study where AWS 31 proved to be a less useful data source in modelling flow during snow events. The number and duration of periods used here for Trout Beck was limited in order to avoid frequent winter snow periods.

The quality of **flow data** was assessed by plotting hydrographs to identify periods where data were missing or seemed to be poor. Most periods where this was the case were excluded, particularly where problems occurred over a long time. For the Roch and the Witham only level data were provided and rating equations were used to convert levels to flows. A revised rating developed by HR for the Roch at Blackford Bridge was used and for the Witham the existing EA rating for Claypole Mill was employed.

Data from three **radars** were used in the study at Clee Hill, Chenies and Hameldon Hill. Both uncalibrated and raingauge-calibrated radar data have been used. Anomalies in the radar field due, for example, to blockages or radar proximity were identified using anomaly maps constructed from long-term radar rainfall totals over 2 and 5 km pixel areas. Figures 2.7.1 and 2.7.2 show anomaly maps for the Hameldon Hill and Clee Hill radars. The figures indicate the position of the Roch and Stour catchments within the 2 and 5 km radar fields and show that the Stour is located at the edge of the Clee 2 km radar circle whilst the Roch is located close to the centre of Hameldon Hill 2 km radar circle. Examination of the position of the study catchments with respect to these anomaly maps revealed that they were not affected by anomalies. Uncalibrated radar data were used in raw form, whilst clutter removal techniques were applied to the radar data prior to calibration using software within HYRAD (Moore *et al.*, 1991; 1994b). Where a catchment was found to lie at the boundary between the 2 and 5 km radar fields, the radar fields were combined using methods within HYRAD. This was the case for the Stour and was made worse by an area of missing data in the 2 km field. Figure 2.7.3(a) reveals the spatial extent of missing 2 km data for Clee Hill, whilst Figure 2.7.3(b) shows the combined 2 and 5 km radar fields. The reason for these missing data stems from the Met. Office's use of the 10 km grid to centre the data grid for all radars whilst the Clee Hill radar is sited in the NE extremity of its 10 km grid square. This offset results in missing data on the southwest periphery of the 2 km radar circle. The area of missing 2 km data was replaced with 5 km data and the resulting fields shown in Figure 2.7.3(c) with the discontinuity in the 2 km circle completely removed.

(a) 2 km field



(b) 5 km field

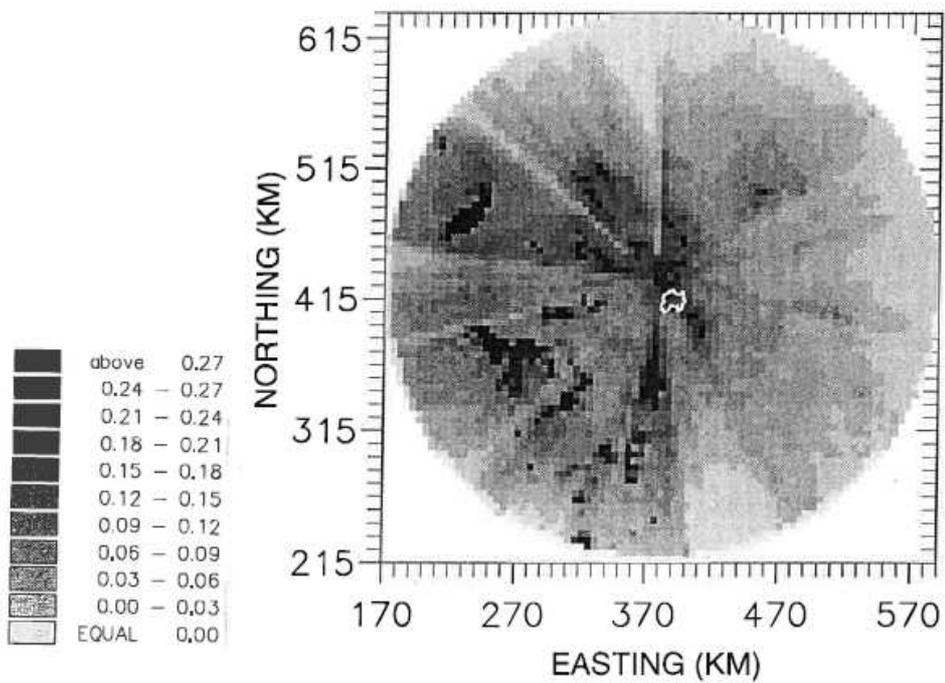
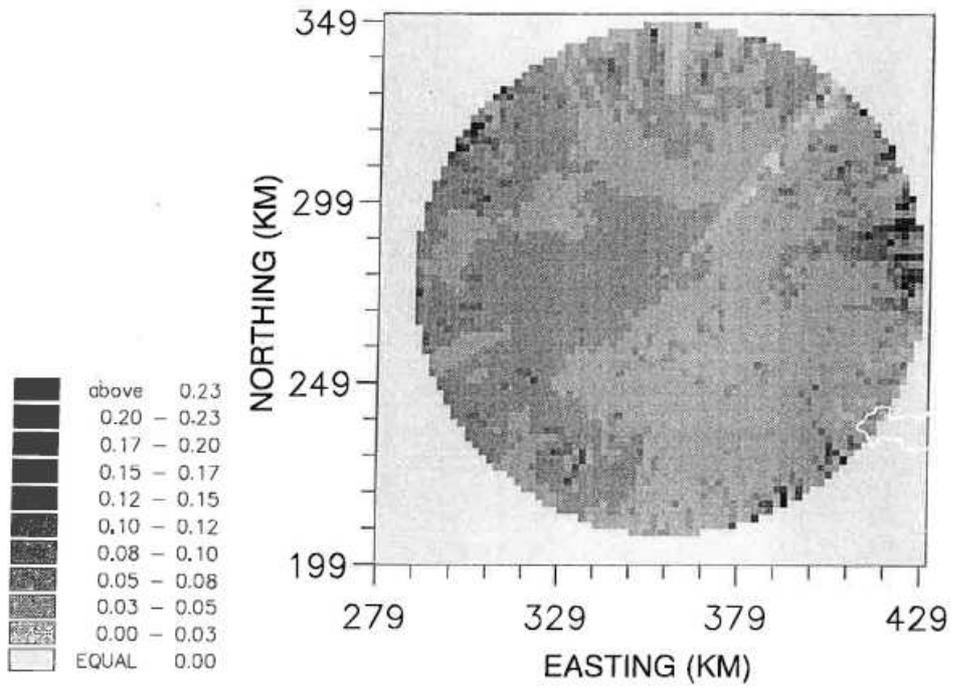


Figure 2.7.1 Anomaly maps constructed from long-term radar rainfall totals: Hameldon Hill radar. (Roch catchment superimposed)

(a) 2 km field



(b) 5 km field

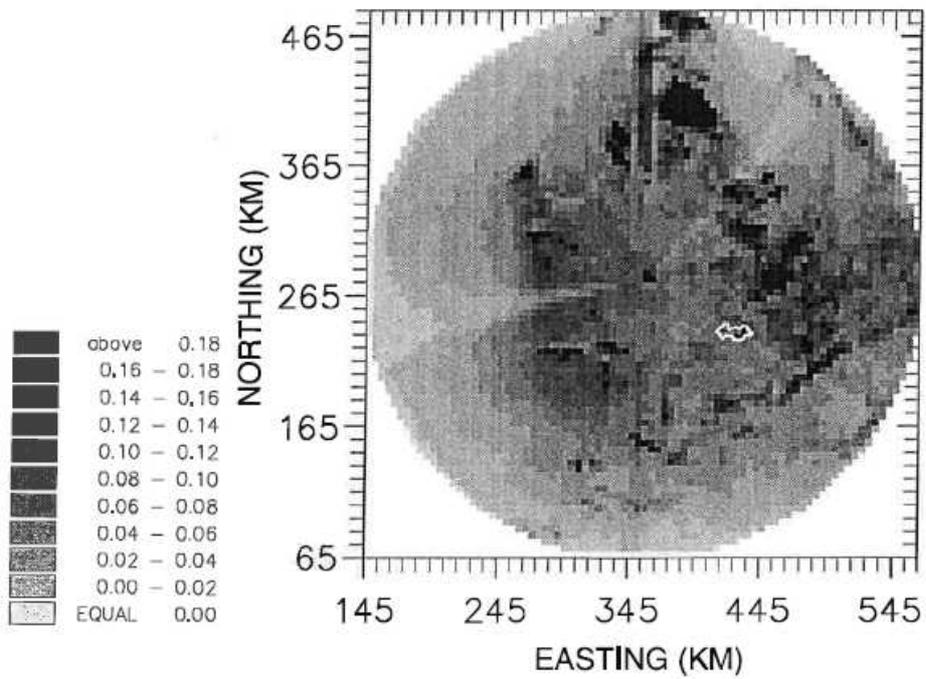
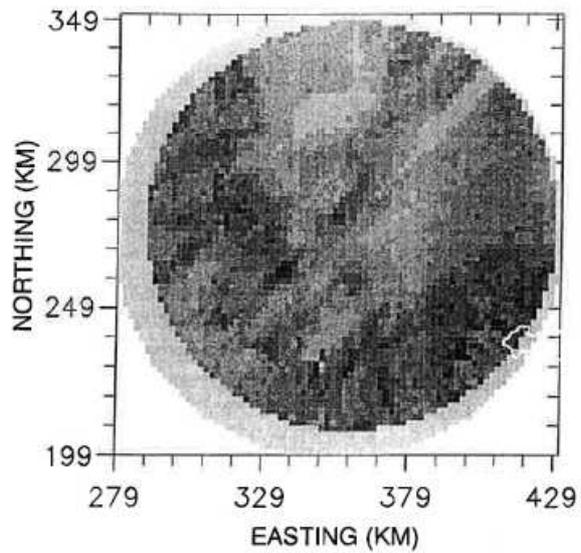
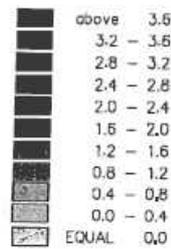
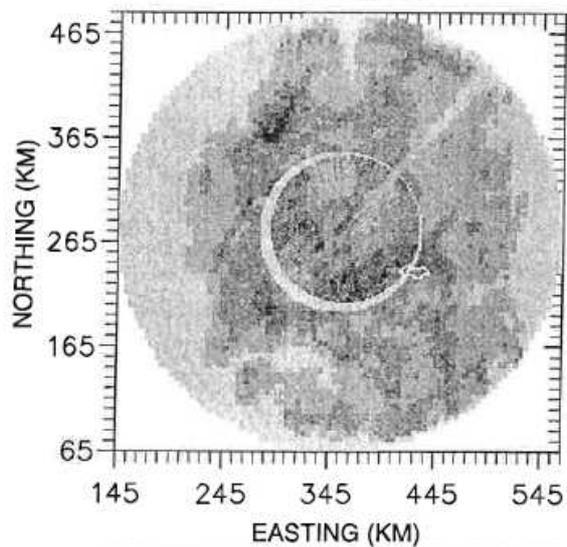
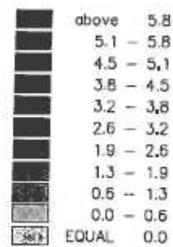


Figure 2.7.2 Anomaly maps constructed from long-term radar rainfall totals: Clee Hill radar. (Stour catchment superimposed)

(a) 2 km field



(b) Combined 2 and 5 km field



(c) Compositing 2 and 5 km field

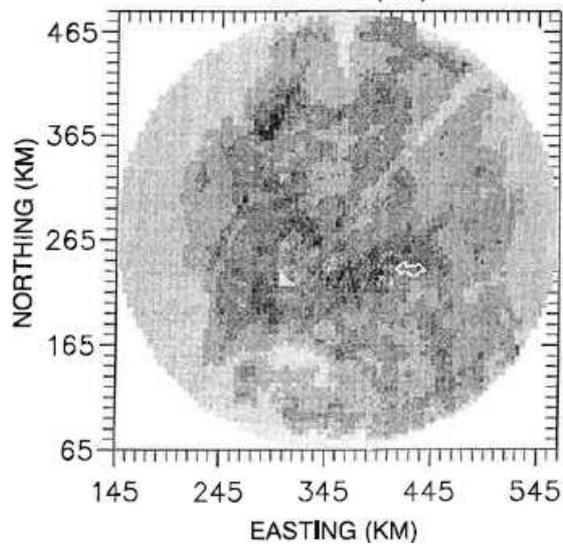
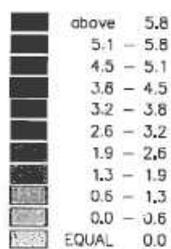


Figure 2.7.3 Compositing of 2 and 5 km radar fields from Clee Hill. (Stour catchment superimposed)

Plots of radar rainfall and raingauge totals were compared over the periods for which radar data were available and any significant periods where radar data were missing were excluded. Radar data would frequently be missing intermittently for a few hours and because of the limited extent of the data it was impractical to exclude every period where this occurred. For those times where enough radar data existed to perform a reasonable analysis of its performance, the radar data were infilled with raingauge data during its downtime. For a particular catchment the raingauge-calibrated radar data were infilled with the same catchment average raingauge rainfalls used in the model assessment. The uncalibrated radar data were infilled with the same catchment averages but adjusted for the mean-field-bias (see Section 2.4.3).

Figure 2.7.4 presents for April 1998 cumulative hyetographs of catchment average rainfall estimated from uncalibrated radar, calibrated radar and raingauge data over the Stour catchment. The cumulative hyetographs highlight the effect of raingauge calibration in correcting for the underestimation in the uncalibrated radar data. Note that the rainfall that resulted in the Easter 1998 Flood occurred on 8 April.

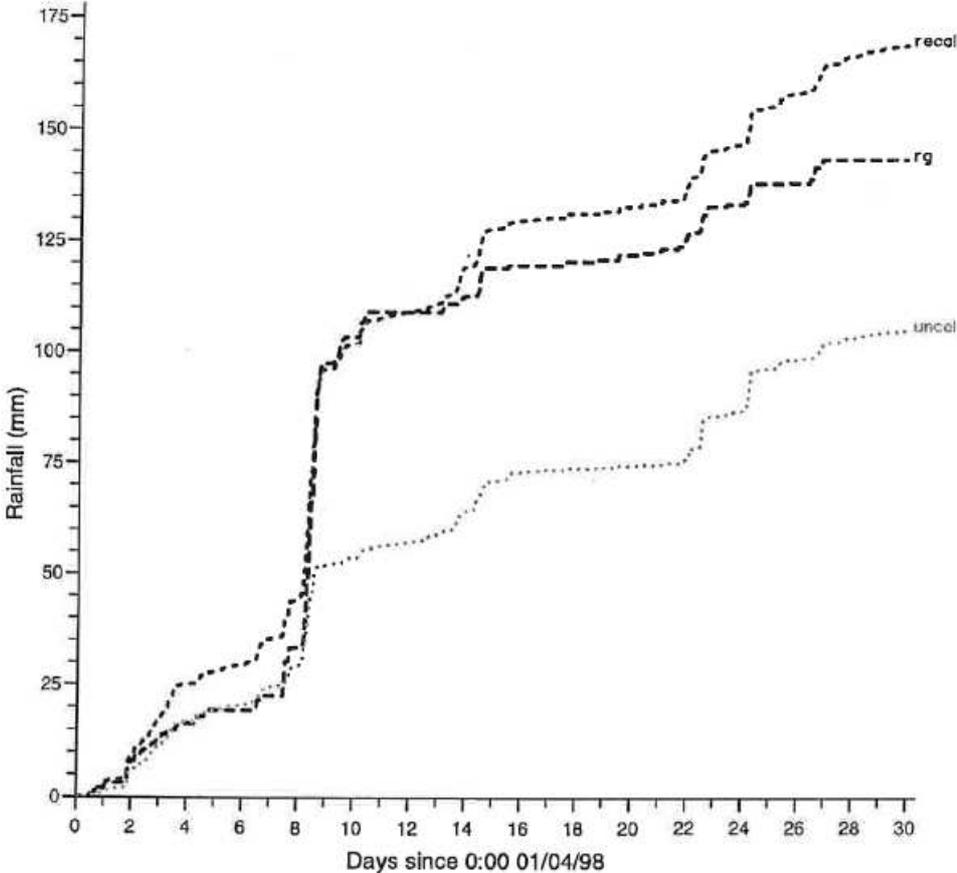


Figure 2.7.4 Cumulative hyetographs of catchment average rainfall over the Stour estimated using uncalibrated and calibrated radar and raingauge data: 1 to 30 April 1998.

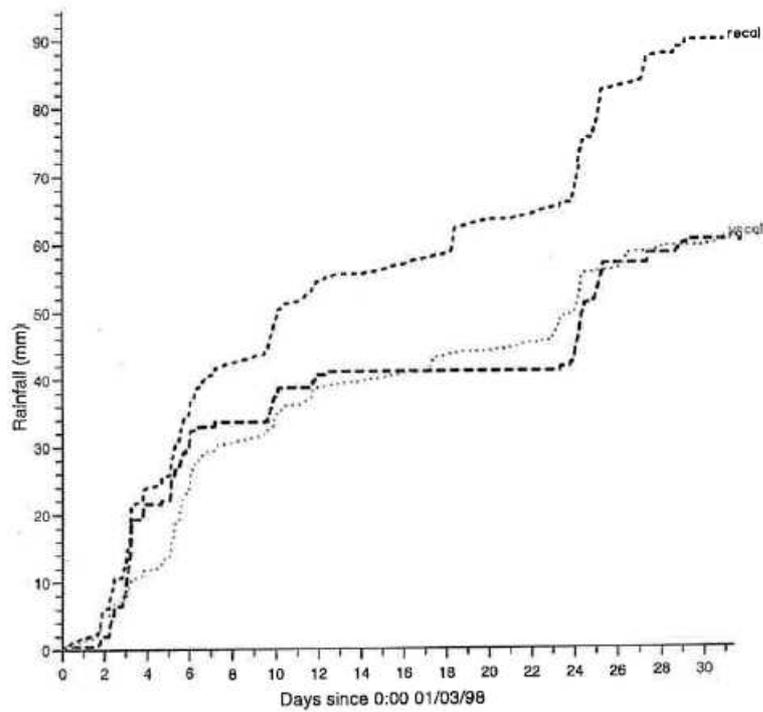
Similar cumulative hyetographs of rainfall over the Stour for January 1994 reveal a slight “drift” in the cumulative radar totals over periods when the raingauge measured no rainfall (Figure 2.7.5(a)). Further investigation has revealed that Cleve Hill radar experiences low levels of clutter that are not removed by the radar pre-processing algorithms. The radar rainfall field in Figure 2.7.6 shows six days accumulation of clutter across the field, and highlights that the Stour catchment lies in an area affected by the clutter. This low-level clutter is thought unlikely to have a significant affect on the assessment of model performance for the Stour and has not been removed. However, the clutter may be easily removed by introducing a threshold of 0.05 mm on each radar pixel value below which zero is used. The result of this correction is illustrated in Figure 2.7.5(b).

2.7.2 Database management

River flow and rainfall data provided by the EA regions were loaded onto the Model Database based on Oracle, allowing fast and easy access via the RFFS Model Calibration Facility. The database was originally developed for the NRA to provide a national archive of quality controlled snowmelt data to support future snowmelt studies. The database component of the Water Information System (WIS), a water-related GIS developed at the Institute of Hydrology in association with ICL, was chosen as the basis of the database. The underlying Oracle database provides multi-platform support via SQL, with the facility to import data in standard National Transfer Format (NTF).

The database holds raingauge, flow/level and evaporation data (where available) for all catchments at a 15 minute resolution. Daily raingauge data are also available to support calibration of the slow response parameters at a daily time-step. Spatially distributed uncalibrated radar data are stored in binary files before being corrected for static anomalies and calibrated with reference to raingauge data. The radar data are converted to catchment average rainfall values before being loaded onto the database. Distributed radar data over catchment areas required by the Grid Model are stored in ASCII files at a 15 minute resolution.

(a) Low-level clutter present



(b) Low-level clutter removed

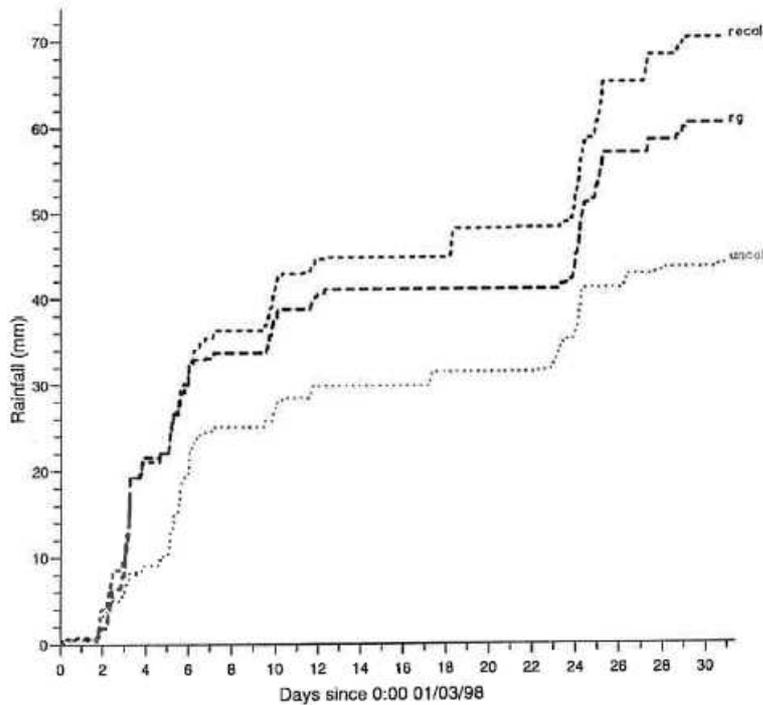


Figure 2.7.5 Cumulative hyetographs for the Stour showing uncalibrated and calibrated radar data and raingauge data. Clee Hill radar, 1 to 31 March 1998.

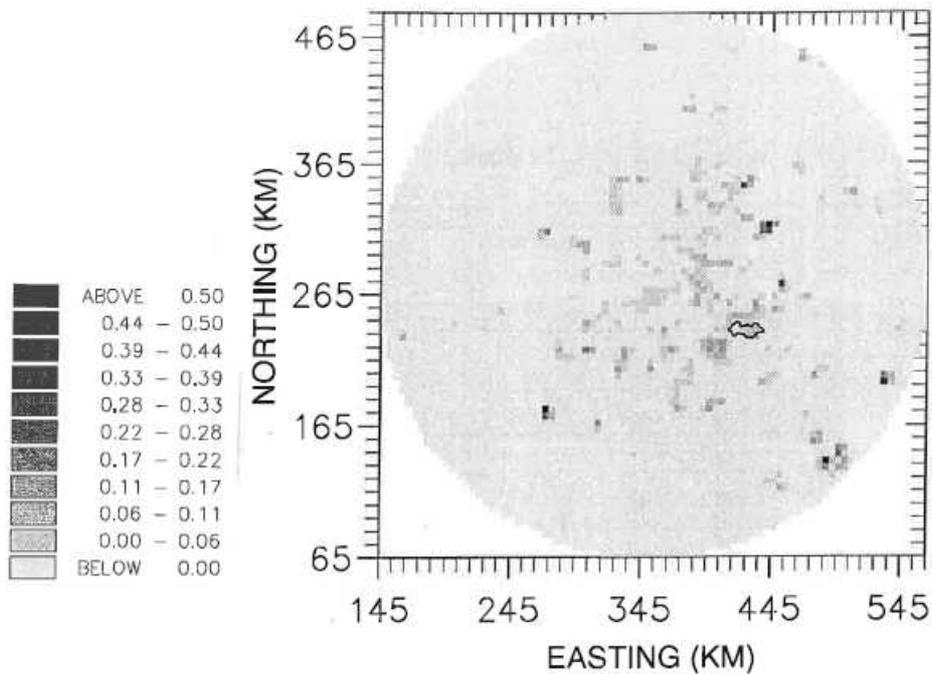


Figure 2.7.6 Radar rainfall field resulting from spurious clutter for the six day accumulation 13 to 19 March 1998 for Cleve Hill.

2.8 Selection of Calibration and Evaluation Periods

Periods of data to be used for calibration and evaluation were mostly chosen to be between 4 and 8 months duration. In the case of the Lavant, where groundwater levels respond over long periods of time, two of the three assessment periods were of the order of two years duration. For each catchment between 2 and 4 periods were selected for calibration and 2 to 5 for evaluation. Use of periods comprising of several months of data rather than single flood events ensured that calibrations would be more robust and that evaluations would give a more accurate indication of model performance for different catchments. The aim was to make as much use of good quality data as possible without individual periods becoming too large to work with efficiently. In general, the first part of a dataset was used for calibration and the second for evaluation. An exception was for the Lavant, so that the performance of the models over the important “Chichester Flood” could be assessed. Calibration was performed for each period and the parameter set that performed best across all periods was adopted.

Periods affected by snow for Trout Beck were identified by examining snow survey forms and excluded from the assessment. Trout Beck receives significant snowfall each year and this has meant that most of the periods included in the assessment occur between April and December, with the exception of one starting in late March. Snow survey forms exist for sites in the Dove and the Stour, but these were used in a less stringent way because of the requirement not to exclude more interesting events (particularly those covering Easter 1998). For other catchments where snow survey data are not available, any periods used in calibration which seemed significantly affected by snow were excluded. In evaluation, the freedom to change periods once defined did not exist. Examination of ‘Weather’, ‘Journal of Meteorology’ and the ‘Daily Weather Summary’ were used to exclude any obvious periods affected by snow before formal model evaluation began.

The periods finally adopted for model assessment using raingauge data are summarised in Table 2.8.1.

Table 2.8.1 Periods used for model calibration and evaluation: raingauge data

Catchment	Calibration periods	Evaluation periods
Trout Beck	1 July – 31 October 1991 16 April – 15 October 1992 1 April – 20 November 1993 2 August – 15 December 1994	13 April – 20 December 1995 15 May – 19 November 1996 20 March – 1 June 1998
Silk Stream	1 September 1989 – 15 January 1990 15 January – 15 September 1990 15 September 1990 – 15 May 1991	10 July 1991 – 1 January 1992 1 April – 7 October 1992
Dove	15 January – 15 March 1990 20 August 1992 – 31 January 1993 1 June – 1 December 1993 1 December 1993 – 30 June 1994	1 September 1994 – 22 March 1995 14 October 1996 – 14 July 1997 15 February 1998 – 20 July 1998
Lavant	1 June 1994 – 1 May 1996 25 December 1997 – 1 April 1998	8 December 1992 – 1 June 1994
Rhondda	15 October 1989 – 15 June 1990 15 September 1990 – 15 February 1991 12 March – 12 August 1991	10 September – 10 December 1991 22 February – 2 August 1992 31 October 1992 – 30 April 1993
Brue	1 October 1993 – 1 January 1994 1 January – 1 August 1994 1 August 1994 – 1 April 1995 1 November 1995 – 1 May 1996	1 November 1996 – 1 April 1997 1 November 1997 – 20 March 1998 20 March 1998 – 23 June 1998
Stour	8 January – 1 September 1990 1 January – 1 September 1991 1 November 1991 – 1 July 1992 1 July 1992 – 1 March 1993	24 March – 1 July 1993 28 September 1993 – 15 March 1994 15 March – 1 July 1994 22 October 1994 – 1 April 1995 15 November 1997 – 1 May 1998
Roch	15 July 1990 – 15 March 1991 1 August 1991 – 1 April 1992 1 November 1992 – 30 June 1993	15 June – 30 November 1994 30 November 1994 – 25 April 1995 1 May 1997 – 20 March 1998
Witham	1 February – 15 September 1991 1 May – 1 December 1992 1 January – 1 March 1993	1 September 1993 – 1 May 1994 15 November 1997 – 15 March 1998

The strategy for model assessment using weather radar data needed to be constrained due to the time-consuming nature and cost of handling large amounts of radar data. It was decided to focus on three catchments involving three radars: Chenies (London) for the Silk Stream, Hameldon Hill for the Roch, and Clee Hill for the Stour. For the Silk Stream an existing radar database at IH was made use of. However, no extensive database existed at IH for the Hameldon Hill and Clee Hill radars and data needed to be extracted from the Met. Office archive. A set of four or five wet months for winter and summer periods were identified for each of the two catchments with the purpose of encompassing both stratiform and convective storm periods. This menu of events is presented in Table 2.8.2 with M indicating the whole month is of interest. In practice, due to problems with missing or unreliable data the final set of periods used in the model assessment was much less extensive, as indicated in Table 2.8.3. Nimrod corrected data (Golding, 1998) have not been used since these data are for a 5 km resolution and are only available since circa 1998 as a stable product. The data used throughout are single site (Type 2) 2 km data, or 5 km data outside the 75 km range of the radar. The processing of these data to remove anomalies and their use in combination with raingauge data have been discussed in Sections 2.7.1 and 2.4.2 respectively.

Table 2.8.2 Menu of events for possible use with weather radar data

(a) Roch at Blackford Bridge (Hameldon Hill radar)

Summer (April - September)		Winter (October - March)	
Event	Peak Flow $\text{m}^3 \text{s}^{-1}$ (approx)	Event	Peak Flow $\text{m}^3 \text{s}^{-1}$ (approx)
April 1994 (M)	38	15 January - 15 February 1995 (M)	150
September 1994 (M)	25	December 1991 (16-27)	115
July 1994 (M)	23	January 1992 (3-10)	70
May 1997 (M)	23	November 1996 (M)	68
September 1993 (M)	23	December 1993 (M)	62

Standby level: $0.75 (27.8 \text{ m}^3 \text{ s}^{-1})$

(b) Stour at Shipston (Clee Hill radar)

Summer (April - September)		Winter (October - March)	
Event	Peak Flow $\text{m}^3 \text{s}^{-1}$ (approx)	Event	Peak Flow $\text{m}^3 \text{s}^{-1}$ (approx)
April 1998 (M)	90	January 1993 (M)	40+
15 May - 15 June 1992 (M)	13	December 1992 (M)	30+
September 1992 (M)	13	15 December 1997 - 15 January 1998 (M)	20
April 1993 (M)	13	February 1990 (M)	27
		January 1994 (M)	23

Alarm levels: Red: $3.6 \text{ m} (57.2 \text{ m}^3 \text{ s}^{-1})$; Amber: $2.5 \text{ m} (20 \text{ m}^3 \text{ s}^{-1})$; Yellow: $2.1 \text{ m} (15.5 \text{ m}^3 \text{ s}^{-1})$

Table 2.8.3 Periods used for model calibration and evaluation: radar data

Catchment	Calibration periods	Evaluation periods
Silk Stream	1 September 1989 – 15 January 1990 15 January – 15 July 1990 15 October 1990 – 15 May 1991	10 July 1991 – 1 January 1992 1 April – 30 November 1992
Stour	15 September – 1 October 1992 8 – 28 January 1993	1 – 31 January 1994 1 March – 30 April 1998
Roch	16 – 27 December 1991 5 – 30 April 1994	1 – 31 December 1993 1 – 30 September 1994 1 – 10 May 1997

3. MODEL CALIBRATION

3.1 Introduction

Eight rainfall-runoff models have been calibrated for nine catchments using approximately four years of data for each. All models have been calibrated using raingauge data, and for three catchments, the Roch, Silk Stream and the Stour, radar data have been used in the assessment alongside raingauge data. Typically, the four-year data record available for calibration has been split into three or four flow/rainfall records of 6-8 months duration. Poor quality or missing data have been avoided where possible. Following calibration in simulation mode, all models have been calibrated in updating mode using either ARMA error-prediction or state-updating, or, in the case of the PDM, both. Calibration results using raingauge and radar data are presented in Sections 3.2 and 3.3 respectively.

3.2 Model Calibration using Raingauge Data

3.2.1 Simulation mode calibration

The calibration procedure followed the guidelines outlined in Section 2 of this report. Results are presented in the form of a table of R^2 values in Table 3.2.1. The R^2 calibration results are presented in the Table with catchment size increasing down the page, and model complexity decreasing from left to right, with the distributed Grid Model to the far right as a special case. In addition, Figure 3.2.1 shows a graph of median R^2 (across periods) across the range of catchments using eight rainfall-runoff models. The lines connecting the results between catchments, while not formally correct, serve to highlight models that are performing well and catchments for which there is a large variation in model performance.

Table 3.2.1 and Figure 3.2.1 both indicate that, while no one model performs consistently best, overall the TCM is best, with the NWS, MCRM and PDM close behind and the TF and PRTF performing least well. The IEM performs well on small or fast catchments such as the Trout Beck, the Rhondda and the Roch, but lags behind the more complex conceptual models such as the MCRM on larger catchments with a significant baseflow component. For the three catchments on which the Grid Model was calibrated - the Stour, Roch and Witham - model performance was comparable to that obtained from the NWS, MCRM and PDM.

Figures 3.2.2 and 3.2.3 give examples of observed and simulated hydrographs for one of the calibration events in each of two catchments, the Silk Stream and the Stour. The event selection is biased towards better performance. The observed hydrographs are shown by *bold* lines and the simulated flows and their components (baseflow for the PDM and zonal flows for the TCM) by *fine* lines. The subsidiary plots show rainfall on an arbitrary scale, and (for the NWS, PDM and IEM) soil moisture deficit.

Table 3.2.1 Calibration performance (R^2 statistic) for different models, catchments and periods of record using raingauge data (best results in bold)

Catchment and period of record	Model							Grid
	TCM	NWS	MCRM	PDM	IEM	TF	PRTF	
<i>Trout Beck</i>								
1 July - 31 October 1991	0.900	0.873	0.862	0.853	0.892	0.748	0.648	
16 April - 15 October 1992	0.828	0.802	0.810	0.779	0.832	0.683	0.613	
1 April - 20 November 1993	0.815	0.814	0.809	0.820	0.830	0.743	0.688	
2 August - 15 December 1994	0.900	0.822	0.851	0.872	0.903	0.808	0.766	
<i>Silk Stream</i>								
1 September 1989 – 15 January 1990	0.825	0.824	0.847	0.880	0.842	0.802	0.806	
15 January - 15 September 1990	0.804	0.746	0.756	0.747	0.777	0.568	0.579	
15 September 1990 - 15 May 1991	0.757	0.556	0.773	0.643	0.702	0.515	0.489	
<i>Dove</i>								
15 January - 15 March 1990	0.826	0.688	0.788	0.813	0.739	0.719	-0.029	
20 August 1992 - 31 January 1993	0.815	0.831	0.834	0.839	0.702	0.362	0.267	
1 June - 1 December 1993	0.657	0.658	0.751	0.742	0.417	-2.797	-2.842	
1 December 1993 - 30 June 1994	0.893	0.834	0.843	0.778	0.625	0.660	0.442	
<i>Lavant</i>								
1 June 1994 - 1 May 1996	0.926	-	-	0.925	-	0.395	-	
25 December 1997 - 1 April 1998	0.631	-	-	-0.194	-	-0.056	-	
<i>Rhondda</i>								
15 October 1989 - 15 June 1990	0.929	0.922	0.899	0.917	0.905	0.818	0.818	
15 September 1990 - 15 February 1991	0.904	0.886	0.882	0.909	0.868	0.732	0.734	
12 March - 12 August 1991	0.879	0.872	0.871	0.875	0.795	0.625	0.626	
<i>Brue</i>								
1 October 1993 – 1 January 1994	0.868	0.825	0.944	0.854	0.743	0.693	0.644	
1 January – 1 August 1994	0.812	0.812	0.804	0.815	0.729	0.273	0.386	
1 August 1994 – 1 April 1995	0.912	0.905	0.850	0.834	0.804	0.503	0.568	
1 November 1995 – 1 May 1996	0.803	0.704	0.783	0.687	0.765	0.144	0.413	
<i>Stour</i>								
8 January - 1 September 1990	0.969	0.805	0.921	0.928	0.824	0.770	0.568	0.847
1 January - 1 September 1991	0.897	0.518	0.844	0.731	0.812	-0.764	-0.002	0.774
1 November 1991 - 1 July 1992	0.704	0.634	0.879	0.702	0.620	-0.511	0.447	0.577
1 July 1992 - 1 March 1993	0.701	0.795	0.741	0.743	0.519	0.476	0.337	0.729
<i>Roch</i>								
15 July 1990 - 15 March 1991	0.874	0.870	0.822	0.848	0.813	0.723	0.722	0.818
1 August 1991 - 1 April 1992	0.886	0.866	0.900	0.910	0.869	0.697	0.704	0.930
1 November 1992 - 30 June 1993	0.825	0.784	0.793	0.804	0.774	0.679	0.671	0.776
<i>Witham</i>								
1 February - 15 September 1991	0.903	0.601	0.858	0.795	0.633	0.032	-0.372	0.456
1 May - 1 December 1992	0.706	0.615	0.876	0.820	0.595	0.227	0.027	0.720
1 January - 1 March 1993	0.903	0.913	0.778	-0.201	0.555	-1.052	-1.894	0.851

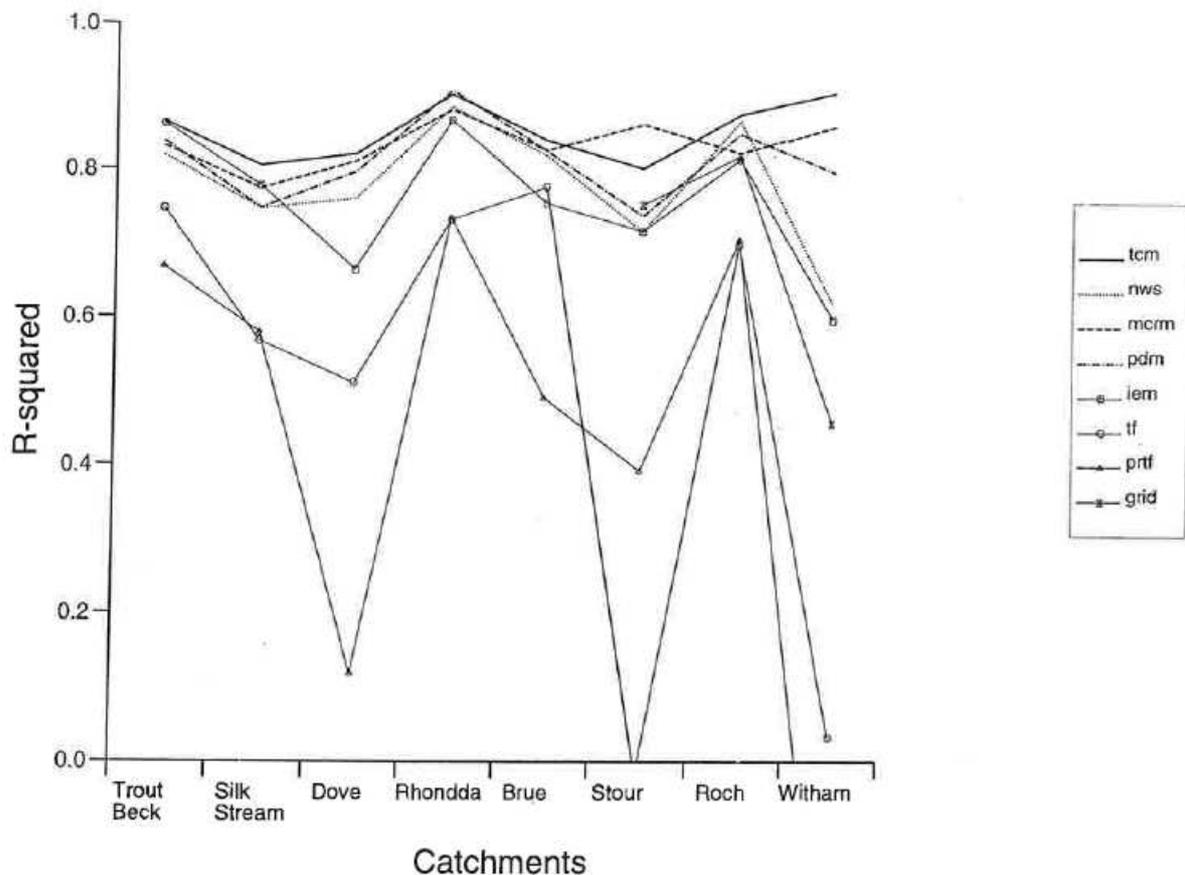


Figure 3.2.1 Variation in model performance (median R^2 statistic) across catchments.

Figure 3.2.2 illustrates results for an eight month calibration period for the Silk Stream from 15 January to 15 September 1990. The main flow peak of $16.9 \text{ m}^3 \text{ s}^{-1}$ occurred on 3 February and exceeded the amber flood warning threshold. Table 3.2.1 shows that, in terms of R^2 , the TCM performed best over this period, followed by the IEM. The NWS, PDM and MCRM were very similar in performance, being slightly worse than the IEM, and the TF and PRTF models were least successful. The PDM modelled the main flow peak of $16.9 \text{ m}^3 \text{ s}^{-1}$ most successfully, with an estimated $13.9 \text{ m}^3 \text{ s}^{-1}$. The TCM, NWS, MCRM and IEM all estimated the peak at approximately $12 \text{ m}^3 \text{ s}^{-1}$, compared with 8 to $9 \text{ m}^3 \text{ s}^{-1}$ using the TF and PRTF models. The main reason for the success of the TCM appears to be that although it was less good at predicting the main peak, it modelled the smaller spring and summer peaks very well.

Figure 3.2.3 illustrates results for the Stour over the calibration period 8 January 1990 to 1 September 1990. The main flow peak of $26.7 \text{ m}^3 \text{ s}^{-1}$ occurred on 2 February as a result of the wet weather that also resulted in the Silkstream event on 3 February. The best simulation result was obtained using the TCM, with an R^2 of 0.969, and the TF and PRTF performed least well. The highest flow peak on 2 February, which exceeded the amber warning level, was modelled most accurately by the NWS model ($28.1 \text{ m}^3 \text{ s}^{-1}$) with the PDM second best ($29.5 \text{ m}^3 \text{ s}^{-1}$). The Grid Model overestimated the peak at $33 \text{ m}^3 \text{ s}^{-1}$. The TF and PRTF performed reasonably well on the January flood, but the models' inability to model soil moisture/groundwater effects meant they were unable to reproduce the low flow conditions in the summer months, resulting in spurious flow peaks.

Impulse response functions for the TF and PRTF are shown in Figure 3.2.4. The impulse response functions (IRF's) are displayed in order of increasing catchment size, and IRF's for the PRTF are displayed alongside those derived for the TF models. Note that the IRF's for the larger catchments are displayed up to 200 time-steps (two days) ahead, while smaller catchments are displayed out to 100 time-steps. The figure shows how the smaller faster catchments have narrow, peaky impulse response functions, whereas the larger, slower catchments are modelled using longer, flatter IRF's. The Dove, which is placed between the Silk Stream and Rhondda in terms of catchment area, has a much flatter impulse response function than those catchments, because of its greater groundwater component. Calibrating a TF model for the groundwater-dominated Lavant proved almost impossible, and resulted in an unrealistic impulse response function which has been excluded from Figure 3.2.4.

In general, the IRFs from the TF models are more realistic looking than those for the PRTF models. For the Dove, the IRF for the PRTF is very flat and there are spikes in the responses for the Brue and the Stour. The PRTF equal root parameterisation restricts the flexibility of the form that the IRF can take. Also, the primary use of IV estimation for the TF models and simplex for the PRTF models will also have some effect. Factors which are thought to influence these results are use of a model structure which does not incorporate a variable model gain (runoff coefficient) with changing catchment wetness and use of estimation procedures which are not tailored to identify parallel flow processes with markedly contrasting volumes and speeds of response. The most well behaved IRF's are for the Rhondda and the Silk Stream where these two factors are arguably least influential.

3.2.2 Updating mode calibration

ARMA error prediction models were calibrated for all models except for the PRTF and the IEM for which state-updating schemes are used operationally. The order (p,q) of each ARMA model was identified from the autocorrelation and partial autocorrelation functions of the time series of model errors from the simulation mode rainfall-runoff models. Table 3.2.2 lists the model order of the ARMA(p,q) error-predictors for the different models and catchments. Although some of the model orders seem high, these were necessary to achieve stable and positive impulse response functions. An Approximate Maximum Likelihood for Events (AMLE) procedure was used to fit the resulting ARMA(p,q) models to the error series, and the resulting impulse response functions are shown in Figure 3.2.5.

Both the TF and PRTF models are used operationally in state-updating mode, where forecast flows are based on past rainfall and flow values. The parameters of the state-corrected form of the TF model are estimated using a recursive least squares algorithm, while the PRTF is recalibrated in updating mode using an embedded recursive least squares estimation of MA parameters within a simplex optimisation of t_{peak} (providing estimates of the AR parameters). Updating-mode impulse response functions (IRFs) for the TF and PRTF models are shown in Figure 3.2.6. The updating-mode IRFs resemble those obtained in simulation mode, except notably for the Rhondda and the Stour for which smoother IRFs are obtained in updating mode, and for the Brue, where recursive least squares results in an oscillatory impulse response function. Also a somewhat more physically realistic IRF is obtained in updating mode for the groundwater-dominated Lavant, which was extremely difficult to calibrate in simulation mode, but is not included in Figure 3.2.6. The model structure and method of estimation employed here again influence the form of IRF identified, as discussed in the

previous section with reference to the importance of catchment wetness and parallel flow processes on catchment response to rainfall.

Calibration of the state-updating parameters for PDM and IEM models was carried out using the automatic optimisation procedure, and proved to be straightforward. The PDM updating gains tended to be between 0.0 and 2.0, whereas the IEM updating gains were between 1.0 and 2.0. The model-gain updating smoothing parameter used by the TF and PRTF models was calibrated using automatic optimisation. The results, presented in Table 3.2.3, indicate that faster catchments such as Silk Stream and Trout Beck require a high smoothing factor, whilst catchments where soil moisture effects are more important, such as the Roch and Stour, need a much smaller smoothing factor, resulting in a more time-varying model-gain.

3.3 Model Calibration using Radar Data

3.3.1 Simulation-mode calibration

Model calibration using radar data was undertaken on three catchments only: the Silk Stream, the Stour and the Roch. The calibration procedures using uncalibrated and recalibrated radar data were identical to those used with raingauge data. Results are presented in the form of tables of R^2 values in Tables 3.3.1(a) and (b). The Tables highlight the improvement in model performance for the Silk Stream following calibration of the radar using raingauge data, although the beneficial effects are less obvious for the Stour and Roch.

Figure 3.3.1 presents a set of eight calibration hydrographs for the Silk Stream, comparing uncalibrated and recalibrated radar data for the MCRM, TCM, PDM and IEM rainfall-runoff models. The set of hydrographs highlights the improvement obtained using the raingauge calibrated radar data. When uncalibrated radar data are used in the models, an erroneous peak of up to $12 \text{ m}^3 \text{ s}^{-1}$ is forecast a few days after the main flow peak of $13 \text{ m}^3 \text{ s}^{-1}$, while the main flow peak is underestimated by approximately 50%. When calibrated radar data are used, the main flow peak is simulated well by most models, and the spurious second large peak vanishes.

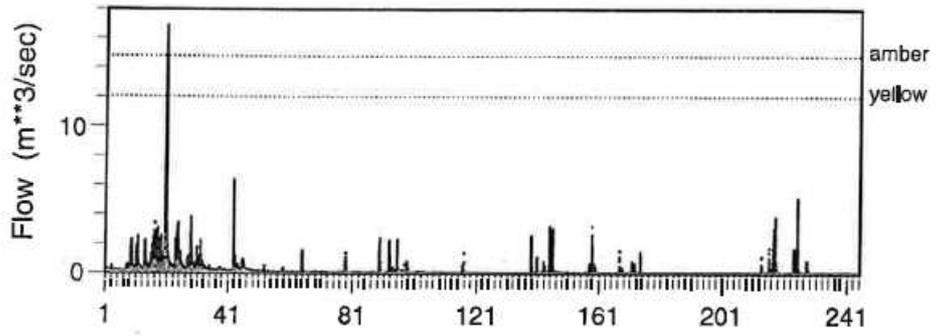
Figure 3.3.2 displays a similar set of hydrographs for the Stour. The main peak of $45 \text{ m}^3 \text{ s}^{-1}$ exceeds the amber flood warning threshold and was modelled most accurately by the MCRM using *uncalibrated* radar data. However, as Table 3.3.1 shows, use of calibrated radar data gives the best overall performance over both calibration periods for the Stour.

Although it appears that the most successful models at the model calibration stage are the MCRM and the TCM, the short periods of radar data used for calibration (typically 2 weeks) may have resulted in a degree of overfitting. Firmer conclusions should therefore be drawn from the evaluation results.

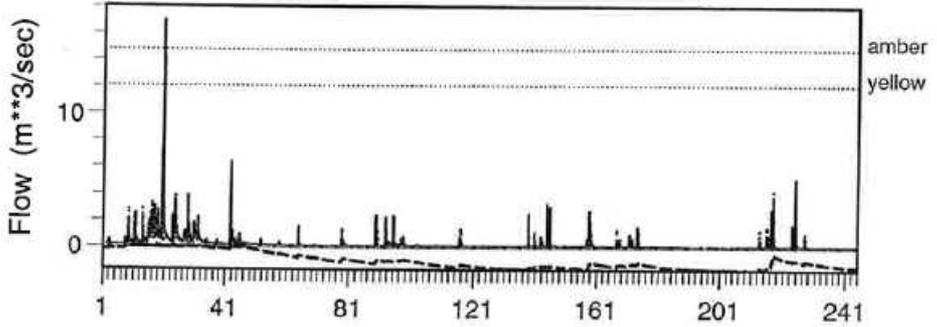
3.3.2 Updating-mode calibration

ARMA error prediction models were calibrated for all models except for the TF, PRTF, and the IEM, for which state-updating schemes are used operationally. Table 3.2.2(b) lists the model order of the ARMA error-predictors for the different models and catchments. An Approximate Maximum Likelihood for Events (AMLE) procedure was used to estimate the model parameters.

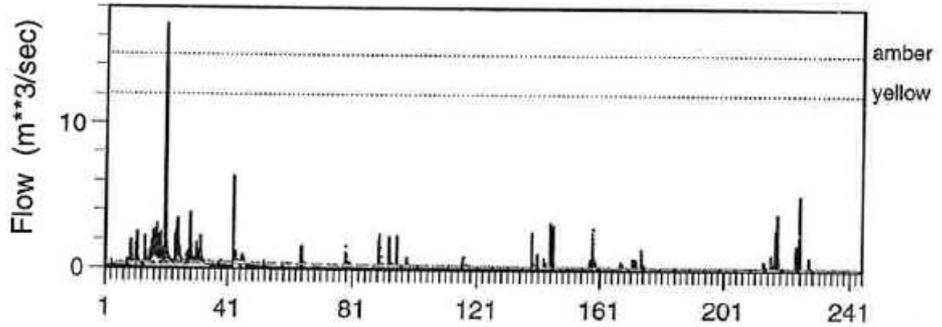
(a) TCM



(b) NWS



(c) MCRM



(d) PDM

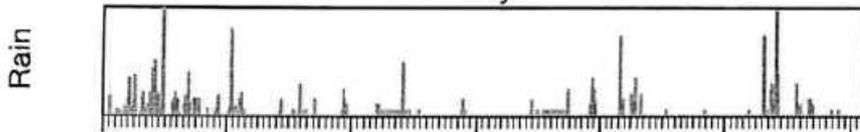
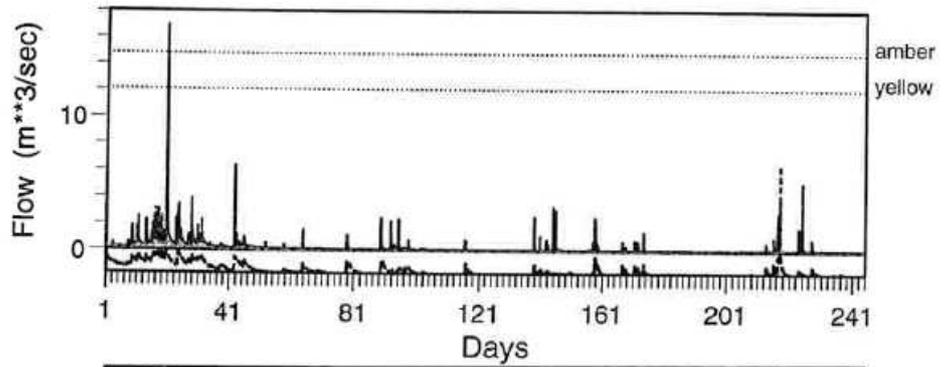
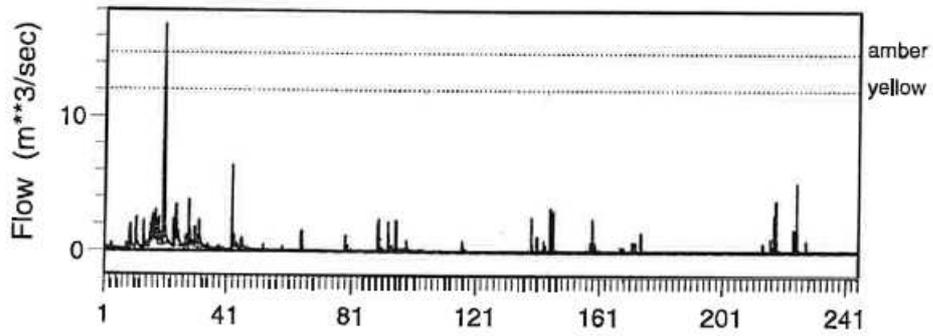
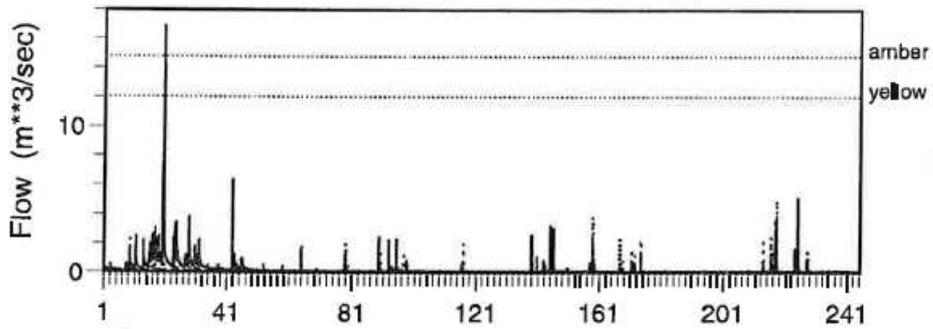


Figure 3.2.2 Model calibration: Flow hydrographs for the Silk Stream, 15/1/90 – 15/9/90.

(e) IEM



(f) TF



(g) PRTF

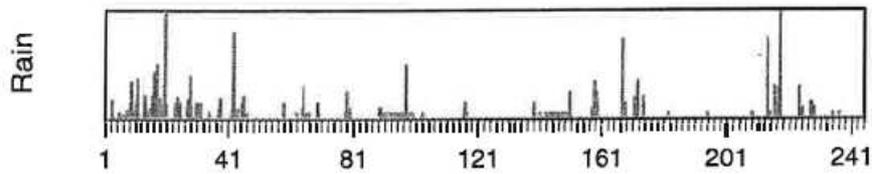
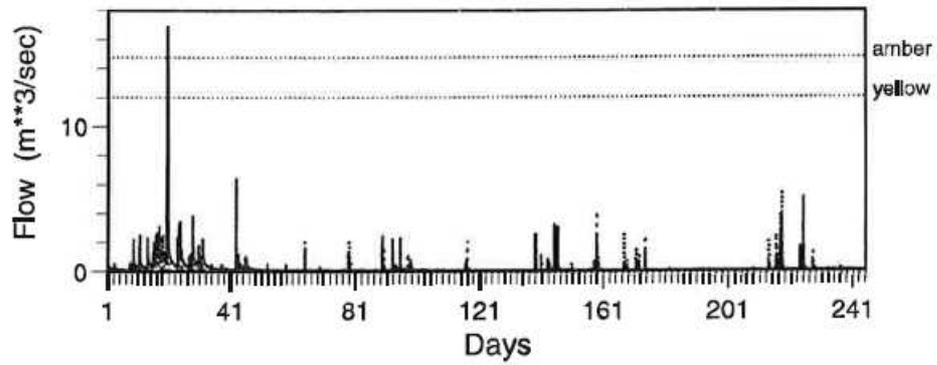
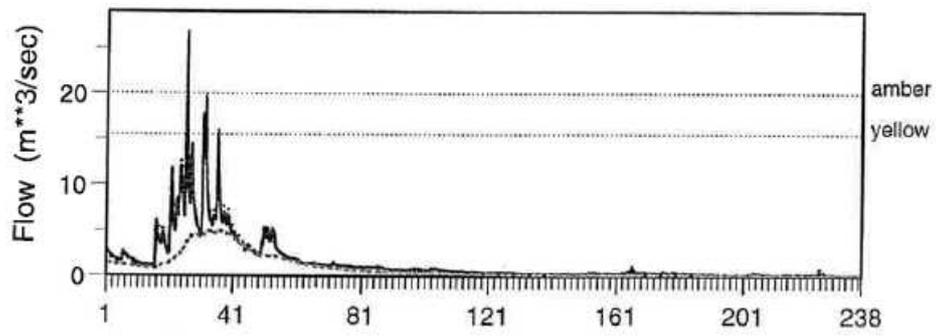
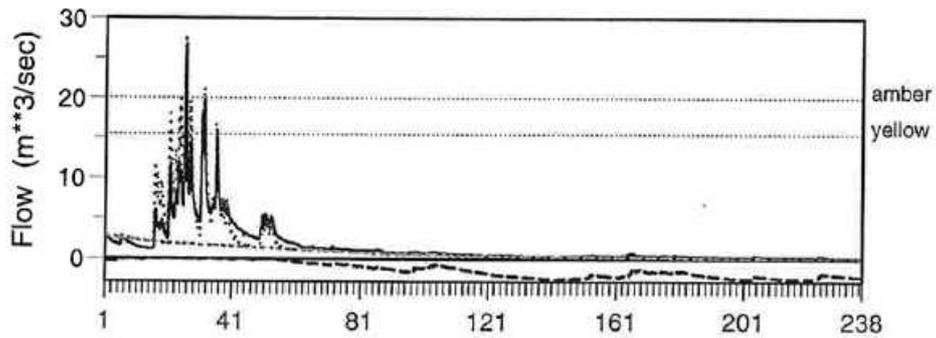


Figure 3.2.2 (cont...) Model calibration: Flow hydrographs for the Silk Stream, 15/1/90 – 15/9/90.

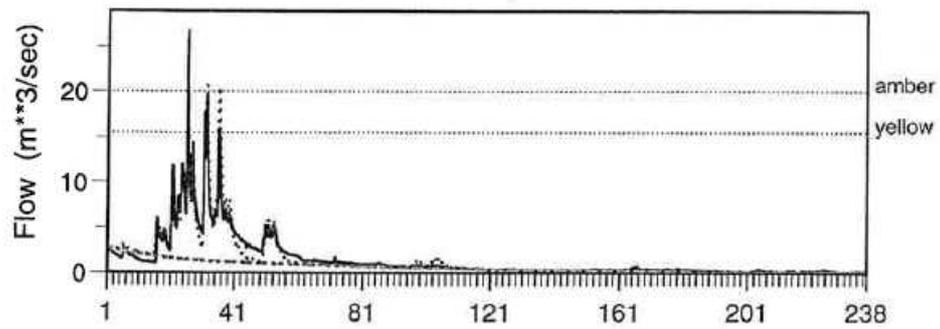
(a) TCM



(b) NWS



(c) MCRM



(d) PDM

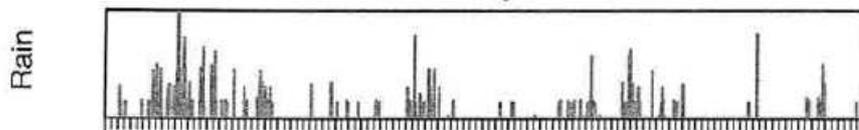
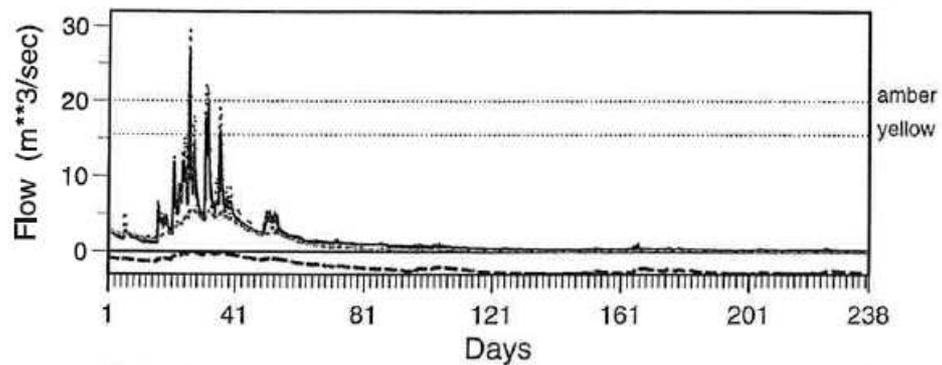
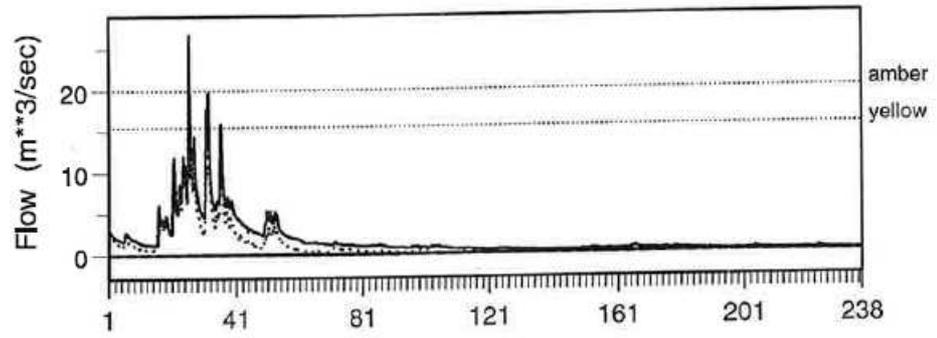
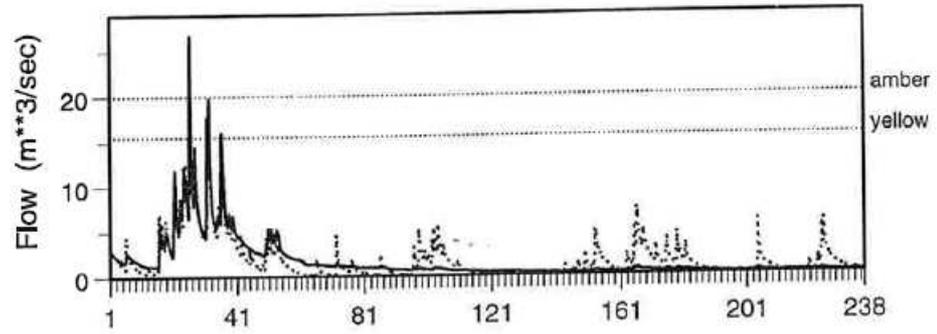


Figure 3.2.3 Model calibration: Flow hydrographs for the Stour, 8/1/90 – 1/9/90.

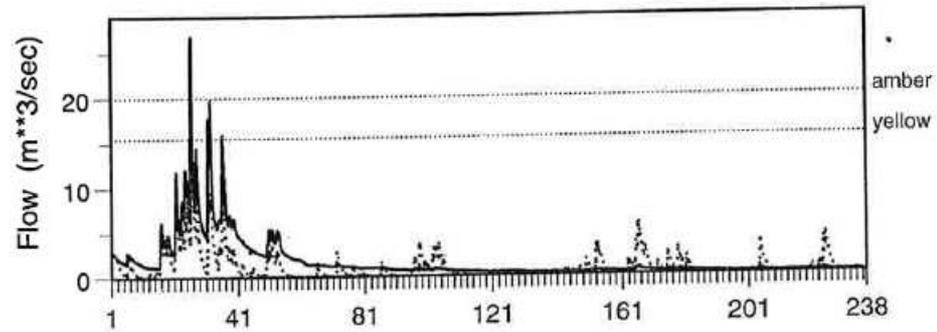
(e) IEM



(f) TF



(g) PRTF



(h) Grid Model

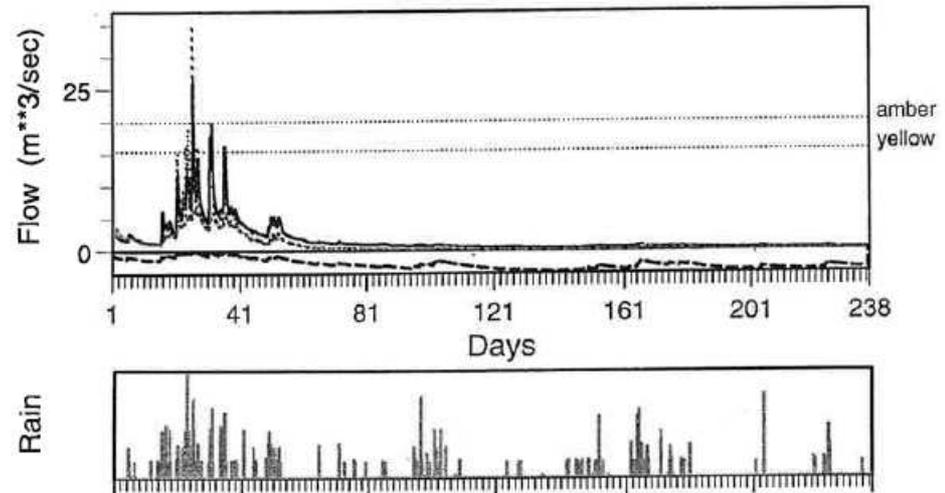


Figure 3.2.3 (cont...) Model calibration: Flow hydrographs for the Stour, 8/1/90 – 1/9/90.

