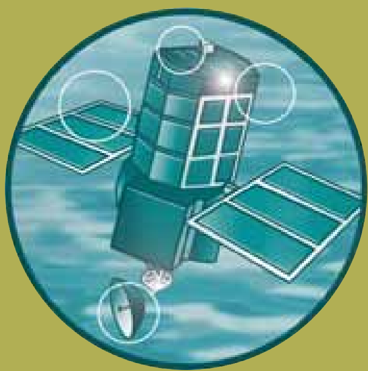
Joint Defra/EA Flood and Coastal Erosion Risk Management R&D Programme

National river catchment flood frequency method using continuous simulation

R&D Technical Report FD2106/TR



Environment Agency

defra
Department for Environment Food and Rural Affairs

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

National river catchment flood frequency method using continuous simulation

R&D Technical Report FD2106/TR

Authors: Ann Calver, Sue Crooks, David Jones, Alison Kay, Thomas Kjeldsen and Nick Reynard

Produced: March 2005

Statement of use
Quantification of river flood frequencies, and associated uncertainties, at research code level.

Dissemination status
Internal: Released Internally.
External: Released to Public Domain.

Keywords – River floods; modelling; continuous simulation; ungauged sites; flood frequency; uncertainty.

Research contractor – CEH Wallingford (formerly the Institute of Hydrology), Maclean Building, Crowmarsh Gifford, Wallingford, Oxon OX10 8BB anc@eh.ac.uk

Defra Project Officer – Professor Edward Evans, Halcrow

Publishing organisation

Defra - Flood Management Division
Ergon House
Horseferry Road
London SW1P 2AL
Tel: 020 7238 3000
Fax: 020 7238 6187
www.defra.gov.uk/environ/fcd

© Crown copyright (Defra); September 2005

Copyright in the typographical arrangement and design rests with the Crown.
This publication (excluding the logo) may be reproduced free of charge in any format or medium provided that it is reproduced accurately and not used in a misleading context. The material must be acknowledged as Crown copyright with the title and source of the publication specified. The views expressed in this document are not necessarily those of Defra or 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 on views contained herein.

Published by the Department for Environment, Food and Rural Affairs. Printed in the UK, September 2005 on recycled material containing 80% post-consumer waste and 20% totally chlorine free virgin pulp.

PB No. 11208 ISBN: 0-85521-164-4

CONTENTS

Technical Report

SUMMARY

This is the Technical Report for project FD2106 *National river catchment flood frequency method using continuous simulation*. It also serves a reporting role to the Scottish Executive who funded inclusion of Scotland in the work, the Defra – Environment Agency remit covering England and Wales. The term 'national' therefore refers to Britain.

The Technical Report contains the main text which describes the development of research approaches, analyses of results and recommendations. The figures and tables of the Appendices (Project Record) comprehensively detail the results.

The first chapter of the main text introduces the aims of the project, the potential advantages of a continuous simulation approach to river flood frequency quantification, and the specific project objectives. These are set in the context of other recent and current research projects of the Flood Management R&D programme.

Data, modelling and extrapolation issues are plainly all interlinked: for convenience they are handled in separate chapters of the report. Chapter 2 covers selection of the data-rich catchments (119 in total) which serve as the basis for extension of the method to full spatial coverage across Britain. The sources and nature of precipitation, evaporation and river flow data are detailed, together with the methods established for quality control and infilling. Both hourly and daily data were used, in order to define detail but also widen availability. Data required for the whole domain of potential application of the method, the more widely-available catchment properties, are also defined and their derivation discussed.

Chapter 3 describes the runoff models used in representing catchment processes defining the translation of rainfall to continuous river discharge series. Because of the need to ultimately apply these models to ungauged sites, with model parameters derived from catchment properties, it is currently pragmatic to use parameter-sparse runoff model formulations, and two such models (the 'Probability Distributed Model' and the 'Time-Area Topographic Extension' model) with a track record of suitable performance are described. Chapter 4 on runoff model calibration details an automated procedure developed to establish suitable parameter sets, as the large number of catchments involved precludes manual calibration. Attention was paid to the performance of earlier-established calibrations in the light of more recent data from the winter 2000-2001 floods. Methods for establishing flood statistics from time series are briefly noted. Calibration objective functions included measures relating to frequency curves as well as to discharge time series.

Spatial generalisation of the method, namely the extrapolation from calibrated catchments to data-poor or ungauged sites, is discussed in detail in Chapter 5. Three methods, together with their variants, are developed in detail: these are, in brief, a univariate regression, a sequential regression and a site-similarity approach. Tests of these methods were undertaken on the sample catchments, withholding observations and treating them as if ungauged. Results are reported and serve as a basis for recommendations of the final chapter.

Chapter 6 develops theory and methods to derive uncertainty bands around spatially-generalised flood frequency curves arising from runoff model parameter uncertainty. Results are presented for the univariate regression and site-similarity generalisation methods. In Chapter 7 an overview comparison is made of issues in flood frequency estimation using continuous simulation and using the 1999 Flood Estimation Handbook. Chapter 8 indicates research directions of likely benefit to the continuous simulation method.

The final chapter brings together the experience of the research and its testing to suggest the best recommendations at the current state of knowledge. The procedures for gauged and ungauged sites are described, with the latter using site-similarity generalisation with the PDM and/or univariate regression with the TATE. Use to long recurrence intervals is demonstrated, and inclusion of climate variability commented upon. The chapter ends with a recommended dissemination strategy.

FAST TRACK TEXT BOXES

Blue text boxes offer an overview fast track through the main report and are to be found at the beginning of each chapter of the Technical Report.

ACKNOWLEDGEMENTS

The project team is grateful for research funding from the Defra – Environment Agency Flood and Coastal Erosion Risk Management R&D Programme (Broad Scale Modelling Theme) and from the Scottish Executive.

The Environment Agency, Scottish Environment Protection Agency and the National Water Archive at Wallingford are thanked for access to rainfall and flow data.

The core project team (see title page) also acknowledges input by current and former CEH Wallingford colleagues, particularly Chak Fai Fung, Sonja Folwell, Helen Davies, Sandra Smith, Rob Lamb and Simon Dadson.

1 AIMS AND BACKGROUND

Nick Reynard, Ann Calver

FAST TRACK TEXT BOX
CHAPTER 1 AIMS AND BACKGROUND

Project FD2106 addresses the issue of river flood frequency quantification. It seeks to develop a method for assessing flood frequency at gauged and ungauged locations using modelling of the continuous river discharge time series. Defra / Environment Agency project funding was augmented by Scottish Executive funding such that the work has been able to address this question across England, Wales and Scotland.

The forerunner pilot project, FD0404, was completed in 2001 and tested aspects of the potential of this approach. FD2106 was designed to substantially increase the sample of catchments on which the method is developed, to explore more than the one 'spatial generalisation' method (whereby the method is made available for ungauged site use) developed in the pilot, and to incorporate quantification of uncertainty arising from model parameter value choice. FD1604 had previously offered some ways forward for uncertainty estimation. In the practical application of the method rainfall data and/or modelling is required: project FD2105 is researching suitable rainfall modelling tools.

This approach of 'continuous simulation' offers potential advantages to flood frequency quantification over flood *event* methods and these advantages are detailed in the chapter text. The method is seen as a next-generation approach which will work alongside methods such as the Flood Estimation Handbook and which may, in time, be preferred in certain types of situations. The research emphasis of the project is on innovative development and testing of approaches and methods. Software is an issue for the future: this project works to 'research level code'.

The essence of the overall approach is to calibrate hydrological catchment runoff models for a representative group of sites with river flow and rainfall time series data. These sites are thus characterised by sets of model parameter values and also by values of 'catchment properties' (such as physiographic, geometric and material property indices) which must be available across the whole domain of concern for which ungauged estimates are required. For the ungauged site, therefore, the catchment properties are used to derive model parameter values which can then be used to generate flow series, and subsequent flood statistics and hydrographs, for the site of interest. This is the spatial generalisation of the method. The use of long rainfall time series allows temporal extension to longer than calibrated time series thereby encompassing larger, rarer floods. For this type of approach it is currently more appropriate to use parameter-sparse runoff models than complex multi-parameter approaches.

FAST TRACK TEXT BOX CONTINUED ON NEXT PAGE

FAST TRACK TEXT BOX CONTINUED FROM PREVIOUS PAGE

The chapter text expands on this overview of the approach and how components of it are systematically detailed through the report. It stresses an important contextual setting of the project in that, in developing new methods to time and budget, the exploration of variations of methodologies cannot be exhaustive: pragmatic choices have necessarily been made.

NEXT FAST TRACK BOX ON PAGE 7

1.1 Overall project aim

This report details the methodological developments and results from project FD2106 of the joint Defra / Environment Agency Flood Management R&D programme, with additional funding from the Scottish Executive. The project exploits advances in hydrological runoff modelling techniques for the advantages they offer design practice and planning through river flood frequency estimation.

It delivers methods for river flood frequency estimation by 'continuous flow simulation', that is, by catchment modelling of the complete flow time series (as opposed to flood events alone). The methods apply to the whole of Britain, for ungauged as well as gauged locations, and include explicit quantification of aspects of uncertainty.

1.2 Setting

Continuous flow simulation for flood estimation capitalises on advances in hydrological modelling, together with computing technology and the increasing availability of good quality fine-time-resolution data. It provides catchment modelling of the whole time series (including peaks, durations and hydrograph shapes), including effects of antecedent wetness conditions, river junctions and, if required, can incorporate changes in climate drivers. It is free of a need to relate recurrence intervals of rainfalls to recurrence intervals of floods.

Hydrological research, including the pilot project FD0404 (see below), indicates that hydrological models can be used, and hence flood statistics derived, where there are no flow data for calibration. In effect this allows transfer of information from data-rich sites to data-poor sites and this is achieved via relationships of model parameters to more widely-available spatial data ('catchment properties'). The forerunner pilot project was completed in 2001 (Calver *et al*. 2001). The approach is to calibrate runoff models for a set of catchments for which rainfall and river flow data are available. These sets of calibrated model parameters are related to properties of the river catchment which are widely available in the spatial sense. For ungauged sites, catchment properties offer, therefore, a way of deriving runoff model parameters, using the regression relationships developed from the calibrated catchment set. This allowed the subsequent modelling of discharge time series and the calculation of flood statistics at ungauged locations. This is a considerable achievement and is at the forefront of hydrological modelling research. Another project, FD1604 (Lamb *et al.* 2000, Lamb and Kay 2004), was innovative in establishing quantitative uncertainty levels for flood frequencies at ungauged sites resulting from model parameter uncertainty.

The current project is aligned with, although not dependent upon, project FD2105 *Improved methods for national spatial-temporal rainfall and evaporation modelling*. FD2105 is designed to deliver a national procedure for both single-site and areal rainfall modelling at daily and sub-daily timescales. There will also be an element of joint testing of these procedures by integrating the rainfall models with the runoff models developed in this project.

1.3 Specific project objectives

FD2106 objectives are to:-

- establish a substantial sample number of gauged catchments which serve as the basis for establishing predictive equations for ungauged sites;
- establish predictive equations for ungauged locations capitalising on exploration of numerical approaches to spatial generalisation;
- incorporate quantitative uncertainty levels resulting from model parameter uncertainty into frequency estimates, extending the prototype approaches of FD1604;
- describe the development and results of the methods and their mode of use.

1.4 Report structure

This report is divided into two volumes. The Technical Report is the main report describing and illustrating the development of the methods. A complete set of project results is presented as Appendices (Project Record). This report presents the final results of the project; it does not present a chronological account of project activity. This may be obtained by reference to the full set of project milestones listed in Table 1.1. The project results may be accessed at three levels: there are short, “fast track” text boxes at the beginning of the chapters that summarise the methods and results outlined in more detail in each of the chapters. There are then the chapters themselves, which are described below. Additionally there are the milestone reports listed in Table 1.1.

The Technical Report follows the logic of the project flow chart in Figure 1.1. There is a description of the background to the project and a brief history of recent research and methods for flood frequency estimation, particularly in the UK, in Chapter 1. Chapter 2 then describes the development of the data archives used, including the selection of the study catchments, the driving climate data, the flow data used for calibration and the data on catchment properties. Chapter 3 describes the two runoff models used throughout the project, with Chapter 4 detailing the calibration methods and results for each of these models. The methods assessed for spatial generalisation are presented in Chapter 5, together with performance results. While issues of uncertainty are present in both Chapters 4 and 5, it is in Chapter 6 where all these are drawn together and the combined (calibration and generalisation) uncertainty method presented.

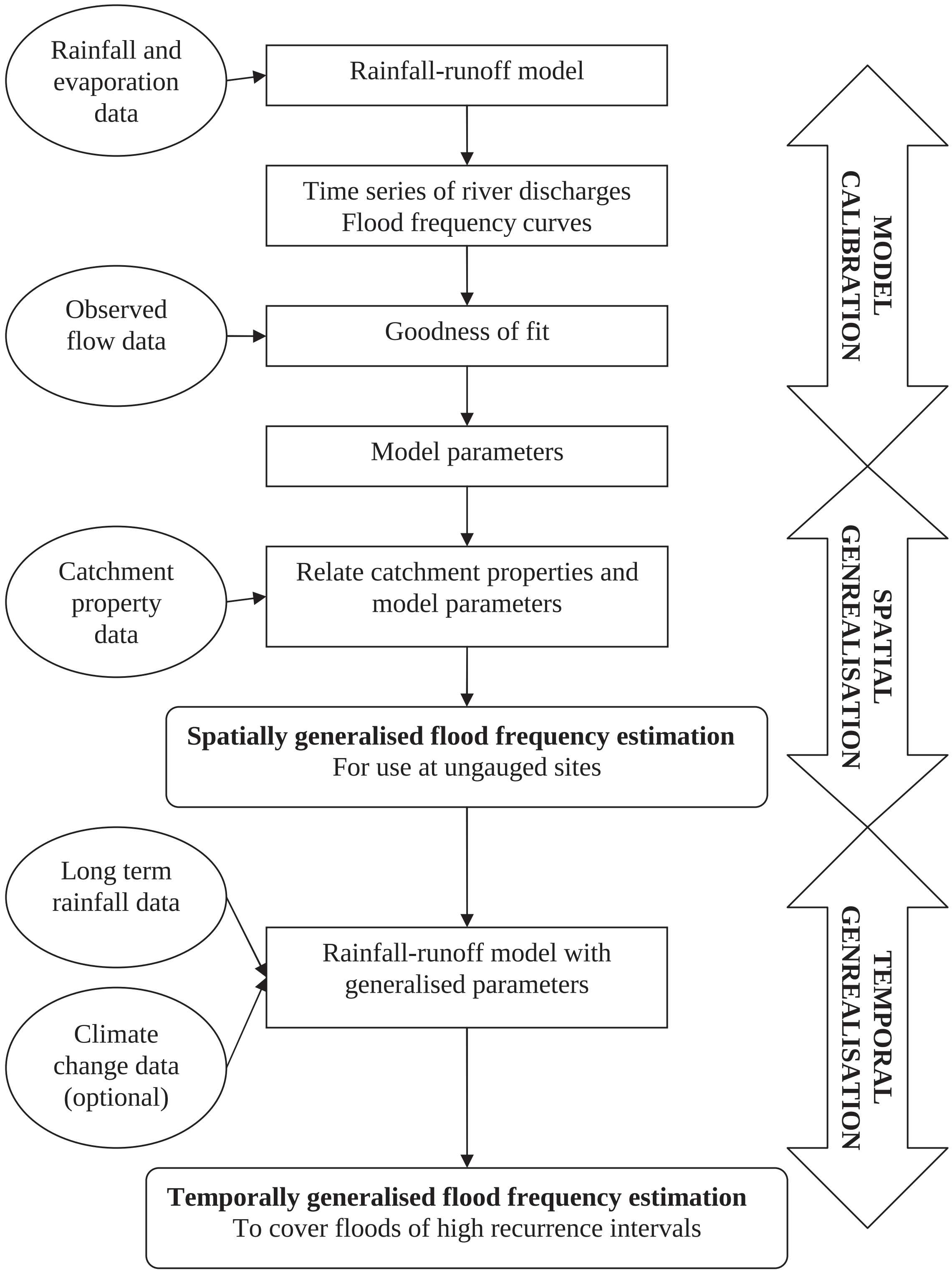


Figure 1.1 Structure of continuous simulation methodology for generalised flood frequency estimation

Table 1.1 List of milestones for project FD2106

Milestone	Title	Authors	Date
	Project inception report	Calver, A., Lamb, R., Kay, A..L., Crooks, S., Jones, D.A., Stewart, E.J.	January 2002
1	Assembly of daily data	Kay, A.L., Calver, A., Jones, D.A., Lamb, R.L., Scarrott, C.J., Stewart. E.J.	May 2002
2	Model testing in the light of extended data series	Crooks, S.M., Kay, A.L. & Calver, A.	November 2002
3	Including the benefits of daily data	Kay, A.L., Jones, D.A. & Calver, A.	May 2003
4	Performance of site-similarity approaches	Kjeldsen, T.R., Kay, A.L. & Jones, D.A.	August 2003
5	Preferred spatial generalisation procedures	Kay, A.L., Jones, D.A., Kjeldsen, T.R., Fung, C.F., Folwell, S., Calver, A., Reynard, N.S. & Crooks, S.M.	July 2004
6	Inclusion of uncertainty estimation methods	Jones, D.A., Kjeldsen, T.R., Kay, A.L. & Calver, A.	June 2004
	Flood frequency quantification for ungauged sites using continuous simulation: a UK approach	Calver, A., Kay, A.L., Jones, D.A., Kjeldsen, T., Reynard, N.S. & Crooks, S.	June 2004 *Proceedings of the International Environmental Modelling and Software Society Conference, Osnabruck.*
	Scottish catchments: data assembly, processing and runoff model calibrations	Crooks, S.M., Kay, A.L. & Calver, A.	August 2004

Chapter 7 describes the major issues as they effect both the continuous simulation and the Flood Estimation Handbook (FEH) methods for flood frequency estimation in the UK: these include issues such as ease of use, catchment properties and the treatment of uncertainty. In Chapter 8 there is the necessary description of possible future work in that the production of a new method based on a developing and relatively new area of science will always produce possible alternatives to meet the specified project objectives.

The final chapter draws the report to a conclusion. It is in Chapter 9 that the best approaches are recommended based on the results described in previous chapters. This chapter also presents an illustration of the method to long recurrence intervals and a worked example of the method.

2 DATA

Sue Crooks, Alison Kay

PREVIOUS FAST TRACK BOX ON PAGE 1

CHAPTER 2 DATA

This chapter describes the extensive data sets that have been collated for this project. This includes the selection of catchments for which the runoff models have been calibrated and the data needed for this, as well as information on the catchment properties used in the spatial generalisation procedures.

A major concern when selecting catchments was the availability of good quality, continuous rainfall and river flow data of at least eight years duration. This resulted in a set of 119 catchments, 46 of which have the required data at an hourly time-step, with remaining 73 having daily data.

A total of 24 catchment properties have been used in the spatial generalisation procedures. The selection was based on previous experience, the independence of descriptors and their ease of calculation and availability. They cover physiographic characteristics of the catchment and river network, and soil, land use and geological properties.

NEXT FAST TRACK BOX ON PAGE 21

2.1 Introduction

The majority of the hourly catchments used in this project were calibrated and tested in the forerunner pilot project FD0404, mostly for the period 1985 to 1995. The observed rainfall and runoff data collected for this period constituted a valuable data set of some 400 station-years of continuous hourly data, infilled and quality-checked (Lamb and Gannon 1996). The data period was in general defined by the beginning of suitable recording and by the date of the data-gathering phase of the pilot project. However, the period was not especially rich in high magnitude flow events.

The current project provided an opportunity to

- update the hourly database to the end of 2001, which extends the flood range in many catchments,
- considerably increase the number of catchments used in the development of the continuous simulation method, by including catchments with only daily data, and
- increase the number of hourly catchments in Scotland.

The extended data period for the hourly catchments included the notable flood events which occurred in England and Wales during the winter of 2000-01. The extended data series were used to test the robustness of the calibrated parameter sets from FD0404 to simulate flood events beyond the flow range for which they were determined, which is important for the credibility of the continuous simulation method for flood frequency estimation. Details of the data testing are given in Section 4.5.

The inclusion of catchments with daily data was considered important as

- Fewer sites are classed as 'ungauged' when daily rather than hourly flow data are required.
- A larger set of gauged catchments should be more representative of the UK, with regard to the ranges of catchment properties.
- Using a larger set of gauged catchments may allow the production of more reliable equations relating model parameters to catchment properties.
- Daily records are generally significantly longer than hourly records, meaning that flood frequency curves can more readily be obtained to higher return periods.

The original, FD0404, data set was relatively sparse in Scotland, particularly western Scotland. Aligned funding from the Scottish Executive allowed the addition of further hourly catchments in Scotland, as well as the updating of some Scottish catchments used in FD0404 (Crooks *et al.* 2004). The advantage of this extended coverage of Scottish catchments is not simply to improve the working of the method on ungauged catchments in Scotland, but the increased catchment property coverage provided by these catchments can potentially improve the performance of the spatial generalisation procedures in general.

2.2 Catchment selection

The main requirement for selection of a catchment was that it had good quality continuous flow and rainfall data for a minimum of eight to ten years. This period is necessary to ensure that the time series of flows provides enough flood peaks to enable the derivation of a flood frequency distribution up to magnitudes and return periods of interest. Note that, although longer flow time series are preferred, this must be balanced by a consideration of the increased possibility of non-stationarity of flows, whether due to human impacts on the catchment, climate variability, gauging station alterations or any other factor. The statistical techniques, which enable the production of flood frequency distributions from flow time series, make the assumption of stationarity of flows.

It is also important that the calibration flow series is long enough to cover a wide range of catchment conditions, as some of the model parameters may only be significant in determining simulated flow series under certain conditions. For instance, the parameters of a 'fast flow' routing function may only affect modelled flows during major flood events.

Continuity of record is also important for the continuous simulation approach. Continuity of input rainfall data is vital: the models cannot be run over periods without corresponding rainfall data. If it were necessary to run the model for a catchment over two separate time periods (for example to get a sufficient length of record), then a separate run-in time would have to be allowed at the beginning of each period, to allow the model to adjust appropriately.

A small number of reasonably isolated missing flow values can be tolerated, as the matching flows in the simulated series are simply set as missing too, before the observed and simulated flows are compared. However, large numbers of missing flows are best avoided.

The 39 catchments with hourly data were selected in FD0404 on data availability and quality. Seven of these catchments were in Scotland; a further seven were added in this project to improve the geographical distribution across the country. Details of the hourly catchments are given in Table 2.1.

An additional factor in selection of daily catchments was catchment area, as small catchments have a fast response time that is not suitable to model at a daily time step. Soil type and hydrogeology of the catchment are also important in determining response times so a threshold of 50 km^2 was chosen as a simple but reasonable compromise for selecting suitability at the daily time step. 73 catchments across Great Britain were selected to give a relatively even geographical distribution. Details of the daily catchments are given in Table 2.2 with the locations of all catchments shown in Figure 2.1. All hourly data sets are to the end of 2001; end dates for the daily data sets are indicated in Table 2.2.

2.3 Rainfall

The National Water Archive (NWA) at CEH Wallingford holds daily rainfall data for around 13,000 rain gauges across the United Kingdom, obtained via the UK Met Office. These data have been quality checked by the Met Office and are considered to be reliable.

The lumped or semi-distributed runoff models used in the project require catchment-averaged rainfall as a forcing input. Data from a number of individual rain gauges are thus combined, using the triangle method (Jones 1983), to produce the time series of Catchment Average Daily Rainfall (CADR) for a catchment. All rain gauges within a certain area encompassing the catchment are included. The method is designed to eliminate, as much as possible, the effects of an irregular density or distribution of rain gauges. It also takes some account of topography, through scaling using Standard Average Annual Rainfall (SAAR) data. On the basis of acceptable pilot project performance, distinction was not explicitly made between precipitation occurring as rain or as snow (cf. Chapter 8 comment).

Data from recording raingauges were obtained directly from Regional Offices of the Environment Agency (EA) and the Scottish Environmental Protection Agency (SEPA). The raw data were received in a number of formats, the most common being 15-minute, hourly and event. With event data, the time of every tip is recorded to the nearest second, where a tip is usually 0.2 mm but in some cases 0.5 mm, with no record made for zero rainfall. With hourly and 15-minute data, the rainfall in every time period is recorded, based on the number of tips, which includes many zero entries; though some records only give the time when the rainfall is non-zero. The data were converted to a

Table 2.1 Details of the hourly catchments

Catchment Number	River	Location	Data start	Catchment Area (km^2)	$SAAR_{61-90}$ (mm)	BFI
03003	Oykel	Easter Turnaig	01-Jan-82	330.7	1895	0.23
07001	Findhorn	Shenachie	01-Jan-85	415.6	1219	0.36
07004	Nairn	Firhall	01-Feb-85	313.0	940	0.45
10003	Ythan	Ellon	22-Feb-89	523.0	826	0.73
12007	Dee	Mar Lodge	13-Oct-89	289.0	1335	0.45
21013	Gala Water	Galashiels	01-Jan-86	207.0	930	0.52
21017	Ettrick Water	Brockhoperig	01-Jan-86	37.5	1733	0.34
22006	Blyth	Hartford Bridge	01-Jan-85	269.4	696	0.35
23011	Kielder Burn	Kielder	01-Jan-85	58.8	1199	0.34
24005	Browney	Burn Hall	01-Jan-82	178.5	743	0.51
25006	Greta	Rutherford Bridge	01-Jan-85	86.1	1128	0.22
27051	Crimple	Burn Bridge	01-Jan-85	8.1	856	0.31
28008	Dove	Rocester Weir	01-Jan-85	399.0	1021	0.62
28039	Rea	Calthorpe Park	01-Jan-85	74.0	781	0.47
28046	Dove	Izaak Walton	01-Jan-85	83.0	1096	0.79
29001	Waithe Beck	Brigsley	01-Jan-85	108.3	690	0.85
30004	Lymn	Partney Mill	01-Jan-85	61.6	685	0.65
36008	Stour	Westmill	01-Jan-85	224.5	589	0.43
36010	Bumpstead Brook	Broad Green	01-Jan-84	28.3	589	0.23
38007	Canons Brook	Elizabeth Way	01-Jan-85	21.4	601	0.41
38020	Cobbins Brook	Sewardstone Road	01-Nov-87	38.4	616	0.26
39007	Blackwater	Swallowfield	01-Jan-85	354.8	707	0.67
39017	Ray	Grendon Underwood	01-Jan-87	18.8	622	0.17
39037	Kennet	Marlborough	01-Jan-85	142.0	772	0.94
39073	Churn	Cirencester	01-Feb-87	84.0	854	0.89
40005	Beult	Stile Bridge	01-Jan-85	277.1	690	0.24
42008	Cheriton Stream	Sewards Bridge	01-Jan-85	75.1	889	0.97
45003	Culm	Wood Mill	02-Apr-85	226.1	971	0.53
54027	Frome	Ebley Mill	01-Jan-85	198.0	827	0.87
54034	Dowles Brook	Oak Cottage, Dowles	01-Jan-85	40.8	715	0.40
54090	Tanllwyth	Tanllwyth Flume	01-Jan-74	0.9	2425	0.30
55008	Wye	Cefn Brwyn	01-Jan-69	10.6	2453	0.31
55013	Arrow	Titley Mill	01-Jan-85	126.4	962	0.55
57005	Taff	Pontypridd	01-Jan-85	454.8	1830	0.47
57006	Rhondda	Trehafod	01-Jan-85	100.5	2184	0.41
58006	Mellte	Pontneddfechan	01-Jan-85	65.8	1979	0.38
60002	Cothi	Felin Mynachdy	01-Jan-85	297.8	1551	0.44
60003	Taf	Clog-y-Fran	29-Mar-85	217.3	1420	0.56
74001	Duddon	Duddon Hall	01-Jan-85	85.7	2265	0.29
79005	Cluden Water	Fiddlers Ford	01-Jan-88	238.0	1423	0.38
81006	Minnoch Water	Minnoch Bridge	01-Jan-88	141.0	1993	0.28
84030	White Cart Water	Overlee	01-Jan-92	111.8	1367	0.32
86001	Little Eachaig	Dalinlongart	01-Jan-92	30.8	2341	0.23
90003	Nevis	Claggan	01-Jan-93	76.8	2913	0.26
93001	Carron	New Kelso	01-Jan-92	137.8	2615	0.26
96001	Halladale	Halladale	01-Jan-85	204.6	1102	0.26

Table 2.2 Details of the daily catchments

Catchment Number	River	Location	Data start	Catchment Area (km^2)	$SAAR_{61-90}$ (mm)	BFI
02001	Helmsdale	Kilphedir	01-Jan-75	551.4	1117	0.48
04005	Meig	Glenmeannie	01-Jan-86	120.5	2145	0.26
06008	Enrick	Mill of Tore	01-Dec-79[2]	105.9	1294	0.32
07002	Findhorn	Forres	01-Jan-61[2]	781.9	1064	0.41
08004	Avon	Delnashaugh	01-Jan-61[1]	542.8	1111	0.56
10002	Ugie	Inverugie	01-Feb-71	325.0	812	0.64
13001	Bervie	Inverbervie	01-Jan-79[2]	123.0	890	0.56
13005	Lunan Water	Kirkton Mill	01-Feb-81	124.0	771	0.52
14001	Eden	Kemback	01-Jan-67[2]	307.4	799	0.62
16003	Ruchill Water	Cultybraggan	01-Jan-70[2]	99.5	1889	0.30
17005	Avon	Polmonthill	01-Jan-71[2]	195.3	989	0.41
19011	North Esk	Dalkeith Palace	01-Jan-63[2]	137.0	907	0.52
20001	Tyne	East Linton	01-Jan-61[2]	307.0	713	0.52
21023	Leet Water	Coldstream	01-Jan-70[2]	113.0	671	0.35
22001	Coquet	Morwick	31-Jan-63[2]	569.8	850	0.45
27007	Ure	Westwick Lock	01-Jan-61[2]	914.6	1118	0.39
27021	Don	Doncaster	01-Jan-61	1256.2	799	0.56
27043	Wharfe	Addingham	01-Jan-73[2]	427.0	1383	0.33
27049	Rye	Ness	01-Jan-74[2]	238.7	839	0.68
28015	Idle	Mattersey	01-Nov-82	529.0	650	0.79
28066	Cole	Coleshill	01-Jan-73[2]	130.0	722	0.44
30017	Witham	Colsterworth	01-Jan-78[1]	51.3	642	0.50
31002	Glen	Kates and King St Brs	01-Jan-61[3]	341.9	608	0.59
32003	Harpers Brook	Old Mill Bridge	01-Jan-61	74.3	623	0.49
33012	Kym	Meagre Farm	01-Jan-61	137.5	585	0.26
33019	Thet	Melford Bridge	01-Jan-62[1]	316.0	620	0.78
33029	Stringside	Whitebridge	22-Jul-65	98.8	629	0.85
34003	Bure	Ingworth	01-Jan-61	164.7	669	0.83
34006	Waveney	Needham Mill	01-Jan-63[2]	370.0	594	0.47
36005	Brett	Hadleigh	01-Oct-62[1]	156.0	580	0.46
37001	Roding	Redbridge	01-Jan-61	303.3	606	0.39
37031	Crouch	Wickford	01-Jan-76[1]	71.8	572	0.30
38003	Mimram	Panshanger Park	01-Jan-61	133.9	656	0.94
39069	Mole	Kinnersley Manor	01-Jan-61	142.0	795	0.39
39105	Thame	Wheatley	18-May-89	533.8	644	0.63
40011	Great Stour	Horton	01-Oct-64	345.0	747	0.70
43005	Avon	Amesbury	01-Feb-65[3]	323.7	745	0.91
43007	Stour	Throop	01-Jan-73	1073.0	861	0.67
44002	Piddle	Baggs Mill	01-Oct-63	183.1	943	0.89
45005	Otter	Dotton	30-Sep-62	202.5	976	0.53
47007	Yealm	Puslinch	01-Oct-63	54.9	1410	0.56
47008	Thrushel	Tinhay	01-Jan-69[2]	112.7	1143	0.39
48003	Fal	Tregony	08-Jun-78	87.0	1210	0.68
50002	Torridge	Torrington	28-Feb-62	663.0	1186	0.39
50006	Mole	Woodleigh	11-Jan-65	327.5	1306	0.47
52010	Brue	Lovington	01-Jan-64[2]	135.2	867	0.47
53009	Wellow Brook	Wellow	01-Jan-66[2]	72.6	998	0.62
54008	Teme	Tenbury	01-Jan-61	1134.4	841	0.57
54018	Rea Brook	Hookagate	01-Jan-62[2]	178.0	756	0.51
54025	Dulas	Rhos-y-pentref	01-Jan-69[2]	52.7	1269	0.37
55029	Monnow	Grosmont	01-Jan-61[2]	354.0	955	0.59
58005	Ogmore	Brynmenyn	01-Jan-70[2]	74.3	1976	0.49
61001	Western Cleddau	Prendergast Mill	01-Oct-65	197.6	1275	0.65

continued…….

Catchment Number	River	Location	Data start	Catchment Area (km^2)	$SAAR_{61\text{-}90}$ (mm)	BFI
64001	Dyfi	Dyfi Bridge	01-Sep-75	471.3	1834	0.38
65006	Seiont	Peblig Mill	01-Aug-76	74.4	2278	0.40
66011	Conwy	Cwm Llanerch	01-Jan-64[2]	344.5	2055	0.28
67009	Alyn	Rhydymwyn	01-Jan-67[2]	77.8	969	0.40
68001	Weaver	Ashbrook	01-Jan-61	622.0	731	0.53
68005	Weaver	Audlem	01-Jan-61[2]	207.0	719	0.50
69040	Irwell	Stubbins	30-Mar-76[3]	105.0	~1405	0.44
73005	Kent	Sedgwick	01-Jan-68[2]	209.0	1732	0.46
75017	Ellen	Bullgill	01-Jan-76[3]	96.0	1110	0.49
76014	Eden	Kirkby Stephen	01-Jan-71[2]	69.4	1483	0.24
78003	Annan	Brydekirk	01-Jan-67[2]	925.0	1351	0.44
79002	Nith	Friars Carse	01-Jan-61[2]	799.0	1460	0.39
79003	Nith	Hall Bridge	01-Jan-61[2]	155.0	1505	0.27
81002	Cree	Newton Stewart	01-Jan-63[2]	368.0	1760	0.27
83005	Irvine	Shewalton	01-Jan-72[2]	380.7	1228	0.26
84012	White Cart Water	Hawkhead	01-Jan-61[2]	234.9	1314	0.35
85003	Falloch	Glen Falloch	01-Oct-70[3]	80.3	2842	0.17
94001	Ewe	Poolewe	01-Oct-70	441.1	2273	0.65
95001	Inver	Little Assynt	01-Aug-77[3]	137.5	2211	0.64
97002	Thurso	Halkirk	01-Jan-72[3]	412.8	1057	0.46

End date: [1] 31-Dec-96, [2] 31-Dec-97, [3] 31-Dec-00, all others 31-Dec-01.

standard hourly format where the rain for any given hour is that falling in the previous 60 minutes.

The hourly rainfall data were quality checked, as described in Lamb and Gannon (1996). Briefly, the data from each recording rain gauge were assessed by comparing the sum of the hourly values over each rain day (9am to 9am) with the total rainfall recorded on that day by a gauge with data in the National Water Archive (NWA). All daily gauges within 10 km of the recording gauge are located, and for each day the nearest working daily gauge is used for the comparison. A quality code is then assigned to the day (the same code for each hour of the day) according to the percentage difference between the two daily totals. If the difference is less than 40% the recording raingauge data are considered acceptable for calculating Catchment Average Hourly Rainfall (CAHR).

If good quality recording rain gauge data exist for a day, then this is used to distribute the CADR. Where data from only one recording gauge are available, this profile is used directly to distribute the CADR. Where data from more than one gauge are available, weights are allocated to each according to the triangle method, and the profile of the weighted sum of the hourly rain gauge data is used to distribute the CADR. For days where no good-quality recording rain gauge data are available, the CADR is distributed according to an average profile of rainfall for the catchment. The average variability method (Pilgrim *et al.* 1969) is used to define three average profiles for each catchment, from the good-quality recording data available for the catchment. The three profiles correspond to days with less than 10 mm of rain; between 10 and 20 mm of rain; and above 20 mm of rain. The processed data are stored on ORACLE tables, accompanied by quality codes.

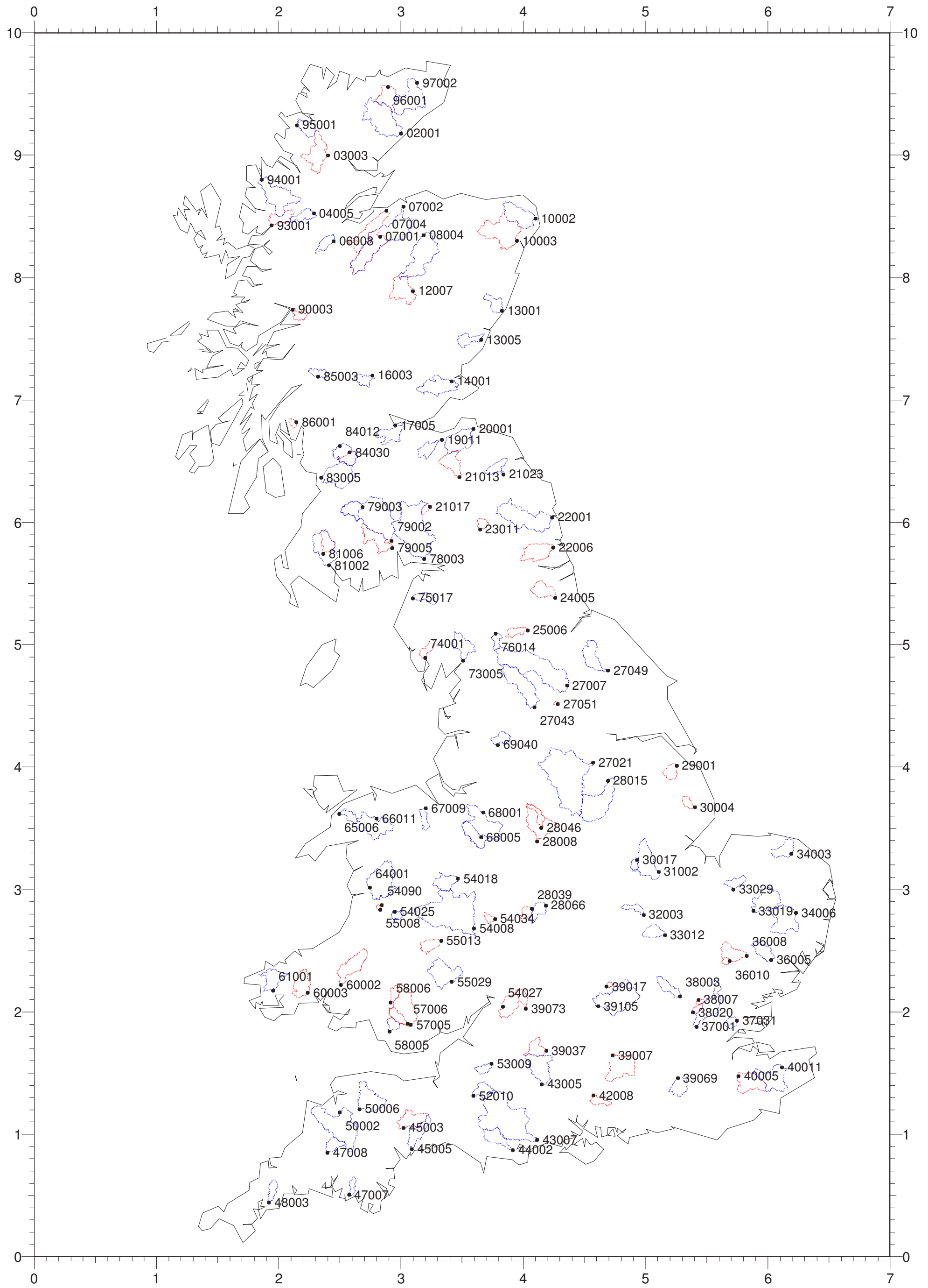


Figure 2.1 Map showing catchment boundaries and outlets. Catchments with hourly data are indicated by red boundaries while those with daily data are indicated by blue boundaries.

2.4 Potential evaporation

As well as rainfall data, the runoff models also require potential evapotranspiration (PE) data. Daily PE data are available from the UK Met Office for synoptic sites. However, the sparse coverage of these sites across the UK was a problem in designing a consistent approach to the provision of spatially averaged PE data for all the project catchments. Consequently, MORECS (Meteorological Office Rainfall and Evaporation Calculation System, Thompson *et al*. 1982) monthly data, which are readily available as average values for 190 40 km $\times$ 40 km grid squares across Great Britain, have been used for PE. The monthly PE data for a catchment were determined by weighting the PE data for each grid square by the proportion of the catchment in that square, and then summating over the squares. The monthly PE data were downscaled to a daily time-step by dividing equally over the number of days in the month; these were further downscaled to an hourly time-step for the hourly catchments, by dividing equally over 24 hours.

A recent paper (Fowler 2002) looked at the effect of using different forms of PE data as input to a long-term soil water balance model run at a daily time-step. It concluded that very little advantage was gained by using input data based on daily observations, compared to that based on readily available monthly observations. Similar conclusions were reached by Andersson and Harding (1991). As the effect of PE in the runoff models is very conservative, and is mainly used to maintain the water balance, there is justification for using monthly data in the runoff models.

2.5 River flows

There is a central archive of mean daily flow data, held in the National River Flow Archive (NRFA), a component part of the NWA. The archive holds data for over 1600 gauging stations across England, Wales, Scotland and Northern Ireland, obtained through a number of measuring authorities. This reflects over 39,000 station-years of data, with an average record length of over 23 years. A quality checking procedure is applied to these data before they are archived. A summary of the data held for each gauge, along with a variety of other information, can be found in the series *Hydrological Data UK: Hydrometric Register and Statistics*, published annually by CEH Wallingford (formerly the Institute of Hydrology) and the British Geological Survey.

Sub-daily flow data were obtained from Regional Offices of the EA and SEPA as instantaneous flow rates at 15-minute, 30-minute or hourly data intervals. The hourly flow can be defined as either the mean rate over the preceding hour or the instantaneous rate observed every hour. The latter definition was used, as this was consistent with that used in FD0404, and the 15-minute data only provides four sampling points from which to calculate the mean, which does not necessarily provide a good definition of mean hourly flow. Instantaneous flows for on-the-hour readings were selected to give point values at hourly intervals.

2.6 Catchment properties

A total of 24 catchment properties, or descriptors, have been considered for use in the spatial generalisation procedures, covering aspects of topography, soil and geology, rainfall, drainage networks, lakes and reservoirs and land cover. A description of each catchment property is given in Table 2.3 together with the source from which they were obtained. These catchment properties were selected from the list of 58 properties compiled for FD0404, many of which were calculated specifically for that project. Selection was based on experiences of FD0404, independence of properties, ease of calculation and availability to the user, whilst providing measures of all aspects of catchments which might affect runoff response and generation of floods.

2.6.1 Transformation and distribution of catchment properties

To try to improve the potential to gain information from catchment properties, some of the catchment properties are transformed before use in generalisation. The transformations are generally chosen to make the distribution of the quantity less skewed, and are given in Table 2.4. Plots, in the form of histograms, showing the distributions of the transformed catchment properties are given in Figure 2.2, which distinguishes between the distributions for the hourly catchments and those for hourly and daily catchments combined.

One potential benefit of including catchments with daily data was to expand the range of catchment properties and combinations of properties to ensure that generalisation of model parameters from catchment properties was not based on too small a sample. The histograms show an extended range for some of the properties, particularly catchment area, but other ranges have not significantly changed. The catchment properties which derive from the Flood Estimation Handbook (see Table 2.3) show similar distributions to those derived from FEH’s set of over 900 catchments (Institute of Hydrology 1999, Volume 5), suggesting that the set used should be reasonably representative of UK catchments as a whole, at least in terms of those properties.

The majority of the transformed catchment properties approximate to a normal distribution, but some are still highly skewed, for example the distributions for FARL and URBEXT. Care may be needed with such distributions, where in many catchments the value of the transformed property is zero, to ensure that the properties are adequately represented in generalisation procedures.

