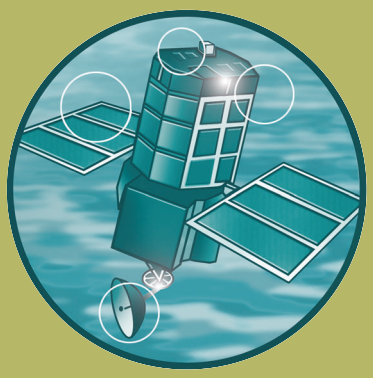
Joint Defra/EA Flood and Coastal Erosion Risk Management R&D Programme

Improved methods for national spatial-temporal rainfall and evaporation modelling for BSM

R&D Technical Report FD2105/TR



Environment Agency

defra
Department for Environment Food and Rural Affairs

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

Improved methods for national spatial-temporal rainfall and evaporation modelling for BSM

R&D Technical Report F2105/TR

Produced: March 2006.

Authors:

Principal Investigators:
HS Wheater[1], VS Isham[2], RE Chandler[2], CJ Onof[1] and EJ Stewart[3]

Research Staff:
E Bellone[1,2], C Yang[1,2], D Lekkas[1], G Lourmas[2], M-L Segond[1], AJ Frost[1,2], C Prudhomme[3] and S Crooks[3]

1. Department of Civil and Environmental Engineering, Imperial College London
2. Department of Statistical Science, University College London
3. Centre for Ecology and Hydrology, Wallingford

Statement of use
This research provides guidance and modelling tools for the simulation of single site and spatial rainfall and evaporation for use in continuous simulation rainfall-runoff modelling for flood design.

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

Keywords: Rainfall simulation, Poisson cluster processes, Generalized linear models, Spatial temporal disaggregation, Rainfall-runoff simulation.

Research contractors:
Prof Howard Wheater, Department of Civil and Environmental Engineering, Skempton Building, Imperial College London, London SW7 2AZ (h.wheater@imperial.ac.uk).

Prof Valerie Isham, Department of Statistical Science, University College London, 1-19 Torrington Place, London WC1E 6BT (valerie@stats.ucl.ac.uk).

Dr Ann Calver, Centre for Ecology and Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford, Oxon OX10 8BB, (anc@ceh.ac.uk).

Defra Project Officer: Prof Edward Evans, Bevis, Great Somerford, Nr Chippenham, Wiltshire SN15 5JA (evansep@onetel.com).

Publishing organisation
Department for Environment, Food and Rural Affairs,
Flood management Division
Ergon House, Horseferry Road,
London SW1P 2AL

Tel: 0207 238 3000 Fax: 0207 238 6178

www.defra.gov.uk/environ/fcd

Published by the Department for Environment, Food and Rural Affairs. Printed in the UK, March 2006, on material containing 80% post-consumer waste and 20% totally chlorine free virgin pulp.

PB 12527/9TS

Executive Summary

Background

The use of continuous simulation rainfall-runoff modelling has several important advantages over conventional event-based methods for flood design. These include the ability to represent explicitly the joint probability of occurrence of rainfall and antecedent conditions, and hence to represent impacts of climate change. However, to generate long series of flows, appropriate input time series are required. This report addresses the need to provide continuous inputs of rainfall and potential evaporation to rainfall-runoff models to support continuous simulation for flood risk assessment.

Objectives

The project aims are:

- to develop a procedure for single-site (or catchment average) continuous rainfall and evaporation modelling;
- to develop a procedure to model spatial variability of daily rainfall and to obtain sub-daily spatial rainfall by disaggregation;
- to develop a methodology to simulate detailed radar rainfall fields, accommodating spatial-temporal heterogeneity;
- to provide guidance to users concerning the relative importance of spatial rainfall properties and appropriateness of methods.

An important element is model validation of the integrated modelling procedures, i.e. testing the combined performance of the rainfall and rainfall-runoff models. This project has therefore been closely linked to a parallel continuous simulation rainfall-runoff modelling project (FD2106) with respect to joint model testing.

Results

Part I Single site modelling of rainfall and potential evaporation.
The most basic requirement for continuous simulation modelling is a long time-series of single-site rainfall, for example representing a single raingauge or an areal average. In Part I of the report, alternative stochastic rainfall models were tested using long raingauge data records representative of the UK. The models aim to provide a simplified conceptual representation of the rainfall process using parameters that have physical significance. A new two-phase fitting method was developed that results in significant improvements to the reproduction of extreme-values from raingauge data. Joint testing of the preferred rainfall model was carried out with the rainfall-runoff model from FD2106, using shorter data sets representing catchment average rainfall.

A problem arose with long simulations in that a small number of infeasibly large events could arise. This was largely corrected by a refinement to the simulation procedure. Performance of the combined models shows strengths and some weaknesses. Extreme value performance for some catchments was well reproduced, but sequences of low flows were under represented. More extensive testing would be required on a wider range of catchments before national application to ungauged catchments could be recommended.