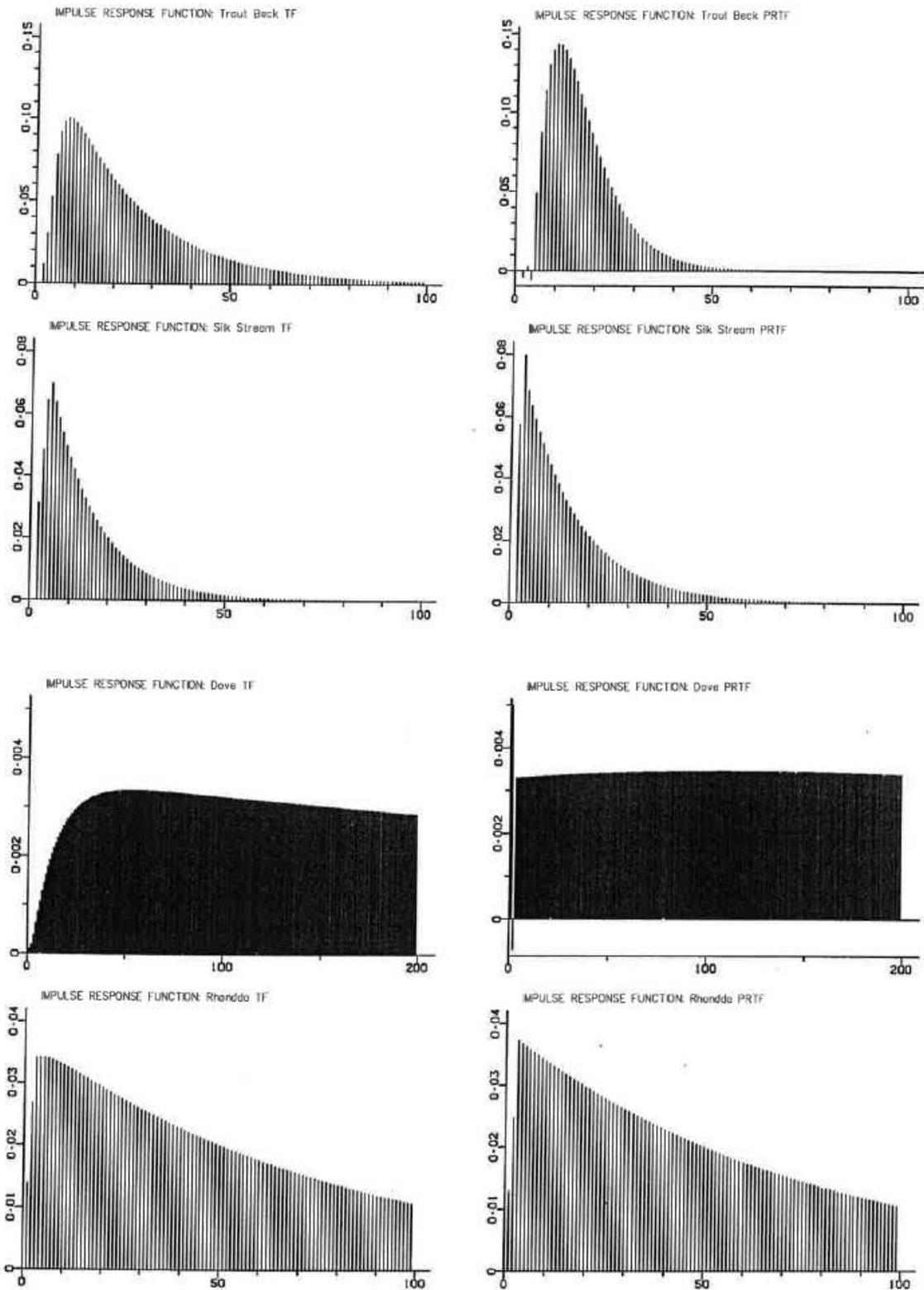


Figure 3.2.4 Impulse response functions for the TF and PRTF models calibrated in simulation mode.

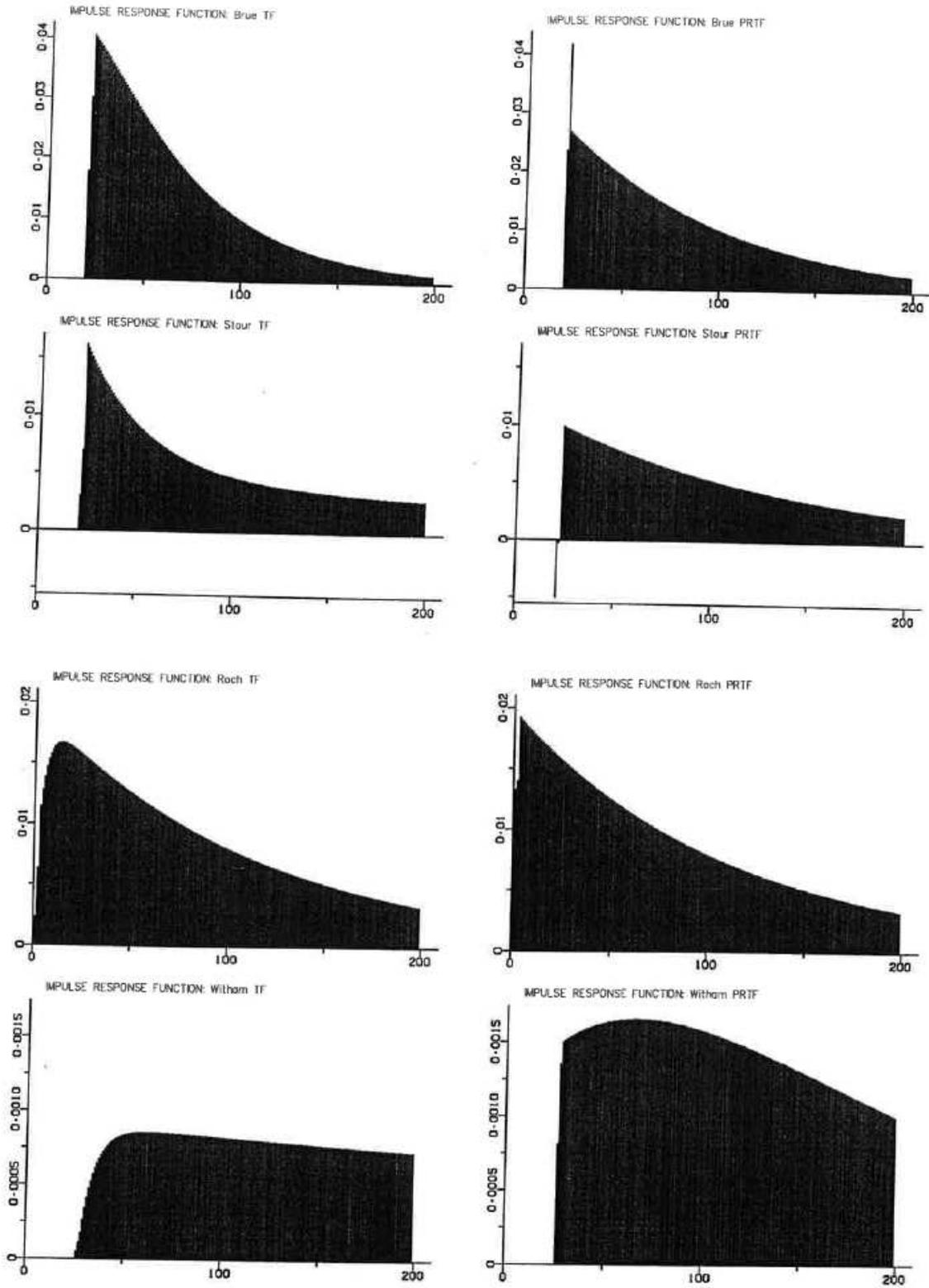


Figure 3.2.4 (cont...) Impulse response functions for the TF and PRTF models calibrated in simulation mode.

Table 3.2.2 Model order of ARMA(p,q) error predictors for different models and catchments (SU indicates where state updating has been used in preference to ARMA error prediction)

(a) Raingauge data

Catchment	Model							
	TCM	NWS	MCRM	PDM	IEM	TF	PRTF	Grid
Trout Beck	(2,0)	(6,0)	(7,0)	(2,0)	SU	(2,0)	SU	-
Silk Stream	(4,0)	(5,0)	(5,0)	(5,0)	SU	(4,0)	SU	-
Dove	(4,0)	(4,0)	(4,0)	(4,0)	SU	(5,0)	SU	-
Lavant	(6,0)	-	-	(6,0)	-	SU	-	-
Rhondda	(2,0)	(2,0)	(2,0)	(2,0)	SU	(2,0)	SU	-
Brue	(4,0)	(7,0)	(8,0)	(4,0)	SU	(5,0)	SU	-
Stour	(7,0)	(4,0)	(7,0)	(7,0)	SU	(5,0)	SU	(5,0)
Roch	(2,0)	(4,0)	(2,0)	(2,0)	SU	(5,0)	SU	(2,0)
Witham	(5,0)	(3,0)	(6,0)	(4,0)	SU	(5,0)	SU	(3,1)

(b) Radar data

Catchment	Model							
	TCM	NWS	MCRM	PDM	IEM	TF	PRTF	Grid
Silk Stream	(2,0)	(3,0)	(3,0)	(2,0)	SU	SU	SU	-
Stour	(3,0)	(2,2)	(2,2)	(3,0)	SU	SU	SU	(3,0)
Roch	-	-	-	-	-	-	-	-

Table 3.2.3 Comparison of model-gain updating smoothing factors across catchments

Model	Trout Beck	Silk Stream	Dove	Lavant	Rhondda	Brue	Stour	Roch	Witham
TF	0.999	1.000	0.536	0.967	0.997	0.612	0.143	0.278	0.531
PRTF	0.905	0.995	0.536	-	0.073	0.989	0.240	0.223	0.004

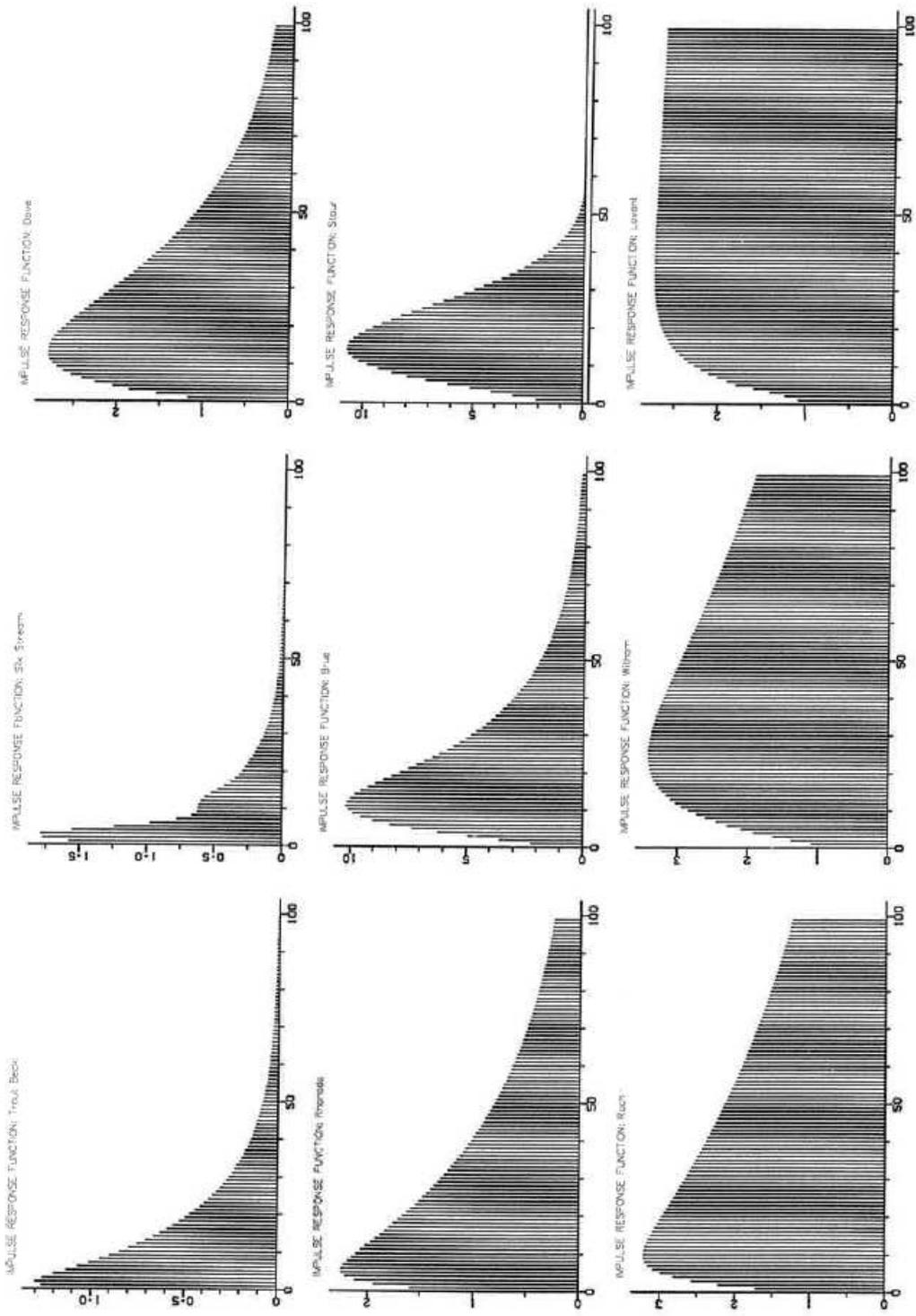


Figure 3.2.5 Impulse response functions of ARMA error-predictors.

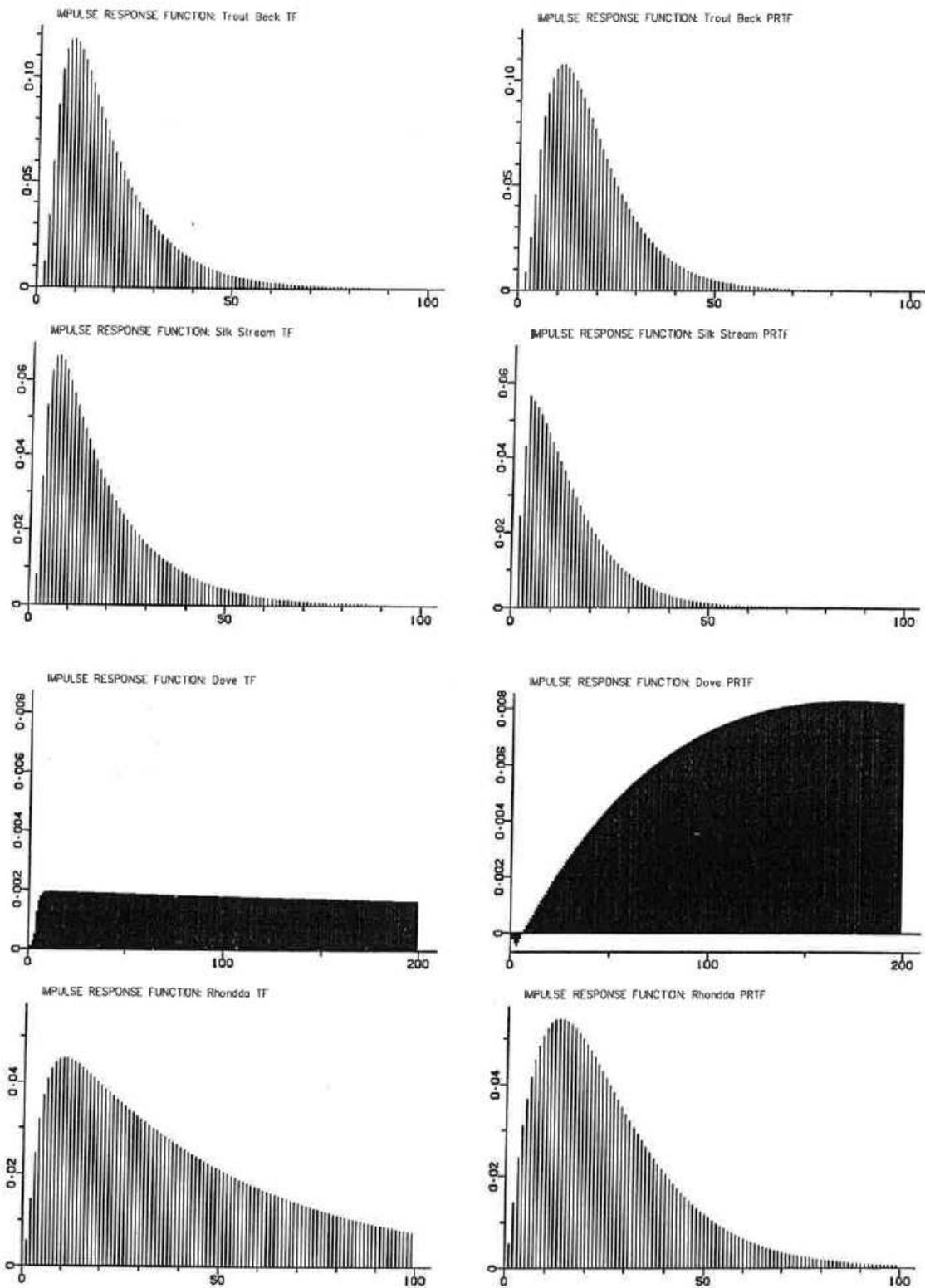


Figure 3.2.6 Impulse response functions of TF and PRTF models calibrated in updating mode.

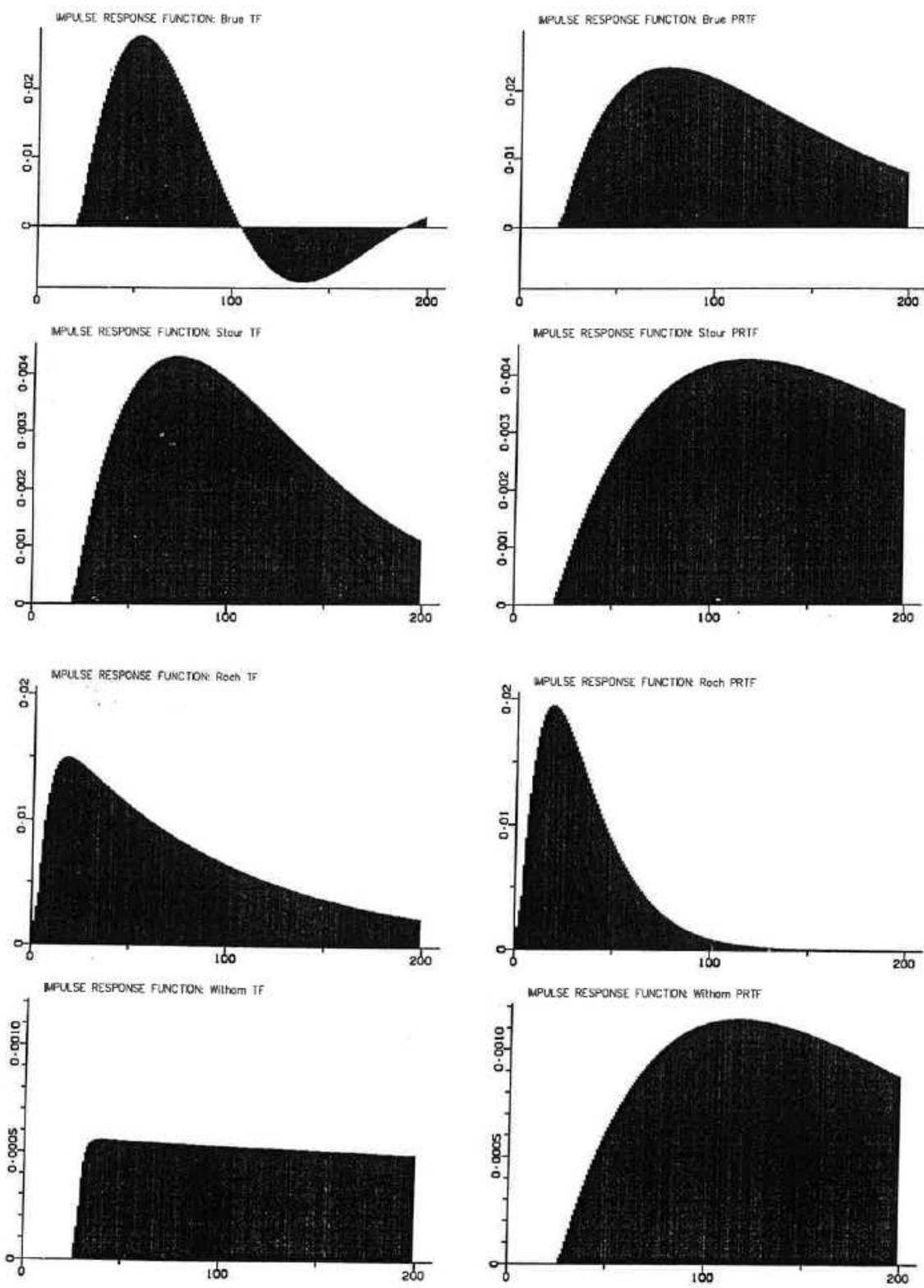


Figure 3.2.6 (cont...) Impulse response functions of TF and PRTF models calibrated in updating mode.

Table 3.3.1 Calibration performance (R^2 statistic) for different models, catchments and periods of record using radar data (best results in bold)

(a) Uncalibrated radar data

Catchment and event	Catchment							Grid
	TCM	NWS	MCRM	PDM	IEM	TF	PRTF	
<i>Silk Stream</i>								
1 September 1989 – 15 January 1990	0.610	0.549	0.639	0.582	0.582	0.534	0.544	-
15 January - 15 July 1990	0.633	0.707	0.668	0.696	0.608	0.591	0.584	-
10 October 1990 - 15 May 1991	0.400	-0.247	0.434	0.302	0.300	-0.141	-0.018	-
<i>Stour</i>								
15 September - 1 October 1992	0.756	0.832	0.893	0.821	0.625	0.820	0.769	0.733
8 - 28 January 1993	0.914	0.841	0.916	0.872	0.836	0.345	0.264	0.894
<i>Roch</i>								
16 - 27 December 1991	0.970	0.938	0.985	0.972	0.952	0.561	0.536	0.902
5 - 30 April 1994	0.655	0.917	0.810	0.831	0.692	0.664	0.677	0.740

(b) Recalibrated radar data

Catchment and event	Catchment							Grid
	TCM	NWS	MCRM	PDM	IEM	TF	PRTF	
<i>Silk Stream</i>								
1 September 1989 – 15 January 1990	0.885	0.820	0.881	0.872	0.832	0.704	0.746	-
15 January - 15 July 1990	0.842	0.820	0.782	0.828	0.818	0.543	0.666	-
10 October 1990 - 15 May 1991	0.783	0.424	0.795	0.718	0.726	0.691	0.484	-
<i>Stour</i>								
15 September - 1 October 1992	0.676	0.843	0.912	0.782	0.612	0.767	0.734	0.771
8 - 28 January 1993	0.860	0.873	0.943	0.896	0.680	0.252	0.184	0.870
<i>Roch</i>								
16 - 27 December 1991	0.958	0.956	0.988	0.984	0.956	0.464	0.435	0.919
5 - 30 April 1994	0.650	0.929	0.845	0.808	0.683	0.678	0.703	0.718

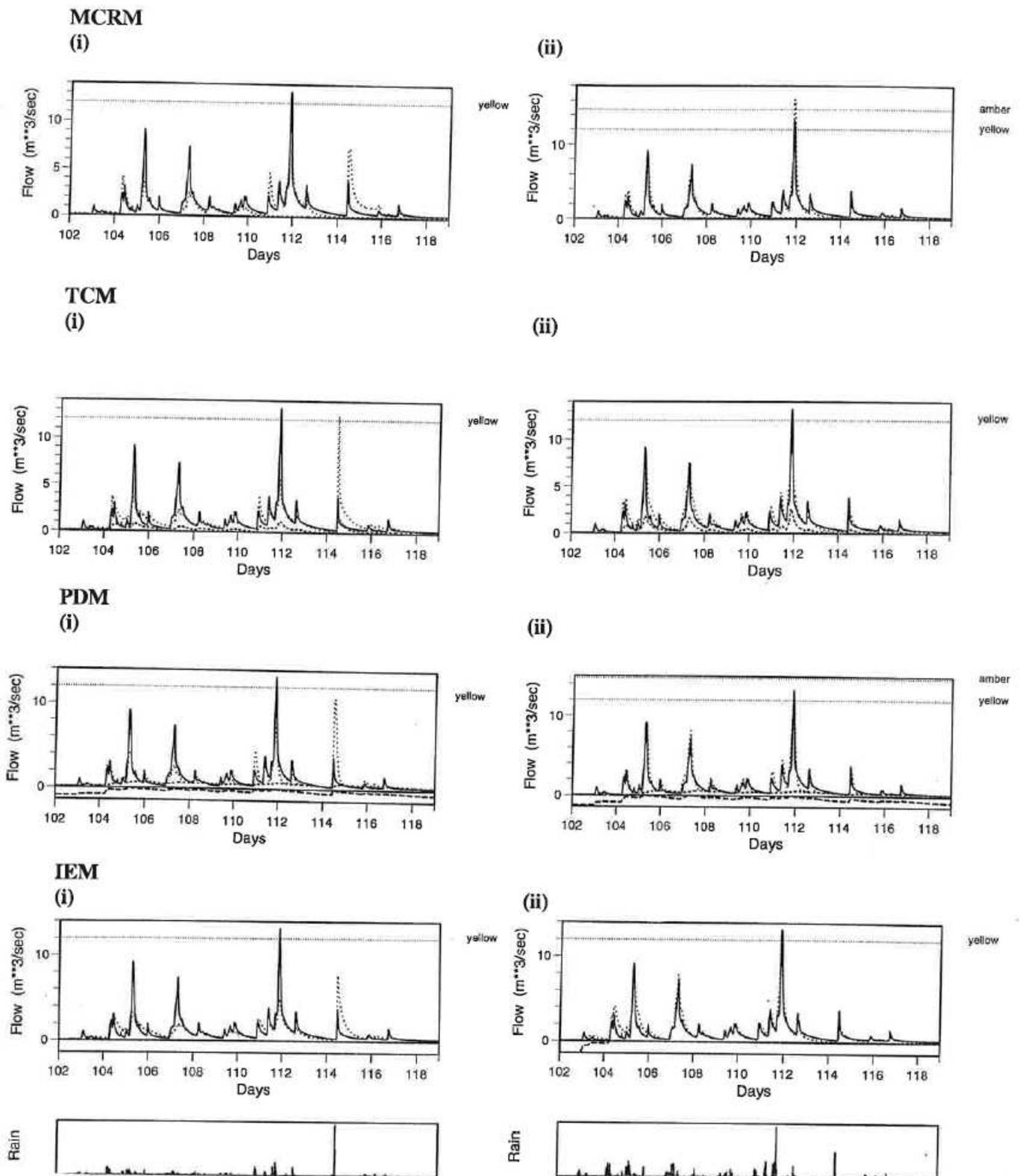


Figure 3.3.1 Simulation-mode hydrographs using (i) uncalibrated and (ii) calibrated radar data: Silk Stream, 12 to 28 December 1989.

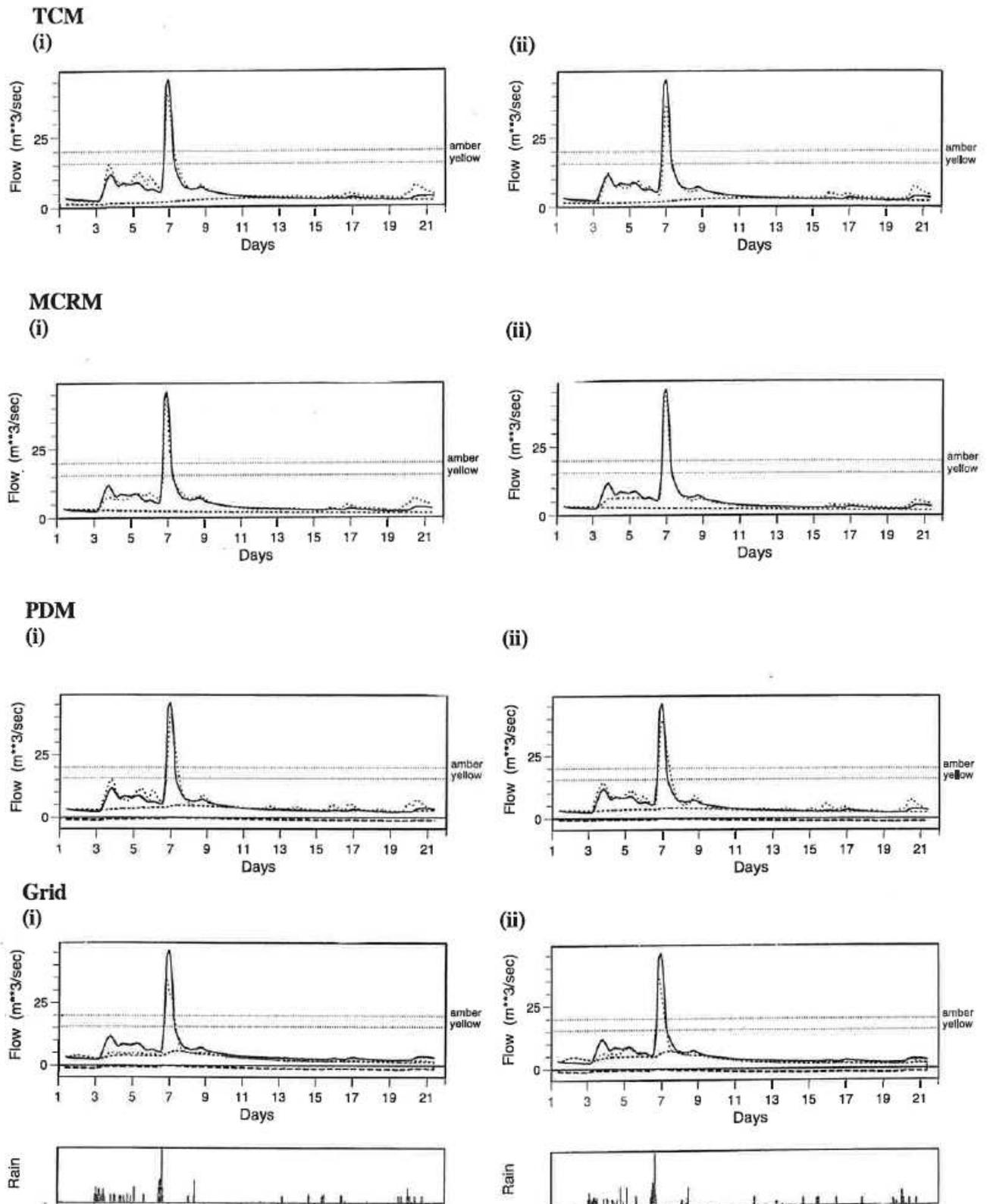


Figure 3.3.2 Simulation-mode hydrographs using (i) uncalibrated and (ii) calibrated radar data: Stour, 8 to 28 January 1993.

Calibration of the state-correction gain parameters for PDM and IEM models was carried out using the automatic optimisation procedure, and proved to be straightforward. The TF and PRTF models were recalibrated in an updating mode using recursive least squares and the smoothing factor in the time-varying model-gain was calibrated by simplex optimisation.

3.4 Overview

Eight rainfall-runoff models, with complexity ranging from conceptually-based models to simple transfer function models, have been calibrated on a total of nine catchments using both raingauge and radar rainfall data. For both types of rainfall data the models have been calibrated in simulation and updating mode.

When raingauge data are used, the simulation-mode calibrations indicate that no one model provides consistently superior results. The TCM model performed best overall, with the PDM, MCRM and NWS models close behind. As expected, the TF and PRTF models were least successful in simulation mode; transfer function models are generally only used in updating mode where observed flows are used to correct the model. The IEM was found to work particularly well on small catchments such as the Trout Beck and Silk Stream, but less well on larger catchments with a significant baseflow component.

The model calibration using radar data highlighted the beneficial effects of calibrating radar data using raingauges, particularly for the Chenies radar, for which a large number of raingauges were available for the calibration. Raingauge recalibration of the radar data resulted in a modest improvement in flow simulations for the catchments covered by the Clee Hill and Hameldon Hill radars (the Stour and the Roch). The model evaluation in Section 4 will determine whether radar or raingauge data provide the best source of rainfall data for flow-forecasting purposes.

ARMA error predictors were calibrated for all models except the PRTF and the IEM, for which state-updating schemes are used operationally. The TF and PRTF models were recalibrated in state-updating mode using recursive least squares. The resulting impulse response functions were found to be broadly similar to those obtained in the simulation-mode calibration, and resulted in improved, smoother IRFs for the Rhondda and the Stour. Calibration of the model-gain updating smoothing factors for the TF and PRTF suggested that smaller, faster catchments may require a larger smoothing factor (more constant model-gain), while catchments for which soil-moisture effects are more dominant require model gains with greater time variation.

The calibration results seem to suggest that the TCM, with its flexible model structure and greater number of parameters, is the most successful rainfall-runoff model. However, this result may reflect a degree of overfitting, and under evaluation conditions the TCM may be found to be less successful. The independent and large dataset used in evaluation should clarify the situation.

4. MODEL EVALUATION

4.1 Introduction

As part of a split-sample testing procedure, the models calibrated in the previous section are evaluated here using independent periods of data not used for model calibration. For model calibration, the focus of attention was to establish a good process model for simulation and less attention was paid to model performance in updating mode. Here, whilst the simulation mode performance is of interest, greater attention is paid to performance in forecast (updating) mode which is of direct relevance to operational flood forecasting and warning.

In simulation mode the eight models calibrated to each of the nine study catchments are evaluated using raingauge data for continuous periods of up to eight months in length. In addition, the relative benefits of using raingauge, uncalibrated radar and calibrated radar data are assessed for three of the catchments, the Silk Stream, the Stour and the Roch. Results are presented formally as tables of performance statistics, and visually as hydrographs and scatter plots of simulated against observed peak flows.

In forecast-mode the models are updated using either state-updating or ARMA error prediction, and fixed lead time forecasts are produced every 15 minutes out to 6 hours ahead. Results are presented both as tables, and graphs showing the variation in forecast accuracy with lead-time. A selection of hydrographs is used to compare model performance visually, and errors in timing and resolution of peak flows are presented as scatter plots. Results of the radar versus raingauge simulation-mode evaluations are used to establish the preferred type of rainfall data to be used for the three catchments for which radar data are available. Where use of radar data is shown to be beneficial, these data have been used in the production of additional flow forecasts, and results are compared to those obtained using raingauge data.

4.2 Results Using Raingauge Data

4.2.1 Simulation-mode results

Tables 4.2.1(a) and (b) present the pooled R^2 and CSI statistics obtained when the eight models for each of the nine catchments are run in simulation-mode over the dataset used for model evaluation. The R^2 statistic measures provides an overall measure of the accuracy of simulation, whilst the threshold CSI statistic focuses on a model's ability to successfully simulate exceedences of chosen flow thresholds (Section 2.3.3). Figure 4.2.1 summarises the results as a graph of R^2 model performance across catchments of increasing size, (the groundwater-dominated Lavant is excluded here as it will be treated as a special case and modelled using only three of the rainfall-runoff models). The use of straight lines joining R^2 values across catchments, whilst not formally appropriate, serves to highlight those catchments most difficult to model and which models are most successful. The graph reveals immediately that there is least variation in model performance for responsive or urbanised catchments such as the Rhondda, Trout Beck, Silk Stream and the Roch. For these catchments, reasonable model performance is obtained independent of the rainfall-runoff model used, particularly if the TF and PRTF are excluded. (Transfer function models are used

Table 4.2.1 Evaluation performance for different models and catchments: simulation mode using raingauge data as input

(a) R² statistic

Catchment	Model								Catchment Median
	TCM	NWS	MCRM	PDM	IEM	TF	PRTF	Grid	
Trout Beck	0.779	0.760	0.753	0.754	0.802	0.692	0.598	-	0.753
Silk Stream	0.678	0.548	0.755	0.630	0.670	0.561	0.572	-	0.630
Dove	0.849	0.835	0.870	0.816	0.605	0.283	0.231	-	0.816
Lavant	0.888	-	-	0.874	-	0.277	-	-	0.874
Rhondda	0.918	0.925	0.911	0.919	0.874	0.818	0.819	-	0.911
Brue	0.835	0.713	0.709	0.805	0.733	0.407	0.571	-	0.713
Stour	0.804	0.442	0.595	0.686	0.637	0.496	0.306	0.703	0.616
Roch	0.831	0.830	0.858	0.828	0.795	0.638	0.637	0.852	0.829
Witham	0.579	0.261	0.362	0.535	0.438	-0.615	-1.075	0.537	0.400
Model Median (Lavant excluded)	0.818	0.737	0.754	0.780	0.702	0.529	0.572	0.703	

(b) Threshold CSI statistic

Catchment	Model								Catchment Median
	TCM	NWS	MCRM	PDM	IEM	TF	PRTF	Grid	
Trout Beck	0.587	0.503	0.529	0.546	0.579	0.476	0.446	-	0.529
Silk Stream	0.578	0.402	0.512	0.517	0.506	0.479	0.404	-	0.506
Dove	0.346	0.263	0.432	0.396	0.217	0.139	0.084	-	0.263
Lavant	0.370	-	-	0.393	-	0.033	-	-	0.370
Rhondda	0.638	0.580	0.626	0.621	0.519	0.455	0.464	-	0.580
Brue	0.524	0.421	0.461	0.467	0.412	0.387	0.354	-	0.421
Stour	0.370	0.331	0.382	0.399	0.246	0.272	0.182	0.348	0.340
Roch	0.489	0.469	0.551	0.563	0.433	0.334	0.331	0.524	0.479
Witham	0.384	0.175	0.300	0.307	0.297	0.053	0.074	0.323	0.299
Model Median (Lavant excluded)	0.507	0.412	0.487	0.492	0.423	0.363	0.343	0.348	

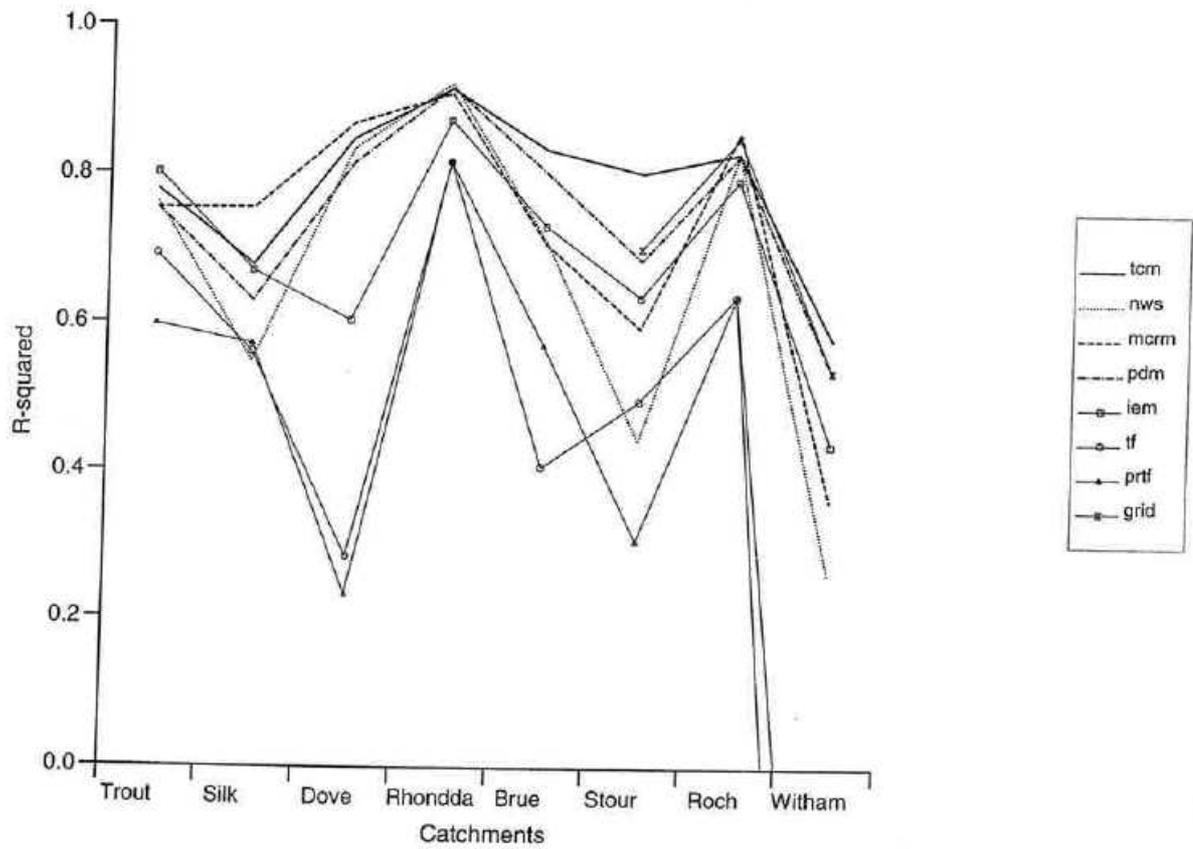


Figure 4.2.1 Simulation-mode R^2 performance across catchments: evaluation using raingauge data.

operationally in updating mode, and are not expected to perform well under simulation-mode conditions). Greater variation in model performance is apparent for catchments such as the Dove, the Brue and the Witham, where groundwater and soil moisture effects are more dominant in the hydrograph response.

Table 4.2.1, which summarises the R^2 and Threshold CSI statistics, indicates that no one model performs consistently best across catchments. The TCM performs best overall with the MCRM, PDM and NWS not far behind. The TCM has the highest CSI for five of the nine catchments, and has the highest R^2 for four of the catchments. For those catchments where another model performs better, such as the Dove and the Roch, the TCM is generally not far behind. Figure 4.2.2 shows scatter plots of observed and simulated peak flows for up to two eight-month periods for each catchment. The selection of periods for plotting is biased towards those with a large range of flows, or a period containing a flood of major significance. Flood warning levels are displayed on each plot. As expected the scatter plots tend to a 1:1 relationship between observed and simulated flow, particularly for the more responsive and urbanised catchments which are easier to model. The plots confirm that the TCM is modelling the peaks reasonably well on all catchments, with the PDM, MCRM, and Grid Model performing well on most peaks. The TF and PRTF tend to underestimate the peak flows particularly for the larger catchments such as the Roch and Witham.

Figures D.1 to D.8 in Appendix D present sets of flow hydrographs obtained from all models for typical events in each catchment. In general the sets of graphs highlight the good performance of the TCM in simulation-mode. The paragraphs that now follow focus on an extreme flood event in four of the study catchments.

The Easter 1998 flooding of the Stour at Shipston was estimated to have a return period of 40 to 80 years. The main flow peak of $91.4 \text{ m}^3 \text{ s}^{-1}$ occurred on 9 April 1998 and resulted in the flooding of 20 properties in Shipston. Figure 4.2.3(a) presents the set of flow hydrographs obtained from all models for the period 15 January 1997 to 1 May 1998, which encompasses the Easter flood. The second scatter plot of peak flows for the Stour in Figure 4.2.2 highlights the range of flows simulated for the Easter 1998 flood. Clearly the two most successful models for that flood event were the MCRM and the Grid Model, simulating flows of 95.2 and $84.6 \text{ m}^3 \text{ s}^{-1}$ respectively. The least successful models are the NWS and PRTF, predicting flows of 15.7 and $27.8 \text{ m}^3 \text{ s}^{-1}$.

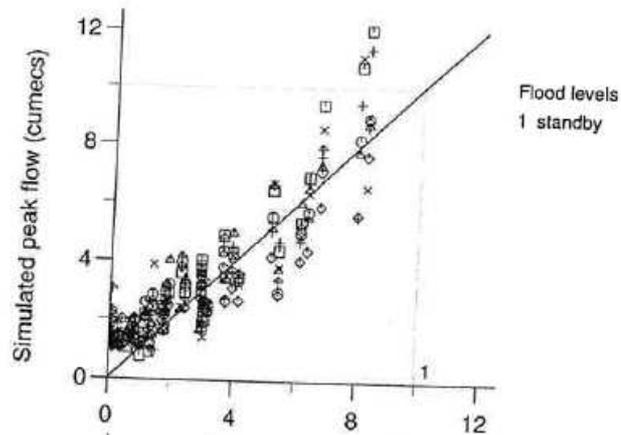
On 23 September 1992 a peak flow of $28.7 \text{ m}^3 \text{ s}^{-1}$ resulted in flooding of the Silk Stream at Colindeep Lane (Plate 2.6.3). Figure 4.2.3(b) presents the range of flow hydrographs obtained from the 7 models applied to the Silk Stream. Most models perform though a large range of peak flows is obtained for the September 1992 flood, as shown in the scatter plot of Figure 4.2.2. The models that simulated the peak flow best were the IEM and the PDM with peak flows of 21.2 and $21.6 \text{ m}^3 \text{ s}^{-1}$ respectively, and the MCRM, which overestimated the peak as $36.8 \text{ m}^3 \text{ s}^{-1}$.

The Roch at Blackford Bridge experienced a peak flow of $149.7 \text{ m}^3 \text{ s}^{-1}$ on 1 February 1995. Figure 4.2.3(c) shows the range of flow hydrographs obtained for this flood event, and the scatter plot of Figure 4.2.2 shows the range of predicted flows. The MCRM and Grid Model came closest to modelling the peak correctly, at 106.8 and $96.2 \text{ m}^3 \text{ s}^{-1}$ respectively. The threshold CSI and R^2 statistics confirm that the MCRM and Grid Model were the most successful models overall for the Roch, although most conceptual models performed well on

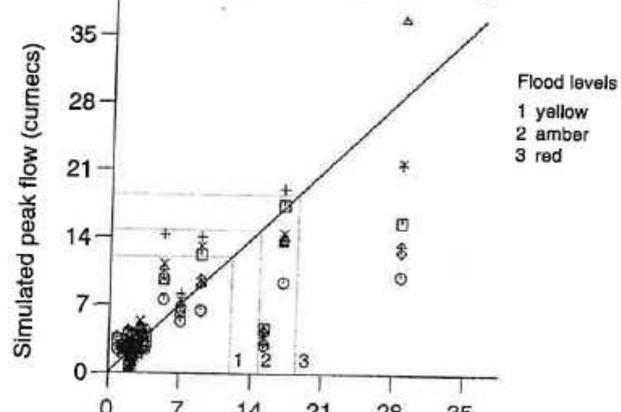
this catchment. The success of the TF and PRTF in updating mode on the Roch will be determined in Section 4.2.2.

Flooding of the River Lavant in January 1994, associated with significant flooding of the Chichester, gave rise to a peak flow of $8.1 \text{ m}^3 \text{ s}^{-1}$. Figure 4.2.3(d) presents simulated hydrographs from the three models for this catchment: the TCM, PDM and TF models. The TCM and PDM have been configured for use with groundwater abstraction data and the TF model is included in the evaluation for comparison. The TCM and PDM simulated hydrographs show that the new model formulations result in a good flow simulation of the River Lavant. The TF model was extremely difficult to calibrate for the Lavant, and the simulated hydrograph shows how poorly the TF model performs for this catchment.

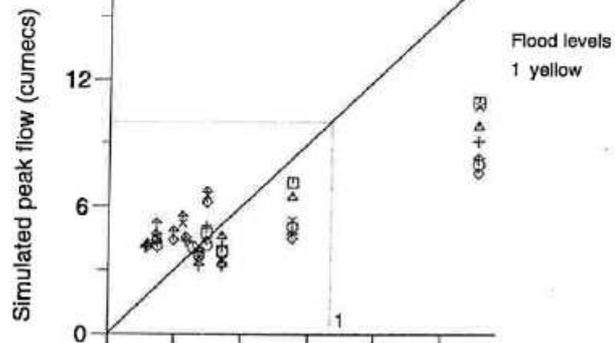
Trout Beck
15/5/96 – 19/11/96



Silk Stream
1/4 – 7/10/92



Dove
15/2/98 – 20/7/98



Rhondda
31/10/92 – 30/4/93

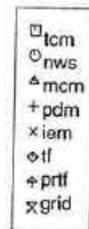
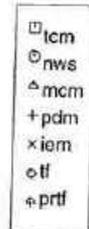
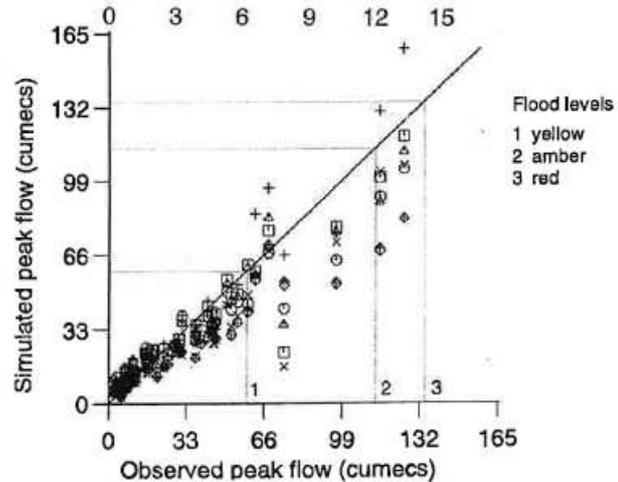
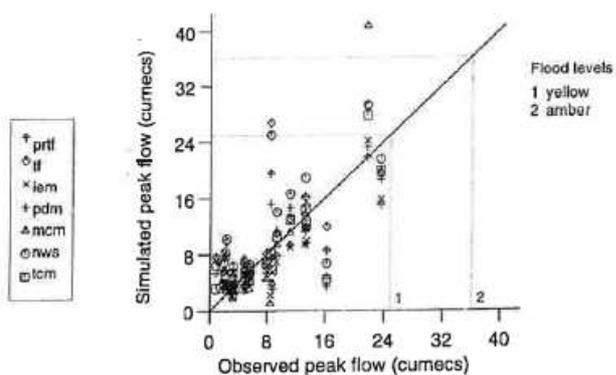
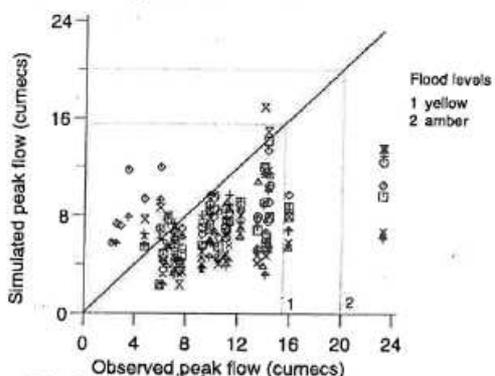


Figure 4.2.2 Scatter plots of observed and simulated peak flows: evaluation using raingauge data.

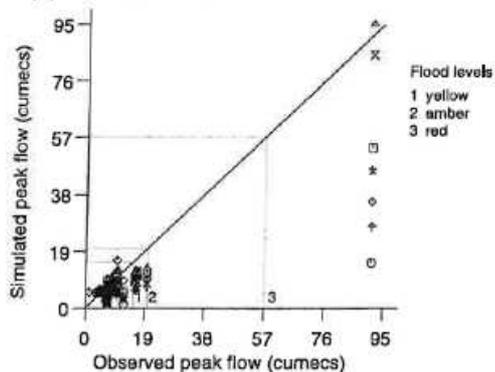
Brue 1/11/96 – 1/4/97



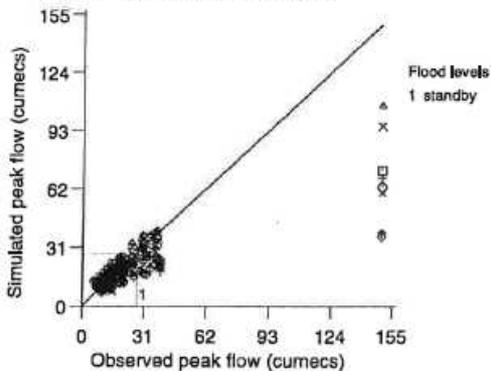
Stour (i) 28/9/93 – 15/3/94



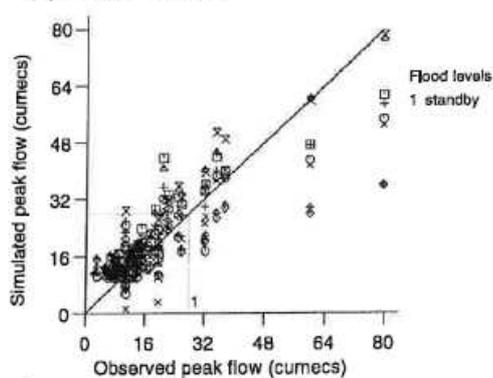
(ii) 15/11/97 – 1/5/98



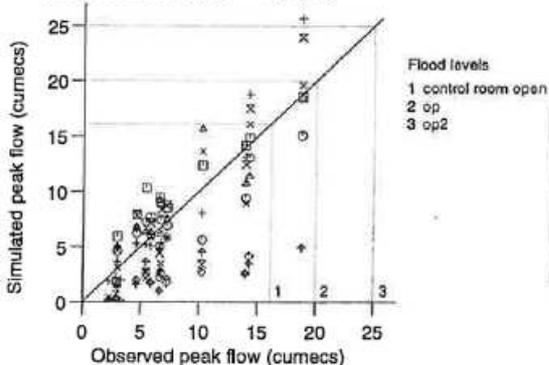
Roch (i) 30/11/94 – 25/4/95



(ii) 1/5/97 – 20/3/98



Witham 15/11/97 – 15/3/98



Lavant 1/6/94 – 1/5/96

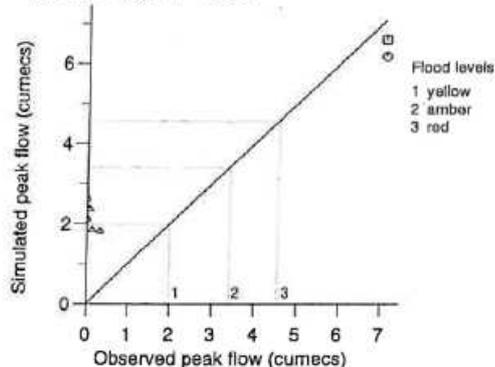


Figure 4.2.2 (cont...) Scatter plots of observed and simulated peak flows: evaluation using raingauge data.

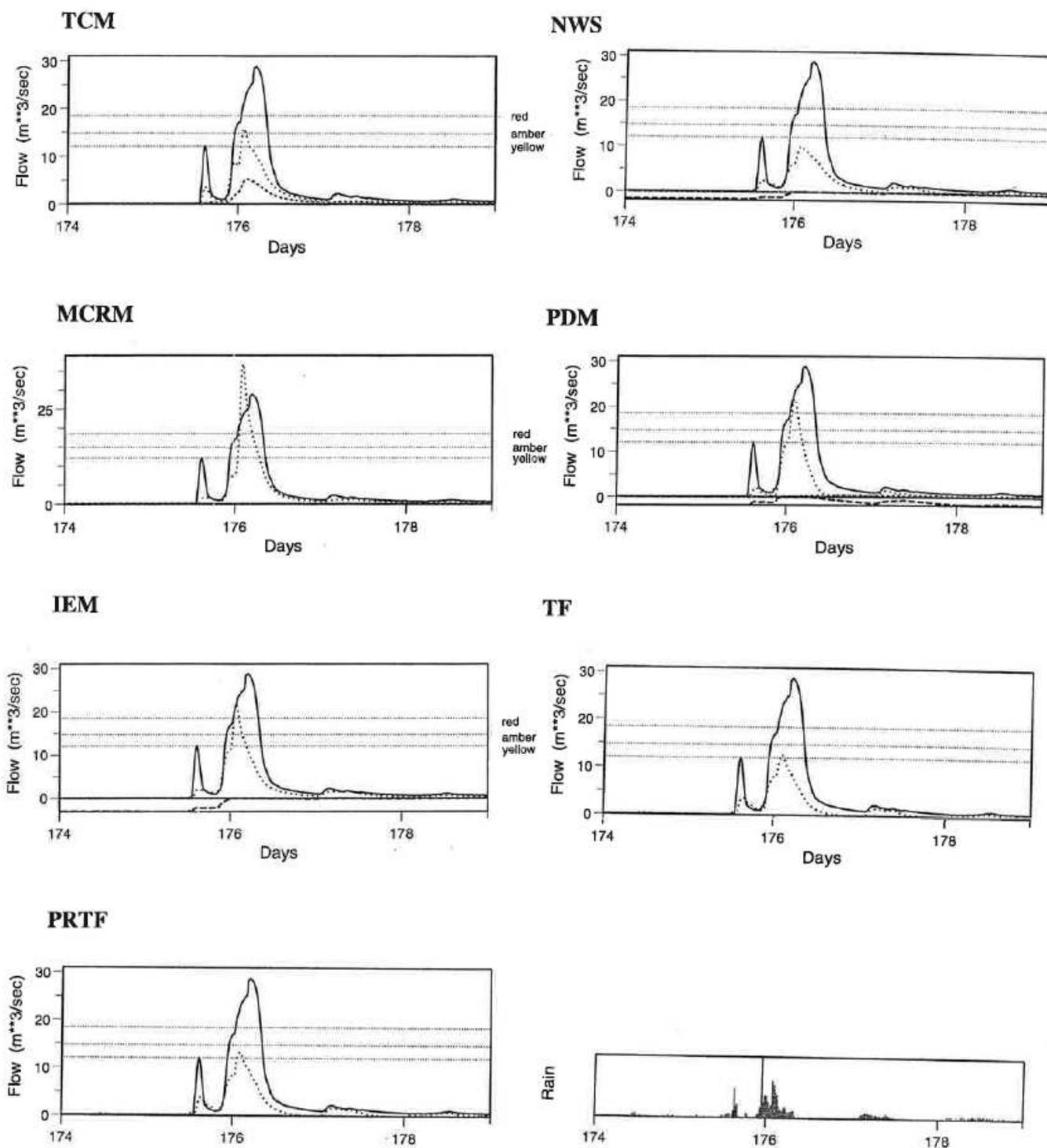


Figure 4.2.3(a) Simulated flow hydrographs from different models: Silk Stream, 21–25 September 1992. Observed flow: bold line; simulated flow: dotted line; simulated baseflow: dashed line (TCM, PDM only); soil moisture deficit: long-dashed line (NWS, PDM, IEM only).

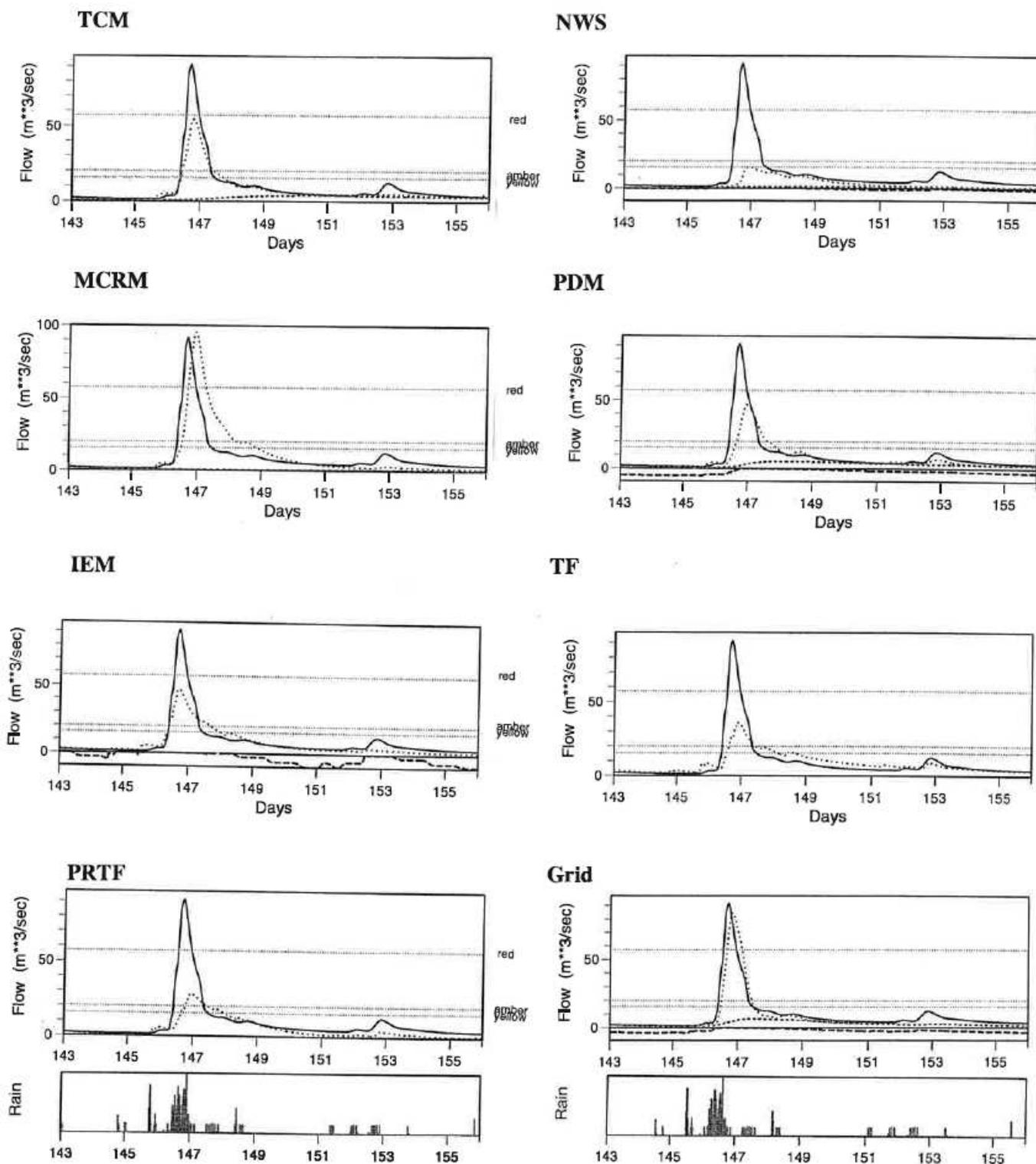


Figure 4.2.3(b) Simulated flow hydrographs from different models: The Stour, 6–18 April 1998 (Easter 1998 flood). Observed flow: bold line; simulated flow: dotted line; simulated baseflow: dashed line (TCM, PDM only); soil moisture deficit: long-dashed line (NWS, PDM, IEM only).