Table 2.3. Definitions of catchment properties used in the spatial generalisation methods

CP Name	Range, units	Source	Notes
AREA	$[0,\infty]$ km^2	FEH	DTM-derived
ALTBAR	$[0,\infty]$ m	FEH	Mean altitude
BFIHOST	[0,1] -	FEH	Base flow index, calculated from weighted average of HOST classes over the catchment
DPLBAR	$[0,\infty]$ km	FEH	Mean drainage path length
DPLCV	$[0,\infty]$ -	FEH	CV drainage path length
DPSBAR	$[0,\infty]$ m/km	FEH	Mean slope of DTM drainage paths to site
FARL	[0,1]	FEH	Index of flood attenuation due to reservoirs and lakes
PROPWET	[0,1] -	FEH	Proportion of time catchment wet (SMD<6mm)
SAAR	$[0,\infty]$ mm	FEH	Standard average annual rainfall, 1961-90
SPRHOST	[0,100] -	FEH	Standard percentage runoff derived from weighted average of HOST classes over catchment
URBEXT	[0,1] -	FEH	Extent of urban/suburban land cover ($URBEXT=URB_{FRAC}+0.5\times SUBURB_{FRAC}$)
HOSTGMIN	[0, 100] %	HOST	% of catchment area covered by HOST 1-10,13,14 (mineral soils with underlying groundwater)
HOSTPEAT	[0, 100] %	HOST	% of catchment area covered by HOST 11,12,15 ('peat soils with groundwater')
HOSTNG	[0, 100] %	HOST	% of catchment area covered by HOST classes 16-29 (essentially 'non-groundwater')
HOSTP	[0,1]	HOST	Index of porosity as a weighted average of values inferred from HOST classes.
FIELDC	[0,100] %	SEISMIC/ HOST	Volumetric soil water content at 5 kPa, as weighted average of values inferred from HOST classes.
RESIDM	[0,100] %	SEISMIC/ HOST	Residual soil moisture, as weighted average of values inferred from HOST classes.
PORO	[0,100] %	SEISMIC/ HOST	Total soil porosity, as weighted average of values inferred from HOST classes.
HYDC	$[0,\infty]$ cm/d	SEISMIC/ HOST	Saturated soil hydraulic conductivity, as weighted average of values inferred from HOST classes
LANDA	[0, 100] %	ITE	% of catchment area covered by grassland based on ITE land cover data (classes 5-8,19,23)
LANDB	[0, 100] %	ITE	% of catchment area covered by upland based on ITE land cover data (classes 9-13,17,24,25)
LANDC	[0, 100] %	ITE	% of catchment area covered by trees based on ITE land cover data (classes 14-16)
LANDD	[0, 100] %	ITE	% of catchment area covered by 'arable' based on ITE land cover data (class 18)
DRAIN2	$[0,\infty]$ km/km^2	DTM	Drainage density (total length of river (km) divided by the catchment area (km^2))

Notes on sources:

FEH Properties appearing on the FEH CD-ROM or based on FEH catchment properties
HOST Properties derived from the HOST soil classification system (Boorman *et al*. 1995)
SEISMIC/HOST Properties derived from the SEISMIC soils characteristics database for each HOST class
ITE Properties derived from the ITE 1990 land cover classification (Fuller, 1993)
DTM Properties derived from the CEH-Wallingford 'Integrated Hydrological Digital Terrain Model' (IHDTM) (Morris and Flavin, 1990)

Table 2.4 Transformations of the catchment properties

Catchment property	Transformation	Catchment property	Transformation
AREA	Sqrt(CP)	HOSTPEAT	Sqrt(CP/100)
ALTBAR	Sqrt(CP)	HOSTNG	-
BFIHOST	-	HOSTP	Sqrt(CP)
DPLBAR	-	FIELDC	-
DPLCV	-	RESIDM	-
DPSBAR	Sqrt(CP)	PORO	-
FARL	Sqrt(1-CP)	HYDC	-
PROPWET	-	LANDA	Sqrt(CP/100)
SAAR	Sqrt(CP)	LANDB	Sqrt(CP/100)
SPRHOST	-	LANDC	Sqrt(CP/100)
URBEXT	Sqrt(CP)	LANDD	Sqrt(CP/100)
HOSTGMIN	Sqrt(CP/100)	DRAIN2	-

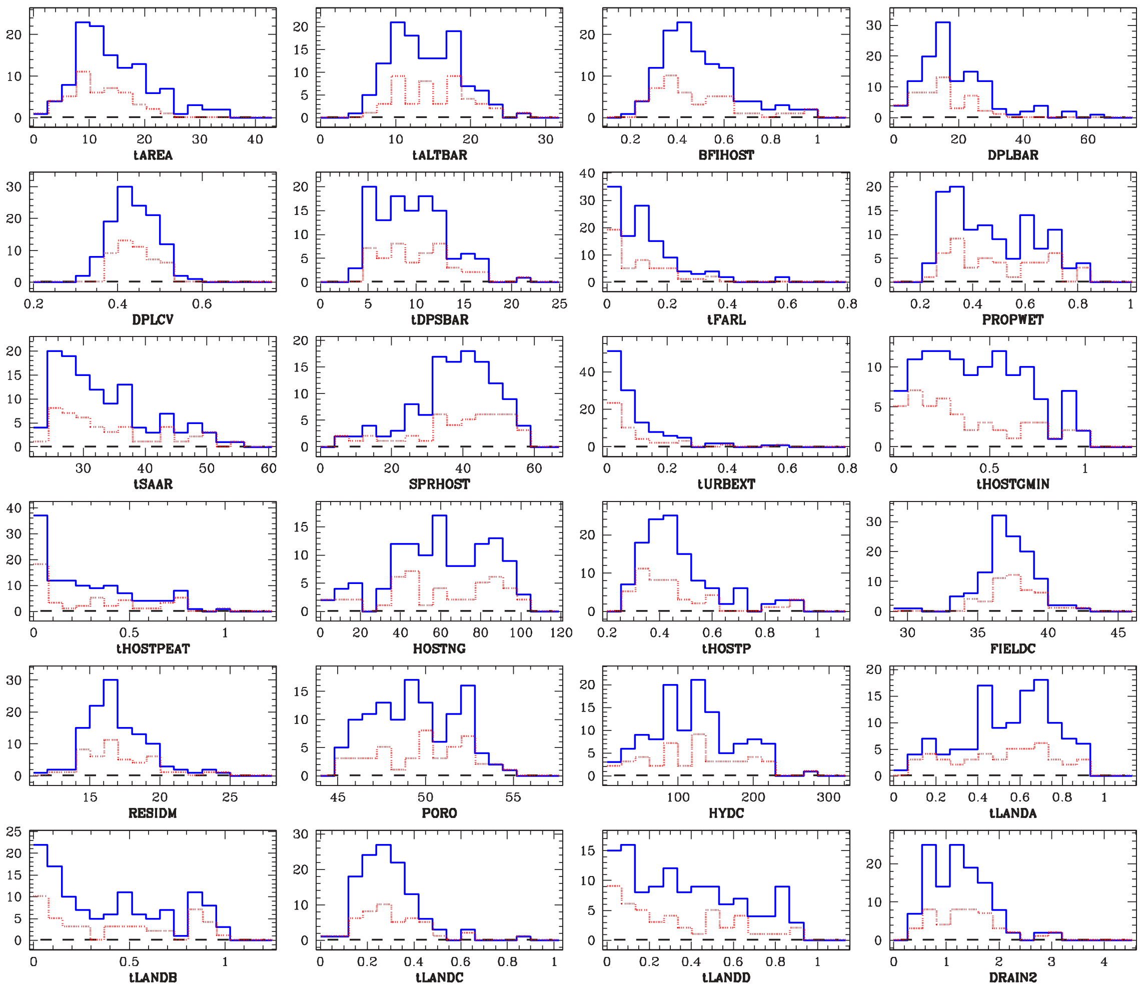


Figure 2.2 Distributions of the catchment properties. Those shown in red are for just the hourly catchments, while those in blue include both hourly and daily catchments. Catchment properties prefixed by a 't' are those that have been transformed, in the way listed in Table 2.4.

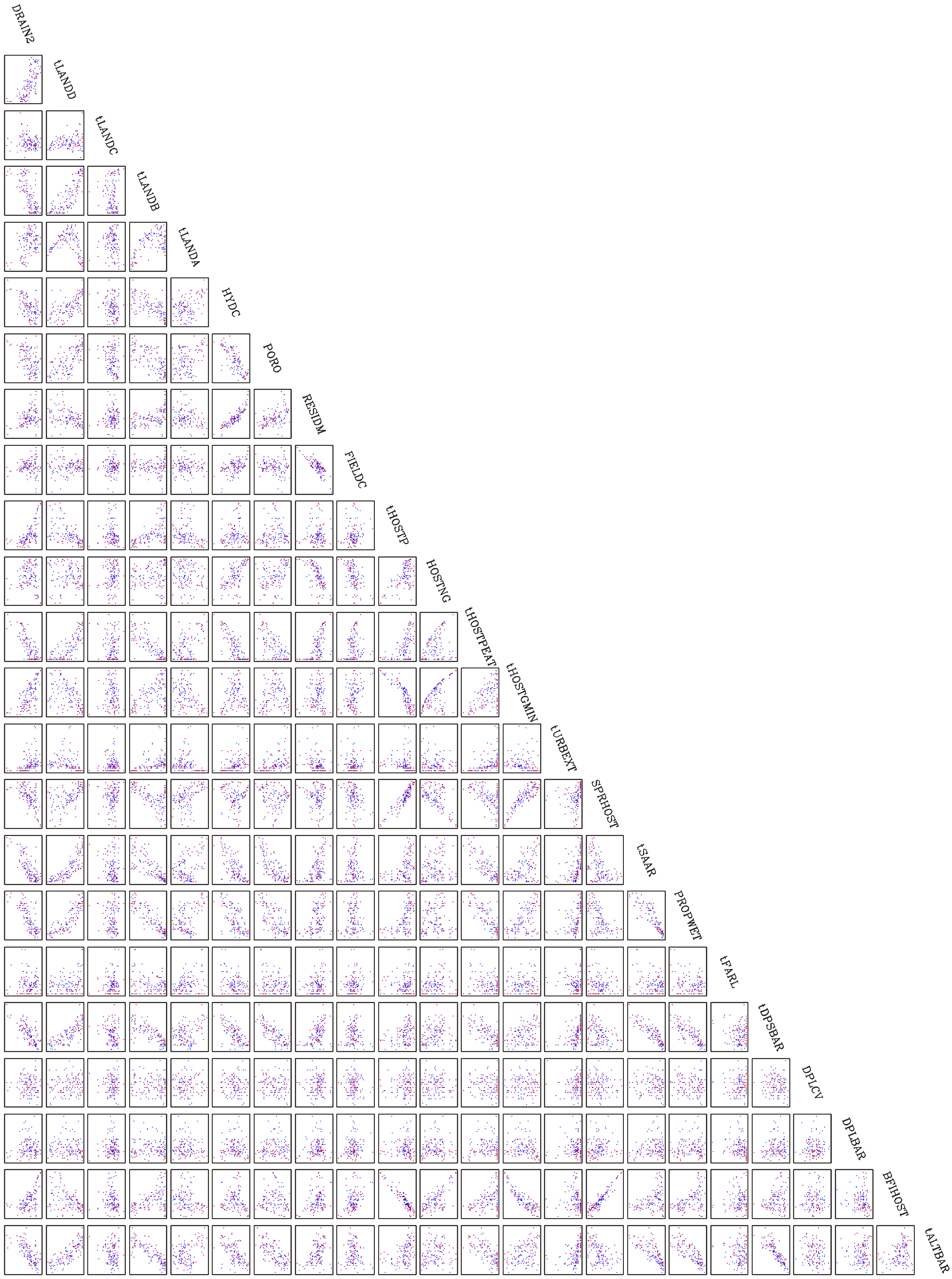


Figure 2.3 Coverage of the catchment property space. Red dots are for hourly catchments while blue dots are for daily catchments. Catchment properties prefixed by a 't' are those that have been transformed, in the way listed in Table 2.4.

2.6.2 Correlations between catchment properties

A number of groups of catchment properties under consideration are highly correlated, before and/or after the transformations given in Table 2.4 are taken into account. The correlated groups are as follows, where those catchment properties in brackets are slightly less correlated to the rest of their group than the others:

- AREA, DPLBAR
- BFIHOST, HOSTGMIN, SPRHOST, HOSTP
- SAAR, DPSBAR, PROPWET, ALTBAR, LANDB, LANDD, (HOSTPEAT)
- RESIDM, HYDC, (FIELDC)

These correlations can be seen in Figure 2.3, which shows the coverage of the catchment property space in terms of hourly and daily catchments. The improvement in coverage of the whole property space provided by the inclusion of the daily catchments can also be seen in Figure 2.3.

It is generally not desirable to have more than one catchment property from a correlated group in a regression equation, or used in defining the pooling group for the site-similarity approach (see Chapter 5), for stability reasons. Therefore care must be taken to limit this, whilst not wishing to make an *ad hoc* choice between properties.

3 RUNOFF MODELS

Ann Calver, Alison Kay

PREVIOUS FAST TRACK BOX ON PAGE 7

CHAPTER 3 RUNOFF MODELS

These models convert input precipitation time series to river discharge time series, encapsulating numerically the processes of catchment hydrology. The need to derive the parameter values for these runoff models from catchment properties for ungauged sites means that it is more appropriate that simpler runoff models are used, rather than those which, although physically or statistically more descriptive, employ more parameters. The models, albeit implicitly, cover both surface and subsurface (throughflow and groundwater) conditions.

Two conceptual hydrological models, both with a concession to catchment spatial configuration (rather than full three-dimensional description) were used. These are the Probability Distributed Model (PDM) and the Time-Area Topographic Extension (TATE) model. Both have evidence of suitable performance in this context: the use of two models is a pragmatic decision between the risk associated with use of a single model and the time overhead of using a range of models. Different formulations of both models exist: in practice a five-parameter version of the PDM was used and a three-parameter version of the TATE. The structure of both models and the roles their parameters serve are detailed in the chapter text.

NEXT FAST TRACK BOX ON PAGE 25

3.1 Introduction

Chapter 1 outlined the type of modelling approach to be adopted in deriving the continuous time series of river flows. The requirement of the runoff modelling is the conversion of rainfall time series to river discharges, representing catchment hydrological processes and encapsulating in particular those responses which are important in the generation of river floods. The model(s) should also show good performance across the range of catchment sizes and conditions - geological, physiographic, climatological and degree of anthropogenic influence - for which the frequency estimation method is sought, for both gauged and ungauged sites. This includes surface and both shallow and deep subsurface flows.

There is, of course, a wide variety of hydrological modelling approaches available with, arguably, more *types* of formulation than, say, for atmospheric or river hydraulic modelling. Highly physically-descriptive formulations may appear to hold certain attractions in describing the fast-changing and varied flood response but, in the present context, it is an important consideration that model parameters need to be established for the ungauged sites, namely those with no gauged river flows and in which other data may be limited. At this stage of hydrological research in the challenging and valuable area of modelling of data-sparse areas, a practical system is more likely to be attained by the restriction of the number of parameters which are established (at the ungauged site) from the widely-available catchment property data. It is the *spatial generalisation* – as opposed to model use on specific gauged catchments – which currently indicates this approach. Attention is therefore directed towards the parameter-sparse approaches where experience of hydrological behaviour is distilled into a low number of powerful parameters with relatively independent roles. Usually this means that the spatial dimensions of catchments are not fully preserved, and either a lumped or semi-distributed scheme is used. It is also important that models used have a track record of suitable performance.

Whilst one would choose to test a range of formulations, time and budget considerations mean one must in practice work with one or two models one judges to be suitable for the context. The structures of the two models used are described below. The way in which they are calibrated is detailed in chapter 4 below.

3.2 PDM

The PDM (Probability Distributed Model; Moore 1985, 1999) is based on conceptual water stores, and represents non-linearity in the transformation of rainfall to runoff by using a probability distribution of soil moisture storage. This determines the time-varying proportion of the catchment which contributes to runoff, through either ‘fast’ or ‘slow’ pathways. The form of the PDM used here (and in FD0404), with five parameters, is illustrated in Figure 3.1 and described briefly below.

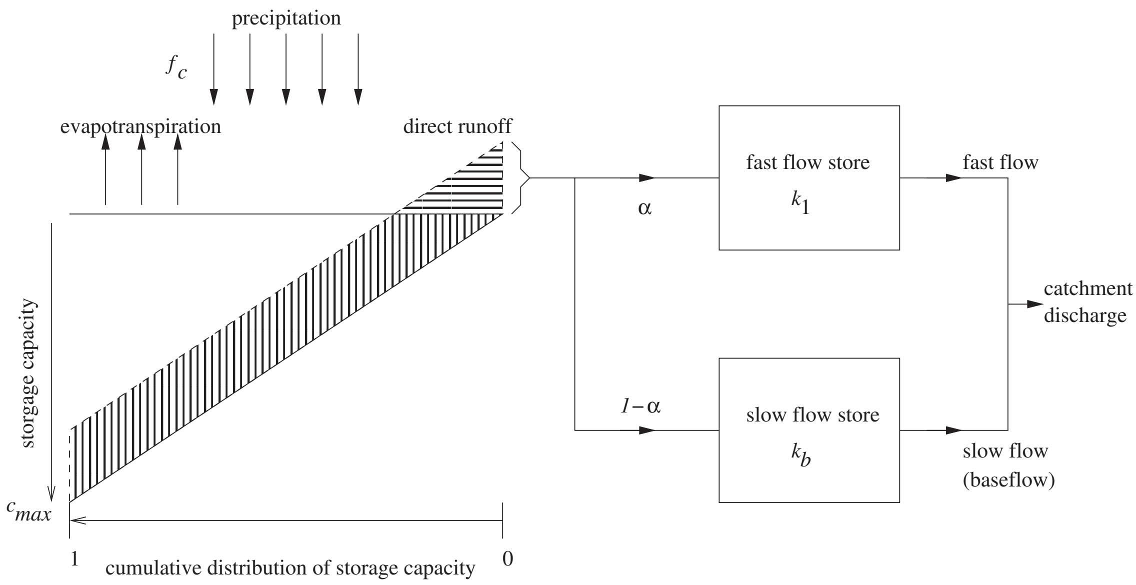


Figure 3.1 Structure of the 5-parameter version of the PDM runoff model.

Rainfall inputs to the soil store are first multiplied by a rainfall correction factor f_c, which can also, if required, compensate for loss or gain of water via lateral, sub-surface routes. The soil store can be depleted through evaporation, with content of the store determining the proportion of the potential evaporation which actually occurs. The distribution of the soil moisture storage capacity with store depth is assumed to be uniform, and the minimum store depth is set to zero. The maximum store depth is given by the parameter c_{max}. The soil store then generates direct runoff from a varying proportion of the catchment area, depending on how full it is. It is assumed that a proportion α of the direct runoff goes to the (linear) fast flow store, whilst $1-\alpha$ goes to the (cubic) slow flow store. The time constants of the fast and slow flow stores are k_1 and k_b respectively. The catchment discharge is then produced from a combination of fast flow (which may be conceptualised as [near-] surface runoff) and slow flow (base flow).

Although the model as described here has five remaining parameters, only four of these require specific calibration or generalisation for a catchment. The parameter determining the split between the fast and slow flow stores, α, is set as SPRHOST/100, where SPRHOST is a readily-available catchment property (standard percentage runoff inferred from soil information, see section 2.6). This was deemed an appropriate simplification, because of the directly comparable meanings of α and SPRHOST.

3.3 TATE

The TATE (Time Area Topographic Extension) model of Calver (1993, 1996) also involves conceptual stores, but routing is achieved using response functions which are based on the distribution of the drainage area with respect to distance from the river channel network. The structure of the three-parameter version of the model is illustrated in Figure 3.2 and described briefly below.

Parameter reduction was effected from earlier versions of the model by reference to experience of pre-project performance in a flood frequency context. Parameters to which results were less sensitive and those which showed a degree of correlation with others were reconsidered. The three parameters of the formulation used here are, broadly, a water balance control, a maximum storage capacity for soil and vegetation, and a control on the rate of fast runoff routing. Other aspects of the model which were formerly dealt with using specified parameters have been accommodated by using weighted mean values based on previous performance, or by a relationship to an internal state variable within the model. An example of the latter is the split between 'fast' and 'slow' runoff which was formerly a model parameter and is now a function of the (time-dependent) storage.

Precipitation enters the soil store, and evaporation depletes the store, the rate of evaporation being a function of the level of water in the soil store. The soil store has a maximum depth, given by the parameter *csm*. Overflow from the soil store, and downward drainage from the soil store, are functions of the content of the soil store. Drainage and any overflow are then combined at each time step and a proportion *1-crm* transfers to a routing mechanism with fast and slow components. The proportion *crm* does not reach the river above the gauging point, for instance because of consumptive use or percolation to deep groundwater. (Note that *crm* can be negative representing an overall gain of water to the catchment, for instance through lateral inflow of deep groundwater). Routing is via an area convolution, using the distribution of drainage area

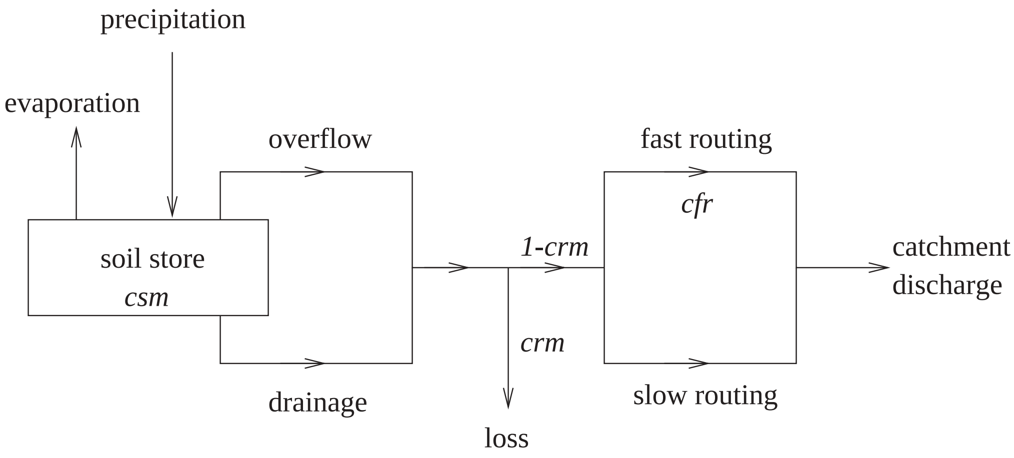


Figure 3.2 Structure of the 3-parameter version of the TATE runoff model.

with respect to time-of-travel from the river channel, with spatially-variable velocity, the parameter *cfr* defining the pattern of response of the fast flow. A concession to catchment spatial configuration in the TATE routing is provided by the use of a catchment area-distance function, whereby the incremental addition of drainage area is determined with increasing distance from the channel network. This was calculated from digital terrain data, as too was catchment drainage area.

4 MODEL CALIBRATION

Alison Kay, David Jones, Sue Crooks

PREVIOUS FAST TRACK BOX ON PAGE 21

CHAPTER 4 MODEL CALIBRATION

The chapter describes the calibration of the runoff models for the 119 catchments for which adequate river flow data are available. The calibration is an important step, not only in terms of assessing runoff model performance where data are available for validation, but also because it provides the basic data set of calibrated model parameters for subsequent use in the spatial generalisation procedures.

An automated calibration process was undertaken sequentially, taking each parameter in turn in a series of two passes, allowing each parameter to have its own objective function, determined by hydrological judgement and plots of objective functions against parameter values. The method allows more weight to be given to certain aspects of the flow regime, according to the parameter being calibrated. Lessons learned from the use of daily, as opposed to hourly data, and the performance of the calibrated models when compared with an extended data series, in this case including the flooding of autumn and winter 2000, were incorporated into the procedures.

The overall performance of the calibrations shows an absolute mean error of 5% (at the 10 year return period) for both runoff models, with a standard deviation about this mean of 7% for the TATE model and 5% for the PDM.

NEXT FAST TRACK BOX ON PAGE 43

4.1 Introduction

Calibration of the runoff models is an important aspect of the development and application of the continuous simulation approach to flood estimation, not just for site-specific applications but in the context of spatial generalisation: the calibrated parameter sets provide the basis for the establishment of relationships to catchment properties.

The calibration process aims to adjust model parameter values until a satisfactory or 'best' agreement is obtained between simulated and observed flows. This is, however, not as straightforward a process as it may sound. There are many methods and techniques which could be applied, from manual to fully-automated, and countless measures of fit of simulated and observed flows (objective functions). Different objective functions emphasise different aspects of the flow hydrograph, and so result in different optimum parameter sets. Inevitably then, the choice of method will depend on the nature of the application, as will the calibrated parameter sets. Wagener *et al.* (2004) provide a discussion of various calibration techniques used for conceptual runoff models as well as potential calibration problems, which include non-identifiability (where performance appears to be insensitive to the value of a particular parameter, at least with respect to a given objective function) and equifinality (where a number of parameter sets, sometimes widely-distributed across the parameter space, appear to perform equally well, at least with respect to a given objective function).

An additional factor which could affect calibration performance is data quality, for both rainfall and observed flows. Chapter 2 described the quality checking procedures applied during the production of catchment average rainfall data, which were designed to limit the effect of rainfall data errors, but the potential effect of some infilling of rainfall data, using catchment average profiles, needs to be borne in mind. The quality of flow data during floods is also a potential problem, due to difficulty in gauging, and there is little that can be done once the data are obtained from the gauging authorities. As described in Chapter 2, data quantity is also important; there must be a sufficient length of data, covering a range of flow regimes, for calibration to be effective.

Because of the larger number of catchments involved in this project compared to the pilot FD0404, it was considered infeasible to manually calibrate each catchment. Instead, a sequential method of automatic calibration was developed, bearing in mind the nature of the current application of flood frequency estimation. Although using automatic calibration removes some of the flexibility inherent in manual calibration, it also removes the subjectivity, resulting in a more consistent set of parameter sets that will, it is hoped, aid generalisation. Automatic calibration is also an advantage in that it allows the estimation of calibration uncertainty for each catchment, which can then be utilised within the generalisation procedure (so that catchments with less certain calibrations receive less weight in the generalisation).

The automatic calibration method decided upon in this project is described below, along with details of how the daily-data catchments were used alongside the hourly-data catchments. The testing of FD0404's calibrated parameter sets with the extended data acquired for this project is also summarised. Finally, details are given of the performance of the new calibrations.

4.2 Methods

4.2.1 Automatic sequential calibration

The parameters for both the PDM and the TATE model for each catchment have been estimated through an automatic calibration procedure which is an extension of that described by Kay *et al*. (2003). The procedure involves sequential calibration of each of the model parameters, in two passes. The method is outlined in the flow chart in Figure 4.1 and described below.

For the first pass, each parameter is calibrated in turn, following Monte-Carlo sampling of the parameter space of the so-far uncalibrated parameters. A different objective function is chosen for fitting each parameter, according to hydrological judgement and plots of objective functions versus parameter values. Each objective function used here considers the whole flow time-series, but some give more weight to certain aspects of it.

For example, when calibrating the parameter determining the overall water availability (f_c for the PDM and *crm* for TATE) only 30-day average flows are considered, whereas when calibrating the parameter of the fast flow store (k_1 for the PDM and *cfr* for TATE) more weight is given to the higher hourly flows. The objective functions used for each parameter are given in Table 4.1.

For the second pass, each parameter is calibrated in turn following Monte-Carlo sampling of its own value, with the values of the other parameters held at their previous calibrated values. This second pass allows a re-adjustment of parameter values, once other values have been estimated, and uses the same objective functions as the first pass. Finally, it was found helpful to allow a final re-calibration of the first parameter for each model (f_c for the PDM and *crm* for TATE), this time using an objective function that concentrates on fit of the flood frequency curve (above the 2-year return period).

As this method is sequential, alternative ordering of parameters in the calibration is possible, and no testing has been done to assess the effect of this. The order used for each pass has been f_c, k_b, c_{max}, k_1 for the PDM and *crm, csm, cfr* for TATE. The choice of order has been based on hydrological judgement and on the behaviour of the response surfaces in the objective function dot plots.

4.2.2 Calibration uncertainty

Automatic calibration allows the use of a jack-knifing procedure (Shao and Tu 1995) to estimate calibration uncertainty. Briefly, after each catchment is calibrated with all of its *N* years of rainfall and flow data, *N* re-calibrations are performed. The *i*th re-calibration is performed by treating the *i*th year of observed flow data as missing. (Note that all of the rainfall data are used each time, to maintain continuity of the simulated time-series.) In this way, *N* new sets of calibrated parameters are derived for each catchment, and the spread of these values gives an indication of the calibration uncertainty for each catchment due to the finite amount of data available for calibration. For a given catchment, the variance of the estimation error of the *p*th model parameter, $\sigma^2(p)$, is calculated from the N values for that parameter, $\alpha(p)$, using the formula

$$\sigma^2(p) = \frac{N-1}{N}\sum_{i=1}^{N}\left(\alpha_i(p) - \frac{1}{N}\sum_{j=1}^{N}\alpha_j(p)\right)^2 .$$

Here $\alpha_i(p)$ is the estimate of parameter $\alpha(p)$ treating the *i*th year of flow data as missing. Similarly, the covariances of the errors of pairs of parameters $\alpha(p)$ and $\alpha(q)$ for a catchment can be calculated as

$$c(p,q) = \frac{N-1}{N}\sum_{i=1}^{N}\left(\alpha_i(p) - \frac{1}{N}\sum_{j=1}^{N}\alpha_j(p)\right)\left(\alpha_i(q) - \frac{1}{N}\sum_{j=1}^{N}\alpha_j(q)\right).$$

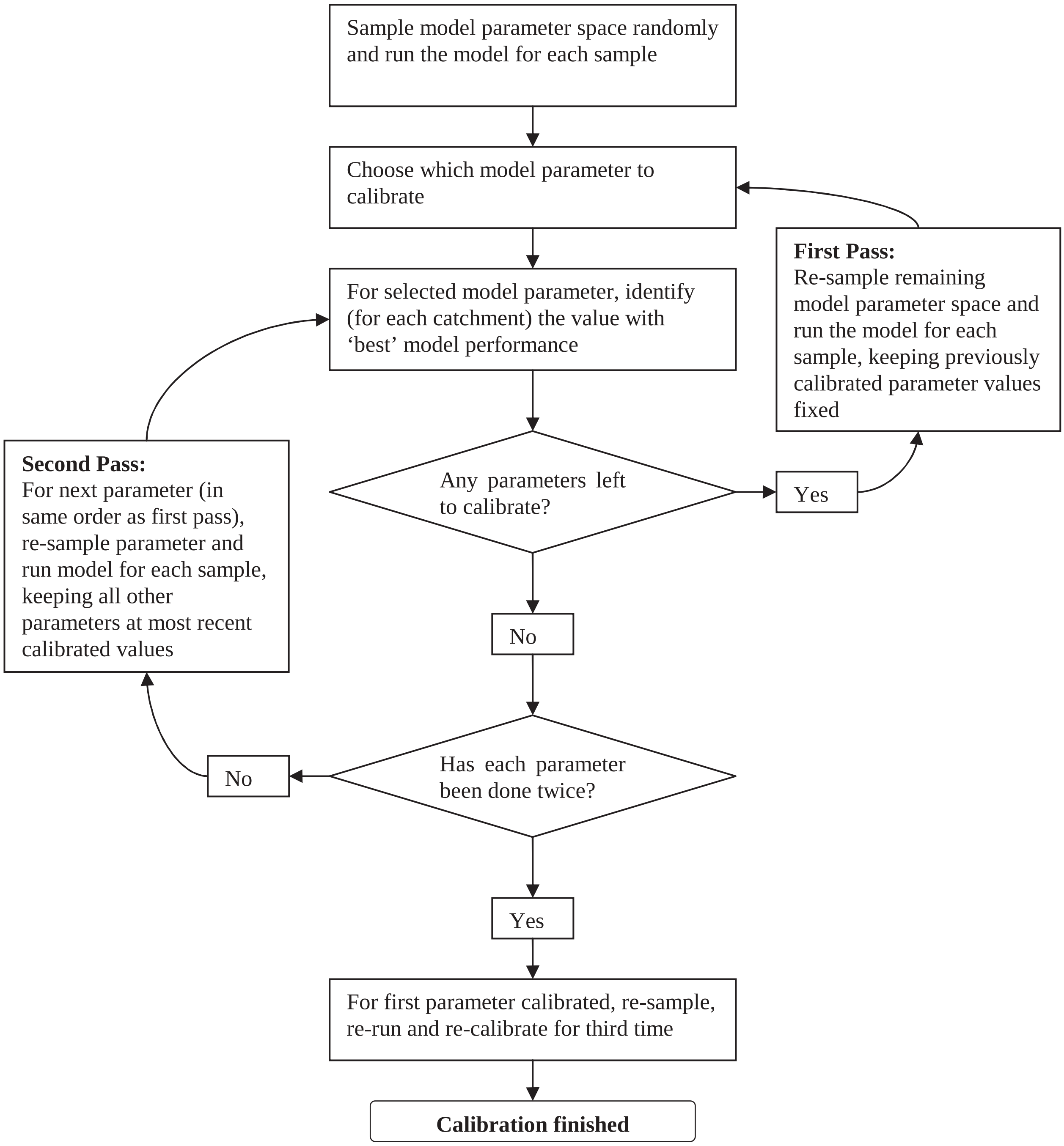


Figure 4.1 Outline of the automatic sequential calibration method.

Table 4.1 Objective functions used for the calibration of each model parameter in each of the first two passes.

PDM		TATE	
parameter	objective function	parameter	objective function
f_c	N&S on 30-day mean flows	*crm*	N&S on 30-day mean flows
c_{max}	N&S on output flows	*csm*	Sum of the absolute errors in flows
k_1	Sum of the absolute errors in flows, weighted to high flows	*cfr*	Sum of the absolute errors in flows, weighted to high flows
k_b	Sum of the absolute errors in flows, weighted to low flows		

NB: N&S = Nash Sutcliffe efficiency measure (Nash and Sutcliffe 1970)

The variance of the estimation error is used in Chapter 5 in an uncertainty weighting scheme in which calibrated parameters for different catchments are combined with higher or lower weight being given to catchments with lower or higher calibration uncertainty. The covariances will be used in Chapter 6 for the development of a scheme for providing uncertainty bounds on generalised flood frequency curves for ungauged catchments.

4.3 Use of catchments with daily rainfall and flow data

A key consideration was how to merge information from catchments with data (both rainfall and flow data) at different time steps. The hourly time step was originally chosen as a compromise between the desirability but practical difficulties of, say, 15-minute data, and the wider availability but possible loss of information, particularly about peak flows, from daily data. Although only larger catchments have been chosen for use with daily data (a lower limit for catchment area of 50 km^2 was chosen), there will still be some loss of information, and this must be borne in mind. Note that the daily flow data is a daily *mean* flow, but the hourly flow data is instantaneous, on-the-hour.

The specific use of the data and the way in which the data interacts with the runoff models needs to be taken into account when considering how to use daily sites together with the original hourly sites. There are three ways in which time-step lengths affect the implementation of the runoff models. These relate to

- the time-step length of the input rainfall data (input time-step);

- the time-step length required to match the recording interval of calibration flow data, or the interval at which simulated flow data are to be saved (output time-step);
- the time-step length at which calculations are performed within the model (internal time-step).

Frequently in conceptual hydrological models all three of these time-steps will be the same. This was the case for the hourly sites in FD0404, and the first consideration was whether the daily sites could follow in the same way (with all three time-step lengths being 1 day).

In theory, both of the runoff models (PDM and TATE) are structured in such a way that their parameters have the same intrinsic meaning (in particular, the same units) regardless of the length of the internal time-step. This is vital for the current application, where many catchments are being treated simultaneously and ungauged catchments are treated using information from gauged catchments. However, tests on a small number of the hourly catchments (Kay *et al.* 2003) showed that the parameters would be calibrated differently were they treated at a daily time-step (e.g. Figure 4.2). This is probably due to the fact that daily mean flow data contain less information for calibration than does the hourly flow data. Thus using a daily internal time-step for the daily sites was not appropriate for this application.

Instead, a common internal time-step of 1 hour was adopted for all of the catchments, whether they have hourly or daily rainfall and flow data. However, the issue then arises of how to convert daily input rainfall data to an hourly time-step, as the input time-step cannot be longer than the internal time-step. By far the simplest option is to uniformly spread the daily rainfall over each of the 24 hours of the day. The use of non-uniform profile(s) to disaggregate daily rainfall would have required a number of arbitrary choices (for instance, whether the profile varied by location, rainfall amount, season etc.) which would be difficult to justify.

It is also necessary to convert simulated flows from the internal time-step to the output time-step, to enable a comparison with calibration flow data. This is simple for the case of an hourly internal time-step with daily mean calibration data, as the 24 simulated hourly values for a day are simply averaged to produce a simulated daily mean flow. Tests showed that calibrations using this simple, uniform disaggregation and comparison, method led to values matching those using full hourly data sufficiently closely (e.g. Figure 4.2).

4.4 Deriving flood frequency curves

To gain the most information from the available hourly data records, the partial duration series or peaks-over-threshold (POT) method was used to fit flood frequency distributions to the observed and simulated flow data. Single-site POT analyses have been investigated theoretically by a number of authors (Davison, 1984; Davison and Smith, 1990; Smith, 1984; Wang, 1991). The methods for POT analysis reported by Naden (1992) were adopted.

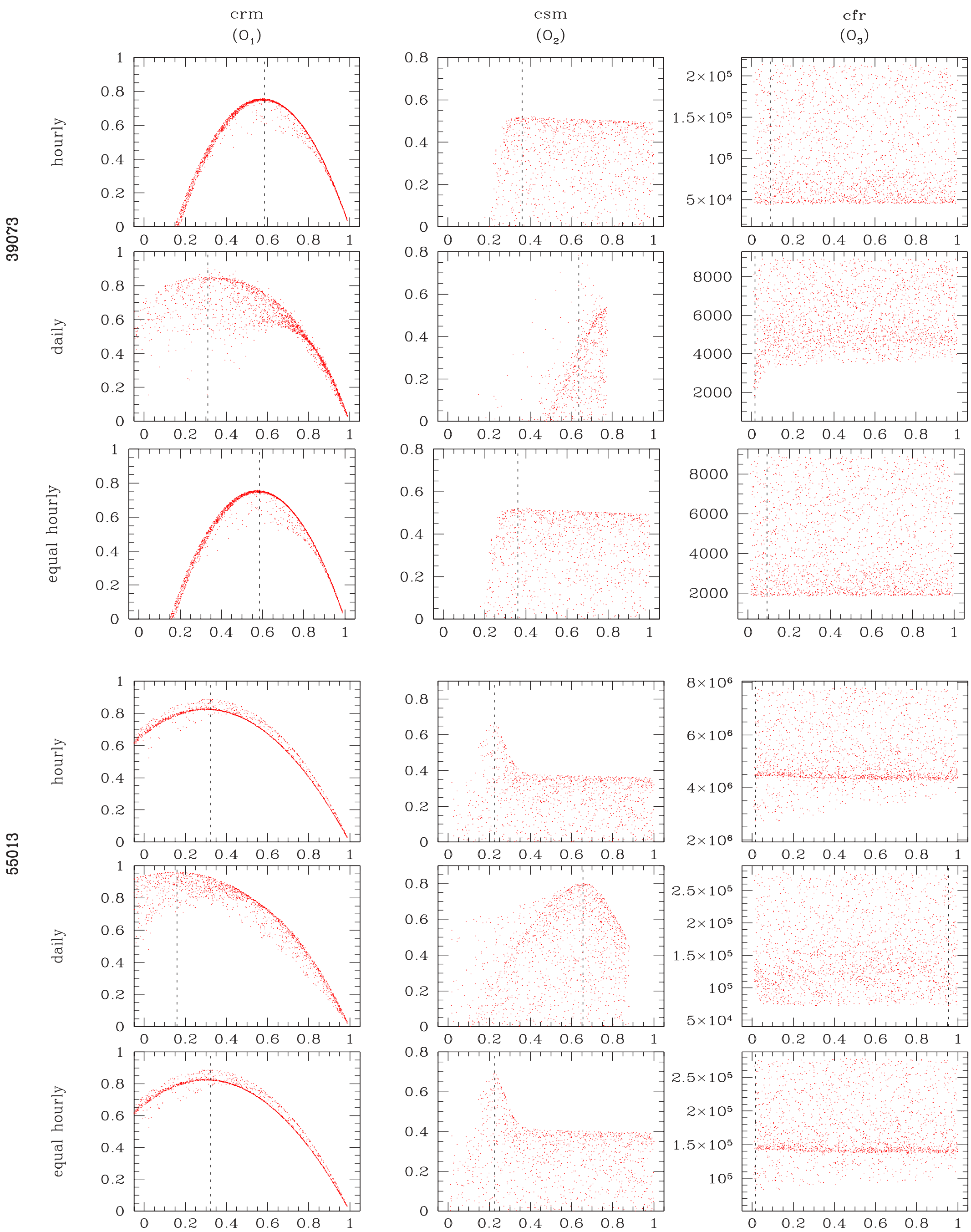


Figure 4.2 **Dot plots of objective functions for the three parameters of the TATE model, comparing calibration performance for two catchments (39073 and 55013) using observed hourly data (top), daily data (middle), and hourly data uniformly disaggregated from daily data (bottom). Optimum values for the parameters are indicated by the vertical dashed lines, and show that calibration using daily data gives very different calibrated values to that using observed hourly data, but that uniform hourly data is much closer. (O_1=Nash Sutcliffe on 30-day mean flows [maximised], O_2=Nash Sutcliffe on daily mean flows [maximised] and O_3=weighted sum of absolute errors [minimised]).**

For both simulated and observed flow series, the magnitudes of the POT data were fitted using the generalised Pareto distribution (GPD), with the peak arrival times assumed to correspond to a Poisson distribution. The combination of these two assumptions is equivalent to the use of the Generalised Extreme Value distribution for annual maxima. Fitting was carried out using the method of probability weighted moments (Hosking and Wallis, 1987). Rather than specifying an arbitrary flow threshold for the extraction of peaks for each catchment, peaks were instead extracted at an average rate of three per year (i.e. 30 peaks would be extracted from a period of 10 years, but not necessarily three peaks from every calendar year). To ensure that the extracted peaks represent independent events, a minimum separation period and a flow criterion are imposed, as defined in the Flood Studies Report (FSR) and the Flood Estimation Handbook (FEH). The flow criterion specifies that the flow between two peaks must drop to at least two thirds of the first peak. The minimum separation period is specified as three times a typical event time-to-rise for a catchment. However, calculation of time-to-rise requires gauged data for the catchment. As a method is required for ungauged catchments, time-to-peak will be used instead of time-to-rise as there is an existing method of estimating time-to-peak from catchment properties (Institute of Hydrology 1999, Volume 4). Figure 4.3 compares the two values for a set of about 40 catchments. The FSR conditions are in fact set somewhat arbitrarily, and it is not believed that the use of time-to-peak rather than time-to-rise will introduce any particular bias; the most important thing is consistency, and since a method for ungauged catchments is required then the use of time-to-peak is considered a sensible and straightforward choice.

It is perhaps worth emphasising that the POT analysis is used here as a convenient method for presenting flood peak data, rather than specifying statistical models *per se* for the flood data. The fitted GPD distributions are not necessarily intended to be applied for extrapolation or as ‘true’ underlying flood peak distributions, but serve rather to interpolate between modelled peaks. For this reason, no detailed analysis is presented to assess the suitability of the distributions fitted to the POT data, nor the implicit flow thresholds resulting from the average extraction rate of 3 peaks per year. However, the GPD is a flexible distribution capable of fitting data that lie in the form of Extreme Value distribution Types I, II and III (i.e. flood frequency curves that plot as a straight line, curve upwards or tend to ‘flatten out’ on Gumbel paper). Also, Naden (1992, 1993) has presented evidence, based on the analysis of POT data from 826 stations, that these techniques are likely to be reasonable for the purposes of this part of the method, i.e. single-site analyses in the UK for relatively short return periods.

4.5 Testing FD0404 calibrations on extended data period

An early task (Crooks *et al.* 2002) was to test the earlier (FD0404) runoff model calibrations for the hourly catchments on the extended data period available to this project; generally 1985-2001 as against the previous end date of around 1995, in particular considering the impact on the fit of the flood frequency curve, and whether the floods of Autumn 2000 were well simulated.