Methods were also developed for the modelling of Potential Evaporation (PE). The aim of this work was to provide simulated sequences of potential evaporation at a daily timescale for input to rainfall-runoff models. The project has developed methods:

- to model weekly potential evaporation time series given rainfall;
- to model daily potential evaporation time series given rainfall; and
- to downscale weekly potential evaporation to daily values.

Part II Spatial-temporal modelling based on daily rainfall modelling and spatial-temporal disaggregation

Long radar sequences are required to identify the spatial structure of rainfall for the methods of Part III above, and are not yet routinely available - nor are the continuous space-time models sufficiently developed for immediate application in engineering practice. Hence in Part II we have developed alternative approaches, based on readily available daily rain gauge data. Generalized Linear Models (GLMs) provide a powerful and flexible technique for spatial modelling of daily rainfall. These allow nonstationarities in time and in space to be quantified and incorporated. The basic idea, an extension of linear regression, is to use the values of various predictors to forecast a probability distribution for the amount of rainfall at a site on a given day. This distribution is specified in two parts - we model the probability of rainfall occurrence separately from the amount of rain if non-zero. Previous days' rainfall amounts can be used as predictors, to account for autocorrelation. Other predictors might include quantities representing regional variation (such as site altitude), seasonal variability, long-term trends and 'external' climatological factors such as the North Atlantic Oscillation (NAO). By defining suitable dependence structures between sites, it is possible to build a multivariate GLM that allows simulation of non-stationary daily rainfall sequences over a network of sites. In this project we have refined and extended the GLM software and report case study applications to two UK data sets.

GLMs allow us to simulate daily precipitation, including spatial dependence, location effects, and climate variability. However, for flood applications in the UK, hourly data are needed. Part II therefore also addresses the problem of spatial-temporal disaggregation from daily to hourly time intervals. A simple approach is developed, whereby models of the form adopted in Part I are used to disaggregate daily data to hourly at a representative gauge location (when measured sub-daily data are not available). The resulting sub-daily time series is then applied to all sites over the catchment, linearly scaled to reproduce the correct spatially-varying daily totals. This scheme is tested for the River Lee catchment north of London.

Part III Development and testing of a non-stationary spatial-temporal rainfall model.
In Part III we consider a more general approach to spatial-temporal rainfall modelling, in a spatial extension of the single-site models of Part I. Rainfall is represented as a fully spatial-temporal process where both space and time vary continuously. The ultimate goal is to develop a model that can be simulated to produce high-resolution rainfall sequences. These can then be aggregated as appropriate and used as input to rainfall-runoff models for flood risk assessment and design. Previous models of this type have assumed statistical stationarity (i.e. topography and location effects were not represented, nor climate variability), and their development was based on short radar sequences. Here we have derived a long radar sequence for model testing from the Chenies radar and explored a new approach to represent spatial non-stationarity.

In this report, a model for rain event interiors has been extensively validated and provides a good representation of the basic statistical properties of the fine scale spatial and temporal structure of rainfall. In addition, a methodology to produce a continuous stationary simulation of rainfall across a catchment has been further developed. This has been applied in conjunction with the Chenies radar data with some success; the simulated realisations of the model have been shown to have statistical properties that are generally a good representation of those of the empirical data. Further, for the first time, a way has been demonstrated to extend the stationary spatial-temporal model to allow topographic effects and possible climate change scenarios to be included. An interesting observation was of an apparent increasing frequency of rainfall event occurrence over the period of the available radar record.

Part IV Effect of spatial structure of catchment-scale rainfall on runoff estimation.
In Part IV we address the issue of the significance of spatial rainfall and its representation for flood estimation. A literature review is reported, and a numerical rainfall-runoff modelling study, also based on the Lee catchment, in which alternative rainfall descriptors (i.e. raingauge networks of different density, radar data and disaggregated daily data) are considered and their effects on simulated runoff are evaluated. Finally in this section of the report we report an associated piece of work in which the performance of GLM-simulated daily rainfall as input to a rainfall-runoff model is investigated.

Conclusions / Recommendations

This project has explored a a broad range of options for continuous rainfall simulation. In our view, some of these (particularly those developed in Part II of the report) are now at the stage where they can be considered for routine national use; others should be used with caution until further testing has been carried out.