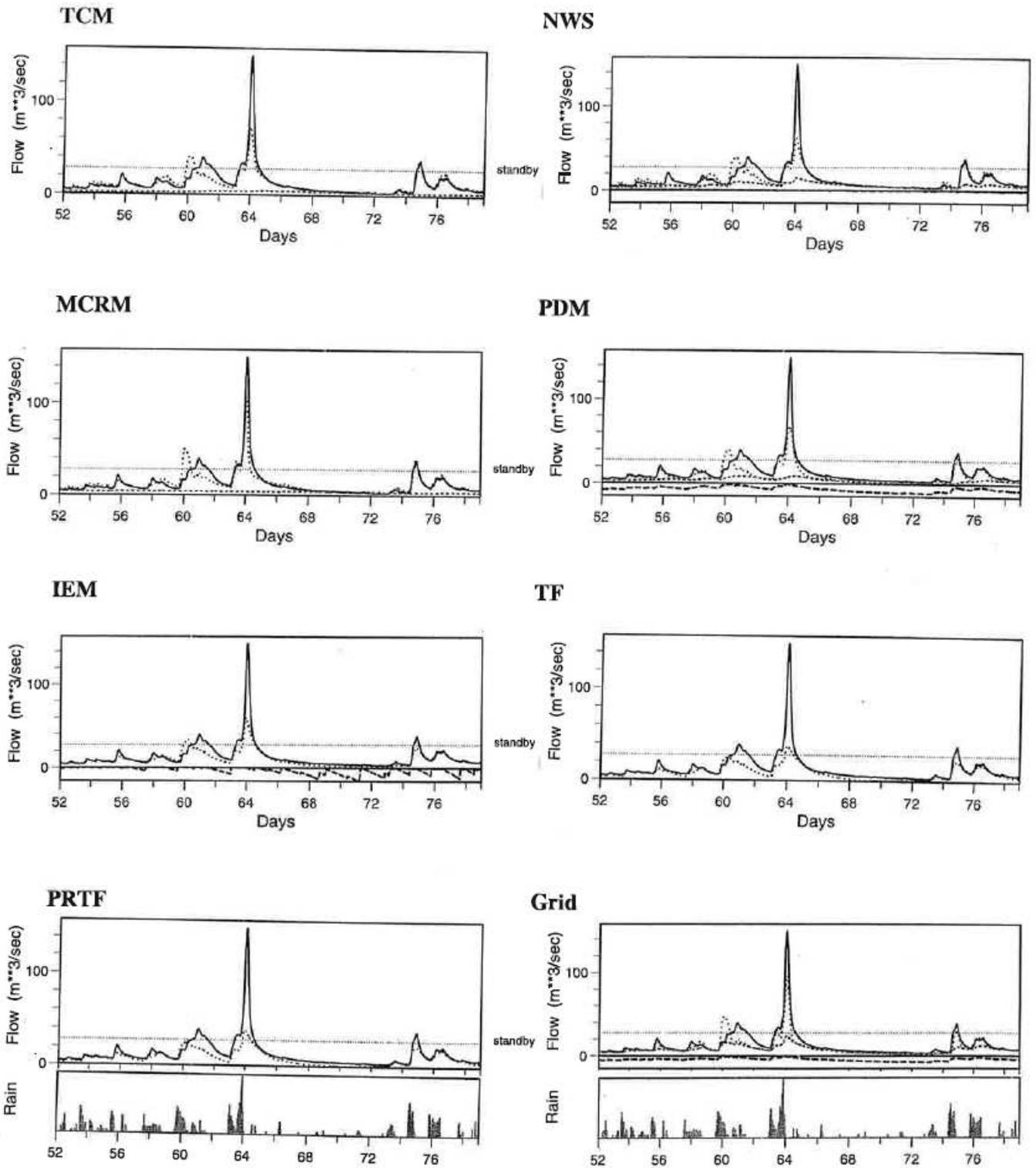
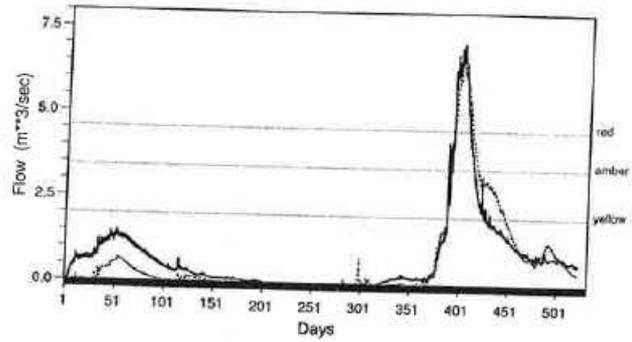
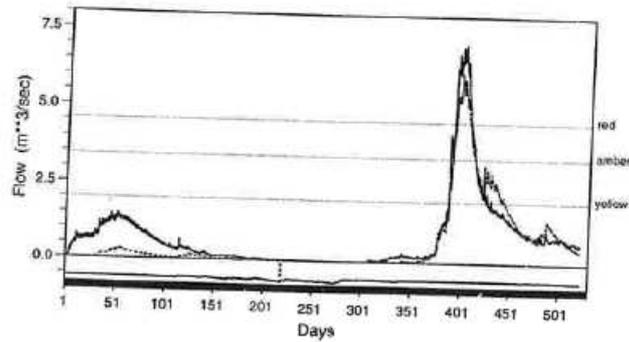


Figure 4.2.3(c) Simulated flow hydrographs from different models: The Roch, 20 January – 15 February 1995. Observed flow: bold line; simulated flow: dotted line; simulated baseflow: dashed line (TCM, PDM only); soil moisture deficit: long-dashed line (NWS, PDM, IEM only).

TCM



PDM



TF

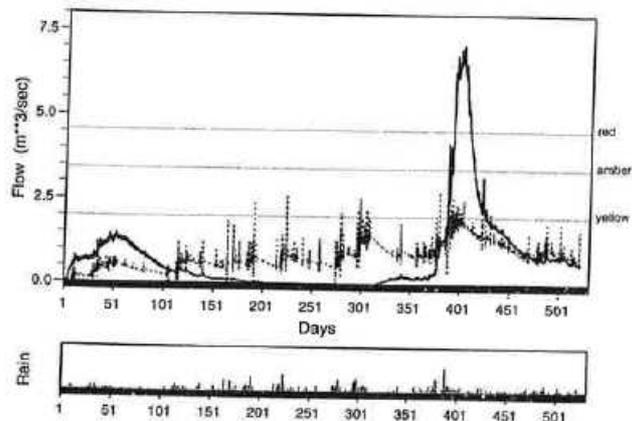


Figure 4.2.3(d) Simulated flow hydrographs from different models: The Lavant, 8 December 1992 – 1 June 1994 (Chichester flood). Observed flow: bold line; simulated flow: dotted line; simulated baseflow: dashed line (TCM, PDM only); soil moisture deficit: long-dashed line (PDM only).

4.2.2 Forecast-mode results

The rainfall-runoff models were evaluated in forecast-mode by comparing fixed lead-time flow forecasts for a number of lead times. On eight of the nine catchments, forecasts were produced for lead times of 0.25, 0.5, 0.75, 1.0, ..., 5.75 and 6.0 hours ahead. For the slow-responding groundwater-dominated Lavant, forecasts were produced for lead times of 0.25, 0.5, 0.75, 1.0, ..., 5.75 and 6.0 days ahead. Results are presented as tables of performance statistics and sets of graphs highlighting different aspects of the forecasting process, which allow model intercomparison across a range of criteria. A selection of fixed origin forecasts superimposed on observed hydrographs allows a visual comparison of model performance.

Table 4.2.2 summarises the performance statistics obtained over the evaluation periods for the different models applied across catchments for a particular lead-time and provides a general overview of model performance. Figure 4.2.4 illustrates how models performance varies with forecast lead time and Figure 4.2.5 compares observed and forecast peak flow for a significant event in each catchment and compares peak timing error. In conjunction with Figure 4.2.5, Figure 4.2.8 shows how peak forecast accuracy varies with forecast lead time for all models, and reveals that although a particular model may perform best at one lead time, it may perform less well at other lead times. Illustrative fixed lead-time forecasts for the Silk Stream and the Stour are shown in Figure 4.2.6. Figure 4.2.7 presents a set of fixed origin forecasts for all models for significant flood events in the Silk Stream and the Stour. Figures D.9 to D.16 in Appendix D present typical sets of fixed lead-time forecasts for all models in each catchment. The figures suggest that there is often little to choose between models in updating-mode.

The following provides an overview of the performance of each model in turn.

Thames Catchment Model (TCM)

Table 4.2.2 presents the R^2 and threshold CSI statistics for a representative lead-time for each model and catchment. The table reveals that the TCM is one of the best performing models in forecasting mode using the R^2 statistic, although the performance is only average when the threshold CSI criterion is used. These results are confirmed by Figure 4.2.4, which shows the *variation* in R^2 and threshold CSI with lead time. The poorer performance in terms of threshold CSI indicates that whilst the TCM performs well, it is less well able to forecast the flow peaks than other models such as the PDM. Figure 4.2.5 confirms that although the TCM is reasonably good in terms of forecast peak flow and timing, other models (such as the PDM and IEM) are as good or better. Figure 4.3.9, which shows how peak forecast accuracy with lead time, indicates that the TCM performs well at all lead times on the Trout Beck, the Dove, Rhondda, Brue and Roch, but less well on the Silk Stream, compared to model such as the IEM and PDM. The fixed-origin plots of Figure 4.2.7 and 4.2.8 for the Silk Stream and the Stour indicate the TCM forecasting reasonably well, particularly for the Stour.

National Weather Service (NWS) Model

Table 4.2.2 and Figures 4.2.4 and 4.2.5 show that the NWS is a sound conceptual rainfall-runoff model, with an overall performance on a par with the IEM and MCRM. The NWS is generally poorer than the TCM in terms of R^2 performance and poorer than the PDM and IEM in terms of threshold CSI. Figure 4.3.9 indicates that the NWS forecasts peak flows at a range

of lead times more successfully on larger catchments such as the Stour and Witham, but is less accurate on smaller, faster catchments such as Trout Beck and the Silk Stream. Figure 4.2.7 confirms that compared to other models, the NWS underestimates the peak flow in the Silk Stream in September 1992.

Midlands Conceptual Runoff Model (MCRM)

Overall, the MCRM gives a comparable performance to the IEM and NWS. The MCRM performs the best of all models on the Silk Stream and the Dove at longer lead times, although the threshold CSI results for larger catchments such as the Roch and Witham are poor. The scatter plot of observed against forecast peak flow for the Easter flood on the Stour shows that at a 3 hour lead time the MCRM overestimates the peak flow of $92 \text{ m}^3 \text{ s}^{-1}$ by as much as $50 \text{ m}^3 \text{ s}^{-1}$, and the overestimation increases with forecast lead time (Figure 4.3.9). The fixed-origin hydrographs of Figure 4.2.8 illustrate this overestimation in the Stour, and Figure 4.2.9 show that the MCRM also overestimated the September 1992 flow in the Silk Stream.

Probability Distributed Moisture (PDM) Model

The PDM performs well in forecast mode for all catchments; it often gives the best performance in terms of threshold CSI, showing that it is a good model for forecasting peaks, although the TCM generally gives a better overall performance as measured by the R^2 statistic. At all but the shortest lead times the threshold CSI performance of the PDM is particularly good for the Rhondda, the Stour and the Roch. At a 3 hour lead time the PDM forecasts the Easter 1998 flood in the Stour most accurately, although it underestimates the February 1995 flood in the Roch by approximately $20 \text{ m}^3 \text{ s}^{-1}$. Figure 4.3.9 indicates that the PDM forecasts the Easter flood well at a range of lead times in the Stour, and gave good performance on the Silk Stream, Lavant and Witham, but forecasted the example peaks for the Rhondda, Roch and Brue with less accuracy. The illustrative fixed-origin plots of Figure 4.2.7 for the Silk Stream show the PDM and IEM forecasting the September 1992 flow peak with greatest accuracy.

Isolated Event Model (IEM)

The forecast performance of the IEM on larger catchments is surprisingly good for a single non-linear storage model with a soil-moisture dependent loss factor. The model gives the best performance in terms of threshold CSI for all lead times on the Witham and the Brue, and also performs well on the Roch and Rhondda. In terms of the R^2 criterion, the IEM is the best model for the quickly responding Trout Beck catchment. The scatter plots of Figure 4.2.5 suggest that the IEM has a tendency to underestimate the higher flow peaks, although the plots of Figure 4.3.9 of forecast peak flow forecast accuracy with lead time suggest that the IEM is forecasting the significant peaks reasonably well for catchments other than the Roch. The fixed-origin hydrographs of Figures 4.2.7 and 4.2.8 show the IEM providing good peak flow forecasts for significant flood events in the Stour and Silk Stream.

Transfer Function Models (TF, PRTF)

The TF and PRTF models give a comparable, though slightly poorer performance to the more complex, conceptually based rainfall-runoff models in terms of R^2 for larger catchments such as the Witham, the Roch and the Stour. However, in terms of threshold CSI, which measures

the models' ability to correctly forecast flow exceeding a number of thresholds, the TF and PRTF models are very poor. The scatter plots of observed and forecast peak flow in Figure 4.2.5 confirm that the TF and, in particular, the PRTF models tend to overestimate the flow. This is confirmed by Figure 4.2.9 which shows overestimation in the PRTF peak flow forecasts increasing with lead time for the Dove, Brue, Witham and Stour. Figure 4.2.7 indicates that use of the TF and PRTF for fixed origin forecasts in the Silk Stream results in a reasonably good set of peak flow hydrographs, but on the Stour (Figure 4.2.8) use of the TF and PRTF for forecasts leads to overestimation of the peak by a factor of two to three. When used operationally, the PRTF is updated using a manual-updating scheme, which may improve upon the model performance indicated here.

The results discussed above were obtained using state updating (full state-correction, with model-gain updating) in preference to error prediction to more closely emulate EA operational practice. Results obtained by using error prediction are summarised in Table 4.2.3. These results suggest that for some catchments forecast accuracy can be improved using a full TFN (Transfer Function Noise) model, incorporating ARMA error prediction, instead of state updating. This is particularly the case for smaller catchments whilst less clear-cut for larger catchments with state updating proving sometimes beneficial for the Stour and Rhondda. Figure 4.2.9 shows examples, for four of the catchments, of the variation in forecast performance obtained using the two methods of updating. Figure 4.2.10 shows for the four smallest catchments, where error prediction gave the greatest improvement, how R^2 and Threshold CSI statistics for different lead times compare with those from the other models. Relating these with those previously shown in Figure 4.2.4, it is clear that error prediction used with TF models provides performance comparable with other models. Notably, the poor Threshold CSI performance has been substantially corrected.

Grid Model

The Grid Model is a distributed rainfall-runoff model designed for use with larger catchments which might be expected to exhibit a distributed response to spatially distributed rainfall data. The Grid Model was evaluated on the three largest of the eight study catchments, the Stour Roch and Witham. Table 4.2.2, which shows R^2 and threshold CSI forecast statistics for all models at a particular lead time, indicates that the Grid Model is the *second best* model in terms of R^2 on all three catchments (the TCM or MCRM perform better), and gives an average model performance in terms of the threshold CSI. Scatter plots of observed and forecast peak flow indicate that the Grid Model forecasts peak flow well on the Roch and Witham, but can overestimate peaks on the Stour. Figure 4.3.9, which plots forecast peak flow against lead time, confirms that the Grid Model performs reasonably well across a range of lead times on the Witham and the Roch. However, Figure 4.3.9 indicates that the Grid Model overestimates the Easter flood on the Stour, and this is illustrated in the set of fixed lead-time hydrographs of Figure 4.2.6, which shows a set of peaky fixed lead-time forecast hydrographs ranging from one to six hours at hourly intervals (the one hour ahead forecast is closest to the observed flow).

Table 4.2.2 Evaluation performance for different models and catchments: forecast mode using raingauge data as input

(a) R² statistic

Catchment	Model									Catchment Median
	TCM	NWS	MCRM	PDM	IEM	TF	PRTF	Grid	Lead time (hrs)	
Trout Beck	0.962	0.945	0.947	0.959	0.961	0.897	0.889	-	1	0.947
Silk Stream	0.910	0.904	0.908	0.910	0.903	0.895	0.820	-	1	0.904
Dove	0.988	0.988	0.990	0.982	0.978	0.976	0.945	-	2	0.982
Lavant	0.978	-	-	0.977	-	0.925	-	-	96	0.977
Rhondda	0.987	0.980	0.974	0.976	0.974	0.973	0.962	-	2	0.974
Brue	0.978	0.950	0.970	0.968	0.961	0.950	0.950	-	2	0.961
Stour	0.971	0.950	0.965	0.943	0.953	0.958	0.953	0.950	3	0.950
Roch	0.963	0.951	0.961	0.954	0.939	0.923	0.901	0.957	3	0.953
Witham	0.976	0.958	0.957	0.972	0.966	0.957	0.919	0.930	4	0.958
Model Median (Lavant excluded)	0.974	0.951	0.963	0.964	0.961	0.954	0.932	0.950		

(b) Threshold CSI statistic

Catchment	Model									Catchment Median
	TCM	NWS	MCRM	PDM	IEM	TF	PRTF	Grid	Lead time (hrs)	
Trout Beck	0.741	0.640	0.634	0.624	0.631	0.247	0.239	-	1	0.631
Silk Stream	0.602	0.628	0.710	0.609	0.643	0.472	0.296	-	1	0.609
Dove	0.468	0.465	0.496	0.654	0.493	0.358	0.054	-	2	0.468
Lavant	0.242	-	-	0.219	-	0.293	-	-	96	0.242
Rhondda	0.658	0.624	0.554	0.761	0.667	0.327	0.205	-	2	0.624
Brue	0.304	0.250	0.361	0.560	0.644	0.180	0.181	-	2	0.304
Stour	0.460	0.347	0.525	0.706	0.546	0.261	0.234	0.296	3	0.404
Roch	0.453	0.447	0.339	0.607	0.557	0.178	0.129	0.367	3	0.407
Witham	0.395	0.290	0.185	0.593	0.685	0.291	0.034	0.290	4	0.290
Model Median (Lavant excluded)	0.464	0.456	0.511	0.617	0.637	0.276	0.193	0.296		

Table 4.2.3 Best TF model updating procedure for each catchment as judged using the R^2 statistic and Threshold CSI criterion. State Updating: SU; Error prediction: ARMA

Catchment	R^2	Best updating type	Threshold CSI
Trout Beck	ARMA		ARMA
Silk Stream	ARMA		ARMA
Dove	ARMA		ARMA
Lavant	-		-
Rhondda	SU (little difference between SU & ARMA)		ARMA
Brue	< 3.5 hr ahead, SU > 3.5 hr ahead, ARMA		< 1.5 hr ahead, SU > 1.5 hr ahead, ARMA
Stour	SU (little difference between SU & ARMA)		< 2 hr ahead, SU > 2 hr ahead, ARMA
Roch	No difference ARMA/SU		SU
Witham	ARMA		< 2.5 hr ahead, ARMA > 2.5 hr ahead, SU

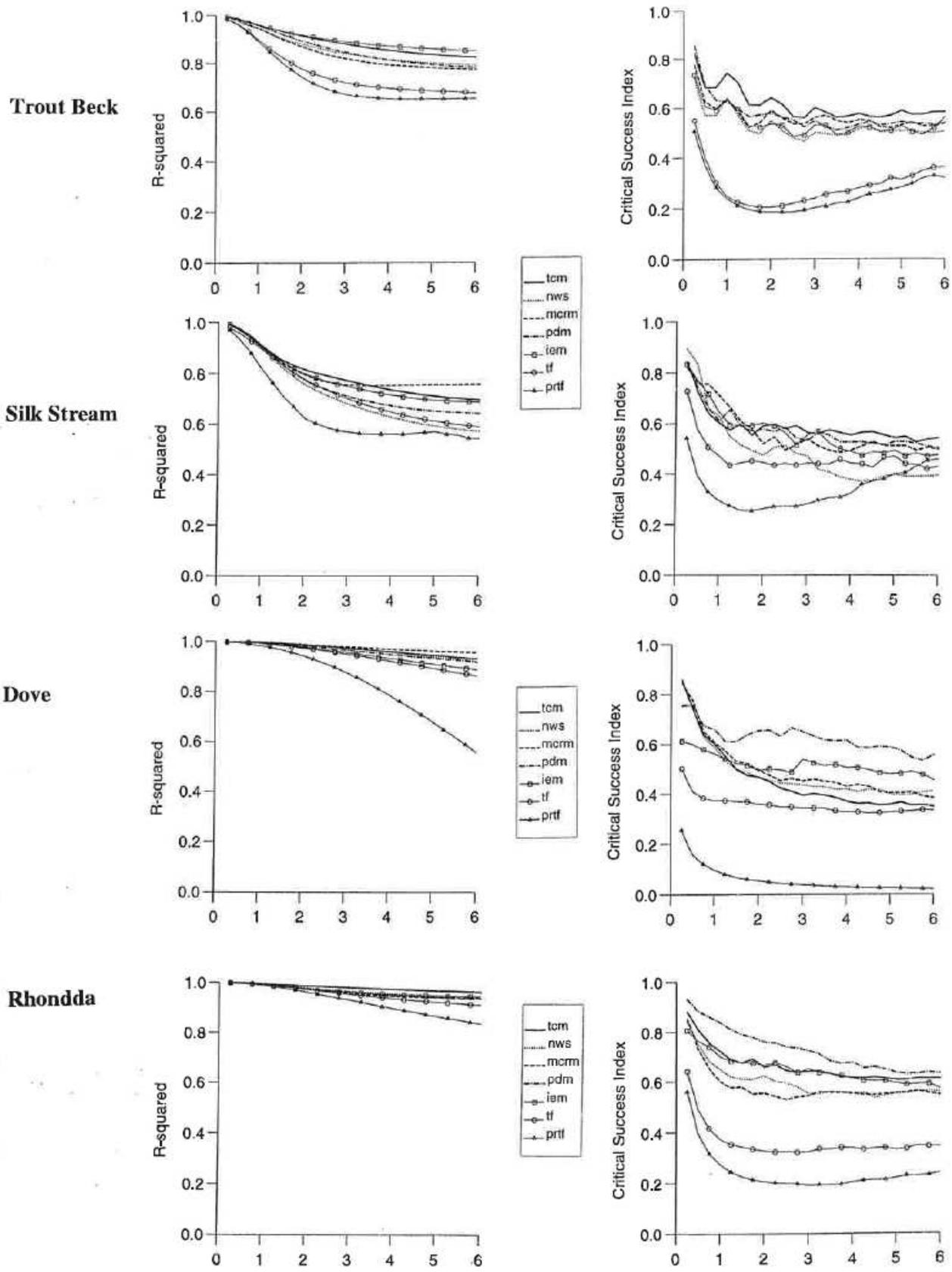


Figure 4.2.4 Variation in forecast performance (R^2 and Threshold CSI) with forecast lead time for different models.

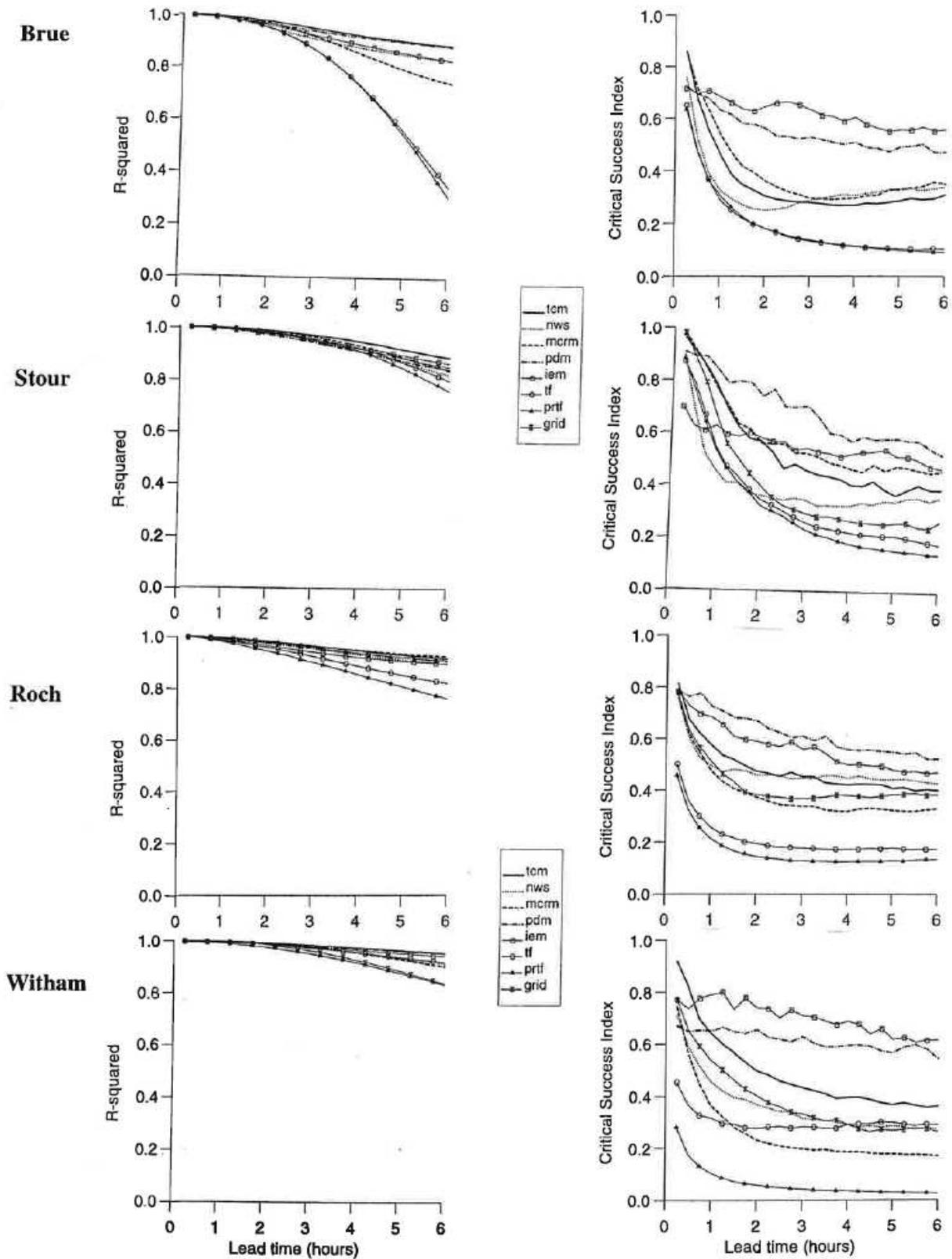


Figure 4.2.4 (cont...) Variation in forecast performance (R^2 and Threshold CSI) with forecast lead time for different models.

Lavant

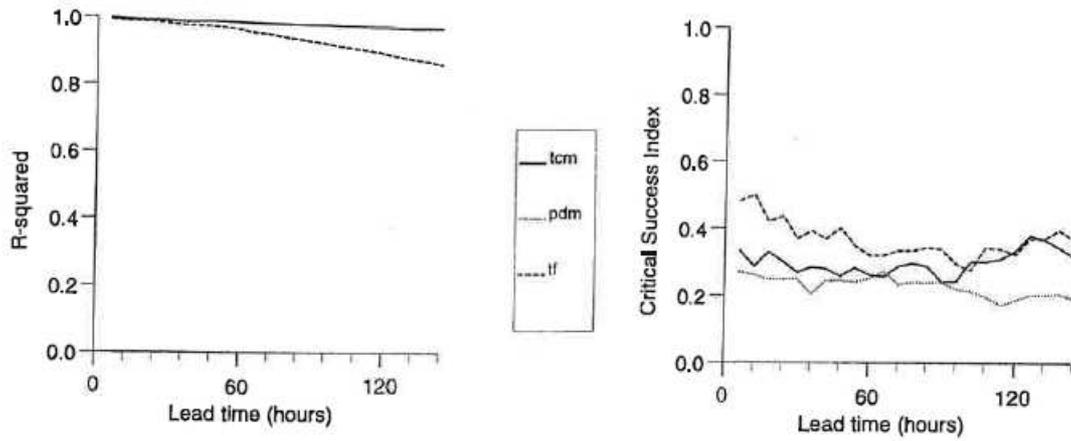


Figure 4.2.4 (cont...) Variation in forecast performance (R^2 and Threshold CSI) with forecast lead time for different models.

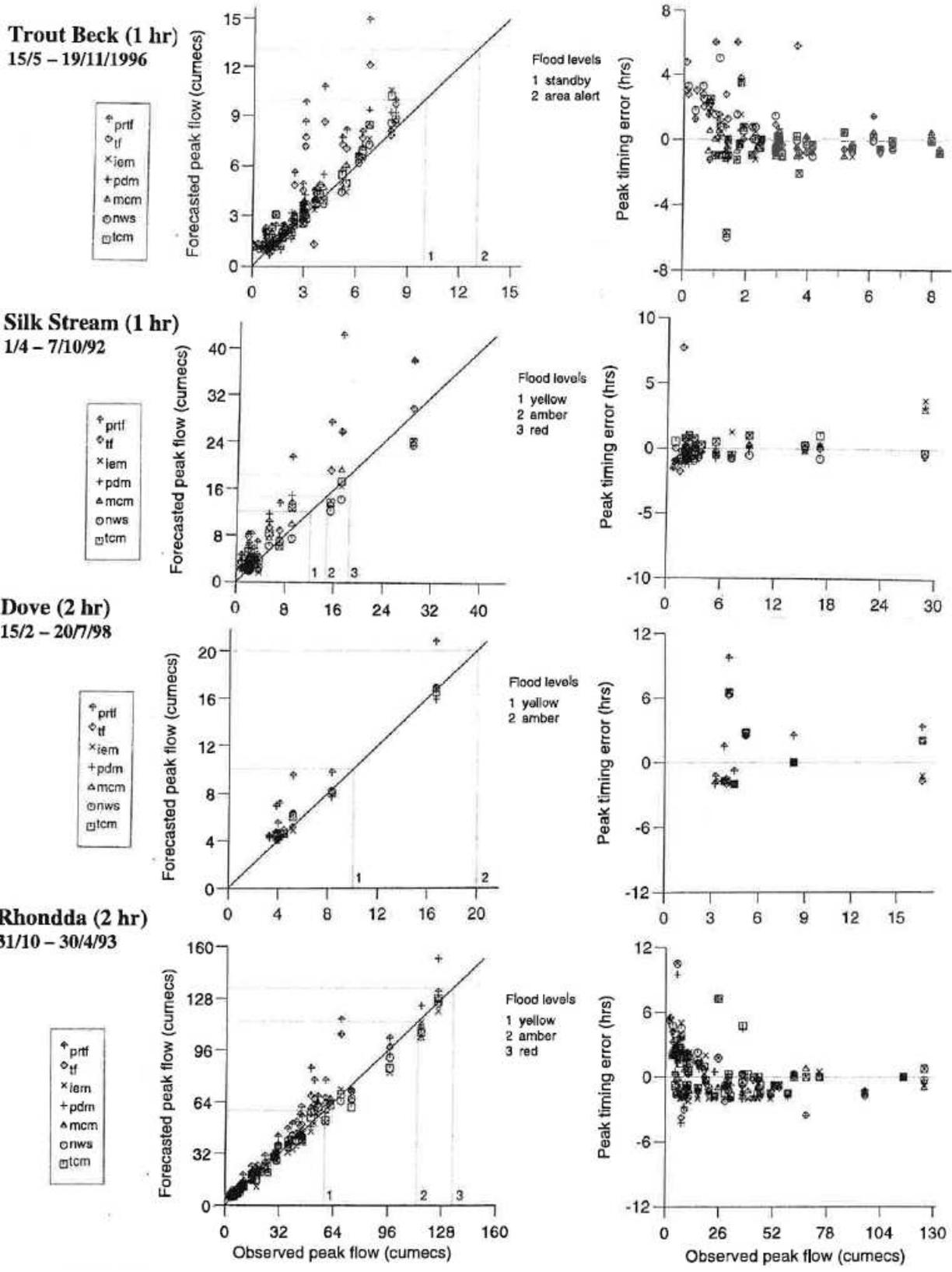
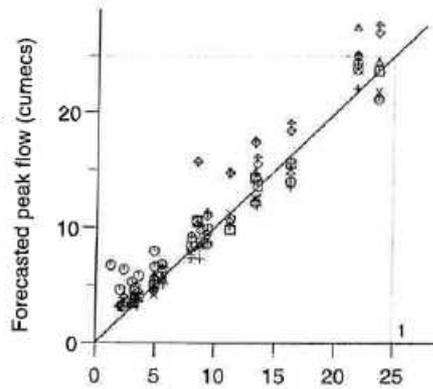


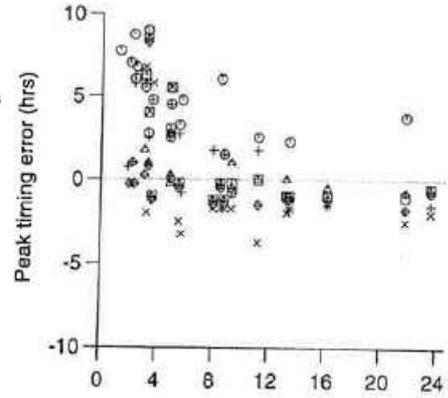
Figure 4.2.5 Scatter plots of observed and peak flow, and peak timing error for different models at a specified forecast lead time.

Brue (2 hr)
1/11/96 – 1/4/93

- ♦ prtf
- tf
- × iem
- + pdm
- △ mcm
- nws
- tcm

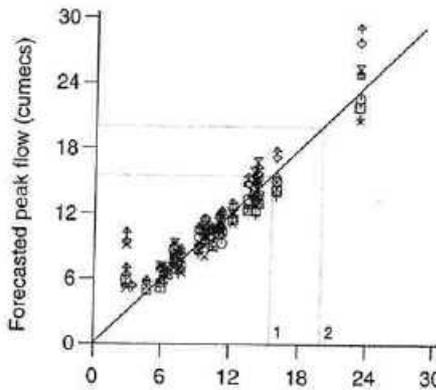


Flood levels
1 yellow

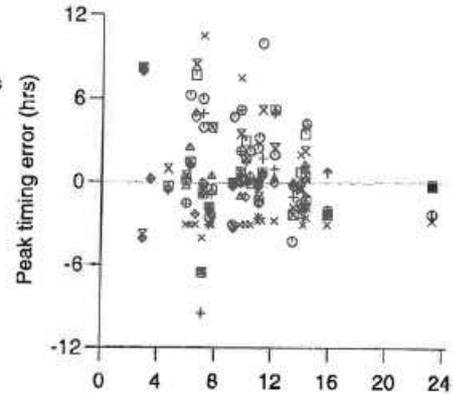


Stour (3 hr)
28/9/93 – 15/3/94

- × grid
- ♦ prtf
- tf
- × iem
- + pdm
- △ mcm
- nws
- tcm

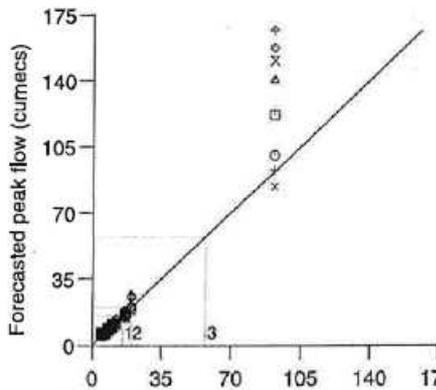


Flood levels
1 yellow
2 amber

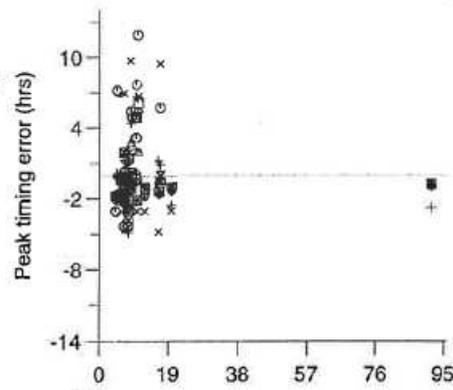


Stour (3 hr)
15/11/97 – 1/5/98

- × grid
- ♦ prtf
- tf
- × iem
- + pdm
- △ mcm
- nws
- tcm

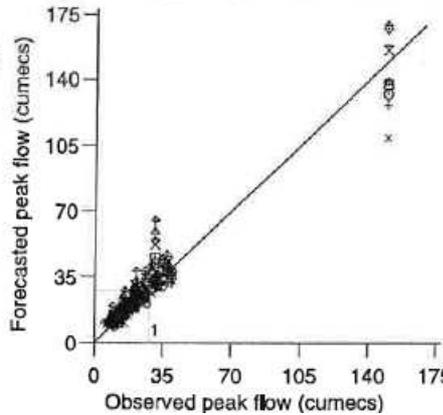


Flood levels
1 yellow
2 amber
3 red



Roch (3 hr)
30/11/94 – 25/4/95

- × grid
- ♦ prtf
- tf
- × iem
- + pdm
- △ mcm
- nws
- tcm



Flood levels
1 standby

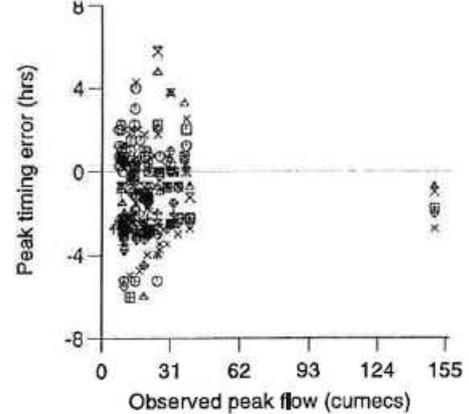
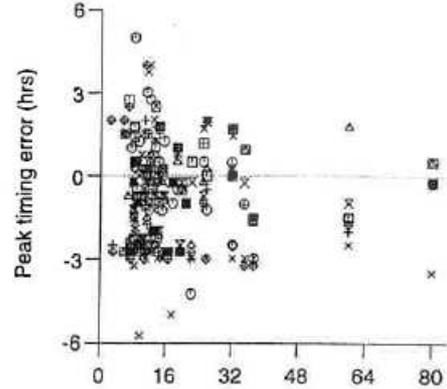
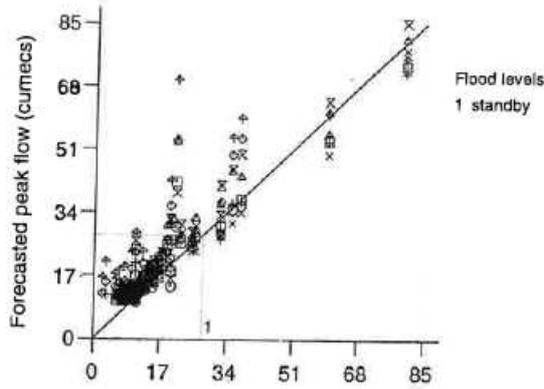


Figure 4.2.5 (cont...) Scatter plots of observed and peak flow, and peak timing error for different models at a specified forecast lead time.

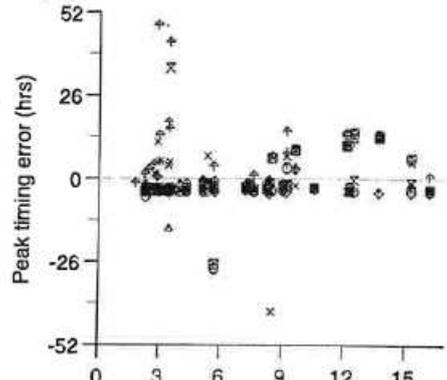
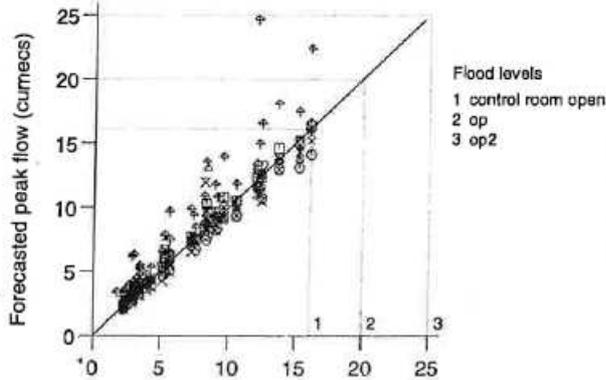
Roch (3 hr)
1/5/97 – 28/5/98

- x grid
- + prtfl
- o tf
- x iem
- + pdm
- △ mcm
- o nws
- tcm



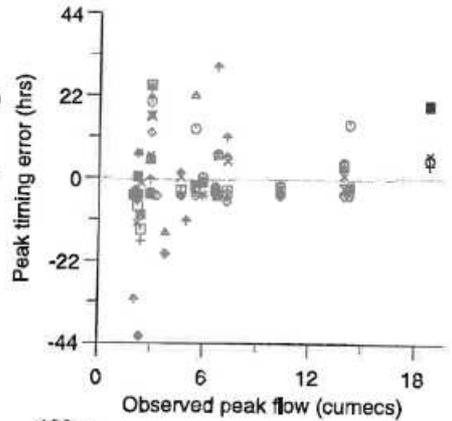
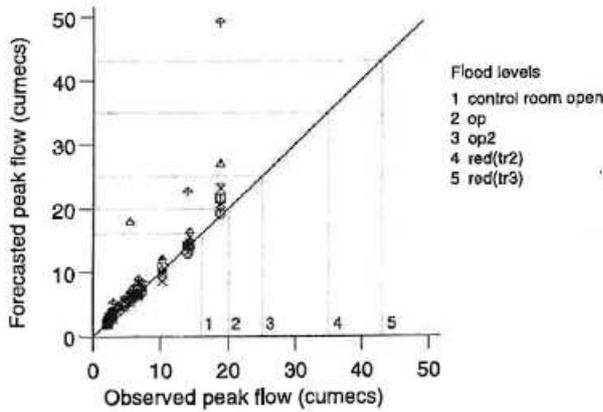
Witham (4 hr)
1/9/93 – 1/5/94

- x grid
- + prtfl
- o tf
- x iem
- + pdm
- △ mcm
- o nws
- tcm



Witham (4 hr)
15/11/97 – 15/3/98

- x grid
- + prtfl
- o tf
- x iem
- + pdm
- △ mcm
- o nws
- tcm



Lavant (96 hr)
8/12/92 – 1/6/94

- △ tf
- o pdm
- tcm

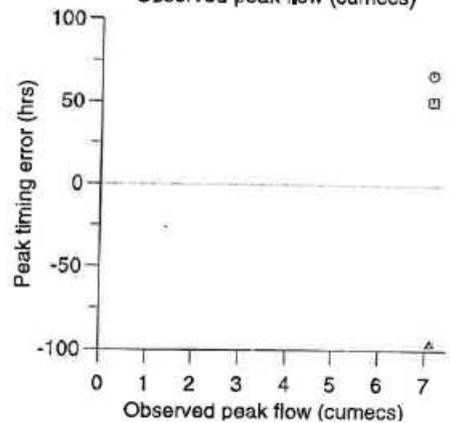
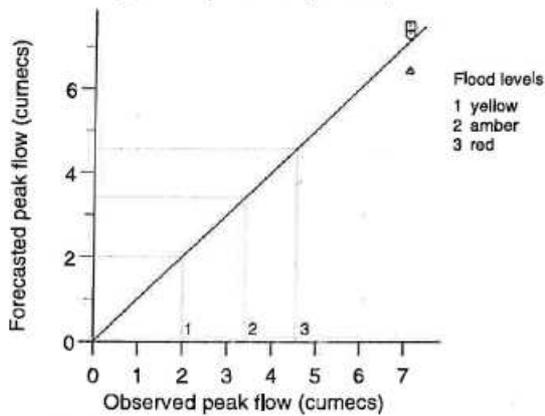
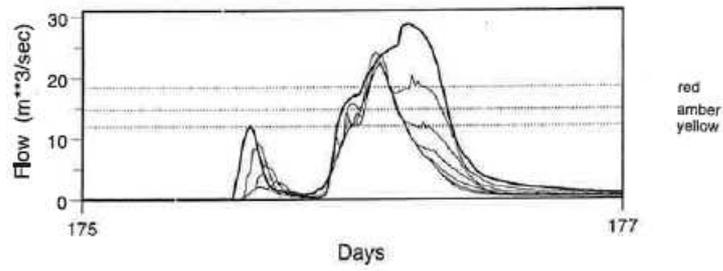


Figure 4.2.5 (cont...) Scatter plots of observed and peak flow, and peak timing error for different models at a specified forecast lead time.

(a) Silk Stream (PDM): 22 – 23 September 1992



(b) The Stour (Grid Model): 9 – 10 April 1998

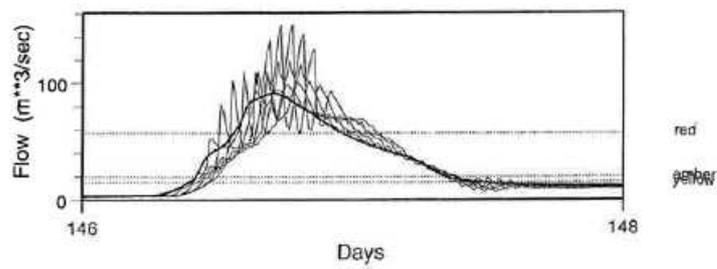


Figure 4.2.6 Examples of fixed lead time evaluation forecasts.

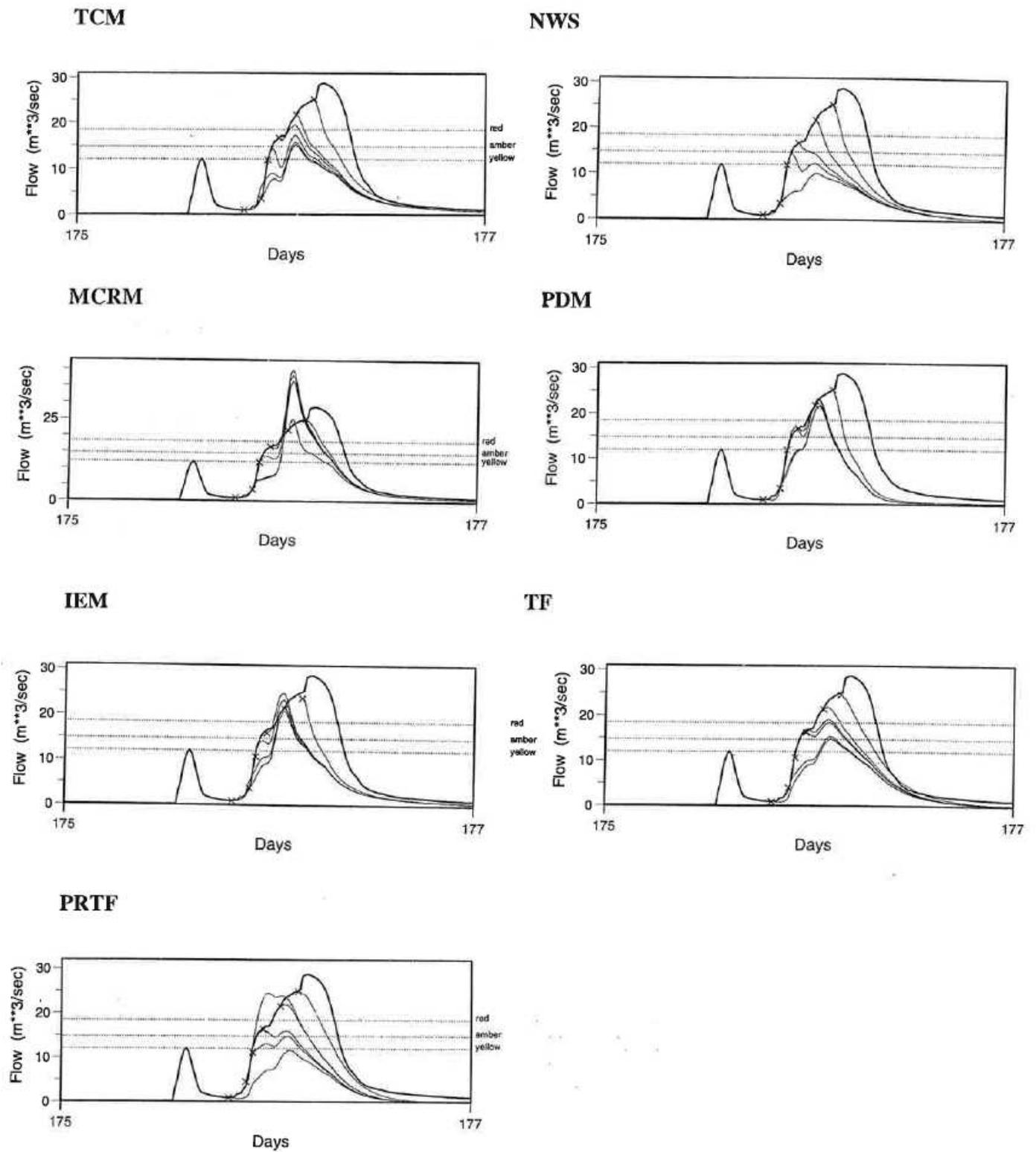


Figure 4.2.7 Evaluation mode fixed-origin forecasts (a) Silk Stream: 22–23 September 1992. (Observed flow: bold line, forecast flow: fine line)

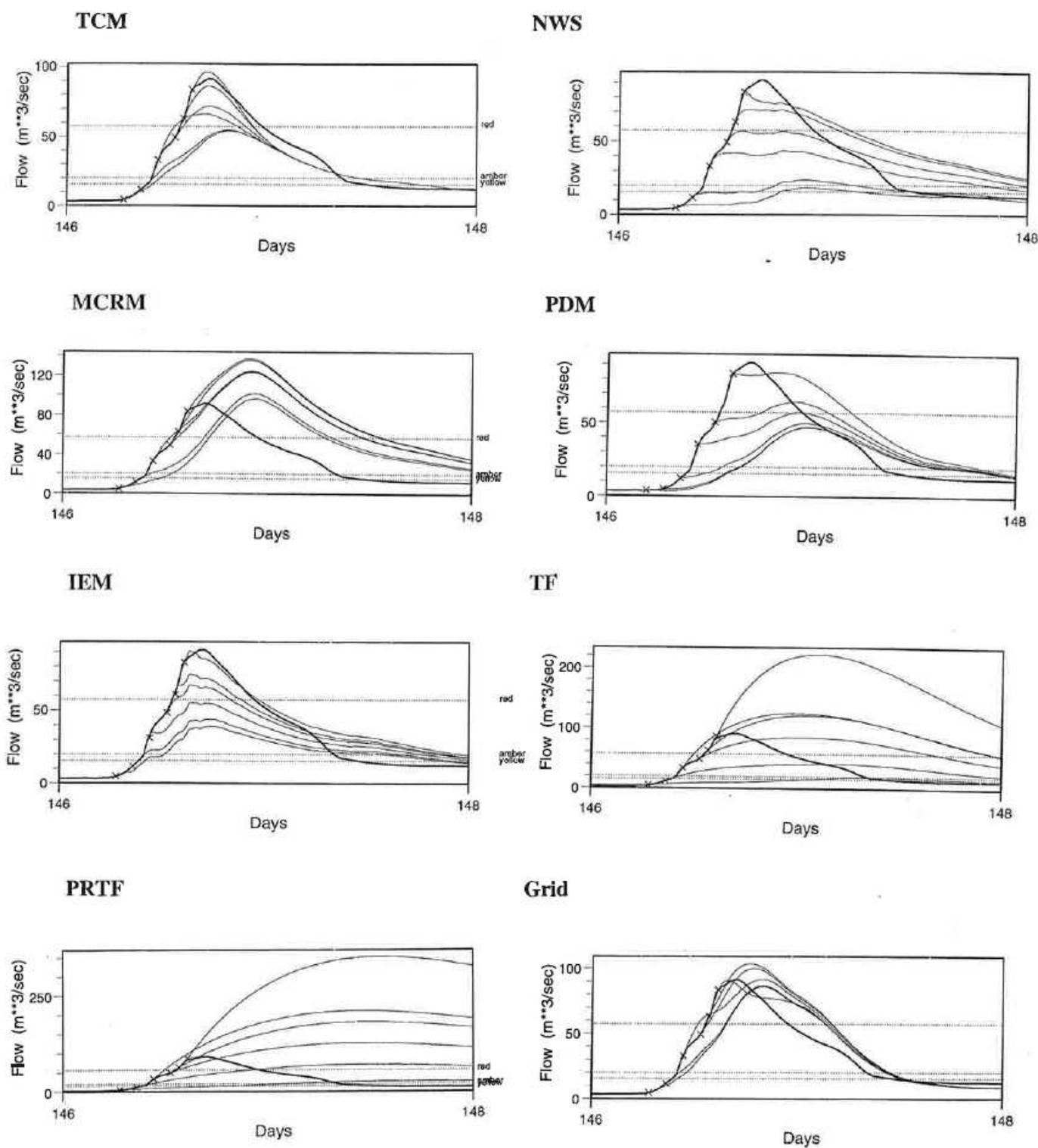
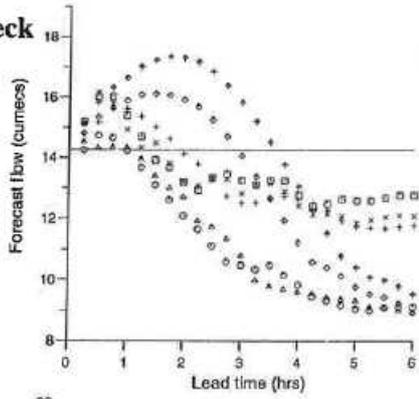
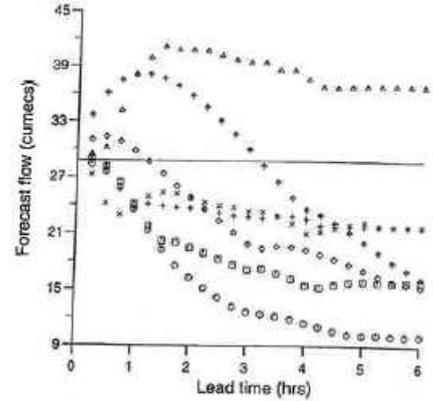
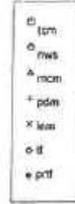


Figure 4.2.7 (cont...) Evaluation mode fixed-origin forecasts (b) The Stour: 9–10 April 1998 (Easter 1998 Flood). (Observed flow: bold line, forecast flow: fine line)

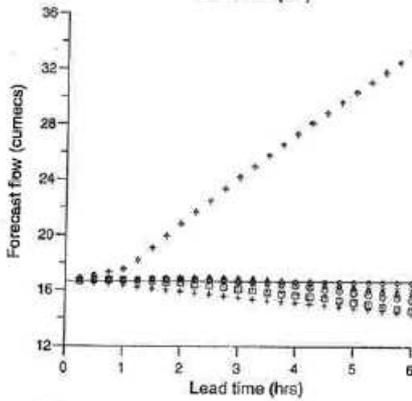
Trout Beck
16/11/95



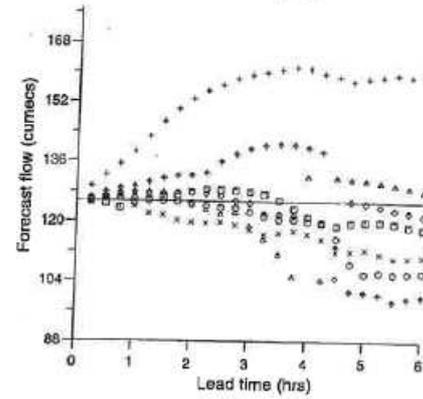
Silk Stream
23/9/92



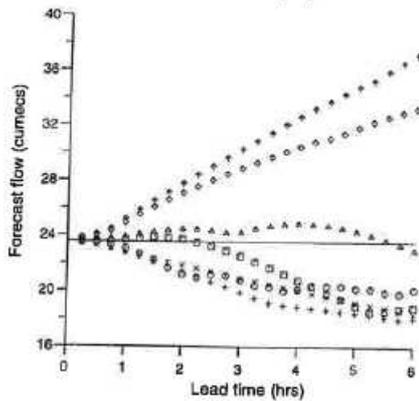
Dove
7/3/98



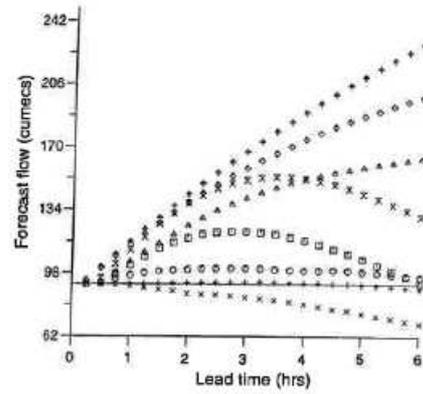
Rhondda
2/12/92



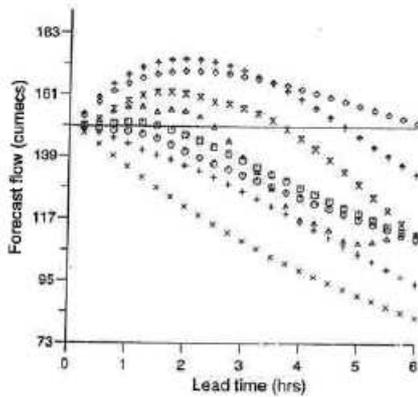
Brue
25/11/96



Stour
9/4/98



Roch
1/2/95



Witham
3/1/98

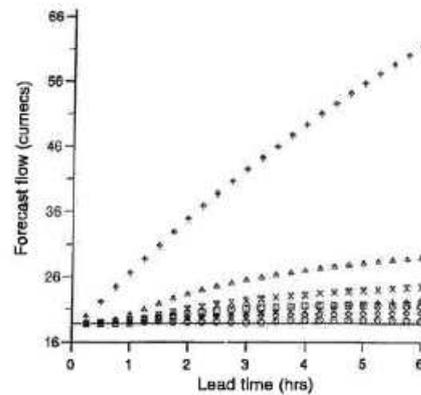
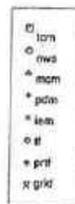


Figure 4.2.8 Variation in peak forecast accuracy with lead time for a significant flood event in each catchment.

Lavant
12/1/94

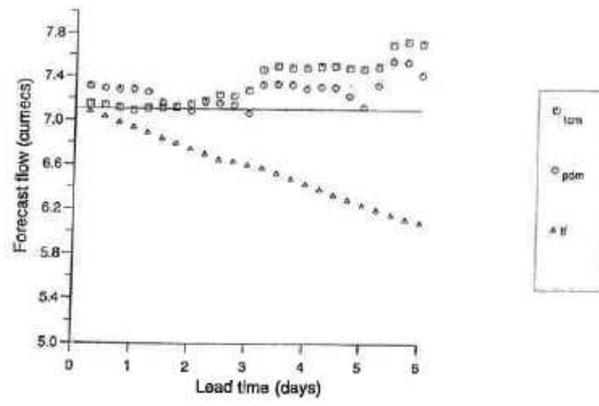


Figure 4.2.8 (cont...) Variation in peak forecast accuracy with lead time for a significant flood event in each catchment.

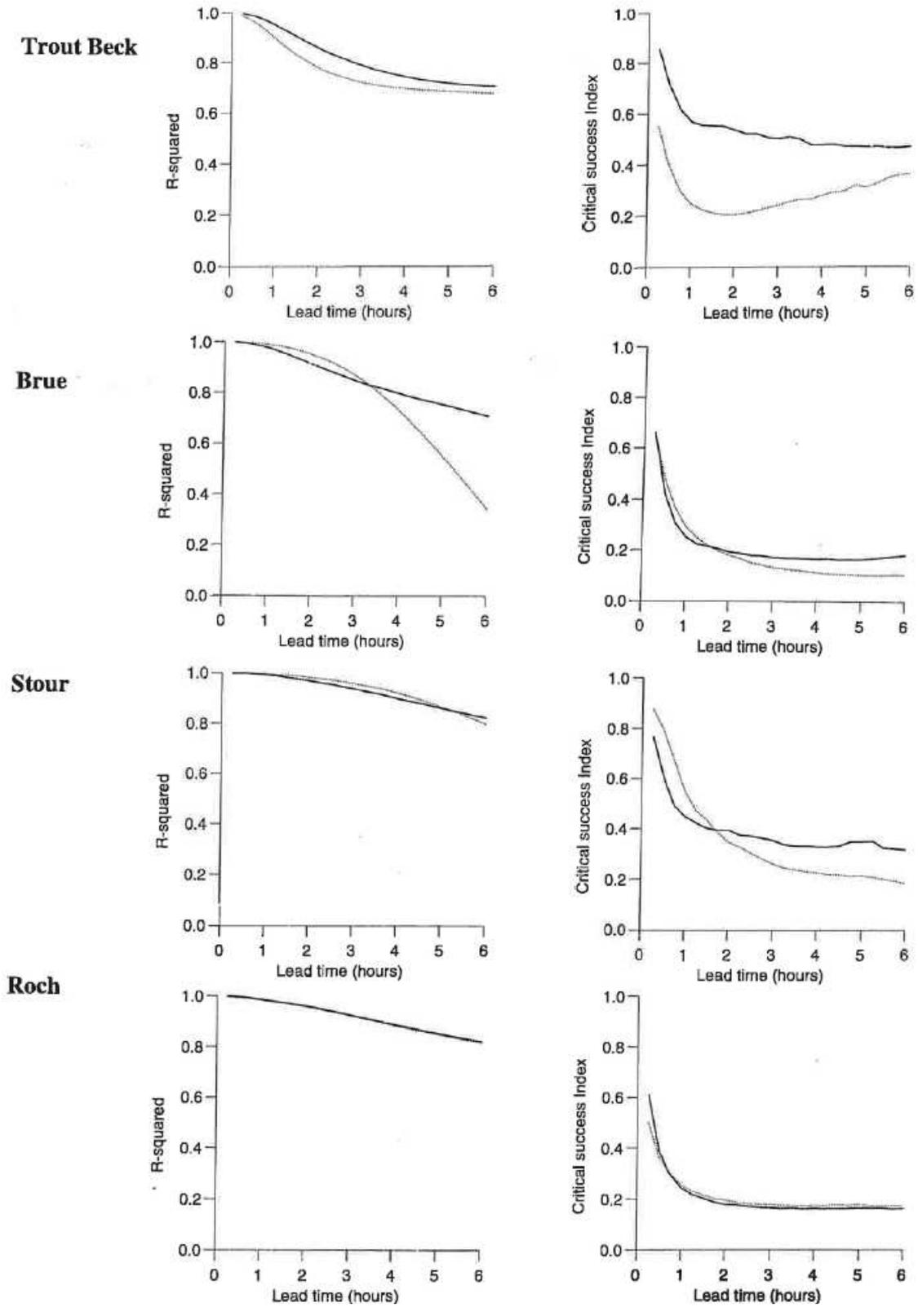


Figure 4.2.9 Variation in forecast performance (R^2 and Threshold CSI) with lead time: TF model using error prediction (bold line) and state updating (dotted line).

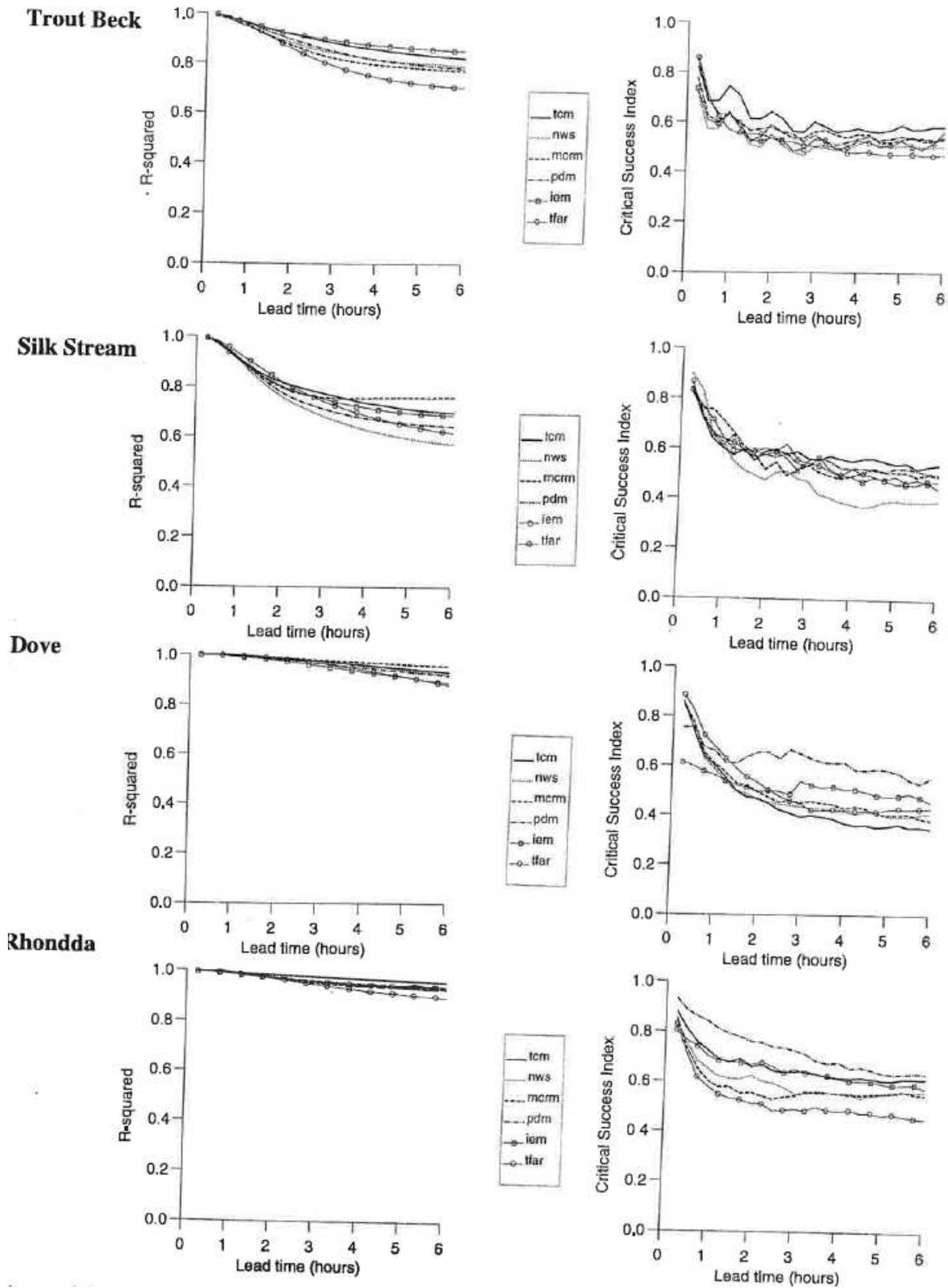


Figure 4.2.10 Variation in forecast performance (R^2 and Threshold CSI) with forecast lead time for different models: TF model with error prediction.

4.3 Results Using Weather Radar Data

4.3.1 Simulation-mode results

Tables 4.3.1(a) and (b) present the pooled R^2 and threshold CSI statistics obtained from all models operated in simulation mode on the three catchments for which radar data are available. Results are presented using three forms of rainfall input: uncalibrated radar data, calibrated radar data and raingauge data. The tables indicate that for the Silk Stream and the Stour, although no one source of rainfall data is best for rainfall-runoff modelling, use of calibrated radar data generally provides good flow simulation results, particularly when judged on the R^2 performance statistic. Use of uncalibrated radar data generally leads to poorer model performance than that obtained with raingauge or calibrated radar data. For the Roch, when model performance using raingauge and radar data is compared, it is clear that use of raingauge data results in superior model performance in all cases.