For the FD0404 catchments manually calibrated parameter sets were available for the TATE model, whereas for the PDM the parameter set used for testing was actually a combination of generalised and calibrated values, arising from the sequential regression

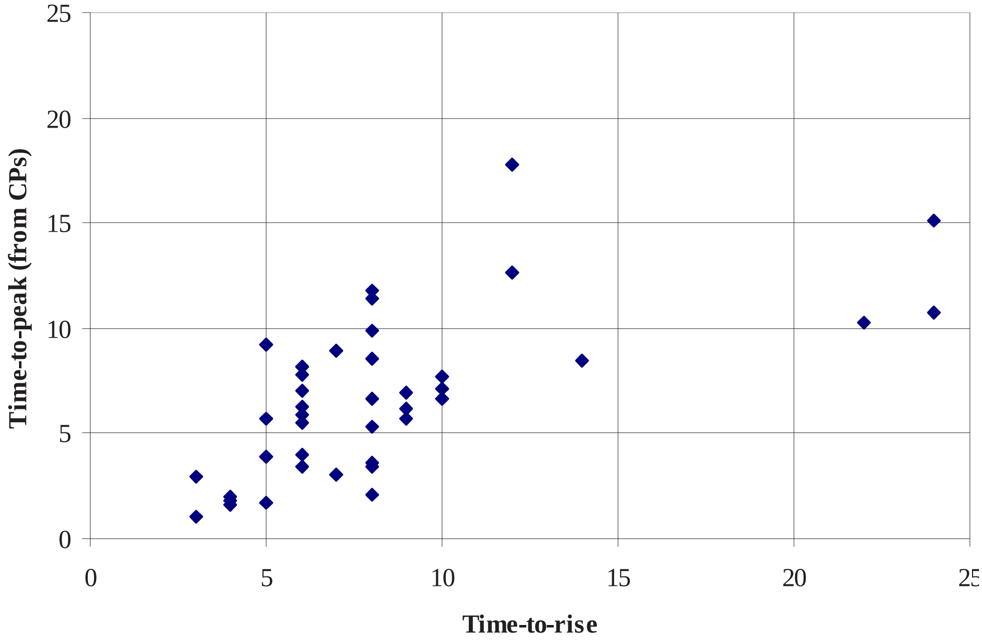


Figure 4.3 Time-to-peak (calculated from catchment properties) versus time-to-rise (calculated from data).

technique. Although the latter set may not provide the best fit that can be achieved with the model, it nevertheless provided a way of assessing the effect of using a parameter set derived for a given data period on an extended data period.

For the analysis, the set of catchments was divided into two groups; one where the extension of the data period increased the range of the high flow values over that in the original data period (group A), and one where there was no such increase, i.e. where the highest flow value still occurred in the original period (group B).

4.5.1 Impact on fit of flood frequency curve

A qualitative assessment was made of the fit of the flood frequency curve for the extended data period, using the parameter sets derived for the original data period. For each catchment the performance was categorised as one of over-estimated, reasonable or under-estimated. The performance for the original data period was categorised in the same way, to assess whether catchments had moved between categories. The results are summarised by group in Table 4.2. An example of simulated flood frequency curves compared against observed for each data period is given in Figure 4.4.

The results show that, whether or not the flow range was increased by the extended data period, the transposed parameter sets provide a good fit of flood frequency curves for around 50% of catchments. However, under-estimation is more likely in the group where the flow range has been increased. Catchments most likely to need re-calibration with extended flow range were those with a fast response time. Manual calibration of

Table 4.2 Catchment analysis of flood frequency curve fit, indicating the number of catchments falling within each performance category for the extended data period, as well as the performance of these catchments for the original data period (Over-estimated, Reasonable, Under-estimated). The latter are colour-coded according to whether the impact of the extended data period has been good (blue), bad (red) or neither (green).

Group (number of catchments)	TATE			PDM		
	Over-estimated	Reasonable	Under-estimated	Over-estimated	Reasonable	Under-estimated
A: Range increased (20)	1	10	9	3	9	7
	1 O	2 O 7 R 1 U	7 R 2 U	3 R	3 O 5 R 1 U	1 O 6 R
B: Range not increased (15)	4	8	3	5	8	2
	3 O 1 R	7 R 1 U	3 U	1 O 4 R	8 R	1 R 1 U

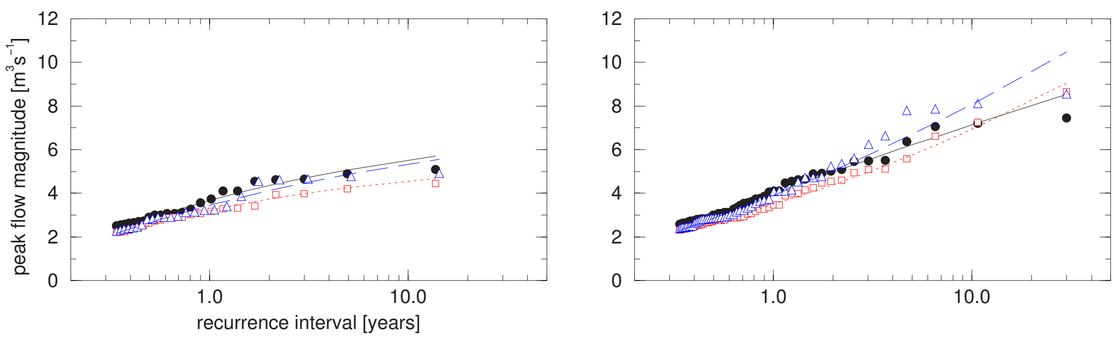


Figure 4.4 Observed (black, filled circles and solid line) and simulated flood frequency curves for catchment 27051 for the PDM (blue, open triangles and dashed line) and TATE (red, open squares and dotted line). The graph on the left shows the results for the original data period and that on the right shows the results for the extended data period.

TATE may have provided more temporally consistent parameter sets but these sets were more likely to underestimate peak flows than the automated sequential method used for the PDM.

4.5.2 Performance during autumn/winter 2000 floods

The winter of 2000-01 resulted in floods in many of the catchments, though for most it was the spatial extent of the flooding and duration of elevated discharges which was notable rather than the peak flow. The highest recorded peak occurred in only 10 of the 37 catchments. To assess model performance during this period (October 2000 to March 2001) the percentage difference between the maximum observed and simulated peak flows was determined. The results are summarised in Table 4.3, arranged both by group A (range increased) and group B (range not increased) as before, and arranged according to whether the peak occurring in the period was the highest in the whole record or not. (Some catchments are excluded from this summary because of missing data during the crucial period).

Table 4.3 Summary of model performance in terms of percentage difference between modelled and observed maximum flow during October 2000 to March 2001.

Model	Group (number in group)	Percentage difference (modelled – observed)		
		Minimum	Mean	Maximum
TATE				
	A (19)	-50.0	-15.1	15.0
	B (13)	-57.8	-12.4	57.3
	Highest peak (10)	-50.0	-12.6	15.0
	Non-highest peak (22)	-57.8	-15.1	57.3
PDM				
	A (18)	-39.0	-7.6	34.0
	B (13)	-61.6	12.2	56.5
	Highest peak (9)	-36.1	-5.5	34.0
	Non-highest peak (22)	-61.6	3.3	56.5

The results show that model performance for the winter 2000-01 flood events was not affected by the extreme conditions, as the high peaks are no more likely to be under- or over-estimated than peaks which were more ‘middle-of-the-range’. Also, the PDM simulated peak flows tend to underestimate less than those from TATE, confirming the difference found between the two models from the flood frequency analyses.

4.5.3 Conclusions arising from new data periods

Although the performance of the transposed parameter sets for the extended period was generally encouraging, the increased likelihood of under-estimation for catchments where the flow range was increased meant that re-calibration was recommended.

The performance of TATE and PDM under transposition of original parameter sets is similar despite the fact that those parameter sets were derived, in FD0404, in quite different ways (one manual the other using a sequential regression/calibration method). This suggests that there is unlikely to be a disadvantage in using the sequential method, which can be automated, and that such a sequential method can therefore be safely used for re-calibrating the catchments using the longer data series.

The automatic calibration method described earlier in this chapter was thus applied to all catchments (old and new, hourly and daily), as it was considered potentially important that calibrated parameter sets be derived in a consistent way that can also be used in the quantification of uncertainty.

4.6 Performance of new calibrations

Examples of calibrated flows and flood frequency curves are given in Figure 4.5 for five catchments, of differing types and locations, for each runoff model. Calibrated flows and flood frequency curves for all catchments and both models are given in Appendix A. (All appendices are in the Project Record.)

Overall, the performance of the automatic calibration method was good. Table 4.4 summarises the performance across sets of catchments in terms of absolute percentage error (simulated compared to observed flood frequency curve) at various different return periods. The performance for just the hourly catchments is given as well as that for all of the catchments, for better comparability with the performance of earlier (FD0404) calibrations, which are also given in the table (taken from Table 4.3 of Calver *et al.* 2001). However, it should be noted that these values are still not directly comparable, as the set of hourly catchments has increased (with the addition of further catchments in Scotland) and the data for the catchments used in FD0404 has been extended in time. Also, for the PDM the calibration performance is given for the seven-parameter version of the model used in FD0404 (prior to the simplification of the model), rather than the five-parameter version used here. No direct calibration of the five-parameter PDM was performed in FD0404, as it was used there only within the sequential regression method.

Despite this, the performance of the new calibrations is similar to or better than those used previously: for the TATE there is a particular improvement at higher return periods (5 years and above). For the PDM the improvement is less obvious, but performance is still good considering that one could expect a model with more parameters to perform better during calibration, due to the greater flexibility afforded by those extra parameters. It is to be expected that the performance averaged over just the

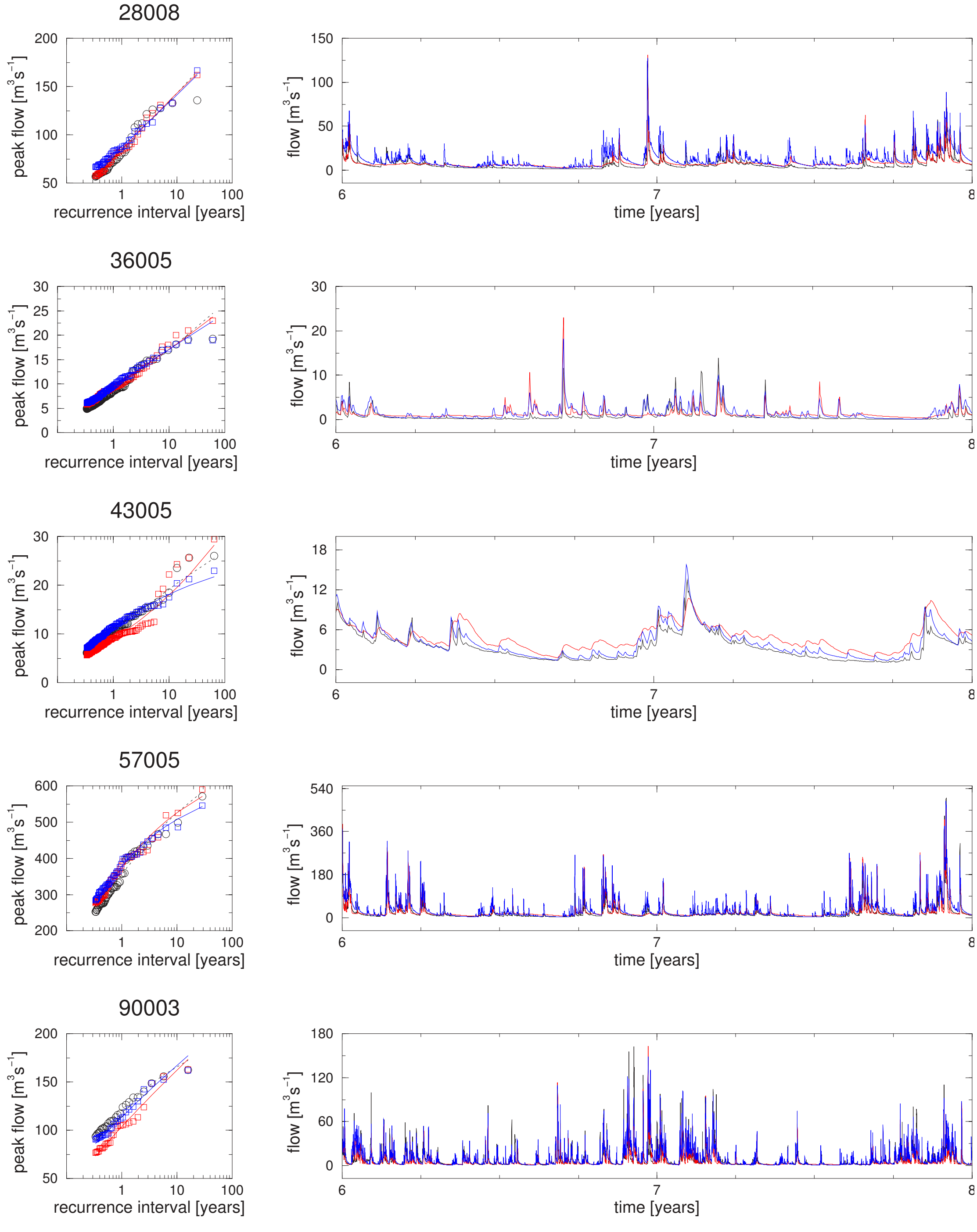


Figure 4.5 Example calibrated flows and flood frequency curves modelled with the PDM (blue) and TATE (red), compared to observations (black), for five catchments of differing types and locations.

Table 4.4 Overall calibration performance, in terms of the mean and standard deviation (SD) of absolute percentage errors at different return periods. The performance of the new calibrations is averaged over all the catchments, and just for the catchments with hourly data, for better comparability with performance of earlier (FD0404) calibrations.

		Mean and SD of absolute % errors	Return Period (years) 1.0	2.0	2.33	5.0	10.0	20.0
TATE	3-parameter, new calibration (all sites)	Mean	11	9	8	5	5	9
		SD	10	9	8	7	7	8
	3-parameter, new calibration (hourly sites)	Mean	12	10	10	7	6	10
		SD	12	10	10	8	7	8
	3-parameter, old calibration (FD0404 sites)	Mean	12	10	9	8	10	13
		SD	13	13	13	12	12	14
PDM	5-parameter, new calibration (all sites)	Mean	9	6	6	4	5	7
		SD	7	5	4	4	5	7
	5-parameter, new calibration (hourly sites)	Mean	9	7	7	5	6	9
		SD	9	6	5	5	6	8
	7-parameter, old calibration (FD0404 sites)	Mean	8	6	6	5	5	7
		SD	9	7	7	6	6	9

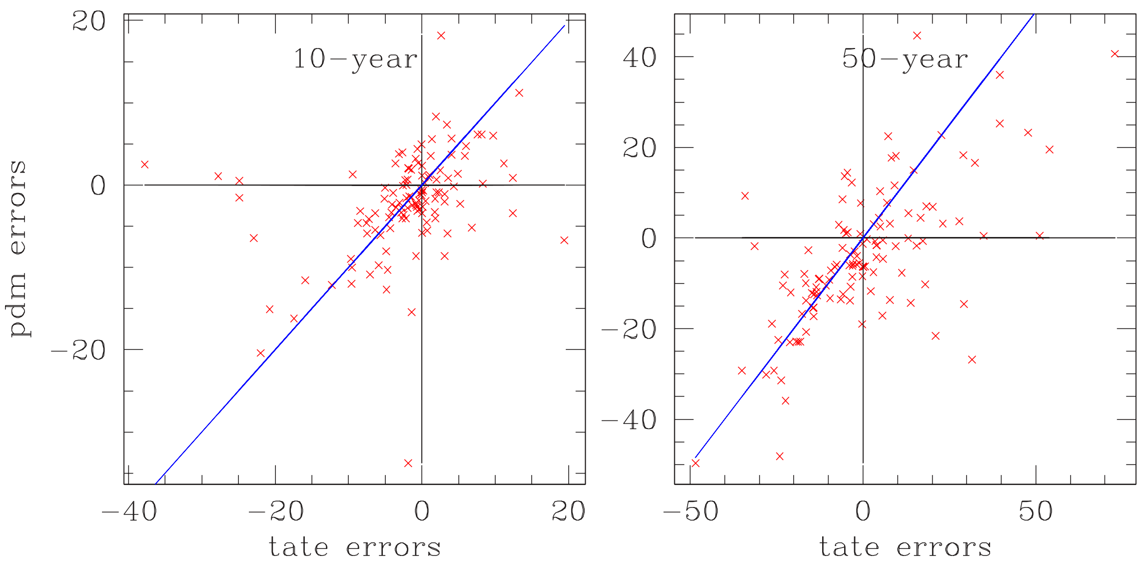


Figure 4.6 Calibration errors for PDM versus those for TATE, at two return periods (10 and 50 years).

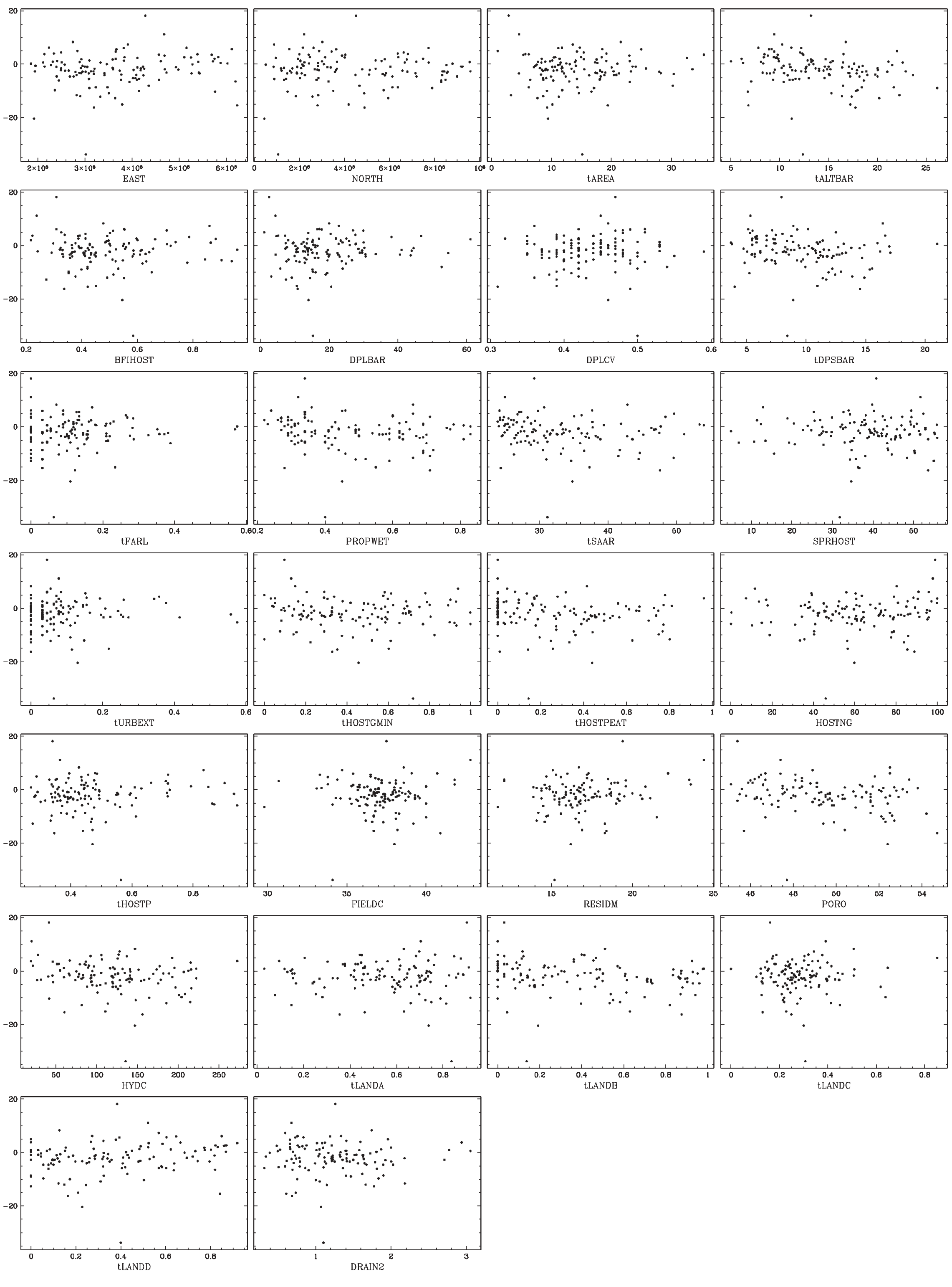


Figure 4.7 PDM calibration errors (percentages) versus catchment properties.

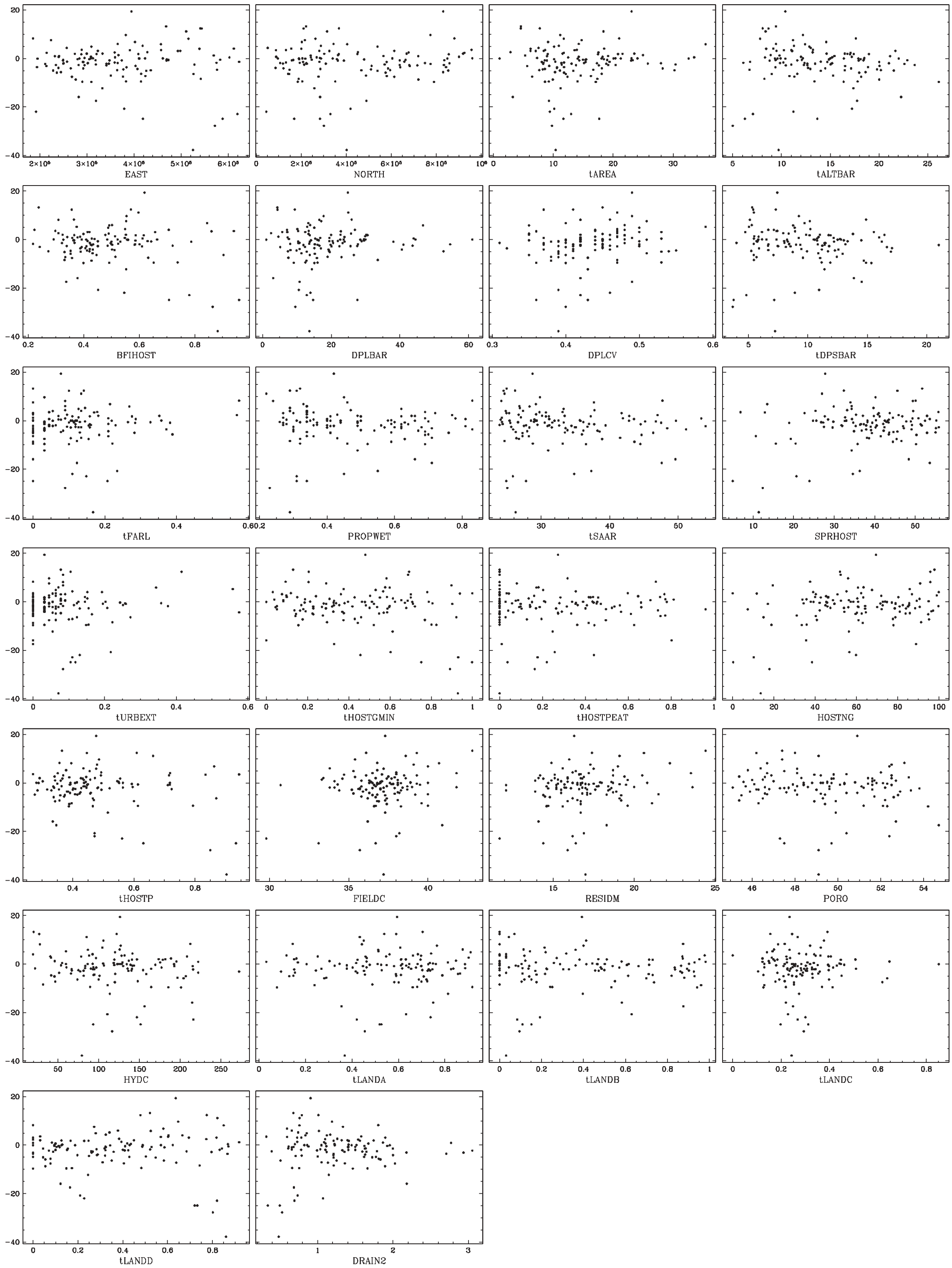


Figure 4.8 TATE calibration errors (percentages) versus catchment properties.

Figure 4.9 **Plots of catchment properties against calibrated parameters. Catchment properties prefixed with a 't' are transformed according to the functions given in Table 2.4. Note that some of the parameters have been log-transformed (see Section 5.3).**

hourly catchments is marginally worse than that over all the catchments, as the models may be harder to fit at the hourly than daily time-step.

Figure 4.6 compares the calibration errors for the two models, at two return periods. It shows that there is no obvious advantage of one model over the other. In general, if one model has a higher error for a particular catchment then the other model has a higher error for that catchment too.

Figure 4.7 and Figure 4.8 illustrate the PDM and TATE calibration errors (respectively) versus the 24 catchment properties (Table 2.3) as well as the Easting and Northing of the catchment outlet. These figures show that there is no obvious dependence of calibration performance on catchment properties or location. Figure 4.9 illustrates plots of catchment properties versus calibrated parameter values, showing possible relationships between the two and so demonstrating the potential of generalisation using catchment properties: the subject of the next chapter.

5 SPATIAL GENERALISATION

Alison Kay, David Jones, Thomas Kjeldsen, Sue Crooks, Ann Calver

PREVIOUS FAST TRACK BOX ON PAGE 25

CHAPTER 5 SPATIAL GENERALISATION

The subject of this chapter is the means by which the practitioner can derive flood frequency estimates for catchments which have little or no runoff data, the so-called ungauged sites which in fact represent much of the country. The essence of the approach is the derivation of runoff model parameters from the widely-available catchment properties (or their transformations) described in Chapter 2.

Whilst there are many variants of spatial generalisation approaches, decisions had to be made as to which would, in a time-limited project, be explored. Attention has been focused on

- (i) multiple univariate regression, whereby each runoff model parameter is defined as a function of catchment properties
- (ii) 'sequential' regression whereby such predictive equations for model parameters are derived in a sequence, rather than independently, in order to try to account for the effect that already-generalised parameters have on the remainder of the parameter set
- (iii) 'site-similarity' approaches in which runoff model parameters are weighted means of parameter values from a suite of catchments of similar hydrological response, as characterised by key catchment properties.

The chapter details the methods developed and explored, together with the test results from variants of the core methods. It is on the basis of these results that issues such as the selection of catchment properties to include in (i) and (ii) above, and the number of 'similar' catchments to include in (iii) above, have been decided.

The performance of the methods developed was compared by treating sites as if they were ungauged and comparing generalised ungauged flood frequency results with (withheld) observational data up to the frequencies appropriate to the observed record length. The overall result from the three types method developed, used with the two runoff models (Chapter 3), was a general similarity in performance level. This may be seen as an indication of having extracted as much information from the data as possible in terms of transferring hydrological response prediction to data-sparse sites: strictly, however, this comment is informed speculation rather than proof of that fact.

FAST TRACK BOX CONTINUED ON NEXT PAGE

FAST TRACK BOX CONTINUED FROM PREVIOUS PAGE

Inspection and analysis of the errors of different generalisation approaches have not permitted a comprehensive deterministic error model to be formulated for this extremely complex issue. The choice of a specific preferred generalisation approach can only, at this stage, be endorsed at an overall level: because there is little to choose between a number of generalisation methods: some catchments, or types of catchments, may be better addressed by particular approaches other than the one which performs best overall on the national sample of catchments tested. The recommended methods are defined in Chapter 9 after consideration (Chapter 8) of associated uncertainties in flood frequency.

NEXT FAST TRACK BOX ON PAGE 77

5.1 Introduction to generalisation

The overall aim of this project is to provide a method for estimating parameters of lumped conceptual runoff models for any catchment in Great Britain to enable a flood frequency analysis to be carried out through continuous simulation modelling. This chapter provides a review of previous work related to spatial generalisation of runoff model parameters for lumped conceptual runoff models, followed by development and comprehensive comparison of three different methods. The methods are:

- multiple univariate regression, relating individual model parameters to selected catchment properties through linear statistical models,
- sequential regression, developing predictor equations sequentially, thereby accounting for the effect that already generalised parameters have on the following parameters, and
- a site-similarity approach, where model parameters are weighted averages of parameter values obtained at hydrologically similar gauged catchments.

It should be noted that although specific approaches of regression and site-similarity have been highlighted, they can be considered as special cases of approaches from a whole spectrum of related approaches, which could include local regression or kriging for instance. Figure 5.1 presents the site-similarity and regression approaches in terms of catchment similarity and linkage functions, alongside the approaches taken in other work, including the Flood Estimation Handbook (FEH; Institute of Hydrology 1999). The forerunner project FD0404 (Calver *et al.* 2001) researched some regression approaches, which have been enhanced in FD2106, particularly in the context of incorporating uncertainty estimates. Site-similarity methods have been introduced in FD2106.

A key point to note is that this field of research is a major hydrological modelling challenge and can be seen, at a pure research level, to be open-ended. In order, therefore, to reach closure for a practical methodology a considerable element of judgement needs to be brought to bear. In short, the research, given the time scale, cannot be exhaustive. Testing of approaches to spatial generalisation is, by the nature of the problem being addressed, an incomplete matter. Comparisons can, of course, be drawn with site-specific calibrations: these are, however, constrained in the extent of their coverage by the range of recurrence intervals and magnitudes of events covered in periods of records. The performance measure of the generalisation methods has been significantly enhanced by introducing a more comprehensive framework for analysis of different sources of uncertainty. The uncertainty framework is described in its entirety in Chapter 6, but implications of the work have been adopted and used for decision-making in this chapter.

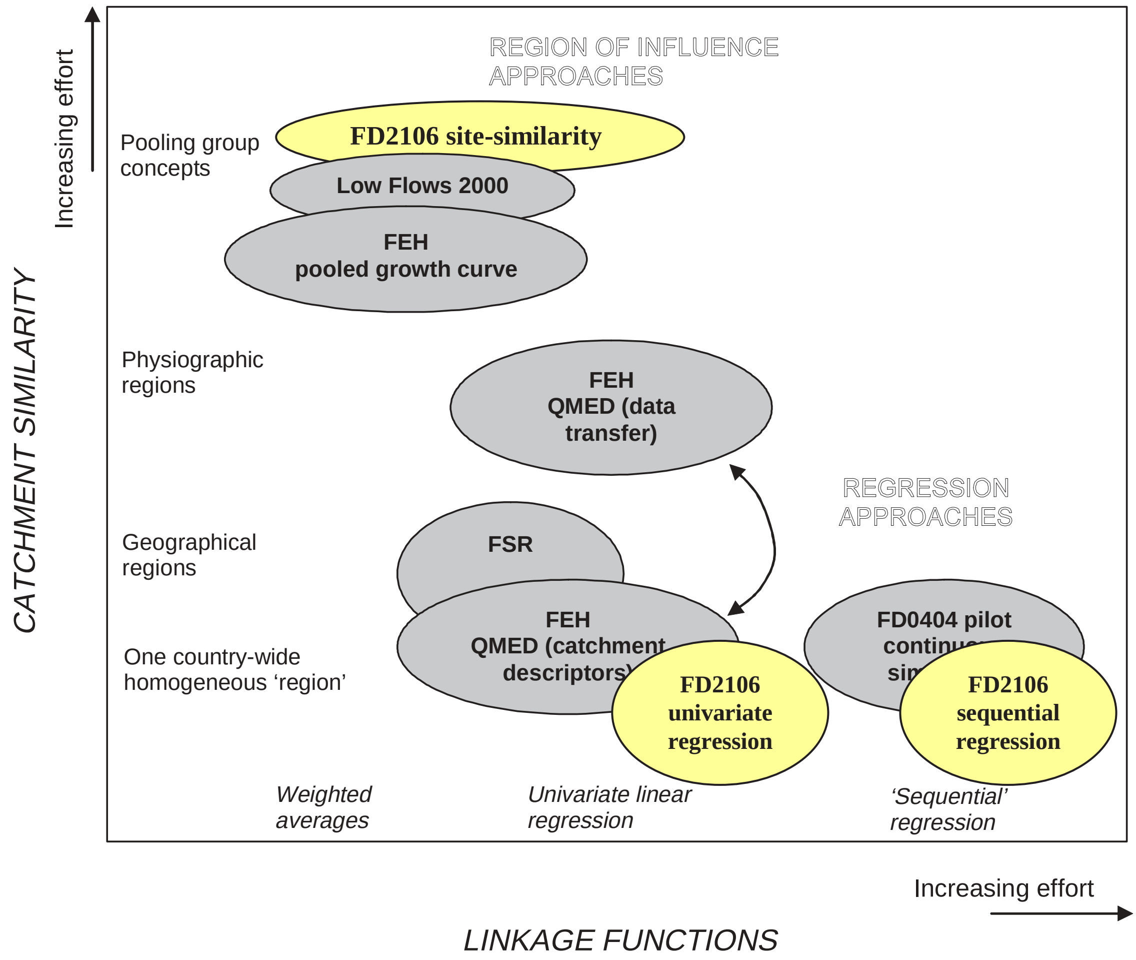


Figure 5.1 Schematic outline of spatial generalisation approaches for estimating hydrological responses at ungauged catchments, in terms of catchment similarity and linkage functions, indicating the 'location' of the site-similarity and regression approaches of this project alongside approaches taken in other work

5.2 Methods of generalisation

5.2.1 Regression approaches

Regression analysis has traditionally been a profitable approach when developing tools for estimation of model parameters at ungauged sites. Examples of the development of spatial generalisation of runoff model parameters through regression analysis are available both in the UK (NERC 1975, Sefton and Howarth 1998, Houghton-Carr 1999) and elsewhere, such as Western Europe (Niggli *et al.* 2002), the USA (Abdulla and Lettenmaier 1997) and Australia (Post and Jakeman 1999). The coefficents of the regression model are estimated from coherent sets of calibrated runoff model parameters and available catchment properties, frequently through least square techniques.

In the regression approach adopted here, each parameter of the two runoff models (PDM or TATE) is treated separately: α is used to denote a typical value for the parameter being considered, while α_a is used to denote the calibrated parameter value for catchment a. In the regression approach, a model parameter α is estimated using an equation of the form

$$\alpha^R = \beta_0 + \sum_{j=1}^{J} \beta_j X_j$$

where

β_j = jth regression coefficient, $j = 0,\ldots, J$,
X_j = jth catchment property, $j = 1,\ldots, J$.

In particular, the regression based estimate of α for catchment a is given by

$$\alpha_a^R = \beta_0 + \sum_{j=1}^{J} \beta_j X_{a,j} \; , \tag{5.1}$$

where $X_{a,j}$ is the value of the jth catchment property for catchment a.

The superscript R refers to an estimate obtained using linear regression. The choice of the relevant catchment properties to use in the regression equation, including any transformations of catchments properties, is a critical factor in the performance of the approach. To ensure that no useful combinations of catchment properties were excluded, an exhaustive search procedure was used to assess all possible combinations of up to six catchment properties (see Section 5.4 for more information).

Both standard and weighted least-squares regression have been applied, treating each of the runoff model parameters separately. The weights in weighted regression are given by

$$w_i = \frac{1}{1 + k\sigma_{i,\varepsilon}^2}$$

where the $\sigma_{i,\varepsilon}^2$ are values describing the uncertainty (variance) in the calibrated model parameters for catchment i, with k a constant that can be estimated iteratively. The values of $\sigma_{i,\varepsilon}^2$ are estimated through jack-knifing (see Chapter 4). On convergence, $1/k$ gives an estimate of the generalisation uncertainty for each parameter (σ_η^2). Alternatively, setting k=0 corresponds to unweighted regression, and performing the first step of the iterative procedure gives an estimate of generalisation uncertainty in this case: see Chapter 6 (Section 6.4.3) for further details.

5.2.2 Sequential regression

The multiple univariate regression models described above do not account for the dependence between the individual runoff model parameters. Alternatively, multivariate regression could be adopted, which recognises inter-dependence between runoff model parameters. However, the requirements of assumptions concerning the form of the dependence structure make this method difficult to apply in this situation. As an alternative to multivariate regression, Crewett *et al.* (1999, 2000) and Lamb *et al*. (2000a,b) adopted a sequential regression approach, designed to account for parameter dependence with very few restrictions. In this approach the model parameter prediction equations are derived sequentially, accounting for the effect that generalisation of earlier parameters has on later parameters. For the sample of around 40 catchments, sequential regression was found to perform better than univariate regression (Calver *et al.* 2001).

The exact form of sequential regression used here has been adapted from that used in FD0404, and is outlined in the flow chart in Figure 5.2. The new developments of the method reflect the alignment of the sequential regression method with the automatic calibration method developed during this project (see Section 4.2.1). That is, the sequential regression method now involves two-passes rather than one, with each parameter being calibrated and then generalised once in each pass. For each model, the ordering of the parameters in sequential regression is the same as that used for automatic calibration, as is the choice of objective functions on which to calibrate each parameter (see Table 4.1). As in automatic calibration, a final calibration and generalisation is then performed for the first parameter for each model (the parameter controlling water balance), to allow for a final adjustment of the flood frequency curve.

Due to the intensive nature of the exhaustive search procedure used for univariate regression, it was considered impractical to attempt this same search procedure at each step of the sequential regression. Instead, the combination of catchment properties that is chosen for each model parameter during univariate regression is retained for sequential regression, but with the coefficients of those catchment properties determined separately, within the sequential regression procedure. Assuming that there are real relationships between model parameter values and catchment properties, then this simplification will not be too restrictive. The sequential regression method is thus not considered again until Section 5.6.

5.2.3 Site-similarity

In addition to the regression approach, the use of site-similarity approaches for spatial generalisation of the model parameters has been investigated. The idea behind site-similarity has evolved from regional flood frequency analysis and was presented as the region-of-influence approach by Burn (1990) based on work by Acreman and Wiltshire (1987, 1989). Burn and Boorman (1992) tested a number of site-similarity type approaches for estimating model parameters in the runoff models developed in the Flood Studies Report (FSR) (NERC 1975) and found the best of these methods to

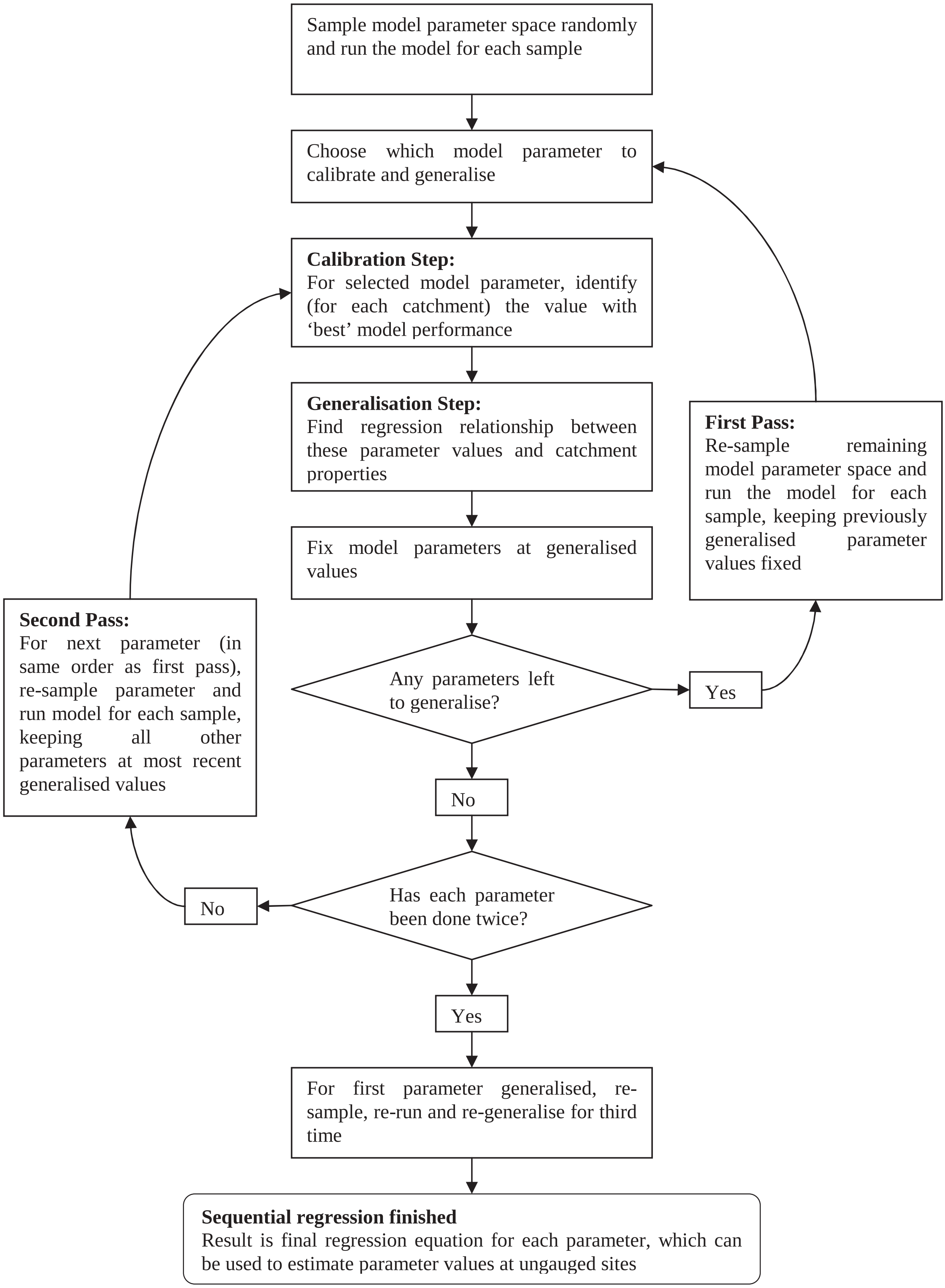


Figure 5.2 Outline of the expanded sequential regression method

outperform the linear regression approach originally suggested in FSR. The site-similarity approach is an important part of the statistical method in the Flood Estimation Handbook (Robson and Reed 1999) and in Low Flows 2000 (Holmes *et al.* 2002).

As for the ordinary regression approach (Section 5.2.1), each of the runoff model parameters are treated separately, and different choices of details within the method are made for each parameter. In a site-similarity approach, a given model parameter for a candidate site can be estimated via the following steps:

1. Obtain the relevant catchment properties at the ungauged site under consideration. The definition of relevant catchment properties is a critical factor in the performance of this approach, just as it is for the regression approach.

2. Identify the distance to all available gauged catchments. The definition of distance is based on Euclidean distance in a given space of catchment properties and derived as

$$dist_{a,b} = \sqrt{\sum_{j=1}^{J} \lambda_j \left(\frac{X_{a,j} - X_{b,j}}{\sigma_{X,j}} \right)^2} \,, \tag{5.2}$$

   where J is the number of catchment properties, j indicates a particular one of the catchment properties, $X_{a,j}$ is the value of that catchment property at the ath site, $\sigma_{X,j,}$ is the standard deviation of the property across all the sites, and λ_j is the coefficient assigned to the particular catchment property. The catchment properties can be transformed (for example using natural logarithms) or untransformed. (Note that the λ_j have in fact been set to 1 throughout, once a choice of the catchment properties to be used for the particular runoff model parameter has been made).

3. The M closest neighbours (minimum distance) are selected to create a pooling group for the candidate site, consisting of the calibrated parameters for the selected catchments $\{\alpha_m ; m = 1, \ldots M\}$. It should be noted that the pooling groups used for a given target catchment can be different for each model parameter.

4. Having formed the pooling group, the estimate α_a^{PG} of the model parameter at target a site is calculated as a weighted average of the corresponding parameters from the sites in the pooling group. That is,

$$\alpha_a^{PG} = \frac{\sum_{m=1}^{M} h_m \alpha_m}{\sum_{m=1}^{M} h_m} \,. \tag{5.3}$$

   Here the superscript PG refers to an estimate obtained using a pooling group. The weight h_m is assigned to the mth gauged catchment in the pooling proup to reflect its importance.

Initial work on site-similarity approaches (Kjeldsen *et al.* 2003) applied the following weighting schemes in Equation (5.3):

- Equal weights: $h_m = 1,$
- Linearly decreasing weights: $h_m = 1 - dist / dist_{\max}$,
- Quadratically decreasing weights: $h_m = 1 - \left(dist / dist_{\max}\right)^2$,

where $dist_{\max}$ is set to be 10% larger than the maximum distance of a pooling group member from the target site being considered. The results suggested that the use of linearly decreasing weights led to a marginal improvement in performance, although this factor was much less important than others. In later work (Kay *et al.* 2004; Jones *et al.* 2004) a slightly different form of weighting was developed and applied, which is able to incorporate both distance-weighting and weighting derived from calibration uncertainty (uncertainty weighting). That is,

$$h_m = \frac{1 - s}{1 + k\sigma_{m,\varepsilon}^2}$$

where s can be zero (equal distance weighting), $dist/dist_{\max}$ (linearly decreasing weights) or $(dist/dist_{\max})^2$ (quadratically decreasing weights) as before, and where the $\sigma_{m,\varepsilon}{}^2$ and k are as described for weighted regression (Section 5.2.1). As for regression, k can be estimated iteratively and, on convergence, $1/k$ gives an estimate of the generalisation uncertainty for each parameter. Alternatively, setting k=0 means that only distance-weighting is used, and performing the first step of the iterative procedure gives an estimate of the generalisation uncertainty in this case. The full theory of this form of uncertainty-weighting is described in Chapter 6 (Section 6.4.3) for both the regression and site-similarity approaches.