Single site rainfall and evaporation modelling

The six parameter Random-Parameter Bartlett-Lewis model is recommended as the preferred single site rainfall model for national application, although in practice there is

little difference between the corresponding Bartlett-Lewis and Neyman-Scott formulations. A minimum of 15-20 years of hourly data is recommended for model fitting. Performance for individual raingauges is variable - some results are good, some less so. Extreme value performance is often difficult to assess given the limited lengths of available sub-daily rainfall data for some sites and the associated large confidence limits. However, the model tends to underestimate hourly and overestimate daily extreme values, and performance for individual months was often better than for the annual maxima. A problem arose with the simulation of very long (1000 year) sequences, in that unrealistically-large events would very occasionally be simulated. This problem has largely been rectified by imposing bounds on parameters used in simulation, but users must inspect simulations to check for infeasible extremes. Model parameter uncertainty results using various record lengths also appear to identify genuine opportunities for parameter bounding, which with further work will aid application of these models in a national setting.

Although the performance of the models were improved with the introduction of a new two-phase fitting method, some deficiencies remain in terms of reproduction of extreme statistics. We conclude that the model and optimisation technique are appropriate for use with caution, with the user undertaking careful verification of the simulated outputs. More extensive testing is required on a wider range of sites before the model can be recommended for general application across the UK. It is recommended that alternate means of improving the extreme value performance of the model (such as the use of skewness in fitting) are compared to the newly introduced fitting technique.

The single site rainfall models were applied to represent catchment-average rainfall and then as input to the simulation of river flows, using the model of FD2106. Some theoretical issues were identified, in that catchment-average sequences were often based on individual gauge sub-daily time-series, giving rise to statistical discrepancies between the moments estimation for daily and hourly data. As with the evaluation for individual raingauges, performances were mixed. Some catchments gave very encouraging results, others less so. One generic problem was that the flow duration curves tended to be underestimated.

We conclude that the models are appropriate to use with caution, but careful assessment on a large number of catchments would be needed before the single-site Random-Parameter Bartlett-Lewis rainfall generator could be used with confidence to provide data for flood risk assessment for ungauged catchments. Any user should validate the model using local data to have sufficient confidence for practical application. More work is recommended to pursue the issue of parameter constraints and non-feasible simulations. There are also ideas for model extension that should be explored to seek improvements in performance.

The project has also developed a new set of modelling approaches for potential evaporation. These include:

- modelling weekly potential evaporation time series given rainfall;
- modelling daily potential evaporation time series given rainfall; and,
- downscaling weekly potential evaporation to daily values.

These potential evaporation models have had limited testing, and as yet have not been used with rainfall-runoff models. Nevertheless, the methods appear robust and defensible, and in our opinion can be applied in practice with a reasonable level of confidence.

Spatial-temporal modelling based on daily rainfall modelling and spatial-temporal disaggregation

Realistic multisite daily rainfall sequences can be generated using Generalised Linear Models (GLMs) fitted to daily raingauge data. A simple disaggregation scheme can be applied if subdaily data are required.

Daily rainfall records should be scrutinised carefully before use in any modelling exercise. Common problems include the presence of unflagged monthly totals (which, if undetected, will seriously bias any analysis of extremes) and undocumented changes in recording resolution (which can produce spurious trends in rainfall sequences). We also recommend routine thresholding of daily rainfall observations by setting to zero any values below 0.5mm, to allow for differences in observer practice in the recording of small rainfall amounts.

The GLM framework enables systematic structures to be incorporated into simulated rainfall sequences. In addition corrections can, and should, be made for features such as changes in recording resolution and measurement device.

A number of contrasting case studies have yielded GLMs for daily rainfall with a broadly similar structure. This basic structure is therefore expected to apply quite generally across the UK.

GLMs are designed to be fitted to point raingauge data. If catchment averages or gridded data are required, models should be fitted to observed raingauge data and then simulated on the required grid. For calibration of rainfall-runoff models, historical gridded values can be obtained by conditioning on all available observations. This avoids artefacts associated with the common procedure of gridding the observations prior to analysis.

As developed, the GLMs are designed primarily for use in small to medium sized catchments. At scales above 2000km^2, the methodology may still produce reasonable results; however, users should check carefully to ensure that the results of any simulation exercise appear reasonable. In this and other respects, warning messages from the simulation software should be taken seriously. Further research is required to extend the range of applicability of the GLMs to larger scales.

Some testing has been carried out to assess the suitability of the combined GLM-disaggregation scheme for use in rainfall-runoff modelling applications. Results from a single case study indicate that the scheme has a lot of potential. At present, however, a limitation is that the disaggregation scheme overestimates spatial correlation in rainfall - this, in turn, leads to some overestimation of peak flows. Further research is required to investigate this, to develop the disaggregation methodology further and to test the performance in a wider range of catchments.

The GLM software has been extensively tested by a number of people and found to be

generally easy to use. However, it would be convenient to add a graphical user interface which would speed up the process of model fitting, testing and simulation.