Figure 4.3.1 presents model hydrographs for a selection of more significant flood events in the three catchments obtained using raingauge and calibrated radar data. Illustrative radar fields at the radar field, storm and catchment scales are shown in Figure 4.3.2. Figure 4.3.1(a) shows hydrographs for the September 1992 flood event in the Silk Stream obtained using the TCM, the MCRM, the PDM and the PRTF model. Overall, the ability of the models to simulate this flow event is disappointing. However, it should be stressed that this is an unusually large flow peak which exceeded the red flood warning threshold of $18 \text{ m}^3 \text{ s}^{-1}$ by approximately $10 \text{ m}^3 \text{ s}^{-1}$. The PDM came closest to modelling the peak correctly using both radar and raingauge data, whilst the MCRM overestimated the main peak by $20 \text{ m}^3 \text{ s}^{-1}$. Radar rainfall fields displayed at different scales for 21:30 22 September 1992 in Figure 4.3.2(a) serve to illustrate the associated storm field.

The 1998 Easter flood event in the Stour is shown in Figure 4.3.1(b), which presents hydrographs from the four best models on this event, the TCM, the MCRM, the IEM and the Grid Model. In all cases model performance is improved through the use of calibrated radar data, although the improvement is minor in the case of the Grid Model. Note that the differences in these hydrographs from those shown in Figure 4.2.3(b), are due to models being run over a shorter period for the evaluation using radar data, resulting in a shorter warm-up than that used in the evaluation using raingauge data. The best models on this event are the Grid Model, using both raingauge and calibrated radar data, and the TCM with radar data.

Figure 4.3.1(c) illustrates the poor model simulation results obtained for the Roch in December 1993 using radar data, compared to hydrographs obtained using raingauge data. Use of radar data resulted in overestimation of a number of peaks and a poor overall model performance. Figure 4.3.2(c) illustrates the associated rainfall field estimated from Hameldon Hill radar on 8 December but fails to reveal evidence of bright band or other forms of anomaly.

Scatter plots of simulated and observed peak flow for the three catchments are shown in Figure 4.3.3(a) – (d). The peak scatter plots for the Roch confirm that many of the models overestimate the peaks when radar data are used, although there is a tendency for the models to underestimate the peaks slightly when raingauge data are used. Figure 4.3.3(a) shows a similar set of scatter plots for the Silk Stream and serves to illustrate how calibration of radar

using raingauges can reduce errors in flow simulation when compared with the use of uncalibrated radar data. Comparing parts (i) and (ii) of Figure 4.3.3(a) shows that overestimation in smaller peaks and underestimation in the main flow peak are lessened following raingauge calibration of radar. Although the scatter plots show a considerable spread of results, the TCM and IEM seem to model flow peaks well, and although the PDM performs well for the main flood peak it tends to overestimate smaller peaks.

Figure 4.3.3(c) presents a set of peak-flow scatter plots over the period 1 to 31 January 1994 for the Stour. The plots highlight the beneficial effect of radar calibration on flow simulation performance, reducing overestimation of flow peaks. Figure 4.3.3(d) shows the corresponding set of scatter plots obtained for the Easter 1998 Flood on the Stour. The plots show that the best peak simulation results were obtained using the Grid Model and the TCM with calibrated radar data.

Summary and Discussion

The simulation results presented in the previous section suggest that improvements in model performance for the Silk Stream and the Stour may be achieved through the use of calibrated weather radar data instead of rainfall estimates from raingauges. For the Silk Stream the median R^2 flow performance statistic increases from 0.635 to 0.672 when calibrated radar data are used, and for the Stour, this statistic increases from 0.622 to 0.703. Improvements in rainfall estimation and flow simulation in the Thames Region through the use of Chenies weather radar data are well documented (Moore *et al.*, 1989, 1993, Carrington and Moore, 1996). However there have been no detailed studies of the performance of the Hameldon and Clee Hill weather radars when used to provide rainfall estimates across a range of rainfall-runoff model types. Cluckie and Owens (1987) incorporated radar rainfall estimates from Hameldon Hill in their evaluation of the TF model delta-updating scheme over the Roch, but Moore *et al.* (1994) obtained mixed results when radar data were used instead of raingauge data for flow simulation in the Wyre catchment. The poor results obtained for the Roch when radar data were used could be attributed to a number of causes:

- (i) During the first two evaluation events in the Roch, radar data were available only at a resolution of 5 km. Data at a resolution of 2 km were only available during the May 1997 evaluation period. The lower data resolution may have an adverse influence on the results.
- (ii) The presence or otherwise of bright band in the data was uncertain. Bright band flags were present in much of the raw radar data, but the value for the height of the bright band was missing.
- (iii) The three periods used in the evaluation employing radar data on the Roch were dominated by stratiform rainfall, which is typically more spatially uniform than convective rain, and is likely to be measured reasonably well by raingauges. The beneficial effects of radar over a large catchment may be more apparent during convective storms, which are associated with rainfall fields that are more spatially heterogeneous.

The long periods of radar and raingauge data available for the Silk Stream (eight years of data were available) have ensured that a reliable comparison of the use of radar and raingauge

rainfall for flow simulation in the catchment can be made. However, significantly shorter periods of radar data are available for use in the Roch and Stour, which may have resulted in less robust model calibrations using radar data than were obtained using raingauge data. Typically the period of data used for the radar model calibration on these catchments was 10 to 30 days, while the length of raingauge rainfall record used for calibration was typically 6 to 8 months. Although the results suggest that the model calibrations using radar data were not adversely affected by the short periods of radar data (flow simulation results in the Stour are improved through the use of calibrated radar data), the short periods of record do not allow the models to “warm up” so well. This may have had an effect on model parameter estimation and resulted in a less robust calibration. For example, Figure 4.3.1(b) for the MCRM shows the poor model performance obtained for the Easter 1998 flood in the Stour when raingauge data are used over a short period of record without a warm-up period. The hydrograph for this period obtained for the evaluation using raingauge data (Figure 4.2.3(b)) presents a significantly better result for the MCRM model because of the longer period of warm-up used. These results highlight the difficulty of a fair comparison between radar and raingauge data for use in flow simulation, and the sensitivity of models such as the MCRM to initial conditions. It is not clear how the correction procedures now available in Nimrod might have led to improvements in the performance reported here; scope for further work in this area clearly exists. However, Nimrod at present cannot enjoy the potential benefit of the local raingauge data employed in this study and has a resolution of only 5 km.

Table 4.3.1 Evaluation performance for different models and catchments: simulation mode using raingauge, uncalibrated radar and calibrated radar estimates of rainfall as input

(a) R² statistic

Catchment	Rainfall Input Type	Model								Catchment Median
		TCM	NWS	MCRM	PDM	IEM	TF	PRTF	Grid	
Silk Stream										
	Raingauge	0.688	0.556	0.762	0.635	0.681	0.566	0.577	-	0.635
	Uncalibrated radar	-0.010	0.180	0.374	-0.237	-0.155	0.093	0.142	-	0.093
	Recalibrated radar	0.717	0.662	0.411	0.710	0.672	0.591	0.664	-	0.664
Stour										
	Raingauge	0.805	0.521	0.603	0.640	0.692	0.554	0.354	0.788	0.622
	Uncalibrated radar	0.276	0.361	-1.387	0.015	0.528	0.469	0.432	0.656	0.397
	Recalibrated radar	0.730	0.676	0.797	0.571	0.839	0.463	0.422	0.892	0.703
Roch										
	Raingauge	0.862	0.843	0.853	0.880	0.833	0.737	0.730	0.866	0.848
	Uncalibrated radar	0.682	0.804	0.690	0.734	0.759	0.715	0.644	0.685	0.703
	Recalibrated radar	0.751	0.682	0.773	0.770	0.750	0.697	0.583	0.736	0.743

(b) Threshold CSI statistic

Catchment	Rainfall Input Type	Model								Catchment Median
		TCM	NWS	MCRM	PDM	IEM	TF	PRTF	Grid	
Silk Stream										
	Raingauge	0.589	0.404	0.510	0.519	0.505	0.500	0.424	-	0.505
	Uncalibrated radar	0.257	0.261	0.228	0.226	0.186	0.331	0.309	-	0.257
	Recalibrated radar	0.514	0.510	0.484	0.546	0.430	0.311	0.495	-	0.495
Stour										
	Raingauge	0.400	0.404	0.342	0.442	0.326	0.271	0.195	0.548	0.371
	Uncalibrated radar	0.326	0.237	0.172	0.254	0.391	0.210	0.184	0.400	0.246
	Recalibrated radar	0.418	0.342	0.484	0.446	0.521	0.254	0.261	0.646	0.432
Roch										
	Raingauge	0.506	0.437	0.522	0.517	0.465	0.305	0.333	0.533	0.486
	Uncalibrated radar	0.417	0.420	0.402	0.438	0.454	0.301	0.228	0.440	0.419
	Recalibrated radar	0.475	0.338	0.406	0.476	0.444	0.266	0.159	0.438	0.422

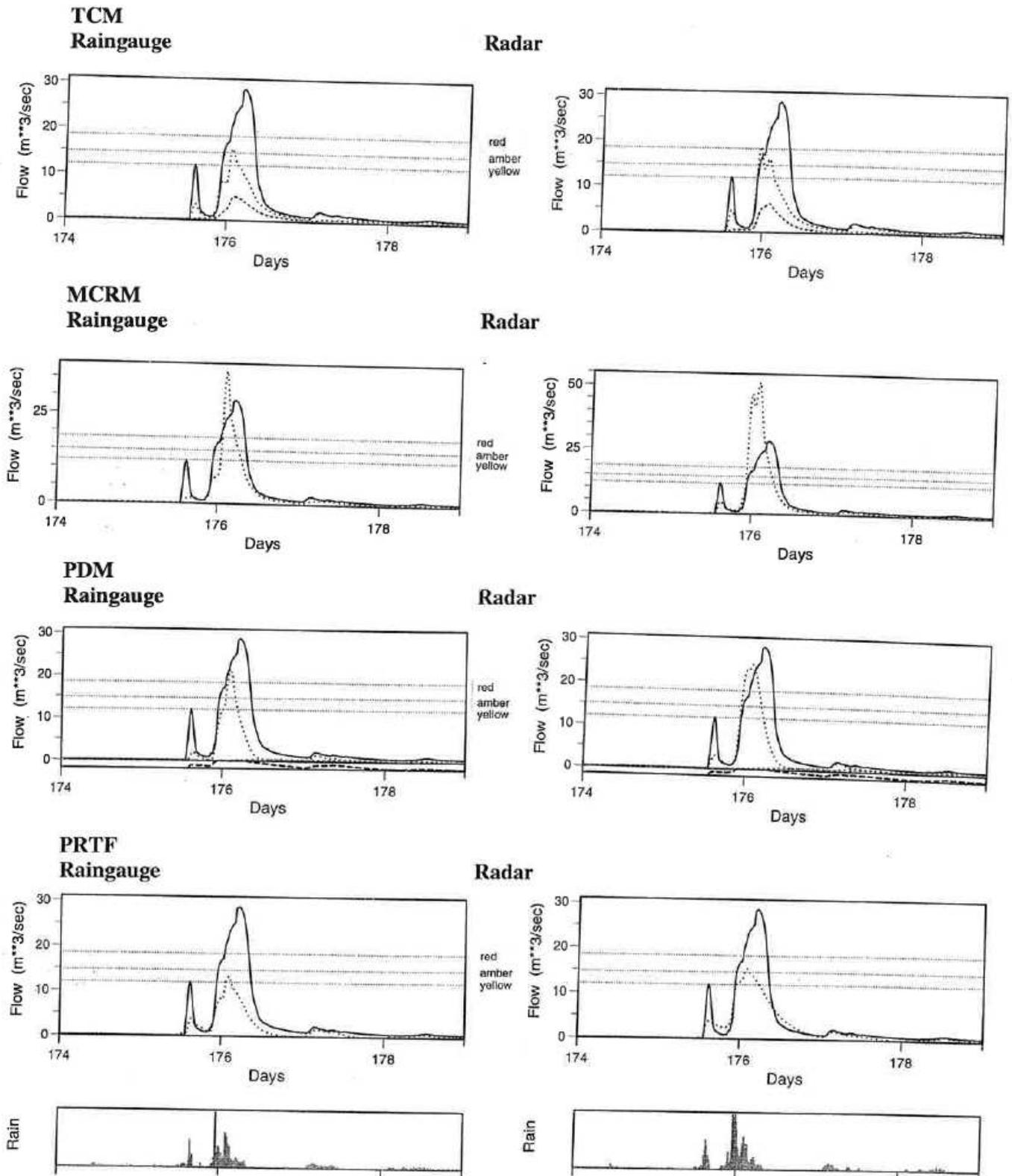


Figure 4.3.1 Simulation-mode model hydrographs using raingauge and calibrated radar data: (a) Silk Stream, 21 to 26 September 1989.

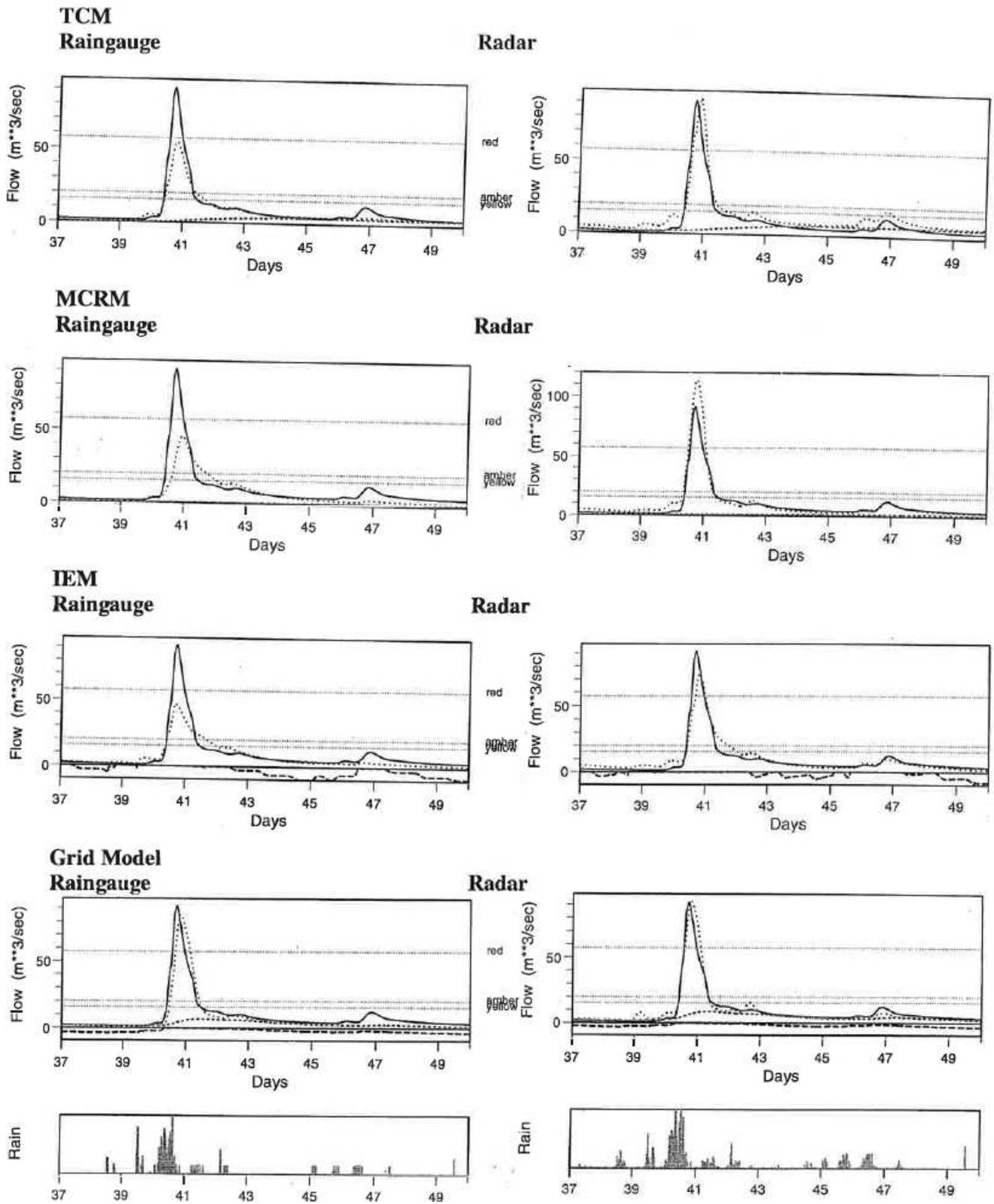


Figure 4.3.1 (cont...) Simulation-mode model hydrographs using rain gauge and calibrated radar data: (b) Stour, 6 to 19 April 1998.

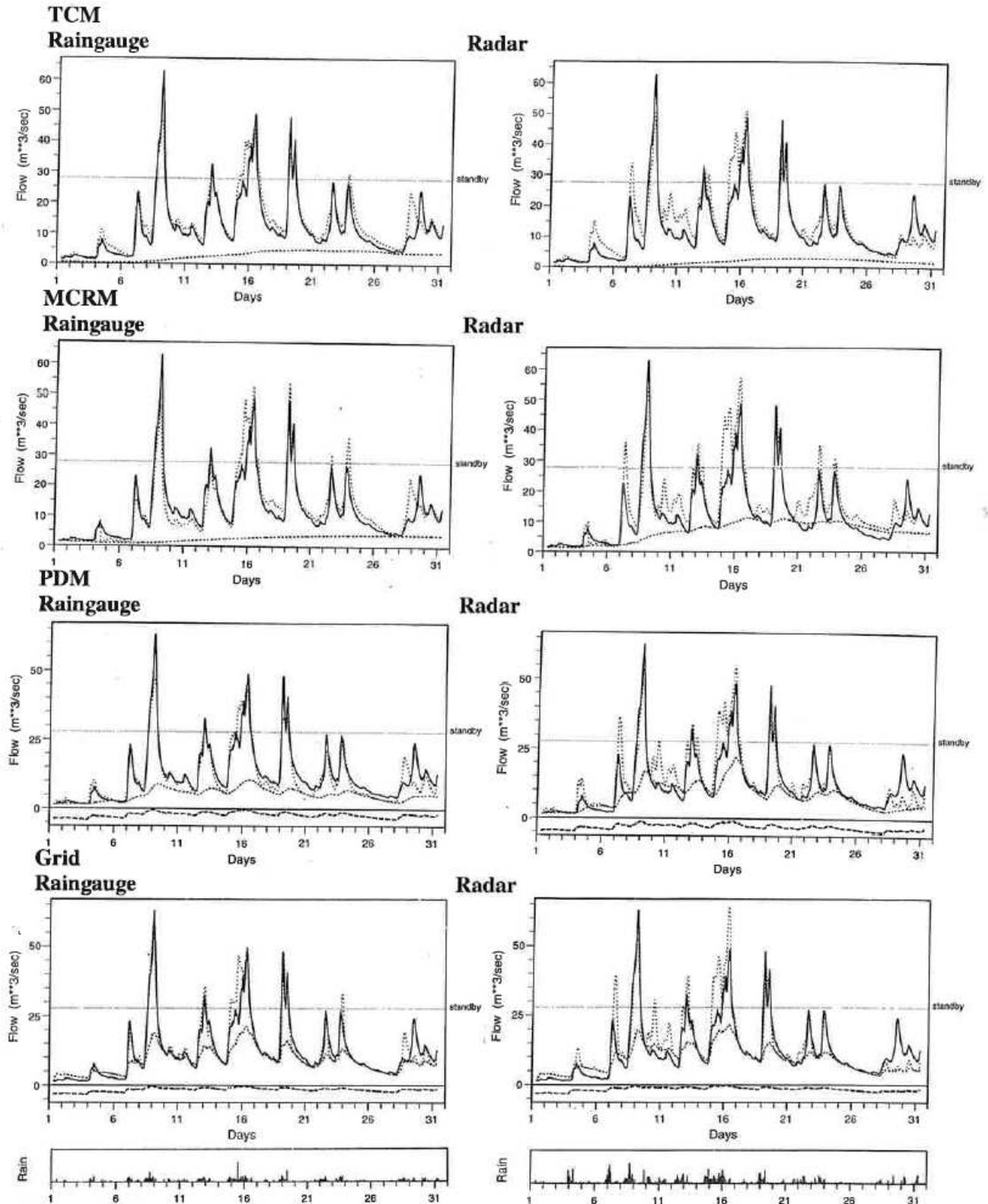
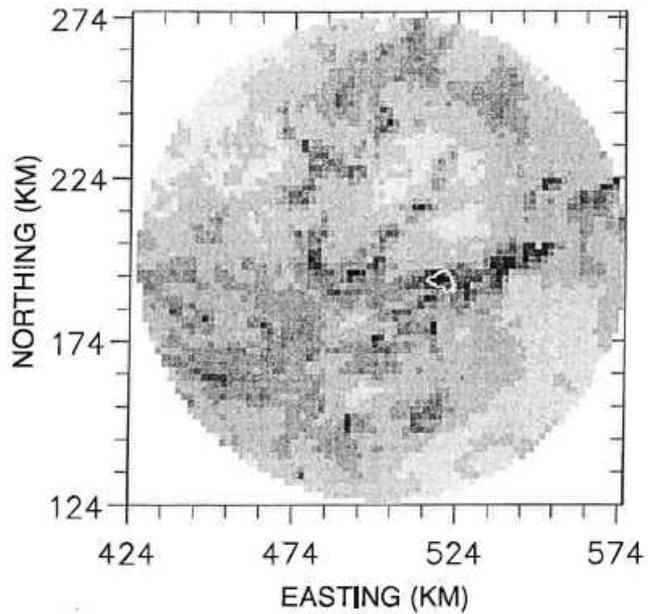
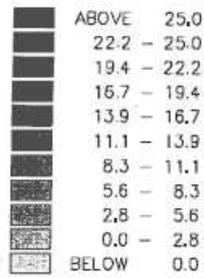
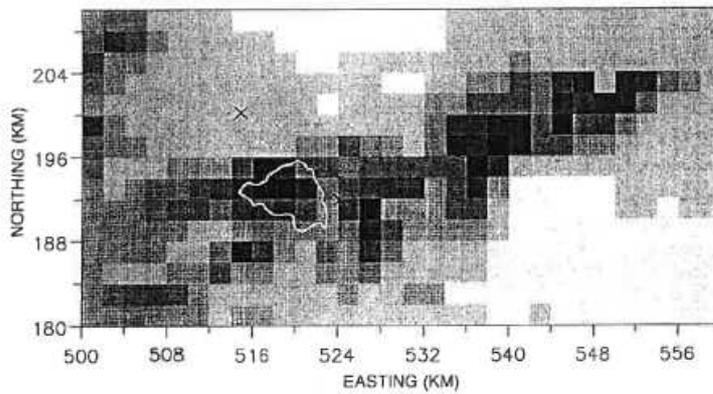


Figure 4.3.1 (cont...) Simulation-mode model hydrographs using raingauge and calibrated radar data: (c) Roch, 1 to 31 December 1993.

(a) 5 km radar field



(b) Storm field



(c) Catchment field

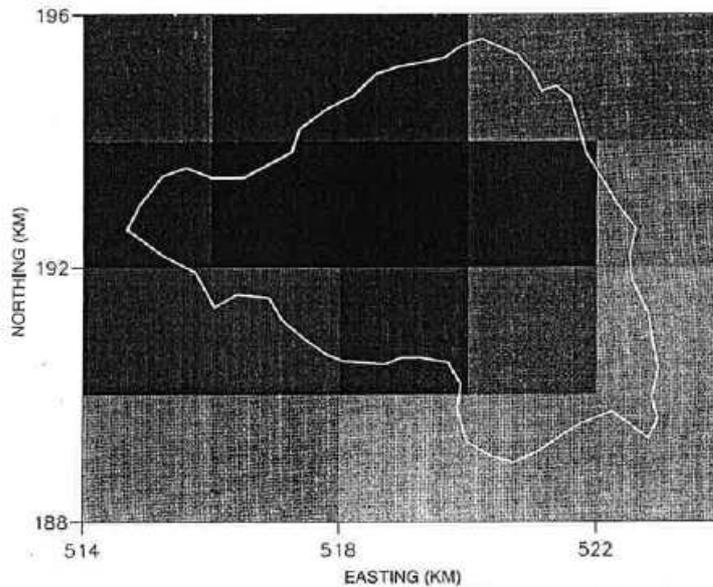
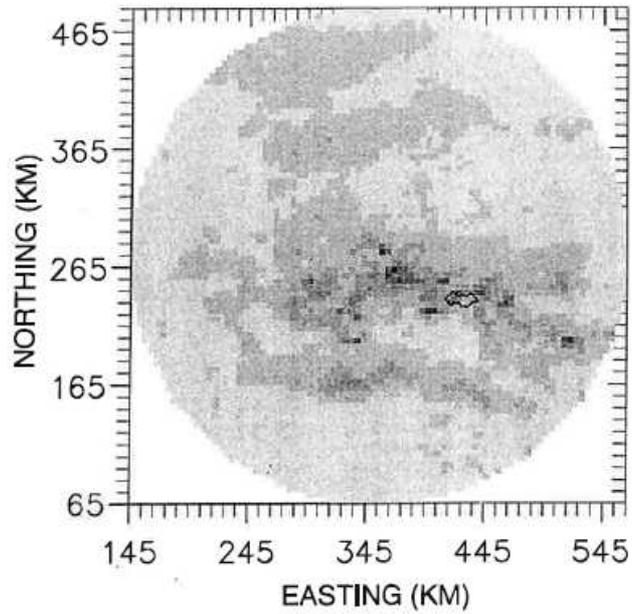
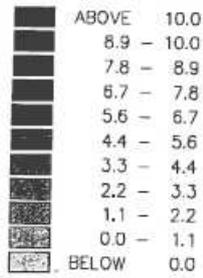
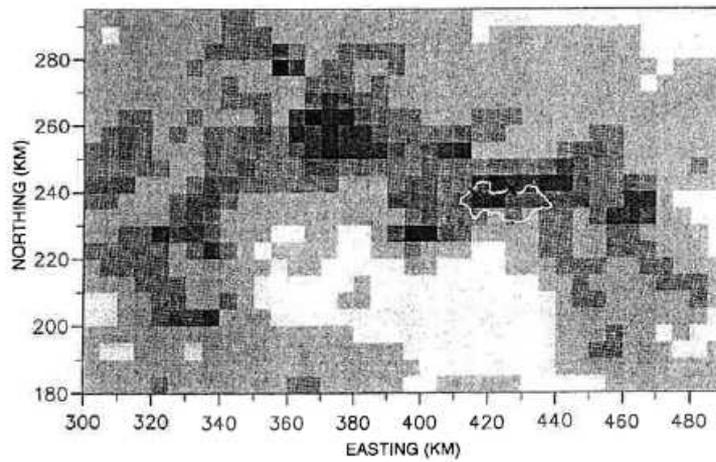


Figure 4.3.2(a) Radar rainfall estimates at different scales from Chenies radar for Silk Stream: 21:30 22 September 1992. (Positions of raingauges denoted by a cross)

(a) 5 km radar field



(b) Storm field



(c) Catchment field

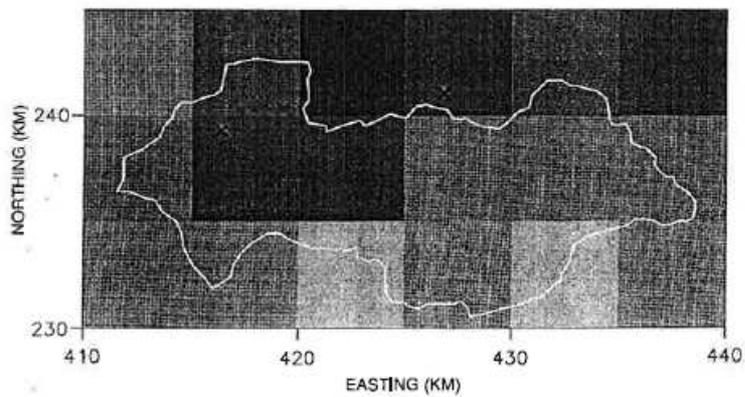
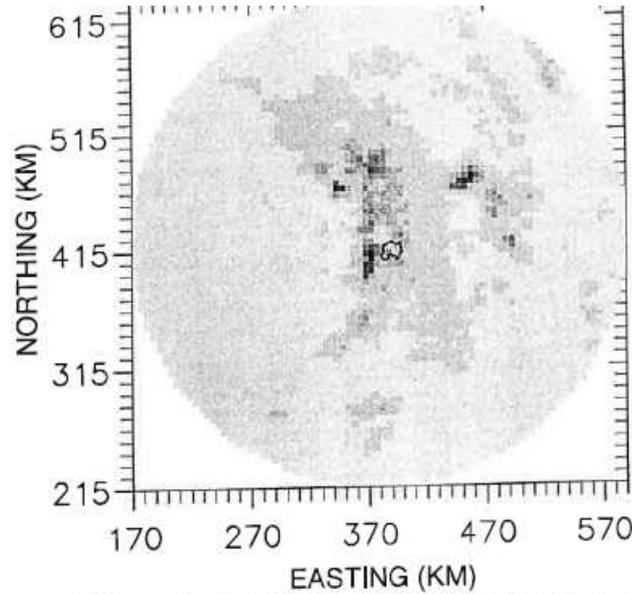
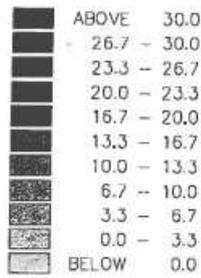
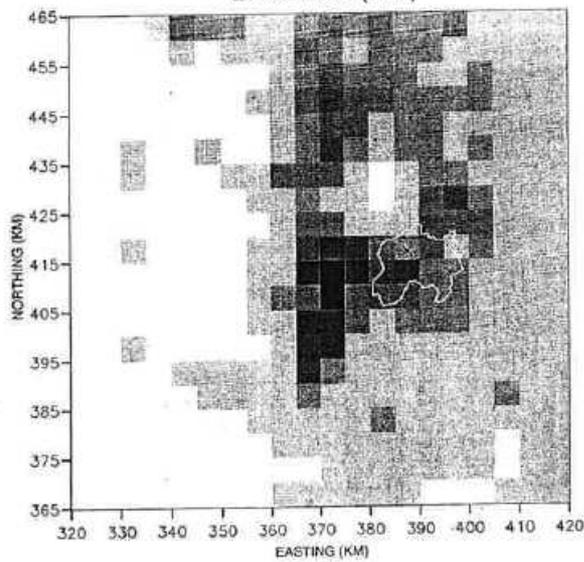


Figure 4.3.2(b) Radar rainfall estimates at different scales from Clee Hill radar for the Stour: 06:30 9 April 1998. (Positions of raingauges denoted by a cross)

(a) 5 km radar field



(b) Storm field



(c) Catchment field

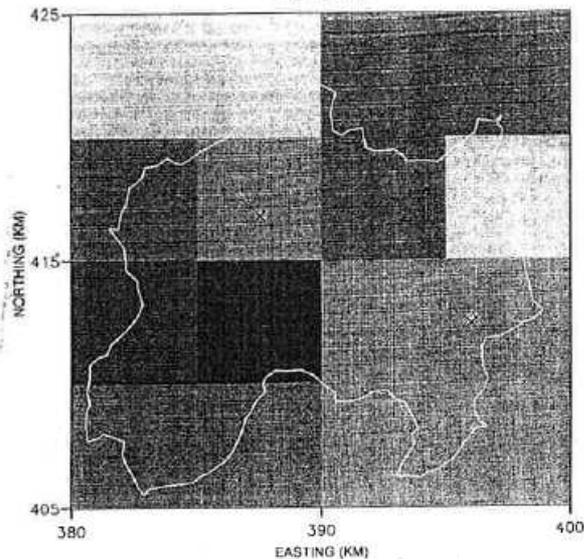
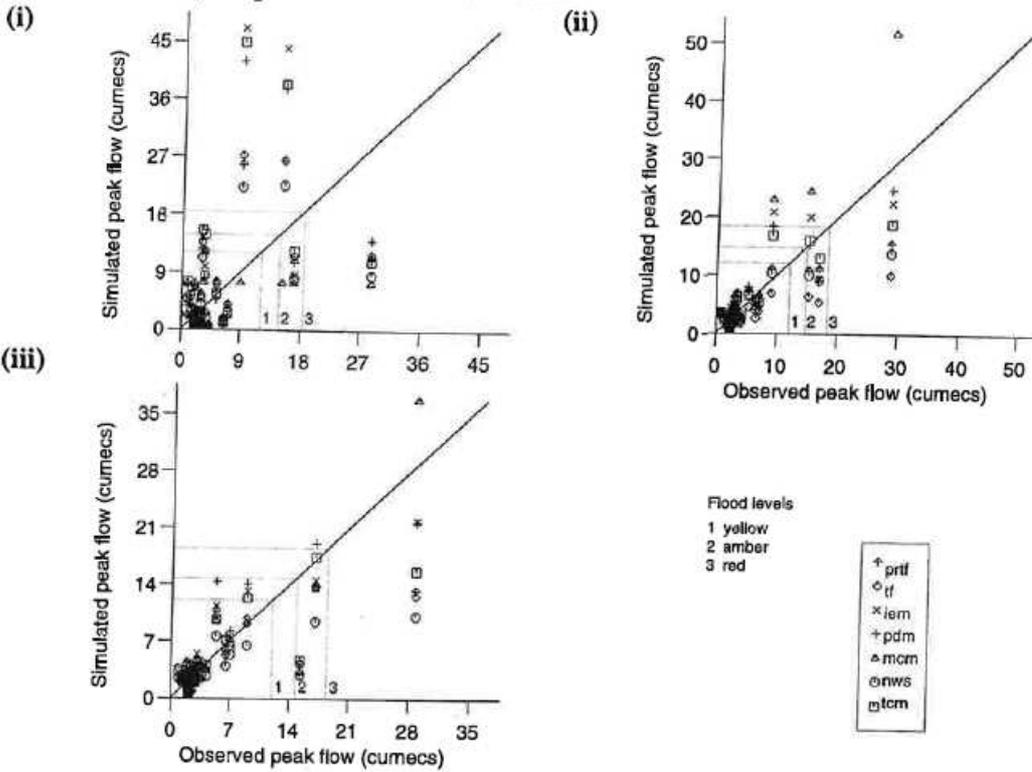


Figure 4.3.2(c) Radar rainfall estimates at different scales from Hameldon Hill radar over the Roch: 16:15 8 December 1993. (Positions of raingauges denoted by a cross)

(a) Silk Stream, 1 April to 30 November 1992



(b) Roch, 1 to 31 December 1993

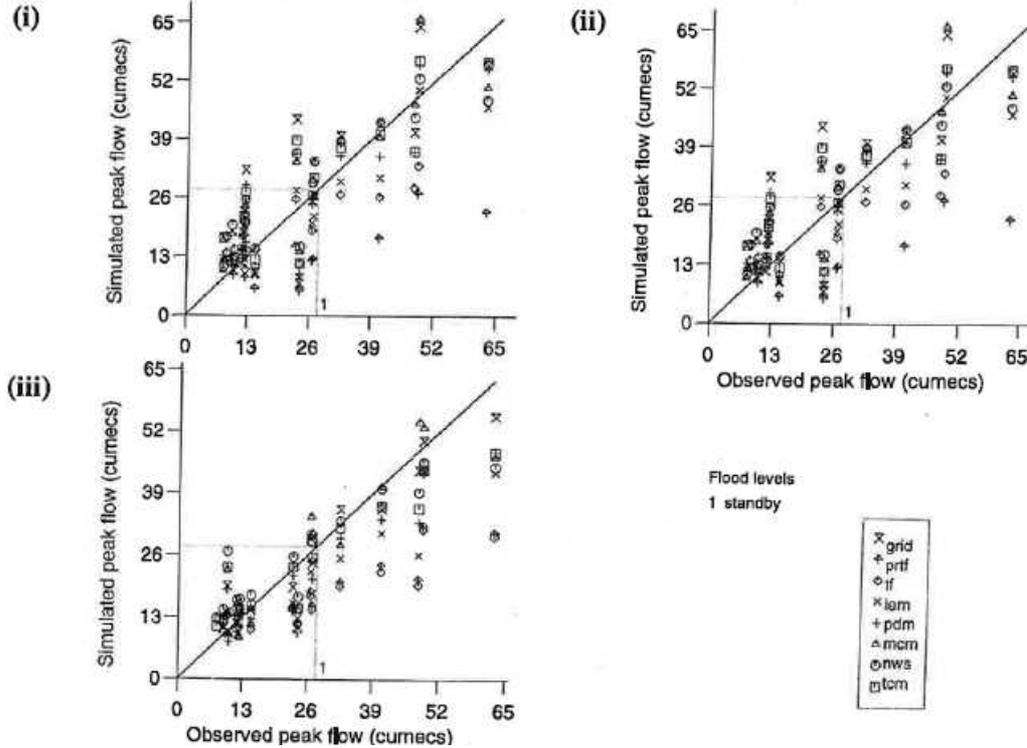
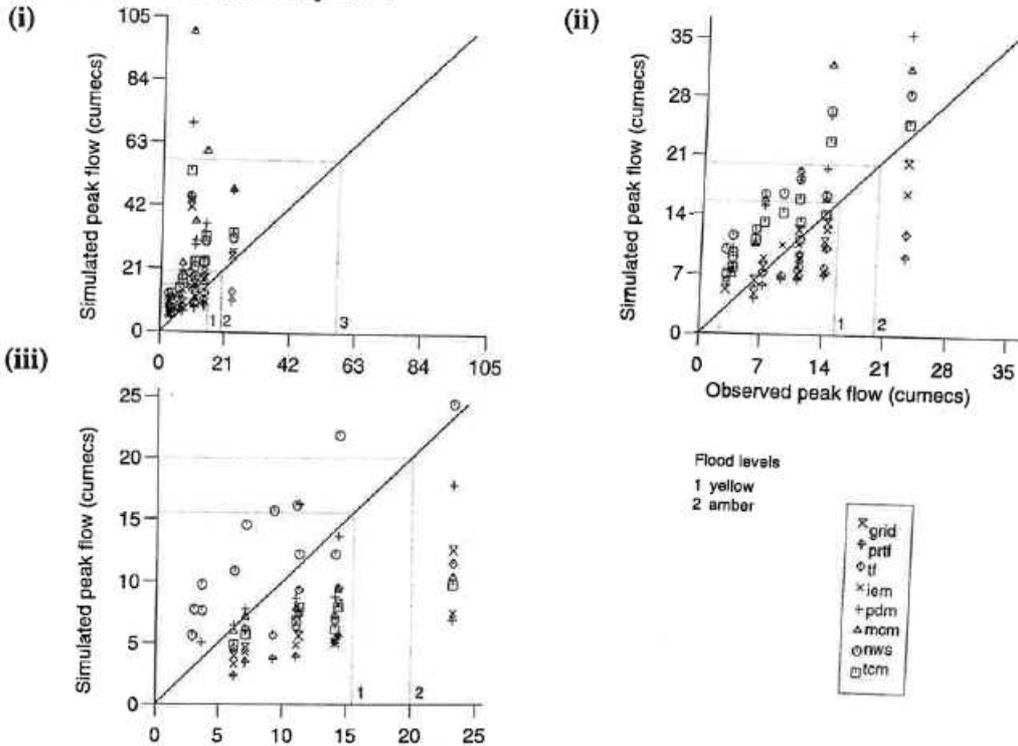


Figure 4.3.3 Scatter plots of simulated and observed peak flow using (i) uncalibrated radar, (ii) calibrated radar and (iii) raingauge rainfall data for all models.

(c) Stour, 1 to 31 January 1994



(d) Stour, 1 March to 31 April 1998 (Easter Flood)

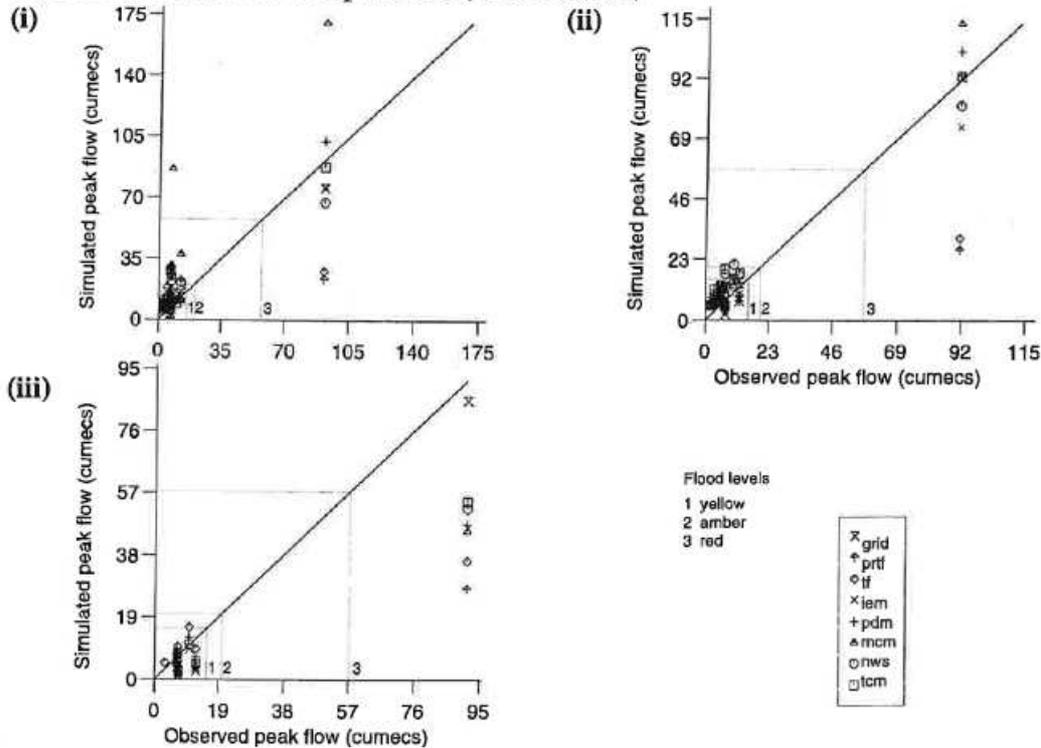


Figure 4.3.3 (cont...) Scatter plots of simulated and observed peak flow using (i) uncalibrated radar, (ii) calibrated radar and (iii) raingauge rainfall data for all models.

4.3.2 Forecast-mode results using radar data

The rainfall-runoff models are evaluated here in updating mode using both radar and raingauge rainfall data by comparing fixed lead-time flow forecasts for a number of lead times. The models are evaluated on two catchments: the Silk Stream and the Stour; simulation results using radar data for the Roch were so poor that the forecast evaluation using radar data has not been undertaken. Forecasts are presented for lead times of 0.25, 0.5, 0.75, 1.0, ..., 5.75 and 6.0 hours ahead. Results are summarised as tables of performance statistics and sets of graphs highlighting different aspects of the forecasting process. Selection of fixed origin and fixed lead-time forecasts superimposed on observed hydrographs allows a visual comparison of model performance to be made.

Table 4.3.2 summarises the performance statistics obtained over the evaluation periods for the different models applied to the Silk Stream and the Stour for a particular lead-time and provides a general overview of model performance. Figures 4.3.4 and 4.3.5 illustrate how model performance varies with forecast lead time and Figures 4.3.6 and 4.3.7 compare observed and forecast peak flow and peak timing error for a significant event in each catchment. Illustrative fixed lead-time forecasts for the Silk Stream and the Stour are shown in Figure 4.3.6. Figures 4.3.7 and 4.3.10 present a set of fixed origin forecasts for all models for significant flood events in the Silk Stream and the Stour. Figures D.17 and D.18 in Appendix D present more typical sets of fixed lead-time forecasts for all models evaluated with radar data on the Silk Stream and the Stour.

The following sections compare model forecast performance for the Silk Stream and the Stour obtained using radar and raingauge data.

The Silk Stream

Table 4.3.2 and Figure 4.3.4 compare pooled R^2 and threshold CSI performance statistics for the two events for which radar and raingauge data are used as model inputs. They suggest that the best results for the Silk Stream are obtained using the TCM, MCRM or PDM models with calibrated radar data, although results obtained using raingauge data are not dissimilar to those obtained using radar data. By way of example, Figure 4.3.8 presents a set of fixed lead-time forecasts for the Silk Stream, using radar and raingauge data, and shows how similar the forecasts are from both types of rainfall data. (Forecasts are shown at hourly intervals, the one-hour ahead forecast being closest to the observed flow). The scatter plots of observed and forecast peak flow, and peak timing error (Figure 4.3.6) confirm that there is little to choose between the use of radar and raingauge data for the Silk Stream. Fixed origin forecasts for the 22 to 23 September 1992 flood are presented in Figure 4.3.9. The IEM performs best on the rising limb of the peak, while the MCRM over predicts by a factor of two. There seems little to choose between the other models for this event

The Stour

Table 4.3.2 and Figure 4.3.5 indicate that although no one model or type of rainfall estimate is best, the MCRM, TCM, IEM and PDM all provide good forecasts at a number of lead times. Use of calibrated radar data results in slightly superior R^2 values, while use of raingauge data results in higher values of CSI. The Grid Model performs well overall in terms of R^2 but achieves better values of threshold CSI than other conceptually-based models, suggesting that

it is not forecasting the peak flows accurately. Figure 4.3.8 provides illustrative fixed lead-time forecasts for the Easter 1998 Flood on the Stour using the MCRM with raingauge and calibrated radar data. (The one-hour ahead forecast is closest to the observed flow). The forecasts made using radar data seem to result in a less 'spiky' hydrograph and result in better model performance. Figure 4.3.10 presents for all models a set of fixed-origin forecast hydrographs for the Stour using radar data as input. Visually, the Grid Model and MCRM result in a more accurate forecast hydrograph, and good results are also obtained from the TCM and NWS. The TF and PRTF models overestimate the peak flow by a factor of two resulting in a poor set of fixed-origin hydrographs. However, use of error prediction might lead to improvement. Peak errors arising from the different models using radar and raingauge data are displayed in Figure 4.3.7 for the Stour. The scatter plots suggest that there is little difference in forecast peak flow accuracy arising from the use of radar or raingauge data, although using radar data for the January 1993 period results in some overestimation of the peaks, which is not apparent when raingauge data are used.

Table 4.3.2 Evaluation performance for different models and catchments using raingauge and calibrated radar estimates of rainfall as input

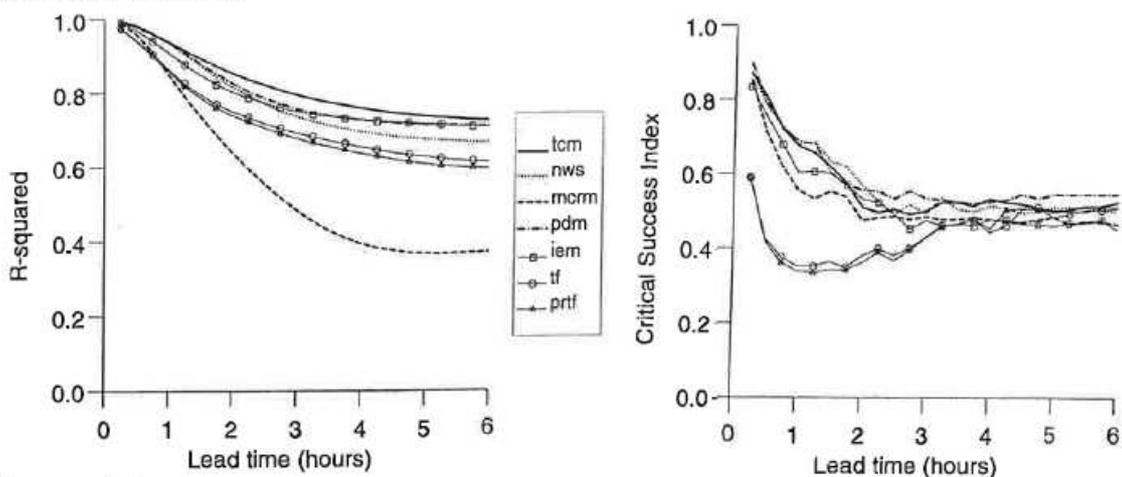
(a) Forecast mode R^2

Catchment	Rainfall Input Type	Model									Catchment Median
		TCM	NWS	MCRM	PDM	IEM	TF	PRTF	Grid	Lead time (hours)	
Silk Stream											
	Raingauge	0.913	0.906	0.911	0.912	0.906	0.898	0.821	-	1	0.906
	Recalibrated radar	0.939	0.935	0.852	0.938	0.909	0.865	0.860	-	1	0.909
Stour											
	Raingauge	0.960	0.927	0.947	0.911	0.950	0.924	0.917	0.929	3	0.929
	Recalibrated radar	0.948	0.940	0.969	0.892	0.958	0.904	0.911	0.964	3	0.944

(b) Forecast mode Threshold CSI

Catchment	Rainfall Input Type	Model									Catchment Median
		TCM	NWS	MCRM	PDM	IEM	TF	PRTF	Grid	Lead time (hours)	
Silk Stream											
	Raingauge	0.605	0.629	0.723	0.621	0.658	0.471	0.302	-	1	0.621
	Recalibrated radar	0.675	0.680	0.556	0.687	0.603	0.352	0.338	-	1	0.603
Stour											
	Raingauge	0.507	0.366	0.421	0.771	0.597	0.230	0.206	0.285	3	0.394
	Recalibrated radar	0.297	0.278	0.391	0.561	0.761	0.600	0.201	0.346	3	0.369

Calibrated radar data



Raingauge data

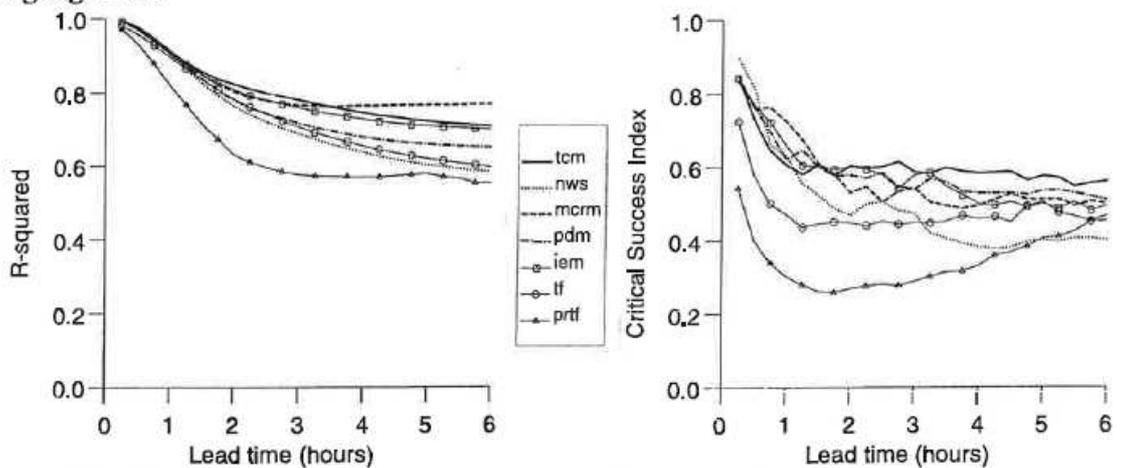


Figure 4.3.4 Variation of R^2 and Threshold CSI with forecast lead-time using raingauge and calibrated radar data: Silk Stream.

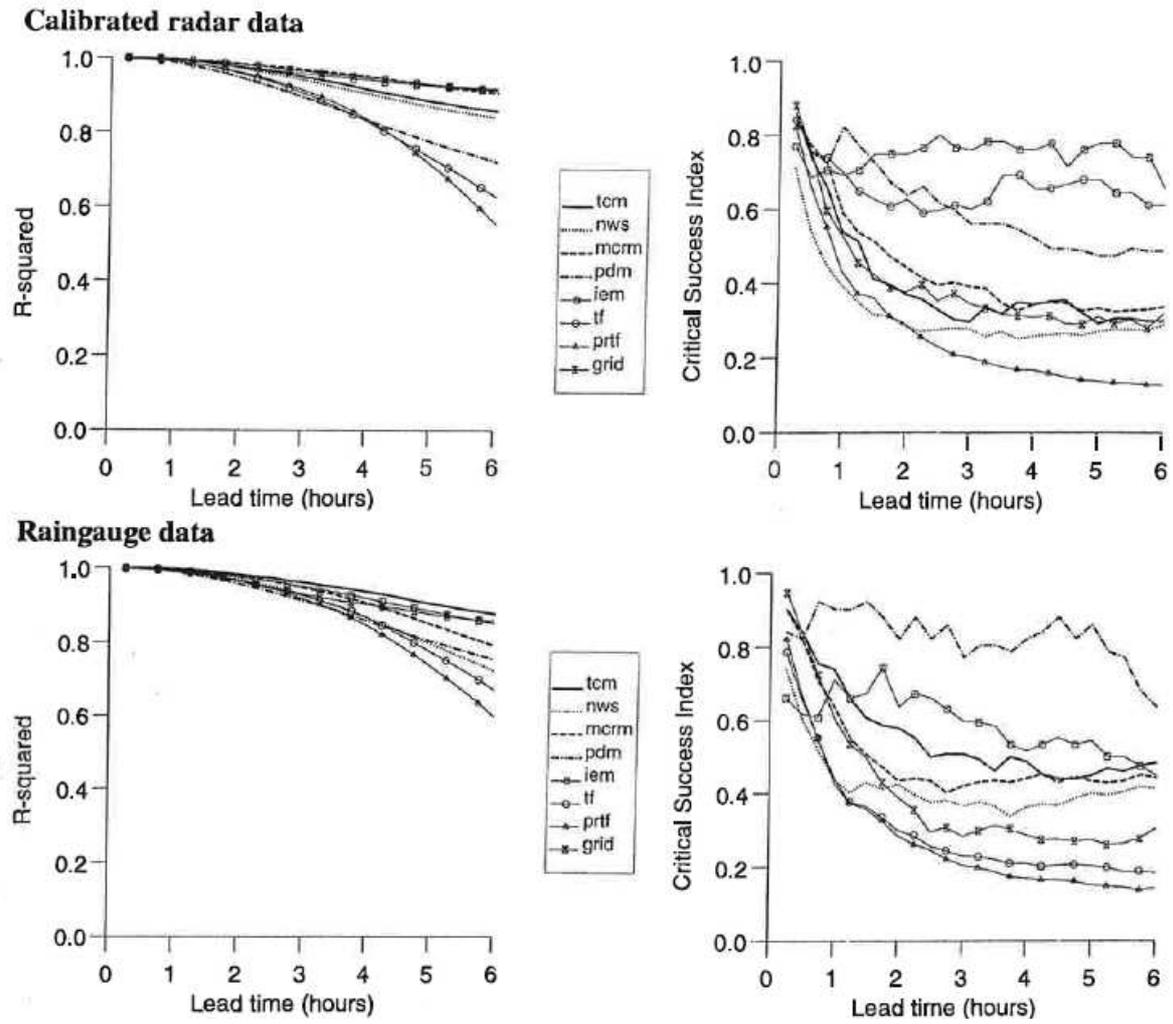
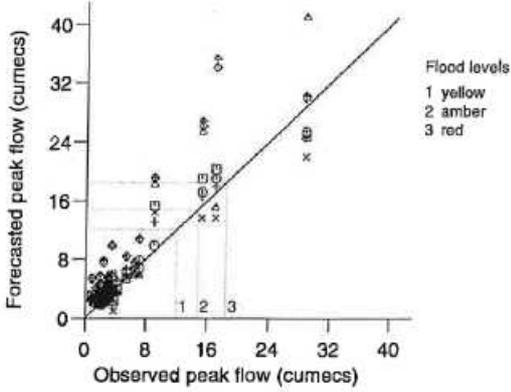


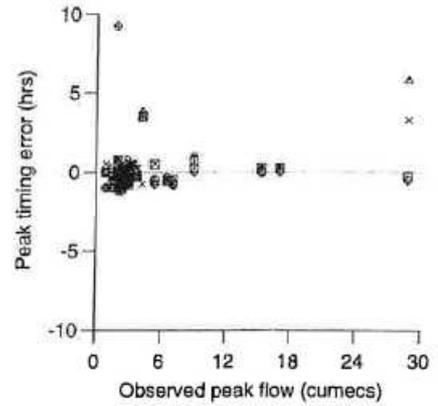
Figure 4.3.5 Variation of R^2 and Threshold CSI with forecast lead-time using raingauge and calibrated radar data: Stour.

Calibrated radar data

(i)

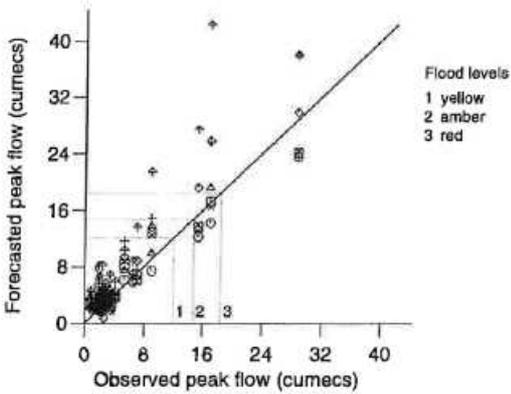


(ii)



Raingauge data

(i)



(ii)

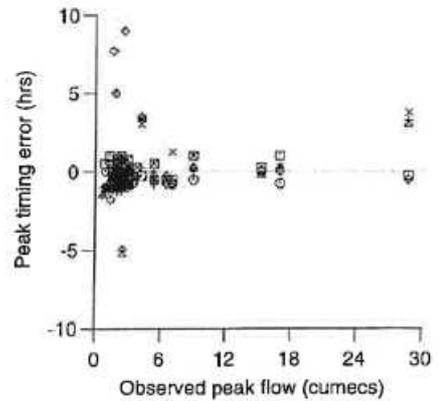
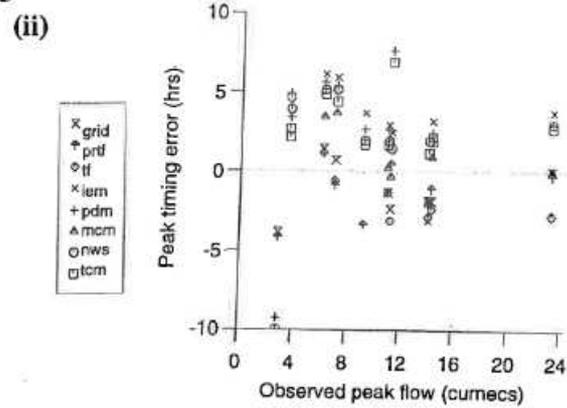
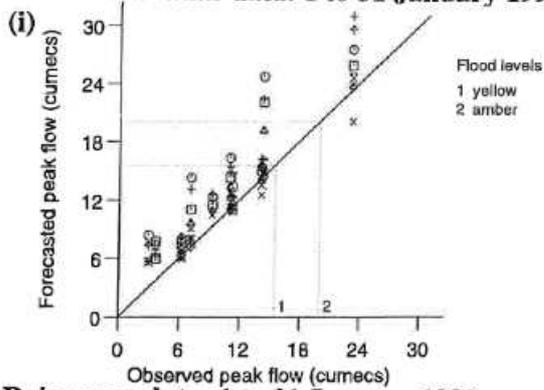
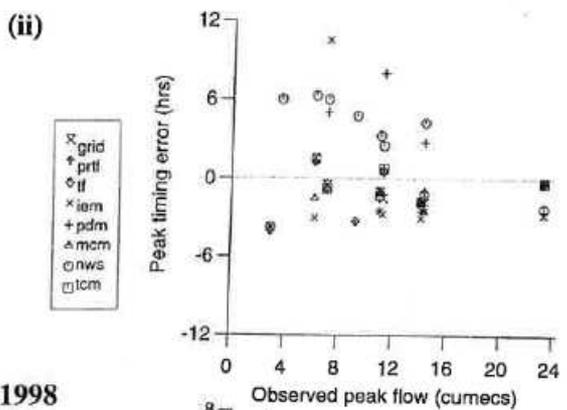
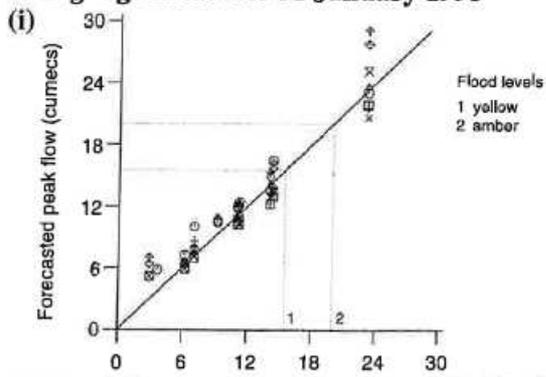


Figure 4.3.6 Scatter plots of (i) observed and forecast peak flow and (ii) peak timing error, using calibrated radar and raingauge data: Silk Stream, September 1992.

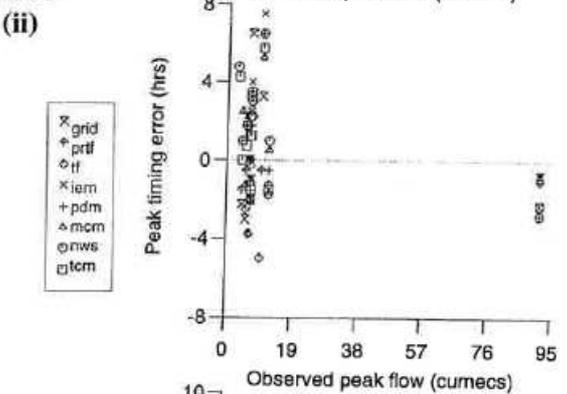
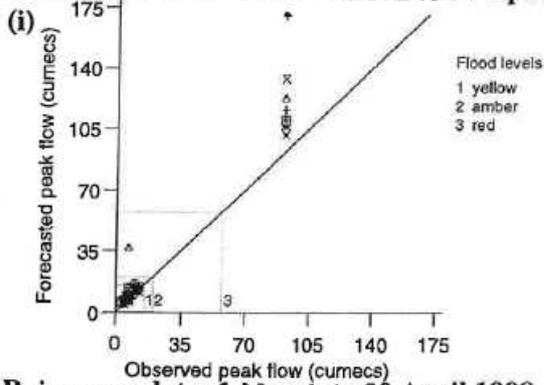
Calibrated radar data: 1 to 31 January 1993



Raingauge data: 1 to 31 January 1993



Calibrated radar data: 1 March to 30 April 1998



Raingauge data: 1 March to 30 April 1998

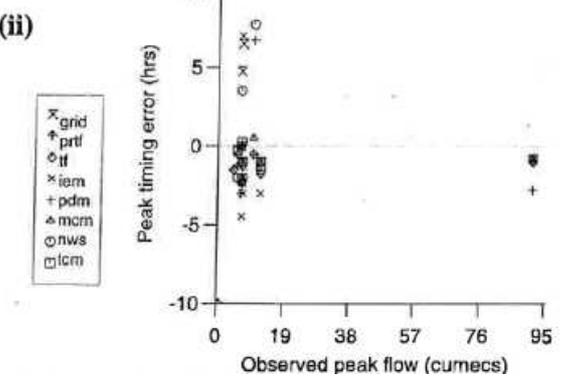
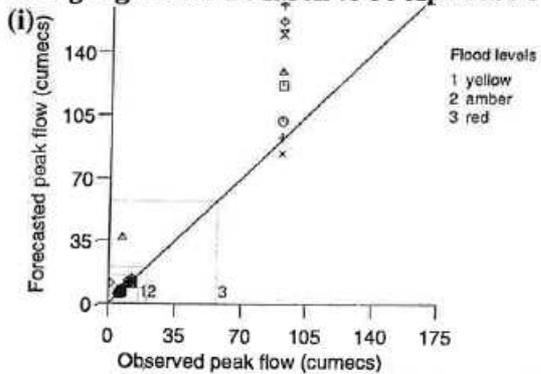
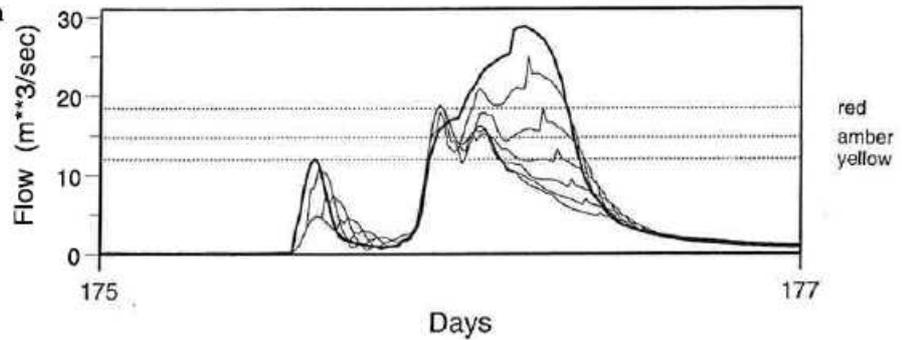


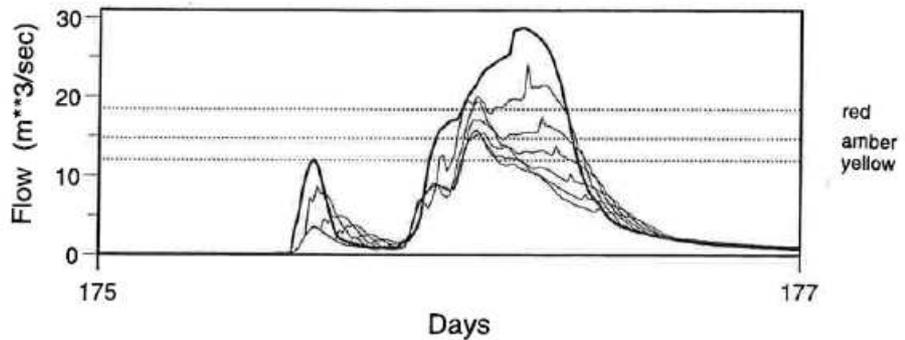
Figure 4.3.7 Scatter plots of (i) observed and forecast peak flow and (ii) peak timing error, using calibrated radar and raingauge data: Stour.