The estimated parameter for a target catchment i could be derived by either including or excluding the information we have from calibration for catchment i. The results from the regression approach would generally include the information from each catchment in its own estimate and, as a parallel to this, the main analyses here for the site-similarity approach have allowed each catchment to be included in its own pooling group. When individual parameters of the runoff models are being considered, the formal statistical measures of the performance of the generalisation approaches do take into account this decision in order for the measures to properly estimate the performance for catchments not in the calibration set. However, there is a need for some more informal assessments of the generalisation performance, particularly where attention switches from the model parameters to the flood frequency curves using the generalised parameters. In the latter case a more formal approach is forestalled because of the un-met need for the theoretical adjustments necessary to account for the inclusion or exclusion of target catchments. The inclusion of the target catchment within a pooling group can lead to an over-optimistic assessment, to an extent that is much larger for the site-similarity approach than for regression methods. Hence some of the informal assessments of site-similarity have been undertaken using two versions of result sets which either include or exclude each target site from its pooling group.

5.3 Data and transformation

To try to increase the potential to gain information about parameter values from catchment properties, some of the parameters and some of the catchment properties are transformed before use in one or more of the generalisation procedures. Calculations are then performed with the transformed parameters/properties. For parameters that are transformed the calibration uncertainty should be calculated with transformed parameter values, rather than untransformed values.

The transformations are generally chosen to make the distribution of the relevant quantity less skewed. An additional benefit of using a logarithmic transformation on certain model parameter values is that it ensures no negative values can be obtained from the regression generalisation equations. The transformations used in this report for the model parameters are given in Table 5.1 and those for the catchment properties were given in Chapter 2, Table 2.4. The effect of the catchment property transformations needs to be borne in mind when considering the descriptive ability of the catchment properties in predictive equations. When using transformations in conjunction with regression approaches it is important to check that relationships between transformed variables are approximately linear: graphical checks have been made for this project.

Table 5.1 Transformations of the model parameters

PDM parameter	Transformation	TATE parameter	Transformation
f_c	-	*crm*	-
$c_{\max}$	-	*csm*	Log
k_1	Log	*cfr*	Log
k_b	Log		

Note that certain catchments are excluded from the generalisation (at least initially, see later), either for all parameters of a model or just for certain parameters: A small number of catchments are excluded from the generalisation for all parameters of a model, because the model does not calibrate satisfactorily for the catchment, so the calibrated parameter values are unlikely to contain any information useful for generalisation. For other catchments, the overall performance of calibration is good but one or more parameters calibrates right at the edge of its allowed range. This would not, in itself, necessarily be a problem. However, if all the jack-knife parameter values also calibrate on the edge of the allowed range then the calibration uncertainty for that parameter for that catchment could be unrealistically low, leading to a high weight in the generalisation. This could skew results towards those edges, and so catchments with this problem for a parameter were (initially) left out of the generalisation process for that parameter, for all generalisation methods.

5.4 Performance criteria

The choice of catchment properties on which to base either approach is difficult. The desired output in each case, for each parameter, is a limited number of catchment properties which provide a good estimate of the model parameters and which give a hydrologically sound model. The final choice of catchment properties, together with other aspects of the method, emerges from an exploratory statistical process underpinned by hydrological judgement. At each stage, exhaustive search techniques are used to ensure that no useful model is omitted from the analysis. To assess the performance of the approaches, three different criteria have been used at different stages of the work.

The initial study of site-similarity approaches to generalisation (Kjeldsen *et al.* 2003) used the performance indicator S_0, which was the (unweighted) leave-one-out cross validation measure defined by

$$S_0 = \sqrt{n^{-1}\sum_{i=1}^{n}(y_i - \hat{y}_{(i)})^2} \tag{5.4}$$

where

n = number of catchments,
y_i = calibrated parameter estimate for catchment i,
$\hat{y}_{(i)}$ = pooling group estimate for catchment i, treating catchment i as missing.

The second criterion, σ_η^2, derives from the weighted sum of squared residuals

$$S^2 = \sum_{i=1}^{n} w_i (y_i - \hat{y}_i)^2$$

where

n = number of catchments,
y_i = calibrated parameter estimate for catchment i,
$\hat{y}_i$ = "*generalised*" estimate for catchment i, and
$w_i = 1/(1 + k\sigma_{i,\varepsilon}^2)$.

Here the generalised estimate $\hat{y}_i$ is either the weighted regression estimate α_i^R in Equation (5.1) or the pooling-group estimate α_i^{PG} in Equation (5.3): in both cases the estimate may depend on the uncertainty-weighting constant, k. The value of S^2 is used to construct an estimate of the variance of the generalisation error

$$\sigma_\eta^2 = \frac{S^2}{n^*} \tag{5.5}$$

Here n^* is a theoretically-derived quantity, related to how the generalised estimate is constructed, which adjusts the number of catchments n for the number of internal parameters contained within the generalisation procedure: see Chapter 6 (Section 6.3.3, Equation (6.4.3.10)) for details. When the full iterative procedure with uncertainty-weighting is used, the weighting constant k is reset at the end of each iteration to

$k=1/\sigma_\eta^2$. In this case the value of σ_η^2 at convergence is reported. If uncertainty weighting is not used, σ_η^2 is calculated for $k=0$, corresponding to an unweighted estimation and an estimated variance derived from an unweighted sum of squared residuals.

A previous report (Kay *et al.* 2004) concentrated on comparisons using σ_η^2. This is a measure which reflects the underlying potential of a given generalisation procedure. Specifically, it measures how well the procedure is likely to do if problems arising from data deficiencies can be eliminated in future. In the case of regression methods, σ_η^2 is a measure of the remaining error given that the regression coefficients become effectively 'perfectly' known as the number of catchments increases. For site-similarity, σ_η^2 is a measure of the remaining error given that the pooling group method is applied in such a way that the number of catchments in a pooling group increases as the overall number of catchments increases, but also so that the size of the region covered by the group (in catchment property space) decreases as further catchments are added. Where uncertainty weighting is not used in constructing the criterion, σ_η^2 includes a contribution from the calibration uncertainty since this cannot then be treated separately. In this case σ_η^2 cannot allow for improvements in future data-sets arising from longer data-records and better defined parameters estimates from the calibration procedures.

The third performance criterion is closely related to σ_η^2 and aims to encapsulate all the uncertainty in the estimated value produced by the generalisation procedure: the generalisation variance is only part of this. In the case of regression estimates, σ_η^2 relates to the error about an unknown regression line and the overall procedure needs to account for the uncertainty in estimating the position of that line. When site-similarity estimates are used, σ_η^2 relates to variations about an unknown local mean value for the pooling-group and the procedure needs to account for errors in estimating that local mean. For any given target catchment for which no calibrated estimates of the model parameters are available, a generalised estimate $\hat{y}_T$ can be found for each model parameter. The theory for the particular type of estimate then provides an expression for the overall uncertainty of the estimate:

$$\mathrm{var}(\hat{y}_T) = \sigma_\eta^2 \{1 + f_T\}.$$

Here f_T is a term for which theoretical expression are available: see Chapter 6 (Section 6.4.4; Equation (6.4.4.4)) for further details. The variance derived in this way corresponds to the amount of uncertainty that needs to be added to the generalised estimate when used in a continuous simulation procedure. In the present context, the presence of the factor f_T is inconvenient since the value varies with the particular target catchment chosen. When regression estimates are used, the value is larger or smaller depending on whether the target catchment is further from or closer to the centre of the catchments used for fitting the regression (when judged according to the catchment properties used). For site-similarity, f_T depends on the number of catchments in the pooling group and, when a distance-based pooling is used, on the closeness of the catchments in the pooling group.

It is possible to convert the estimation variance into a useful performance measure by treating each of the catchments used for establishing the generalisation procedure as if it did not have data for calibration, and then averaging across these catchments:

$$\sigma_{TA}^2 = \sigma_\eta^2 \; n^{-1} \sum_{i=1}^{n} \left\{ 1 + f_i \right\} \tag{5.6}$$

Note that lower values of $S_0^{\,2}$, $\sigma_\eta^{\,2}$ and $\sigma_{TA}^{\,2}$ indicate better performance, but that these values are parameter-specific in that their magnitudes are related to those of the model parameters. This means that they cannot be used to compare performance across different model parameters, just for the same model parameter across different generalisation approaches. While a convenient scaling is to divide each of these measures by the sample variance across catchments of the calibrated parameter values, there are weighted and unweighted versions of this sample variance to consider and a different choice might naturally be made according to whether or not the generalisation procedure used a weighting scheme. Employing different scalings would not allow a good comparison to be made of the different generalisation approaches for the same model parameter.

Once the more specific choices within generalisation methods have been made according to the above performance criteria, the final comparison of the performance between different methods also takes into account the fit of generalised flood frequency curves. Whilst plainly the level of performance of the approaches is a major factor in reaching a preferred option, it is also instructive to consider the advantages and disadvantages of the methods themselves from the point of view of the user. These are summarised in Table 5.2.

5.5 Decisions within generalisation procedures

The investigation into the preferred method for spatial generalisation of the runoff model parameters is divided into three main parts. Parts one and two are concerned with the univariate regression and site-similarity methods respectively, and examine the optimal structure of each method (weightings etc.). Part three looks at the choice of catchment properties used for the generalisation of each parameter within each method.

To choose the best versions of the univariate regression and site-similarity generalisation methods, a comprehensive search through many possible combinations of catchment properties was conducted at each step in the investigation. The search routine provides a list of combinations of catchment properties ranked according to the chosen performance measure (Section 5.4). Although a total of 24 catchment properties was available (Section 2.6), for practical reasons the maximum number of catchment properties allowed in any one combination was set to six.

This section compares the performance using different choices within the univariate regression and site-similarity methods, mainly in terms of plots of performance measures against number of catchment properties, for the best three combinations of one to six catchment properties. Also included is the performance using no catchment properties, that is, where the parameter value for a target catchment is estimated using a simple (weighted or un-weighted) average of the calibrated parameter values of each of the other catchments.

Table 5.2 Advantages and disadvantages of the regression and site-similarity approaches to spatial generalisation.

Spatial generalisation method	Advantages	Disadvantages
Univariate regression	• Ease of use (simple application of equation) • Possible use of predictive equation to extrapolate (with care) beyond the range of properties in the calibration set. • Theory for inclusion of uncertainty has been developed in this project	• Often pragmatic, rather than explanatory, choice of catchment properties in predictive equations • Updating with new data, when new catchments are calibrated, requires re-establishment of predictive equations (and testing)
Sequential regression	As first two points for univariate regression, plus: • Sequential process allows for inter-relationship of parameters	As for univariate regression, plus: • More complex, time-consuming procedure, if updating is required • Non-standard theory for inclusion of uncertainty
Site-similarity	• Flexibility of structure, in that an explicit relation of parameters to catchment properties is not required • Flexibility at application stage (sites can be left out of pooling group if not considered reliable) • Easy to include new data, when new catchments are calibrated • Theory for inclusion of uncertainty has been developed in this project	• Large number of possible choices in method • Explanation is implicit only • Application requires access to parameter sets and catchment properties for calibrated catchments • Application should be restricted to catchments not too different from those in the set used for generalisation • Pooling groups may be difficult to establish for 'unusual' situations.

The section also contains an initial comparison between the 'optimal' regression and site-similarity methods (Section 5.5.3), again based on plots of performance against number of catchment properties, for the best three combinations of one to six catchment properties. However, this simple, parameter-by-parameter comparison proves insufficient to identify the best overall method of generalisation, and so further comparisons based on generalised flood frequency curves are required (Section 5.6). This is undertaken following the choice of catchment properties for use in each generalisation method for each parameter (Section 5.5.4).

5.5.1 Univariate regression

Multiple univariate linear regression, as described in Section 5.1.1, involves fewer subjective choices than the site-similarity approach. The two main decisions to be made are which combination of catchment properties to use and how to weight each individual catchment in the weighted least-squares parameter estimation.

Figure 5.3 and Figure 5.4 show, for the PDM and TATE models respectively, that use of weighted regression (using weights estimated iteratively, as described in Section 5.1.1 and Chapter 6) is an improvement over un-weighted regression.

5.5.2 Site-similarity

A number of factors can influence the site-similarity approach, but the main ones are

- which catchment properties to use,
- the number of catchments in the pooling group, and
- the weights used for deriving average model parameters.

Initial work developing the site-similarity generalisation technique (Kjeldsen *et al*. 2003) suggested that, although the size of the pooling group was not as important as other factors, the use of a pooling group with around 10 members was preferable to either a much larger or smaller pooling group (Table 5.3). Therefore a size of 10 was selected and used in subsequent development of the technique.

As for the regression approach, investigations (Kay *et al.* 2004) showed that the use of uncertainty weighting within the site-similarity approach gave an improvement over the original weights based only on distance as used in Kjeldsen *et al.* (2003) (e.g. Figure 5.5 for the PDM and Figure 5.6 for TATE). This work also supported a previous conclusion that the effect of using different distance weightings (equal, linear or quadratic) was relatively minor in comparison. This conclusion is confirmed here, using an alternative performance measure to that used previously (Figure 5.7 for the PDM and Figure 5.8 for TATE). It was decided to use linear distance weighting, as the use of weights which decrease according to distance from the target site was considered appropriate, and linear distance weighting generally performs very similarly to, or better than, quadratic distance weighting.

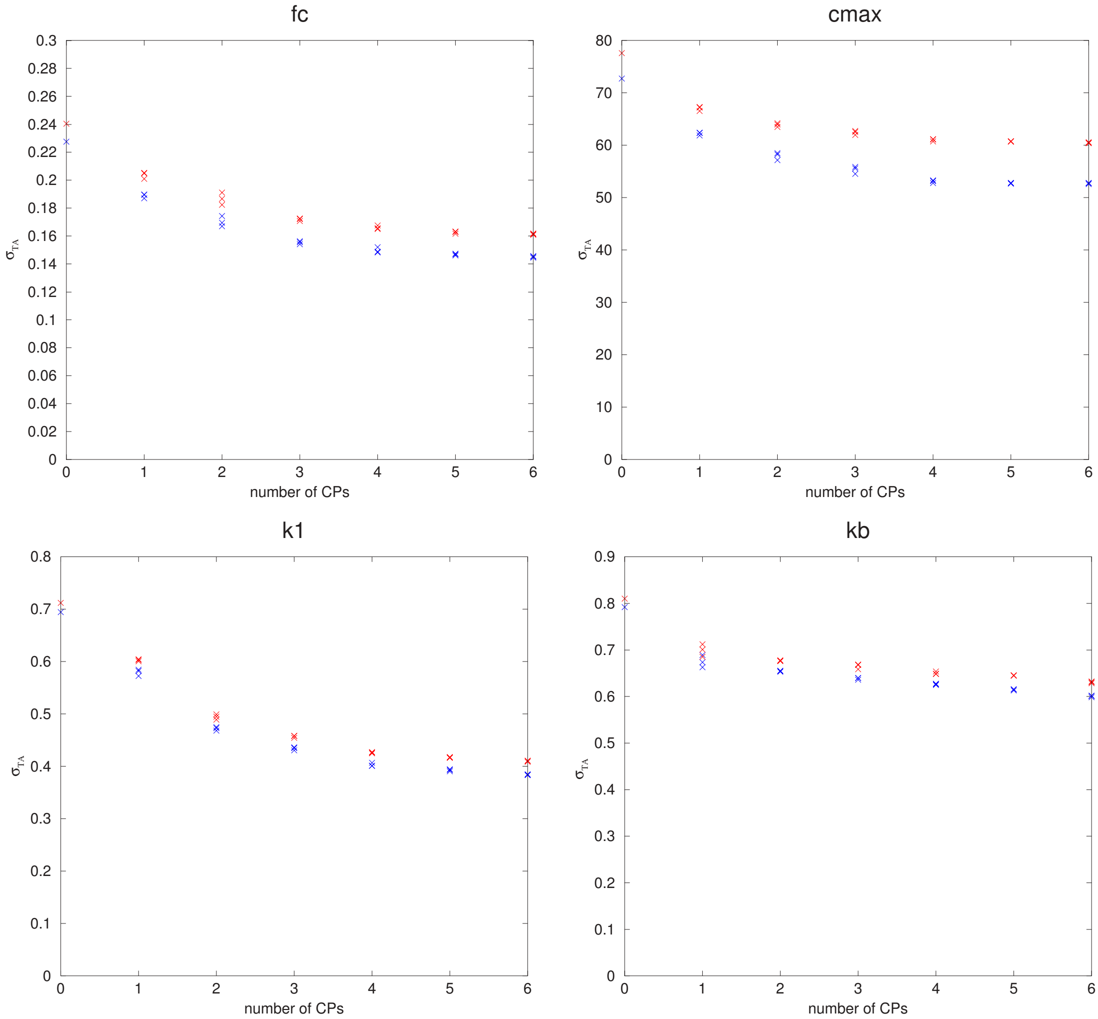


Figure 5.3 Comparison of weighted regression (blue crosses) with un-weighted regression (red crosses) for the four PDM parameters (f_c, c_{max}, k_1 and k_b). The performance measure σ_{TA} is plotted against the number of catchment properties used in the regression equation, for the top three combinations of catchment properties (CPs) in each case.

5.5.3 Initial comparison of regression and site-similarity performance

Figure 5.9 and Figure 5.10 compare the performance of site-similarity and univariate regression, for the PDM and TATE parameters respectively, in terms of the performance measure σ_{TA}. They illustrate that there is no single preferred method, with one method being better for some parameters and the other method being better for some other parameters. It is necessary, therefore, to carry each method forward to a choice of catchment properties for the generalisation of each parameter, in order to compare performance in terms of generalised flood frequency curves.

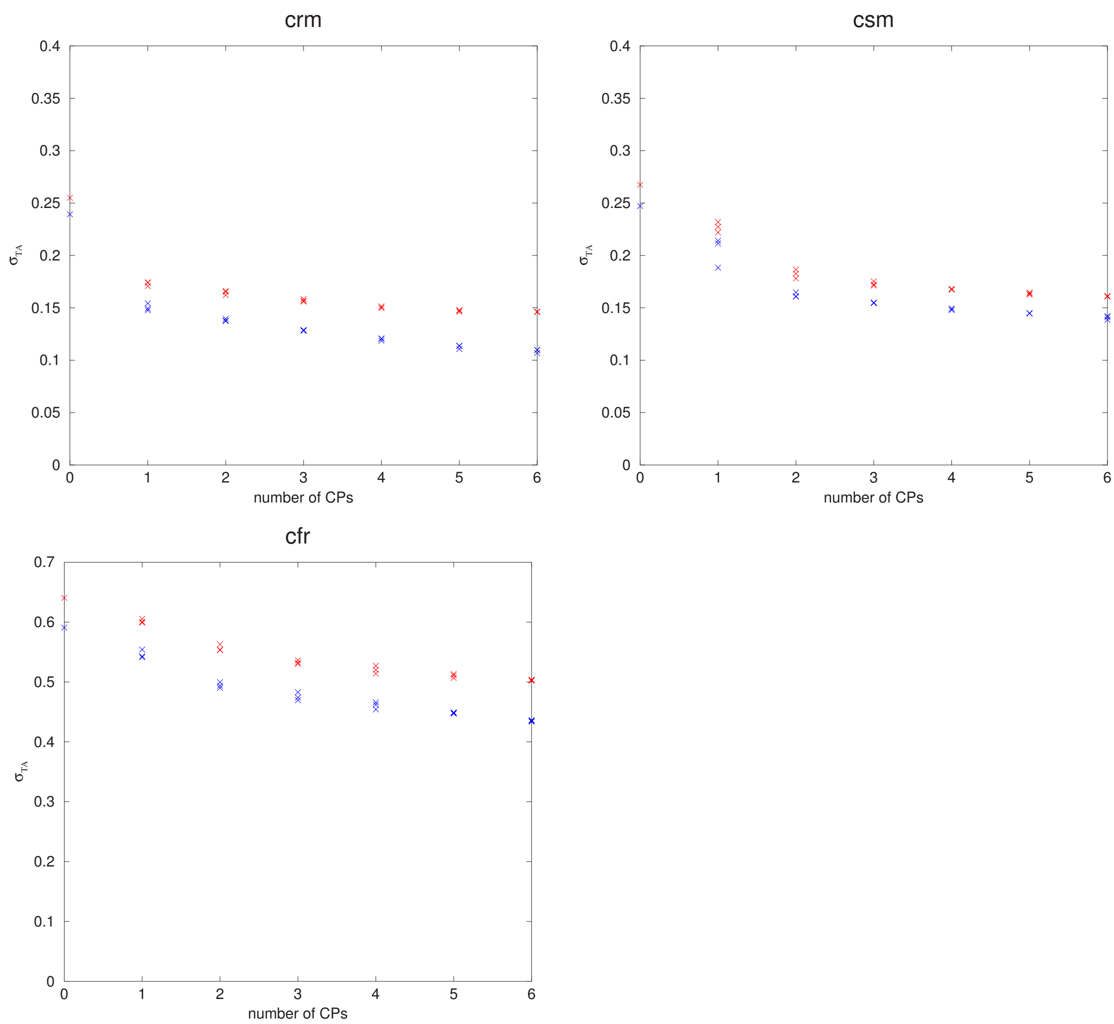


Figure 5.4 As Figure 5.3, but for the three TATE parameters (*crm, csm* and *cfr*).

5.5.4 Choice of catchment properties for each parameter

As stated in Section 5.5.1, a comprehensive search routine was used to provide lists of combinations of catchment properties ranked according to chosen performance measures. The combinations of catchment properties were either those used in univariate regression equations or those used to define the pooling group in the site-similarity method. Lists were produced for the best 10 combinations for each parameter within each model for three to six catchment properties (Appendix B.1). A decision was then required as to which combination, and number, of properties to select as performance criteria were not the sole factors determining the final choice. Other main considerations were the hydrological relevance of the catchment property to the function of the parameter in the runoff model and, if possible, choice of a combination of properties that constituted a cohesive group avoiding correlated parameters (Section 2.6.2). An additional consideration for regression methods was that the direction of change of a catchment property on a parameter was as expected hydrologically. These directions of change are also given in the lists in Appendix B.1.

Table 5.3 Performance measure S_0 as a function of the number of gauged catchments in a pooling group (based in this case on BFIHOST and PROPWET). The numbers in blue indicate the best performance (lowest value of S_0) for each model parameter, and those in red indicate a performance more than 5% greater (worse) than the best one.

Number of catchments	PDM				TATE		
	f_c	c_{max}	k_1	k_b	*crm*	*csm*	*cfr*
All	0.2427	87.5	16.35	132.31	0.2769	0.0693	0.2327
50	0.2160	78.7	13.61	113.29	0.1925	0.0612	0.2289
25	0.1902	76.9	12.82	108.43	0.1856	0.0557	0.2268
15	0.1806	77.2	12.46	108.78	0.1826	0.0542	0.2358
12	0.1775	75.3	12.53	111.14	0.1857	0.0540	0.2376
10	0.1775	74.2	12.58	111.49	0.1858	0.0542	0.2404
8	0.1799	74.0	12.52	110.06	0.1859	0.0533	0.2467
5	0.1918	76.6	12.95	113.15	0.1960	0.0537	0.2553
3	0.2048	80.2	14.48	119.68	0.2124	0.0605	0.2643
2	0.2108	82.0	15.14	125.75	0.2315	0.0605	0.2788
1	0.2481	105.6	16.73	151.33	0.2651	0.0603	0.3297

A guide to determining the number of catchment properties was provided by the plots of number of properties against performance (Figures 5.3 to 5.8). For most parameters there was a marked improvement in performance for both regression and site-similarity methods using between one and three properties, with a smaller, or negligible, improvement as the number increased from three to six. A minimum of four properties was used as an initial guideline, with patterns of dominant properties and stable groupings identified for further guidance. The former are those which appear in almost all equations or site-similarity groups while the latter are combinations of properties which pass on through the lists for each number of properties but with an additional property as the number is increased. The inclusion of a comprehensive band of catchment properties within the groups for each parameter for each model was also considered desirable, with the inclusion of the catchment properties URBEXT and FARL highly desirable because of their particular impact on flood peaks, URBEXT with the potential to increase peaks and FARL to decrease peaks due to the attenuating effect of storage from reservoirs and lakes. However, because they are not present in many catchments a particular requirement was made to include these properties for relevant parameters. Suitable groupings were in all cases found from within the top 10 combinations defined by performance (Appendix B.1).

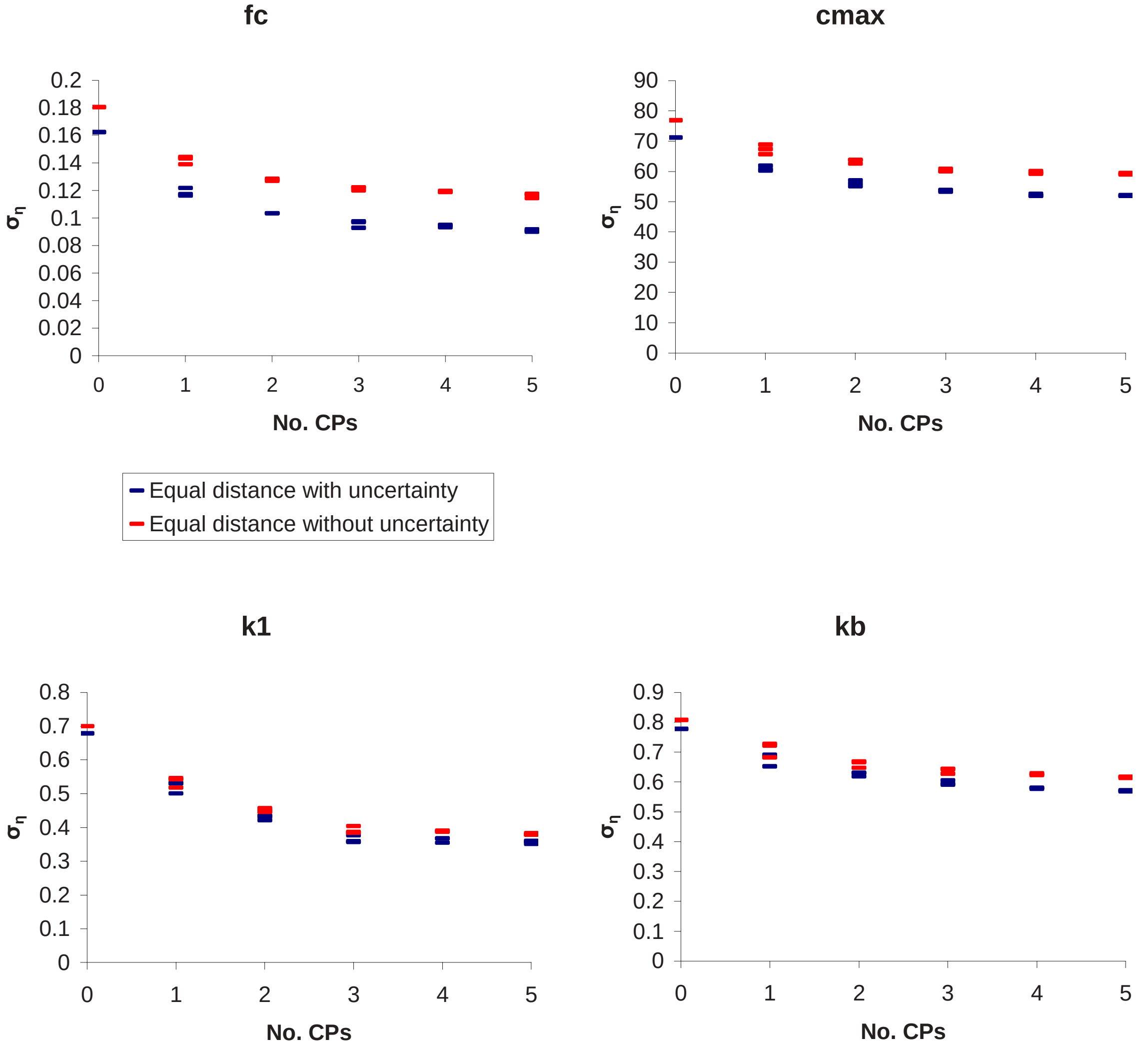


Figure 5.5 Comparison of site-similarity generalisation with and without uncertainty weighting (with equal distance weighting) for the four PDM parameters.

Plots of calibrated against generalised parameter values for possible property combinations were inspected to identify if there were particular parameter ranges which were not well generalised. In general, there was little overall difference between the plots for different combinations. However, the plots did inform consideration of the generation of parameter values for those catchments where the calibrated values were close to the edge of the theoretical, or allowable, domain. Initially, these catchments, as well as those with poor calibrations, were excluded from the analyses for both the univariate regression equations and site-similarity groups to ensure that such calibrations did not impact on the final equations or property groups. This is a reasonable restriction in the determination of regression equations, as parameter values

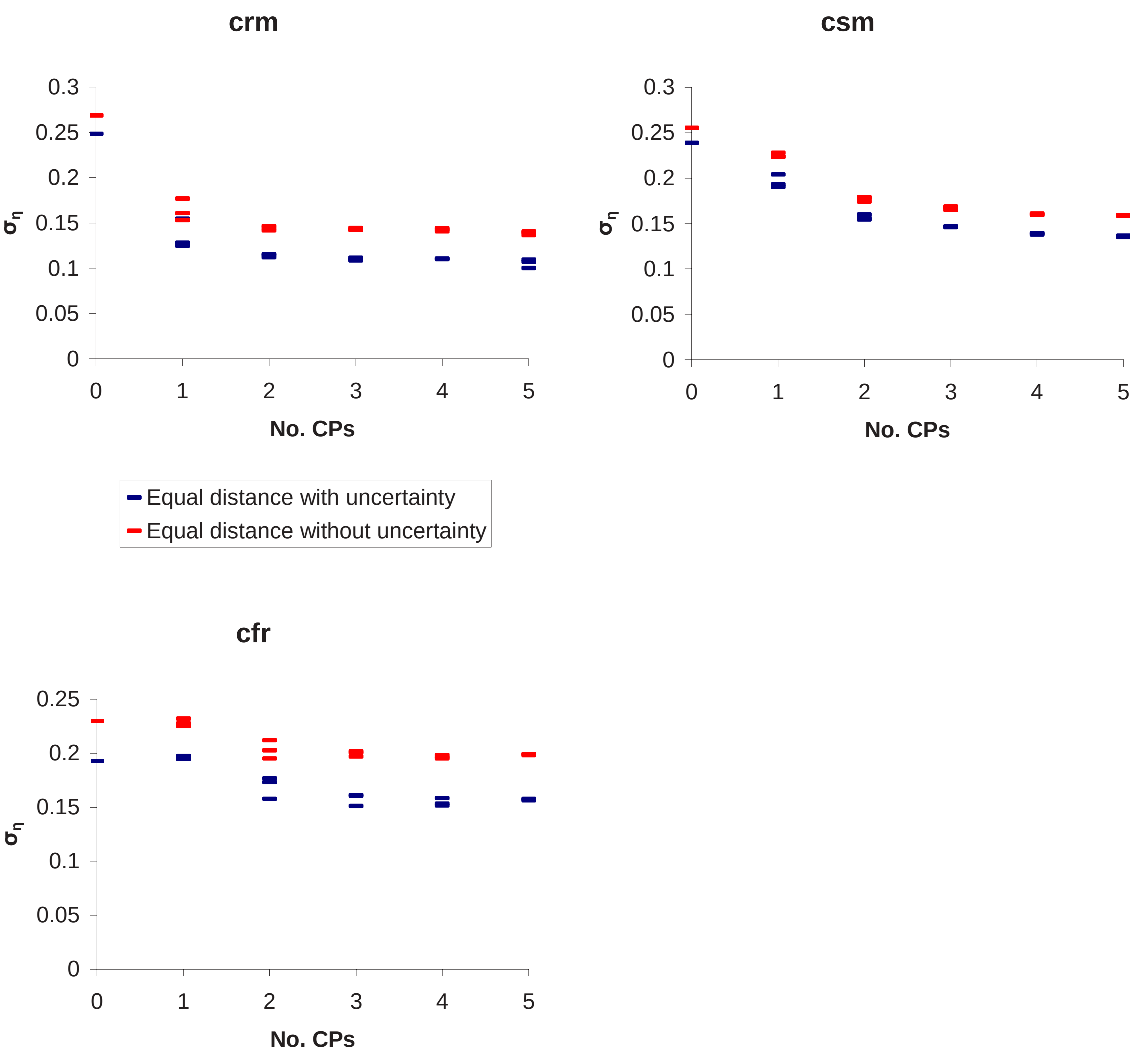


Figure 5.6 Comparison of site-similarity generalisation with and without uncertainty weighting (with equal distance weighting) for the three TATE parameters.

beyond the range of the sample catchments are generated by extrapolation. However, for site-similarity, where generated parameter values are averages, the exclusion prevents generation of values outside those of the sample catchments. Accordingly, a further list of catchment property combinations was generated for the site-similarity method including catchments with 'boundary' parameter values but still excluding those with poor calibrations. This resulted in improved performance for the PDM but not for the TATE.

The catchment properties selected for each model parameter using the univariate regression and site-similarity generalisation methods are given in Table 5.4. The actual univariate regression equations for each parameter are given in Table 5.5. The site-

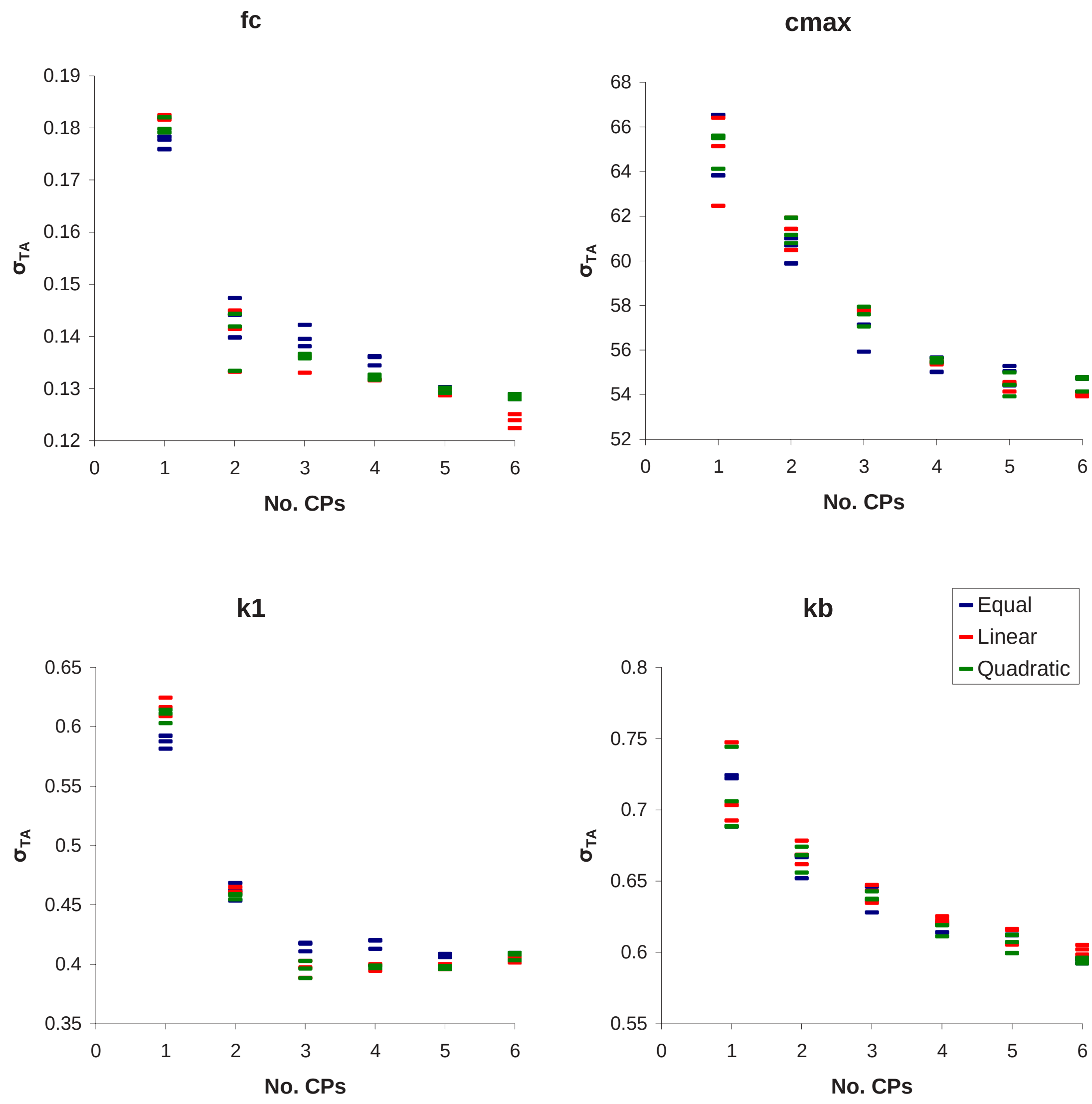


Figure 5.7 Comparison of site-similarity generalisation with equal, linear and quadratic distance weighting (with uncertainty weighting) for the four PDM parameters.

similarity groups for the PDM were determined from lists for catchments including boundary values while those for the TATE exclude boundary values. Five catchment properties were chosen for most parameters, but six properties were selected where performance was noticeably improved or where it was felt desirable to include particular properties.

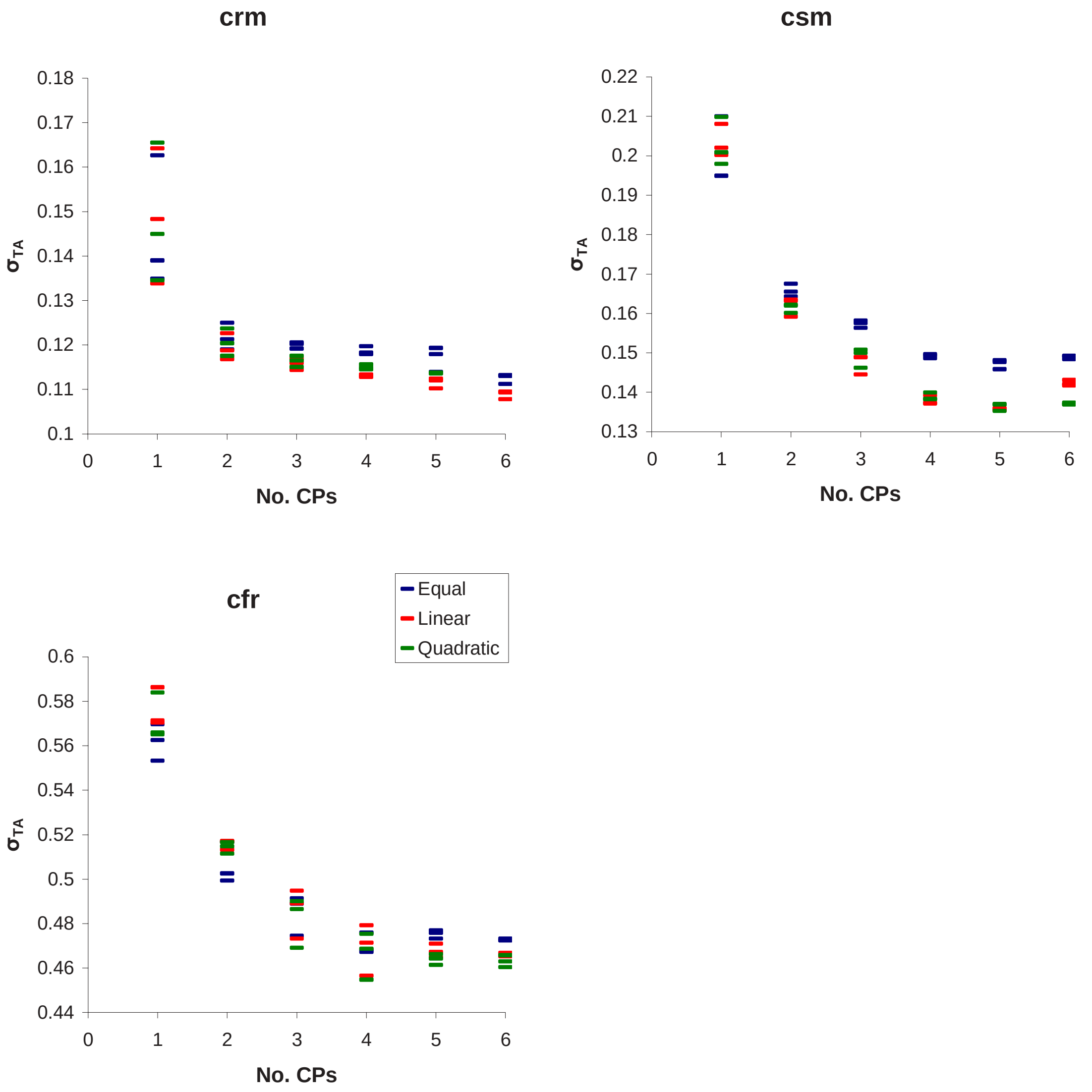


Figure 5.8 **Comparison of site-similarity generalisation with equal, linear and quadratic distance weighting (with uncertainty weighting) for the three TATE parameters.**

Table 5.4 highlights catchment properties that are used within each generalisation method for the parameter, or those where closely correlated properties are used for each method, and shows a high degree of similarity in catchment properties between the methods. Although some of these properties were specifically chosen (that is, combinations including them were considered necessary for particular parameters), others were not, and all still had to appear in the top 10 catchment property combinations: none was ‘forced’ in. This similarity in catchment properties between the methods suggests that they are making use of real relationships between parameters and catchment properties.

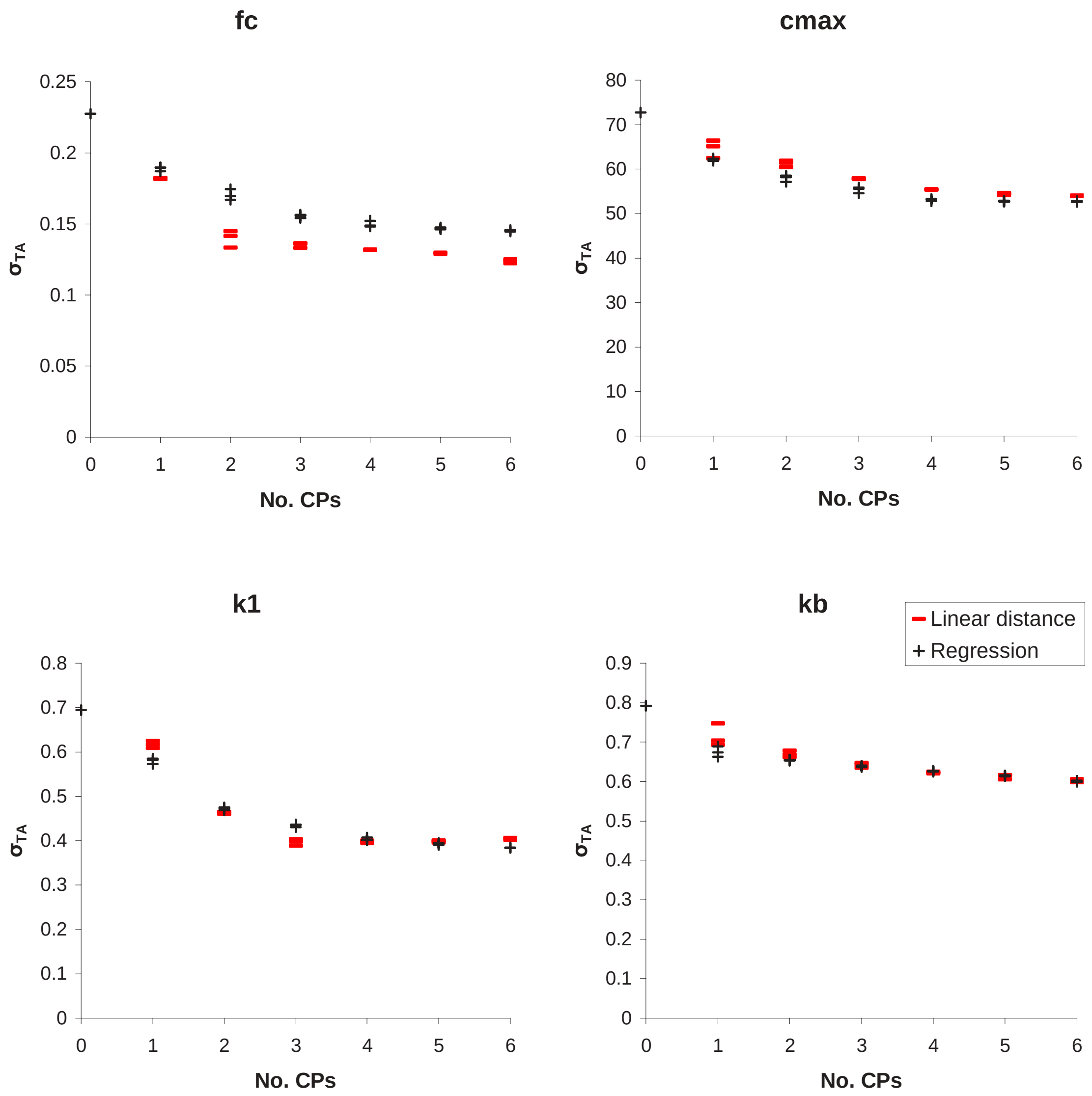


Figure 5.9 Comparison of the performance of site-similarity (with linear-distance and uncertainty weighting) and univariate regression (with uncertainty weighting), for the four PDM parameters.