(a) Silk Stream using TCM: 22 to 23 September 1992

Calibrated radar data

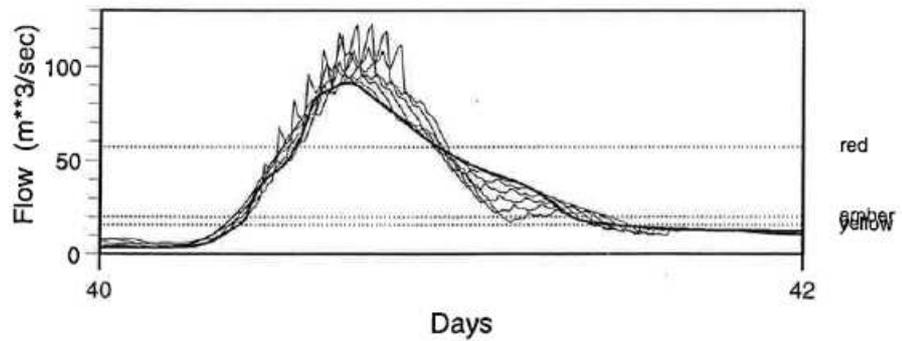


Raingauge data



(b) Stour using MCRM: 9 to 10 April 1998

Calibrated radar data



Raingauge data

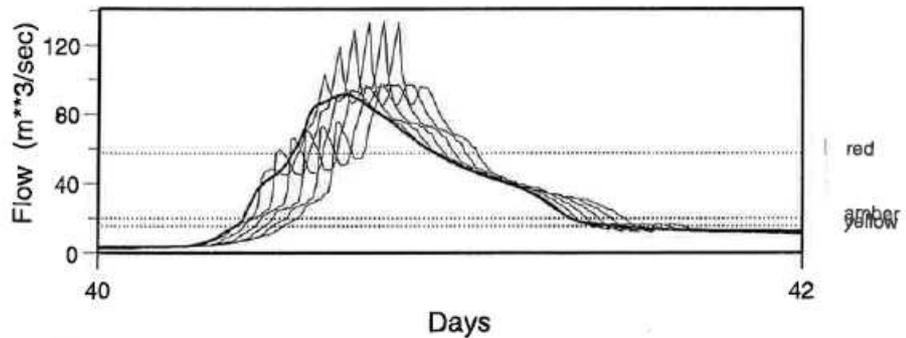


Figure 4.3.8 Fixed lead-time hydrographs (1 to 6 hours ahead) using calibrated radar and raingauge data.

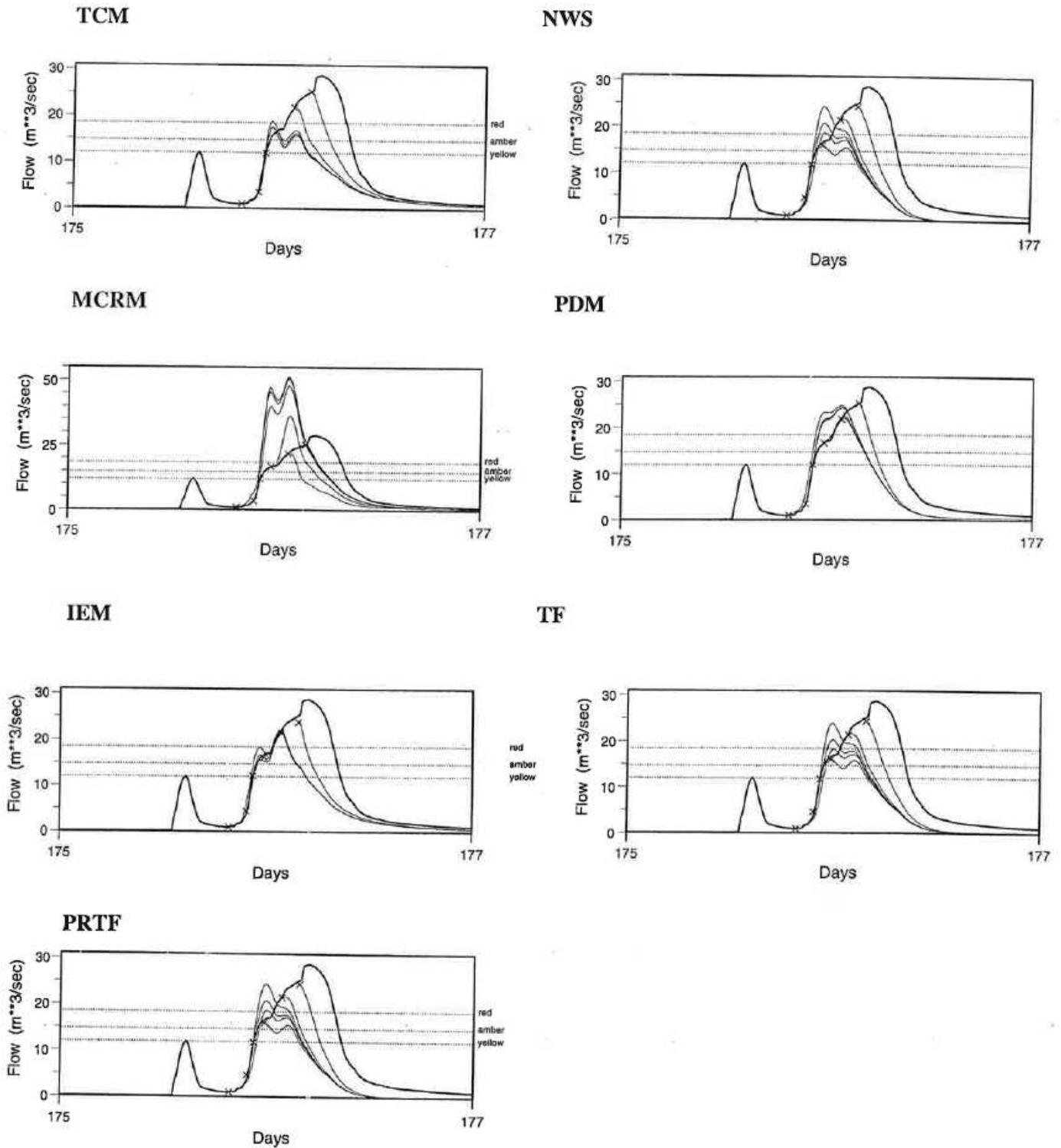


Figure 4.3.9 Evaluation mode fixed-origin forecasts using calibrated radar data: Silk Stream, 22 to 23 September 1992. (time origins are denoted by a cross, forecasts are shown with a fine line and the observed flow is shown with a bold continuous line)

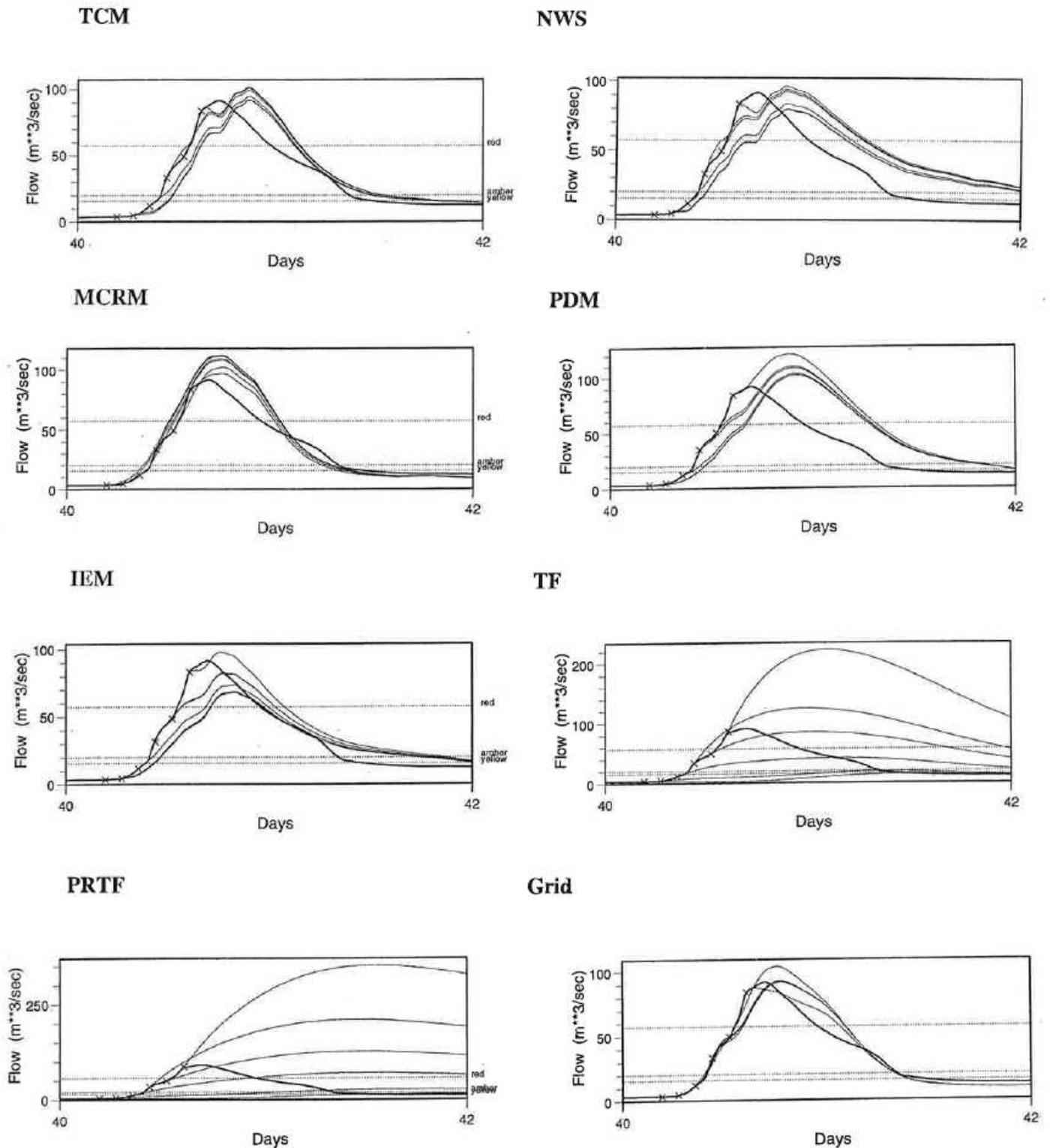


Figure 4.3.10 Evaluation mode fixed-origin forecasts using calibrated radar data: Stour: 9 to 10 April 1998. (time origins are denoted by a cross, forecasts are shown with a fine line and the observed flow is shown with a bold continuous line)

4.4 Overview

In simulation-mode the eight models calibrated to each of the nine study catchments have been evaluated using raingauge data for continuous periods of up to eight months in length. The results indicate that there is least variation in model performance across model types for responsive or urbanised catchments such as the Rhondda, Trout Beck, Silk Stream and the Roch. For these catchments reasonable model performance is obtained independent of which model is used. Greater variation in model performance is apparent for catchments such as the Dove, the Brue and the Witham, where groundwater and the effect of catchment wetness are more dominant in the hydrograph response. Although no one model out-performs the others consistently across all catchments, the TCM performs best overall with the MCRM, PDM and NWS not far behind. On the three catchments for which the Grid Model was evaluated, it performed well, consistently coming second in overall simulation performance. The MCRM was found to be particularly sensitive to initial soil moisture conditions, and initialising the model could be problematic.

During extreme flood events, such as the Easter 1998 flood in the Stour, the majority of models have a tendency to underestimate the peak flow. The NWS underestimates peak flows more than most and simulations often do not exceed the flood warning thresholds, while the TCM, PDM, IEM, TF and PRTF often underestimate the peak flows, but generally model simulations exceed at least one warning threshold. In contrast, the MCRM tends to achieve, or even exceed the flood peak, and generally crosses all flood warning thresholds, although the time of peak can be out by a few hours. On the three catchments tested, the Grid Model performs well during extreme flood events and is best overall on the Easter 1998 flood in the Stour.

When models are compared in forecast-mode, although no one model out-performs the others consistently across all catchments, the TCM performs particularly well when the R^2 statistic is used to compare models. When the Threshold CSI criterion is used the PDM is very successful, indicating that it is good at forecasting crossing of flow thresholds, such as alarm levels. For a simple model, the IEM is surprisingly successful in forecast-mode particularly in terms of the Threshold CSI. In many cases the forecast performance of the TF model can be substantially improved through the use of error prediction instead of state updating.

Three different types of rainfall input – raingauge data, uncalibrated radar data and radar data calibrated using raingauges - have been evaluated in simulation and updating mode on three catchments: the Silk Stream, the Roch and the Stour. The simulation-mode results indicate that for the Silk Stream and the Stour, although no one source of rainfall data is best for rainfall-runoff modelling, use of calibrated radar data generally provides good flow simulation results, particularly when judged on the R^2 performance statistic. Use of uncalibrated radar data generally gives poorer model performance than that obtained with raingauge or calibrated radar data. For the Roch, when model performance using raingauge and radar data is compared, it is clear that use of raingauge data results in superior model performance in all cases.

Flow forecasts have been produced for the Silk Stream and the Stour using calibrated radar data and results compared to those obtained for the same periods using raingauge data. Overall, the results indicate that when radar is estimating rainfall well, use of calibrated radar data can give as good as, or sometimes slightly better results using the R^2 performance

criterion, compared with those obtained using raingauge data. However, in order to obtain these results, considerable effort had been put into selecting calibration and evaluation periods for which radar data were not missing, unreliable, or affected by significant amounts of snow or anomalies. It is important to note that over the events tested, the extra time and effort put into evaluating large volumes of radar data resulted in, at best, a modest improvement in flow forecast accuracy in two out of the three study catchments. Over the third catchment, better simulation accuracy was obtained using raingauge data.

5. CHOICE OF CATCHMENT MODEL

Section 3 and 4 have focused on an assessment of models in terms of forecast performance using R^2 and Threshold CSI as performance statistics together with visual aids in the form of hydrographs and scatter plots of peak flows. Whilst performance is arguably the most important criterion to be used in the choice of model, there are other issues such as ease of calibration, speed of execution, ease of comprehension of model structure, strength of model formulation and model data requirements. This section aims to take a broader view on the question of model choice focusing on ease of model use, based on the experience gained in this study, in addition to forecast performance. The data requirements of the different models are also reviewed. This leads to a set of guidelines for choosing a rainfall-runoff model for different types of catchments. The advantages and disadvantages of each of the eight models included in the assessment are set down in a concise summary table.

5.1 Ease of Model Use

The eight rainfall-runoff models included in the assessment, in decreasing order of complexity, are:

- (i) Grid Model;
- (ii) Thames Catchment Model or TCM;
- (iii) National Weather Service model or NWS;
- (iv) Midlands Catchment Runoff Model or MCRM;
- (v) Probability Distributed Moisture, or PDM, model;
- (vi) Isolated Event Model or IEM;
- (vii) Transfer Function model or TF; and its variant,
- (viii) Physically Realisable Transfer Function or PRTF model.

For a detailed review of these models the reader is referred to the Part 1 Report whilst an indication of the model structure of the models may be gleaned from the model schematics presented as Figures 1.2.1 to 1.2.8. The following reviews each model from a user's perspective focusing on ease of model configuration, calibration and initialisation.

5.1.1 TCM

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

The structure of the Thames Catchment Model, or TCM, is based on subdivision of a basin into different response zones representing differing types of land use, soil, geology and topology, for example representing runoff from gravel, clay, aquifer and riparian areas. Identification of these zones has been achieved using the IH DTM for the UK in conjunction with a number of digital spatial datasets (Section 2.4.2). The zonal responses should be sufficiently different both to avoid excessive parameter interaction and also because each zone should have a hydrological justification. This process produces proportions of the

catchment covered by differing hydrological response zones. In the Thames Catchment Model these proportions are multiplied by the area of the catchment to give an initial value for the size of each response zone. Final values for the area of each zone may differ from these initial values, as zone size is a parameter in the TCM and can be adjusted to give optimum model performance. In addition, the zone size can act as a multiplicative rainfall factor, adjusting for the representativeness of the raingauges used. Therefore the total area of the zones may differ from the size of the catchment after calibration, and also their relative sizes compared to each other may change. For some catchments, it may suffice to think simply in terms of a “slow response” zone and a “fast response” zone, analogous to the slow and fast response paths of the PDM.

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

5.1.2 NWS

The NWS is a relatively complex model with a large number of parameters (16), but is surprisingly straightforward to calibrate using a mixture of manual and automatic optimisation. The usual procedure of calibrating the slow runoff component first, followed by the fast component should be followed. The NWS appears to be relatively insensitive to initial soil moisture conditions, so a “warm-up” period has not been used here. The original daily time-step model formulation did not include a time-delay. However, initial trials suggested that model performance is improved with the use of a time-delay and one has consequently been included here.

5.1.3 MCRM

The MCRM has a large number of parameters (22). However, once the initial conditions for an event are set-up correctly the model can be more straightforward to calibrate than the number of parameters might suggest. Several of the parameter values should lie within a narrow range and so can be set to standard values initially. Calibration of the model can be

divided into five parts: the groundwater, the soil store, the timing, the smoothing, and, finally the interception store. Model parameters can be identified by taking each of these parts in turn and using a mixture of manual and automatic optimisation; this procedure was found in practice to be iterative.

The performance of the MCRM model can be very sensitive to the initial soil moisture deficit; it is therefore vital to have a good understanding of the antecedent conditions before calibrating the model. This might involve including a “warm-up” period before the event begins. Operationally, initial conditions are based on previous runs and are adjusted weekly using MORECS data. The use of MORECS soil moisture data was not considered practical in this study and instead the warm-up procedure outlined in Section 2.5.3 has been used.

5.1.4 PDM

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

A cascade of two identical linear reservoirs is usually appropriate for the fast response path. A quadratic or, more usually, cubic storage should be used for the slow response path. In general state-correction is preferred to error prediction for forecast updating. The soil moisture storage, evaporation, recharge and runoff generation mechanisms in the PDM are interlinked and highly non-linear, and the effect of changes in the associated model structures and parameter values can be difficult to predict. Use of the PDM at a daily time-step can be useful in determining slow-response model parameters. In general the PDM was found to be relatively insensitive (robust) to initial soil moisture conditions. Early model trials suggested that although an initial period of warm-up was beneficial to models such as the MCRM, model performance for the PDM did not always benefit from a warm-up, and in some cases, got worse.

5.1.5 IEM

Since the IEM structure is essentially fixed the number of parameters for calibration is small. Therefore calibration of the IEM is fast and straightforward and can be achieved largely by automatic optimisation. The IEM was found to be sensitive to initial soil moisture deficits and use of a warm-up period was found to be beneficial.

5.1.6 TF and PRTF

It has not been possible to emulate the most commonly used form of TF/PRTF model employed by the EA since this involves manual adjustment of the model parameters as the flood develops and a method of baseflow separation which is event-based. Calibration of EA models is commonly carried out using flood event data, rather than the continuous records used here, and can involve choosing a single model parameter set from those estimated for each event. Model orders and time-steps are commonly chosen according to the speed of catchment response and can involve a search for each catchment. Trials here of different orders has led to the use of a fixed TF(2,4,b) model order and time-step of 15 minutes as providing sufficient flexibility in model response across the study catchments. In other respects the TF model closely emulates that used in the Leigh Barrier Flood Forecasting Model (Pollard, undated) for the River Medway in EA Southern Region. The model gain update procedure is also used in parts of Anglian Region (Page, 1991). No special provision has been made to allow for parallel flow processes, comprising of fast and slow (baseflow) components, or nonlinear loss modules, representing the effect of catchment wetness, as these do not feature strongly in EA operational practice.

The simple structure of the TF and PRTF ensures that both models are simple to calibrate and very quick to run. For this study the models were calibrated twice, once in simulation mode and once in forecast mode. In practice, calibration is only required in simulation mode if a full TFN model (incorporating an error predictor) is used. For TF predictors which employ “full state correction” only one simple calibration is required.

For the simulation-mode calibration, an Instrumental Variable (IV) algorithm has been used to estimate the parameters of a TF(2,4,b) model for each catchment. This method of calibration is very fast and should require no user-input. However, the difficulty of fitting a TF model in simulation-mode to some of the more complex catchments meant that user intervention was required for catchments such as the Witham and Stour. For catchments such as these, additional calibration in the form of automatic optimisation was required in order to obtain a reasonable model. The special form of the PRTF meant it could not be fitted using the IV algorithm. Instead, the TF(2,4,b) parameter set was used to provide an initial set of parameter values from which to start automatic calibration.

In forecast-mode, the parameters of the state-corrected form of the TF model (in which observed flows are used to forecast future flows) were estimated using a recursive least squares algorithm. Again, this is a very simple method of calibration to use and requires little intervention from the user. For the PRTF, only the moving average parameters can be determined in this way, given a t_{peak} value, which determines the autoregressive parameters. This is achieved by embedding recursive least squares estimation of the moving average parameters within a simplex optimisation of t_{peak} . The methods of parameter estimation of TF and PRTF models in forecast-mode have been chosen to correspond to those most commonly used in practice by the Environment Agency.

Both error-prediction and gain-updating (Owens, 1986) can be used with the TF/PRTF models. The time-varying model gain updating scheme outlined in the Part 1 Report has been used in preference to ARMA error-prediction because this scheme is used in practice by the EA. The parameter of the gain-updating scheme has been estimated by automatic optimisation. However, trials suggest that a classical Transfer Function Noise model (a TF

model with ARMA error-prediction) can give as good or better results than gain-updating, for some, especially smaller, catchments.

5.1.7 Grid Model

Unlike other models, which were calibrated using a mixture of manual adjustment and automatic optimisation of model parameters, the Grid Model was calibrated by manual adjustment alone. For a large catchment and a model resolution of 2 km, the Grid Model is extremely slow to calibrate using automatic methods and manual adjustment is preferred. However, the model is reasonably straightforward to calibrate manually, and although the Grid Model is sensitive to initial soil moisture conditions, these can be set explicitly to dry or wet conditions in the parameter file.

5.2 Data Requirements of Models

None of the models differ radically in terms of their basic data requirements, all requiring at least records of rainfall and runoff. The continuous water accounting models require an estimate of potential evaporation. A simple sine curve function over the year has been used in most cases or a standard annual profile, although data derived from weather stations could have been used where available. TF/PRTF models as formulated here do not utilise potential evaporation. Also their focus on the routing phase, without explicit representation of the runoff production phase, arguably allows these models to be calibrated using shorter periods of record and data for individual flood events. However, where the model gain (equivalent to the runoff coefficient) varies with catchment wetness across events this can, at the same time, lead to problems of model identification and be seen as a weakness of some forms of implementation of TF/PRTF model. It should be noted that the RFFS Model Calibration Environment used here allows event data to be used with the different forms of continuous water accounting models, with a water budget maintained across events using daily rainfall records. However, it has been found preferable here to use longer periods of continuous record, typically eight months in duration, for model calibration.

The GRID model utilises additional spatial data in support of model structure definition, making use here of DTM data to define its storage and translation characteristics. It can also utilise weather radar data in its original grid form, as well as the catchment average rainfall used by the other models. Definition of the response zones in the TCM can be assisted by a knowledge of the catchment characteristics. In this study use has been made of the DTM along with the following spatial digital datasets: IH urban areas, IH 100 year flood map (for riparian areas) and Winter Rain Acceptance Potential (WRAP) soil classes. Whilst such data are not essential, they have been found to be helpful in defining the number and the extent of the TCM response zones to be used for a given catchment. Note there is both benefit in a model having minimal data requirements and in being able to utilise additional data where these exist to assist model configuration and/or to improve model performance.

5.3 Guidelines for Model Choice

The following is a set of recommendations for choosing a model for a particular type of catchment.

Models should only be as complex as they need to be. The IEM model may suffice for simple catchments dominated by a fast response. It is quick and easy to calibrate by automatic optimisation and requires a minimum knowledge of catchment characteristics. However, its empirical accounting for losses through a Penman store, operated independently of the quadratic storage, is not a good scientific construct. Preference should be given to models that strive to incorporate mass balance in their formulation. The IEM is not one of these. Poor performance can be expected in simulation-mode when applied to catchments having a complex response, such as the Witham, Stour or Dove. However, forecast-mode results indicate that the updating can considerably improve the performance of the IEM on these catchments, and the IEM is surprisingly successful when applied to large and complex catchments.

The TCM, is best suited to more complex basins in which the hydrograph derives from a variety of areal response zones. It resulted in the best overall model performance in terms of R^2 and can be a powerful model in the hands of an experienced user, but can lead to difficulties and wasted effort otherwise. The model may be simplified through the use of a small number of zones. Initial zone configuration may be eased through the use of a DTM and spatial digital datasets. Since the same model structure is used in each zone, parameter identifiability is poor and behaviour must be forced by the user. Automatic optimisation of model parameters is of limited use. A strategy based on informal visual calibration, initially to daily data over a season or more, will help establish the right order of values for the parameters dominating the water balance. Calibration using a 15 minute time-step can then help refine these parameters and establish those dominating the short-term dynamics.

Models such as the PDM and MCRM provide a good compromise between simplicity and complexity. The range of model structures the PDM accommodates allows it to represent a quite large range of catchment behaviours. It also employs model structures which circumvent the adverse effects on parameter estimation of storage-controlled thresholds on runoff production. Model functions are also parsimonious of parameters. Model performance results indicate that in terms of R^2 performance criterion the PDM is in the second tier of models with the MCRM and NWS, but scored highly when the Threshold CSI is used to compare model performance. Threshold CSI measures the ability of the model to forecast exceedences over a range of flow thresholds accurately so is important in a flood warning context. The MCRM is more complex than the PDM, but is more straightforward to handle than its number of parameters might suggest. It seems to be a reasonable all-rounder, but performs best over catchments such as the Silk Stream and the Dove at the small to medium end of the scale. It is crucial that its initial soil moisture is set correctly at the start of any model run as its performance is very sensitive to initial conditions. When used in practice, the soil moisture component of the MCRM is updated with reference to MORECS data, which might indicate that the soil moisture component of the MCRM may be a weakness. The NWS performs similarly well to the MCRM but is less sensitive to initial soil moisture conditions. It was designed more for use at a daily time-step and may require further model development for optimum operational implementation at a 15 minute time-step.

The TF and PRTF have a simple “black-box” model structure and are simple to understand and apply. As applied here, they give poor results in simulation-mode, but updating with respect to observed flow improves performance considerably, particularly if error prediction is used. The R^2 forecast performance of the TF and PRTF was found to be slightly worse, but comparable to conceptually-based models when applied to seven of the nine study catchments, but significantly worse when applied to the Brue and the Dove. The Threshold CSI statistics obtained through the use of the TF and PRTF are also significantly worse than those obtained from conceptual models, indicating that the TF models are less well able to forecast warning of alert levels. However, these poorer results were substantially corrected when error prediction was used instead of state updating. It is important to note that there are many variants of the basic TF/PRTF models, some incorporating the effects of catchment wetness on runoff production and the parallel routing of baseflow. Further, it has not been possible to emulate those forms most frequently used operationally by the EA which involve event-mode baseflow separation and manual updates to model parameters as the flood develops.

The Grid Model is still primarily a research tool and is slow to run and calibrate on the larger catchments for which it was designed. It is the only model tested here that can use radar rainfall data as input preserved in the original distributed grid square form (rather than aggregated to a rainfall zone or catchment scale). The model is expected to be of most benefit when applied to a catchment with a distributed hydrological response, or distributed rainfall input, such as a localised rainfall event. The lack of convective storms in the datasets used for model assessment has not allowed its potential to be fully explored.

5.4 Summary of Model Choice

By way of summary, Table 5.4.1 presents the main advantages and disadvantages of the eight models considered here. Note that for TF and PRTF models the disadvantages relating to antecedent conditions (catchment wetness effects) and groundwater can be addressed through variants, not tested here, involving a nonlinear loss component and representation of parallel flow processes together with appropriate parameter estimation procedures.

Table 5.4.1 Summary of advantages and disadvantages of the rainfall-runoff models

Model	Advantages	Disadvantages
Thames Catchment Model	<p>Very good simulation and forecast-mode performance.</p> <p>Uses hydrological characteristics of the catchment.</p> <p>Good for catchments with multi-modal response and Chalk catchments.</p>	<p>Uses hydrological characteristics of the catchment.</p> <p>Many parameters with considerable interaction and redundancy, resulting in difficult model identification and possible rogue parameters.</p> <p>No satisfactory state-correction scheme if used with flow routing component.</p>
National Weather Service Model	<p>Easy to calibrate using automatic optimisation.</p> <p>Reasonable simulation and forecast-mode performance.</p>	<p>Complex model with large number of parameters. Fairly insensitive to initial soil moisture conditions.</p> <p>Designed more for use at a daily time-step. May require further model development for practical use at a 15 minute time-step.</p>
Midlands Catchment Runoff Model	<p>Good all-rounder. Complex, but not too complex to use straightforwardly.</p> <p>Good performance for small-medium catchments.</p>	<p>Performance becomes more mediocre for larger catchments and can be too over-sensitive to initial conditions. Operationally the soil moisture store is corrected using MORECS data.</p>
Probability Distributed Moisture Model	<p>Good overall performance for most catchments including Chalk. Very good performance on flow peaks and alert level crossings.</p> <p>State-correction methods applicable.</p> <p>Sophisticated without excessive numbers of parameters.</p>	<p>Experimentation with different structures may be required.</p> <p>Effect of parameter changes can be difficult to anticipate.</p>
Isolated Event Model	<p>Quick and easy to calibrate.</p> <p>Requires minimal knowledge of catchment characteristics.</p> <p>Good for catchments with simple, uni-modal response, and performs well in forecast-mode for most catchments</p>	<p>Poor simulation-mode performance for catchments with complex response.</p> <p>Empirical loss function theoretically weak.</p>

Model	Advantages	Disadvantages
Transfer Function Model	<p>Simple to calibrate using Instrumental Variable (IV) algorithm.</p> <p>Good results when applied to simple, fast catchments without significant groundwater component</p> <p>Quick to run, and requires minimal understanding of hydrology to use.</p>	<p>Requires minimal understanding of hydrology to use.</p> <p>Ineffective when applied to complex catchments with a significant groundwater component.</p> <p>Antecedent conditions only accommodated via updating.</p>
Physically Realisable Transfer Function model	<p>Relatively easy to calibrate using a mixture of IV and automatic optimisation</p> <p>Quick to run, reasonable results on simple, fast catchments without significant groundwater component.</p>	<p>Requires minimal understanding of hydrology to use.</p> <p>Ineffective when applied to complex catchments with a significant groundwater component</p> <p>Antecedent conditions only accommodated via updating.</p>
Grid Model	<p>Designed for use with spatially distributed radar data.</p> <p>Good performance on larger catchments with a distributed hydrological response</p> <p>Physically-based, distributed model structure allows for spatial distribution of slope and landuse in the catchment.</p>	<p>Slow to run, making automatic calibration impractical. (Has been calibrated manually in this study.)</p> <p>Primarily an R&D model requiring further development before being implemented in practice.</p>

6. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

6.1 Summary

This study has used data from nine catchments in England and Wales to assess the performance of eight rainfall-runoff models over a wide range of flow events. The selection of catchments was guided by operational importance, ensuring a mix of hydrological characteristics and choosing at least one catchment from each of the eight EA regions. Three of the catchments were selected for model assessment using weather radar data as a replacement for raingauge data, or as a complement to it via raingauge-calibration. Single site 2 km (5 km beyond 75 km range) data have been used rather than Nimrod corrected data which are only available for a 5 km resolution and since 1998 as a stable product. The rainfall-runoff models chosen for assessment range in complexity from simple transfer function models to more complex models such as the Thames Catchment Model (TCM) and a simple distributed model, the Grid Model. The full set of models, arranged in order of complexity except for the Grid Model treated here as a special case, is:

- Thames Catchment Model (TCM);
- National Weather Service (NWS);
- Midlands Catchment Runoff Model (MCRM);
- Probability Distributed Moisture model (PDM);
- Isolated Event Model (IEM);
- Transfer Function Model (TF);
- Physically Realisable Transfer Function Model (PRTF);
- Grid Model.

Of the eight models, only two are not used operationally in the UK: the NWS and the Grid Model.

Data for between six and ten years for all catchments were made available for model assessment. These data have been quality controlled and used in a split sample testing scheme for model calibration and evaluation. Models have been calibrated on a number of long (eight month) periods in both simulation and updating mode. ARMA error-prediction has been used to update those models for which an explicit state-updating scheme is not available. Models have been evaluated in both simulation and forecast mode using a range of performance criteria. The R^2 statistic measures the overall performance of the model while the Threshold CSI has been used to indicate the ability of a model to correctly forecast the crossing of a range of flow thresholds. Scatter plots of observed and forecast peak flow were used to highlight a model's ability to forecast flow peaks, whilst a selection of hydrographs served to provide a visual indication of model performance. Flow forecasts from different models were compared numerically in the form of fixed-lead time forecasts, and visually as fixed-origin forecasts for significant flood events.

The TF/PRTF models are used operationally in an event-mode with baseflow (defined as the flow at the start of the event) normally subtracted and the model time-step varying with the responsiveness of the catchment. Also, manual adjustment of the model parameters as the flood develops is encouraged. It has not been possible to emulate these features in the model assessment. Individual flood events are not identified and performance statistics are pooled

over every time point of a record which typically extends over an eight month period. Instead, total flow has been used, a fixed model time-step of 15 minutes and an automated method of model parameter adjustment that has been used operationally in Anglian and Southern regions. Parameter estimation of TF/PRTF models used in forecast mode has used procedures which are similar to those in the MATH code employed by the EA. The results and conclusions relating to TF/PRTF models should be interpreted against this background. There are also other variants of TF/PRTF model accommodating the effect of catchment wetness on runoff production and the inclusion of baseflow as a TF model in parallel. Since these are not used operationally by the EA these variants have not been assessed here. Special TF model estimation schemes suitable for identifying parallel slow and fast response systems, such as the simplified refined instrumental variable (SRIV) method (for example, see Young, 1992), have not been used since these do not feature in current EA operational practice.

6.2 Main Results

The main results obtained are summarised below.

Model calibration

Performance results obtained for model calibration were found to be broadly similar to those obtained for model evaluation and therefore are not discussed separately here. However, ease of model calibration is an important consideration and some models were found to be more straightforward to calibrate than others. The TF, PRTF and IEM are very quick and easy to calibrate and require minimal user knowledge of catchment characteristics. At the other extreme, the TCM is a more complex model with a large number of parameters and calibration can be more difficult and time-consuming. However, use of a Digital Terrain Model (DTM) and digital spatial datasets to help identify different TCM response zones can ease the task. The PDM, MCRM and NWS were moderately easy to calibrate: the NWS has a large number of parameters, but is easy to calibrate using automatic optimisation, whilst the PDM and MCRM may be calibrated by a mixture of manual and automatic parameter adjustment. The Grid Model is a distributed model designed for use with spatially distributed radar data. As a result it is slow to run on larger catchments and was calibrated manually.

Model evaluation: simulation-mode results

- (i) There is least variation in model performance across model types for responsive or urbanised catchments such as the Rhondda, Trout Beck, Silk Stream and the Roch. For these catchments reasonable model performance is obtained independent of which model is used (particularly if the TF and PRTF are excluded).
- (ii) Although no one model out-performs the others consistently across all catchments, the TCM performs best overall with the MCRM, PDM and NWS not far behind. On the three catchments for which the Grid Model was evaluated, it performed well, consistently coming second in overall simulation performance. The MCRM was found to be particularly sensitive to initial soil moisture conditions, and initialising the model could be problematic.

- (iii) Transfer function models are used operationally in updating mode and are not expected to perform well under simulation-mode conditions. They have not performed well here in simulation-mode particularly for larger, more complex catchments with a significant groundwater response.
- (iv) Moderate improvements in model performance for the Silk Stream and the Stour may be achieved through the use of calibrated weather radar data instead of rainfall estimates from raingauges. For the Silk Stream the median R^2 statistic increases from 0.635 to 0.672 when calibrated weather radar data are used in place of raingauge estimates; for the Stour, this statistic increases from 0.622 to 0.703. Model results are poor for these catchments when uncalibrated radar data are used.
- (v) Use of both uncalibrated and calibrated radar rainfall estimates employing data from Hameldon Hill radar for flow simulation on the Roch resulted in poor model performance compared to that obtained using raingauge rainfall estimates.
- (vi) The use of a groundwater-modelling component in the TCM and PDM, incorporating daily abstraction data, was successful resulting in an R^2 value of 0.88 for the TCM and 0.87 for the PDM for an eighteen-month period including the Chichester flood event.
- (vii) The Grid Model was used with distributed radar rainfall data on the Stour and Roch. The use of the Grid Model on the Stour resulted in the best overall model performance for that catchment in simulation-mode, and in forecast-mode the results were good in terms of R^2 , but disappointing when the Threshold CSI criterion was used to assess models. Performance on the Roch with radar rainfall estimates was moderate because of the poor quality of radar data from the Hameldon Hill radar.

Model evaluation: forecast-mode results

- (viii) Although no one model out-performs the others consistently across all catchments, the TCM is one of the best performing models in forecast-mode when the R^2 statistic is used to compare models. When the Threshold CSI criterion is used the PDM is very successful, indicating that it is good at forecasting crossing of flow thresholds, such as alarm levels. For a simple model, the IEM is surprisingly successful in forecast-mode particularly in terms of the Threshold CSI.
- (ix) Following a poor performance in simulation-mode, the TF and PRTF improved when state-updating was used. The Threshold CSI statistics obtained from these models were poor compared to those calculated for more conceptually-based models. The TF and PRTF performed better in terms of R^2 , particularly on the Rhondda, Stour, Roch and Witham. Trials suggested that somewhat better forecast performance might be obtained with the use of ARMA error-prediction instead of state-updating for some, especially smaller, catchments.
- (x) Flow forecasts produced for the Silk Stream and the Stour using calibrated radar data were compared to those obtained over the same periods using raingauge data. Overall, based on the R^2 performance criterion, the results indicate that when the radar was

working well, use of radar data gave as good, or sometimes slightly better results as those obtained using raingauge data.

6.3 General Conclusions

This study has assessed eight rainfall-runoff models in terms of a variety of performance measures in both simulation and forecast mode, and more subjectively in terms of “ease of use”. The results are complex, and predictably, no one model has been found to be best for all types of catchment and assessment criteria. However, some general conclusions may be drawn from the study which are set down below.

- (i) Rainfall-runoff model performance on small or quickly responding catchments with a significant urban area is moderately insensitive to the type of model used. For catchments such as these, a simple, easy-to-use model is most appropriate, such as the IEM, PDM, or the TF model. The MCRM works well on small-to-medium size catchments, but is very sensitive to initial soil-moisture conditions, and is updated operationally with MORECS soil-moisture data.
- (ii) On larger, or slowly responding catchments, the TCM has been shown to be very effective in providing a good overall model performance. The TCM is a complex model with a large number of parameters and can be a challenge to calibrate, but once successfully calibrated, model results in both simulation- and forecast-mode can be very good. The PDM and IEM also provide good overall model performance and can often be better than other models, including the TCM, when judged using the Threshold CSI.
- (iii) The data requirements of different models do not differ radically, requiring at least records of rainfall and runoff, and do not exert a strong influence on model choice. Arguably, TF/PRTF models might require less extensive records if calibrated on isolated events but this will be less true for variants which aim to accommodate catchment wetness effects. There is both benefit in a model having minimal data requirements and in being able to utilise additional data (grid square weather radar, DTM and spatial datasets) where these exist to assist model configuration and/or to improve model performance.
- (iv) When good quality radar data are available, calibration of the radar data with reference to a network of raingauges is essential. Once good quality radar data has been calibrated in this way, it can lead to as good, and sometimes better flow forecasts than those obtained using raingauge data alone. However, if the radar data are of poor quality (for example, Hameldon Hill), use of raingauge data is likely to result in more reliable flow forecasts.
- (v) The TF and PRTF models as applied here performed poorly in simulation-mode; these models are used operationally in forecast-mode where they are updated using recent observations of flow. However, their performance in forecast mode was also found to be poor, particularly for larger catchments with a significant groundwater component. Trials suggested that forecast performance of TF models is improved by using error prediction instead of state updating for some, particularly smaller, catchments. When

updating is carried out operationally the model response can be adjusted manually to take into account catchment soil moisture conditions; this may improve upon the results reported here. Also, TF/PRTF models are operated in practice in event-mode with flow at the start of the event being commonly subtracted as “baseflow” prior to modelling. The continuous form of assessment used here, with events not explicitly defined, has precluded the emulation of this feature. This, together with model calibration on single short events, differs from the approach employed here and the results need to be interpreted against this background. There are also variants which incorporate the effect of catchment wetness on runoff production and incorporate baseflow as an integral model component (see, for example, Lees, 2000). These have not been assessed as they do not feature in EA present practice.

6.4 Operational Recommendations

This section provides operational recommendations for rainfall-runoff modelling for different EA regions. Due to the heterogeneity of some regions it is not always easy to provide strong or clear guidance, and results from some types of catchment can be more relevant to particular areas rather than specific EA regions. For example, recommendations deriving from the Lavant catchment apply particularly to catchments on the Chalk, encompassing the Yorkshire Wolds, East Anglian Heights, Chilterns and the North and South Downs. In response to these concerns, operational recommendations are also presented in terms of catchment size and dominant characteristics.

Table 6.4.1 presents operational recommendations for each EA region alongside a list of models currently in operational use. Where a region is familiar with a certain model this has been used as a reason for recommending retaining the model unless the performance assessment strongly favours use of another model.

Table 6.4.2 provides operational recommendations in terms of catchment size and dominant characteristics.

Table 6.4.1 Summary of current operational use of rainfall-runoff models by EA regions in the UK and operational recommendations for each

EA Region	Models currently in use operationally	Operational recommendations
Anglian	TF, PDM (Lincoln)	<p>The TCM was shown to perform best overall on the Witham, which is a large lowland catchment typical of the region. The PDM, of which there is some familiarity in Anglian region, also performed well.</p> <p>Recommend Anglian region to consider greater use of PDM or similar conceptually based rainfall-runoff model.</p>
Midlands	MCRM	<p>This study showed that the most successful model on the Dove was the MCRM, whilst on the Stour, the TCM or PDM were best.</p> <p>Soil moisture accounting in MCRM shown to be a weakness.</p> <p>Recommend Midlands region consider use of alternative conceptually based model (e.g. TCM, PDM) in larger catchments.</p>
North East	PDM	<p>Good results were obtained for Trout Beck using the TCM and IEM. However, the complex TCM model formulation is probably wasted on small upland catchments. Recommend investigating use of IEM, or continued use of PDM, which performed reasonably well on Trout Beck.</p>
North West	TF/PRTF	<p>Best results for the Roch were obtained using TCM and PDM, but most models, including the TF and PRTF, performed reasonably well on this fast, urbanised catchment.</p> <p>Suggest use of PDM may benefit accurate forecasting of flow peaks and alert level crossings. Where TF models are in use, consideration of ARMA error prediction might be beneficial.</p>
Southern	TF (PDM currently being trialled on the Medway to replace TF)	<p>Trials of the PDM and TCM incorporating a new groundwater-modelling component were successful on the Lavant. The new model components currently use daily abstraction data, and well-level data could be easily incorporated.</p> <p>Recommend operational trials of TCM or PDM with new groundwater modelling component. General use of TCM or PDM may bring benefits.</p>
South West	TF/PRTF	<p>Best model performance on the Brue achieved through use of TCM. Use of IEM/PDM in forecast mode lead to more accurate forecast peak flow and flow threshold crossings.</p> <p>Recommend investigation of use of TCM, PDM or IEM in the region. Where TF models are in use, consideration of ARMA error prediction might be beneficial.</p>
Thames	TCM, IEM, PDM	<p>TCM and MCRM performed best on the Silk Stream.</p> <p>Recommend Thames region to continue use of the TCM and trials of PDM.</p>
Welsh	ISO Many areas have no operational flood-forecasting system.	<p>Most models, including the IEM, perform well on the Rhondda in both simulation and forecast mode.</p> <p>Recommend increased use of flood-forecasting in Welsh Region. The PDM provides good compromise between complexity and accurate model performance, although use of IEM should lead to satisfactory model performance.</p>

Table 6.4.2 Operational recommendations for rainfall-runoff models in terms of catchment size and dominant characteristics

Catchment size	Catchment characteristics	Example	Operational recommendations
Small	Upland impervious, rural or urban	Trout Beck	Good results were obtained for Trout Beck using the TCM and IEM. However, the complex TCM model formulation is probably wasted on small upland catchments. Recommend use of model with a simple structure such as the IEM, or a TF model formulation.
	Urban clay	Silk Stream	TCM and MCRM performed best on the Silk Stream. Recommend use of TCM for its urban zoning, though simpler models such as the IEM, PDM or a TF formulation may be adequate.
Medium	High relief, impervious	Rhondda at Trehafod	Most models, including the IEM, perform well on the Rhondda in both simulation and forecast mode. The PDM provides good compromise between complexity and accurate model performance, although use of the simpler IEM should also lead to good model performance.
	High relief, mixed geology	Dove at Izaak Walton	This study showed that the most successful model on the Dove was the MCRM, although soil moisture accounting in MCRM thought in general to be a weakness. Recommend use of alternative conceptually based models (e.g. TCM, PDM).
	Lowland permeable (Chalk), groundwater dominated	Lavant at Graylingwell	Trials of the TCM and PDM incorporating a new groundwater-modelling component were successful on the Lavant. The new model components currently use daily abstraction data, and well-level data could be easily incorporated. Recommend operational trials of TCM or PDM with new groundwater modelling component.
	Modest relief, rural	Stour at Shipston, Brue at Lovington	Good performance may be achieved through use of TCM/PDM/IEM. Where TF models are in use, consideration of ARMA error prediction might be beneficial.
	Modest relief, significant urban fraction	Roch at Blackford Bridge	Use of the TCM, with its urban zoning may be most appropriate, although responsive catchments can be modelled reasonably well using most models, including TF formulations
Large	Lowland, clay	Witham at Claypole Mill	The TCM was shown to perform best overall on the Witham. The simpler PDM model also performed well.
	Lowland, Chalk		TCM or PDM expected to perform well on large lowland catchments. New groundwater-modelling components to TCM/PDM may be of benefit

6.5 Opportunities for Further Work

Opportunities for further work, identified in this study, are set out below:

- (i) Following the good overall performance of the TCM and the IEM (which have been unified in the Penman Store Model or PSM), there are opportunities for research into the optimum number of TCM response zones for catchments of differing size and complexity guided by DTMs and digital spatial datasets. Certain catchments may be just as accurately modelled using two response zones as four, which would simplify model calibration.
- (ii) Following the development of new methods of raingauge weighting for calculating catchment average rainfall, there is now the opportunity to investigate the use of Thiessen/SAAR/elevation weights across a range of catchments to develop some simple rules to determine which weighting schemes are appropriate for different circumstances.
- (iii) Investigation of Cleve Hill radar data on no-rain days revealed a large area of low-level clutter covering approximately one quarter of the 2 km radar field. Techniques which remove isolated pixels of transient clutter fail to remove the more spatially coherent clutter. However, the time-varying clutter field can be readily suppressed using a simple threshold cut-off which could be readily incorporated into a radar pre-processing scheme.
- (iv) Performance of the Hameldon Hill radar should be reviewed following the replacement of the radar magnetron. This upgrade may now allow data from this radar to be used with greater confidence than the results for the Roch obtained here suggest.
- (v) Assessment of forecast performance using Nimrod corrected weather radar data.
- (vi) A rigorous assessment of the use of radar rainfall estimates in flow-forecasting has been made difficult by the lack of significant convective rainfall events over the study catchments. Attempts had been made to obtain radar data for convective storms, but in practice much of the data are missing. Radar rainfall estimates are expected to be most beneficial during localised convective rainfall which may not be registered by raingauges. Further research should assess the use of radar data in flow forecasting during convective rainfall events. The assessment should include use of the Grid Model, which was specifically designed for use with spatially distributed radar data.
- (vii) Further development of the TCM and PDM models for permeable catchments with groundwater abstractions, incorporating well level data to support model calibration and updating.
- (viii) This study has highlighted opportunities for research into new state-updating schemes for models that currently do not have them, or further development of schemes for existing models, such as the use of state-updating in the TCM.
- (ix) Difficulties of model identification experienced with TF/PRTF models are considered to stem from an inappropriate model structure which, in its basic form, assumes a

constant gain (or runoff coefficient) and in which variable baseflow is not accommodated as an integral component. Adopting forms which account for the effect of catchment wetness on runoff production and incorporate variable baseflow as an integral part, along with appropriate estimation procedures, would be a step in the right direction. Whilst, these do not feature in EA current practice, steps are being taken in Southwest region to use the effective rainfall transformation used in IHACRES (Jakeman *et al.*, 1990) as one way of accounting for catchment wetness.

The study has focused on a specific set of eight “brand name” models. There is clearly scope for research into the components of these models and their configuration which might help in the formulation of new, improved models or toolkits of models. For example, the spatial zoning used in the TCM to represent different responses has application to other models, albeit at the expense of increased complexity and difficulties of calibration. The use of rainfall zones, available in the TCM model but not invoked here, may offer further scope for improvement, particularly for larger catchments. The efficient PRTF parameterisation may be of value as a component of a conceptual catchment model used to represent flow translation to the basin outlet. Variants of the TF/PRTF models which incorporate catchment wetness effects on runoff production and incorporate baseflow as an integral component, with appropriate parameter estimation schemes, need to be assessed. However, their link with models such as the PDM – which have emerged as toolkits encompassing TF models in conceptual form, nonlinear storage components and soil moisture accounting procedures – needs to be appreciated. Exhaustive data-based model intercomparisons of “brand name” models should be initiated with caution and only undertaken when necessary against a strong theoretical understanding of what is to be achieved. New variants of the PDM theory (e.g. Senbeta *et al.* 1999) require to be evaluated. There are also other models not included here, some reviewed in the Part 1 Report, that might bring benefits to flood forecasting. For example, the ISO (Input-Storage-Output) models may offer a similar level of performance to the IEM as well as sharing its simplicity and ease of use. Emerging technologies – such as neural networks, nearest neighbour forecasting and fuzzy logic approaches reviewed in the Part 1 Report – provide further prospects for research aimed at improving flood forecasting via modelling, updating and forecast uncertainty estimation. The model assessment framework of models and data developed here provides a valuable test-bed within which new forecasting methodologies can be judged in the future.

REFERENCES

- Bailey, R.A. and Dobson, C. (1981) Forecasting for floods in the River Severn catchment. *J. Inst. Water Eng. and Scientists*, 35(2), 168-178.
- 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.
- Box, G.E.P. and Jenkins, G.M. (1970) *Time series analysis forecasting and control*, 553pp, Holden-Day.
- Burnash, R.J.C., Ferral, R.L. and McGuire, R.A. (1973) *A generalized streamflow simulation system: conceptual modelling for digital computers*, Report of the Joint Federal State River Forecast Centre, U.S. National Weather Service and California Department of Water Resources, Sacramento.
- Carrington, D.S., Moore, R.J. (1995) *Extension of the London Weather Radar Local Calibration Procedure to the Entire Thames Basin*. Contract Report to NRA Thames Region, Institute of Hydrology, 30pp.
- Cluckie, I.D. and Owens, M.D. (1987) Real-time rainfall-runoff models and use of weather radar information. In: V.C. Collinge and C. Kirby (eds), *Weather Radar and Flood Forecasting*, 171-190, J. Wiley.
- Gelb, A. ed. (1974) *Applied optimal estimation*, 374 pp, MIT Press.
- Gill, P.E., Murray, W. and Wright, M.H. (1981) *Practical Optimisation*, Academic Press.
- Golding, B.W. (1998) Nimrod: A system for generating automated very short range forecasts. *Meteorol. Appl.*, 5, 1-16.
- Greenfield, B.J. (1984) *The Thames Water Catchment Model*. Internal Report, Technology and Development Division, Thames Water, UK.
- Han, D. (1991) *Weather radar information processing and real-time flood forecasting*. PhD Thesis, Water Resources Research Group, Department of Civil Engineering, University of Salford, 277pp.
- Hardy, R.L. (1971) Multiquadric equations of topography and other irregular surfaces. *J. Geophys. Res.*, 76(8), 1905-1915.
- Isermann, R. (1980) Practical aspects of process identification. *Automatica*, 16, 575-587.

Jakeman, A.J. and Hornberger, G.M. (1993) How much complexity is warranted in a rainfall-runoff model? *Water Resources Research*, 29(8), 2637-2649.

Jakeman, A.J., Littlewood, I.G. and Whitehead, P.G. (1990) Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. *J. Hydrology*, 117, 275-300.

Jakeman, A.J. and Young, P.C. (1984) Recursive filtering and smoothing procedures for the inversion of ill-posed causal problems. *Utilitas Mathematica*, 25, 351-376.

Jazwinski, A.H. (1970) *Stochastic processes and filtering theory*, Academic Press, 376 pp.

Jones, D.A. and Moore, R.J. (1980) A simple channel flow routing model for real-time use. Hydrological Forecasting, Proc. Oxford Symp., *IAHS-AISH Publ. No. 129*, 397-408.

Lees, M.J. (2000) Data-based mechanistic modelling and forecasting of hydrological systems. *J. of Hydroinformatics*, 02.1, 15-34.

Lees, M., Young, P.C. and Ferguson, S. (1993) Flood Warning, Adaptive. In: P.C. Young (ed.), *Concise Encyclopaedia of Environmental Systems*, Pergamon.

Moore, R.J. (1980) *Real-time forecasting of flood events using transfer function noise models: Part 2*; Contract report to Water Research Centre, 115pp, Institute of Hydrology.

Moore, R.J. (1982) Transfer functions, noise predictors and the forecasting of flood events in real-time. In: Singh, V.P. (ed.), *Statistical analysis of rainfall and runoff*, Water Resources Publ., 229-250.

Moore, R.J. (1982) Advances in real-time flood forecasting practice, Invited paper, *Symposium on Flood Warning Systems*, Winter Meeting of the River Engineering Section, The Institution of Water Engineers and Scientists, 23pp.

Moore, R.J. (1999) Real-time flood forecasting systems: Perspectives and prospects. In: R. Casale and C. Margottini (eds), *Floods and landslides: Integrated Risk Assessment*, Chapter 11, 147-189, Springer.

Moore, R.J., Austin, R.M., Carrington, D.S. (1993) *Evaluation of FRONTIERS and Local Radar Rainfall Forecasts for use in Flood Forecasting Models*. R&D Note 225, Research Contractor: Institute of Hydrology, National Rivers Authority, 156pp.

Moore, R.J., Bell, V.A., Roberts, G.A. and Morris, D.G. (1994) *Development of distributed flood forecasting models using weather radar and digital terrain data*. R&D Note 252, Research Contractor: Institute of Hydrology, National Rivers Authority, 144pp

Moore, R.J., Hotchkiss, D.S., Jones, D.A. and Black, K.B. (1991) *London Weather Radar Local Rainfall Forecasting Study: Final Report*. Contract Report to the National Rivers Authority Thames Region, Institute of Hydrology, September 1991, 124 pp.

Moore, R.J. and Jones, D.A. (1978) An adaptive finite-difference approach to real-time channel flow routing. In G.C. Vansteenkiste (ed.), *Modelling and Control in Environmental Systems*, North Holland.

Moore, R.J., Jones, D.A., Black, K.B., Austin, R.M., Carrington, D.S., Tinnion, M. and Akhondi, A. (1994b) RFFS and HYRAD: Integrated systems for rainfall and river flow forecasting in real-time and their application in Yorkshire. "Analytical techniques for the development and operations planning of water resource and supply systems", BHS National Meeting, University of Newcastle, 16 November 1994, *BHS Occasional Paper No. 4*, British Hydrological Society, 12pp.

Moore, R.J., May, B.C., Jones, D.A. and Black, K.B. (1994a). Local calibration of weather radar over London. In: M.E. Almeida-Teixeira, R. Fantechi, R. Moore and V.M. Silva (eds), *Advances in Radar Hydrology*. Proc. Int. Workshop, Lisbon, Portugal, 11-13 November 1991, Report EUR 14334 EN, European Commission, 186-195.

Moore, R.J., Watson, B.C. Jones, D.A. and Black, K.B. (1989) *London Weather Radar Local Calibration Study: Final Report*. Contract report prepared for the National Rivers Authority Thames Region, 85pp, September 1989, Institute of Hydrology.

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

Moore, R.J. and Weiss, G. (1980a) Recursive parameter estimation of a non-linear flow forecasting model using the extended Kalman filter, in O'Connell, P.E. (ed.), *Real-time hydrological forecasting and control*, Proc. 1st Int. Workshop, July 1977, pp 264, Institute of Hydrology.

Natural Environment Research Council (1975) *Flood Studies Report*, Volume 1 Hydrological Studies, 550pp.

Nelder, J.A. and Mead, R. (1965) A simplex method for function minimisation, *Computer Journal*, 7, 308-313.

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

Owens, M.D. (1986) *Real-time flood forecasting using weather radar data*. Ph.D. Thesis, Dept. Civil Engineering, University of Birmingham, 224pp.

Owens, M.D. (1987) Rainfall runoff models for flood forecasting. *National Hydrology Symposium*, University of Hull, 14-16 September 1987, British Hydrological Society, Poster Paper, 3pp.

Page, C. (1991) *Flow Forecasting System User Guide*, NRA Anglian Region, 13pp.

Penman, H.L. (1949) The dependence of transpiration on weather and soil conditions, *J. Soil Sci.*, 1, 74-89.

- Pollard, O. (undated) *Development of a flow forecasting model for the Leigh Barrier Scheme, River Medway*. Technical Report submission for IWEM Membership, 11pp plus figures.
- Powell, S.M. and Cluckie, I.D. (1985) On the sampling interval of discrete transfer function models of the rainfall-runoff process. *7th IFAC/IFORS Symp. on Identification and System Parameter Estimation*, York, 1119-1124.
- Senbeta, D.A., Shamseldin, A.Y. and O'Connor, K.M. (1999) Modification of the probability-distributed interacting storage capacity model. *J. Hydrology*, 224, 149-168.
- Thiessen, A.H. (1911) Precipitation for large areas. *Monthly Weather Review*, 39, 1082-1084.
- Wallingford Water (1994) *A flood forecasting and warning system for the River Soar: Stage 2 Report*, contract report to the National Rivers Authority Severn Trent Region, November 1994, 91 pp plus Appendices, Wallingford Water, Wallingford, UK.
- Wedgwood, O. (1993) *A knowledge based approach to modelling fast response catchments*. PhD thesis, Water Resources Research Group, University of Salford, 318pp.
- Wood, S.J., Jones, D.A., Moore, R.J. (2000) Climatological and dynamic calibration of radar data for hydrological use. HYREX Special Issue, *Hydrology and Earth System Sciences*, submitted.
- Young, P.C. (1974) Recursive approaches to time-series analysis. *Bull. Inst. Maths. Appl.*, 10, 209-224.
- Young, P.C. (1992) Parallel processes in hydrology and water quality: objective inference from hydrological data. In: Falconer, R.A. (ed.), *Water Quality Modelling*, Ashgate, Vermont, 10-52.

APPENDIX A: RADAR CALIBRATION USING MULTIQUADRICS

A.1 Introduction

This Appendix reviews the theory of radar calibration using multiquadric surface fitting to the calibration factors formed from the adjusted ratios of gauge to coincident radar pixel values. Initially the method used in HYRAD (Moore *et al.*, 1994b) is outlined as background to multiquadrics and to document the method used for the Silk Stream catchment. Then the modified form used for the remaining study catchments is presented.

A.2 HYRAD Local Calibration Procedure

The surface fitting method employed by the HYRAD Local Calibration procedure is based on an extended form of the multiquadric presented by Hardy (1971). First, let z_i be the calibration factor values defined at the n raingauge locations, having grid coordinates $\underline{x}_i=(u_i,v_i)$. Here, the calibration factor is defined as the ratio of the raingauge value, R_g , to the coincident radar pixel value, R_r , such that

$$z_i = \frac{R_g + \varepsilon_g}{R_r + \varepsilon_r} \quad (\text{A.1})$$

where ε_g and ε_r are small positive constants introduced to ensure that the calibration factor is sensibly defined for zero values of radar rainfall. The values used here are $\varepsilon_g=\varepsilon_r=1$ mm with the calibration factor values calculated over intervals of 15 minutes.

The multiquadric calibration surface is defined as the weighted sum of n distance, or basis functions centred on each gauge; that is

$$s(\underline{x}) = \sum_{j=1}^n a_j g(\underline{x} - \underline{x}_j) + a_0 \quad (\text{A.2})$$

where $\{a_j, j=0,1,2,\dots,n\}$ are parameters of the surface. The distance function used is the simple Euclidean distance

$$g(\underline{x}) = \|\underline{x}\| = \sqrt{(u^2 + v^2)} \quad (\text{A.3})$$

which corresponds to building up the surface from a set of n right-sided cones, each centred on one of the n raingauge locations.

Formally, estimation of the a_j weights is achieved as follows. Equation (A.2) for

$$s(\underline{x}_i) = \sum_{j=1}^n a_j g(\underline{x}_i - \underline{x}_j) + a_0 = z_i \quad (i = 1,2,\dots,n) \quad (\text{A.4})$$

expressed in matrix form is

$$\underline{G}\underline{a} + a_0\underline{1} = \underline{z} \quad (\text{A.5})$$

where \underline{G} is an n by n matrix with the (i,j) 'th element given by $G_{ij}=g(\underline{x}_i-\underline{x}_j)$, $\underline{1}$ is a unit vector of order n , and \underline{z} is the vector containing the n calibration factor values. To avoid anomalies in the surface form away from n raingauge locations, a requirement for flatness at large distances is imposed through the constraint

$$\underline{a}^T \underline{1} = 0. \quad (\text{A.6})$$

For the Euclidean distance function of cone type this constraint corresponds to a requirement of zero-slope with increasing distance from the raingauge network. Solution of equation (A.5) subject to constraint (A.6) for the weighting coefficients gives

$$a_0 = (\underline{1}^T \underline{G}^{-1} \underline{z}) / (\underline{1}^T \underline{G}^{-1} \underline{1}) \quad (\text{A.7})$$

$$\underline{a} = \underline{G}^{-1} (\underline{z} - a_0 \underline{1})$$

An important feature of the HYRAD Local Calibration procedure is the forming of a conservative calibration factor surface by adopting a fitting method which allows the surface to depart from the actual calibration factor values. This is achieved by allowing the Euclidean distance $g(\underline{x}_i-\underline{x}_i)=g(\underline{0})$, normally zero, to take a value $-K$. The result is a surface which passes within a distance $a_i K$ of the calibration factor value for the i 'th raingauge; K is referred to as the offset parameter. The problem of discontinuities is avoided by using this form in the estimation of the weights, a_j , and using the normal form in calculating the surface values for radar calibration of the full field. The constraint of equation (A.6) ensures that the "errors", introduced by using $g(\underline{0})=0$ in equation (A.2) (and not $-K$) when forming the surface for calibration, add up to zero.

A.3 Hybrid Calibration Procedure

The new Hybrid calibration procedure is based on first removing the mean field bias from the radar field. A special form of multiquadric surface is then fitted to the calibration factors, calculated from the bias-adjusted radar rainfall field, which tends to unity with increasing distance from the raingauge locations used for calibration.

The mean field calibration factor bias is calculated as

$$b = \frac{1}{n_g} \sum \left(\frac{1}{n_i} \sum \frac{R_g}{R_r} \right) \quad (\text{A.8})$$

where R_g and R_r denote the raingauge and coincident radar value respectively. It is seen from (A.8) that the ratio of these are summed over n_i time periods to obtain the climatological mean bias for each radar pixel containing a raingauge. For n_g raingauges in the field then the mean

of these provides an estimate of the climatological mean field bias, b , of the calibration factors. (Additional notation on R_g and R_r to indicate the time period and gauge are omitted for simplicity.) In practice, values of the ratios R_g/R_r are only included in the summation when both R_g and R_r are at least 1 mm in magnitude. This helps suppress any effects of discretisation errors and the influence of anaprop.