Table 5.4 includes values for the R^2 measure of fit of the generalised estimates to the calibrated parameter values. These R^2 values are calculated in the basic unweighted form to avoid difficulties in making comparisons when the weighted form uses different weights. In the case of the site-similarity method, two values of R^2 have been calculated: for the first the generalised estimate for a target catchment has been calculated including the target site in the pooling group while, for the second, the target site is excluded. It

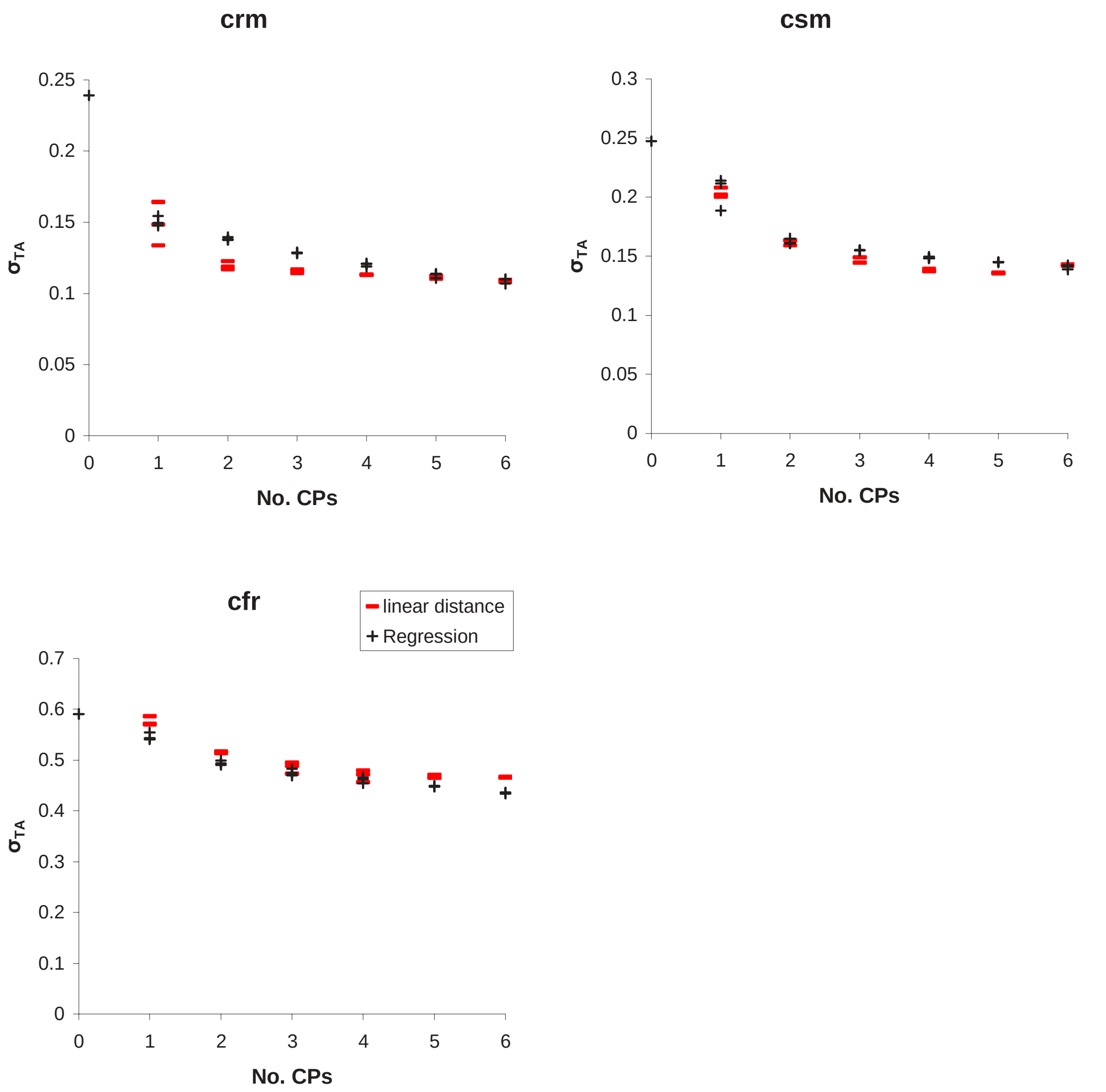


Figure 5.10 Comparison of the performance of site-similarity (with linear-distance and uncertainty weighting) and univariate regression (with uncertainty weighting), for the three TATE parameters.

can be argued that the first type of R^2 is unrepresentative of the actual performance since high weight is given to the target site whereas in practice, for an ungauged catchment, calibrated parameters for the target site are not available. Values for the second form of R^2 for the site-similarity method are of a similar size to those obtained for the univariate regression method. However, the large difference between including and excluding the target site for site-similarity suggests that strong caution is needed when judging the performance of the generalisation procedure in terms of reproducing the flood frequency curves.

Table 5.4 The catchment properties used for each model parameter in each generalisation method (‘t’ indicates properties that are transformed according to Table 2.4). For each parameter, properties in red are used for both generalisation methods , while those in blue are ones where correlated properties (Section 2.6.2) are used in the other generalisation method.

Method	PDM				TATE		
	f_c	c_{max}	k_1	k_b	*crm*	*csm*	*cfr*
Univariate regression	tDPSBAR	PROPWET	tALTBAR	BFIHOST	tFARL	BFIHOST	tAREA
	tHOSTGMIN	SPRHOST	DPLBAR	DPLBAR	tSAAR	tFARL	tFARL
	tHOSTPEAT	tURBEXT	tFARL	tDPSBAR	SPRHOST	tSAAR	PROPWET
	HOSTNG	HOSTNG	PROPWET	tURBEXT	tLANDA	tURBEXT	SPRHOST
	tLANDA	FIELDC	SPRHOST	tHOSTPEAT	tLANDB	DRAIN2	tURBEXT
			tURBEXT	DRAIN2			PORO
R^2	0.51	0.46	0.68	0.51	0.69	0.65	0.39
Site-similarity	tAREA	PROPWET	BFIHOST	BFIHOST	DPLCV	tALTBAR	DPLBAR
	BFIHOST	tURBEXT	DPLBAR	tURBEXT	tFARL	BFIHOST	tALTBAR
	tDPSBAR	HOSTNG	tFARL	HYDC	PROPWET	tDPSBAR	BFIHOST
	tSAAR	tHOSTP	tURBEXT	tLANDC	tSAAR	PORO	tHOSTGMIN
	tLANDB	tLANDA	tLANDB	DRAIN2	SPRHOST	DRAIN2	tURBEXT
	tLANDC						
R^2 (inc. target)	0.78	0.70	0.82	0.74	0.74	0.76	0.35
R^2 (exc. target)	0.55	0.41	0.66	0.50	0.64	0.63	0.21

Table 5.5 Multiple univariate regression equations

Model	Parameter		Regression equation	
PDM	f_c	=	-0.241	+0.021√ DPSBAR +0.668√ (HOSTGMIN/100) +0.919√ (HOSTPEAT/100) +0.0093 HOSTNG +0.217√ (LANDA/100)
	c_{max}	=	-70.46	-231.1 PROPWET -2.588 SPRHOST -270.3√ URBEXT +0.399 HOSTNG 11.62 FIELDC
	Log(k_1)	=	4.270	-0.049√ ALTBAR +0.023 DPLBAR +1.479√ (1-FARL) -1.595 PROPWET -0.016 SPRHOST -2.423√ URBEXT
	Log(k_b)	=	3.237	+2.154 BFIHOST +0.015 DPLBAR +0.085√ DPSBAR +1.852√ URBEXT +0.986√ (HOSTPEAT/100) -0.845 DRAIN2
TATE	*crm*	=	1.381	+0.457√ (1-FARL) -0.018√ SAAR -0.0077 SPRHOST -0.412√ (LANDA/100) -0.331√ (LANDB/100)
	Log(*csm*)	=	-2.695	+0.987 BFIHOST +0.440√ (1-FARL) +0.023√ SAAR -0.466√ URBEXT -0.160 DRAIN2
	Log(*cfr*)	=	0.915	-0.042√ AREA -2.009√ (1-FARL) +2.263 PROPWET +0.031 SPRHOST +1.376√ URBEXT -0.074 PORO

5.6 Results

Having decided on the structure of the regression and site-similarity methods, and on the combinations of catchment properties to be used for the each parameter within each method, the sets of generalised parameters for each catchment can be used to simulate time series of river flows and thus produce generalised flood frequency curves. The generalised flood frequency curves are created by using the observed rainfall series for a given catchment to drive the runoff models with the parameters set to values from the generalisation procedure as if this were an ungauged catchment with properties that happen to coincide with one of those in the calibration set. The third generalisation method, sequential regression, is also included, based on the choices made for univariate regression (see Section 5.2.2). Figure 5.11 shows examples of generalised flood frequency curves for each method, for five catchments. The generalised curves for all catchments are given in Appendix B.2.

5.6.1 Overall performance

Table 5.6 summarises the performance of each generalisation method for each model, in terms of the mean and standard deviation of the absolute percentage error at various return periods (simulated compared to observed flood frequency curve). In each case, the best-performing generalisation gives average errors two to three times those for calibration. For the TATE, it is the univariate regression method which performs best (although only marginally better than sequential regression), with site-similarity performing the worst. For the PDM it is the opposite way around, with site-similarity best and univariate regression worst (with sequential regression marginally better than univariate). The earlier discussion (Section 5.2.3) suggested that the assessment of site-similarity may be unfairly biased by the inclusion of information from the target site in the generalisation. Because of this, Table 5.6 includes the summary information for two versions of site-similarity for the PDM, one allowing the target site to be included in each pooling group and one excluding it. The results show that the effect of including the target site within the information used for the generalisation estimate is rather large: with the target site excluded the apparent performance of the site-similarity method for the PDM model is much closer to that of univariate regression, although still marginally better. For the TATE model, both sets of site-similarity results (including and excluding the target site) are shown for completeness. Again there is a large difference in apparent performance between the two assessments but, for this model, the generalisation of the flood frequency curve provided by univariate regression has closest similarity to the at-site calibration results.

Table 5.7 compares the best-performing generalisation for each model here with the best model from the pilot project, FD0404. In order to provide results more comparable to the datasets used in this previous project, the table presents results for the subset of

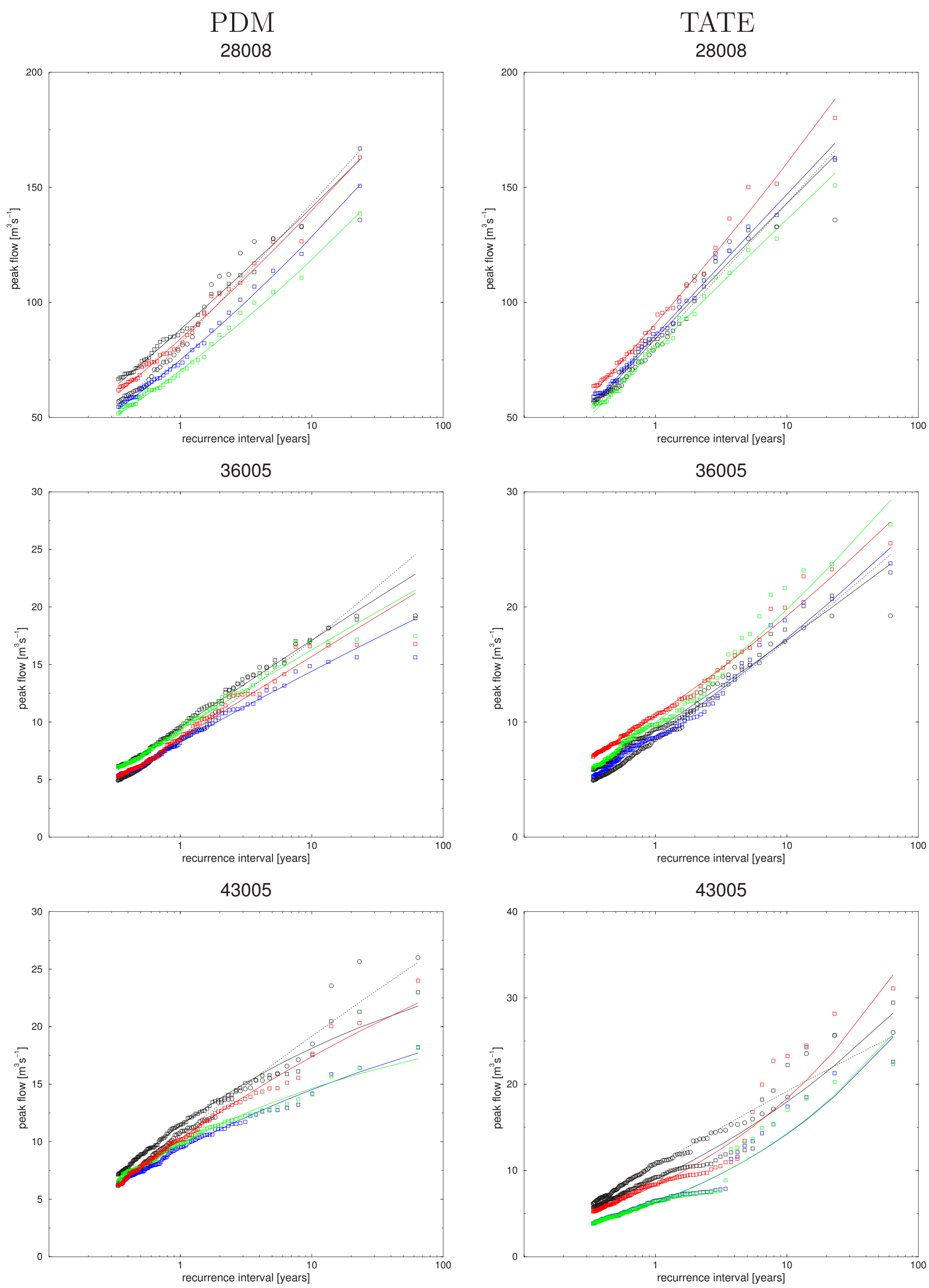


Figure 5.11 Example generalised flood frequency curves from each of the different methods (blue – univariate regression, green – sequential regression, red – site-similarity), compared to using calibrated parameter sets (black solid lines) and flood frequency from observed flows (black dotted lines), for PDM and TATE, for five catchments.

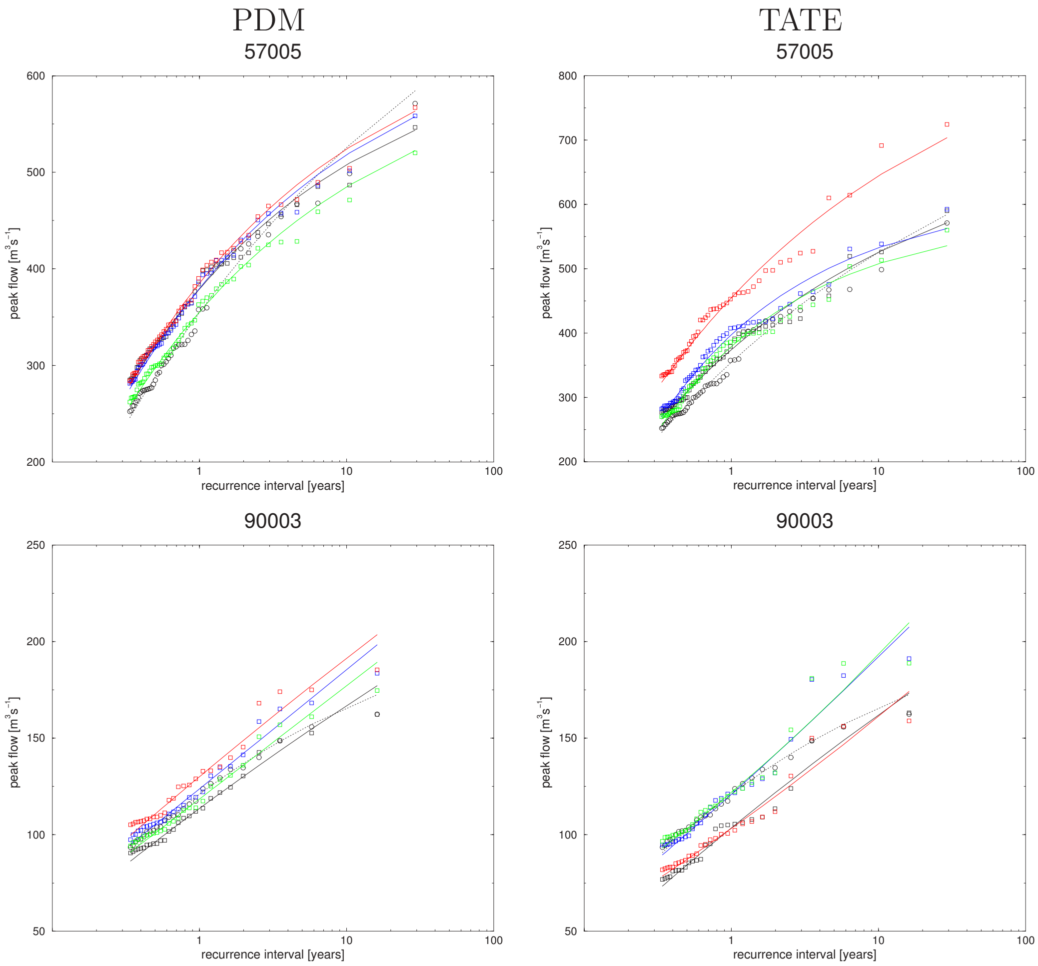


Figure 5.11 continued

catchments having hourly data in addition to repeating the results from Table 5.6. However, it should be borne in mind that the results are not directly comparable due to the use here of additional catchments and an extended data period. As the extended data period includes higher flood peaks for a number of catchments, it could be expected that the generalisation based on the extended data period would work better for flood peaks than if the old generalisation (based on the original data period) were used for the whole data period. Evidence for this was given in Tables 4.3 and 4.4 of Kay *et al*. 2003. For the PDM the new, site-similarity, generalisation shows roughly the same performance as the previous best-performing generalisation (sequential regression) if the “target-excluded” results are considered. For TATE there is very little difference in performance.

Table 5.6 Overall performance of each generalisation method for each model. Values are the mean and standard deviation (SD) across catchments of the absolute percentage errors in the estimated flood at a given return period. The performance using directly calibrated parameters is also given, for comparison.

		Mean and SD of absolute % errors	Return Period (years) 1.0	2.0	2.33	5.0	10.0	20.0
TATE	Calibration	Mean	11	9	8	5	5	9
		SD	10	9	8	7	7	8
	Univariate regression	Mean	21	21	21	22	24	27
		SD	19	22	22	27	32	37
	Sequential regression	Mean	22	21	21	22	24	27
		SD	20	23	24	28	33	38
	Site-similarity: target included (excluding boundary points)	Mean	26	28	28	30	32	36
		SD	38	45	46	50	53	56
	Site-similarity: target excluded (excluding boundary points)	Mean	32	33	34	35	37	41
		SD	50	55	56	59	61	62
PDM	Calibration	Mean	9	6	6	4	5	7
		SD	7	5	4	4	5	7
	Univariate regression	Mean	25	25	25	25	26	27
		SD	42	41	40	39	37	36
	Sequential regression	Mean	23	23	23	24	25	26
		SD	42	41	41	39	38	36
	Site-similarity: target included (including boundary points)	Mean	17	16	16	16	17	18
		SD	22	21	21	20	20	19
	Site-similarity: target excluded (including boundary points)	Mean	23	23	23	23	23	24
		SD	36	36	35	34	34	33

Table 5.7 Comparison of overall performance of the 'best' generalisation methods for each model in this project and in the pilot FD0404.

		Mean and SD of absolute % errors	Return Period (years) 1.0	2.0	2.33	5.0	10.0	20.0
TATE	Univariate regression (all sites)	Mean	21	21	21	22	24	27
		SD	19	22	22	27	32	37
	Univariate regression (hourly sites)	Mean	22	23	23	24	25	28
		SD	15	16	16	17	18	19
	FD0404 sequential regression (FD0404 sites)	Mean	20	21	21	23	24	28
		SD	20	20	20	22	26	35
PDM	Site-similarity: target included (all sites)	Mean	17	16	16	16	17	18
		SD	22	21	21	20	20	19
	Site-similarity: target excluded (all sites)	Mean	23	23	23	23	23	24
		SD	36	36	35	34	34	33
	Site-similarity: target included (hourly sites)	Mean	19	18	18	18	18	19
		SD	17	17	17	16	15	15
	Site-similarity: target excluded (hourly sites)	Mean	24	24	24	24	24	26
		SD	25	23	23	21	20	19
	FD0404 sequential regression (FD0404 sites)	Mean	22	23	24	24	26	27
		SD	18	18	19	20	21	23

Table 5.8 presents the performance of different generalisation methods compared to calibration in terms of the percentage of catchments with absolute errors (at return periods of between 2 and 50 years) classified in different groupings: A; all errors less than 15%, C; one or more errors greater than 30%, or B otherwise. This confirms the general conclusions drawn from Table 5.6. Specifically, site-similarity can be judged to perform best for the PDM since, even if the target site is excluded from the pooling group, it has the highest percentage in group A and the lowest in group C. However, for the TATE, univariate and sequential regression cannot be separated on the criteria in Table 5.8, while the results suggest that these are to be preferred over site-similarity because of the poorer performance of this when the target sites are excluded from the pooling group.

Table 5.8 Percentage of catchments with absolute errors in various classes, for each generalisation method, compared to calibration. Classes summarise errors at return periods of between 2 and 50 years as follows:

A : all less than 15%
B : all less than 30%, but one or more greater than 15%
C : one or more greater than 30%

	Class	Calibration	Generalisation			
			Univariate regression	Sequential regression	Site-similarity	
					Including Target	Excluding Target
PDM	A	80	25	32	45	33
	B	18	37	29	32	37
	C	2	38	39	23	30
TATE	A	69	30	27	30	22
	B	22	31	41	31	32
	C	9	39	32	39	46

Figure 5.12 illustrates the distributions of errors for the different generalisation methods. This again confirms that site-similarity performs best for the PDM - with a distinct peak to the distribution close to zero. However, the 'best' method for TATE is rather less clear, both for a given return period and across different return periods. Figure 5.12 also suggests a tendency towards slight underestimation at higher return periods, particularly for the PDM (both for calibration and generalisation).

5.6.2 Analysis of performance

Figure 5.13 and Figure 5.14 compare the performance of the TATE and PDM models at two different return periods (10 and 50 years respectively), by plotting the calibration and generalisation errors for one model against the other. These figures show that there is a degree of consistency of performance under generalisation, in that for a given catchment the direction of the error (positive or negative) is generally the same for each model. For univariate regression these errors are also similar in size (that is, the points are arranged close to the 1:1 line), but for site-similarity it is clear that TATE performs less well than the PDM. It is also interesting that, although the calibration errors clearly expand at the higher return period, the generalisation errors are not affected to the same extent. Note that the errors plotted here are logarithmic errors in the estimated flood at the given return period, rather than the percentage errors referred to previously: a logarithmic error of +1 means that the estimated peak flow at that return period is 2.7 times the observed peak flow, while -1 means that the peak flow is underestimated by a factor of 2.7.

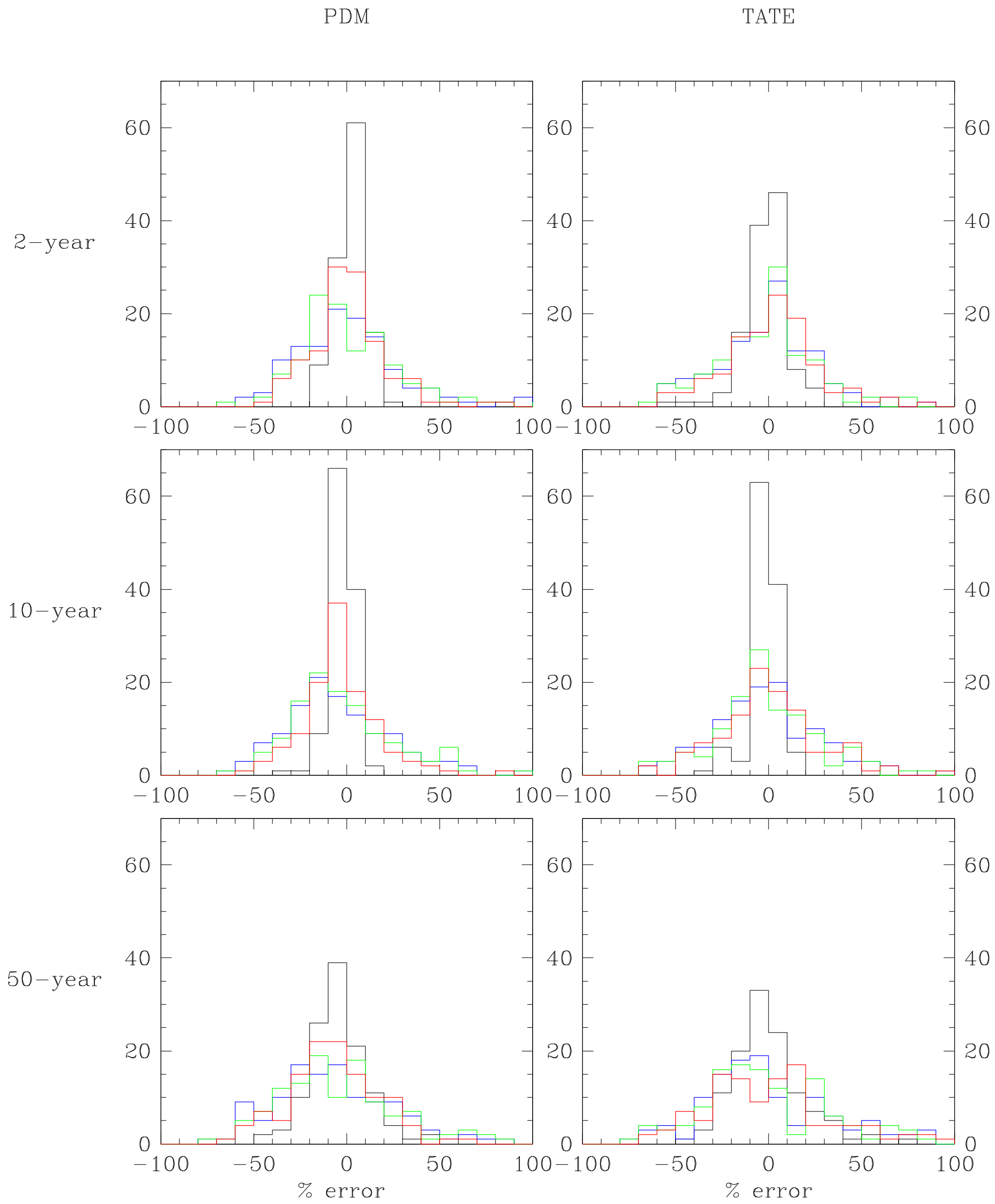


Figure 5.12 Distributions of the flood frequency curve errors at return periods of 2, 10 and 50 years, using calibrated parameters (black) and parameters generalised using univariate regression (blue), sequential regression (green) and site-similarity (red).

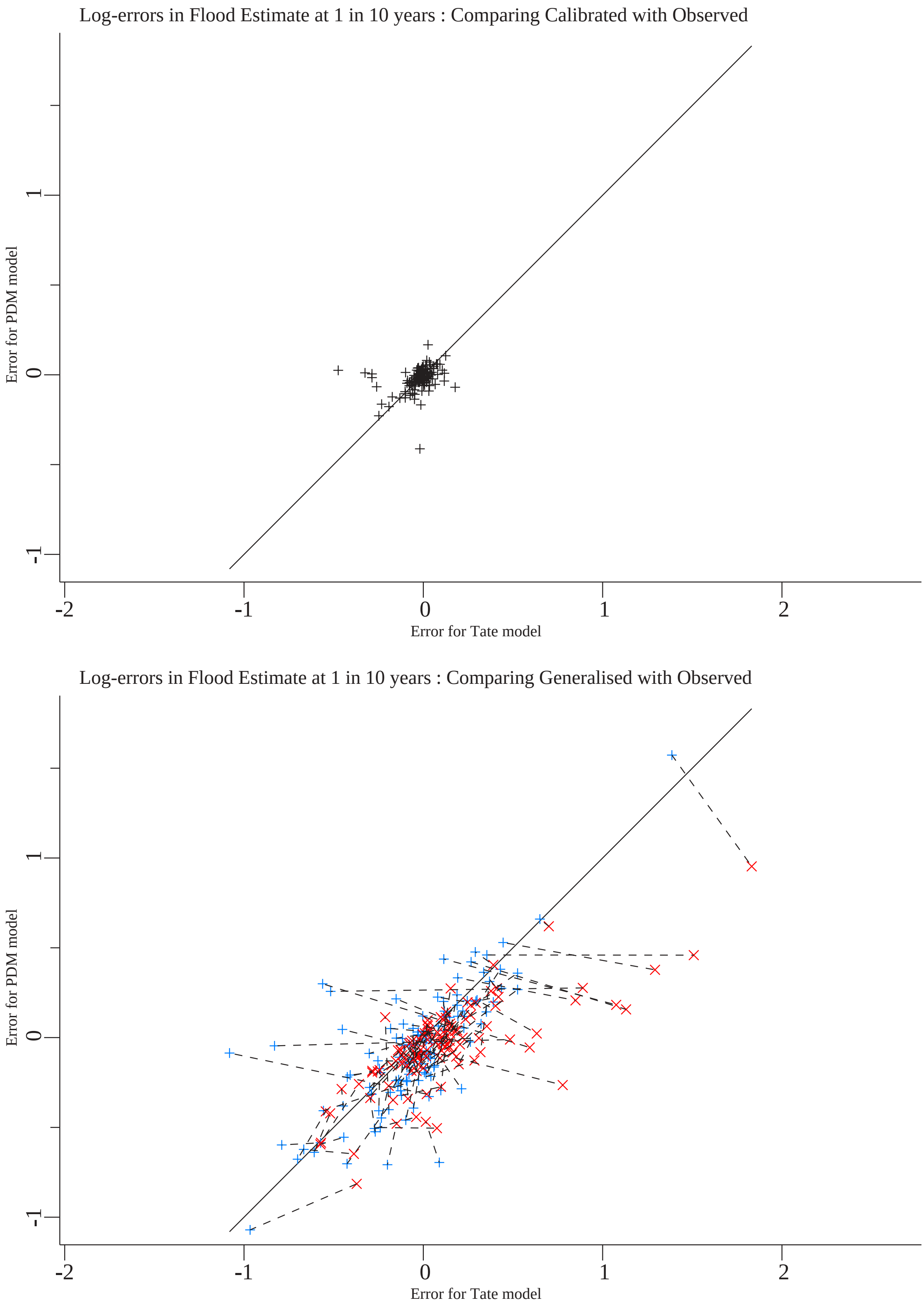
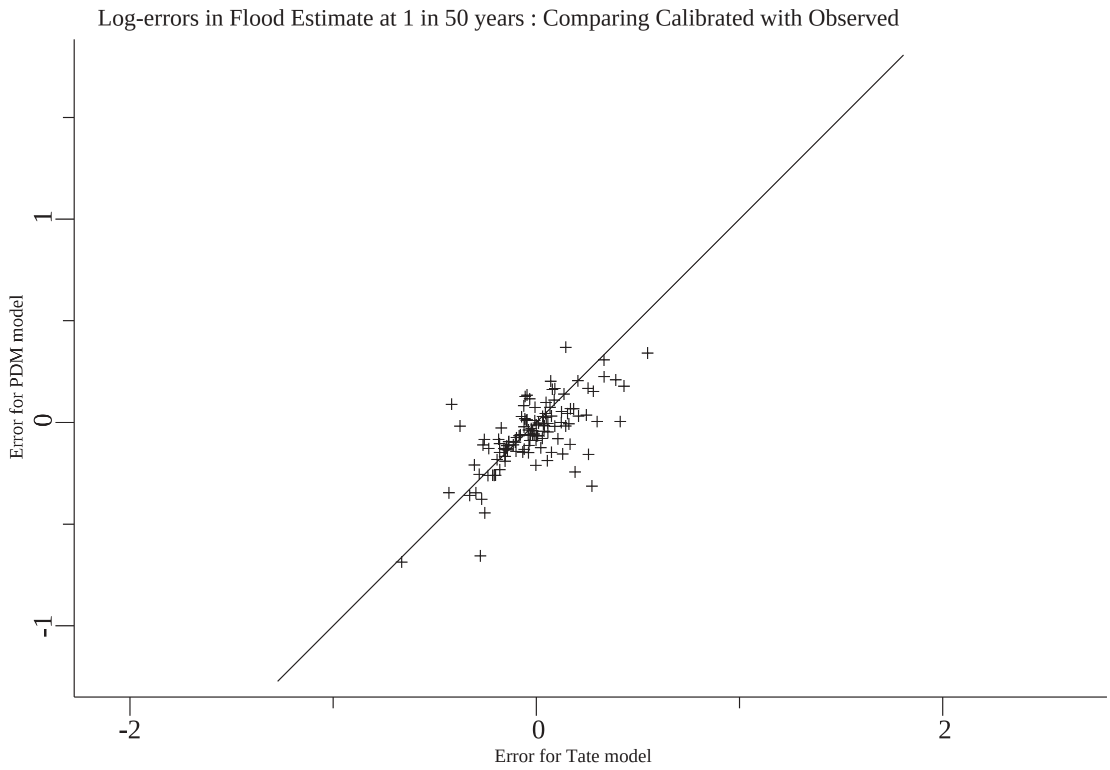


Figure 5.13 **Comparison of calibration and generalisation errors for TATE and PDM, at the 10 year return period. Top graph; calibration errors - black plus signs. Bottom graph; univariate regression - blue plus signs, site-similarity - red crosses, joined for each catchment by a black dashed line.**



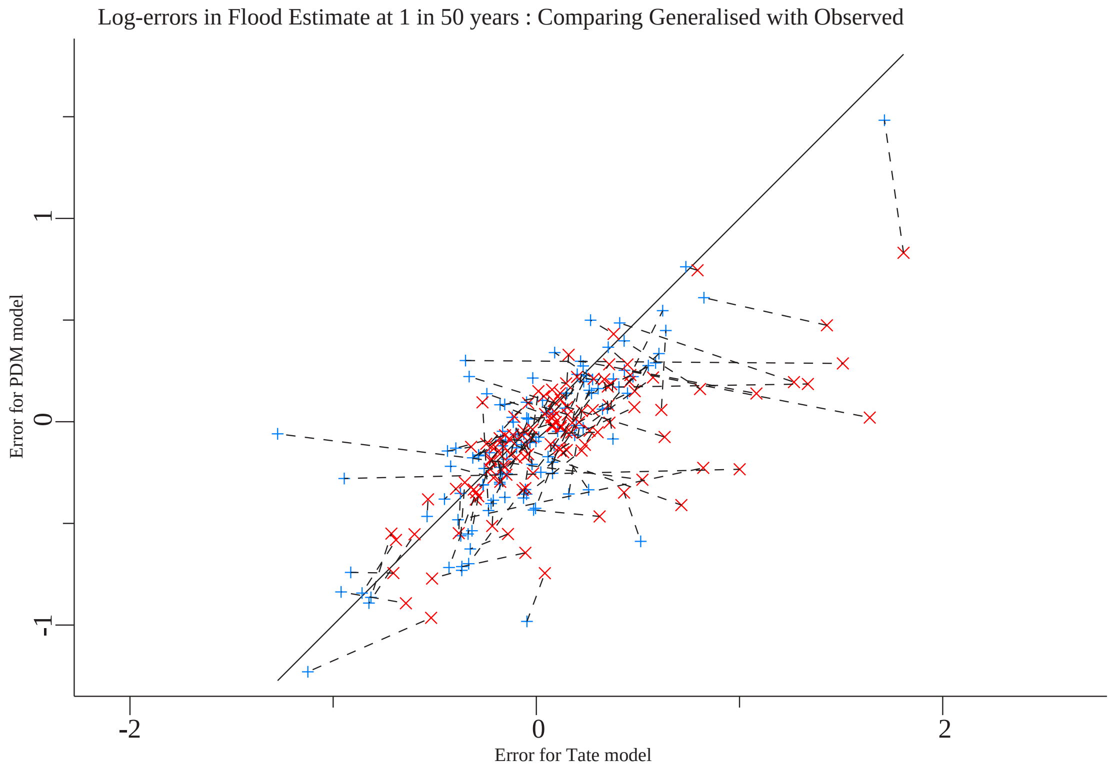


Figure 5.14 As Figure 5.13, but at the 50 year return period.

Appendix B.3 presents plots of univariate regression and site-similarity generalisation errors (at return periods of 10 and 50 years) versus catchment properties: these show no clear relationships. The only potential dependence appears to be between HOSTBFI and errors using TATE site-similarity.

The overall results of this chapter plainly pave the way for the recommendations of the final chapter of this report. It is, however, necessary to also consider associated uncertainty issues: these are the subject of the next chapter.

6 TREATMENT OF UNCERTAINTY

David Jones, Alison Kay

PREVIOUS FAST TRACK BOX ON PAGE 43

CHAPTER 6 TREATMENT OF UNCERTAINTY

Chapter 5 presented the methods used for calculating estimates of the flood frequency curves for ungauged catchments. This chapter is concerned with the likely errors in such estimates resulting from model parameter uncertainty. It does not cover data and model structure uncertainty. The final results of Chapter 6 are uncertainty bands for the estimated flood frequency curves: example results are contained in this chapter with a more extensive collection being presented in Appendix D.

This chapter describes the methods used for calculating the uncertainty bands for the estimated flood frequency curves. The initial sections discuss the various sources of uncertainty associated with the generalisation procedures and outline statistical models for representing these sources. Once such models are framed, it is possible to develop improved versions of the generalisation procedures, allowing higher weights to be attributed to the calibration results where the calibration uncertainty is low. These improved procedures (uncertainty-weighted estimates) have been implemented in earlier chapters but are derived from the models described here. This chapter outlines how the statistical models for uncertainty are fitted to the calibration results for each catchment. These methods lead to a model that describes how well the parameters of the runoff model for an ungauged catchment are determined from the catchment properties using the generalisation procedure.

The practical implementation of procedures for dealing with the parameter uncertainty uses an approach based on creating a collection of randomly-selected sets of model parameters which together have a variation equivalent to the uncertainty in the set of generalised estimates. Each of the random sets of parameters is used to derive an estimated flood frequency curve based on the result of driving the runoff model with a sequence of rainfalls. The variation among this collection of flood-frequency curves represents the uncertainty in the result of applying the generalisation procedure to an ungauged catchment.

NEXT FAST TRACK BOX ON PAGE 109

6.1 Introduction

This chapter outlines how uncertainty has been treated within the present project. Several different aspects of the project are affected by considerations of uncertainty and it is convenient to discuss these in a unified context, rather than piecemeal throughout the report. Chapter 5 has described the methods used for undertaking the spatial generalisation of the parameters of runoff models. The variants of these methods that have been recommended in the final chapter derive from the uncertainty considerations discussed in this chapter. The aspects of uncertainty considered are those resulting from model parameter identification in both calibration and generalisation. This project does not cover data and model structure uncertainty. Results for the uncertainty associated with estimates of flood frequency curves derived from generalised runoff models are presented Section 6.6.

As discussed earlier, the main target for the present project is to provide a methodology for estimating the flood frequency curve at sites for which there is not sufficient information to allow the direct calibration of a runoff model relevant for that site. In order to apply the continuous simulation principle, estimates for the parameters of the catchment model have to be derived from the parameter values calibrated for other catchments, relating these parameters to catchment properties.

As indicated in Figure 6.1, the estimate for a flood frequency curve will be subject to uncertainties inherent in the generalisation step. An important aim of the project is to provide a means of assessing these uncertainties. This analysis of uncertainty has two major benefits:

(i) it allows the development of generalisation procedures that take account of the different sources of uncertainty and which, at a theoretical level, are better than procedures which do not;

(ii) it allows the amount of uncertainty attributed to the final generalisation results to be compared to an empirical approach not based on an *analysis* of uncertainty.

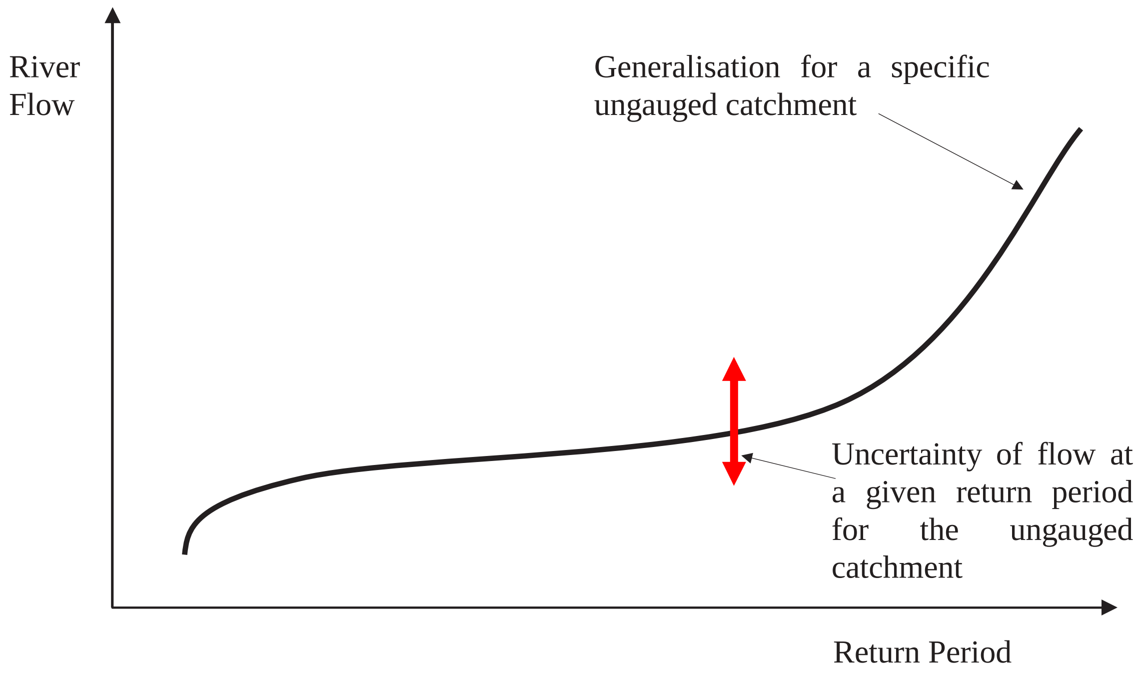


Figure 6.1 Uncertainty of the target results

To clarify this last point, while there are some minor reductions in uncertainty arising from point (i) above, effectively all of the reduction in the width of an uncertainty band placed about the flood frequency curve arises because the analysis of uncertainty is able to omit an unnecessary part of the uncertainty that would otherwise be included in a more empirical approach.

The approach based on the analysis of uncertainty distinguishes between the notional true value for a catchment, and the value that would be obtained for the same catchment based on a model calibrated using a limited length of data record. The difference between the two is one source of uncertainty. Use of the more complete analysis of uncertainty allows an assessment to be made of the likely error of the generalisation procedure in estimating the true catchment value, rather than in estimating a value that would notionally be obtained from an analysis of limited data for that catchment, if such data were available. More direct empirical approaches are essentially directed at errors in estimating the latter type of quantity.

In a practical context, the main benefit of using generalisation procedures based on an analysis of uncertainty is not that estimation errors are reduced compared to simpler procedures, although there may be a small improvement. However, the attribution of error sources inherent in the analysis of uncertainty does lead to narrower uncertainty bounds for estimates from the generalisation procedure.