The calibration factors are calculated as before but using the radar pixel values scaled by the mean field bias of the calibration factors. In this case it is sensible to adopt a form of multiquadric surface which tends to a fixed value of 1 with increasing distance from the raingauge locations; in practice, unity is subtracted from the calibration factor values so that the fixed value required is zero. This requirement is met by replacing the normal Euclidean measure of distance by the exponential distance function

$$g(\underline{x}) = \exp(-\|\underline{x}\|/\ell) \quad (\text{A.9})$$

where P is the scaling length parameter. For this distance measure, allowing the surface to depart from the calibration factor (less unity) values is achieved by setting $g(\underline{0})=1+K$. Also, $a_0=0$ and \underline{a} is given simply by

$$\underline{a} = \underline{G}^{-1} \underline{z}. \quad (\text{A.10})$$

A.4 Application to the Study Catchments

Table A.1 presents the results of mean field bias calculations for raingauges over the Stour catchment using the Clee Hill radar and over the Roch catchment using the Hameldon Hill radar. Calculations for different periods allow the temporal stability of the bias to be assessed. Periods used for model calibration and evaluation, and where only 5 km data are available, are identified separately in the table. The average bias across periods and gauges for a given catchment is used in the Hybrid calibration procedure. The values used are 2.518 and 1.359 for the Stour (Clee Hill radar) and the Roch (Hameldon Hill radar) catchments respectively.

The scaling length parameter, P , and the offset parameter, K , of the multiquadric surface have been estimated by removing one raingauge at a time and using the Hybrid calibration procedure to estimate its value. The errors from this selective deletion procedure are pooled to form a root mean square error criterion and combinations of P and K values trialed to obtain a minimum criterion value. Since no more than two gauges are available for each catchment, the scope for parameter optimisation was limited. Initial estimates were obtained by analysing the correlation distance function of calibration factors obtained from the HYREX dense raingauge network over the Brue catchment (Wood *et al.*, 2000). The final estimates used are $P=10000$ km and $K=0.03$ for Clee Hill radar and $P=70$ km and $K=0.03$ for Hameldon Hill radar.

Table A.1 Mean field bias of radar rainfall over catchments for different periods

a) Stour catchment and Clee Hill radar

Period	Chipping and Shipston raingauges		Chipping raingauge		Shipston raingauge	
	Mean field bias	Number of values	Mean field bias	Number of values	Mean field bias	Number of values
<i>Calibration and evaluation periods: 5 km or 2 and 5 km data</i>						
Sep. 92	2.718	125	2.703	70	2.737	55
Jan. 93	2.443	84	2.516	50	2.335	34
Jan. 94	2.703	132	2.782	63	2.631	69
Mar. 98	2.482	100	2.793	52	2.144	48
Apr. 98	2.603	238	3.053	124	2.111	114
<i>Other periods: composite 2 and 5 km data</i>						
Feb. 90	1.527	83	1.281	48	1.863	35
Dec. 91	2.072	12	1.785	7	2.474	5
May 92	2.384	7	2.085	4	2.784	3
Jun. 92	3.537	14	3.571	9	3.477	5
Dec. 92	2.233	99	2.214	55	2.256	44
Apr. 93	2.946	45	3.155	25	2.685	20
<i>Other periods: 5 km data only</i>						
Dec. 93	3.359	96	3.490	48	3.228	48
Mar. 94	-	-	-	-	-	-
Apr. 94	2.8	108	3.235	43	2.512	65
Jun. 94	2.25	15	1.673	1	2.292	14
Sep. 94	2.112	231	2.225	120	1.991	111
Jan. 95	2.572	184	2.760	88	2.399	96

b) Roch catchment and Hameldon Hill radar

Period	Spring Mill and Kitcliffe raingauge		Spring Mill raingauge		Kitcliffe raingauge	
	Mean field bias	Number of values	Mean field bias	Number of values	Mean field bias	Number of values
<i>Calibration and evaluation periods: 5 km or 2 and 5 km data</i>						
Dec. 91	1.1699	294	1.5301	157	1.1893	137
Dec. 93	1.3383	783	1.2358	446	1.474	337
Apr. 94	1.2613	239	1.3878	121	1.316	118
Sep. 94	1.670	239	1.7747	115	1.5738	124
May 97	1.4425	179	1.3835	88	1.4996	91
<i>Other periods: composite 2 and 5 km data</i>						
Jan. 90			no gauge data			
Jan. 92	-	-	-	-	-	-
Nov. 96	1.4656	143	1.4842	75	1.4450	68
<i>Other periods: 5 km data only</i>						
Sep. 93	1.2283	60	0.8283	30	1.6282	30
July 94	1.339	54	1.2126	32	1.5227	22

APPENDIX B: CATCHMENT MAPS

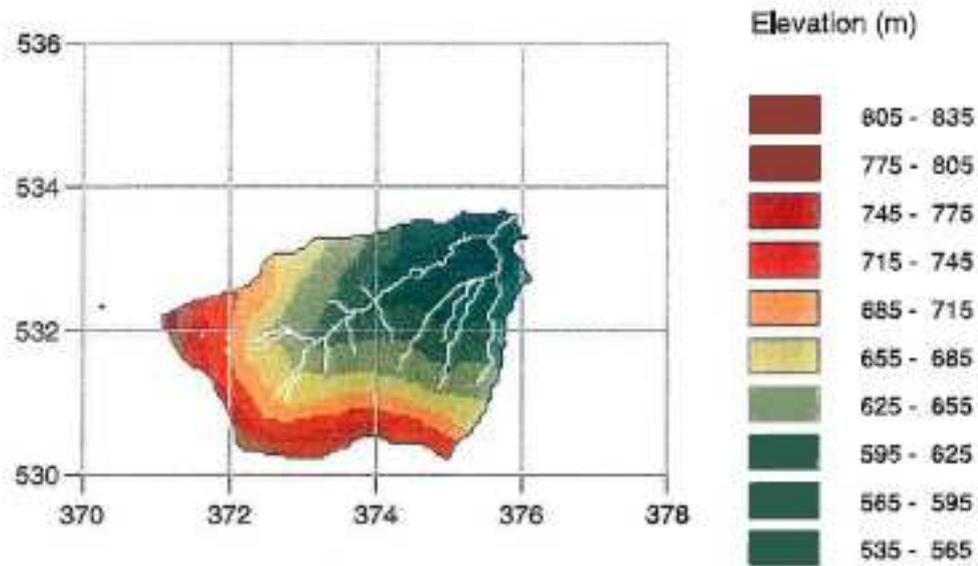


Figure B.1.1 Map of catchment relief: Trout Beck.

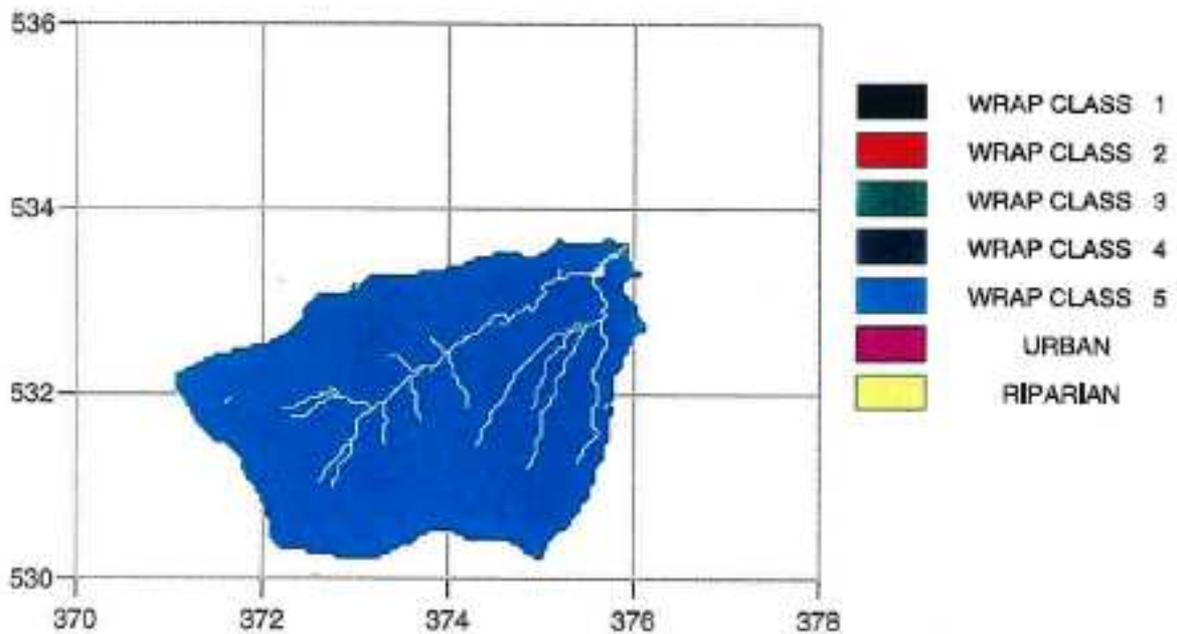


Figure B.1.2 Map of DTM-derived TCM response zones: Trout Beck.

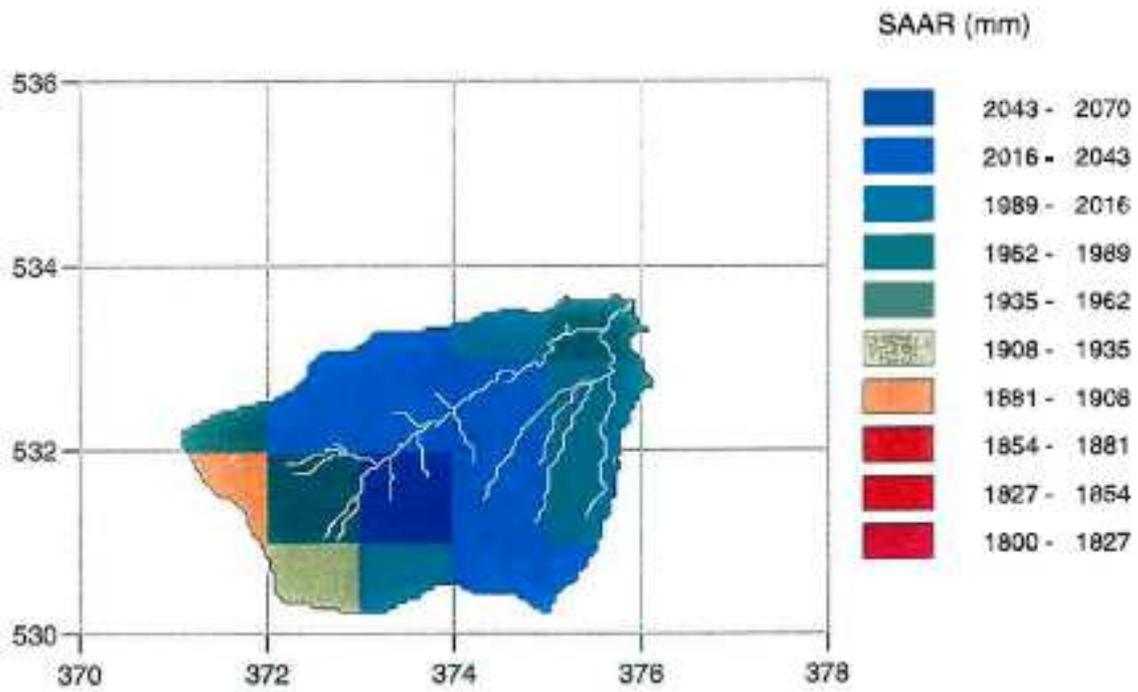


Figure B.1.3 Map showing distribution of SAAR across the catchment: Trout Beck.

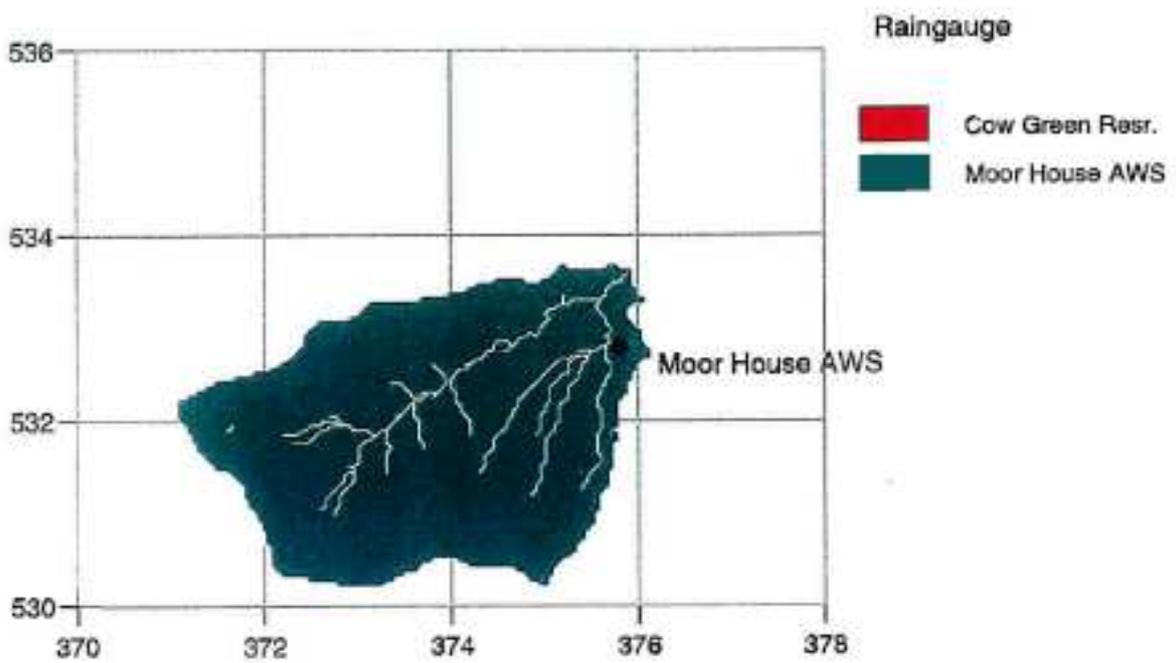


Figure B.1.4 Areas used to define raingauge weights: Trout Beck.

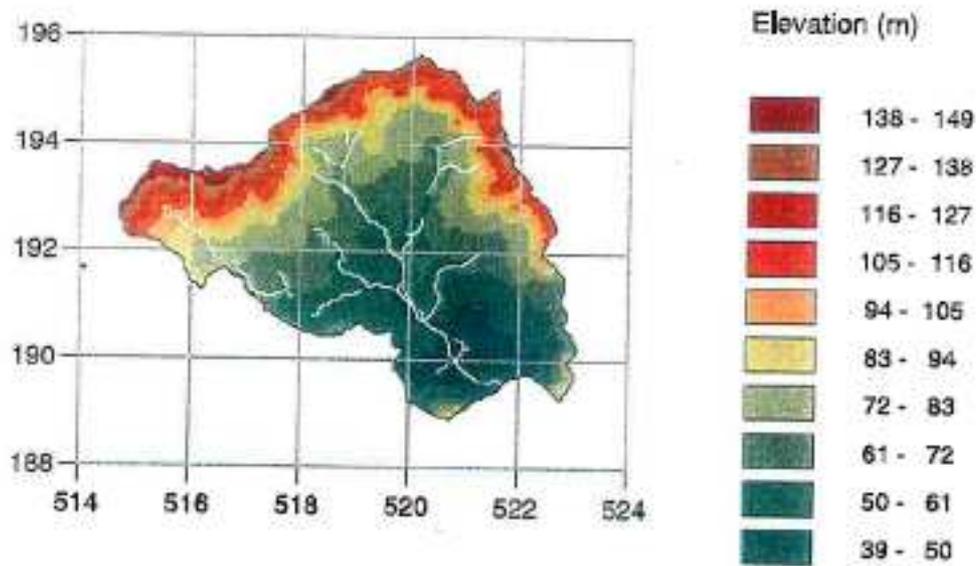


Figure B.2.1 Map of catchment relief: Silk Stream.

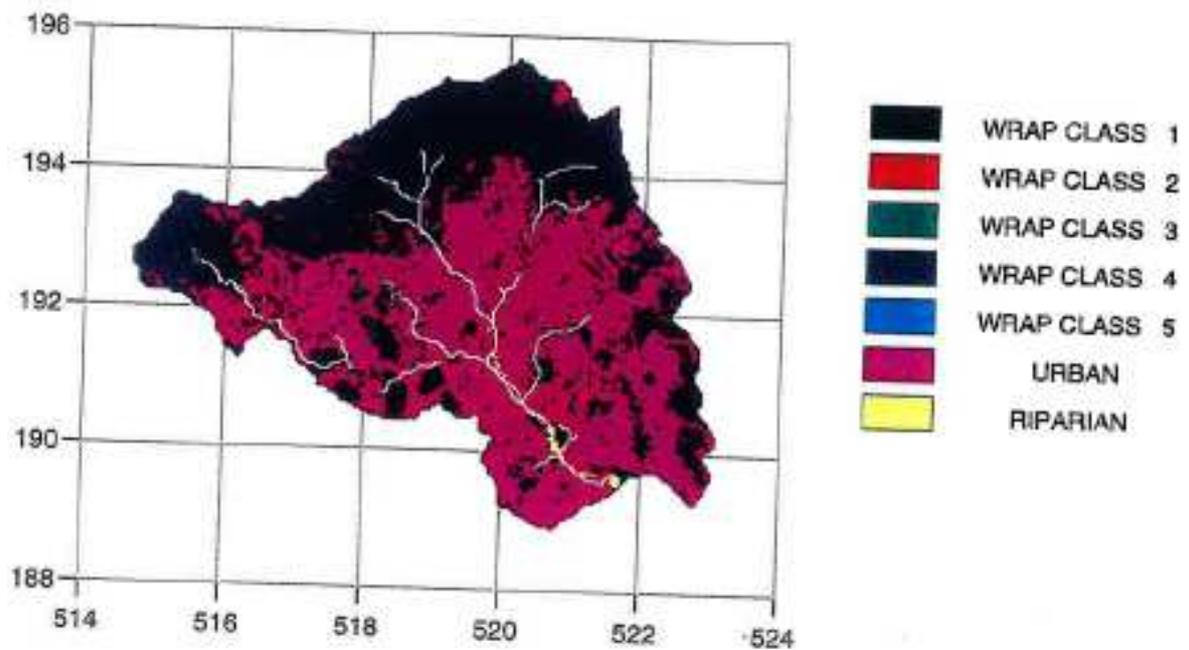


Figure B.2.2 Map of DTM-derived TCM response zones: Silk Stream.

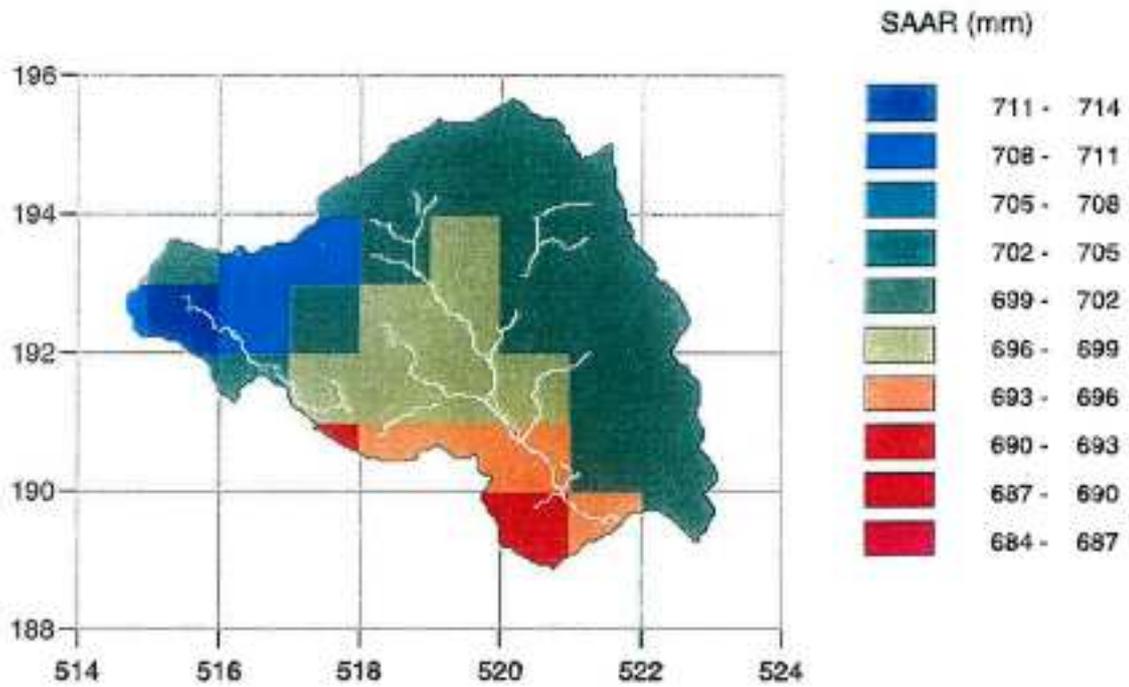


Figure B.2.3 Map showing distribution of SAAR across the catchment: Silk Stream.

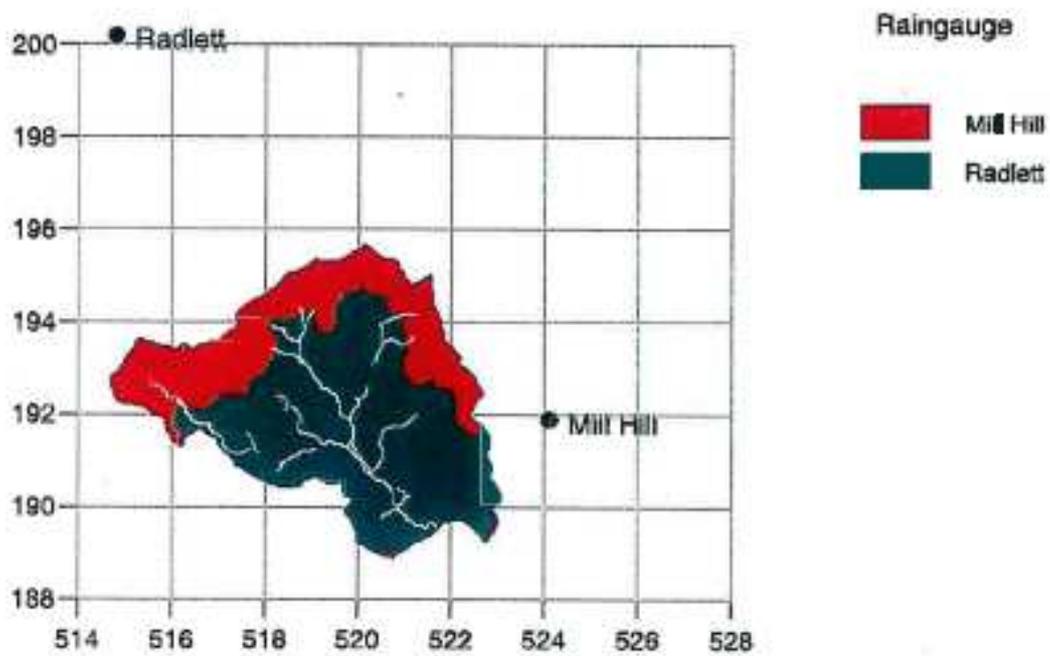


Figure B.2.4 Areas used to define raingauge weights: Silk Stream.

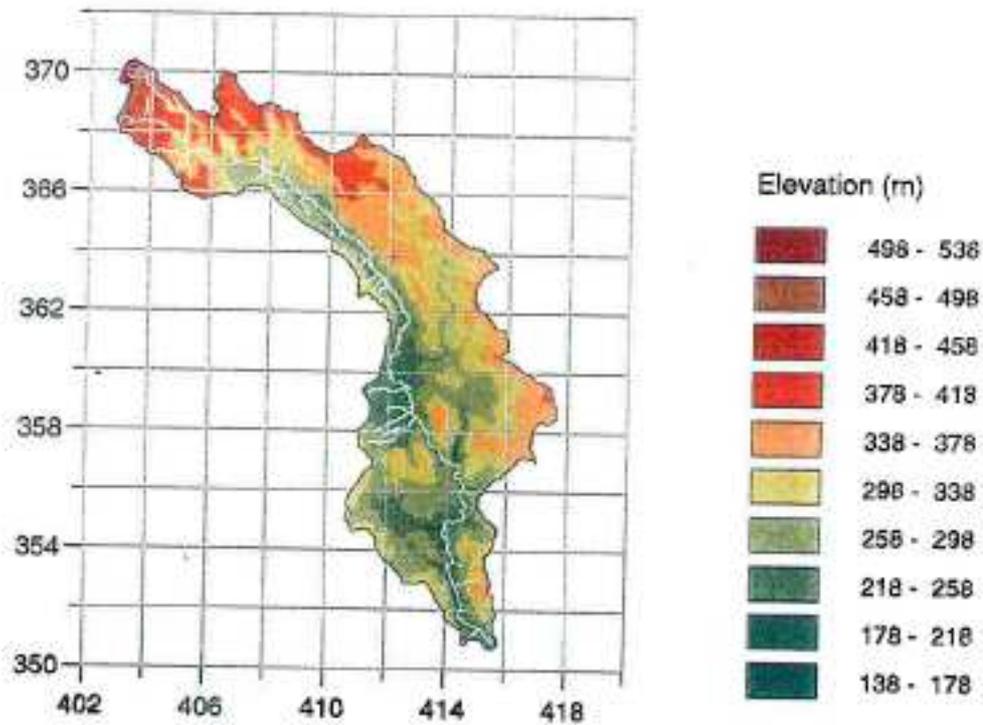


Figure B.3.1 Map of catchment relief: Dove.

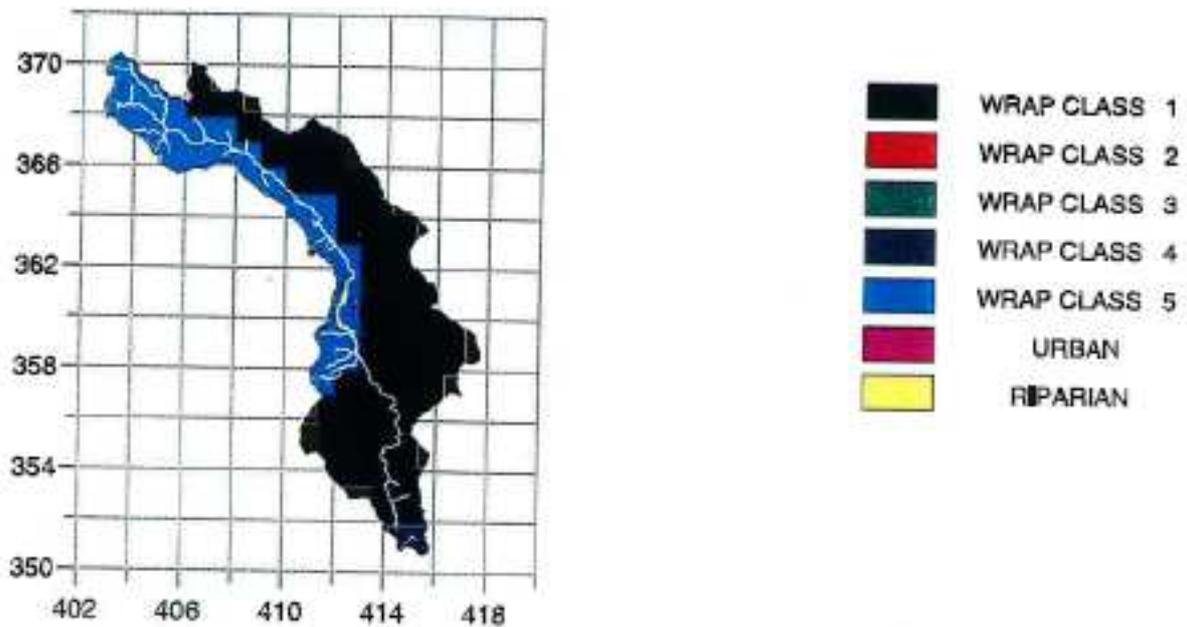


Figure B.3.2 Map of DTM-derived TCM response zones: Dove.

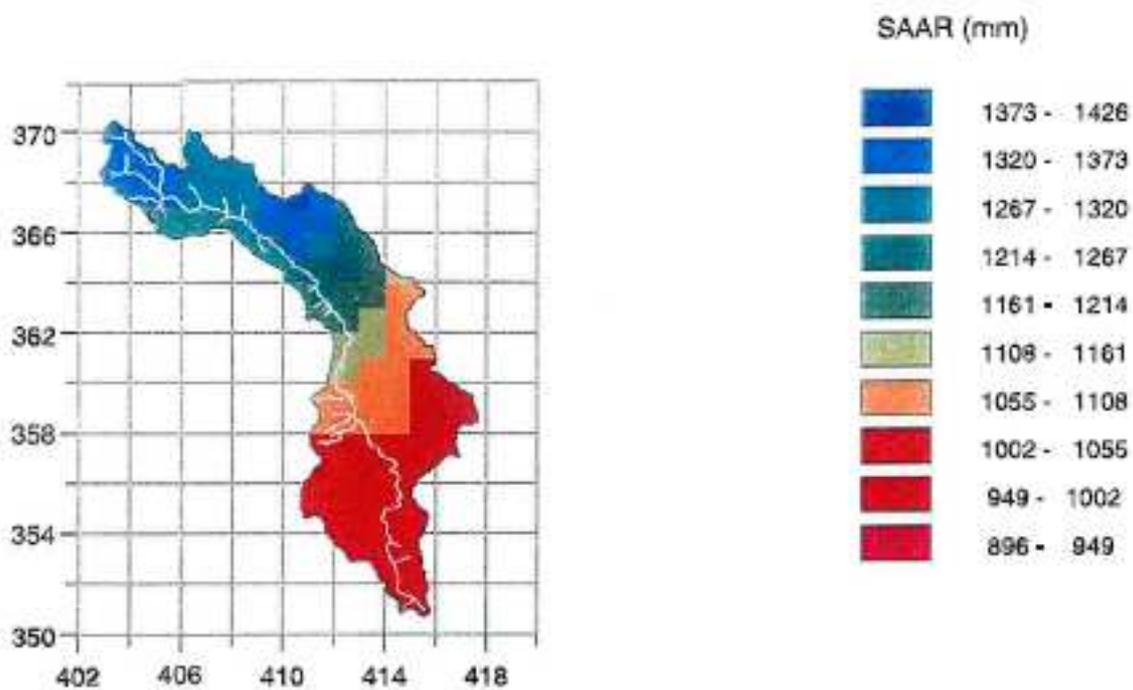


Figure B.3.3 Map showing distribution of SAAR across the catchment: Dove.

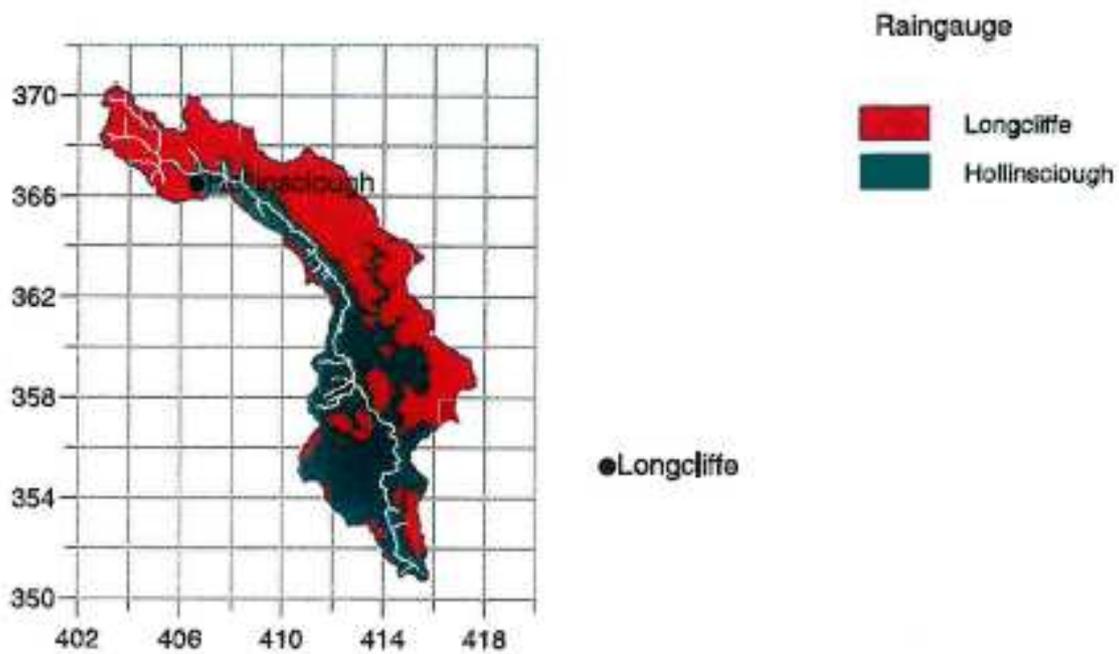


Figure B.3.4 Areas used to define raingauge weights: Dove.

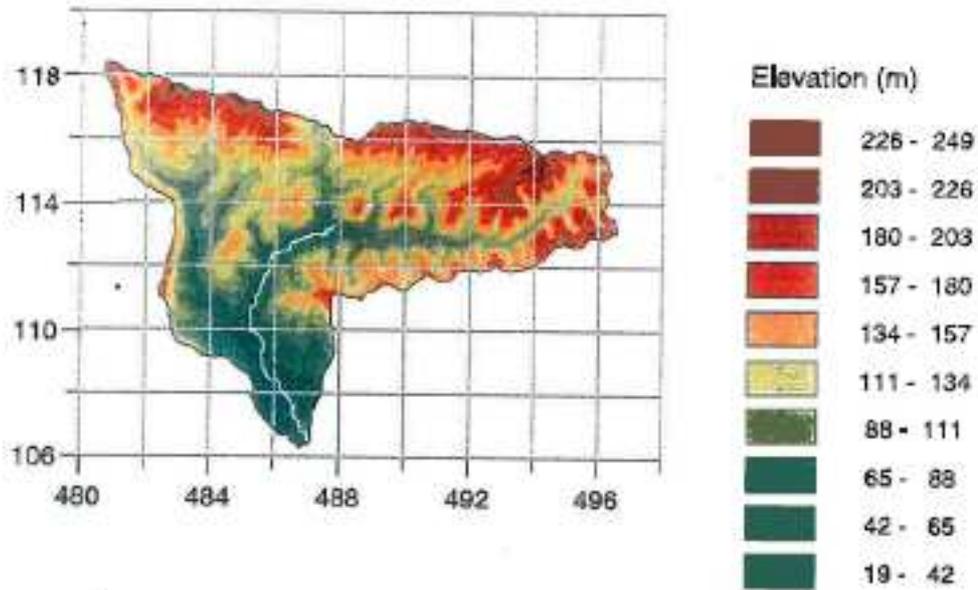


Figure B.4.1 Map of catchment relief: Lavant.



Figure B.4.2 Map of DTM-derived TCM response zones: Lavant.

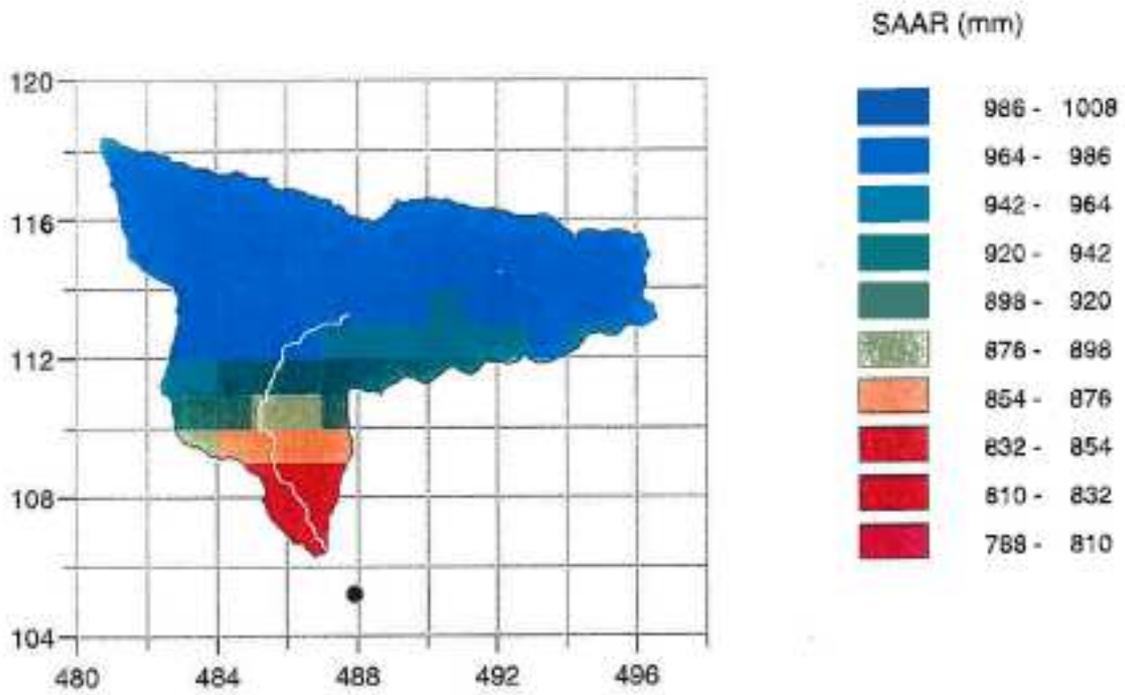


Figure B.4.3 Map showing distribution of SAAR across the catchment: Lavant.

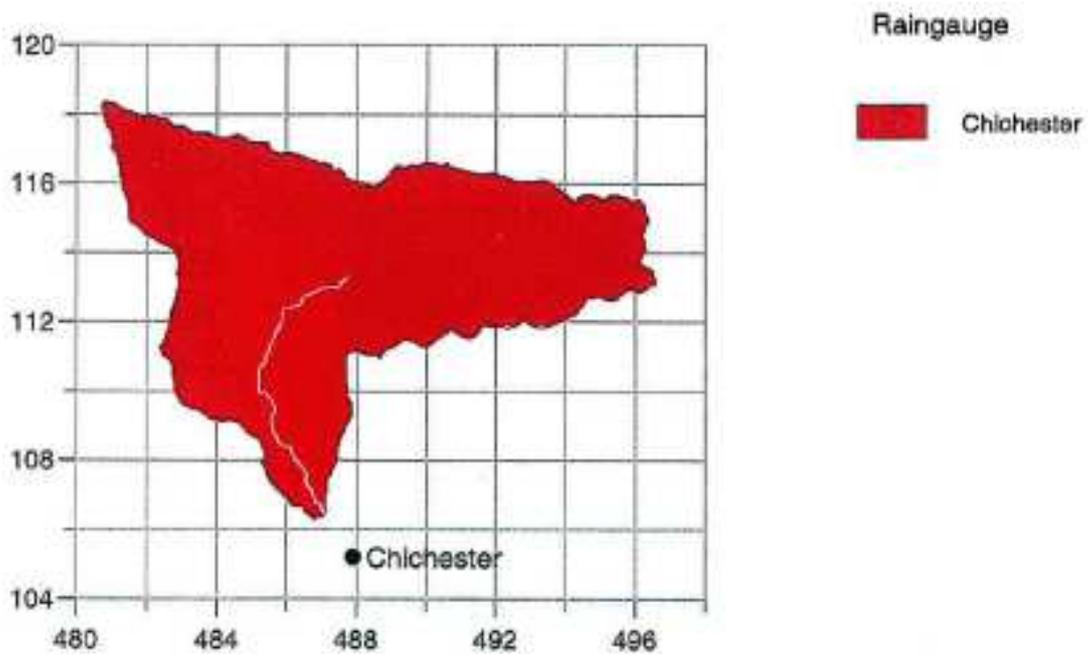


Figure B.4.4 Areas used to define raingauge weights: Lavant.

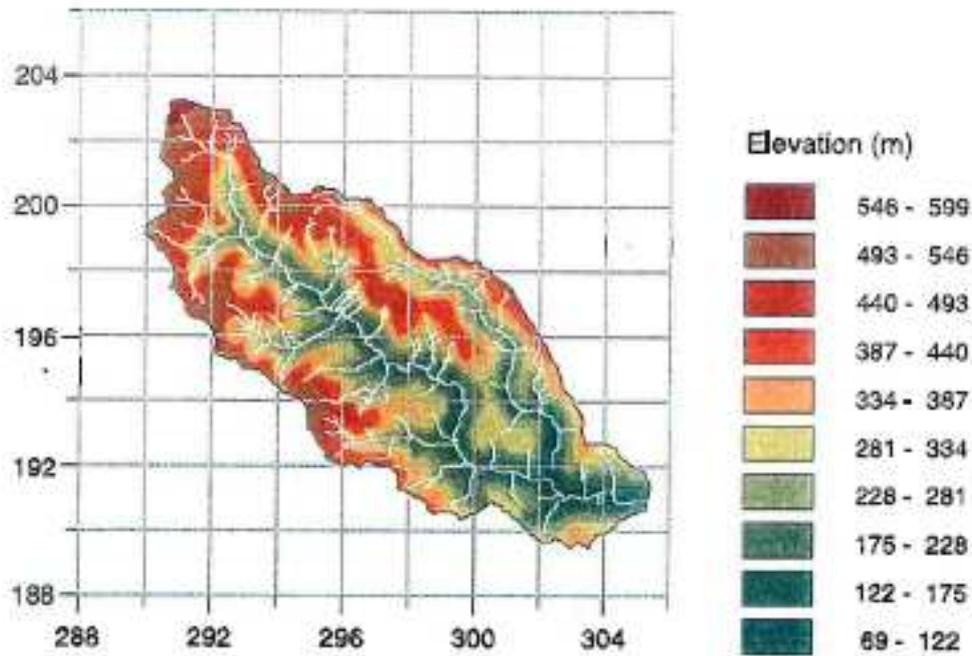


Figure B.5.1 Map of catchment relief: Rhondda.

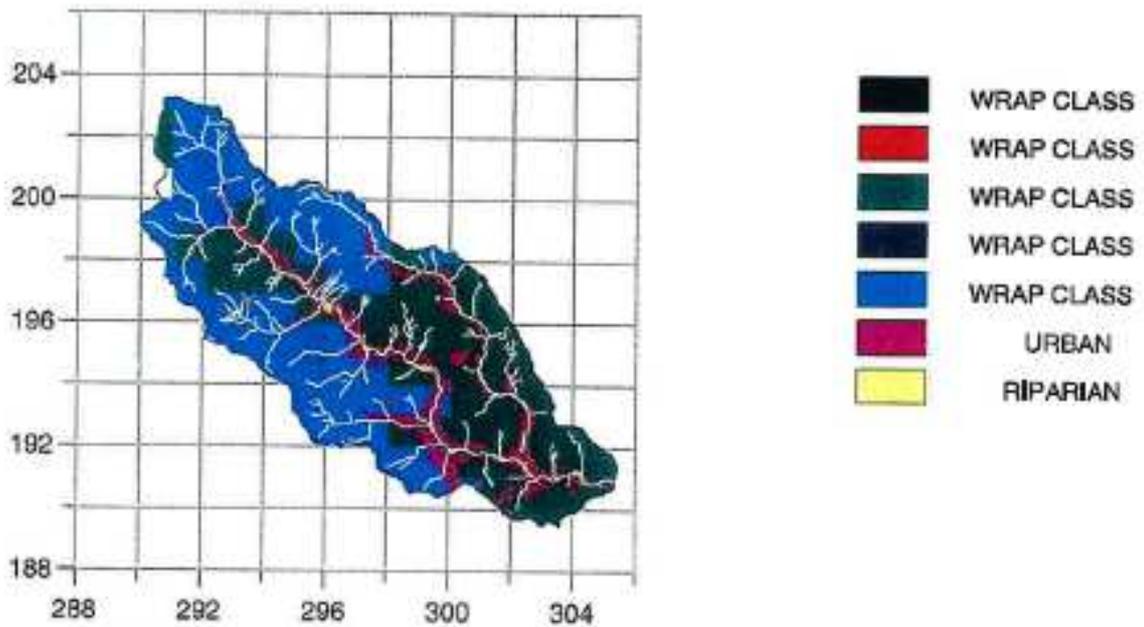


Figure B.5.2 Map of DTM-derived TCM response zones: Rhondda.

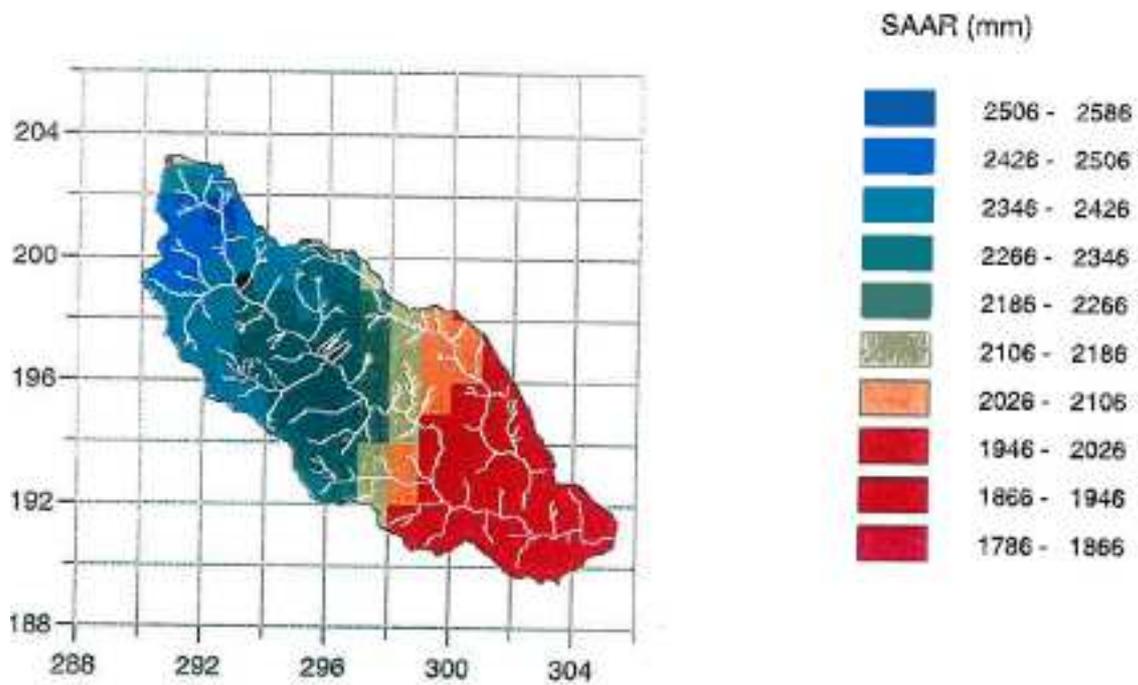


Figure B.5.3 Map showing distribution of SAAR across the catchment: Rhondda.

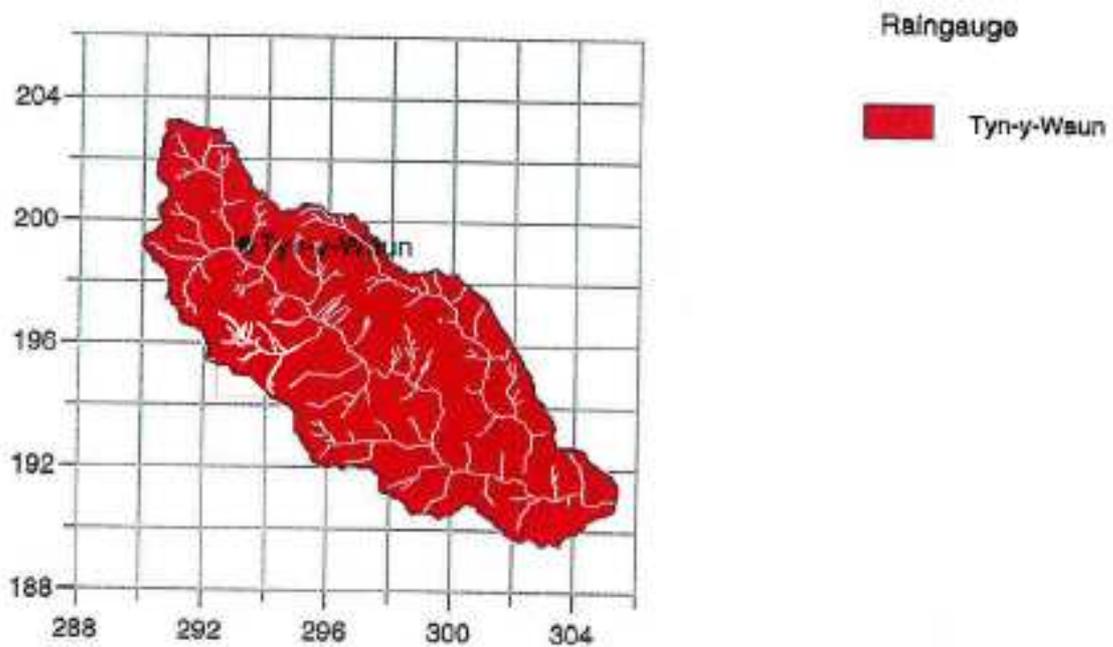


Figure B.5.4 Areas used to define raingauge weights: Rhondda.

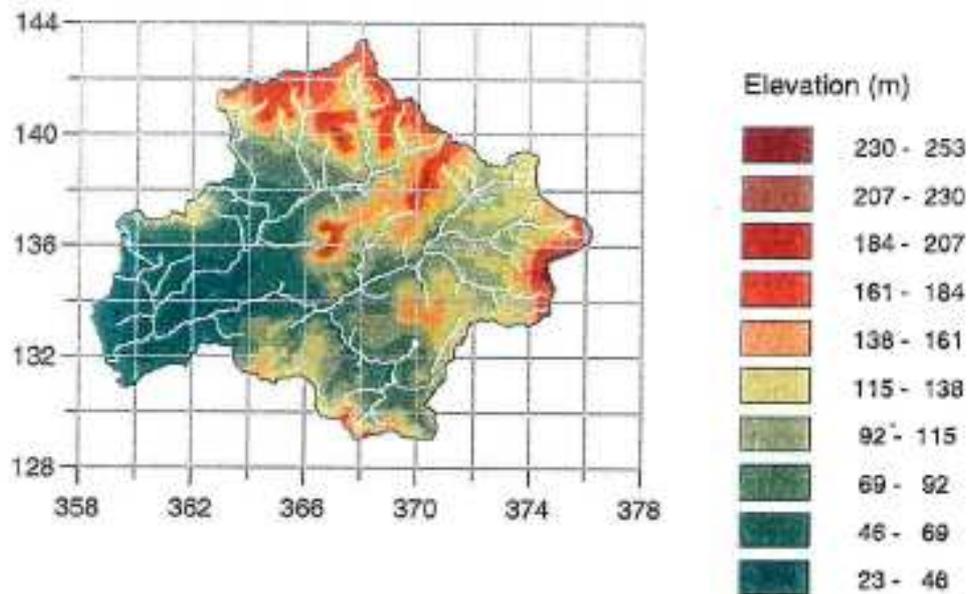


Figure B.6.1 Map of catchment relief: Brue.

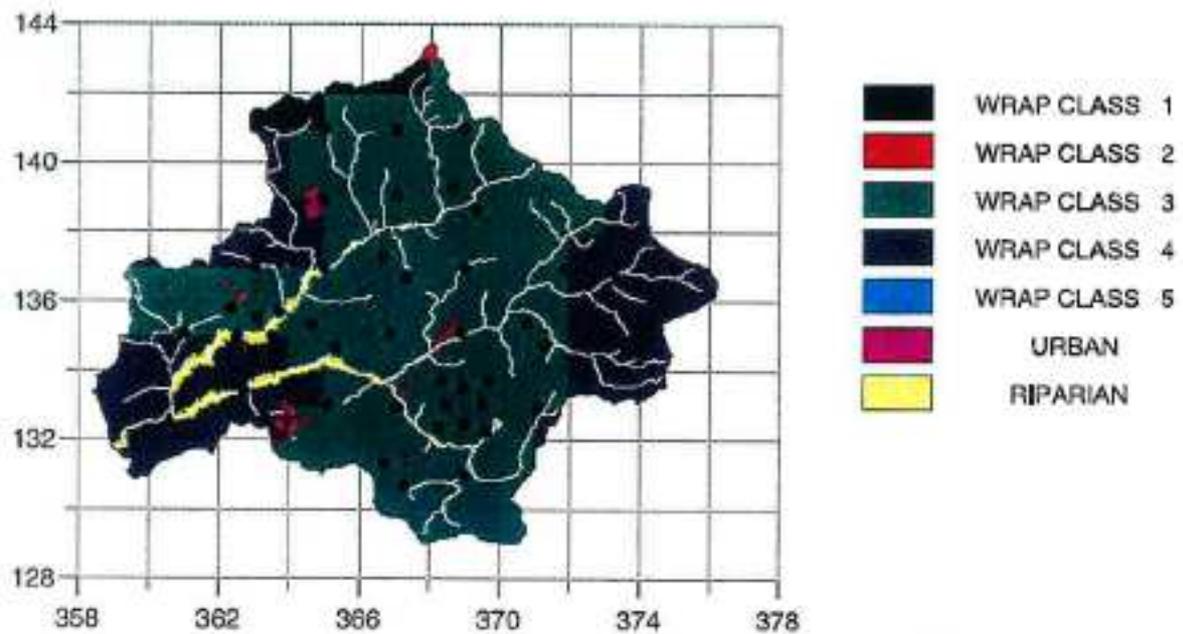


Figure B.6.2 Map of DTM-derived TCM response zones: Brue.

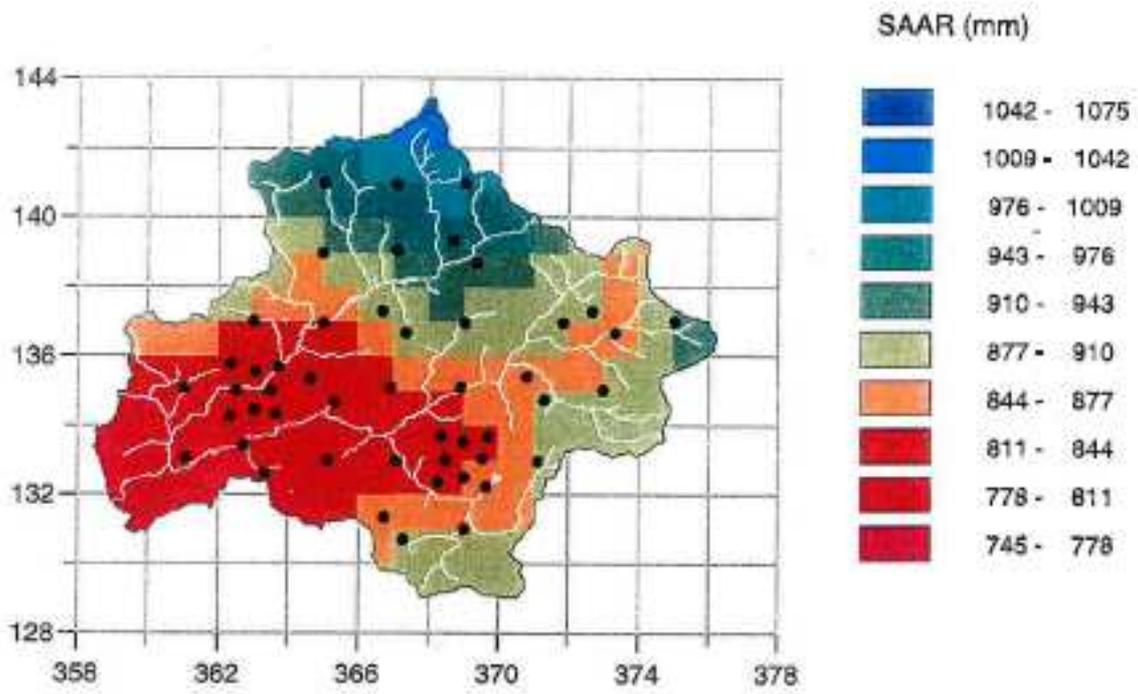


Figure B.6.3 Map showing distribution of SAAR across the catchment: Brue.

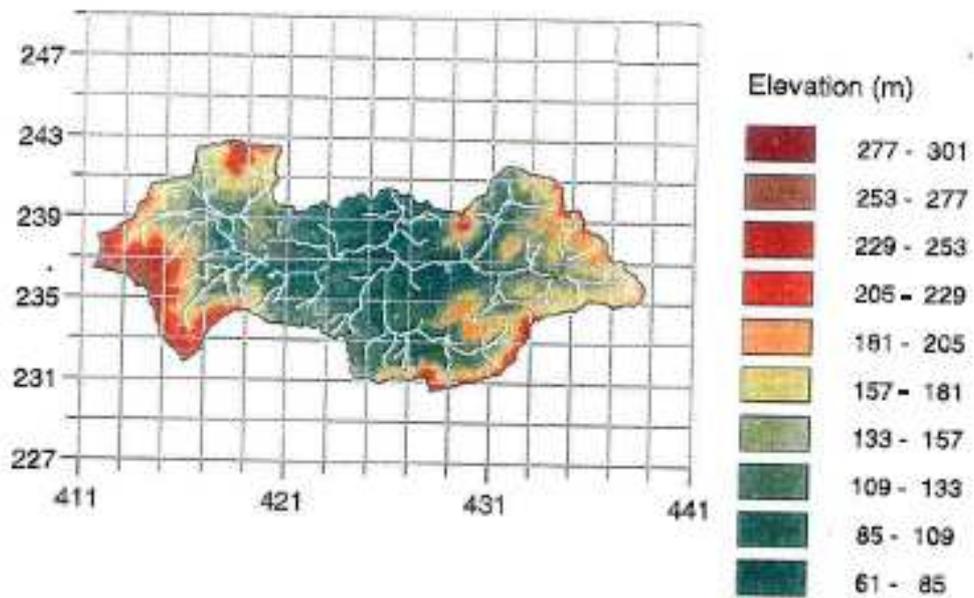


Figure B.7.1 Map of catchment relief: Stour.

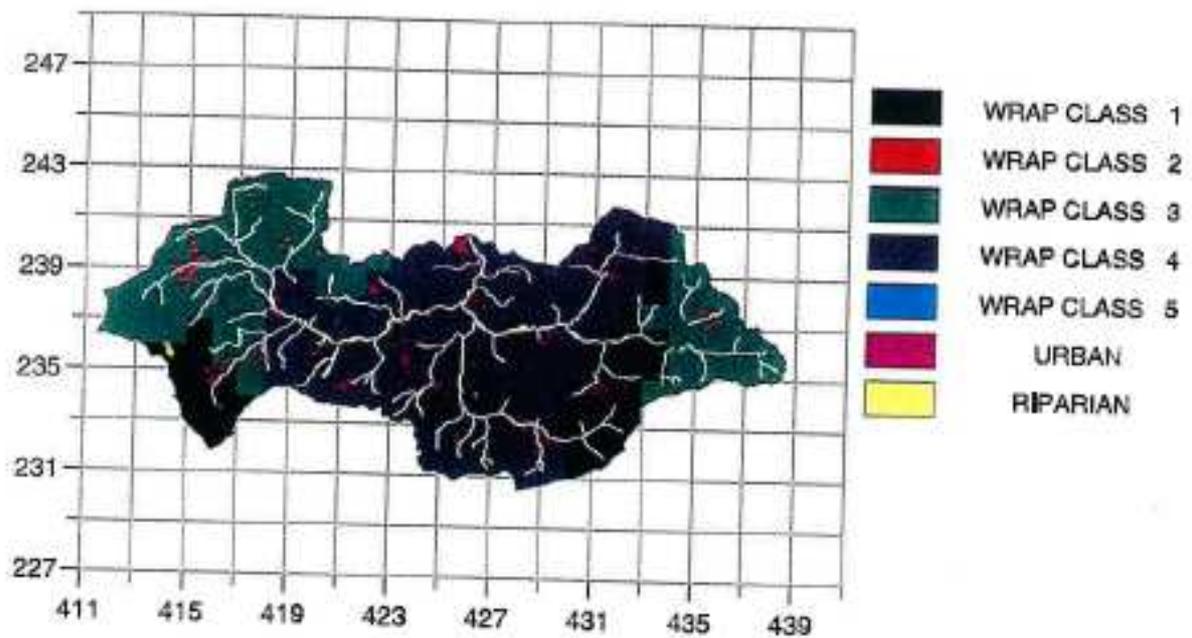


Figure B.7.2 Map of DTM-derived TCM response zones: Stour.

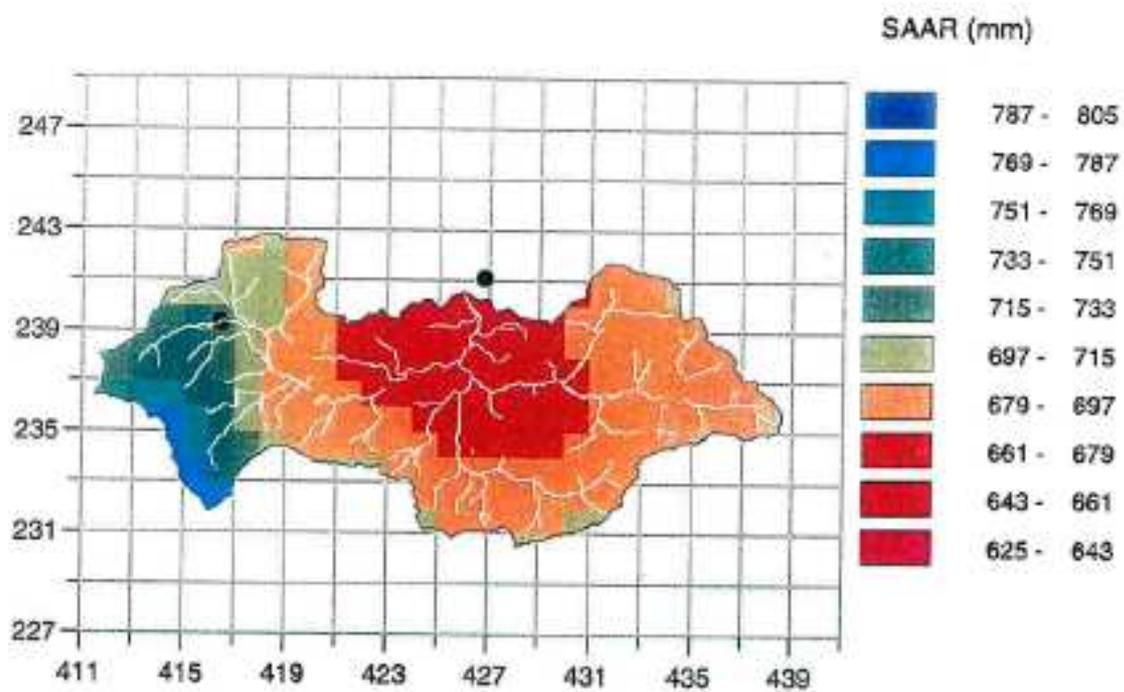


Figure B.7.3 Map showing distribution of SAAR across the catchment: Stour.

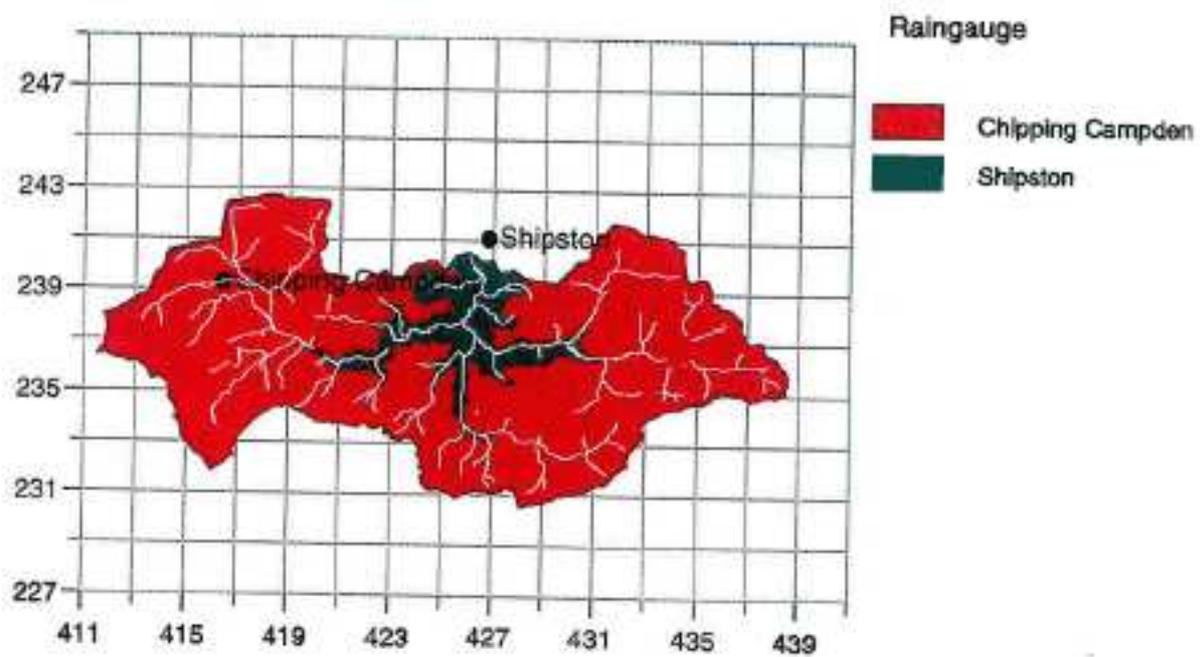


Figure B.7.4 Areas used to define raingauge weights: Stour.