Section 6.2 outlines the sources of error or uncertainty associated with generalisation procedures for continuous simulation, while Section 6.3 sets out a formal statistical model for how these sources of error combine within the present types of generalisation procedures. Details are provided in Section 6.4 of how the sizes of the uncertainties have been estimated. Section 6.5 outlines how these estimates for the sizes of the uncertainty components are used to generate randomised sets of runoff model parameters for an ungauged catchment that reflect the information obtainable via the generalisation model: these sets are centred on the straightforward generalised estimates and the variation within each set reflects the error likely to exist in the generalised estimate for an ungauged catchment. Section 6.5 also provides results comparing the sizes and hence relative importance of the different components of uncertainty. Finally, Section 6.6 outlines results for the uncertainty of the flood frequency curves for ungauged catchments estimated using the continuous simulation based on generalised estimates of the parameters of a runoff model.

6.2 Sources of uncertainty

Two of the sources of uncertainty that arise in generalisation are illustrated in Figure 6.2.

The first of these sources might be called "generalisation uncertainty". If one thinks of having several catchments all with identical catchment properties (as far as can be measured), the generalisation procedure would inevitably produce the same "best estimates" for the catchment model parameters, yet the catchments would not really have the same responses and hence the model parameters really should be different. The variation between the model parameters for catchments having the same properties is the “generalisation uncertainty”.

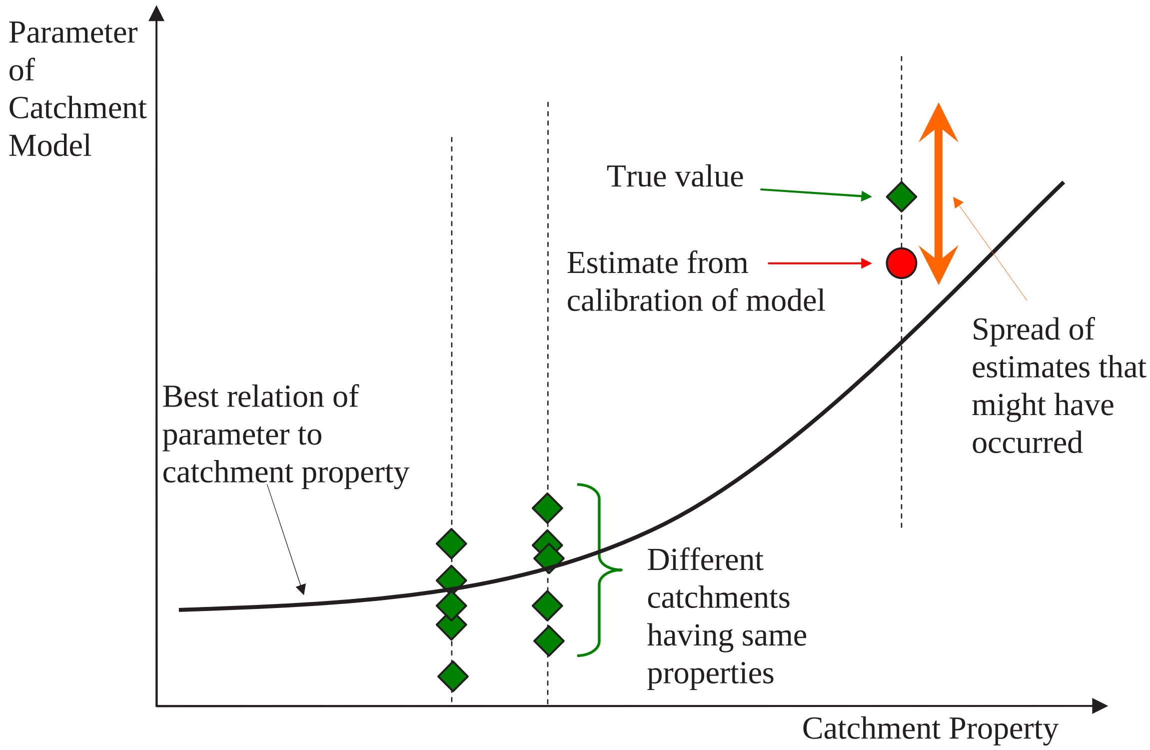


Figure 6.2 Sources of uncertainty in the relationship of a catchment model parameter to catchment properties

The second source of variation relates to the calibration of the catchment models for catchments where data are available: this procedure produces estimates of model parameters that would change if a longer data record for calibration were supplied. A pragmatic approach to the specification of the "true" model-parameters for a catchment model which is known to be only an approximation, at best, is to define them to be the values to which the parameter estimates would converge if the calibration procedure were applied to the given catchment and if an increasing length of data record were supplied. The difference between the estimates obtained from the existing data set and these "true" values will be called the "calibration error". It should be noted the "true" values defined here would change if a different model calibration procedure were applied: for example, one based on different selections for the objective functions used in the calibration.

The above discussion has not fully defined the "generalisation" rule relating model parameters and catchment properties that appears in Figure 6.2. In fact two versions of this curve are required. Again, a "true" generalisation curve can be defined as the rule that would be derived given an arbitrarily large collection of representative catchments to which the rule is fitted. Secondly, there is the "sample" generalisation curve that is derived from the collection of catchments actually available for catchment-wise model calibration. The "true" generalisation curve is defined without any specific parameterisation: notionally it is what the result would be from fitting generalisation rules whose flexibility (in catchment property space) increases as the number of catchments increases. In contrast, the "sample" generalisation curve may be derived by assuming a particular parametric form for the generalisation rule being fitted. Figure 6.3 is a revised version of Figure 6.2 that illustrates how the various error components inherent in the generalisation procedure relate to each other. Compared with Figure 6.2, Figure 6.3 is more closely related to the actual implementation of the generalisation procedure, since it depicts only one instance of a catchment with any given set of values for the catchment properties.

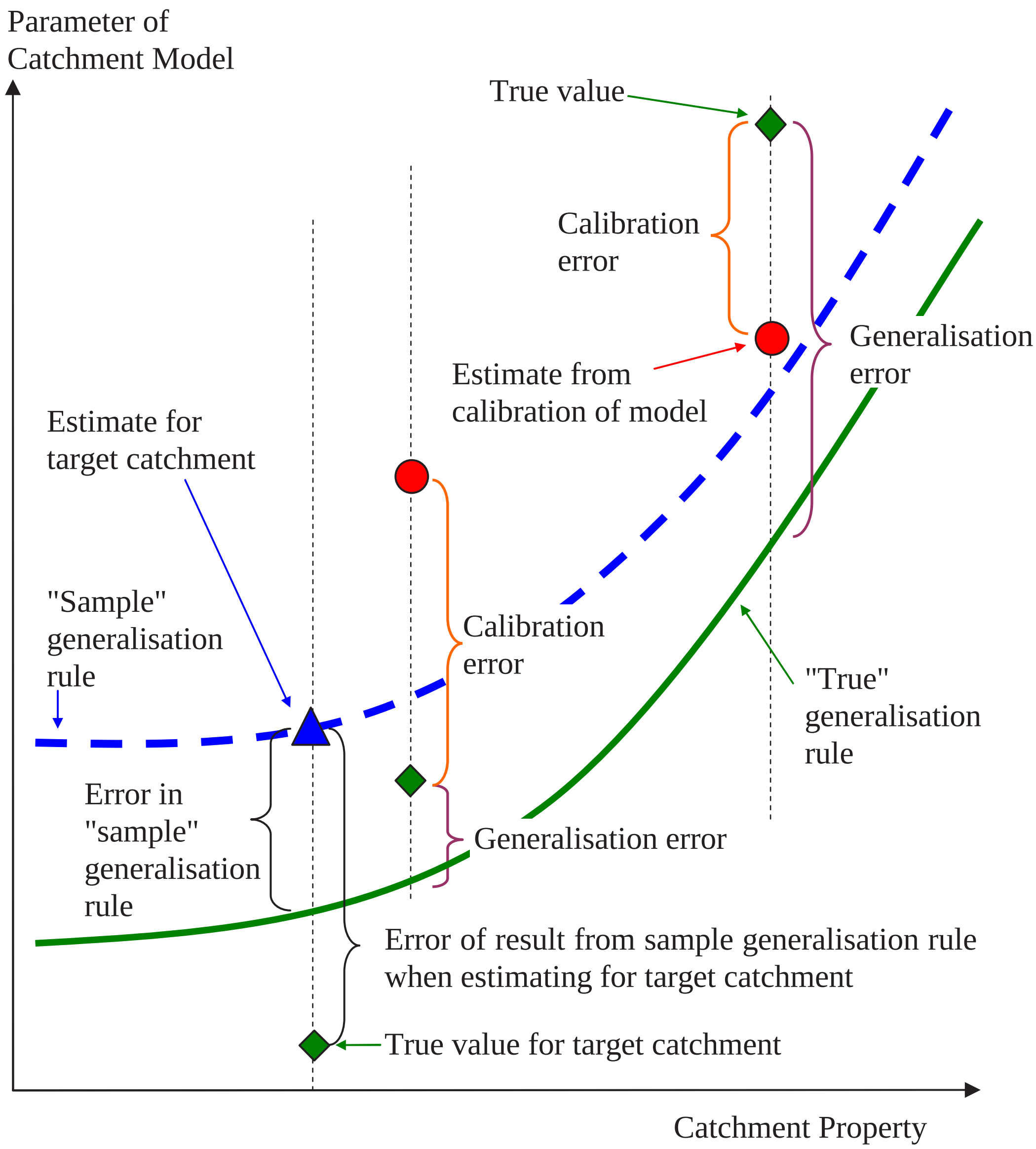


Figure 6.3 Relationship of errors in the generalisation procedure

6.3 Modelling of uncertainty

6.3.1 Treating model parameters separately

For simplicity, the discussion here treats a single model parameter of the catchment model at a time. For applications within this project, this forms the main basis of the methodology for handling uncertainty. While the uncertainties of all the model parameters are, in fact, considered jointly (a multivariate analysis), the main types of generalisation methods considered are constructed by dealing with each parameter separately (*i.e.* using univariate techniques). Section 6.3.3 outlines how the univariate models extend to the multivariate case for the simple generalisation methods. The treatment of uncertainty in the special case of the sequential regression method of generalisation (outlined in Section 5.1.2) is considered in Section 6.3.4, while Section 6.3.5 describes some possibilities for more fully-multivariate methods of parameter generalisation.

Consider a fixed set of values for catchment properties and consider a number of "very similar" catchments having these catchment properties (that is, they would be indistinguishable according to hydrological catchment descriptors). If the catchment model were calibrated separately to these catchments, the mean value of the selected parameters across all of these "very similar" catchments is defined as the "true" generalisation parameter value μ. The "true" parameter value for a given catchment is

$$T = \mu + \eta , \qquad (6.3.1.1)$$

where T represents the model parameters that would be calibrated for a given site if there were an infinitely long data series available. The random error term η differs between instances of catchments having the same properties and represents variations in the "true" model parameters for catchments that would be judged to be very similar. The random error term η is assumed to have a mean of zero, given that μ represents the mean value across all similar catchments.

The parameter value obtained by calibrating the catchment model for a single catchment is denoted by Y, where

$$Y = T + \varepsilon = \mu + \eta + \varepsilon . \qquad (6.3.1.2)$$

Here ε represents the calibration error for the catchment. This random variable is assumed to have a zero mean, but its distribution will typically have a spread related to the length of record available for calibration and other catchment-specific factors.

In the following, it is assumed that a number n of catchments are available for which the catchment model has been calibrated by a well-defined approach, and for which sets of catchment properties are available. Thus the data to be used for generalisation consist of the calibrated parameter values for sites $i = 1,...,n$ which can be put into a similar representation to Equation (6.3.1.2)

$$Y_i = T_i + \varepsilon_i = \mu_i + \eta_i + \varepsilon_i , \qquad (6.3.1.3)$$

where the following notation is used:

μ_i	mean value of the parameter across catchments "very similar" to catchment i;
η_i	the generalisation error for catchment i;
T_i	$(=\mu_i+\eta_i)$ the parameter value that would be calibrated at catchment i, given an infinitely long calibration data set;
ε_i	calibration error at catchment i due to having a limited data set for calibration.

The problem of generalisation typically involves a catchment for which no calibration data are available: values for such a catchment will be indicated with a subscript *. For the present discussion it is convenient to assume that a generalisation procedure exists and that a generalised estimate for the model parameter can be constructed: it is also assumed that there are sufficiently many calibration catchments with sufficiently extensive datasets for the generalised value, μ_*, to be determined essentially without error. The task of the generalisation procedure is to calculate an estimate for T_*, where

$$T_* = \mu_* + \eta_* . \tag{6.3.1.4}$$

Here η_* is the unknown generalisation error for the target catchment: given the assumption that the catchments used for calibrating the generalisation model are representative of all target catchments to which the calibration model might be applied, η_* is assumed to have the same statistical characteristics as the η_i for the calibration catchments. The size of the generalisation error is therefore characterised by the variance of η_*. The analysis of uncertainty to be undertaken allows this variance to be estimated.

In contrast, a direct empirical treatment of uncertainty can be viewed as involving the quantity

$$Y_* = T_* + \varepsilon_* = \mu_* + \eta_* + \varepsilon_* , \tag{6.3.1.5}$$

instead of T_* as the target quantity. Here ε_* represents a notional calibration error for the target catchment. The direct approach provides an estimate for the variance of $(\eta_* + \varepsilon_*)$, which is not what is wanted. Specifically, under the above assumptions, the direct approach would base its assessment of uncertainty on

$$s^2 = n^{-1}\sum_{i=1}^{n}\left(Y_i - \hat{\mu}_i\right)^2 = n^{-1}\sum_{i=1}^{n}\left(Y_i - \mu_i\right)^2 = n^{-1}\sum_{i=1}^{n}\left(\eta_i + \varepsilon_i\right)^2 . \tag{6.3.1.6}$$

This provides an estimate of the variance of $(\eta_* + \varepsilon_*)$ in cases where the extent of the availability of calibration data can be treated as randomly determined in a statistically equivalent way for different catchments, and where this includes data that is notionally available for the target catchment (but is not) in order to provide the notional value of Y_* that the generalisation estimate is to be judged against. Note that this particular assumption is not made for the more complete analysis of uncertainty used in this project. For this analysis, the quantities Y_* and ε_* are not involved at all.

The above discussion has assumed that the generalised value, μ_*, can be estimated without error, which is clearly not the case. The uncertainty associated with this generalisation error can readily be taken into account in the analyses of uncertainty: this is outlined later.

It should be noted that, in earlier sections, the notation used has employed the Greek letters $\alpha, \beta, \ldots$ *etc*. for the calibrated parameters of the runoff models. In that context, these quantities control the behaviour of the (runoff) model. In the present context, the calibrated parameter-values are treated differently, and there are other quantities which control the behaviour of the model. Here the model relates the calibrated values to catchment properties using another set of parameters, and the calibrated values take on the role of "observed" values. The models being used for this generalisation step are essentially the same as regression models taken from statistical theory, where there is a common notation of using *y* or *Y* for observed values of the quantities being modelled.

6.3.2 Some assumptions in the uncertainty model

The model structure outlined in Equation (6.3.1.3) needs to be extended by making some assumptions about the statistical properties of the sets of error components $\{\varepsilon_i\}$, $\{\eta_i\}$ and η_*. Given the way these error components have been defined, the assumption follows that they all have a mean value of zero. A major assumption is that the typical sizes of the generalisation errors $\{\eta_i\}$, as measured by the variance, do not vary in any predictable way in terms of any set of known catchment properties: the assumption is that the variance of the generalisation error is constant. In practice the generalisation procedures are applied to transformed versions of the parameters of runoff models (for example, by constructing a generalisation rule for the logarithm of a parameter and then transforming back with the exponential function to derive the final generalised value). Such transformations are chosen as a matter of judgement on the combination of several criteria. For the present study, the assumption of constant variance has been checked for the transformed parameters on the basis of scatter plots involving the calibrated values, the generalised values once derived, and the individual catchment properties.

Another set of assumptions made is that the error components are uncorrelated between the different types of component and are uncorrelated for errors of the same type at different catchments. These are partly justified by the conceptual basis of the error components in the above model. It can be further argued that, if there were any correlation in the generalisation errors with a given set of catchment properties as the basis of the generalisation, it could be used to create an improved generalisation rule either using the same set or an extended set of catchment properties. Thus the usual checks on the behaviour of generalisation rules, which involve checking that there is no benefit from obvious modification to the rule, serve to check on this assumption. The assumption of no-correlation for the calibration errors $\{\varepsilon_i\}$ is more problematic. It can be argued that neighbouring catchments will be affected by the same weather events within their calibration periods, and thus the calibration errors might be expected to be correlated on this basis. The likely extent of this correlation is unknown: adjacent catchments might be sufficiently different for them to be sensitive to different aspects of rainfall patterns and thus the calibration errors might be quite unrelated. For the present study, the catchments being considered are widely dispersed geographically, although there are some pairs of catchments that are rather close together. Results from the present study (Section 6.5.2) suggest that the calibration errors make only a small contribution to the uncertainty with which the generalised values predict model parameters for ungauged catchments. Thus the assumption of uncorrelated calibration errors may be relatively unimportant.

6.3.3 Treating model parameters jointly

When all the parameters of a runoff model are considered together, the same uncertainty structure as used in Section 6.3.1 can be used except that now all the quantities involved are vector-quantities. Thus, if a catchment model has P parameters, then the quantity Y in the earlier section is a (column) vector with elements $\{Y(p), p=1,\ldots,P\}$, where $Y(p)$ is the calibrated value of the *p*th parameter. Similarly μ is the vector of "true" generalised values for catchments similar to a given catchment. Thus μ is now the (column) vector with elements $\{\mu(p), p=1,\ldots,P\}$, where $\mu(p)$ is the "true" generalised value for the *p*th parameter.

The error-components η and ε identified in Section 6.3.1 also become vectors. Section 6.3.2 described the assumptions about the statistical properties of these components that are required for the analysis of uncertainty in the case of univariate modelling. These assumptions remain essentially unchanged for the multivariate case. In particular, errors of the same type and for the same catchment but for different model parameters are allowed to be cross-correlated. All other cross-correlations are assumed to be zero. It is clearly important to include the possible cross-correlation of the generalisation errors for different parameters within any assessment of the uncertainty of quantities derived from runoff models using generalised parameters.

Two of the methods of model parameter generalisation that are considered in this project are such that the generalised values for a given parameter are derived using only the calibrated values for the same model parameter. Specifically, the regression and site-similarity approaches to generalisation treat each model parameter separately and, in fact, allow different choices for the structure of the generalisation rule to be made for each parameter. In the case of the regression method, the generalisation rule can be taken to be equivalent to the regression parameters relating the calibrated values of the particular model-parameter to catchment properties: these regression parameters do not make used of the calibrated values for the other model parameters. In such cases the generalisation rules are effectively univariate, but it is still possible to undertake a multivariate analysis of the generalisation uncertainty, as outlined in Section 6.4.5.

The third method of generalisation, sequential regression, does not lie neatly within either a multiple univariate or multivariate modelling context, and it is necessary to extend the multivariate model to consider additional quantities. This is described later in Section 6.5.3 since it is helpful to have seen the uses made of the uncertainty model in the simpler cases. Section 6.5.4 outlines how the multivariate model for uncertainty might be used to construct more fully-multivariate generalisation rules, and says why this approach has not been undertaken in this project.

6.4 Estimation of uncertainty components

6.4.1 Analysis of uncertainty

The analysis of uncertainty involves the variances of the two different types of errors: the generalisation error η and the calibration error ε. Because the extent of calibration error will vary between catchments, partly due to differences in record lengths, the variances of the calibration error will change from catchment to catchment. The quantities involved in the analysis of uncertainty are therefore:

σ_η^2, the variance of the generalisation error, and

$\sigma_{i,\varepsilon}^2$, the variance of the calibration error for catchment i.

The analysis proceeds in three stages which are outlined in the following sections. First, the variances $\sigma_{i,\varepsilon}^2$ are estimated separately for each catchment (Section 6.4.2) and these values are then treated as fixed. Secondly, the generalisation variance, σ_η^2, is estimated by an iterative procedure (Section 6.4.3) which involves constructing generalisation estimates on the basis of assuming that the variance components are all known and then deriving a new estimate of σ_η^2 from a comparison of these with the calibrated values of the parameters. Finally, the uncertainty of the generalised values for a given target catchment can be evaluated (Section 6.4.4): this involves augmenting the basic generalisation uncertainty, represented by σ_η^2, with a contribution from the uncertainty arising in using the "sample" generalisation rule instead of the unknown "true" generalisation rule. This calculation is summarised in Table 6.1.

Table 6.1 Summary of the calculation of the overall estimation variance for generalisation models.

Quantity	Interpretation
$T_* = \mu_* + \eta_*$	The "true" parameter value is the value from the "true" generalisation curve (Figure 6.3) for the target catchment, plus the generalisation error
$\hat{T}_* = \hat{\mu}_*$	The estimated parameter value for the target catchment is the value from the "sample" generalisation curve.
$T_* - \hat{T}_*$ $= \eta_* + (\mu_* - \hat{\mu}_*)$	The estimation error is the sum of the generalisation error, η_*, and the error in the sample generalisation curve, $(\mu_* - \hat{\mu}_*)$.
$\mathrm{var}(T_* - \hat{T}_*)$ $= \sigma_\eta^2 + \mathrm{var}(\hat{\mu}_*)$	The variance of the overall estimation error is the sum of the generalisation variance and the variance of the sample generalisation estimate.

The theory outlined in Sections 6.4.3 and 6.4.4 treats each individual parameter of a runoff model separately. In contrast, the assessment of the uncertainty of the flood-frequency curve derived from spatial generalisation requires that the uncertainty of the complete set of parameters should be assessed jointly. In the present context this is undertaken by estimating the covariance of the estimation errors for pairs of parameters: this is outlined in Section 6.4.5.

The estimation of uncertainty components is based on the usual types of assumptions about the error components: the generalisation errors η_i and calibration errors ε_j are assumed to be uncorrelated between the two types of errors and across all catchments. While it is possible to extend the methodology to include cases where the target catchment for the generalisation step is one of the catchments used as the basis of the generalisation, this is not dealt with here as it is not one of the major concerns of the project: discussion of this case is provided by Jones *et al.* (2004). It can, however, be noted that this extended methodology offers the potential for improving single-catchment calibrations of runoff models by transferring information via a generalisation rule. Any improvement from this type of approach is likely to be small unless the record length available for single-catchment calibration is currently rather short.

6.4.2 Estimation of the calibration variances

For the present project, the variances of the calibration errors, $\{\sigma^2_{i,\varepsilon}\}$, have been estimated as an extension of the procedure for calibrating the runoff models, by applying a jackknifing methodology. This is an established statistical procedure that provides a way of correcting the bias of an estimate of a quantity that is of direct interest, and for estimating the sampling variance of the estimate. The jackknife methodology can be traced back to initial ideas for bias correction by Quenouille (1949, 1956) and, for variance estimation, by Tukey (1958). By the late 1970s jackknifing had been given a firm theoretical justification and practical experience with the methodology was beginning to build up, as exemplified by the papers by Miller (1964, 1974), Bissell and Ferguson (1975). More recent discussions of the methodology include the paper by Peddada (1993). In current text-books, jackknifing is often discussed in association with another related technique called “bootstrapping”: see, for example, the books by Efron and Tibshirani (1993), Shao and Tu (1995) and Davison and Hinkley (1997).

In the usual statistical theory, the jackknife estimate of the sampling variance of an estimate can be formulated as follows. It is assumed that the basic estimate is a function of N items, denoted by $\{X_1,\ldots,X_N\}$ and it is assumed that the N items are statistically independent. The basic estimate is assumed to be defined in a consistent way as the number of items available changes. Let the basic estimate obtained from the N items be denoted by Z_N, and define the estimate that would be produced from the $(N-1)$ items, when item i is omitted from the full list of items, to be $Z_{N-1,i}$. There are N such estimates with one item deleted: the sample mean of these is defined to be

$$\overline{Z}_N = N^{-1}\sum_{i=1}^{N} Z_{N-1,i}\ . \qquad (6.4.2.1)$$

Then the jackknife variance estimate, defined so as to estimate $\mathrm{var}(Z_N)$, is

$$v = \frac{N-1}{N}\sum_{i=1}^{N}\left(Z_{N-1,i} - \overline{Z}_N\right)^2\ . \qquad (6.4.2.2)$$

This estimate is based on the variation between the leave-one-out estimates $\{Z_{N-1,i}\}$ but it contains an adjustment factor that accounts for the non-independence of these estimates. In cases where several quantities are being estimated simultaneously, as for the parameters of the runoff models, the jackknife variance estimation procedure provides an estimate of the covariance matrix of the set of estimated values as

$$V = \frac{N-1}{N} \sum_{i=1}^{N} \left(Z_{N-1,i} - \overline{Z}_N\right)\left(Z_{N-1,i} - \overline{Z}_N\right)^{\mathrm{T}}, \tag{6.4.2.3}$$

where the basic estimate Z_N and the leave-one-out estimates $\{Z_{N-1,i}\}$ are now vector-valued.

For the present project, the jackknife variance estimation procedure has been applied in the way described below, which includes a procedural element intended to overcome problems associated with serial dependence in the time-series being used for the calibration of runoff models. This element is related to the idea of Moran (1975), in which serial correlation is overcome in the estimation of the sampling variance of a mean value by first forming mean values for non-overlapping sub-intervals which are each long enough for these sub-interval means to be effectively uncorrelated. In the present context the "N items" on which the jackknife procedure is based are identified with the modelling error information contained within each of N calendar years. At the leave-one-out stage, the standard runoff-model calibration procedure is applied by treating as missing all of the error contributions to objective functions that arise from a given calendar year. Thus it is assumed that the overall effect on the calibration procedure of the whole set of modelling errors in separate years will be effectively independent across the years. Notionally the assumption of independence being made relates to the errors in modelling flow values, not to the flow values themselves. The jackknife variance estimation procedure for estimating the uncertainty of the calibrated values of the parameters of the runoff model is as follows:

(i) Apply the calibration procedure to the full data set to create the vector of values of fitted model parameters that are carried over into the generalisation procedure: this vector is effectively the basic estimate Z_N above.

(ii) For each of N calendar years covering the data period for which flow information is available, apply the calibration procedure to the data set ignoring contributions to the objective functions used for calibration that arise from a given calendar year. Each such leave-one-year-out calibration creates a vector of parameter values corresponding to $Z_{N-1,i}$ above.

(iii) Use the above formula to calculate the estimate V for the covariance matrix of the calibration errors for a vector of runoff model parameters.

In practice it may be most convenient to implement the leave-one-year-out calibration step by setting to "missing" all of the values of observed flows for a given calendar year. Calibration procedures for runoff models are typically constructed to cope with periods of missing data in observed flow records. Note that the leave-one-year-out procedure involves ignoring the observed flow values within a given year, but the observed rainfall and evaporation data for that year would continue to be used to create the full set of modelled flow values, and their effect would be carried over into subsequent years.

Use of jackknife procedures to assess the uncertainty of the calibrated estimates of runoff model parameters appears not to have been tried before. Other approaches to assessing calibration uncertainty are available, but some are not suited to the calibration scheme adopted for the present project, where the scheme involves using different objective functions for determining each model parameter. In the present context, jackknifing was chosen because it requires less extensive computational resources than the alternative approach of bootstrapping (see references noted at the beginning of this section) and is less complex to implement in conjunction with the calibration procedures. However, the present research indicates that the jackknife estimates of variance may be overly sensitive for use in the context of calibrating runoff models: this sensitivity arises from the highly non-linear effects of the model parameters on the model outputs.

6.4.3 Estimation of the generalisation variance

Chapter 5 considered two major types of spatial generalisation procedure: regression and site-similarity. Although these are distinct procedures, with different conceptual bases, they are sufficiently similar that the same approach to the estimation of the generalisation variance can be adopted. In particular, the two generalisation procedures have the common structural feature that the basic form for the estimated value at a given target site can be represented as a linear combination of the parameter values obtained for the calibration catchments:

$$\hat{\mu}_* = \sum_{j=1}^{n} w_{*j} Y_j \ .$$

For site-similarity, the weights w_{*j} are determined by the rule for defining neighbours to the target site and by the weighting within these neighbours: many of these weights will be zero. For regression, the weights are derived from the estimates of the regression coefficients (which are themselves linear combinations of the calibrated values Y_j) and from the catchment descriptors for the target catchment. In the case of generalisation procedures that take account of the uncertainty model, the “best” weights would also depend upon variance components σ_η^2 and $\{\sigma_{i,\varepsilon}^2\}$: in fact the weights only depend on the ratios, K_i, where

$$K_i = 1 + \frac{\sigma_{i,\varepsilon}^2}{\sigma_\eta^2} . \tag{6.4.3.1}$$

Section 6.4.2 has outlined a method for estimating $\{\sigma_{i,\varepsilon}^2\}$, and these estimates are now treated as fixed, leaving the problem of estimating σ_η^2. This problem is solved by the approach of using iteratively re-weighted estimates. In the approach an initial guess, $k^{(0)}$, for the value of

$$k = \frac{1}{\sigma_\eta^2} , \tag{6.4.3.2}$$

is supplied: the value $k^{(0)} = 0$ is commonly taken for convenience. This allows an initial set of values of $\{K_i\}$ to be constructed, and hence corresponding values for $\{w_{ij}\}$ can be found. Here w_{ij} is the weight used on the value for catchment *j* when constructing the generalised value for a catchment having the same catchment properties as catchment *i*. To align with the usual theory adopted in the case of regression, the approach used here allows the calibrated value for catchment *i* to be used as part of the set of data used to construct the generalised value that would be calculated for a catchment having the same properties as catchment *i*. Either directly, or subsequently to calculating the weights $\{w_{ij}\}$, the set of generalised values $\hat{Y}_i$ can be calculated. This set of generalised values, across the calibration catchments, is then used to construct an estimate of σ_η^2 on the following indirect basis. The approach here is based on calculating the weighted sum of squares of residuals

$$S^2 = \sum_{i=1}^{n} K_i^{-1}\left\{Y_i - \hat{Y}_i\right\}^2. \tag{6.4.3.3}$$

The expected value of S^2 is related to σ_η^2 in a slightly complicated way. First, note that

$$Y_i = \mu_i + \eta_i + \varepsilon_i = \mu_i + \omega_i, \tag{6.4.3.4}$$

where ω_i is a random variable with variance

$$\sigma_\eta^2 + \sigma_{i,\varepsilon}^2 = K_i\sigma_\eta^2. \tag{6.4.3.5}$$

Then

$$\begin{aligned} Y_i - \hat{Y}_i &= \mu_i + \omega_i - \sum_{j=1}^{n} w_{ij}\left(\mu_j + \omega_j\right) \\ &= \omega_i - \sum_{j=1}^{n} w_{ij}\omega_j. \end{aligned} \tag{6.4.3.6}$$

The last step here holds exactly in the case of regression-based generalisation procedures, according to the theory for that case, and it follows in other cases in an approximate sense from the assumption that the generalisation procedure is sufficiently flexible to accommodate all the smooth variation of the "true" model parameters in relation to the catchment properties. Hence

$$Y_i - \hat{Y}_i = \sum_{j=1}^{n} u_{ij}\omega_j, \tag{6.4.3.7}$$

where $\{u_{ij}\}$ is another set of weights. Then

$$\begin{aligned} \mathrm{E}\left(S^2\right) &= \mathrm{E}\left(\sum_{i=1}^{n} K_i^{-1}\left\{Y_i - \hat{Y}_i\right\}^2\right) = \mathrm{E}\left(\sum_{i=1}^{n} K_i^{-1}\left\{\sum_{j=1}^{n} u_{ij}\omega_j\right\}^2\right) = \sum_{i=1}^{n} K_i^{-1}\left\{\sum_{j=1}^{n} u_{ij}^2\mathrm{E}\left(\omega_j^2\right)\right\}, \\ &= \sum_{i=1}^{n} K_i^{-1}\left\{\sum_{j=1}^{n} u_{ij}^2 K_j\right\}\sigma_\eta^2. \end{aligned}$$

It then follows that the quantity s^2, defined by

$$s^2 = \frac{S^2}{\sum_{i=1}^{n} K_i^{-1} \left\{ \sum_{j=1}^{n} u_{ij}^2 K_j \right\}}, \tag{6.4.3.8}$$

proves an unbiased estimate for σ_η^2, at least according to assumption that the values $\{K_i\}$ being used are correct. The value for s^2 is accepted as an improved estimate of σ_η^2, and hence a new value for k is calculated as

$$k^{(1)} = \frac{1}{s^2}. \tag{6.4.3.9}$$

This then allows new sets of values of $\{K_i\}$ and $\{w_{ij}\}$ to be found and new "re-weighted" generalisation estimates $\hat{Y}_i$ are calculated. Eventually a new estimate s^2 for σ_η^2 is obtained. This procedure is repeated iteratively until convergence, when the final value for s^2 is used to provide the estimate for σ_η^2 that is carried forward into other calculations.

It may be noted that the expression

$$n^* = \sum_{i=1}^{n} K_i^{-1} \left\{ \sum_{j=1}^{n} u_{ij}^2 K_j \right\} \tag{6.4.3.10}$$

can be simplified in some cases. For example, if the generalisation procedure is based on weighted least-squares regression, this quantity can be shown to be equal to $(n - p - 1)$, where p is the number of catchment properties used in the regression relation (not counting the constant term in this set).

6.4.4 Estimation of the variance of errors in generalised parameter values

As outlined in Section 6.3, the relevant uncertainty associated with the generalised estimate of a parameter of the runoff model is that which treats the quantity being estimated as the true parameter value for the target catchment. Specifically, this is defined to be the value that would be obtained by a standard calibration procedure if there were an arbitrarily large amount of data available for direct calibration of the runoff model

For a target catchment, denoted by the subscript *, the generalised estimate is constructed as

$$\hat{\mu}_* = \sum_{j=1}^{n} w_{*j} Y_j , \tag{6.4.4.1}$$

where the weights $\{w_{*j}\}$ are determined by the type of generalisation procedure being applied. These weights are constructed according to the same rules as used in Section 6.4.3 except there is now no "iterative" element to the procedure: the weights are calculated using the final value of σ_η^2 from the iterative procedure. The weights also depend on the calibration uncertainties $\{\sigma_{i,\varepsilon}^2\}$, but these are only required for the catchments used in constructing the generalised estimate, not for the target catchment. At this stage, the weights are treated as fixed.

The quantity being estimated is T_*, where

$$T_* = \mu_* + \eta_* . \tag{6.4.4.2}$$

Thus the error in the estimate is

$$\begin{aligned} T_* - \hat{\mu}_* &= T_* - \sum_{j=1}^{n} w_{*j} Y_j = (\mu_* + \eta_*) - \sum_{j=1}^{n} w_{*j} (\mu_j + \omega_j), \\ &= \eta_* - \sum_{j=1}^{n} w_{*j} \omega_j . \end{aligned} \tag{6.4.4.3}$$

As in Section 6.4.3, this last step holds exactly in the case of regression-based generalisation procedures, according to the theory for that case, and it holds for site-similarity procedures if they are sufficiently flexible to accommodate all the smooth variation of the "true" model parameters in relation to the catchment properties.

In the case of most interest to this project, the target catchment is not any of those used to provide information for the generalisation procedure. This means that η_* is uncorrelated with all of the combined errors in the set $\{\omega_j\}$, which are themselves uncorrelated within the set. Therefore the variance of the estimation error is

$$\begin{aligned} \operatorname{var}(T_* - \hat{\mu}_*) &= \operatorname{var}\left(\eta_* - \sum_{j=1}^{n} w_{*j} \omega_j \right), \\ &= \operatorname{var}(\eta_*) + \sum_{j=1}^{n} w_{*j}^2 \operatorname{var}(\omega_j). \end{aligned}$$

Hence

$$\operatorname{var}(T_* - \hat{\mu}_*) = \sigma_\eta^2 \left\{ 1 + \sum_{j=1}^{n} w_{*j}^2 K_j \right\}. \tag{6.4.4.4}$$

6.4.5 Estimation of the covariances of errors in generalised parameter values

Sections 6.4.3 and 6.4.4 have outlined how the uncertainty of generalised parameter values has been treated in this project, dealing with each parameter of the runoff model individually. No equivalent for the iterative procedures used for estimating the generalisation uncertainty has been found that works on a fully multivariate basis across several model parameters simultaneously. One difficulty here arises from the requirement to be able to use different sets of catchment properties in determining the generalisation procedure for each model parameter. The simplest example arises in the use of regression-based generalisation, where it is natural to try to choose which properties "should" appear in the regression for each parameter. There are conceptual difficulties to developing a unified theory which arise from the imposition within the model of the constraints of "knowing" that certain regression parameters are zero, in the case of regression models, and from whatever the equivalent of this is in the case of site-similarity / pooling-group methodologies.

Instead, a pragmatic approach has been taken to estimating the covariance matrix of the overall generalisation errors for the runoff model parameters. The outcome of this approach is such that variances estimated for individual model parameters are unchanged from those obtained by the methods outlined in Sections 6.4.3 and 6.4.4: the covariances are treated in a manner chosen primarily on the basis of guaranteeing that the estimates of the covariance matrices being generated are feasible covariance matrices, so that the step of generating randomised sets of model parameters can proceed.

The method for estimating the covariance matrix Σ_η of the generalisation errors in the model parameters for the same catchment is as follows. First the iterative re-weighting procedure is implemented for each parameter separately, yielding a set of values $\left\{\sigma_\eta^2(p); p=1,\ldots,P\right\}$ for the P parameters. The set of weighting values for each parameter is then available, given by

$$K_i(p) = 1 + \frac{\sigma_{i,\varepsilon}^2(p)}{\sigma_\eta^2(p)}, \tag{6.4.5.1}$$

where $\sigma_{i,\varepsilon}^2(p)$ is the estimated calibration variance for catchment i and parameter p. The weighted sum-of-squared-residuals that is used in the iterative procedure is replaced by a weight sum of cross-products of residuals defined to be

$$S(p,q) = \sum_{i=1}^{n} \frac{\left(Y_i(p) - \hat{Y}_i(p)\right)}{\sqrt{K_i(p)}} \frac{\left(Y_i(q) - \hat{Y}_i(q)\right)}{\sqrt{K_i(q)}} \tag{6.4.5.2}$$

where $Y_i(p), \hat{Y}_i(p)$ are the calibrated and generalised values of the pth parameter. Note that, when $p = q$, this expression is identical to the weighted sum-of-squared-residuals used in Section 6.4.3. Finally, the estimate for the covariance matrix Σ_η is defined to have elements given by

$$\Sigma_\eta(p,q) = \frac{S(p,q)}{\sqrt{n^*(p)n^*(q)}}, \tag{6.4.5.3}$$

where $n^*(p)$ is the "unbiasing factor" for the pth parameter as used in Section 6.4.3 and defined in Equation (6.4.3.10).

The covariances of the overall generalisation error can then be computed as follows. From Equation (6.4.4.3) the generalisation errors of the pth and qth parameters for the same catchment are

$$T_*(p) - \hat{\mu}_*(p) = \eta_*(p) - \sum_{j=1}^{n} w_{*j}(p)\omega_j(p), \tag{6.4.5.4}$$

and

$$T_*(q) - \hat{\mu}_*(q) = \eta_*(q) - \sum_{j=1}^{n} w_{*j}(q)\omega_j(q)\,. \tag{6.4.5.5}$$

Here $\{w_{*j}(p)\}$ and $\{w_{*j}(q)\}$ are the different sets of weights applied to the calibrated values of the *p*th and *q*th parameter of the runoff model. It then follows that

$$\begin{aligned}
&\operatorname{cov}\big(T_*(p) - \hat{\mu}_*(p), T_*(q) - \hat{\mu}_*(q)\big) \\
&\quad = \operatorname{cov}(\eta_*(p), \eta_*(q)) + \sum_{j=1}^{n} w_{*j}(p) w_{*j}(q) \operatorname{cov}\big(\omega_j(p), \omega_j(q)\big), \\
&\quad = \Sigma_\eta(p,q) + \sum_{j=1}^{n} w_{*j}(p) w_{*j}(q) \operatorname{cov}\big(\Sigma_\eta(p,q) + \Sigma_{j,\varepsilon}(p,q)\big),
\end{aligned} \tag{6.4.5.6}$$

where $\Sigma_{j,\varepsilon}(p,q)$ is the covariance of the calibration error for parameters *p* and *q* and catchment *j*.

If it becomes necessary to calculate the covariances between generalised parameters for different catchments, these can be found by the same type of approach. Thus, if subscripts * and # denote different catchments, neither of which are in the set of catchments on which the generalisation is based,

$$\begin{aligned}
&\operatorname{cov}\big(T_*(p) - \hat{\mu}_*(p), T_\#(q) - \hat{\mu}_\#(q)\big) \\
&\quad = \operatorname{cov}(\eta_*(p), \eta_\#(q)) + \sum_{j=1}^{n} w_{*j}(p) w_{\#j}(q) \operatorname{cov}\big(\omega_j(p), \omega_j(q)\big), \\
&\quad = \sum_{j=1}^{n} w_{*j}(p) w_{\#j}(q) \operatorname{cov}\big(\Sigma_\eta(p,q) + \Sigma_{j,\varepsilon}(p,q)\big).
\end{aligned} \tag{6.4.5.7}$$

6.5 Uncertainty of estimated parameters of runoff models

6.5.1 Representation of uncertainty where generalised estimates are used

As discussed earlier (Section 6.1), generalised estimates of model parameters would typically be used to supply information to a subsequent procedure which estimates the flow-frequency relationship for an ungauged catchment. Sophisticated analyses may involve the use of runoff models for several catchments. The most straightforward way of assessing the uncertainty in the final results from such analyses is to complement the results obtained for the “best estimates” of the model parameters with equivalent results obtained for sets of models parameters close to the best estimates but within a range determined by the uncertainty associated with the generalised estimates. If randomised sets of model parameters are generated to have the covariance structure indicated by the analyses of uncertainty outlined in Section 6.4, then the corresponding sets of flow-frequency curves will directly represent the uncertainty arising from generalisation of the model parameters.

In the present context, the use of multivariate normal distributions to represent uncertainty is most convenient. A detailed analysis of distributions representing uncertainty should logically follow an intensive study of the calibration results for individual catchments, comparing these with the generalised values of model parameters and seeking explanations for any large discrepancies, with the possibility of recalibration using amended procedures. There have not been resources for this within the present project. The use of multivariate normal distributions should be adequate because the main requirement is for a simple indication of the extent of uncertainty and no great reliance is placed on the shape of the corresponding distributions. However, several of the parameters of the runoff models being used are subject to natural or imposed constraints: for example, certain model parameters cannot be negative. These cases can reasonably be treated by truncating any parameter values that are generated from the multivariate normal distribution so as to lie within the required range.

The generation of multivariate normal random variates is a well-understood topic, so that no details of this are given here. For the present application, two candidate approaches for actual implementation arise. In the first of these, the set of parameters required across a number of catchments is determined, the covariance matrix of the overall generalisation errors is determined by expressions such as those in Section 6.4.5 and random values are generated corresponding to this covariance matrix. This approach seems be best suited to cases where only a few catchments are involved as target catchments. An alternative approach may be more suited to cases where many target catchments have to be treated simultaneously, since it avoids dealing with a covariance matrix with large dimensions. This alternative involves the representation of the overall generalisation errors in terms of their basic components. For example, based on Equation (6.4.5.4):

$$T_*(p) - \hat{\mu}_*(p) = \eta_*(p) - \sum_{j=1}^{n} w_{*j}(p)\omega_j(p), \tag{6.5.1.1}$$

it is possible to obtain the required randomised versions of overall generalisation error, simultaneously for all parameters and catchments, by generating random versions for the component errors $\{\eta_*(p)\}$ and $\{\omega_j(p)\}$ and then combining them using Equation (6.5.1.1). This may be relatively simple, given the assumptions that these components are uncorrelated between catchments and between the different types of error component.

6.5.2 Comparisons of sizes of uncertainty

Section 6.4.1 has outlined the model of uncertainty in terms of two sources of uncertainty, while Section 6.4.5 has shown how these can be combined together to determine the accuracy of the generalised estimates of parameters of the runoff models. The relative importance of the various contributions only becomes clear after the implementation of a generalisation procedure that takes account of the uncertainty model. This section presents some results that illustrate this relative importance.