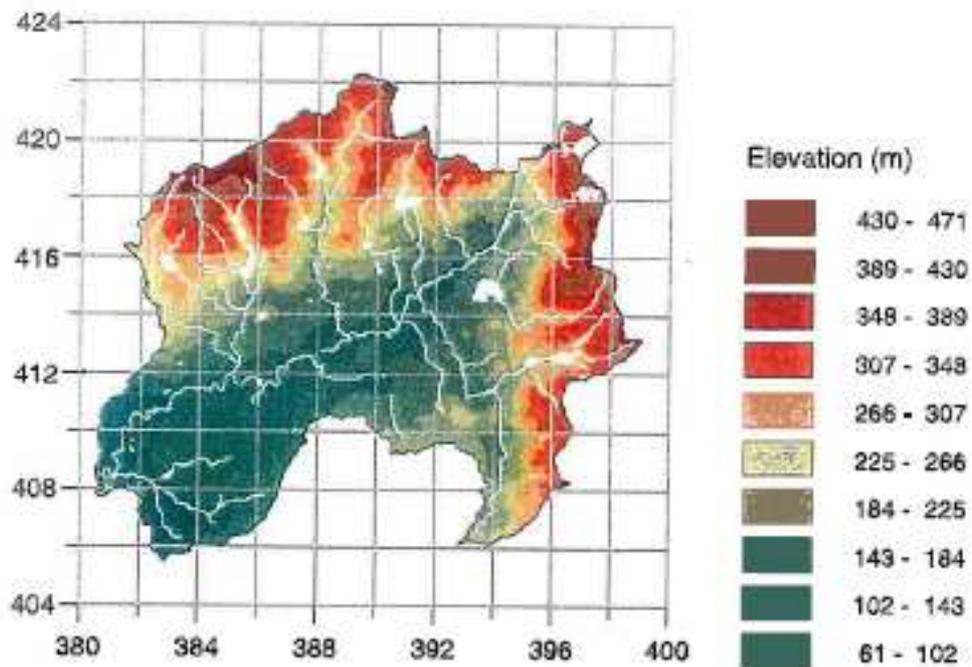


Figure B.8.1 Map of catchment relief: Roch.

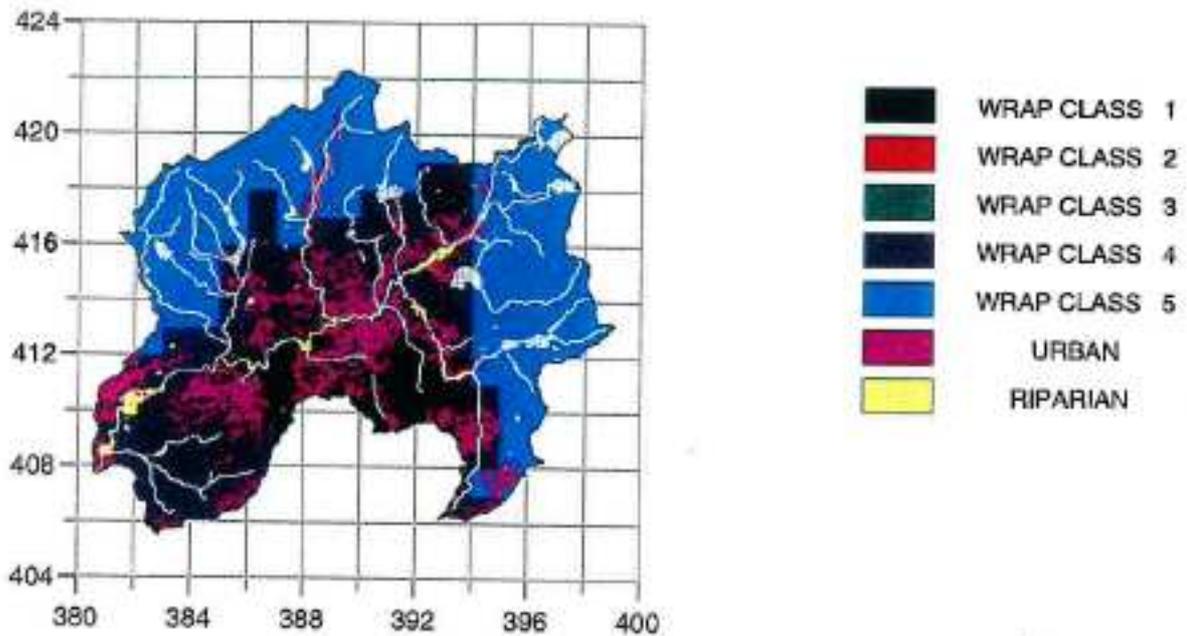


Figure B.8.2 Map of DTM-derived TCM response zones: Roch.

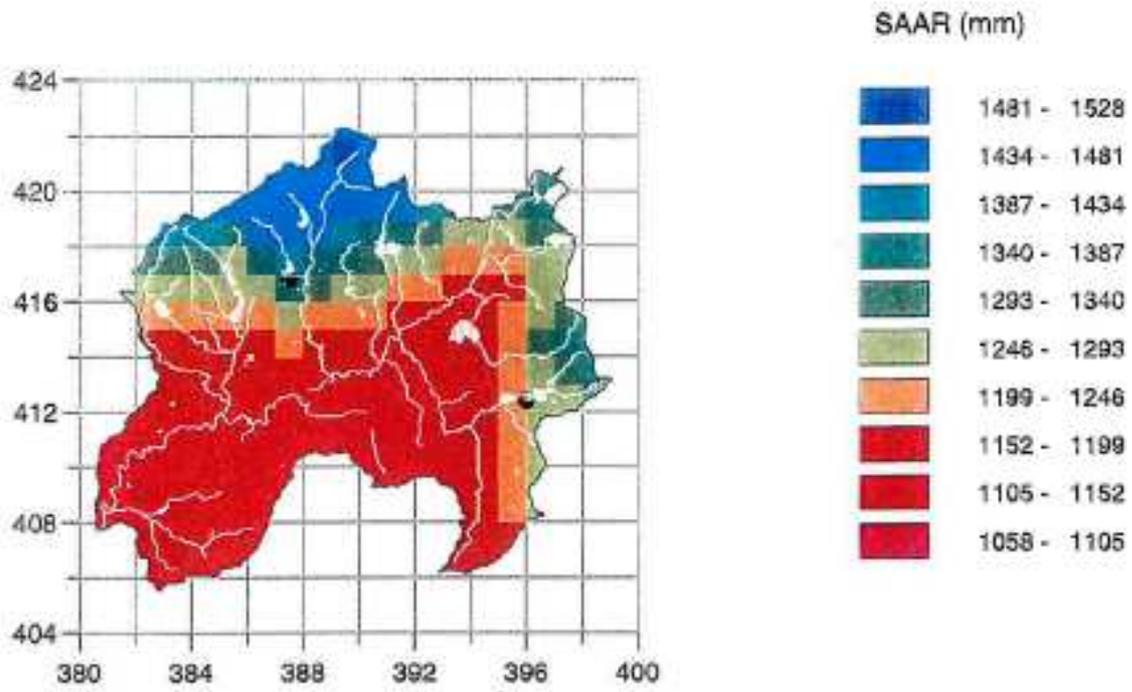


Figure B.8.3 Map showing distribution of SAAR across the catchment: Roch.

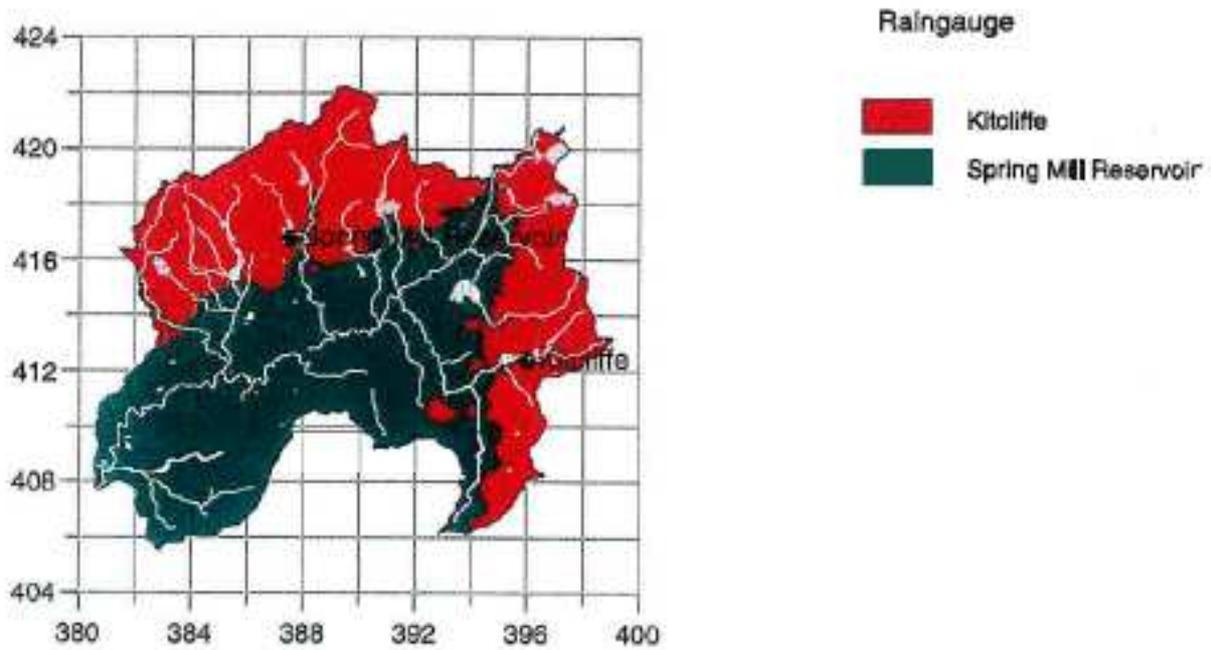


Figure B.8.4 Areas used to define raingauge weights: Roch.

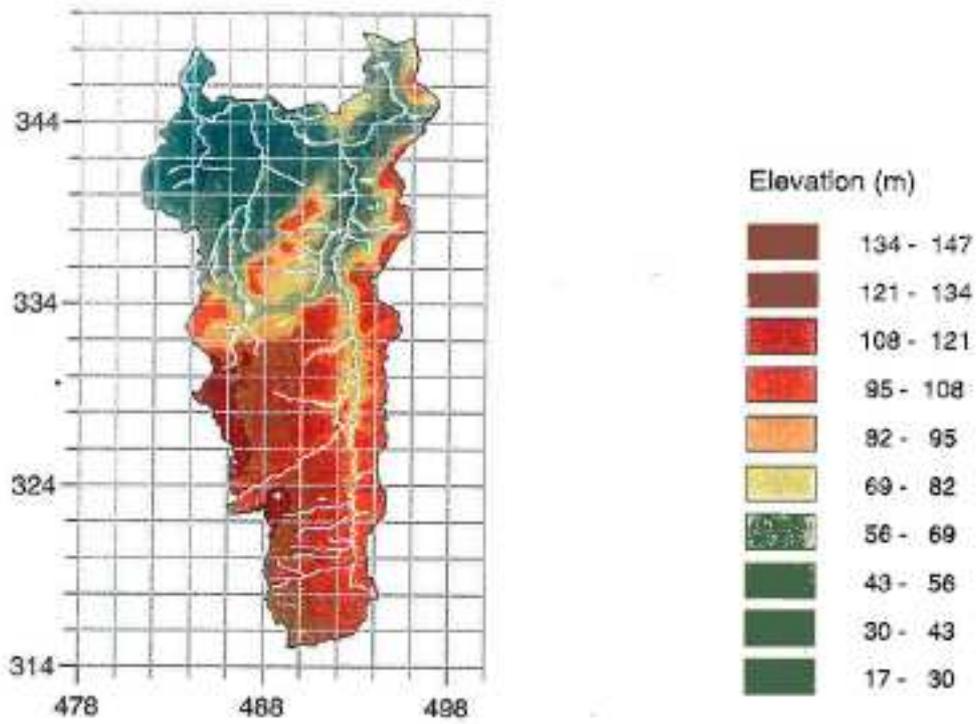


Figure B.9.1 Map of catchment relief: Witham.

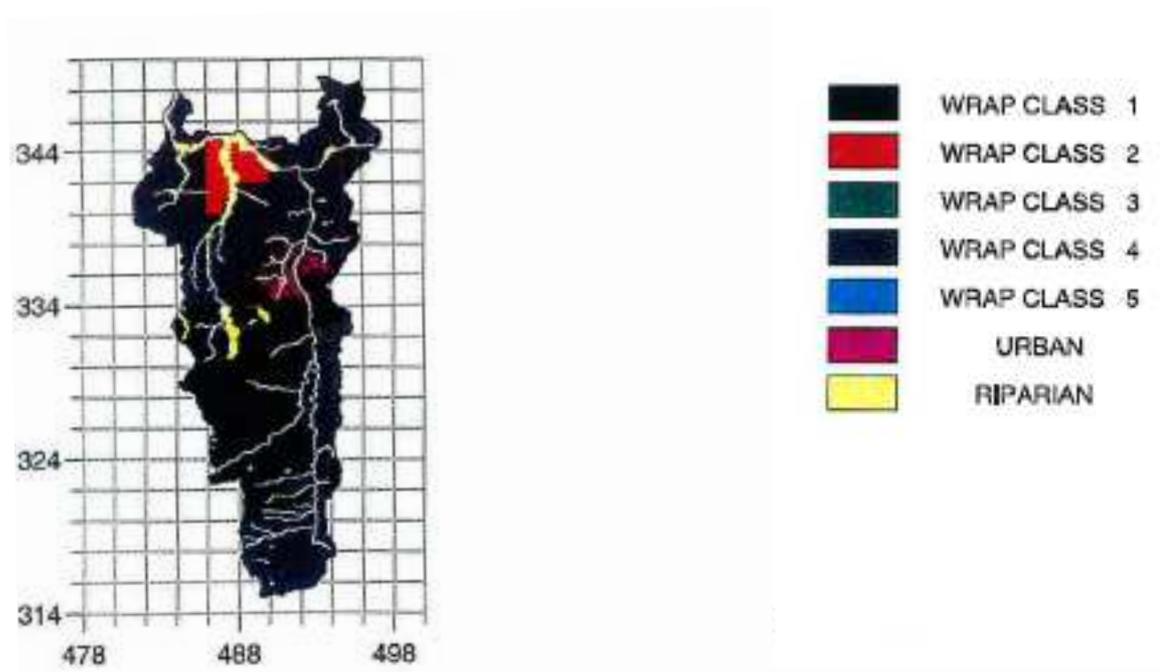


Figure B.9.2 Map of DTM-derived TCM response zones: Witham.

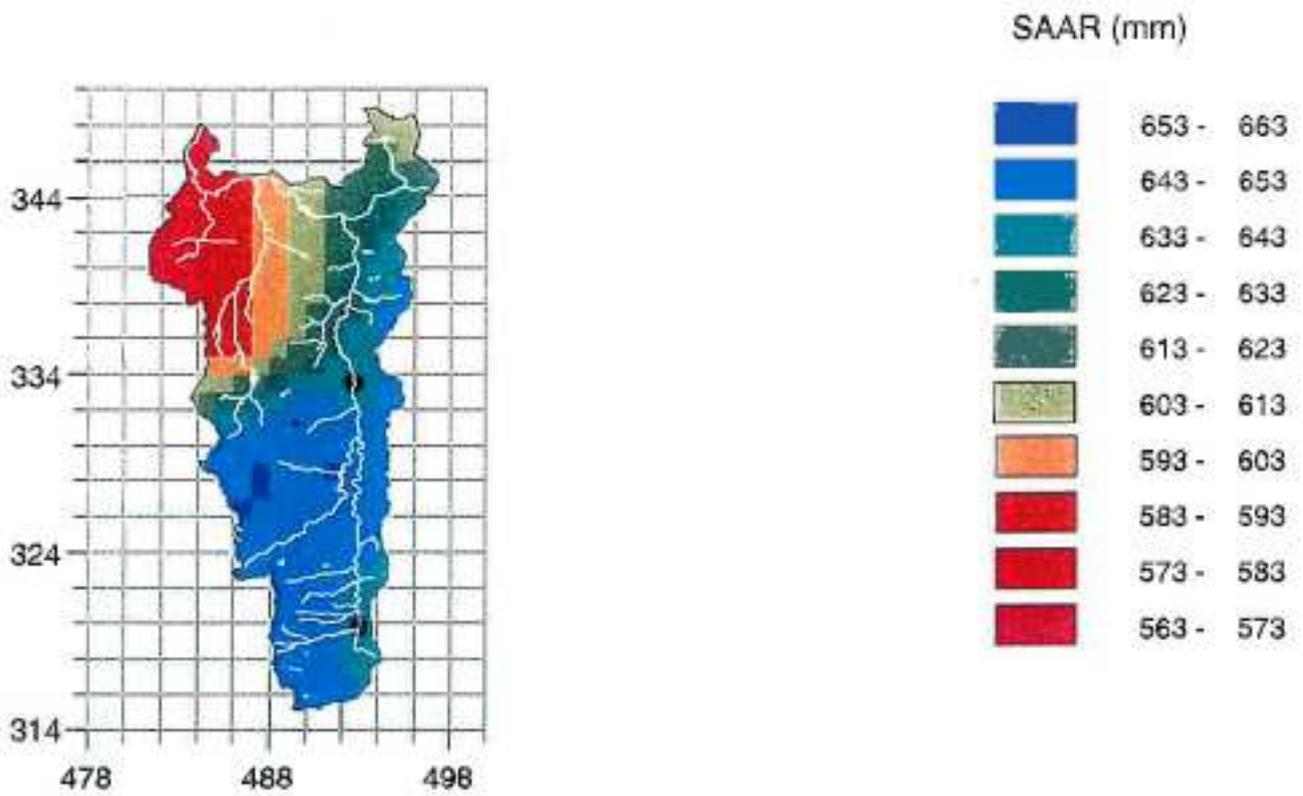


Figure B.9.3 Map showing distribution of SAAR across the catchment: Witham.

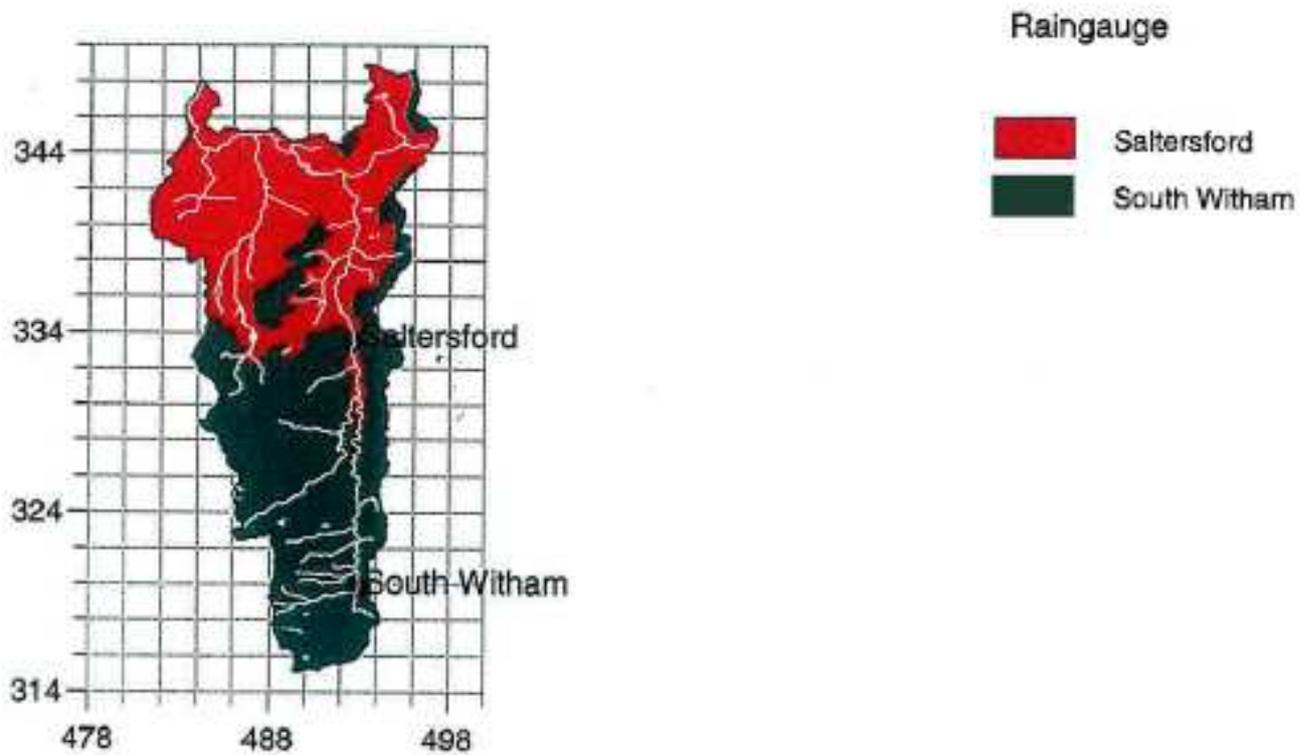


Figure B.9.4 Areas used to define raingauge weights: Witham.

APPENDIX C: MODEL PARAMETER SETS

The tables in this appendix detail the model structures and parameter values for the main set of models calibrated for each of the nine study catchments using raingauge data, uncalibrated radar and calibrated radar rainfall data. The models are presented in the following order: TCM, NWS, MCRM, PDM, IEM, TF, PRTF and Grid Model.

C.1 TCM Models

The units and meaning of the TCM model parameters are outlined in Section 1.2. Tables C.1.1 to C.1.4 summarise the calibrated parameter values obtained for each of the study catchments.

Table C.1.1 Comparison of TCM parameters across catchments using raingauge data

Model parameter		Catchment								
		Trout Beck	Silk Stream	Dove	Lavant	Rhondda	Brue	Stour	Roch	Witham
Number of zones	n_z	1	3	5	2	4	4	5	5	5
Zone parameters										
area (km ²)	A	9.85	4.22	53	86	45.0	70	73	40	178
			4.91	26.4	1.0	31.0	37.0	11.5	35	9.86
			4.2	16.2		1.6	1.7	40.2	19	40
				1.8		15.0	20	1	3.82	30
				5.6				9.1	50	2
lower zone drying rate	γ	0.3	0.3	0.3	0.3	0.3	0.3	0.1	0.3	0.3
			0.3	0.3	0.3	0.1	0.3	0.3	0.3	0.3
			0	0.1		0.3	0.1	0.3	0.3	0.3
				0.3		0	0.0	0.3	0.1	0.3
				0				0	0	0
direct percolation	ϕ	0	0	0.15	0.15	0.15	0	0.15	0.15	0.15
			0	0	0	0	0	0	0	0
			0	0		0	0	0	0	0
				0		0	0	0	0	0
				0				0	0	0
root constant (mm)	R_c	1	0	75	75	75	75	75	75	75
			54.5	75	75	75	75	75	75	75
			0	75		100	100	75	75	75
				45		0	0	100	100	100
				0				0	0	0
linear reservoir time constant	k	0.66	0.94	316	50	67.5	1	45	231.7	178
			0.96	229.1	0.5	0.9	2.1	12	23.7	8
			0.78	1.89		3.0	0.1	4	1.73	28
				44.5		1.6	3.8	6	3.52	10.7
				15.0				8.3	2.41	14.2
quadratic reservoir time constant	K	56.2	206.0	17231	1.8×10^5	350	10000	42200	2413	104751
			133.0	85	500	411	165	409.1	689	12864
			3.86	5081		40	130	320	96.9	488
				1556		39	72	11.02	140	180
				199				170	176	42
Constant flow	q_c	0	0	0	0	0	0	0	0	
Number of reaches	N	0	0	0	0	0	0	0	0	
Reach constant	θ	1	1	1	1	1	1	1	1	
Time delay	τ	0	0	0.16	0	0.14	3.76	0.5	0.33	0.34

Table C.1.2 Comparison of TCM parameters across catchments using uncalibrated radar data

Model parameter		Catchment		
		Silk Stream	Stour	Roch
Number of zones	n_z	3	5	5
Zone parameters				
area (km ²)	A	2.81	195.0	30.0
		4.95	3.55	25.0
		3.6	64.8	15.0
			93.2	10.0
			74.0	62.0
lower zone drying rate	γ	0.3	0.1	0.3
		0.3	0.3	0.3
		0.0	0.3	0.3
			0.3	0.1
			0.0	0.0
direct percolation	ϕ	0.0	0.15	0.15
		0.0	0.0	0.0
		0.0	0.0	0.0
			0.0	0.0
			0.0	0.0
root constant (mm)	R_c	0.0	75.0	75.0
		23.0	75.0	75.0
		0.0	75.0	75.0
			100.0	100.0
			0.0	0.0
linear reservoir time constant	k	1.05	90.4	107.0
		6.37	0.64	1.0
		0.35	2.1	31.9
			1.79	11.0
			4.52	1.33
quadratic reservoir time constant	K	226.0	13173.0	4107
		290.0	334.0	386.5
		17.0	149.2	212.0
			133.4	204.0
			28.3	204.0
Constant flow	q_c	0.0	0.0	0.0
Number of reaches	N	0	0	0
Reach constant	θ	1.0	1.0	1.0
Time delay	τ	0.44	2.87	1.3

Table C.1.3 Comparison of TCM parameters across catchments using calibrated radar data

Model parameter		Catchment		
		Silk Stream	Stour	Roch
Number of zones	n_z	3	5	5
Zone parameters				
area (km ²)	A	3.38	100.0	14.6
		5.8	34.0	10.9
		4.36	12.0	10.5
			38.0	1.5
			14.0	60.0
lower zone drying rate	γ	0.3	0.1	0.3
		0.3	0.3	0.3
		0.0	0.3	0.3
			0.3	0.1
			0.0	0.0
direct percolation	ϕ	0.0	0.15	0.15
		0.0	0.0	0.0
		0.0	0.0	0.0
			0.0	0.0
			0.0	0.0
root constant (mm)	R_c	0.0	75.0	75.0
		45.6	75.0	75.0
		0.0	75.0	75.0
			100.0	100.0
			0.0	0.0
linear reservoir time constant	k	0.73	76	159.0
		1.03	2.14	35.7
		0.72	9.5	22.3
			2.97	1.47
			0.1	0.92
quadratic reservoir time constant	K	150.0	24739	10914
		233.0	242.5	947.0
		7.70	28.9	306.0
			60.9	159.0
			132.8	369.0
Constant flow	q_c	0.0	0.0	0.0
Number of reaches	N	0	0	0
Reach constant	θ	1.0	1.0	1.0
Time delay	τ	0.44	4.9	1.41
ARMA updating				
AR(1)	δ_1	-1.4775	-2.4329	-
AR(2)	δ_2	0.5212	1.9794	-
AR(3)	δ_3	-	-0.5438	-

C.2 NWS Models

Table C.2.1 Comparison of NWS model parameters across catchments using raingauge data

Model parameter		Catchment								
		Trout Beck	Silk Stream	Dove	Lavant	Rhondda	Brue	Stour	Roch	Witham
Rainfall factor	r_f	1.13	0.472	1.0		0.893	0.955	0.736	0.814	0.598
Fraction of the catchment that is impervious	f_i	0.0	0.543	0.0024		0.125	0.013	0.029	0.173	0.033
Maximum additional fraction of impervious area which develops as tension water requirements are met	f_w^{\max}	0.0	0.454	0.016		0.0	0.276	0.503	0.307	0.520
Fraction of the catchment covered by streams, lakes and riparian vegetation	c	0.002	0.0034	0.018		0.015	0.032	0.019	0.02	0.06
Capacity of unsaturated zone tension water store	S_u^{\max}	16.27	25.25	217.0		430.0	13.17	40.39	46.2	48.7
Capacity of unsaturated zone free water store	S_s^{\max}	86.7	86.2	49.0		33.26	24.62	11.69	43.9	19.6
Rate of interflow from saturated zone	k_i	0.608	0.781	0.003		0.120	0.211	0.985	0.05	2.1×10^{-4}
Proportional increase in percolation from saturated to dry conditions	γ	10.56	72.11	60.0		56.8	31.01	1339.0	43.8	31.6
Exponent in equation for percolation rate	δ	0.56	0.196	1.24		5.5	6.463	6.711	1.22	0.847
Capacity of groundwater zone tension water store	S_{gt}^{\max}	153.0	233.4	96.0		144.2	45.58	34.25	55.6	49.3
Capacity of groundwater zone secondary free water storage	S_{gs}^{\max}	422.0	149.4	190.0		196.0	351.2	90.33	187.0	211.0
Lateral drainage rate from secondary groundwater zone	k_{gs}	0.0047	0.183	4.8×10^{-4}		7.5×10^{-4}	0.0014	2.4×10^{-5}	0.0017	1.1×10^{-4}
Capacity of groundwater zone primary free water storage	S_{gp}^{\max}	89.2	157.0	164.0		157.9	221.4	204	38.0	35.4
Lateral drainage rate from primary groundwater zone	k_{gp}	0.0025	0.183	3.1×10^{-4}		5.0×10^{-3}	3.9×10^{-4}	1.5×10^{-4}	0.01	9.2×10^{-5}
Fraction of percolated water going directly to groundwater zone free water store in preference to tension water store	p	0.565	0.569	0.2		0.2	0.2	0.445	0.882	0.484
Fraction of groundwater zone free water not available for resupplying lower zone tension water store	r_s	0.337	0.324	0.3		0.3	0.432	0.318	0.475	0.6

Table C.2.2 Comparison of NWS model parameters across catchments using uncalibrated radar data

Model parameter		Catchment		
		Silk Stream	Stour	Roch
Rainfall factor	r_f	0.467	1.908	0.712
Fraction of the catchment that is impervious	f_i	0.543	0.029	0.173
Additional fraction of impervious area which develops as tension water requirements are met	f_w	0.454	0.335	0.307
Fraction of the catchment covered by streams, lakes and riparian vegetation	c	0.0034	0.019	0.021
Capacity of unsaturated zone tension water store	S_u^{\max}	34.9	47.0	48.0
Capacity of unsaturated zone free water store	S_s^{\max}	87.15	38.5	12.16
Rate of interflow from saturated zone	k_i	0.832	0.145	0.522
Proportional increase in percolation from saturated to dry conditions	γ	72.9	15.22	59.10
Exponent in equation for percolation rate	δ	0.198	11.35	3.57
Capacity of groundwater zone tension water store	S_{gt}^{\max}	231.0	11.36	37.30
Capacity of groundwater zone secondary free water storage	S_{gs}^{\max}	149.0	137.0	223.30
Lateral drainage rate from secondary groundwater zone	k_{gs}	0.185	0.002	0.00096
Capacity of groundwater zone primary free water storage	S_{gp}^{\max}	155.6	531.8	26.2
Lateral drainage rate from primary groundwater zone	k_{gp}	0.232	0.0012	0.0082
Fraction of percolated water going directly to groundwater zone free water store in preference to tension water store	p	0.396	0.504	0.536
Fraction of groundwater zone free water not available for resupplying lower zone tension water store	r_s	0.378	0.119	0.584

Table C.2.3 Comparison of NWS model parameters across catchments using calibrated radar data

Model parameter		Catchment		
		Silk Stream	Stour	Roch
Rainfall factor	r_f	0.498	1.08	0.608
Fraction of the catchment that is impervious	f_i	0.543	0.029	0.173
Additional fraction of impervious area which develops as tension water requirements are met	f_w	0.454	0.29	0.307
Fraction of the catchment covered by streams, lakes and riparian vegetation	c	0.0034	0.019	0.021
Capacity of unsaturated zone tension water store	S_u^{\max}	111.0	47.0	67.0
Capacity of unsaturated zone free water store	S_s^{\max}	64.17	38.5	4.35
Rate of interflow from saturated zone	k_i	0.260	0.120	0.553
Proportional increase in percolation from saturated to dry conditions	γ	64.8	28.6	31.9
Exponent in equation for percolation rate	δ	0.389	5.26	2.94
Capacity of groundwater zone tension water store	S_{gt}^{\max}	259.0	15.24	34.4
Capacity of groundwater zone secondary free water storage	S_{gs}^{\max}	240.0	138.0	82.1
Lateral drainage rate from secondary groundwater zone	k_{gs}	0.353	0.0019	0.0011
Capacity of groundwater zone primary free water storage	S_{gp}^{\max}	102.0	245.0	82.2
Lateral drainage rate from primary groundwater zone	k_{gp}	0.227	7.6×10^{-4}	3.8×10^{-3}
Fraction of percolated water going directly to groundwater zone free water store in preference to tension water store	p	0.399	0.933	0.838
Fraction of groundwater zone free water not available for resupplying lower zone tension water store	r_s	0.397	0.875	0.502

C.3 MCRM Models

Table C.3.1 Comparison of MCRM parameters between catchments using raingauge data

Model parameter		Catchment								
		Trout Beck	Silk Stream	Dove	Lavant	Rhondda	Brue	Stour	Roch	Witham
Rainfall factor	f_c	1.0	0.91	1.04		0.90	0.86	0.86	0.83	1.0
Capacity of interception store	S_{max}	1.0	1.0	2.0		2.0	1.0	2.84	4.3	2.3
Fraction of catchment evaporation potentially met by interception storage	f	1.0	1.0	0.84		1.0	1.0	1.0	1.0	1.0
Minimum value of rapid runoff proportion	c_0	0.445	0.99	0.11		0.165	0.24	0.28	0.93	0.1
Parameter in rapid runoff proportion function	c_1	0.096	0.015	0.099		0.088	0.082	0.016	0.005	0.02
Maximum value of rapid runoff proportion	c_{max}	0.6	0.32	0.2		0.54	0.76	0.56	0.5	0.24
Maximum percolation rate	q_p^{max}	0.024	0.36	27.0		3.53	0.14	0.01	0.08	0.69
Maximum soil store moisture surplus	D_{surp}	0.23	10.0	85		18.5	3.7	48.6	30.0	6.0
Soil function exponent controlling rapid drainage	γ_d	1.3	2.9	2.9		1.89	2.14	2.26	1.0	1.18
Soil function coefficient controlling rapid drainage	k_d	55.0	50.0	21.0		11.8	40.0	10.0	60.0	85.0
Potential transpiration factor	T_p	0.40	0.54	0.3		0.61	0.7	0.3	0.4	1.0
Minimum transpiration factor	T_m	0.058	0.08	0.205		0.079	0.04	0.018	0.04	0.2
Deficit below which potential transpiration factor applies	E_{max}^D	90.0	85.0	27.0		72.4	70.0	56.8	62.0	72.0
Deficit above which minimum transpiration factor applies	E_{min}^D	150.0	149.0	109.0		142.0	131.0	137.0	155.0	165.0
Time constant in baseflow storage function	K_g	0.2	67.0	13.9		0.50	1.5	10.0	0.76	16.5
Time lag applied to total runoff	τ	0.12	0.07	0.5		0.069	2.3	0.95	0.33	0.8
Duration of time spread applied to total runoff	T	5.75	1.25	7.5		7.0	9.0	7.5	4.75	16.0
Channel storage at bankfull	S_{bf}	35.6	26.0	43.0		330.0	55.0	47.0	55.0	43.0
In-channel routing storage coefficient	k_{cr}	0.027	0.013	0.026		0.81	0.029	0.011	0.022	0.053
In-channel routing storage exponent	γ_{cr}	5.0	2.82	1.45		2.28	2.64	1.44	1.46	0.83
Out-of-bank channel routing storage coefficient	k_{or}	0.051	0.095	0.28		0.046	0.11	0.014	0.036	0.068
Out-of-bank channel routing storage exponent	γ_{or}	4.3	2.0	1.73		1.5	1.7	1.51	2.23	1.97

Table C.3.2 Comparison of MCRM parameters between catchments using uncalibrated radar data

Model parameter		Catchment		
		Silk Stream	Stour	Roch
Rainfall factor	f_c	0.63	2.29	0.82
Capacity of interception store	S_{max}	1	2.13	4.77
Fraction of catchment evaporation potentially met by interception storage	f	1	1.0	0.99
Minimum value of rapid runoff proportion	c_0	0.75	0.28	0.93
Parameter in rapid runoff proportion function	c_1	0.038	0.032	0.0077
Maximum value of rapid runoff proportion	c_{max}	0.35	0.75	0.48
Maximum percolation rate	q_p^{max}	0.11	0.07	0.19
Maximum soil store moisture surplus	D_{surp}	9.75	80.0	18.9
Soil function exponent controlling rapid drainage	γ_d	2.64	2.0	1.34
Soil function coefficient controlling rapid drainage	k_d	50.0	30.0	57.0
Potential transpiration factor	T_p	0.14	0.73	1.1
Minimum transpiration factor	T_m	0.084	0.035	0.044
Deficit below which potential transpiration factor applies	E_{max}^D	84.7	66.0	64.0
Deficit above which minimum transpiration factor applies	E_{min}^D	149.0	154.0	156.0
Time constant in baseflow storage function	K_g	54.9	2.3	0.89
Time lag applied to total runoff	τ	0.13	1.22	0.88
Duration of time spread applied to total runoff	T	1.25	7.5	5.0
Channel storage at bankfull	S_{bf}	26.0	47.0	40.0
In-channel routing storage coefficient	k_{cr}	0.0117	0.0087	0.022
In-channel routing storage exponent	γ_{cr}	3.5	1.55	1.62
Out-of-bank channel routing storage coefficient	k_{or}	0.1	0.014	0.04
Out-of-bank channel routing storage exponent	γ_{or}	0.015	2.68	1.96

Table C.3.3 Comparison of MCRM parameters between catchments using calibrated radar data

Model parameter		Catchment		
		Silk Stream	Stour	Roch
Rainfall factor	f_c	0.92	1.0	0.62
Capacity of interception store	S_{max}	1	2.78	4.93
Fraction of catchment evaporation potentially met by interception storage	f	1	1.0	1.0
Minimum value of rapid runoff proportion	c_0	0.87	0.39	0.93
Parameter in rapid runoff proportion function	c_1	0.025	0.035	0.0009
Maximum value of rapid runoff proportion	c_{max}	0.31	0.64	0.48
Maximum percolation rate	q_p^{max}	0.41	0.11	0.42
Maximum soil store moisture surplus	D_{surp}	9.92	76.0	2.4
Soil function exponent controlling rapid drainage	γ_d	2.86	2.28	1.4
Soil function coefficient controlling rapid drainage	k_d	50.0	41.0	59.0
Potential transpiration factor	T_p	0.35	0.14	1.4
Minimum transpiration factor	T_m	0.079	0.06	0.05
Deficit below which potential transpiration factor applies	E_{max}^D	84.7	65.72	67.0
Deficit above which minimum transpiration factor applies	E_{min}^D	149.0	154.0	159.0
Time constant in baseflow storage function	K_g	62.0	3.84	1.96
Time lag applied to total runoff	τ	0.08	1.23	0.77
Duration of time spread applied to total runoff	T	1.25	8.75	4.75
Channel storage at bankfull	S_{bf}	26.0	25.4	33.9
In-channel routing storage coefficient	k_{cr}	0.023	0.00054	0.056
In-channel routing storage exponent	γ_{cr}	2.88	2.40	0.92
Out-of-bank channel routing storage coefficient	k_{or}	0.11	0.072	0.047
Out-of-bank channel routing storage exponent	γ_{or}	1.75	1.52	1.76

C.4 PDM Models

The units and meaning of the PDM model parameters are given in Section 1.2. Tables C.4.1 to and C.4.3 summarise the calibrated parameter values.

Note that all models use a Pareto distribution of store capacity with $c_{\min}=0$, and a cascade of two equal linear reservoirs for the surface storage model. The state-correction scheme employs the super-proportion gain scheme.

Table C.4.1 Comparison of PDM parameters across catchments using raingauge data

Model parameter		Catchment								
		Trout Beck	Silk Stream	Dove	Lavant	Rhondda	Brue	Stour	Roch	Witham
Rainfall factor	f	0.91	0.39	0.92	1.1	0.9	0.86	0.76	0.61	0.6
Soil moisture										
Min. depth	c_{min}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Max. depth	c_{max}	50.0	10.0	266.0	600.0	135.0	105.0	84.8	80.0	100.0
Exponent	b	3.0	1.5	0.2	0.25	0.63	1.25	0.34	4.0	0.25
Evaporation exponent	b_e	1.0	2.5	2.0	1000	2.5	2.5	2.5	2.0	2.0
Recharge model	type	1	1	1	1	1	1	1	1	2
	1 2 3									
parameter 1	k_g α α	500.00	200.0	70000.0	248600	85000.0	90000.0	97596.0	2200.0	0.025
parameter 2	S_t β	0.0	0.0	0.0	85.0	0.0	0.0	0.0	0.0	49.0
parameter 3	b_g q_{sat}	2.0	2.1	1.83	13.0	2.42	2.65	2.24	2.28	4.0
Surface storage coefficient	k_s	1.6	1.0	7.9		3.1	4.04	6.08	4.44	18.3
Baseflow storage										
Exponent	m	3	3	2	3	3	3	3	3	3
coefficient	k_b	5.0	5.0	10.5	460	13.0	8.83	5.0	11.7	82.0
Constant flow	q_c	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Time delay	τ	0.13	0.36	0.25	0.0	0.17	4.2	3.12	0.2	0.45
State-correction										
Surface gain	g_s	1.535	1.591	1.453	0.0	1.530	1.897	1.882	1.644	1.641
baseflow gain	g_b	1.325	0.606	1.126	1.04	1.714	1.736	1.584	1.232	0.670
soil moisture gain	g_g	0	0	0	0	0	0	0	0	0

Table C.4.2 Comparison of PDM parameters across catchments using uncalibrated radar data

Model parameters		Catchment		
		Silk Stream	Stour	Roch
Rainfall factor	f	0.38	1.79	0.68
Soil moisture				
min. depth	c_{\min}	0.0	0.0	0.0
max. depth	c_{\max}	15.0	84.8	120.0
exponent	b	0.95	0.43	3.73
Evaporation exponent	b_e	1.1	2.5	2.0
Recharge model	type	1	1	1
	1 2 3			
parameter 1	k_g α α	370.0	97500	4051
parameter 2	S_t β	0.0	0.0	0.0
parameter 3	b_g q_{sat}	1.75	2.25	2.37
Surface storage coefficient	k_s	1.55	4.59	3.44
Baseflow storage				
exponent	m	3.0	3.0	3.0
coefficient	k_b	2.89	5.0	13.2
Constant flow	q_c	0.0	0.0	0.0
Time delay	τ	0.025	4.3	0.47
State-correction				
surface gain	g_s	-	-	-
baseflow gain	g_b	-	-	-
soil moisture gain	g_g	-	-	-

Table C.4.3 Comparison of PDM parameters across catchments using calibrated radar data

Model parameters		Catchment		
		Silk Stream	Stour	Roch
Rainfall factor	f	0.43	0.93	0.49
Soil moisture				
min. depth	c_{\min}	0.0	0.0	0.0
max. depth	c_{\max}	15.0	83.0	126.0
exponent	b	1.8	0.52	4.79
Evaporation exponent	b_e	1.1	2.5	2.0
Recharge model	type	1	1	1
	1 2 3			
parameter 1	k_g α α	320.0	85923	2722
parameter 2	S_t β	0.0	0.0	0.0
parameter 3	b_g q_{sat}	2.02	2.23	2.45
Surface storage coefficient	k_s	1.4	5.3	2.86
Baseflow storage				
exponent	m	3.0	3.0	3.0
coefficient	k_b	5.0	5.0	13.5
Constant flow	q_c	0.0	0.0	0.0
Time delay	τ	0.16	3.49	1.42
State-correction				
surface gain	g_s	1.466	1.982	-
baseflow gain	g_b	1.745	1.495	-
soil moisture gain	g_g	0	0	-

C.5 IEM Models

The units and meaning of the IEM model parameters are given in Section 1.2. Tables C.5.1 and C.5.2 summarise the calibrated values.

Table C.5.1 Comparison of IEM parameters across catchments using raingauge data

Model parameter		Catchment								
		Trout Beck	Silk Stream	Dove	Lavant	Rhondda	Brue	Stour	Roch	Witham
Runoff										
Coefficient	α	0.95	0.47	0.83		0.99	0.81	0.55	0.68	0.5
Exponent	β	0.02	0.006	0.014		0.05	0.075	0.034	0.03	0.02
Storage coeff.	k	80.00	34.00	15041.00		860.00	998.00	943.00	815.00	2252.00
Smoothing function										
Delay to start	τ_s	0.06	0	0.07		0.08	2.74	0.00	0.06	0.07
Start to peak	τ_p	0.55	0.41	0.60		0.1	1.05	0.491	0.55	0.25
Start to end	τ_e	1.0	0.94	0.61		0.5	4.6	0.493	1.0	10.0
Constant flow	q_c	0	0	0		0	0	0	0	0
Updating gain	g	1.653	1.590	1.143	-	1.593	1.876	1.774	1.649	1.270

Table C.5.2 Comparison of IEM parameters across catchments using radar data

(a) Uncalibrated radar data

Model parameter		Catchment		
		Silk Stream	Stour	Roch
Runoff				
coefficient	α	0.4	1.5	0.64
exponent	β	0.0047	0.22	0.043
Storage coeff.	k	74.0	534.0	661.0
Smoothing function				
delay to start	τ_s	0	0.059	0.001
start to peak	τ_p	0.44	4.85	2.34
start to end	τ_e	1.01	6.91	2.43
Constant flow	q_c	0	0	0
Updating gain	g	-	-	-

(b) Calibrated radar data

Model parameter		Catchment		
		Silk Stream	Stour	Roch
Runoff				
coefficient	α	0.49	0.6	0.46
exponent	β	0.008	0.22	0.063
Storage coeff.	k	41.0	438.0	520.0
Smoothing function				
delay to start	τ_s	0	1.59	0.09
start to peak	τ_p	0.49	3.65	2.34
start to end	τ_e	1.05	3.88	2.67
Constant flow	q_c	0	0	0
Updating gain	g	1.518	1.954	-

C.6 TF Models

Table C.6.1 Comparison of simulation mode TF parameters across catchments using raingauge data

Model parameters		Catchment								
		Trout Beck	Silk Stream	Dove	Lavant	Rhondda	Brue	Stour	Roch	Witham
Autoregressive parameters	δ_1	-1.571	-0.8871	-1.9189	-1.895	-1.755	-1.925	-1.967	-1.791	-1.872
	δ_2	0.589	-0.0301	0.91901	0.895	0.7579	0.926	0.967	0.793	0.873
Moving average parameters	ω_1	0.0114	0.0312	9.6×10^{-5}	0.0147	0.0138	0.0178	0.0030	0.0023	5×10^{-5}
	ω_2	0.0120	0.0206	6.4×10^{-6}	-0.0198	0.0024	-0.0041	0.0010	0.0022	1×10^{-5}
	ω_3	0.0119	0.0205	8.8×10^{-5}	-0.0041	-0.0023	-0.0010	0.0054	0.0018	1.6×10^{-5}
	ω_4	0.0134	0.0110	9.3×10^{-5}	0.0093	-0.0054	-0.0100	-0.0093	-0.0026	3.2×10^{-5}
Time delay	τ	0.5	0.5	0.25	0.5	0.25	5.0	5.25	0.25	6.5

Table C.6.2 Comparison of simulation mode TF parameters across catchments using uncalibrated radar data

Model parameters		Catchment		
		Silk Stream	Stour	Roch
Autoregressive parameters	δ_1	-0.8957	-1.9666	-1.789
	δ_2	-0.0529	0.9666	0.790
Moving average parameters	ω_1	0.0569	0.0102	0.0135
	ω_2	1.49×10^{-5}	9.88×10^{-4}	0.0199
	ω_3	6.45×10^{-3}	1.68×10^{-3}	-0.0182
	ω_4	6.50×10^{-4}	-0.0125	0.0132
Time delay	τ	0.500	7.75	2.5

Table C.6.3 Comparison of simulation mode TF parameters across catchments using calibrated radar data

Model parameters		Catchment		
		Silk Stream	Stour	Roch
Autoregressive parameters	δ_1	-0.8832	-1.967	-1.780
	δ_2	-0.0890	0.9670	0.781
Moving average parameters	ω_1	0.0288	5.345×10^{-3}	0.0120
	ω_2	0.0016	-4.155×10^{-4}	0.0130
	ω_3	1.46×10^{-4}	2.850×10^{-3}	-8.31×10^{-3}
	ω_4	2.69×10^{-3}	-7.6×10^{-3}	-1.561×10^{-2}
Time delay	τ	0.5	8.5	3.0

Table C.6.4 Comparison of forecast mode TF parameters across catchments using raingauge data

Model parameters		Catchment								
		Trout Beck	Silk Stream	Dove	Lavant	Rhondda	Brue	Stour	Roch	Witham
Autoregressive parameters	δ_1	-1.669	-1.540	-1.486	-1.233	-1.708	-1.970	-1.965	-1.814	-1.592
	δ_2	0.688	0.566	0.487	0.233	0.713	0.971	0.965	0.816	0.593
Moving average parameters	ω_1	0.0120	0.0080	9.9×10^{-5}	1.0×10^{-4}	0.0055	9.1×10^{-4}	1.4×10^{-4}	0.0018	7.1×10^{-5}
	ω_2	0.0138	0.0218	1.1×10^{-4}	7.2×10^{-5}	0.0051	4.2×10^{-6}	1.8×10^{-5}	-2.6×10^{-4}	3.4×10^{-5}
	ω_3	0.0113	0.0049	1.5×10^{-4}	5.0×10^{-5}	0.0037	1.1×10^{-5}	-7.6×10^{-6}	5.7×10^{-5}	5.0×10^{-5}
	ω_4	0.0105	0.00014	6.4×10^{-4}	5.2×10^{-5}	5.2×10^{-4}	8.4×10^{-4}	6.4×10^{-5}	0.0015	7.2×10^{-5}
Time delay	τ	0.5	0.25	0.25	0.5	0.25	5.0	5.25	0.25	6.5
Smoothing factor	μ	0.9999	0.990	0.536	0.967	0.997	0.612	0.144	0.278	0.531

Table C.6.5 Comparison of forecast mode TF parameters across catchments using calibrated radar data

Model parameters		Catchment		
		Silk Stream	Stour	Roch
Autoregressive parameters	δ_1	-1.620	-1.751	-
	δ_2	0.652	0.751	-
Moving average parameters	ω_1	0.0500	6.2×10^{-4}	-
	ω_2	0.0187	2.7×10^{-4}	-
	ω_3	0.0194	2.6×10^{-4}	-
	ω_4	-0.0153	-1.3×10^{-4}	-
Time delay	τ	0.5	8.5	-
Smoothing factor	μ	0.999	0.524	-

C.7 PRTF Models

Table C.7.1 Comparison of simulation mode PRTF parameters across catchments using raingauge data

Model parameters		Catchment								
		Trout Beck	Silk Stream	Dove	Lavant	Rhondda	Brue	Stour	Roch	Witham
Time to peak	t_{peak}	5.296	11.48	338.5		77.73	58.92	59.6	1.807	100.2
Moving average parameters	ω_1	-0.0037	0.0572	0.0129		0.013	0.024	4.9×10^{-5}	0.0133	8.1×10^{-4}
	ω_2	0.0090	-0.0258	-0.0265		-6.6×10^{-4}	-0.0045	3.2×10^{-5}	-0.012	-2.6×10^{-4}
	ω_3	-0.0131	-0.0303	0.0175		9.9×10^{-4}	-0.033	9.0×10^{-8}	0.0044	-3.8×10^{-4}
	ω_4	0.0604	0.0055	-0.0040		-0.0127	0.0143	3.6×10^{-4}	-0.005	-1.4×10^{-4}
Time delay	τ	0.5	0.5	0.25		0.25	5.0	5.25	0.25	6.5

Table C.7.2 Comparison of simulation mode PRTF parameters across catchments using uncalibrated radar data

Model parameters		Catchment		
		Silk Stream	Stour	Roch
Time to peak	t_{peak}	17.392	147.1	201.4
Moving average parameters	ω_1	0.0515	-0.0029	0.01165
	ω_2	-0.0369	0.1440	-0.0136
	ω_3	-0.0205	-6.41×10^{-5}	0.00486
	ω_4	0.0093	-0.01132	-0.00283
Time delay	τ	0.5	6.5	1.25

Table C.7.3 Comparison of simulation mode PRTF parameters across catchments using calibrated radar data

Model parameters		Catchment		
		Silk Stream	Stour	Roch
Time to peak	t_{peak}	40.056	114.9	257.7
Moving average parameters	ω_1	0.05416	-0.0044	0.104
	ω_2	-0.0469	0.01233	-0.0151
	ω_3	-0.0161	-1.2887	0.00624
	ω_4	0.0097	-0.0077	-0.00157
Time delay	τ	0.5	7.0	1.25

Table C.7.4 Comparison of forecast mode PRTF parameters across catchments using raingauge data

Model parameters		Catchment								
		Trout Beck	Silk Stream	Dove	Lavant	Rhondda	Brue	Stour	Roch	Witham
Time to peak	t_{peak}	6.982	7.194	164.5		11.69	54.4	97.0	15.9	90.0
Moving average parameters	ω_1	0.0084	0.0243	-2.0×10^{-4}		0.0053	7.3×10^{-4}	1.1×10^{-4}	0.0018	2.7×10^{-5}
	ω_2	0.0101	-4.8×10^{-6}	6.4×10^{-5}		0.0044	-7.9×10^{-5}	-1.1×10^{-6}	-3.4×10^{-4}	-9.3×10^{-6}
	ω_3	0.0074	-6.1×10^{-4}	1.7×10^{-5}		0.0024	-4.9×10^{-5}	-2.7×10^{-5}	3.4×10^{-5}	7.1×10^{-6}
	ω_4	0.0067	0.0111	2.5×10^{-4}		-0.0014	5.3×10^{-4}	3.4×10^{-5}	0.0015	9.3×10^{-6}
Time delay	τ	0.5	0.5	0.25		0.25	5.0	5.25	0.25	6.5
Smoothing factor	μ	0.905	0.995	0.661		0.073	0.989	0.24	0.223	0.0044

Table C.7.5 Comparison of forecast mode PRTF parameters across catchments using calibrated radar data

Model parameters		Catchment		
		Silk Stream	Stour	Roch
Time to peak	t_{peak}	4.029	167.41	
Moving average parameters	ω_1	0.0529	7.8×10^{-5}	
	ω_2	-0.0232	-1.64×10^{-5}	
	ω_3	0.0217	1.2×10^{-4}	
	ω_4	-0.0164	-3.6×10^{-4}	
Time delay	τ	0.5	7.0	
Smoothing factor	μ	0.993	0.484	-

C.8 Grid Models

Table C.8.1 Comparison of Grid Model parameters across catchments using raingauge data

Description	Parameter	Stour	Roch	Witham
Rainfall correction factor	f_r	0.6	0.667	0.57
Storage threshold deficit (or root constant) in evaporation function	D^*	1.0	1.0	1.0
Proportion of total storage capacity initially full	S_0	0.4	0.377	0.85
Regional upper limit of gradient	g_{max}	13.0	4.0	9.0
Regional upper limit of storage capacity	c_{max}	75.0	80.0	110.0
Maximum infiltration rate	i_{max}	5×10^7	5×10^7	5×10^7
Storage constant of (cubic) drainage function	k_d	2.0×10^{-6}	5×10^{-6}	1×10^{-7}
Wave speed parameter for routing direct runoff	θ_s	0.5	0.232	0.3
Wave speed parameter for routing drainage	θ_b	0.4	0.52	0.1
Advection velocity of flow along land path	v_L	6.0	100.0	2.0
Advection velocity of flow along river path	V_R	120.0	357.0	135.0

Table C.8.2 Comparison of Grid Model parameters across catchments using uncalibrated radar data

Description	Parameter	Stour	Roch	Witham
Rainfall correction factor	f_r	1.35	0.667	-
Storage threshold deficit (or root constant) in evaporation function	D^*	1.0	1.0	-
Proportion of total storage capacity initially full	S_0	0.5	0.636	-
Regional upper limit of gradient	g_{max}	11.75	21.25	-
Regional upper limit of storage capacity	c_{max}	80.0	80.0	-
Maximum infiltration rate	i_{max}	5×10^7	5×10^7	-
Storage constant of (cubic) drainage function	k_d	2.8×10^{-6}	5×10^{-6}	-
Wave speed parameter for routing direct runoff	θ_s	0.554	0.225	-
Wave speed parameter for routing drainage	θ_b	0.445	0.298	-
Advection velocity of flow along land path	v_L	5.54	100.0	-
Advection velocity of flow along river path	V_R	102.0	357.0	-

Table C.8.3 Comparison of Grid Model parameters across catchments using calibrated radar data

Description	Parameter	Stour	Roch	Witham
Rainfall correction factor	f_r	0.697	0.5	-
Storage threshold deficit (or root constant) in evaporation function	D^*	1.0	1.0	-
Proportion of total storage capacity initially full	S_0	0.8	0.636	-
Regional upper limit of gradient	g_{max}	11.75	21.25	-
Regional upper limit of storage capacity	c_{max}	80.0	80.0	-
Maximum infiltration rate	i_{max}	5×10^7	5×10^7	-
Storage constant of (cubic) drainage function	k_d	2.9×10^{-6}	5×10^{-6}	-
Wave speed parameter for routing direct runoff	θ_s	0.523	0.225	-
Wave speed parameter for routing drainage	θ_b	0.313	0.298	-
Advection velocity of flow along land path	v_L	6.93	100.0	-
Advection velocity of flow along river path	V_R	114.5	357.0	-

APPENDIX D: SIMULATED AND FORECAST FLOW HYDROGRAPHS

Simulated and forecast flow hydrographs for typical events for each study catchment are presented in this Appendix.

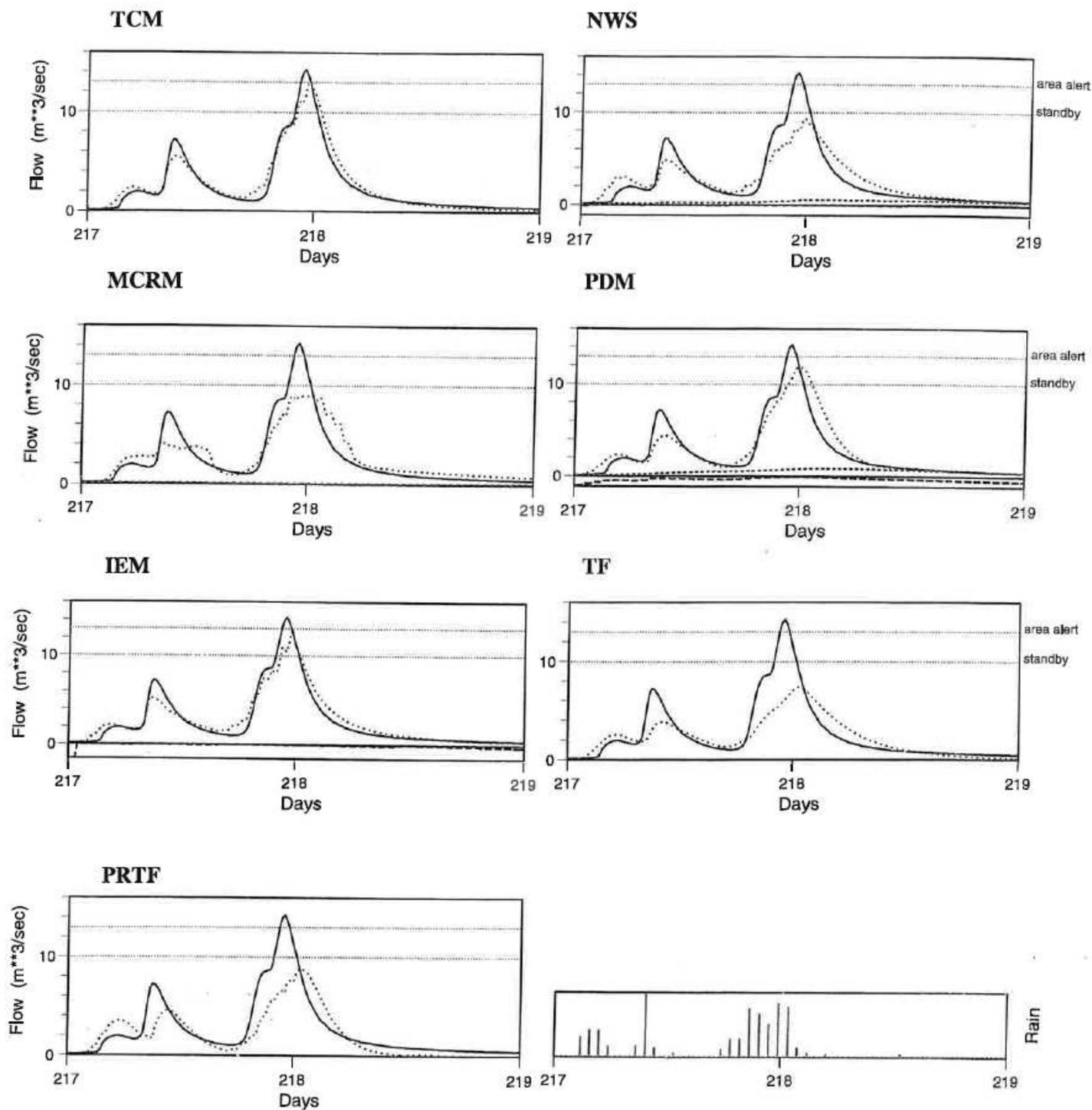


Figure D.1 Simulated flow hydrographs from different models: Trout Beck at Moor House, 16/11/95-17/11/95. Observed flow: bold line; simulated flow: dotted line; simulated baseflow: dashed line (TCM, PDM only); soil moisture deficit: long-dashed line (NWS, PDM, IEM only).

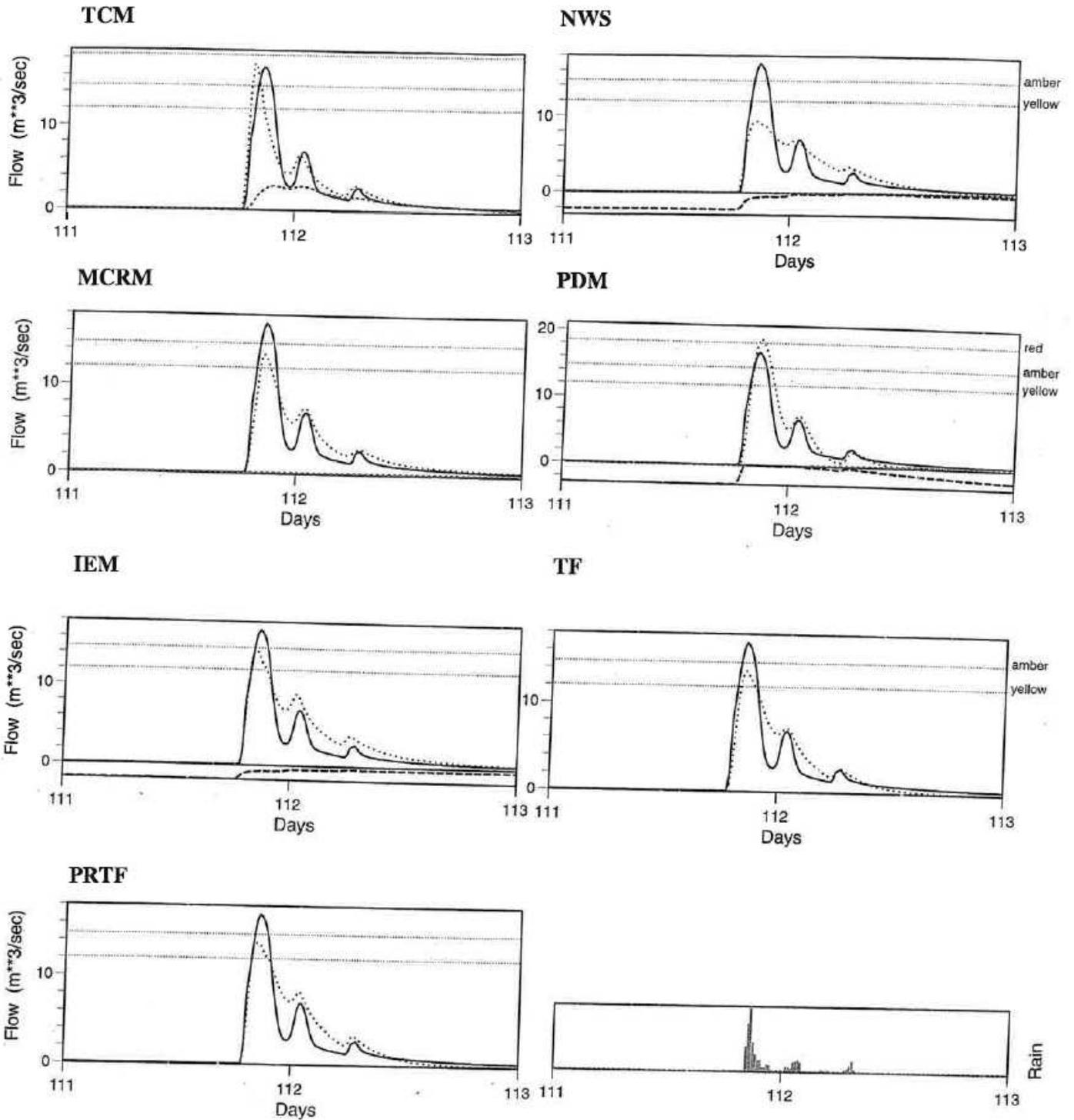


Figure D.2 Simulated flow hydrographs from different models: Silk Stream at Colindeep Lane, 21/6/92-22/6/92. Observed flow: bold line; simulated flow: dotted line; simulated baseflow: dashed line (TCM, PDM only); soil moisture deficit: long-dashed line (NWS, PDM, IEM only).