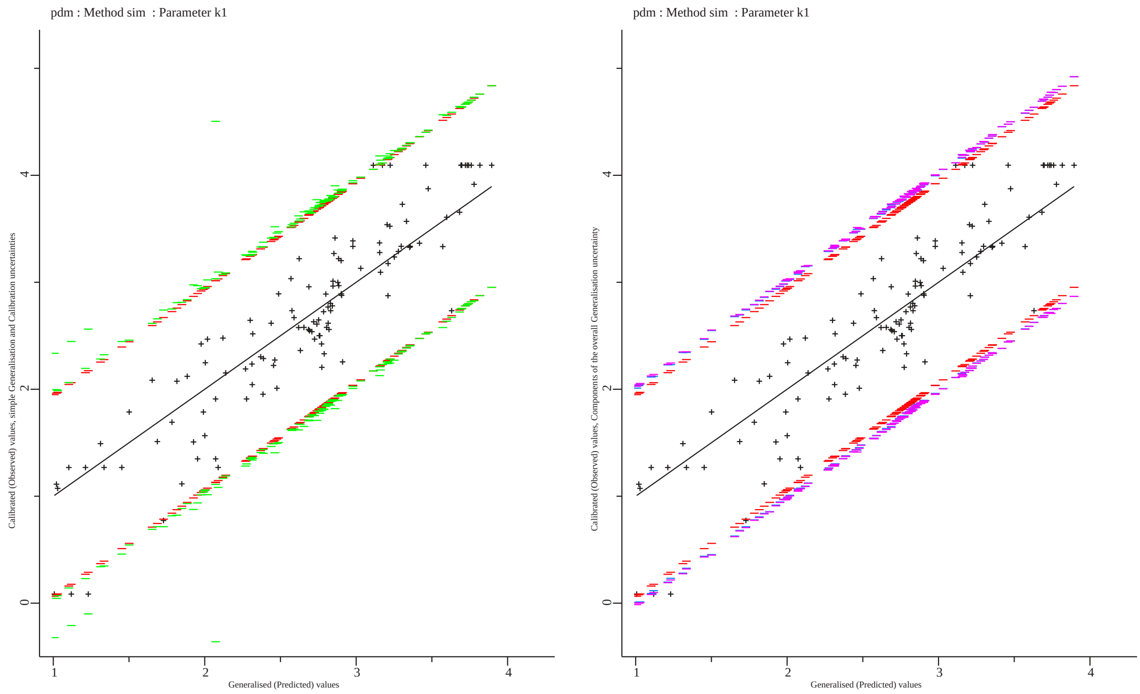
Section 5.1 has outlined generalisation procedures based on site-similarity and regression approaches and Section 5.4 has described how a final choice of structure has been made for these two approaches. For the purposes of comparing the uncertainty components associated with the two types of generalisation procedures, the analysis described here has been undertaken using model structures selected so that the same modelling steps have been included in both cases so as to ease the comparison. Figure 6.4 shows results for these chosen model structures in the case of one model parameter for one of the two runoff models. This case has been chosen arbitrarily as an example and the plots are shown in reduced form here: a set of figures for the complete collection of model parameters for both models is provided in Appendix C. Each figure in the appendix relates to a specific parameter of one of the models, with results from the site-similarity and regression approaches appearing on opposite pages for comparison: in Figure 6.4, the results for site-similarity and regression appear on the upper and lower portions of the page. Within each figure, a scatter plot of the calibrated model parameters against the generalised values is shown twice: the two scatterplots are shown with two different sets of error-bounds associated with: (i) the sizes of the basic error components; and, (ii) the sizes of the overall generalisation uncertainty. These error bounds are shown in the form of ± 2 standard deviation limits about the generalised values. Table 6.2 indicates how the various sets of bounds are determined: the "standard deviation" used for each set of error bounds corresponds to an estimated variance representing the effects of different combinations of errors. For convenience the bounds are drawn centred on the generalised values which here take the role of "estimates from the model".

Figure 6.4 and those in the appendix have been constructed so that the same scales are used for plots relating to the same parameter. In cases where the error-bounds are wide, the lower-limits of the bounds have sometimes been omitted in order to improve the appearance of the plots without loss of information. The following conclusions can be drawn from these plots.

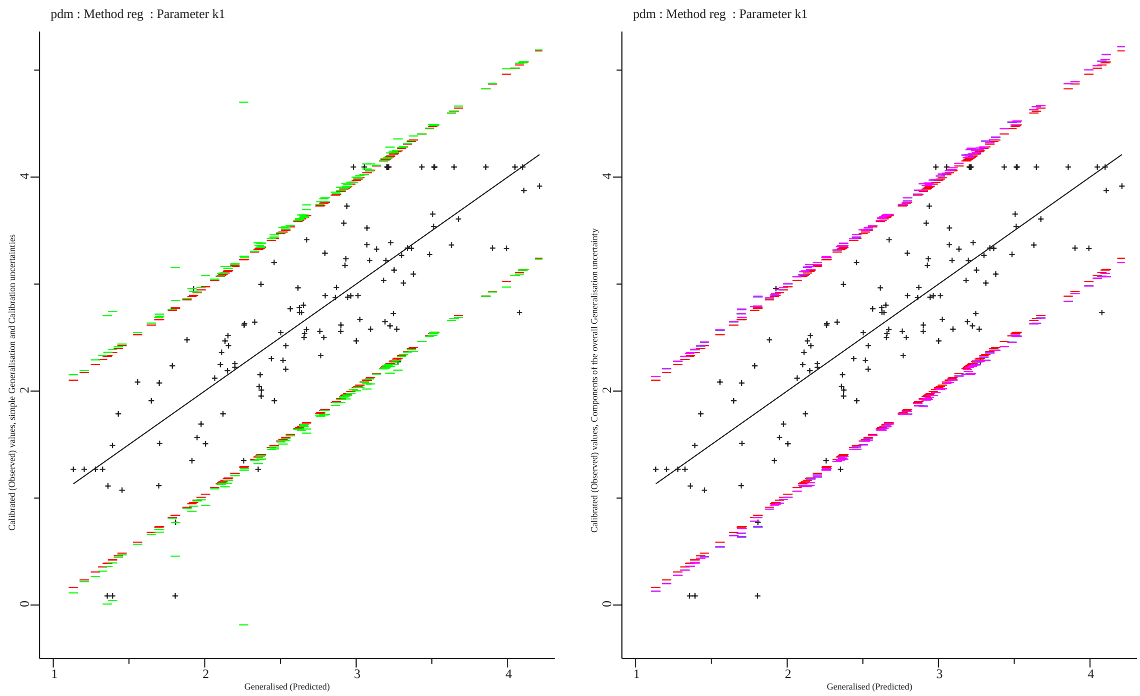
(a) The sizes of the variances of the calibration-error, $\sigma^2_{i,\varepsilon}$, vary considerably from catchment to catchment, and can be comparable to or greater than the variance of the generalisation error in some cases. However, for a reasonable number of catchments the variances of the calibration error is comparatively small.

(b) The scatter of the points in the plots relating calibrated to generalised values is typically smaller for the site-similarity approach than for the regression approach. However, the sizes estimated for the variances of the generalisation errors are more similar for the two approaches. These estimated variances are derived from the squared-errors using factors which are different and which are determined by the details of the generalisation procedures. Typically, the calibrated and generalised values for the site-similarity procedure are closer because the generalised value estimated for a given catchment gives relatively large weight to the calibrated value for that catchment: the factors used to convert the squared-errors to estimates of variance adjust for this preferential weighting.

(c) The component in the total generalisation error related to errors transferred from other catchments is fairly modest, but is larger for the site-similarity approach than for the regression approach. This arises because the generalised values in the site-similarity approach are effectively averages over a smaller number of catchments.



(a)



(b)

Figure 6.4 Calibrated and generalised parameter values, with uncertainty bounds: Parameter "k_1" of the PDM model. See text and Table 6.2 for a description. (a) generalisation by site-similarity; (b) generalisation by regression.

Table 6.2 Composition of the error bounds shown in Figure 6.4 and in Figures C.1 to C.14 of Appendix C

Colour	Position	Description	Formula
Left-hand scatter plots:			
Red	Inner-most	Basic generalisation error	σ_η^2
Green	Outer-most	Generalisation variance plus the calibration error variance	$\sigma_\eta^2 + \sigma_{i,\varepsilon}^2$
Right-hand scatter plots:			
Red	Inner-most	Basic generalisation error	σ_η^2
Pale blue*	Next	Generalisation error variance plus the transferred effect of the generalisation error for sites used in the generalisation procedure	$\sigma_\eta^2 + \sum_{j=1}^{n} w_{*j}^2 \sigma_\eta^2$
Purple	Outer-most	Total error – generalisation error variance plus the transferred effects of generalisation error and calibration errors for sites used in the generalisation procedure	$\sigma_\eta^2 + \sum_{j=1}^{n} w_{*j}^2 \left(\sigma_\eta^2 + \sigma_{j,\varepsilon}^2\right)$

* Note that this bound is often obscured in the figures by the outer-most bound

(d) The variance component relating to transferred calibration errors is typically small (the pale blue and purple marks are very close). This suggests that it may be reasonable to adopt an approximation for the variance of the overall generalisation error in which this term is averaged. The potential advantage of such a scheme is only apparent when all the parameters of a model are considered jointly: the full scheme would require keeping track of the covariance matrices of the calibration errors for all the calibration catchments, while a reduced scheme would require only the variances. Such approximations for the overall generalisation variances have not been considered at this stage.

(e) The jackknife procedure for estimating the variances of the calibration error can sometimes produce unrealistically large variances: this is particularly apparent in cases where the calibration scheme includes placing restriction on the range of values for particular model parameters. One reason for this effect is the scaling-up that occurs, in the jackknife procedure, from the variation between the leave-one-out estimates to the implied sampling variation of the calibration estimates. As discussed in Section 6.4.2, bootstrap procedures provide a possible alternative way of estimating calibration variances.

6.5.3 Uncertainty modelling for the sequential regression approach

The structure for a suitable model for uncertainty for use with the sequential regression method needs to be extended beyond that set-out in Section 6.3.1. Because, overall, the sequential regression method has not been found to perform as well as the other generalisation methods, such an extended framework has not yet been fully formulated. No results for the uncertainty associated with the sequential regression approach are presented in this report.

In order to frame a model for uncertainty, it is necessary to have a clear specification for the results of the sequential regression methodology, treated as a stand-alone procedure. In a formal sense the "result" is a set of regression coefficients, $\left\{\hat{\beta}_{seq}(p), p=1,\ldots,P\right\}$, by which the parameters of the runoff model are estimated for an ungauged catchment:

$$\hat{\mu}_{seq,*}(p) = x^T_{seq,p,*}\ \hat{\beta}_{seq}(p)\,.$$

Here $x_{seq,p,*}$ denotes the set of explanatory variables (catchment properties) chosen for parameter p, evaluated for target catchment *. However, these regression coefficients do not provide enough information for an uncertainty model. Instead it is necessary to consider the set of calibrated values, $\left\{Y_{seq,i}\right\}$, produced by the sequential regression approach. It is important to take on board the fact that the nature of this set of calibration results is rather different from that of the set of values $\left\{Y_i\right\}$ that arise from the ordinary calibration procedure. In particular, the meanings of these two sets may be described as follows. Let $\left\{\alpha(1),...,\alpha(P)\right\}$ be a generic set of parameters of the runoff models for the particular catchment for which i is the label. Then:

(i) taken together $\left\{\alpha(1)=Y_i(1),...,\alpha(P)=Y_i(P)\right\}$ is an optimal set of parameter values for catchment i;

(ii) for any given parameter labelled p, $\alpha(p)=Y_{seq,i}(p)$ is an optimal value of the parameter p when the other parameters are set to their generalised values $\alpha(1)=\hat{\mu}_{seq,i}(1),...,\alpha(p-1)=\hat{\mu}_{seq,i}(p-1),\alpha(p+1)=\hat{\mu}_{seq,i}(p+1),\ldots,\alpha(P)=\hat{\mu}_{seq,i}(P)$.

There is nothing in the sequential calibration procedure that can ensure that the set of values $\left\{\alpha(1)=Y_{seq,i}(1),...,\alpha(P)=Y_{seq,i}(P)\right\}$ provides a good set of parameters for catchment i: this combination of parameter values is not even tried specifically within the overall sequential regression calibration procedure.

The discussion on Section 6.3.1 introduced the idea of "true" values for various quantities. In the case of the generalised values, the role played by μ_i in the earlier section is now taken by the vector $\mu_{seq,i}$, consisting of elements $\left\{\mu_{seq,i}(p)\right\}$ which are defined as the limiting values of $\left\{\hat{\mu}_{seq,i}(p)\right\}$ as the number of catchments available for fitting the generalisation rule increases, along with the lengths of records for each of these. However, the definition of a "true" set of parameter values for a given catchment is more problematic. In the present context this "true" set of values needs to represent the aspiration that the generalisation rule estimates a single set of parameter values which would give good performance of the runoff model if flow data were available for the ungauged site. Thus the "true" values for a given catchment cannot be defined in terms of the limiting behaviour of $Y_{seq,i}$ as the amount of data for the catchment increases. Depending on the approach taken, this leads to one of the following:

(i) a single vector defined as the convergence point of a set of vectors $Y_{seq,i}$, none of which has any claim to provide good performance for the given catchment and thus the convergence point cannot be considered "good";

(ii) a set of P vectors, each of which consists of the limit obtained by setting all but one parameter to the elements of $\mu_{seq,i}$ and optimising the other using an increasing amount of data.

Thus neither approach leads to a definition of a "true" value which notionally gives best-possible performance for the target catchment. It therefore seems that the "true" catchment value needs to be defined in terms of the limiting values of the separate-catchment calibration results $\{Y_i\}$, and this leads on to trying to define a viable uncertainty analysis based on simultaneous consideration of both sets of results $\{Y_{seq,i}\}$ and $\{Y_i\}$.

Figure 6.5 shows a schematic of the basic quantities that would be involved in an uncertainty model for dealing with generalisation by sequential regression for a model with two parameters. Here it is assumed that such a methodology would need to consider both sequential regression and ordinary generalisation results. In addition to these basic quantities it would be necessary to include a set of error components inter-connecting the basic quantities. It is clear that this set would need to be extended beyond that used for analysing the simpler generalisation methods. The specification of the uncertainty model would need to consider the correlation between the different sets of error components: it is certainly not clear that the assumption of no-correlation can be justified for all pairs of components. One place where particular conceptual difficulty arises is the relationship between the two different versions of "true" generalisation values μ_i and $\mu_{seq,i}$. It does not immediately seem possible to argue that these quantities are identical and, if they are not, then some representation of the difference is needed: it may be necessary to allow for some type of bias, possibly depending on catchment properties, as well as a possible random component in the difference. If two separate vector quantities are needed, the relationships of the true catchment values, T_j, to the two true generalisation values inevitably leads to dependence in the corresponding two generalisation errors.

If a well-formulated model for the uncertainty associated with the sequential regression method of generalisation is to be constructed, it seems necessary to gain considerable experience with the empirical results of this approach, jointly with those the ordinary regression approach, so as to help decide on a suitable overall structure for the model.

6.5.4 Fully-multivariate generalisation approaches

As remarked in Section 6.3.3, the site-similarity and regression generalisation methods are implemented by performing separate generalisation analyses for each parameter, and these methods can therefore be considered as multiple-univariate in nature. An important feature of these methods is that separate choices are made for each parameter of the catchment properties that are used within the procedure.

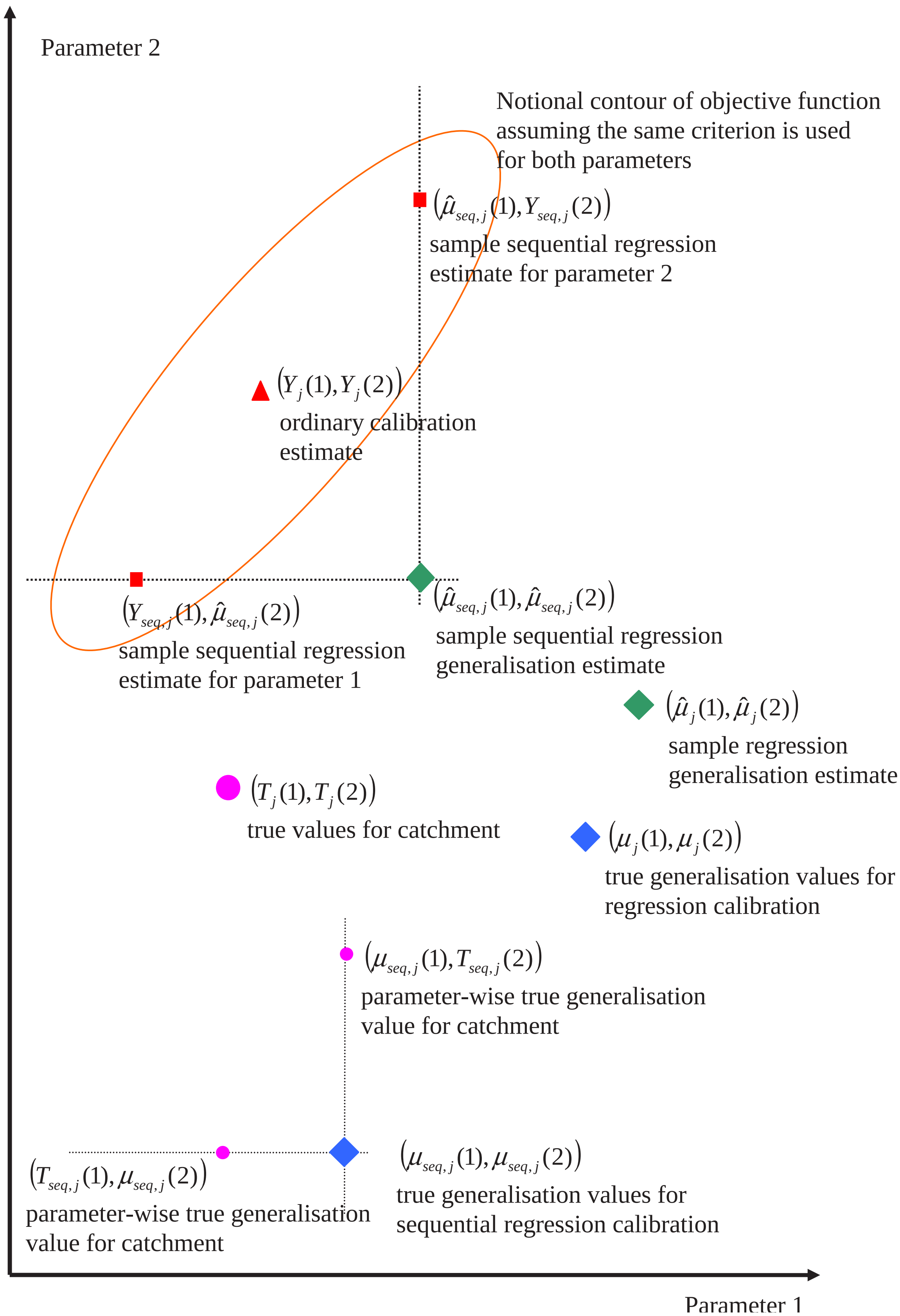


Figure 6.5 Schematic of quantities involved in an uncertainty model for generalisation using sequential regression

In principle, improved modelling procedures are available if fully-multivariate statistical modelling methodologies are applied. However, any such improvement is largely illusory since reliance would need to be placed on assumptions which cannot be justified.

The main candidate for potential improvement is the use of multivariate rather than univariate regression. The basic structure for this model is

$$Y = XB + \varepsilon \,, \qquad (6.5.4.1)$$

where Y is now a matrix of calibrated parameter values consisting of n rows, corresponding to the n catchments used for fitting the generalisation, and P columns, corresponding to the P parameters of the runoff model. As usual, X is a matrix of values of the catchment properties used for the generalisation, having one row for each catchment. The matrix B is a matrix of regression coefficients having P columns: each column corresponds to the usual vector of regression coefficients in the corresponding univariate model. The error term ε is an $n \times P$ matrix of random variables which, in the cases discussed here, are assumed to be uncorrelated between different rows: correlation within the same row is allowed.

In the case that the univariate regressions for all parameters of the runoff model all estimate regression-parameters for exactly the same set of catchment properties, the estimates obtained from separate univariate regression and multivariate regression are identical, given that the cross-parameter covariances of the multivariate parameters are assumed unknown. This result of identical results extends to the results of estimating the covariance matrix of the regression errors from either the multiple-univariate approach or the multivariate approach. Given that the covariance matrix of the errors will always be treated as fully-unknown and to be estimated, the main situation in which the multivariate regression approach can provide improved model performance is when the separate univariate regression models do not all contain exactly the same set of catchment properties. This is readily achieved with separate regressions and, in the case of multivariate regression, is included within the general form of model in Equation (6.5.4.1) by imposing constraints on the elements of the regression-coefficient matrix B, that certain elements are known to be zero. In this situation estimates for the non-zero elements of the matrix B can be obtained from the multivariate model which have improved performance compared to the separate-regression results, in terms of both the estimation variance for the regression coefficients themselves and the predictive performance of the fitted regression estimator for a new catchment. Imposing zero-coefficients in the regression estimates allows certain information to be transferred between the regressions for the individual parameters, which allows improved estimates to be constructed.

Besides the complexity of implementing the methodology, the main reason for not using multivariate regression within the present project is that improvement supposedly gained over applying multiple-univariate regression arises by relying on certain assumptions which can not be justified: the importance of the exact truth of these assumptions within the methodology is unknown. The assumption is that it is known that particular (“true”/population) regression coefficients are zero. Of course this seems to be exactly the same assumption being made when a choice of regression-variables is being made in univariate regression and certain variables omitted. However, there is a radically different stature being given to the assumption. In univariate regression, model-assessment procedures effectively allow the conclusion that not much is lost if certain coefficients are set to zero, given considerations about the extra estimation errors introduced by including unnecessary variables. In multivariate regression, the assumption that a regression coefficient is zero is used in a complicated way to change the estimates of all the other regression coefficients, extending to all the dependent variables. The resulting improved estimates depend strongly on the assumption that certain regression coefficients are zero.

The way in which improved estimates are obtained in multivariate regression may be illustrated as follows. Suppose that a single sample of variables (U,V) is obtained, and that the population mean value of U has a known value μ_U. Then the raw estimate of V for the mean value of the population mean value of the Vs can be improved by considering

$$\hat{\mu}_V(b) = V - b(U - \mu_U)\,. \tag{6.5.4.2}$$

This is clearly unbiased for all values of b, and the variance is minimised by setting

$$b = \frac{\sigma_{UV}}{\sigma_{UU}}$$

and the estimation variance obtained for this value is

$$\operatorname{var}\{\hat{\mu}_V(b)\} = \sigma_{VV}\left(1-\rho_{UV}^2\right)\,, \qquad \text{where } \rho_{UV} = \frac{\sigma_{UV}}{\sqrt{\sigma_{UU}\sigma_{VV}}}\,. \tag{6.5.4.3}$$

Here $\sigma_{UU}, \sigma_{UV}, \sigma_{VV}$ are the variances and covariances of (U,V). The relevance of this to imposing zero-constraints in the multivariate regression model can be seen by noting that the raw estimate of the regression coefficient matrix, $\hat{B}$, given by

$$\hat{B} = \left(X^{\mathrm{T}}X\right)^{-1}X^{\mathrm{T}}Y\,, \tag{6.5.4.4}$$

can be arranged into a set of components (Us) whose true mean values are assumed to be zero, and another set (Vs) whose mean values are to be estimated. The actual estimation method is more complicated than this argument indicates because of the need to estimate the values on which the adjustment depends. Some computational methods for fitting multivariate models in which regression coefficients are forced to be zero proceed by reformulating the model in Equation (6.5.4.1) in a vector form, which then involves working with a structured covariance matrix for new vector of model errors. Pollock (1979, Chapter 13) provides some details of this. Zellner (1971, Chapter 8) discusses multivariate regression models mainly in a Bayesian context, although formula for point estimates are also given. The work by both these authors originated in an econometric context, where the terminology "seemingly unrelated regressions" is used, emphasising that there need be no common explanatory variables across the regression models used for the different quantities being treated.

While it may be possible to argue on physical grounds that certain parameters of runoff models should be more highly related to certain geographically derived variables, or that the relationships with particular variables are likely to be positively or negatively dependent, the assumption of zero-regression coefficients is more problematic. Strictly, the consideration needs to take account of the other catchment properties that are to be included in the multivariate model. It seems unwise to base an estimation procedure on such an uncertain assumption.

6.6 Uncertainty of estimated flood frequency curves

6.6.1 Derivation of uncertainty bounds

Uncertainty bounds were produced around each of the (univariate regression and site-similarity) generalised flood frequency curves for each catchment, by using the methods described above to generate a large number of 'generalised parameter sets with uncertainty'. The runoff model is then run with each of these parameter sets to produce time series of flows, flood frequency curves are derived from each flow series, and the set of flood frequency curves produced is used to estimate bounds. As in Chapter 5, the flood frequency curves here are calculated using the observed record of catchment rainfall for each catchment.

A collection of 1000 parameter sets were used. The uncertainty bounds were estimated at each plotted return period by ranking the 1000 estimated peak flows at the given return period (lowest to highest) and selecting ranked points appropriate to the bounds required. That is, for central Z% uncertainty bounds (where (100-Z)/2% fall above the upper bound and (100-Z)/2% below the lower bound), the 1000*((100-Z)/2)/100th and 1000-1000*((100-Z)/2)/100th points were selected, for the lower and upper bounds respectively. The corresponding points at each return period are joined up, to produce continuous bounds on the flood frequency curve. The 90, 95 and 99% bounds were calculated, to better illustrate the spread and asymmetry of the bounds.

Figure 6.6 shows examples of uncertainty bounds on the generalised flood frequency curves, for two catchments. The bounds for all of the catchments, using each of the two generalisation methods and both of the two runoff models, are given in Appendix D.

6.6.2 Analysis of uncertainty bounds

As for the generalised flood frequency curves themselves, the uncertainty bounds vary by catchment, generalisation method and runoff model. Table 6.3 summarises the performance of the bounds according to whether the flood frequency curve from observed flows lies completely within each of the bounds ('within'), lies outside an outer (99%) bound somewhere ('outside'), or otherwise (that is, lies outside a 90 or 95% bound somewhere, but not outside either 99% bound). This again suggests that PDM site-similarity performs best overall, with the highest number of catchments classified as 'within' and the lowest number classified as 'outside'. However, a closer look at the bounds suggests that those from PDM site-similarity are also the widest for a large

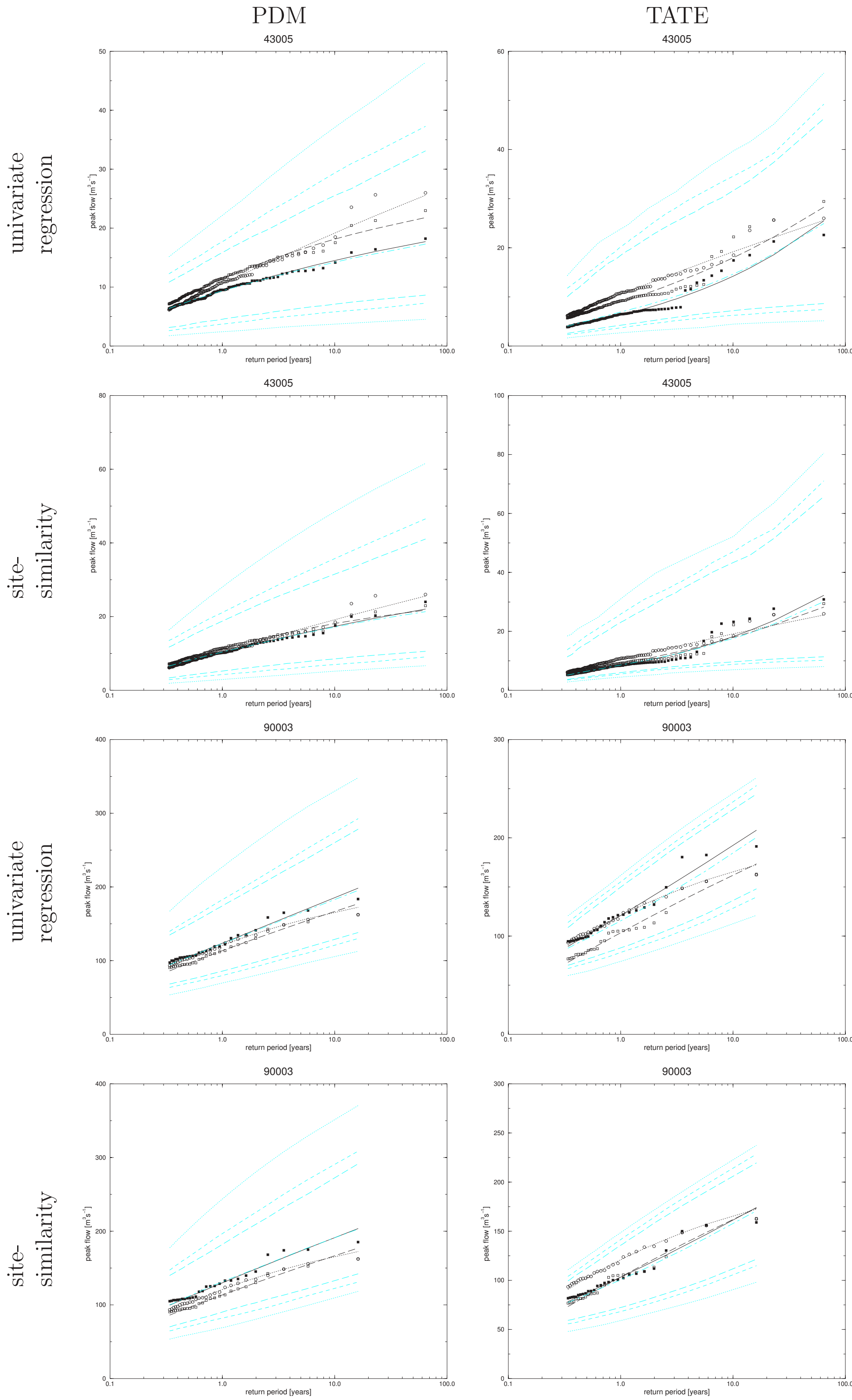


Figure 6.6 **Example of uncertainty bounds (blue; 50% - dot/dash, 90% - long dashed, 95% - short dashed, 99% - dotted) on generalised flood frequency curves (black solid), for two catchments.**

Table 6.3 Summary of performance of uncertainty bounds, by the number of catchments for which the observed flood frequency lies a) totally within the uncertainty bounds, b) outside an outer bound somewhere or c) otherwise.

Classification	PDM		TATE	
	univariate regression	site-similarity	univariate regression	site-similarity
Within all bounds (90, 95 and 99%)	86	103	77	80
Outside an outer (99%) bound somewhere	9	3	17	17
Otherwise	24	13	25	22

number of catchments (about 60%), so it is not surprising that the observed flood frequency curve is more likely to lie completely within the bounds. These impressions are studied in more detail in the analysis that follows.

Table 6.4 shows some results based on the widths of the uncertainty bounds, and these can also be used to summarise the more general properties of the statistical distributions reflecting the uncertainty of the generalised estimates of the flood frequency curves. For a given percentage level, say 95%, and a given return period, say 10 years, a relative width for the upper bound is defined as the ratio of the flood value at the upper 95% bound to the flood value at the 50% point for the same return period. This value is calculated for each catchment and averaged across catchments. Similarly, a relative width for the lower bound is defined as the ratio of the flood value at the 50% point to the flood value at the lower 95% bound for the same return period. Again these values are calculated for each catchment, and averaged. When defined in this way the relative widths have values greater than one and can be interpreted as factors to multiply and divide by to create upper and lower bounds about a central estimate. These are summary measures only and, for actual applications, it will be better to use the catchment-specific bounds derived separately for each target catchment.

The results in Table 6.4 show that the upper uncertainty bounds derived for the PDM model are generally wider, in a relative sense, than those for the TATE, while the opposite is true for the lower uncertainty bounds. For the PDM, the uncertainty bounds associated with site-similarity generalisation are slightly wider than those for generalisation by regression, while the opposite is true for the TATE.

A comparison between the relative widths of the upper and lower uncertainty bounds indicates rather different behaviour for the distributions describing the uncertainty for the PDM and TATE models. The plots in Figure 6.6 and in Appendix D show that these distributions are skewed towards the right when considered straightforwardly in terms of flow. Table 6.4 shows that, for the PDM model, the distributions are close to symmetrical when judged in a logarithmic sense (the relative widths for the upper and lower bounds are reasonably similar), while those for the TATE model are skewed to the left.

Table 6.4 Mean relative widths of uncertainty bounds.

Return period	Bounds	PDM		TATE	
		univariate regression	site-similarity	univariate regression	site-similarity
10-years					
	99% upper	2.05	2.12	1.80	1.77
	99% lower	2.22	2.23	3.12	2.41
	95% upper	1.71	1.75	1.58	1.57
	95% lower	1.74	1.80	1.95	1.85
	90% upper	1.56	1.60	1.49	1.47
	90% lower	1.58	1.62	1.68	1.64
50-years					
	99% upper	2.17	2.23	1.81	1.75
	99% lower	2.26	2.28	3.27	2.52
	95% upper	1.79	1.83	1.61	1.55
	95% lower	1.78	1.83	2.01	1.93
	90% upper	1.62	1.66	1.50	1.46
	90% lower	1.61	1.65	1.73	1.69

The difference between the models here may well be related to the different role played by the parameters for the two models: in particular the PDM model has the parameter f_c which acts multiplicatively on the rainfall before other modelling-elements apply, whereas the closest corresponding parameter with TATE acts once rainfall has been divided into different types of runoff. A general summary of the uncertainty bounds for the two models is that the upper uncertainty bounds are wider for the PDM than for the TATE, but the lower uncertainty bounds are wider for the TATE than the PDM.

7 COMPARISON WITH FLOOD ESTIMATION HANDBOOK METHODS

Sue Crooks

PREVIOUS FAST TRACK BOX ON PAGE 77

CHAPTER 7 COMPARISON WITH FLOOD ESTIMATION HANDBOOK METHODS

It is appropriate to offer a comparison of approaches of the new continuous simulation methods with the current event-based recommended practice of the 1999 Flood Estimation Handbook (FEH), which provided a new statistical flood peak approach, and the 2005 updates to the 1975 Flood Studies Report unit hydrograph, or runoff, approach to frequency estimation.

This chapter therefore compares the two approaches with respect to principles, data requirements, handling of particular hydrological issues and what is known about general error levels. It is recommended that a full numerical comparison is undertaken in the future with the aim of producing guidance as to under what circumstances advantages may accrue from the use of each approach. Currently, more is plainly known about practical use of the FEH but more work has been undertaken on the quantification of errors and uncertainties of the continuous simulation approach as detailed in this report.

NEXT FAST TRACK BOX ON PAGE 119

7.1 Introduction

The continuous simulation method of flood frequency estimation has been developed as a possible next-generation approach following the Flood Studies Report (FSR), (NERC, 1975 and Supplementary Reports) and its successor, the Flood Estimation Handbook (FEH), (Institute of Hydrology, 1999). The purpose of this chapter is to look at the concepts contained within the FEH and continuous simulation (CS) methods which are relevant to their application and usage and which are discussed through a number of key points. No quantitative testing has been undertaken. A brief summary is given first of the FEH methods of flood frequency estimation to facilitate comparisons between the two approaches. Definitions of catchment properties used in the text (e.g. AREA) are given in Table 2.3.

7.2 FEH methods

FEH comprises two main routes for estimating flood frequency, applicable to gauged and ungauged catchments, with choices for determining a particular route based on:

- purpose of flood frequency estimation;
- availability of observed data;

- catchment properties, notably area, permeability, urban extent and catchment storage (reservoirs and lakes).

The two routes are a runoff method originally developed as part of the FSR and statistical analysis of observed peak flows. In all situations the use of observed data within application of FEH methods is emphasised, whether for the target catchment or for donor or analogue catchments. The importance of hydrological judgement is also a factor in use of the FEH. The main features of each method are summarised below.

7.2.1 Runoff

The FSR runoff method convolves a derived unit hydrograph with design storm rainfall with the addition of baseflow to calculate a flood hydrograph of related return period. The method is, thus, event based and produces design flood hydrographs for specified return periods. The method has recently been revitalised (Kjeldsen *et al.*, 2005) partly to take advantage of extended datasets containing larger flood events and partly to improve the representation of hydrological concepts through updated analytical techniques. The revitalisation work has involved the development of the Revitalised Flood Hydrograph (ReFH) model with three new formulations for the loss, routing and baseflow components. The loss model is based on the PDM (Probability Distributed Model; see Chapter 3) assuming a uniform distribution of soil moisture capacity; the routing model retains the concept of a standard instantaneous unit hydrograph scaled to individual catchments but allows a more flexible shape; and the baseflow model allows for change during the flood event.

The ReFH model has four parameters, one for the losses model, one for the routing model and two for the baseflow model, and two boundary conditions, the initial soil moisture and baseflow conditions. The values for the parameters and boundary conditions are estimated from observed event data for a gauged catchment or catchment properties for an ungauged catchment. Nine of the eleven FEH catchment properties listed in Table 2.3 (not ALTBAR and DPLCV) have been used to derive multiple linear regression equations for the parameters and boundary conditions.

The design storm rainfall is calculated from a depth-duration-frequency (DDF) model, based on annual maximum rainfall data, which has six parameters which have been determined for all points on a 1 km grid across the UK (Institute of Hydrology 1999, vol.2). Catchment average sets of parameters are derived as the weighted average of point values which are then adjusted by multiplying by an areal reduction factor, incorporating two further parameters, both functions of catchment area. The design storm rainfall is distributed over a design storm profile which is symmetrical, single peaked and invariant with duration and location. A distinction in profile is made between rural and urban catchments ($0.125 \leq \text{URBEXT} \leq 0.5$) which use winter and summer profiles, respectively.

Additional features of the revitalised runoff method are the application of a seasonal correction factor (summer/winter) to the design rainfall and the assumption that the T-year flood is generated by the T-year rainfall.

7.2.2 Statistical

The statistical method estimates a flood peak of a specified return period, T, as the product of an index flood, QMED, the median annual maximum flood, and a growth factor, where the latter is the ratio between the T-year flood peak and QMED. A growth curve relates growth factors to return period for a range of return periods. QMED and statistical parameters describing the growth curve are either determined from observed data or estimated from a multiple regression equation using six catchment properties for QMED and the use of a pooling group for the growth curve. The pooling group is assembled using observed data from hydrologically similar catchments determined by distance, in terms of AREA, SAAR and BFIHOST, from the subject catchment to give a combined record length of at least 5T years, where T is the target return period. Hydrological judgement is required to review, and possibly revise, the catchments contributing to the pooling group. The component procedures have been automated (Morris, 2003) to enable estimates of flood peaks to be determined for almost any site on the UK river network for return periods between 2 and 200 years.

7.3 Comparison issues

7.3.1 Ease of use

FEH. The FSR/FEH approach to flood frequency has been widely used by practitioners over the last 30 years and has become the standard method within the water industry. It comprises a suite of well defined and tested methods for estimating flood frequency for almost any site on a UK river, whether gauged or ungauged. In response to widespread feedback there are now many choices to make, depending on the purpose of the flood estimates, the availability of observed data and the type of catchment. The FEH approach, although more comprehensive, has become more complex to use than the FSR method, requiring a certain level of hydrological expertise/judgement in its implementation. There is scope for generating different flood frequency estimates depending on the user.

CS. The aim in developing the method has been to provide one overall inclusive method, probably requiring less hydrological knowledge from the user. For this to be a reality it requires all necessary software to be available to a user – for input climatic data (rainfall and PE at appropriate spatial and temporal timescales), runoff model, generation of model parameters, flood peak/event extraction and statistical analysis, uncertainty calculation and presentation of results.

7.3.2 Hydrological representation

CS and the FEH runoff method both use simplified mathematical representations of hydrological processes controlling the relationship between rainfall and runoff to generate flood frequency estimates. One of the main aims of such modelling is to achieve a balance between the level of simplification, availability of input data, observed or generalised, and the accuracy of the modelled flows. Theoretically, the more processes represented and linked to a catchment through calibrated model parameters, the greater the model accuracy. However, in reality, increase in model complexity does not necessarily increase the level of model performance, indeed the opposite may be the case, particularly when the parameters are estimated through generalisation procedures.

FEH. The revitalisation of the FEH runoff method, through the development of the ReFH model, was instigated partly to provide a more physically based procedure for flood modelling. The loss model incorporated within ReFH is based on that of the PDM and the baseflow model now includes a linear reservoir store so, in concept and model parameters, the FEH runoff method has moved towards a continuous simulation approach. Surface runoff and baseflow are modelled separately which requires estimates of initial boundary conditions for soil moisture and baseflow. The revitalisation has increased the number of parameters to be estimated from three in the original runoff method to four in ReFH, plus the two initial boundary conditions. Design rainfall is estimated through a parameterised depth-duration-frequency model.

CS. Use of continuous simulation for flood frequency combines input of rainfall and representation of relevant hydrological processes in an overall model. Continuous accounting of water movement removes the need for separation of surface runoff and baseflow, and joint probability problems between rainfall events and antecedent soil moisture are automatically handled in the simulation. A deliberate policy of the continuous simulation approach to flood frequency has been to use parameter-sparse versions of the PDM and TATE models which are likely to result in more effective spatial generalisation.

7.3.3 Catchment properties

Many catchment properties (see Section 2.6) contribute to the relationship between rainfall and runoff which varies both spatially and temporally. Four of these, which have a particular impact on flood response, are discussed below. Measures of catchment properties are used to represent or determine similarity of catchments within generalisation procedures. However, it should be noted that similar values may not result in similar response characteristics due to differences in spatial distribution and other factors. The properties URBEXT and FARL are of particular note in this context.

- **Area**
  FEH. The minimum area using either method is 0.5 km^2 and applicable to smaller catchments only where sufficient gauged data are available. A maximum of 500 km^2 to 1000 km^2 is recommended using the runoff method, depending on the applicability of a catchment wide design storm. Flood frequency estimates for larger catchments should use the statistical method.

  CS. Catchment areas used in development of the method range from 0.9 km^2 to 1256 km^2 (see Figure 2.2). Calibration and generalisation performance were shown to be not affected by catchment area (Figures 4.7, 4.8 and Appendix B). A lumped runoff model is not recommended for very large catchments (> 1500 km^2) due to the lack of allowance for spatial variability of rainfall and catchment properties.

- **Permeability**

Highly permeable catchments present particular problems for many simple modelling systems, including the determination of contributing groundwater catchment boundaries from digital terrain models.

FEH. A special permeable-catchment variation of the statistical method is recommended when the value of SPRHOST is less than 20%, to allow for the possibility of no flood peaks in dry years. This variation is not currently included in the automated version.

CS. Modelling of groundwater/baseflow is automatically included within the continuous simulation method. Calibration and generalisation performance across the range of catchment values of BFIHOST and SPRHOST are shown in Figures 4.7, 4.8 and Appendix B. The TATE model performs less well for the site-similarity generalisation method with high values of BFIHOST (low values of SPRHOST), representative of permeable catchments.

- **Urban**

FEH. In the statistical method a flood frequency estimate is determined as for a rural catchment and then an urban adjustment factor, related to the value of URBEXT, is applied both to QMED (the median discharge) and the growth curve (applied when URBEXT > 0.025). In the runoff method allowances for urbanisation are incorporated in the procedures. Neither method is recommended when URBEXT exceeds 0.5 (tURBEXT > 0.707). Care is required when using urban catchments as donor or analogue sites and such catchments should not be used in pooling groups.

CS. Allowance for urbanisation has been made by ensuring that URBEXT was included in relevant regression equations or site-similarity groups. However, only six catchments have a value of URBEXT greater than 0.1, two greater than 0.3 and none greater than 0.5, which is too small a sample to fully assess the reliability of flood frequency estimates for medium to heavily urbanised catchments.

- **Reservoirs and lakes**

FEH. The runoff method is recommended for reservoir flood estimation and for taking account of the attenuating effect of reservoir or lake storage for ungauged sites (FARL < 0.9). For gauged sites with FARL less than 0.9 (tFARL > 0.32) the statistical method can be used provided the impact of storage has been the same throughout the period of record. Care is required when using reservoired catchments as donor or analogue sites or when used in pooling groups to ensure that similar values of FARL apply.

CS. Allowance for reservoir/lake storage has been made by ensuring that FARL was included in relevant regression equations or site-similarity groups. However, only eight catchments have a value of FARL less than 0.9, with three less than 0.85 which is too small a sample to fully assess the reliability of flood frequency estimates for reservoired catchments or those with considerable natural storage of surface water.

7.3.4 Output

FEH. Output from the runoff method is an event hydrograph of stated return period determined by the convolution of a design storm rainfall with a unit hydrograph. The generation of design hydrographs is required for particular problems such as reservoir flood estimation and storage routing. Output from the statistical method is normally a flood frequency curve for peak flow to cover a range of return periods. Both methods are limited to providing information for catchments treated individually.

CS. The method produces a complete time series of simulated flows from which, for flood frequency estimation, the highest peaks are extracted for statistical analysis. Potentially, the simulated flows can also be used for analysis of hydrograph shape and durations of flows provided the model calibration is applicable to the complete flow range. In addition, if simultaneous rainfall can be simulated realistically for several catchments within a drainage network, flows can be simulated for these catchments and analysed jointly.

7.3.5 Uncertainty

All methods for flood frequency estimation incur a certain amount of uncertainty, whether in the use of observed data or from modelling and generalisation procedures. Lack of estimates of uncertainty does not imply accuracy of method.

FEH. The FEH suggests that a gauged record twice as long as the target return period is required to be confident that a statistical analysis of flood peaks provides a good estimate of the true flood frequency. Single site gauged records are generally not long enough to provide estimates for likely target return periods, hence the use of pooling-groups, but donor or analogue catchments are never exact replicas. Long data records are also likely to be affected by changes within the catchment so that the record is not stationary. Formal estimates of uncertainty for FEH statistical estimates are still under development but users can explore measures of uncertainty through resampling and other techniques.

CS. The continuous simulation method has been developed with the aim of incorporating measures of uncertainty into the formulation of the calibration and generalisation methods both to enable overall uncertainty in the final flood estimates to be quantified but also to minimise levels of uncertainty throughout the procedure. Although measures of uncertainty may appear quite large, see Figure 5.6 and Appendix D, it is a strength of the CS method that sources of uncertainty have been quantified.

7.3.6 Return period

FEH. The FEH states that the runoff method can be used for return periods between 2 and 2000 years, and with caution can be extrapolated to 10000 years, with the statistical method appropriate for return periods between 2 and 200 years. Estimates for high return periods are based on extrapolation of statistical analyses of observed data – the depth-duration-frequency of rainfall for the runoff method and pooled flow data for the statistical method.

CS. The CS method relies on the availability of suitable long-term rainfall and PE data at an appropriate timescale for the target catchment. While some long daily rainfall records are available, hourly continuous rainfall records are rarely longer than 20 years and neither provides a UK wide distribution. It is proposed that generated rainfall data will be used to provide long datasets (see Chapter 9) to exceed the required target return period, for use with continuous simulation models for flood estimation at high return periods. Generation of such rainfall data series introduces a further element of uncertainty which is propagated through the runoff model.

7.3.7 Observed data

FEH. The incorporation of available observed flood data at the target site, or from donor or analogue catchments, at all stages of either method is emphasised to reduce uncertainty, particularly when estimating from catchment properties, and ensure relevance of flood frequency estimates to the catchment characteristics of the target site.

CS. The method at present does not directly incorporate use of observed data. However, it is possible that observed data from, for example a short record, could be used as a comparison with simulated flows obtained by generalisation, with adjustment of model parameters as necessary.

7.3.8 Climatic variability

Natural climatic variability over the UK tends to produce sequences of flood-rich and flood-poor years, and thus the period of observed record used within the development of flood frequency methods contributes to the overall uncertainty of the derived estimates. Seasonal variation is also a factor in flood frequency; for example, floods in large catchments occur almost exclusively in the winter from spatially extensive, long duration rainfall events, but in small urban catchments floods may be a particular problem in the summer from intense convective storms.

FEH. The FEH recommends a variety of ways to allow for the impact of climatic variability within observed records; for example all records used in pooling-groups must be of at least eight years and observed QMED values for records shorter than 14 years should be adjusted by reference to longer local records. Within the runoff model adjustments for seasonality have been introduced to allow for differences between winter (October-March) and summer (April-September) design rainfall and initial soil moisture.

CS. The main impact of climatic variability within the CS method is the flood characteristics of the observed data used in the model calibration. The (necessarily) short data record length (approximately eight years) and comparative lack of floods within the earlier pilot project (FD0404) were two of the reasons for updating the hourly database and including daily catchments with generally longer records (see Table 2.2) in the current project. The impact of data period on calibration and modelling of major flood events was investigated as part of the model testing (see Chapter 4.6). Seasonality of flood events is automatically incorporated in the modelling method.

7.3.9 Non-stationarity

The concept of relating a return period to a flood event assumes that the conditions pertaining throughout the period of record have remained the same, both in terms of climate and catchment characteristics. Care is needed when using an observed record to ensure it is stationary. Generalisation procedures based on catchment properties and model calibration may limit their application to a particular timeframe.

FEH. Only stationary observed records should be used within either the runoff or statistical method. Statistical analyses are included in both the runoff (design storm rainfall) and statistical methods so neither method is suitable for including impacts of climate change on flood frequency.

CS. It has been assumed that all data records used in development of the method are essentially stationary. Potentially, the method can be readily used to estimate flood frequency with different climate scenarios, by generating alternative rainfall and PE data series. This assumes that model parameters, determined through calibration and generalisation methods relating calibrated parameters to catchment properties for a particular timeframe, are applicable under a changed climate.

7.3.10 Spatial consistency

Flood frequency estimates for locations within a river network (that is, drainage to a single point) should be consistent with each other. This requires that sudden increases in flow occur only at confluences, that flood estimates increase in a downstream direction and that estimates downstream of a confluence are consistent with those of the contributing upstream areas.

FEH. One of the problems with the original FSR methods was the use of regional areas for determining parameter values and statistical analysis at a limited number of point sites which led to discontinuity across area boundaries and spatial inconsistencies of flood estimates. Automation of the statistical method (Morris, 2003) has sought to eliminate, as far as possible, these problems.

CS. Spatial consistency depends on lack of discontinuities when generating parameters through generalisation methods for sub-catchments compared with the catchment as a whole. Chapter 8 indicates that this should be the subject of comprehensive testing.

7.4 Concluding remarks on FEH and continuous simulation

The ten points described above have discussed the main issues which may impact on choice of FEH methods or continuous simulation for determination of flood frequency estimates for a catchment and uncertainties incurred with use of that method. Detailed quantitative comparisons between the FEH and CS approaches are required to explore these issues further and to enable the development of formal guidelines for recommended use with particular catchment types and purpose of the flood frequency estimates.

8 FUTURE DIRECTIONS

Ann Calver

PREVIOUS FAST TRACK BOX ON PAGE 109

CHAPTER 8 FUTURE DIRECTIONS

In this fast track narrative, and to a greater extent in the text of this report, note is made of research issues which could potentially be explored in the context of flood frequency quantification through the modelling of continuous river discharge time series. The key such issues are stated in the text of this chapter. They relate to variations on the approaches developed here, other possible strands of the modelling approach to ungauged site flood frequency estimation, and the 'next steps' which need to be explored between the methods presented here and their practical use. These cover not only the conversion of research code to software and user guidance but also important issues of performance testing of a point estimation technique in a continuous spatial context. The project team has formulated proposals to address the most pressing of these issues: wider hydrological research advances of the coming years may be expected to offer increasing evidence in the field of transfer of information for prediction at data-sparse sites.

NEXT FAST TRACK BOX ON PAGE 121

This short chapter deals not with details of the research developed in the project but with the *context* in which the concluding chapter (Chapter 9 below) is set and should be read. It has been mentioned at various junctures in this report that pragmatic decisions, based on hydrological judgement, have had to be taken in the course of the project (for time and budget reasons), such that it is recognised that by no means all possible alternatives have been explored. One aspect of desirable future activity may therefore be to explore the implications of these choices, unless the practical decisions made attract a wide degree of acceptance and agreement, which may be the case. The main, but not only, such issues relate to alternative choices within both the regression-type and similarity-type spatial generalisation and the possibility of establishing hybrid and split-region methods. Whilst types of errors in spatial generalisation have been quantified (Chapters 5 and 6), attribution to the multiplicity of their sources is not readily achieved but may be expected to assist tailored generalisation methods to particular types of situation. Further techniques such as self-organising mapping may offer potential. The explicit, rather than implicit, treatment of snowmelt, whilst maintaining a parameter-sparse formulation, may also be considered desirable.

There are, too, extra issues which, whilst beyond the current project remit, it would be additionally useful to explore. Amongst these is the testing of what is essentially a point frequency quantification method in the full spatial context for consistency of estimates (in the same way as the Flood Estimation Handbook was, after publication, tested in project FD1603). Also important is the consideration of how specific local knowledge can best be incorporated into the procedures for ungauged sites. Data and model structure uncertainties (in addition to the model parameter uncertainties evaluated in Chapter 6 above) may be considered worthy of exploration.

A third main category of consideration for the future is the follow-on activities for engineering use of the conclusions of this work, modified if appropriate by the above activities. These primarily cover detailed numerical comparison with the Flood Estimation Handbook methods and advice for preferred usage, conversion of research-level code to software, the production of guidance documentation and incorporation of user feedback in testing programmes.

In the time these developments will take, a watching brief should be kept on hydrological research advances to determine whether other approaches are able to offer enhancements – whether, for example, subsets of large detailed areal catchment models can produce locally-ungauged flood frequency estimates, and whether channel (as opposed to catchment) properties can perform in the spatial generalisation context to higher than current levels. The runoff modelling research of this project needs to be accompanied by a sufficiently high level of performance of the long time series rainfall modelling which drives the temporally-extended use of the method: a pertinent issue here may be the merits of spatio-temporal statistical methods in relation to downscaled physical methods.

The following and final chapter of this report presents the best recommendations for the continuous simulation method for flood frequency quantification at the current stage of research.

9 USE OF METHODS AND IMPLEMENTATION

Ann Calver, Nick Reynard, Sue Crooks, Alison Kay, David Jones, Simon Dadson

PREVIOUS FAST TRACK BOX ON PAGE 119

CHAPTER 9 USE OF METHODS AND IMPLEMENTATION

The final chapter of the report recommends how the research findings are currently best used for the quantification of river flood frequencies.

For a gauged site the calibration of the chosen runoff model, or transposition from a good 'matching' catchment, provides the parameters to generate a long series of river flows from which flood statistics and hydrograph characteristics are derived. For an ungauged site, catchment characteristics allow establishment of runoff model parameters to similarly generate long time series of river flows. For users preferring a site-similarity approach (here, with the Probability Distributed Model, showing *overall* best performance to levels of recurrence interval suited to checks by observations), pooling groups are established and weightings applied to their calibrated runoff model parameter values. For those preferring a possibly simpler regression approach (the Time-Area Topographic Extension model in this context, with a univariate approach, gave *overall* best performance), the established predictive equations are applied to derive model parameters from the catchment properties. It is to be noted that 'overall' performance across the 119 sites is not the same as best performance at a specific site. Local knowledge, in terms of prior estimates or observations of large floods, is always a useful guideline. The measures of uncertainty developed can be applied to final estimates in detailed or summary form.

Whilst the provision of long rainfall series for driving the generation of long runoff series is not the remit of this project, a demonstration is provided, using a stochastic rainfall generator, to show how river flood frequencies are established to recurrence intervals greater than those appropriate to direct observation. The way in which continuous simulation methodology readily allows variability in climate conditions is also indicated.

The report ends with recommendations for dissemination of the continuous simulation methodology which include scientific and practitioner meetings and liaison, consideration of any additional research necessary, the trialling of methods, and the feedback of practical experience of use, prior to later consideration of development of software tools.

END OF FAST TRACK BOXES

This final chapter brings together the research developments elaborated in earlier chapters to describe the way in which it is recommended that the continuous simulation method is currently best applied for the estimation of river flood frequencies. Plainly, experience of practical use will, over time, feed into the refining of the method (including comparison of performance with the Flood Estimation Handbook, cf. Chapter 7), in addition to incorporation of ongoing research advances (as for example those indicated in Chapter 8).

This chapter considers first the gauged site case, then the ungauged case. As well as spatial extension of the method, a demonstration of temporal extension of the method to high recurrence intervals is included (recognising, however, that generation of rainfall inputs is not the remit of this project). Comment is also made on the handling of climatic variability. The chapter concludes with a proposed outline structure of dissemination following the release of this report. Overall it is considered that it has been timely and productive to have investigated the potential of the runoff modelling approach to river flood frequency, that the approach has much to offer and can best progress further with informed liaison between practitioners and research developers.

9.1 Recommendations for gauged sites

To consider a site as 'gauged' in this context in practice means that there are sufficient rainfall and river flow data of sufficiently good quality at daily, and preferably sub-daily, discretisation that a catchment conceptual runoff model can be calibrated with a degree of confidence. It is helpful if floods have occurred in the period and the catchment has not experienced great change with respect to dominant runoff generation mechanisms. Whilst this decision is a matter of hydrological judgement in the light of other factors such as purpose of frequency estimate and associated engineering risk, a guideline length of good data record would be a minimum of two years: whilst the continuity of discharge data is important, that of the rainfall record is even more so. For shorter records and those of poor quality, the site should primarily be treated as ungauged.

The next step for a site considered as gauged is to establish the parameter values of the runoff model. This project provides over 100 such sets with high quality calibrations for the PDM and TATE models. There are, in addition, other existing calibrated parameter sets from modelling undertaken in a number of other contexts: these should, however, be checked with respect to purpose of calibration in that it is plainly important that flood peaks have been accorded due weight. The question of which model to use depends, amongst other considerations, on personal preference and experience of use.

There will be gauged sites for which calibrated parameter sets are not available or are of poor or doubtful quality. One route for these catchments is to transpose runoff model parameters from a calibrated catchment deemed, from hydrological experience, to be sufficiently similar in response to the site in question. In this case, of transposing model parameters from a physiographically similar catchment, river flow and rainfall data are not strictly necessary (see also Section 9.2 below on ungauged sites): it is, however, an approach included here since the availability of data substantially assists in establishing transposition possibilities. Note that catchments do not have to be similar in terms of rainfall characteristics in this method: it is the response pattern to rainfall which is the key factor. The potential for transposition is therefore based on catchment configuration and material properties, which need not be confined to geographical proximity. This is

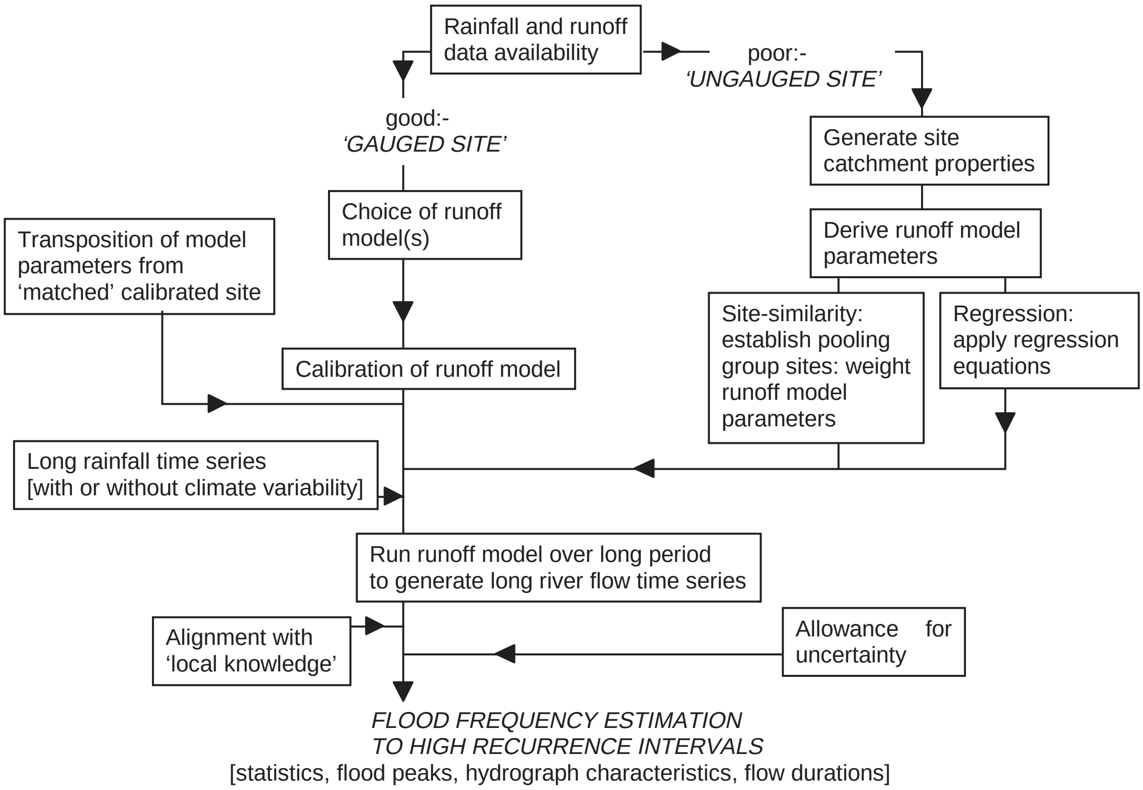


Figure 9.1 Use of the continuous simulation method for gauged and ungauged sites.

an approach requiring hydrological skill and one for which guidelines can be built up as experience of use develops.

For a gauged site for which calibrated parameters are not yet established, and which are not suitable for or thought adequately covered by the transposition route of the above paragraph, the chosen runoff model can be calibrated with the rainfall data as the driving time series and the flow data as the target output. Experienced users may wish to carry out this step themselves: additionally or alternatively the successful automated calibration methods of Chapter 4 of this report for the PDM and TATE models offer the potential to be encapsulated in a user-driven system.

These methods allow the derivation of numerical values of runoff model parameters at the site of interest. With the compiled / packaged code, and the long rainfall time series (sourced elsewhere beyond this project), they can be used to run the model and produce continuous flow time series, or indeed suites of these, which yield the aspects of flood flows of interest, including characteristics of the larger, rarer floods. To demonstrate the output of flood frequency curves including the more extreme floods associated with longer-than-commonly-observed rainfall series, a brief example is included in Section 9.3 below, after the consideration of ungauged sites to which the temporal extension issue also applies. Figure 9.1 summarises these gauged site procedures: it also shows the ungauged procedures which are the subject of the next section.

9.2 Recommendations for the ungauged site

The 'ungauged' site is that with little or no rainfall and river discharge data, or with data which are discontinuous and/or of poor quality. In practice, the majority of sites for which flood frequency estimates are required are likely to fall into this category. It is a recognition, indeed, that data availability is in practice unlikely to be adequate (if only because of length of record with respect to inclusion of rare events) that has made it necessary to explore countrywide flood frequency methods offering estimation for the data-sparse areas.

Section 9.1 above (for the gauged site) indicates that the direct transposition of runoff model parameter values from a sufficiently similar calibrated site is also a possibility for the ungauged site. In this latter case more reliance is necessarily placed on judgement of similarity because of the scarcity or absence of measured hydrological response. It is therefore to be used with extreme caution.

The more general route for ungauged site estimation is the use of catchment properties to spatially-generalise the method and allow derivation of runoff model parameter values. The results reported from generalisation methods developed in Chapter 5, seen also in the light of uncertainty estimates of Chapter 6, offer a choice of ways forward for the ungauged site. Note, especially, that testing of alternative methods can only be done up to the recurrence interval lengths that observations allow: extension to rarer events cannot be conclusively proven to be of a particular quality.

Overall across Britain, for the sample of catchments tested and to the recurrence intervals for which observations afford testing potential, the generalisation method which gave the best results was a site-similarity variant used with the PDM runoff method. This procedure could therefore be encapsulated in a pooling group search tool, with a variety of levels of user-interaction, to provide the runoff model parameters values, suitably weighted, for the target ungauged site.

It is, however, extremely important to note that for many of the catchments used to test the procedures there are methods other than the one with the best overall performance across the country which perform better in establishing runoff model parameter values for the target site (still with the proviso of recurrence interval range). These involve both the TATE and PDM models: as well as a site-similarity route with the TATE, they include both univariate and sequential regression approaches with both runoff models. A valuable procedure to try to establish in the near future is guidance as to when to override the overall-best-performing method by one with evidence of a higher standard of performance in a particular type of ungauged site. It is perhaps unsurprising, given the multivariate nature of river flood generation, that it is not an immediately straightforward matter to characterise the pattern of contributions to generalisation errors, nor at this stage to be able to offer rigorous guidelines. The best current recommendation for departure from the averagely-best route is similarity of target ungauged site to a test catchment in this project which showed enhanced predictive performance with the alternative methods listed at the beginning of this paragraph.

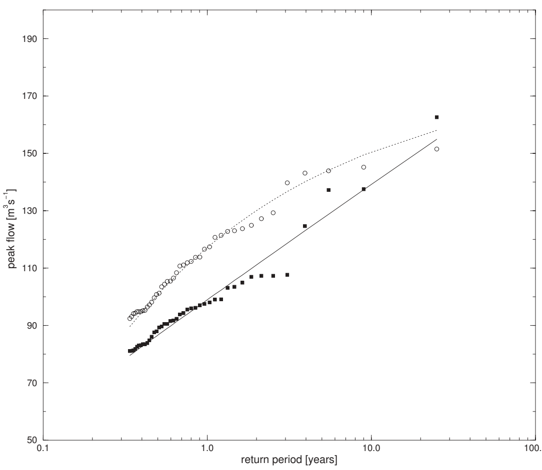
Site-similarity
PDM

Site: 78005, Kinnel Water at Bridgemuir
(229 km^2)

Catchment properties:
AREA=229.0, BFIHOST=0.434, DPLBAR=21.4, DPSBAR=110.7, FARL=0.997, PROPWET=0.62, SAAR=1397, URBEXT=0.000, HOSTNG=44.74, HOSTP=0.138, HYDC=159.5, LANDA=38.6, LANDB=29.4, LANDC=21.2, DRAIN2=1.9835

Pooling group catchments (and their weights) for each parameter:
f_c:
06008 (0.067), 07004 (0.124), 13001 (0.045), 47008 (0.075), 50006 (0.075), 54025 (0.075), 57005 (0.089), 60002 (0.151), 60003 (0.049), 79005 (0.251).
c_{max}:
27043 (0.103), 55008 (0.102), 60002 (0.065), 64001 (0.095), 65006 (0.025), 78003 (0.147), 79002 (0.165), 79003 (0.068), 79005 (0.148), 81002 (0.082).
k_1:
03003 (0.072), 50006 (0.082), 60002 (0.073), 64001 (0.193), 79005 (0.167), 81002 (0.129), 81006 (0.038), 93001 (0.041), 94001(0.115), 95001(0.089).
k_b:
03003 (0.094), 06008 (0.094), 07001 (0.126), 27043 (0.027), 47008 (0.048), 66011 (0.084), 79003 (0.119), 79005 (0.194), 81002 (0.139), 81006 (0.076).

Weighted PDM parameters for target site:
f_c=1.2559
c_{max}=95.8854
k_1=17.2766
k_b=63.9202



Multiple univariate regression
TATE

Site: 53009, Wellow Brook at Wellow
(72.6 km^2)

Catchment properties:
AREA=73.53, BFIHOST=0.643, FARL=0.987, PROPWET=0.37, SAAR=999, SPRHOST=27.3, URBEXT=0.038, PORO=50.1, LANDA=48.8, LANDB=2.0, DRAIN2=0.6805

TATE parameters for target site, from predictive equations:
crm=0.3171
csm=0.2274
cfr=0.2433

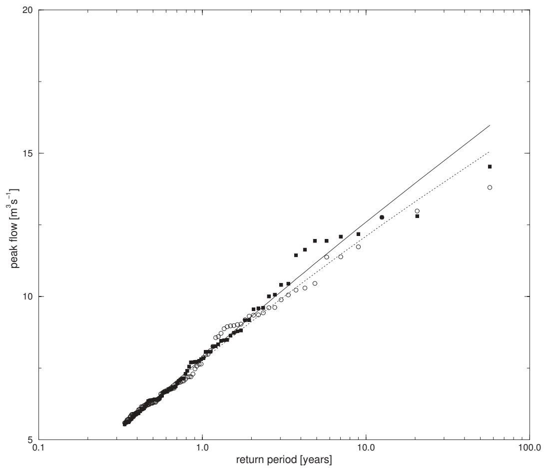


Figure 9.2 Worked examples of ungauged site methodology; solid line – generalised, dotted line – observed (withheld in derivation).

Personal preference of methodology is also to be borne in mind. Users may perceive the application of a regression approach to be a more straightforward procedure than that of a pooling group method and one which may avoid the occasional difficulty of establishing such groups for atypical ungauged sites. Regression approaches generate runoff model parameter values directly from inserting site catchment properties into the predictive equations. If a regression approach is sought, the best overall performance (again with site-specific and recurrence interval provisos) is currently provided by univariate regression with the TATE model.

Once runoff model parameters are established for the ungauged site, the model is run (as in the case of the gauged site above) to derive long continuous river flow time series from which flood statistics and hydrograph characteristics can be readily derived.

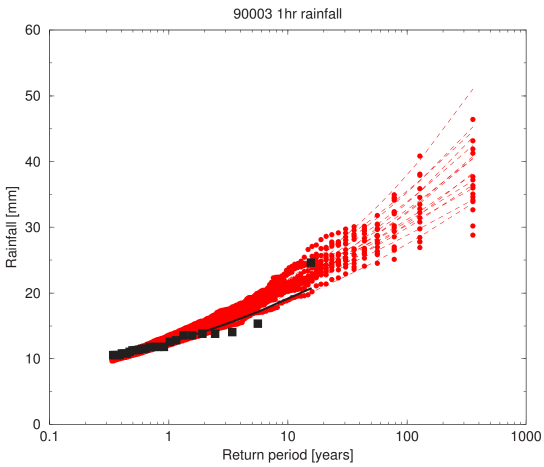
Figure 9.2 gives worked examples of deriving flood frequency curves for ungauged sites for both a site-similarity method using the PDM and a univariate regression approach with the TATE model: details of approaches are as developed in Chapter 5 of this report.

The levels of uncertainty around flood frequency curves for ungauged sites resulting from model parameter uncertainties provide a useful context in which flood frequency results are handled in practice. At this stage of the continuous simulation procedure it is recommended that the overall characteristics of the results (Chapter 6 and Appendix D) serve as quantitative guidelines of a general nature. The procedures can, if appropriate as the uptake of continuous simulation evolves, be encapsulated in software tools for the user or in look-up tables / diagrams for representative types of situations.

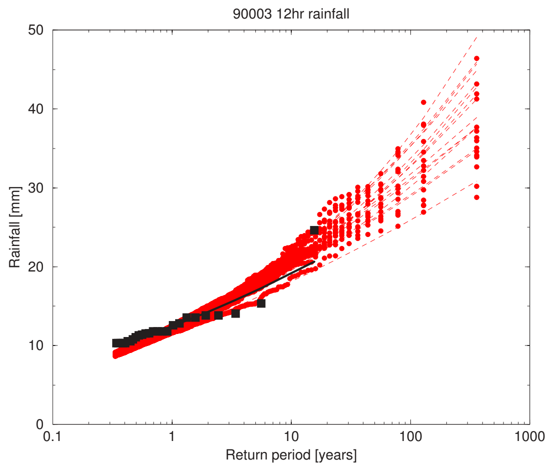
9.3 Temporal extension of the continuous simulation method

Both gauged and ungauged site methods use long rainfall time series to drive the catchment runoff model (see lower part of Figure 9.1 above): the long periods increase the likelihood of inclusion of rarer larger floods. Some sites may have appropriate precipitation observations, but many will not. Defra is elsewhere sponsoring research on temporal-spatial rainfall (and evaporation) modelling with the aim of fulfilling this role. In order to *demonstrate* the runoff modelling procedures of this FD2106 project extended in time to cover higher recurrence interval floods, a simple stochastic rainfall generator has been used.

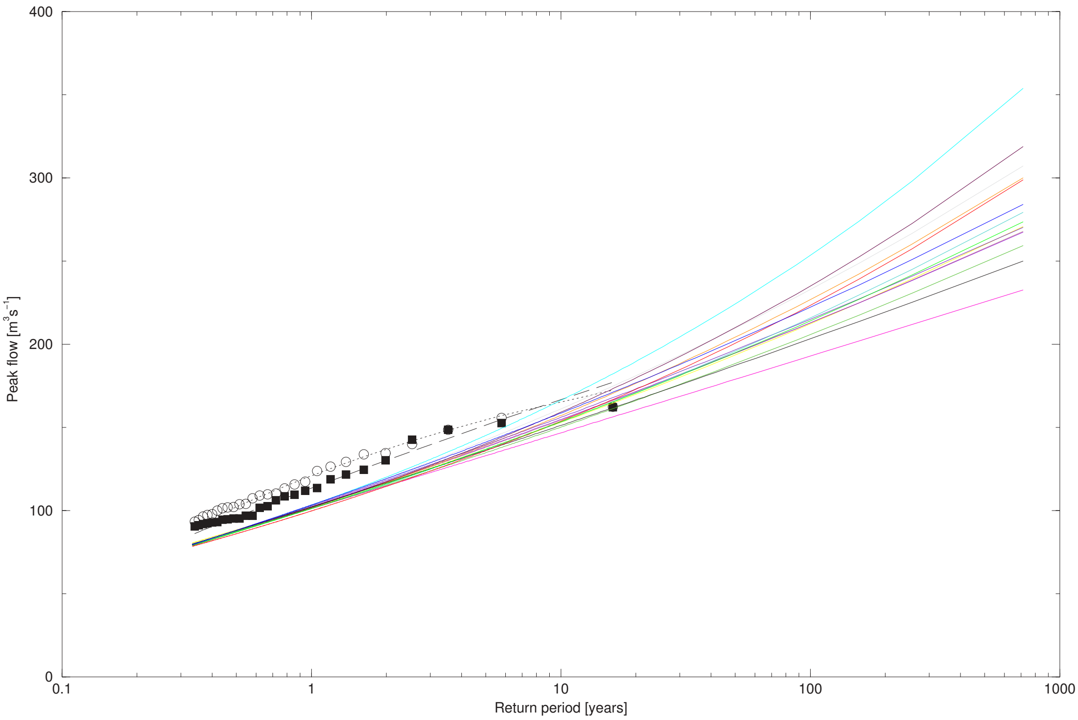
Hourly rainfall data were synthesised using this generator (Goodsell and Lamb, 1999), where parameters describing storm duration, mean intensity and arrival time were derived from analysis of the observed record. The storm profile was based on a dimensionless average profile perturbed by an amount generated from a lognormal distribution and smoothed according to the first-order autocorrelation coefficient. Allowance for seasonality was achieved by simulation of separate series for winter (October to March) and summer (April to September). The model was calibrated by adjustment of the first-order autocorrelation coefficient for the summer to achieve a fit between observed and simulated peak rainfall data taken from an analysis of the whole series. Frequency analyses were compared for hourly data and cumulative 12-hour



a)



b)



c)

Figure 9.3 Temporal extension of the continuous simulation method for catchment 90003: a) simulated (red circles) and observed (black squares) hourly rainfall frequency distribution, with lines showing fitted 3-parameter generalised Pareto distribution (GPD); b) as a) but 12-hourly rainfall frequency distribution; c) flood frequency distribution from observed flows (black, open circles and dotted line), simulated using observed rainfall (black, filled squares and dashed line) and simulated (multi-coloured solid lines) using stochastic rainfall series.

rainfall, the latter ensuring realistic distribution of rainfall within the storm profile. An ensemble of rainfall series was generated using different random seeds to initialise the model. Fifteen rainfall series, each of 400 years, were derived by interleaving six-month sequences of winter and summer rainfall and these series were used to drive a calibrated runoff model. The generated rainfall and the resultant river flood frequency curves to recurrence intervals of 200 years are shown for an example catchment in Figure 9.3.

In using flood frequency estimation for scheme or strategic planning, the question of allowance for climatic variability arises. This is not a main remit of this project: it is, however, useful to note that variability in climate series can very readily be incorporated into the continuous simulation approach to flood frequency quantification if required. In brief, scenarios of change in climate variables (under chosen emission scenarios and derived from available atmospheric circulation models) should be appropriately scaled in space and time before application as driving precipitation and evaporation input series to the catchment runoff models. It is to be borne in mind that uncertainties in data and methods are of a similar order to the magnitude of possible physical changes. Changes can be applied transiently or as established new regimes, with statistical expressions of risk respecting the degree of non-stationarity of the situation.

As with any countrywide generic method, a continuous simulation approach can profitably be enhanced by consideration of local knowledge, whether of historic flood records or of estimates of flood frequency by earlier or alternative methods.

- *Familiarisation with FD2106 Final Report contents*
  - *web postings*
  - *scientific meetings*
  - *journal publications*
  - *practitioner seminars*

- *Consideration of the further research issues (Chapter 8)*
  - *client - researcher liaison*

- *Trialling and feed-back of experience*
  - *development of opportunities for trialling*
  - *researcher / practitioner meetings*
  - *feedback into methodologies*

- *Consideration of specification of tools*
  - *client / researcher / software-house liaison*

Figure 9.4 Chief next-steps in river flood frequency estimation by continuous simulation

9.4 Outline structure of dissemination for continuous simulation

Chapters 1 to 6 of this report have detailed the research of project FD2106 on the continuous simulation method for river flood frequency quantification. Comparison in terms of principles has been made with the Flood Estimation Handbook in Chapter 7. Whilst Chapter 8 gives an overview of what research issues remain of value in taking forward the continuous simulation approach, this final chapter has offered the best recommendations for use at the *current* state of information. An outline plan of the best next-steps recommended for the dissemination and enhancement of the continuous simulation method is proposed in Figure 9.4.

REFERENCES

Abdulla, F.A. and Lettenmaier, D.P. 1997 *Development of regional parameter estimation equations for a macroscale hydrologic model*. Journal of Hydrology, 197, 230-257.

Acreman, M. and Wiltshire, S.E. 1987 *Identification of regions for regional flood frequency analysis*, (abstract), EOS Trans. AGU, 68(44), 1262.

Acreman, M. and Wiltshire, S.E. 1989 *The regions are dead; long live the regions. Method of identifying and dispensing with regions for flood frequency analysis. In: FRIENDs in Hydrology, (eds. L. Roalds et al.),* IAHS Publ. No. 187, 175-188.

Andersson, L. and Harding R.J. 1991 *Soil-moisture deficit simulations with models of varying complexity for forest and grassland sites in Sweden and the U.K.* Water Resources Management, 5, 25-46.

Bissell, A.F. and Ferguson, R.A. 1975 *The jackknife – toy, tool or two-edged weapon?* The Statistician, 24, 79-100.

Boorman, D.B., Hollis, J.M. and Lilly, A. 1995 *Hydrology of soil types: a hydrologically based classification of soils in the United Kingdom.* IH Report No. 126, Institute of Hydrology, Wallingford.

Burn, D.H. 1990 *Evaluation of regional flood frequency analysis with a region of influence approach*. Water Resources Research, 26(10), 2257-2265.

Burn, D.H. and Boorman, D.B. 1992 *Catchment classification applied to the estimation of hydrological parameters at ungauged sites*. IH Report 118, Institute of Hydrology, Wallingford, UK.

Calver, A. 1993 *The time-area formulation revisited.* Proceedings of the Institution of Civil Engineers: Water, Maritime and Energy, 101, 31-36.

Calver, A. 1996 *Development and experience of the 'TATE' rainfall-runoff model.* Proceedings of the Institution of Civil Engineers: Water, Maritime and Energy, 118, 168-176.

Calver, A., Kay, A.L., Jones, D.A., Kjeldsen, T., Reynard, N.S. and Crooks, S. 2004 *Flood frequency quantification for ungauged sites using continuous simulation: a UK approach*. Proceedings of the International Environmental Modelling and Software Society Conference, Osnabruck, June 2004.

Calver A., Lamb, R., Kay, A. L. and Crewett, J. 2001 *The continuous simulation method for river flood frequency estimation.* Defra Project FD0404 Final Report, CEH Wallingford.

Calver, A., Lamb, R., Kay, A.L., Crooks, S., Jones, D.A. and Stewart, E.J. 2002 National river catchment flood frequency method using continuous simulation. Defra – Environment Agency Flood management R&D Project FD2106 Inception Report, 19pp.

Crewett, J., Lamb, R. and Calver, A. 1999 *Spatial generalisation of model parameters for flood frequency estimation using continuous simulation: recent research advances*. Report to Ministry of Agriculture, Fisheries and Food, Institute of Hydrology, Wallingford, UK.

Crewett, J., Lamb, R. and Calver, A. 2000 *Spatial generalisation for flood frequency estimation by continuous simulation.* Geophysical Research Abstracts, 2, EGU, April 2000, Nice, France.

Crooks, S.M., Kay, A.L. and Calver A. 2002 *Model testing in the light of extended data series.* Defra Project FD2106 Report, CEH Wallingford.

Crooks, S., Kay, A.L. and Calver, A. 2004 *Scottish catchments: data assembly, processing and runoff model calibrations*. Defra Project FD2106 Report, CEH Wallingford, August 2004, 6pp.

Davison, A.C. 1984 *Modelling excess over high thresholds, with an application.* In: Statistical Extremes and Applications (ed. J. Tiago de Oliveira), Dordrecht, Reidel, 461-482.

Davison, A.C. and Hinkley, D.V. 1997 *Bootstrap Methods and their Application.* Cambridge University Press, Cambridge.

Davison, A.C. and Smith, R.L. 1990 *Models for exceedances over high thresholds*. Journal of the Royal Statistical Society Ser. B, 52, 393-442.

Efron, B. and Tibshirani, R.J. 1993 *An Introduction to the Bootstrap.* Chapman and Hall, New York.

Fowler A, 2002 *Assessment of the validity of using mean potential evaporation in computations of the long-term soil water balance.* Journal of Hydrology, 256, 248-263.

Fuller, R.M. 1993 *The land cover map of Great Britain*. Earth Space Review, 2**,** 13-18.

Goodsell, G. and Lamb, R. 1999 *Estimating long return period floods by continuous simulation using a stochastic rainfall generator*. Institute of Hydrology Report to Ministry of Agriculture, Fisheries and Food.

Holmes, M.G.R., Young, A.R., Gustard, A.G. and Grew, R. 2002 *region of influence approach to predicting flow duration curves in ungauged catchments.* Hydrology and Earth System Sciences. 6(4), 721-731.

Hosking, J.R.M. and Wallis, J.R. 1987 *Parameter and quantile estimation for the Generalised Pareto distribution.* Technometrics, 29, 339-349.

Houghton-Carr, H. 1999 *Restatement and application of the Flood Studies Report rainfall-runoff method.* Flood Estimation Handbook Vol. 4, Institute of Hydrology, Wallingford, UK.

Institute of Hydrology, 1999 *Flood Estimation Handbook* (5 volumes). Institute of Hydrology, Wallingford.

Jones, S.B. 1983 *The estimation of catchment average point rainfall profiles.* IH Report No. 87, Institute of Hydrology, Wallingford.

Jones, D.A., Kjeldsen, T.R., Kay, A.L. and Calver, A. 2004 *Inclusion of uncertainty estimation methods.* Defra Project FD2106 Report, CEH Wallingford, June 2004, 31pp.

Kay, A.L., Calver, A., Jones, D., Lamb, R., Scarrott, C.J. and Stewart, E.J. 2002 *National river catchment flood frequency method using continuous simulation: Assembly of daily data.* Defra Project FD2106 Report, CEH Wallingford.

Kay, A.L., Jones, D.A., and Calver, A. 2003 *National river catchment flood frequency method using continuous simulation: Including the benefits of daily data*. Defra Project FD2106 Report, CEH Wallingford.

Kay, A.L., Jones, D.A., Kjeldsen, T.R., Fung F., Folwell, S., Calver, A., Reynard, N.S. and Crooks, S. 2004 *National river catchment flood frequency method using continuous simulation: Preferred spatial generalisation procedures.* Defra Project FD2106 Report, CEH Wallingford, July 2004, 22pp.

Kjeldsen, T.R., Kay, A.L. and Jones, D.A. 2003 *National river catchment flood frequency method using continuous simulation: Performance of site-similarity approaches.* Defra Project FD2106 Report, CEH Wallingford, August 2003, 32pp.

Kjeldsen, T.R., Stewart, E.J., Packman, J.C., Folwell, S. and Bayliss, A.C. 2005 *Revitalisation of the FSR/FEH Rainfall Runoff Method.* FD1913 Final Research Report to Defra / Environment Agency, CEH Wallingford, UK.

Lamb, R. Crewett, J. and Calver, A. 2000a *Progress in spatial generalisation of 'continuous simulation' flood frequency modelling.* In: River Flood Defence (eds. F. Toensmann and M. Koch), Kassel Reports on Hydraulic Engineering No. 9/2000, D-117-125, Herkules Verlag. Kassel, Germany.

Lamb, R., Crewett, J. and Calver A. 2000b *Relating hydrological model parameters and catchment properties to estimate flood frequencies from simulated river flows*. In: Proc. BHS 7th National Hydrology Symposium, September 2000, Newcastle, UK, 3.57-3.64.

Lamb, R. and Gannon, B. 1996 *The establishment of a database of hourly catchment average rainfall and flow for flood frequency estimation by continuous simulation.* MAFF Project FD0404 Report, Institute of Hydrology, Wallingford.

Lamb, R. and Kay, A.L. 2004 *Confidence intervals for a spatially generalised, continuous simulation flood frequency model for Great Britain*. Water Resources Research, 40(7), W07501.

Lamb, R., Kay, A. and Crewett, J. 2000 *Accommodating uncertainty in applying broad-scale models for flood estimation.* MAFF Project FD1604 Final Report, CEH Wallingford.

Miller, R.G. 1964 *A trustworthy jackknife*. Annals of Mathematical Statistics, 35, 1584-1605.

Miller, R.G. 1974 *The jackknife – a review*. Biometrika, 61, 1-15.

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

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

Moran, P.A.P. 1975 *The estimation of standard errors on Monte Carlo simulation experiments*. Biometrika, 62, 1-4.

Morris, D.G. and Flavin, R.W. 1990 *A digital terrain model for hydrology*. Proc. 4th International Symposium on Spatial Data Handling, Zurich, Vol 1, 250-262.

Morris, D.G. 2003 *Automation and appraisal of the FEH statistical procedures for flood frequency estimation.* Defra Project FD1603 Final Report, CEH Wallingford.

Naden, P.S. 1992 *Analysis and use of peaks-over-threshold data in flood estimation*. In: A.J. Saul (Ed.) Floods and Flood Management, Kluwer Academic, Dordrecht, p. 131-143.

Naden, P.S. 1993 *Methods and Techniques for Peaks-over-threshold Flood Analysis*. MAFF Report. Institute of Hydrology.

Nash J.E. and Sutcliffe, J.V., 1970 *River flow forecasting through conceptual model, 1. A discussion of principles.* Journal of Hydrology, 10, 282-290.

NERC 1975 *Flood Studies Report*, Natural Environment Research Council, London, UK.

Niggli, M., Hingray, B. and Musy, A. 2002 *Regional estimation of model parameters – a case study in western Europe.* In: Proc. International Conference on Flood Estimation (eds. M Spreafico and R. Weingartner), March 6-8, 2002, Bern, Switzerland, 611-620.

Peddada, S.D. 1993 *Jackknife variance estimation and bias reduction.* In: “Computational Statistics: Handbook of Statistics, Vol. 9 (Ed. Rao, C.R.) North-Holland, London

Pilgrim, D.H., Cordery, I. and French, R. 1969 *Temporal patterns of design rainfall for Sydney*. Civil Engineering Transactions, Institution of Engineers, Australia, CE11, 9-14.

Pollock , D.S.G. 1979 *The Algebra of Econometrics*. Wiley, Chichester.

Post, D.A. and Jakeman, A.J. 1999 *Predicting the daily streamflow of ungauged catchments in S.E. Australia by regionalising the parameters of a lumped conceptual rainfall-runoff model.* Ecological Modelling, 123, 91-104.

Quenouille, M.H. 1949 *Problems in plane sampling*. Annals of Mathematical Statistics, 20, 355-375.

Quenouille, M.H. 1956 *Notes on bias estimation*. Biometrika, 43, 353-360.

Robson, A. and Reed, D.W. 1999 *Statistical procedures for flood frequency estimation*. Flood Estimation Handbook Vol. 3, Institute of Hydrology, Wallingford, UK.

Sefton, C.M. and Howarth, S.M. 1998 *Relationships between dynamic response characteristics and physical catchment descriptors of catchments in England and Wales*. Journal of Hydrology, 211, 1-16.

Shao, J. and Tu, D. 1995 *The Jackknife and Bootstrap.* Springer, New York.

Smith, R.L. 1984 *Threshold methods for sample extremes*. In: Statistical Extremes and Applications (ed. J. Tiago de Oliveira), Dordrecht, Reidel, 621-638.

Thompson, N., Barrie, I. A. and Ayles, M. 1982 *The Meteorological Office Rainfall and Evaporation Calculation System: MORECS* (July 1981). Hydrological Memorandum No. 45, Met Office, Bracknell.

Tukey, J.W. 1958 *Bias and confidence in not-quite large samples*. Annals of Mathematical Statistics, 29, 614 (abstract).

Wagener, T., Wheater, H.S. and Gupta, H.V. 2004 *Rainfall-Runoff Modelling in Gauged and Ungauged Catchments*. Imperial College Press, London.

Wang, Q.J. 1991 *The POT model described by the Generalised Pareto distribution with Poisson arrival rate*. Journal of Hydrology, 129, 263-280.

Zellner, A. 1971 *An Introduction to Bayesian Inference in Ecomometrics*. Wiley, New York.

PB11208

Nobel House
17 Smith Square
London SW1P 3JR

www.defra.gov.uk

ISBN 0-85521-164-4

9 780855 211646

Department for Environment
Food and Rural Affairs