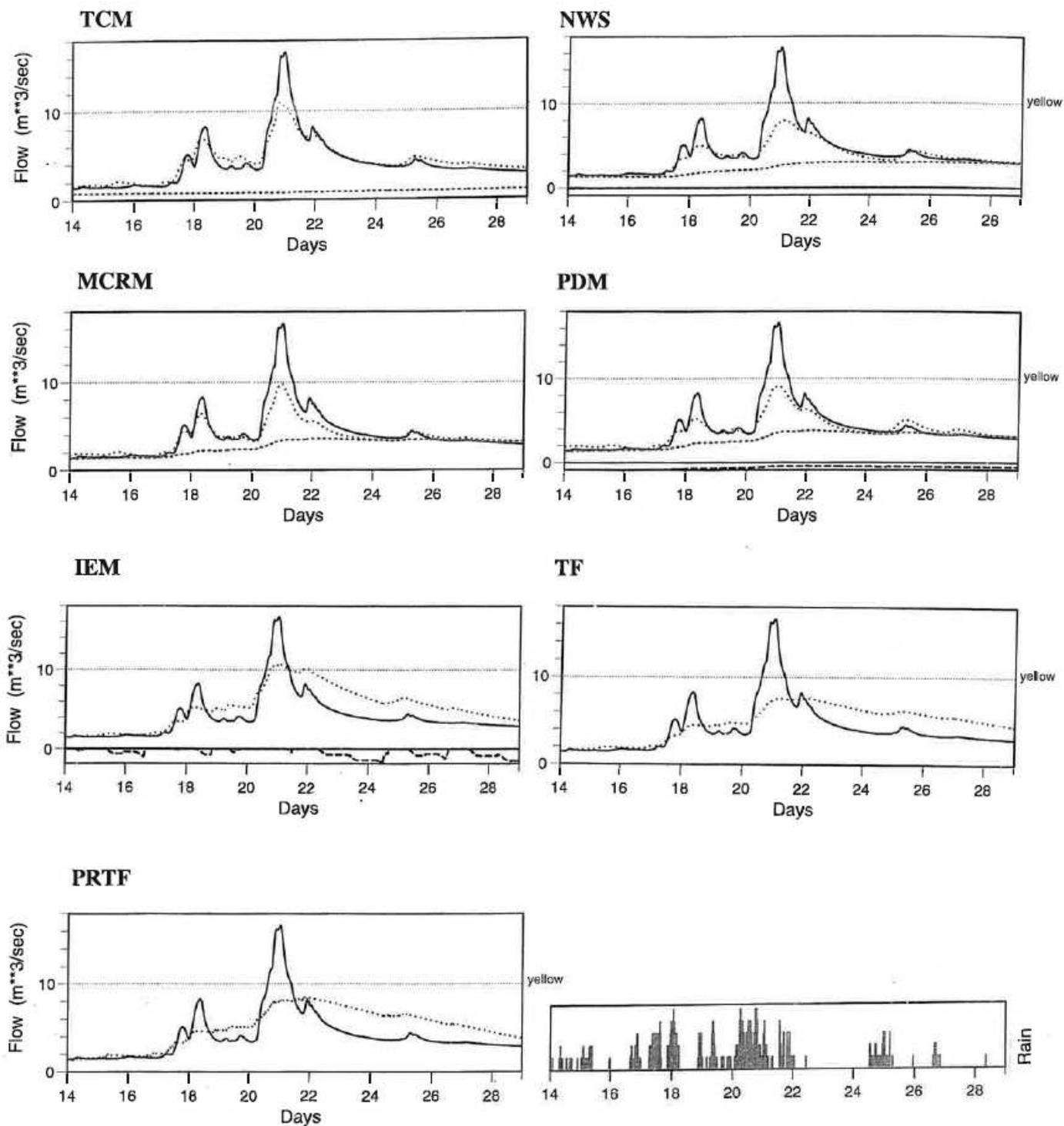


Figure D.3 Simulated flow hydrographs from different models: Dove at Izaak Walton, 1/3/98-15/3/98. Observed flow: bold line; simulated flow: dotted line; simulated baseflow: dashed line (TCM, PDM only); soil moisture deficit: long-dashed line (NWS, PDM, IEM only).

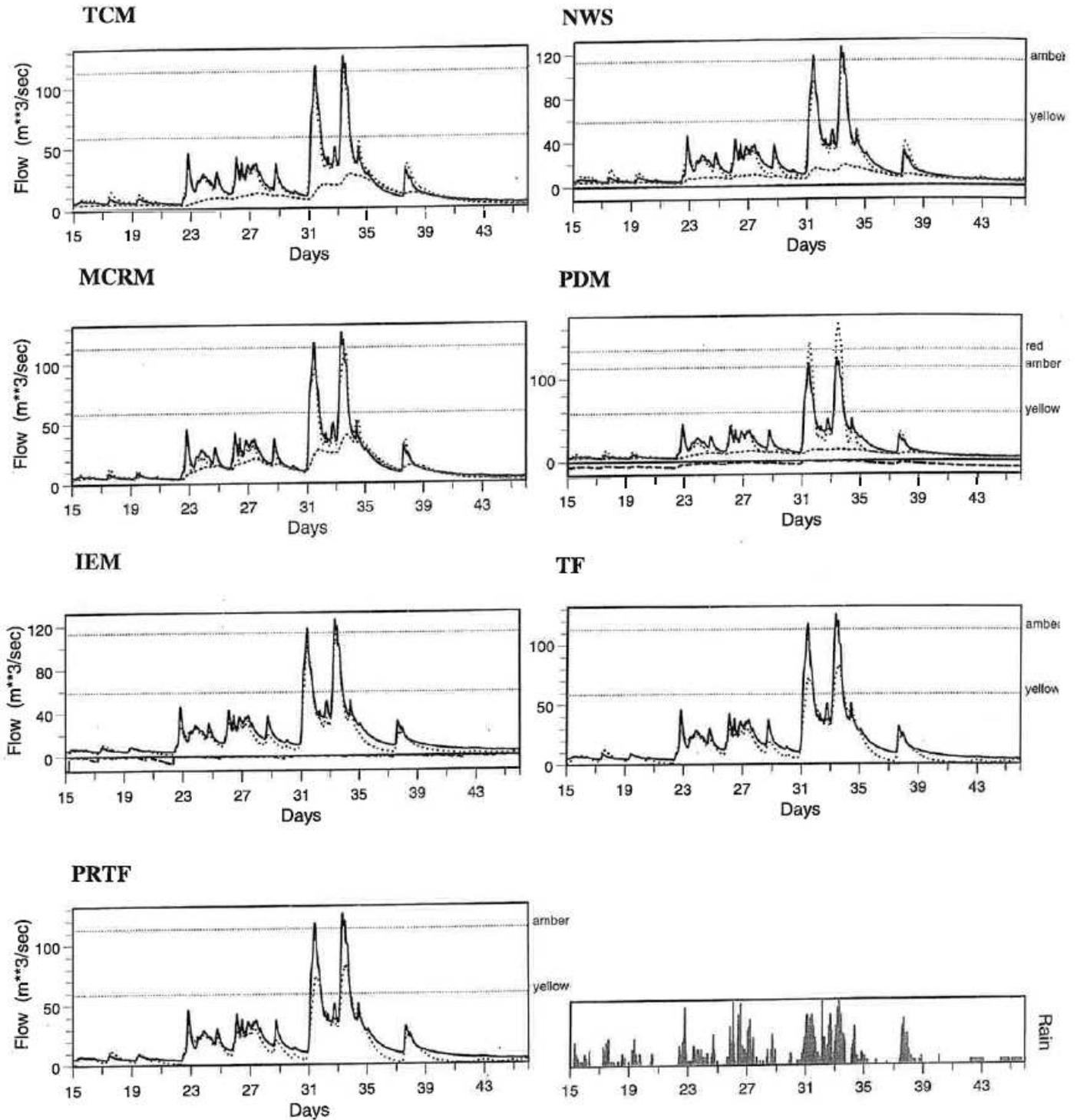


Figure D.4 Simulated flow hydrographs from different models: Rhondda at Trehafod, 15/11/92-15/12/92. Observed flow: bold line; simulated flow: dotted line; simulated baseflow: dashed line (TCM, PDM only); soil moisture deficit: long-dashed line (NWS, PDM, IEM only).

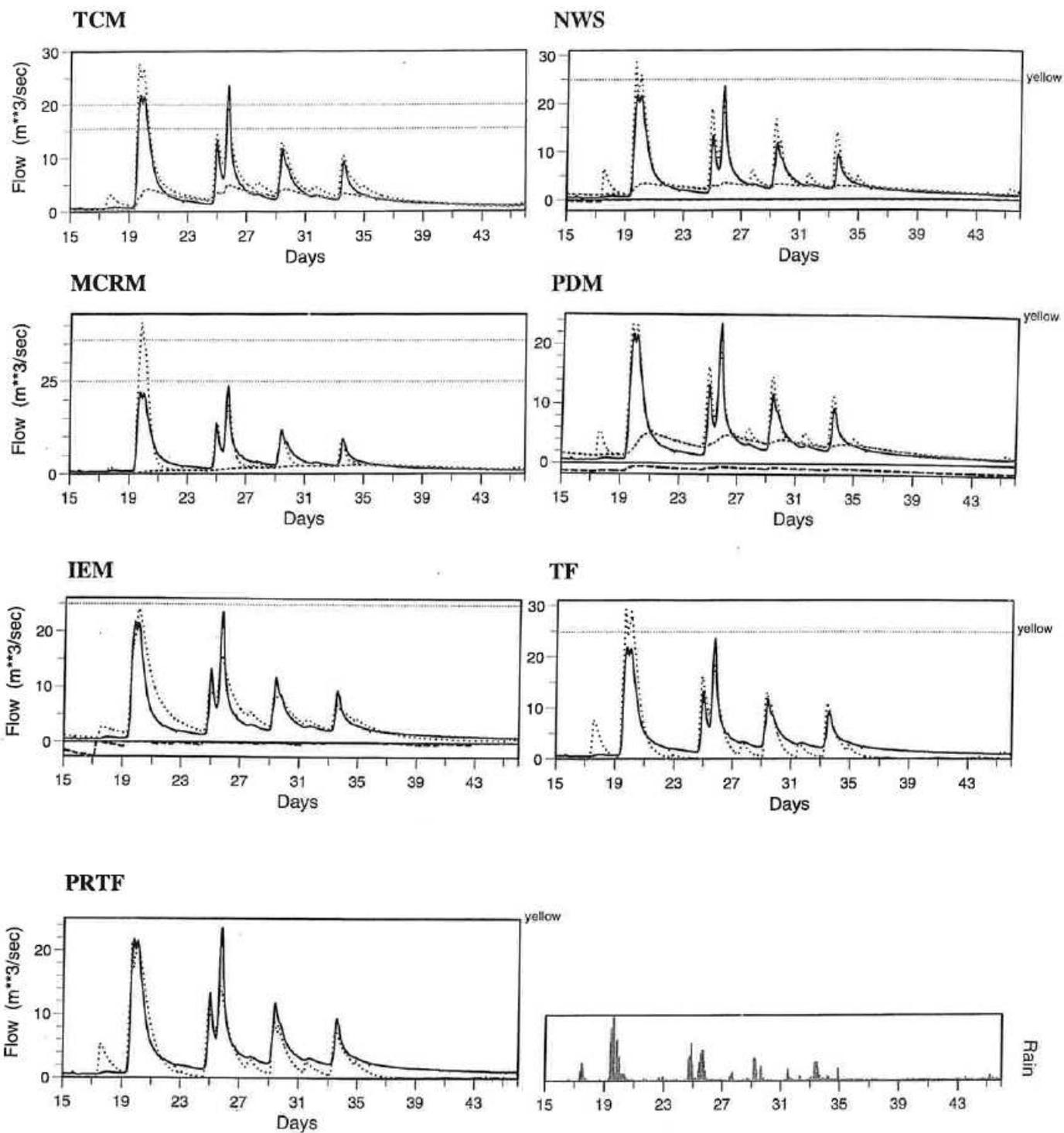


Figure D.5 Simulated flow hydrographs from different models: Brue at Lovington, 16/11/96-16/12/96. Observed flow: bold line; simulated flow: dotted line; simulated baseflow: dashed line (TCM, PDM only); soil moisture deficit: long-dashed line (NWS, PDM, IEM only).

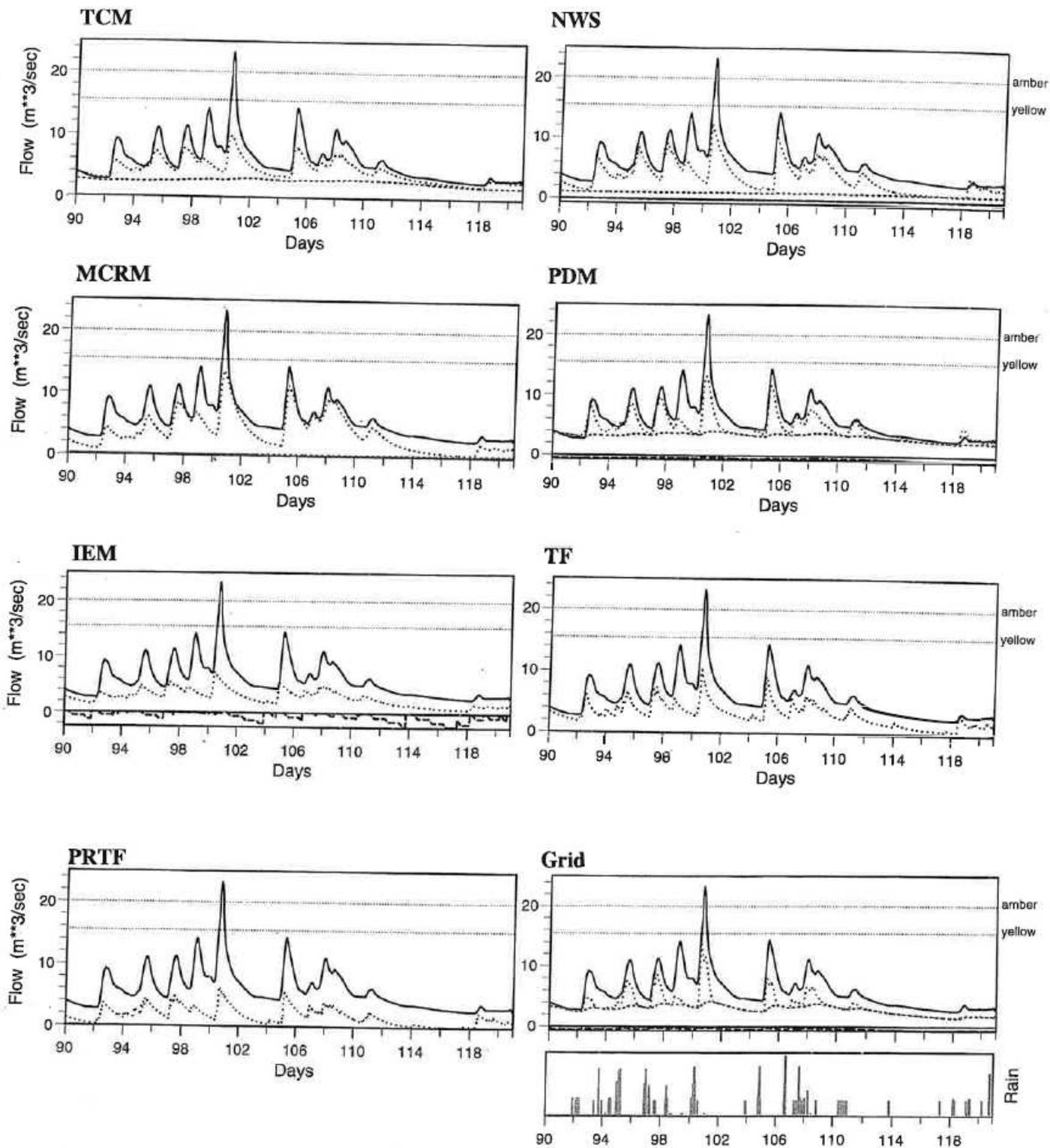


Figure D.6 Simulated flow hydrographs from different models: Stour at Shipston, 27/12/93-26/1/94. Observed flow: bold line; simulated flow: dotted line; simulated baseflow: dashed line (TCM, PDM only); soil moisture deficit: long-dashed line (NWS, PDM, IEM only).

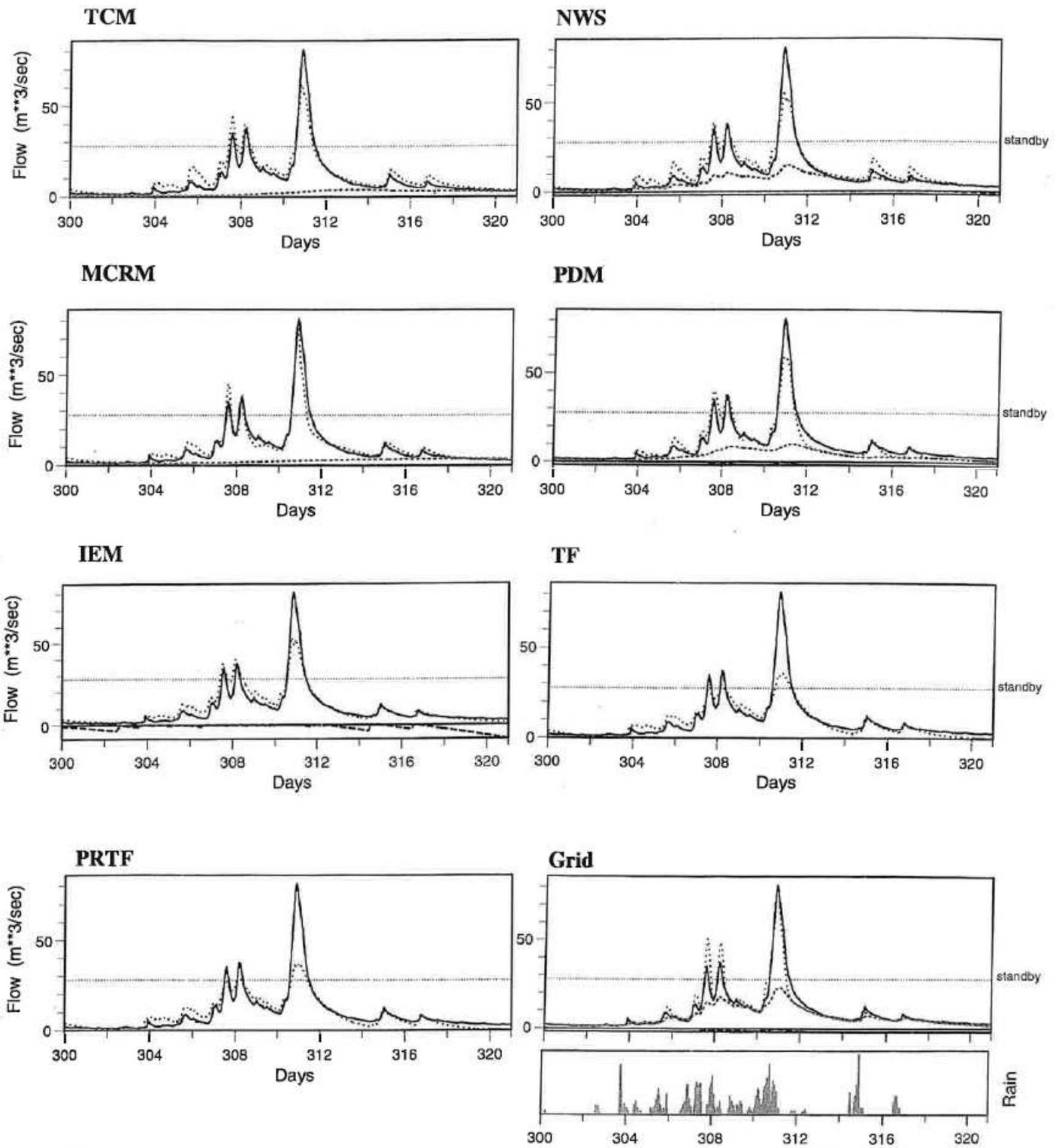


Figure D.7 Simulated flow hydrographs from different models: Roch at Blackford Bridge, 25/2/98-17/3/98. Observed flow: bold line; simulated flow: dotted line; simulated baseflow: dashed line (TCM, PDM only); soil moisture deficit: long-dashed line (NWS, PDM, IEM only).

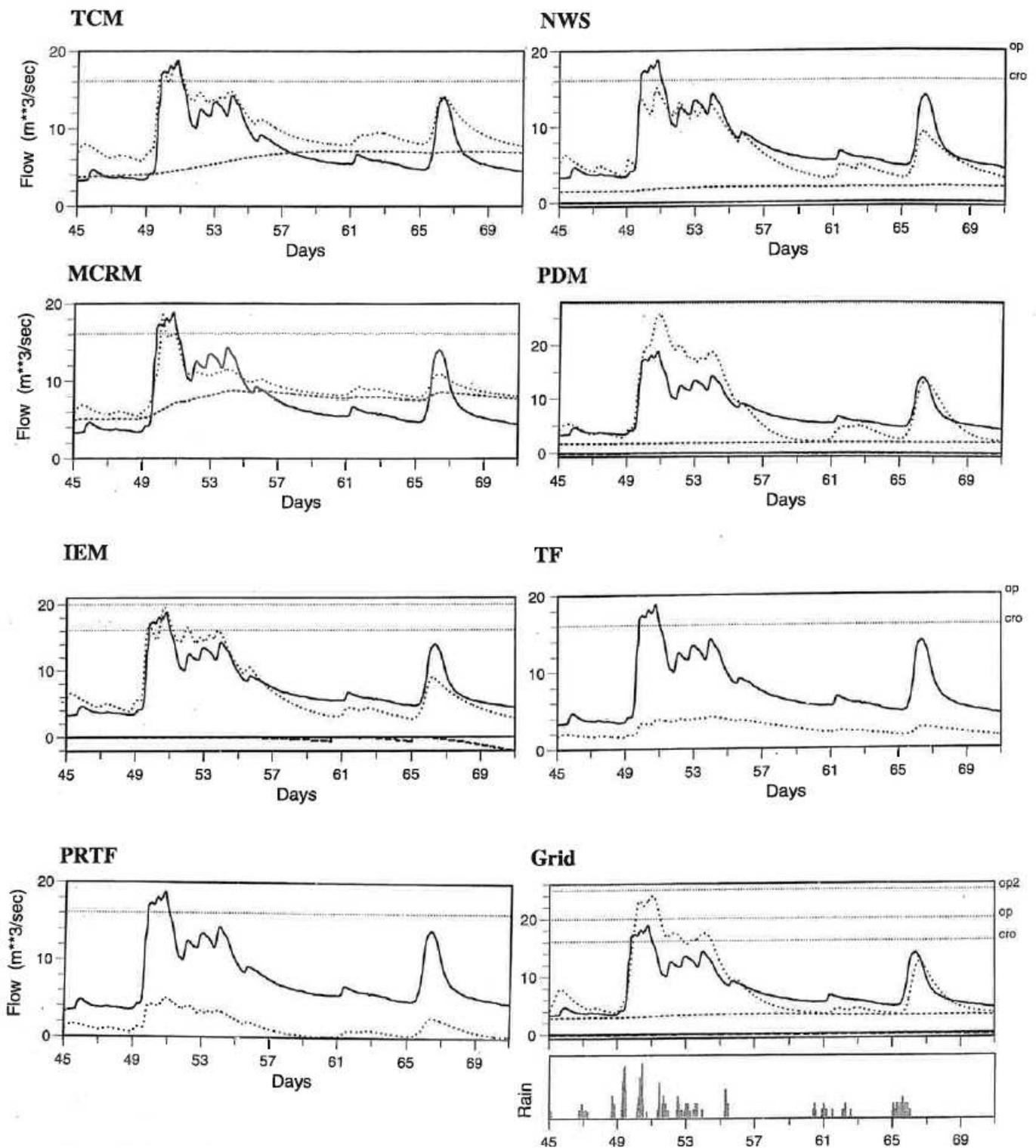


Figure D.8 Simulated flow hydrographs from different models: Witham at Claypole Mill, 30/12/97-24/1/98. Observed flow: bold line; simulated flow: dotted line; simulated baseflow: dashed line (TCM, PDM only); soil moisture deficit: long-dashed line (NWS, PDM, IEM only).

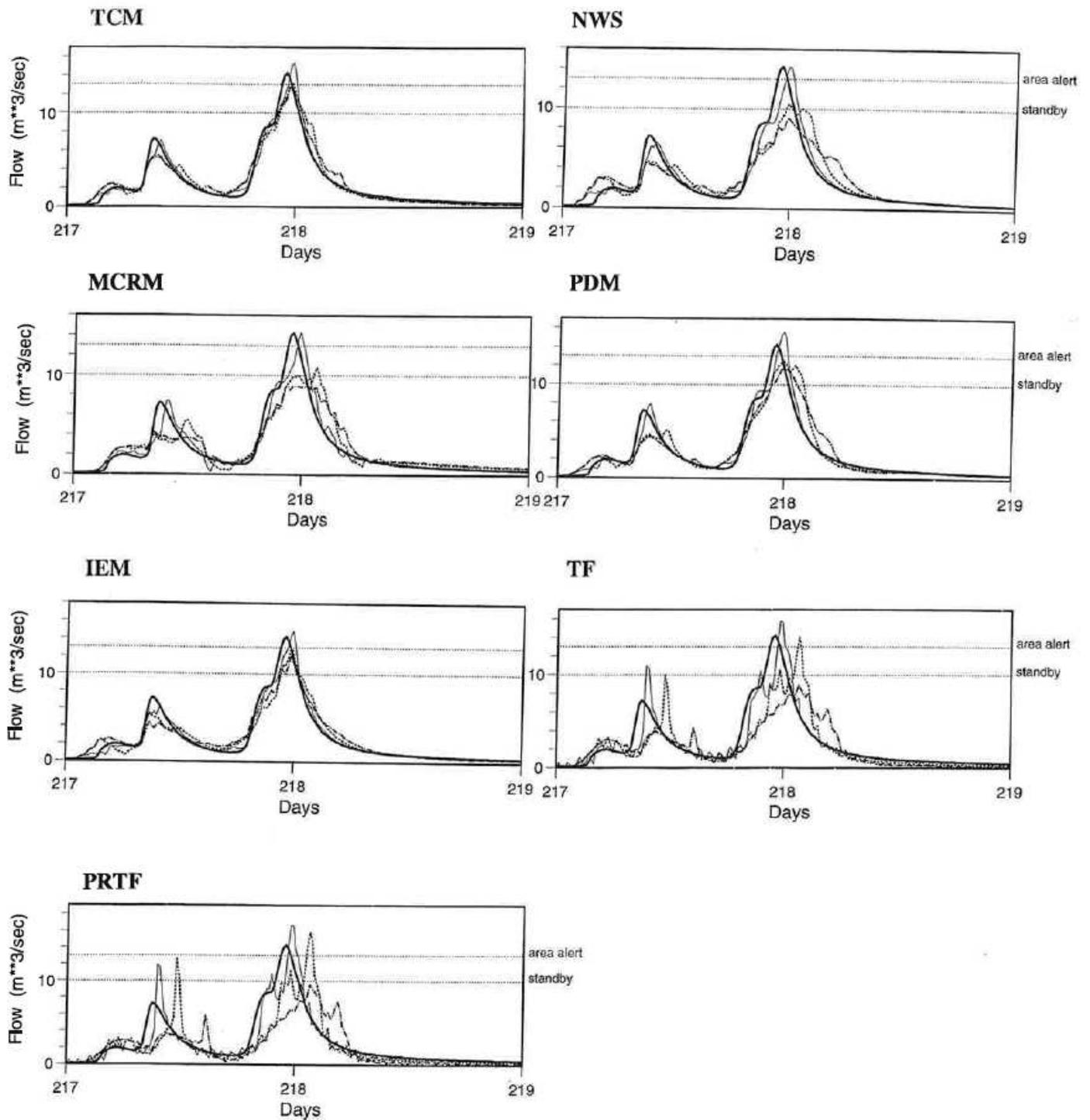


Figure D.9 Fixed lead-time evaluation forecasts for different models: Trout Beck at Moor House, 16/11/95-17/11/95. Observed flow: bold line; 1 hour ahead: thin line; 3 hours ahead: short dashes; 6 hours ahead: dots and dashes.

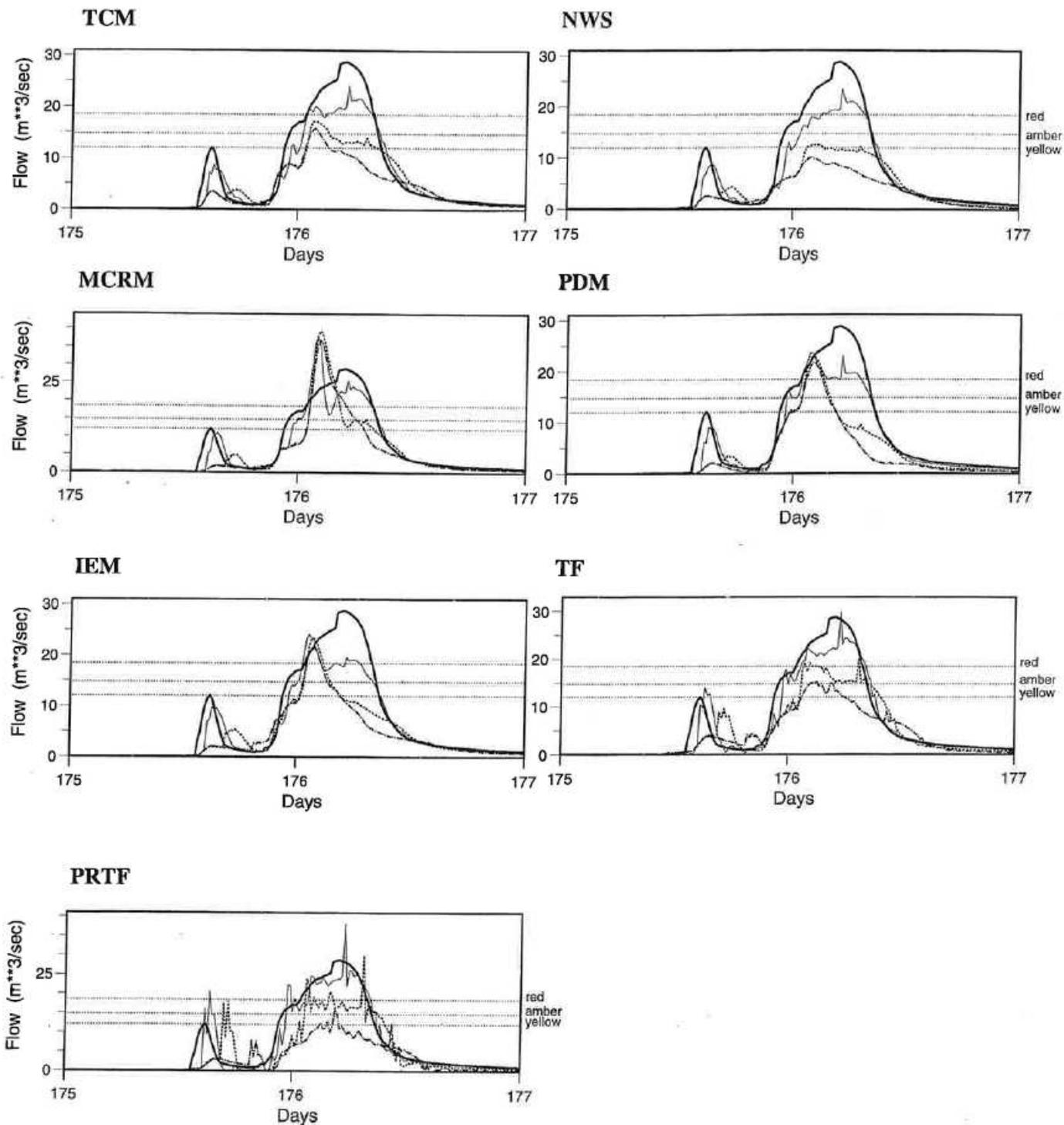


Figure D.10 Fixed lead-time evaluation forecasts for different models: Silk Stream at Colindeep Lane, 23/9/92-24/9/92. Observed flow: bold line; 1 hour ahead: thin line; 3 hours ahead: short dashes; 6 hours ahead: dots and dashes.

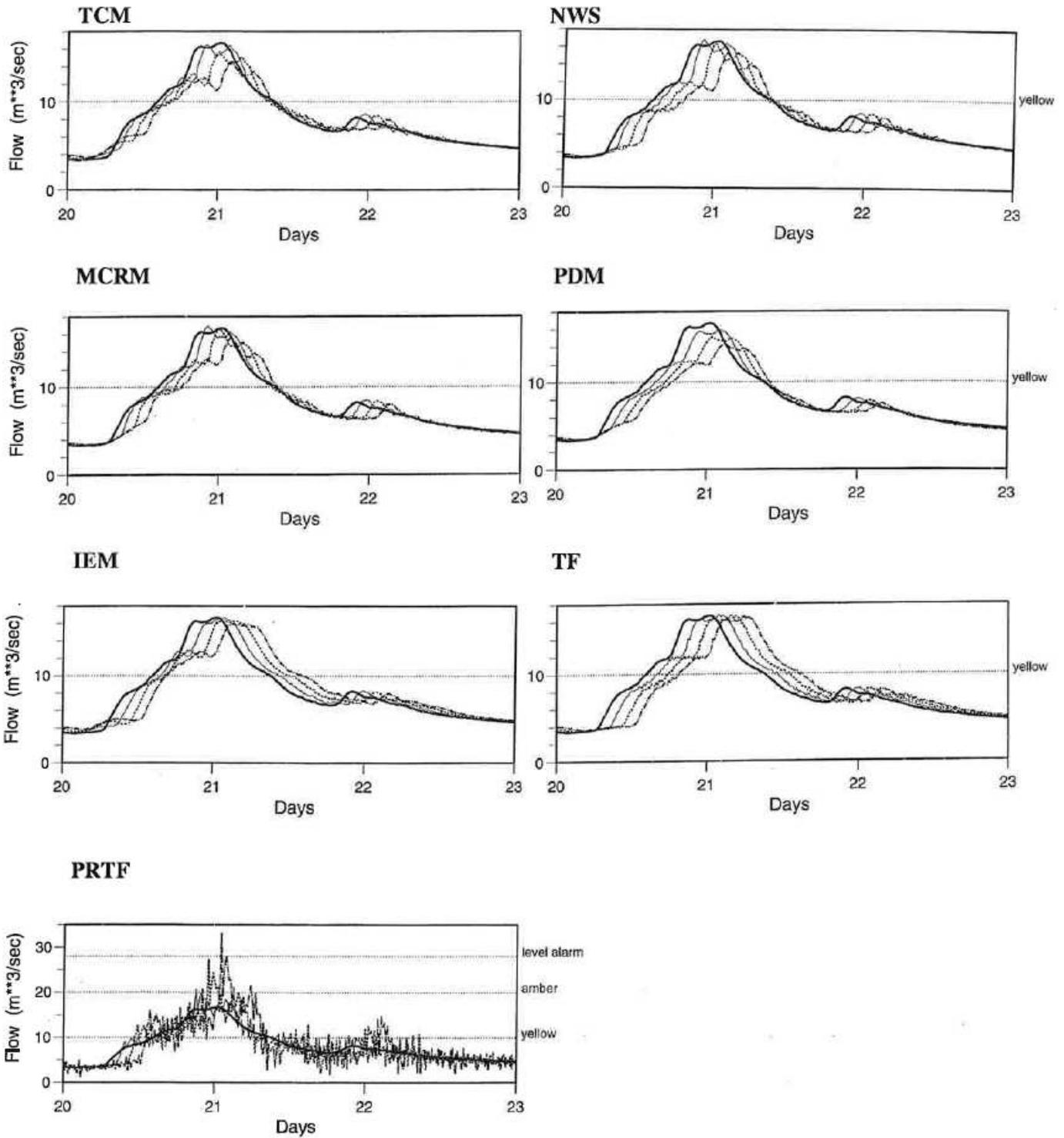


Figure D.11 Fixed lead-time evaluation forecasts for different models: Dove at Izaak Walton, 7/3/98-9/3/98. Observed flow: bold line; 2 hours ahead: thin line; 4 hours ahead: short dashes; 6 hours ahead: dots and dashes.

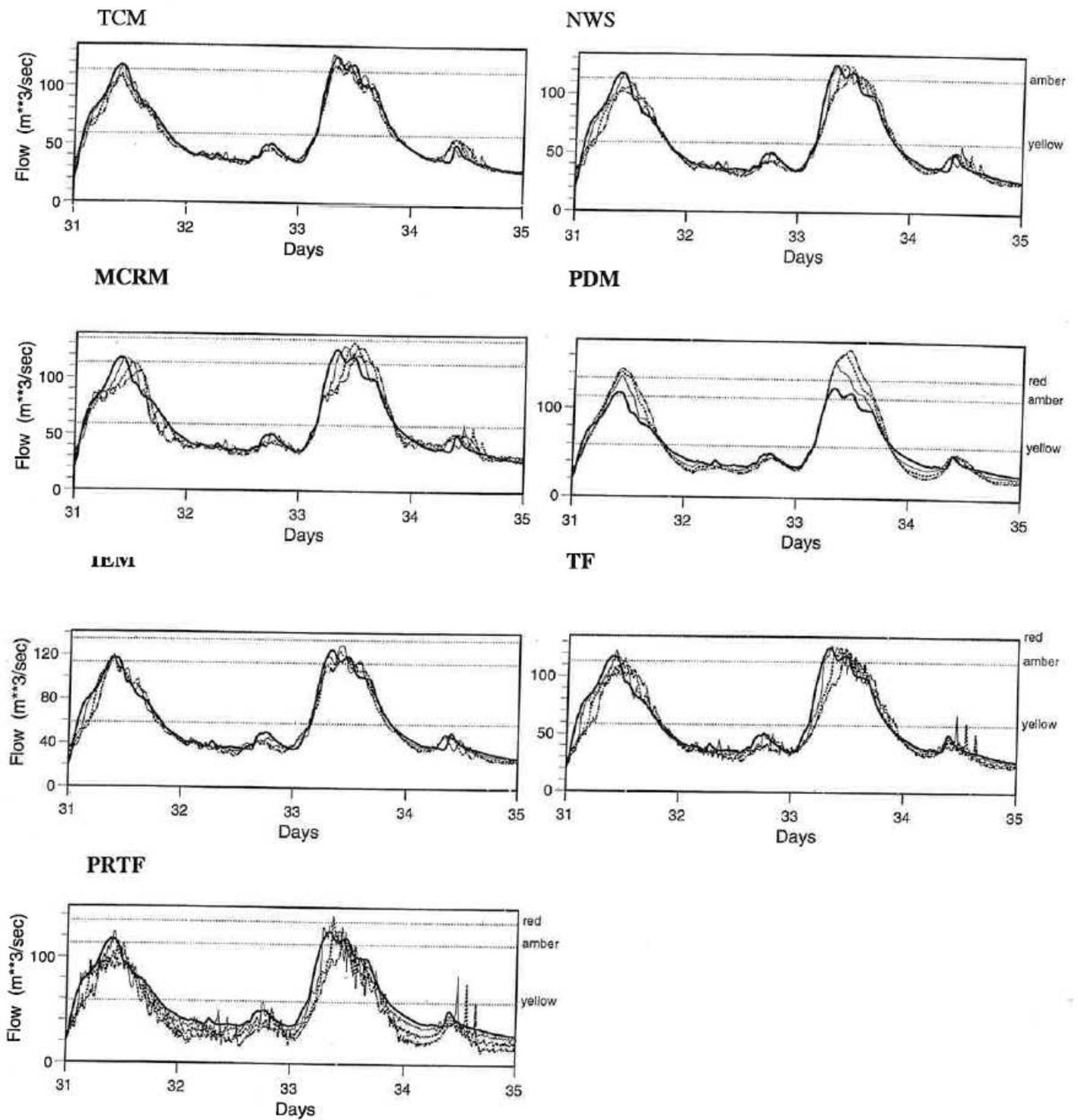


Figure D.12 Fixed lead-time evaluation forecasts for different models: Rhondda at Trehafod, 1/12/92-4/12/92. Observed flow: bold line; 2 hours ahead: thin line; 4 hours ahead: short dashes; 6 hours ahead: dots and dashes.

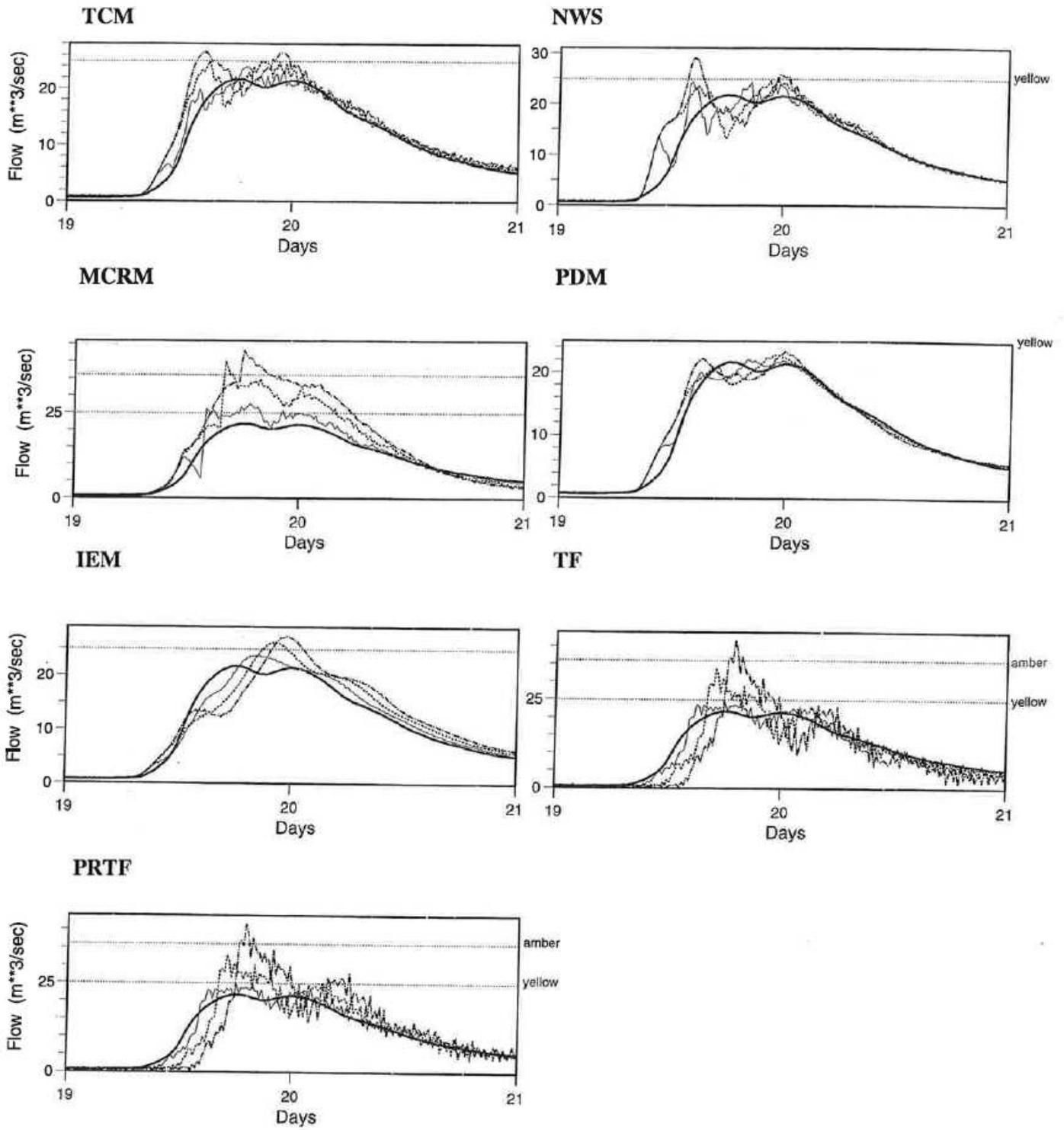


Figure D.13 Fixed lead-time evaluation forecasts for different models: Brue at Lovington, 20/11/96-21/11/96. Observed flow: bold line; 2 hours ahead: thin line; 4 hours ahead: short dashes; 6 hours ahead: dots and dashes.

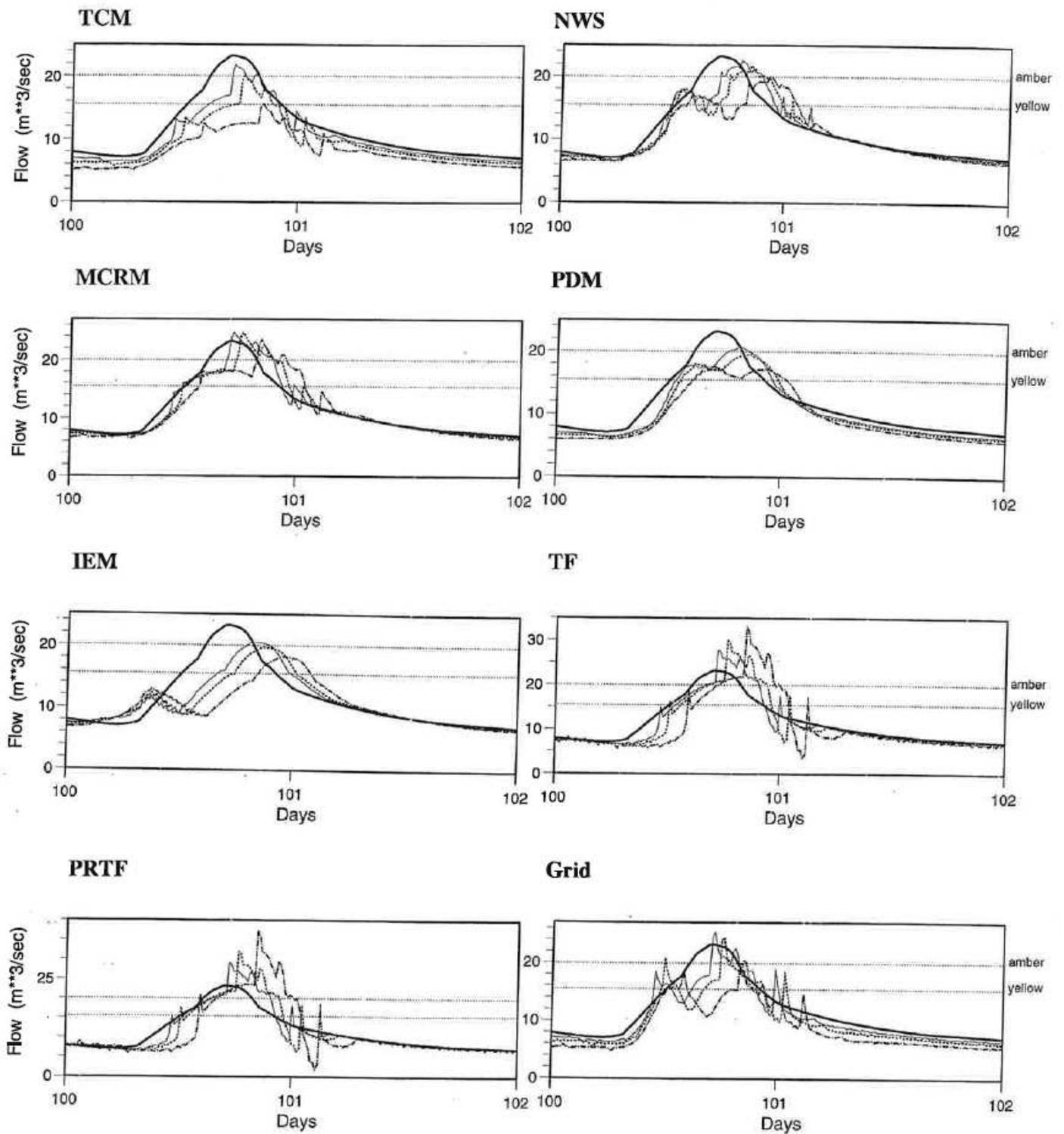


Figure D.14 Fixed lead-time evaluation forecasts for different models: Stour at Shipston, 6/1/94-7/1/94. Observed flow: bold line; 3 hours ahead: thin line; 4 hours ahead: short dashes; 6 hours ahead: dots and dashes.

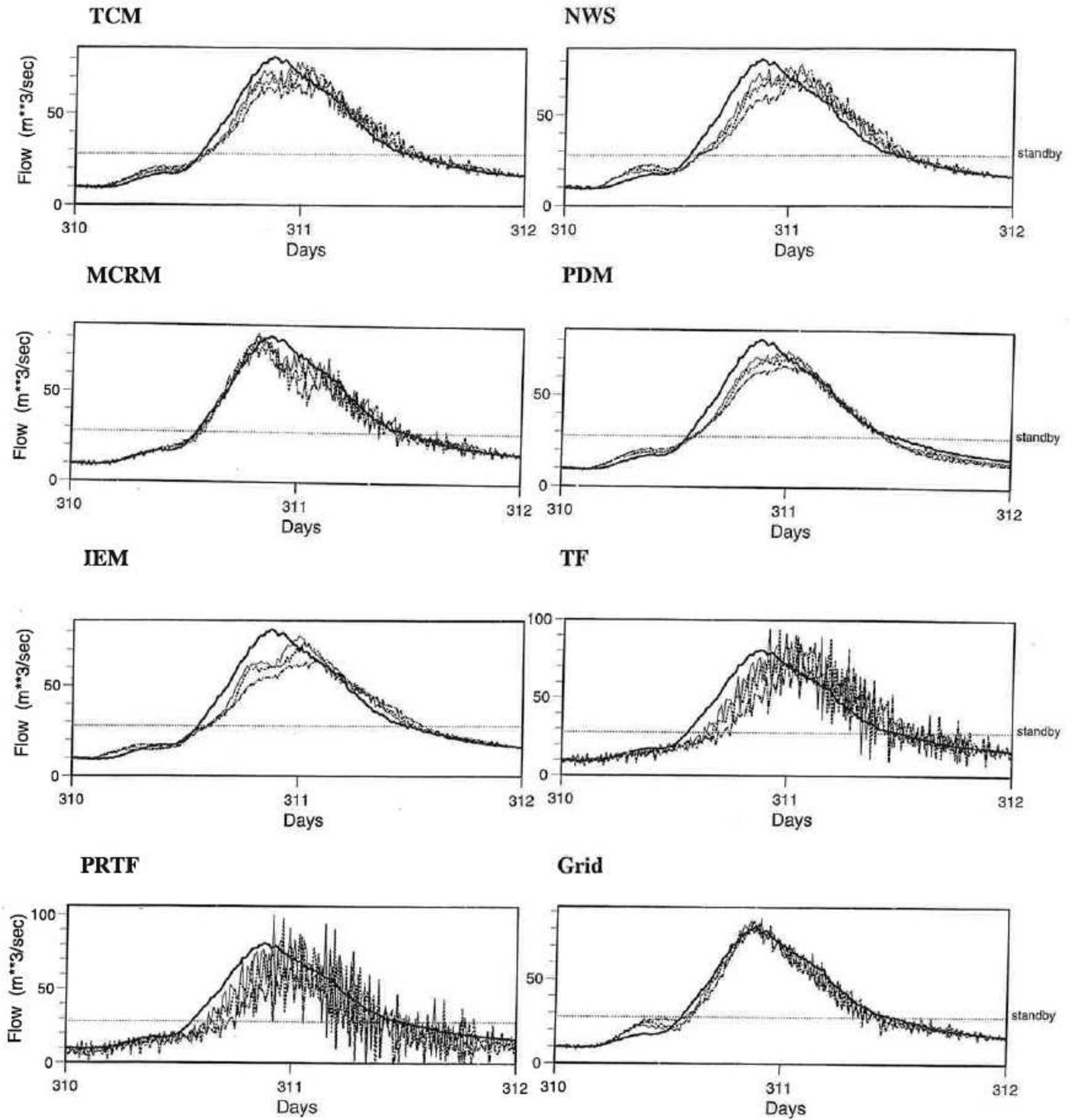


Figure D.15 Fixed lead-time evaluation forecasts for different models: Roch at Blackford Bridge, 7/3/98-8/3/98. Observed flow: bold line; 3 hours ahead: thin line; 4 hours ahead: short dashes; 6 hours ahead: dots and dashes.

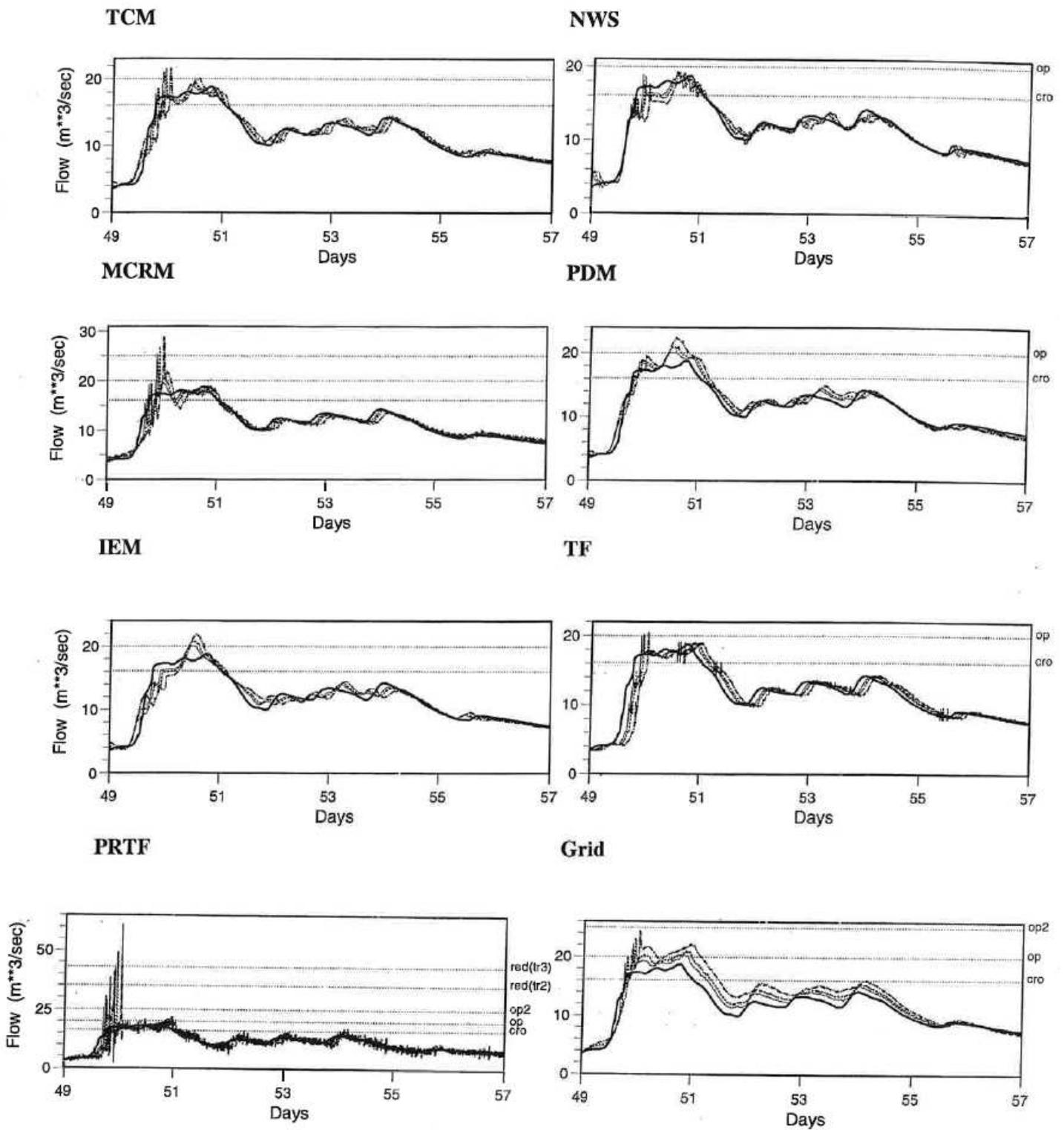


Figure D.16 Fixed lead-time evaluation forecasts for different models: Witham at Claypole Mill, 3/1/98-6/1/98. Observed flow: bold line; 3 hours ahead: thin line; 4 hours ahead: short dashes; 6 hours ahead: dots and dashes.

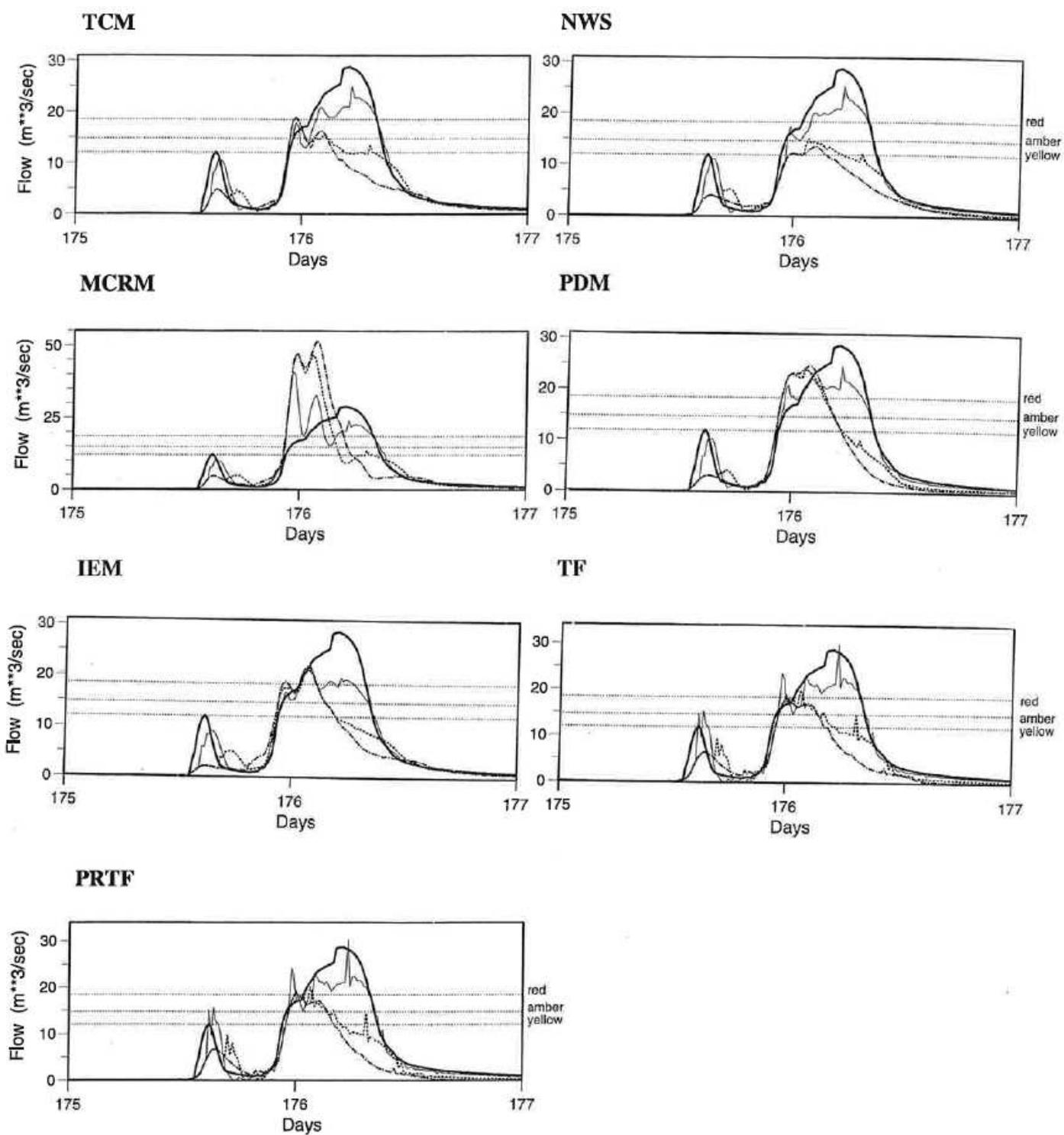


Figure D.17 Fixed lead-time evaluation forecasts for different models using calibrated radar data: Silk Stream at Colindeep Lane, 23/9/92-24/9/92. Observed flow: bold line; 1 hour ahead: thin line; 3 hours ahead: short dashes; 6 hours ahead: dots and dashes.

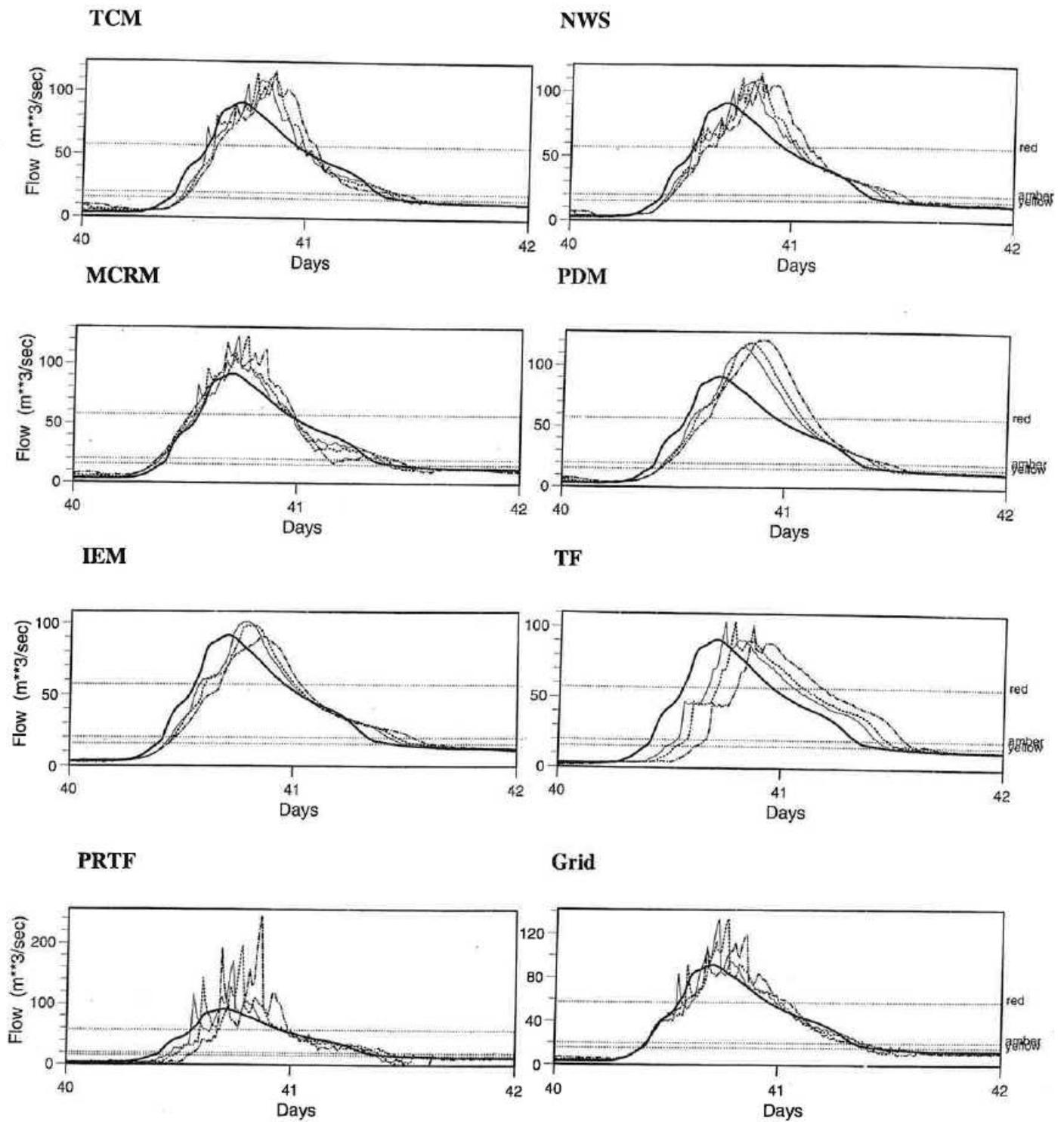


Figure D.18 Fixed lead-time evaluation forecasts for different models using calibrated radar data: Stour at Shipston, 10/4/98-11/4/98. Observed flow: bold line; 3 hours ahead: thin line; 4 hours ahead: short dashes; 6 hours ahead: dots and dashes.