Modelling spatial-temporal rainfall in continuous space and time

Based on our review, we conclude that models for continuous simulation of spatial-temporal rainfall should be based on the Poisson point process approach for medium-term application; alternative models are not yet at the stage where they are readily applicable.The modelling of spatial-temporal rainfall in continuous space and time requires radar data to define the full structure of the spatial rainfall fields. This presents a number of problems; long records of radar data are not yet available, and calibration issues are complex. Radar data need careful quality assessment before use. Calibration procedures have significant effects on the statistical properties of the calibrated fields, and hence on the properties used in the simulation models.

The project has explored the performance of the Poisson process modelling approach for continuous space-time modelling more extensively than has been possible hitherto. Performance is encouraging. The model for event interiors can be used to represent short-term fine scale spatial and temporal rainfall variability of typical rain events. Continuous rainfall space-time fields can be generated by generating storm arrivals and sampling model parameters on a monthly basis from a library of storm parameters that has been derived from extensive analysis of a 13 year radar sequence. A new methodology has been developed to represent spatial non-stationarity and appears promising. However, further work is needed before practical application could be recommended.

The significance of spatial rainfall for runoff estimation

The literature on the significance of spatial rainfall for runoff estimation is complex and sometimes contradictory. Effects can be expected to vary depending on the nature of the rainfall, the nature of the catchment, and the spatial scale of the catchment and rainfall. It is helpful to note that as the scale increases, the importance of spatial rainfall decreases and there is a transfer from spatial variability of rainfall to catchment response time distribution as the dominant factor governing runoff generation. Runoff from urban areas is extremely sensitive to spatial rainfall, and the same is true for arid areas, where rapid flow response is generated from spatially-localised convective rainfall. This is problematic, since the required density of raingauges to capture the spatial variability exceeds that normally available from routine monitoring networks.

Numerical experiments for the Lee catchment (1000 km^2 to Feildes Weir) showed a complex picture. The subcatchments varied greatly in geology and runoff response. Chalk catchments were less sensitive to spatial rainfall description than clay catchments or those with significant urban development. As a result, no clear pattern emerged as a function of catchment scale, or response time, except that the effect of spatial variability was damped at the whole catchment scale of 1000 km^2. The dominant effect was the variability of the rainfall; as this increased, so did the significance of appropriate rainfall characterisation. When the catchments were numerically urbanised, the significance of spatial rainfall was enhanced. It is

concluded that, for largely rural catchments, a network of 16 raingauges seems appropriate at 1000 km^2 scale; between 4 and 7 gauges are required at 80 - 280 km^2, but this recommendation may not be appropriate for summer convective events. As urbanisation increases, the requirements become greater.

A notable result was the relative success of the spatial-temporal disaggregation scheme, when used in conjunction with the rainfall-runoff model. The use of this strategy to disaggregate daily data worked impressively well for the natural catchments, and the procedure can be recommended with some confidence. However, as might be expected, some deterioration in performance occurred when the catchments were artificially urbanised.

The application is reported, at large catchment scale, of the GLM models for daily rainfall to continuous simulation rainfall-runoff modelling using a distributed hydrological model. The GLM results were based on gridded data, which may have affected performance, and showed strengths and weaknesses. They were particularly good for winter rainfall and the areas of the catchment where rainfall was highest. As a result, overall performance in flood frequency estimation at catchment scale was impressive although some over-estimation of lower flows occurred.

Concluding comments

The rainfall models presented here have strengths and weaknesses in aspects of their performance, and the objective of continuous simulation modelling necessitates a trade-off between different aspects of performance. It is a major challenge to seek to represent such a wide range of properties of rainfall and flow, including the extremes of flood frequency and low flow durations, using coupled modelling of the rainfall and runoff processes. Nevertheless the methods presented here have clear potential, and we have attempted to indicate to the user where the models are in our view ready for immediate application. The main requirement at present is for further testing. In particular it is likely that performance could be improved by revisiting aspects of the rainfall modelling with the benefit of the results from the joint model testing.

All of the methods developed here require a reasonable level of user awareness. The methods have the potential to deliver much more than traditional event-based techniques, but with this comes a responsibility on the part of the user to evaluate the results carefully and critically in any application.

The full text of this report, along with selected software developed during this project, can be accessed from the project web site at:
http://www.imperial.ac.uk/ewre/research/currentresearch/hydrology/improvedmethods

Guidance for users

Rainfall modelling is required for continuous simulation of extreme events using rainfall-runoff models, and in this project a suite of models has been developed and tested. The tools presented here include methods that can also be used to infill missing data (from daily raingauge networks) and to maximise the use of available daily and subdaily raingauge data to generate spatially distributed sub-daily rainfall.

The rainfall modelling problem is complex; not only is it necessary to capture the properties of extreme rainfall, but also the full range of rainfall characteristics including low intensity rainfall and interstorm rainfall properties. These determine the antecedent catchment wetness prior to a flood event that is an important control on runoff response. This study represents one of the most extensive analyses of rainfall simulation methods yet undertaken, but nevertheless some important limitations of the models are apparent under certain circumstances. Rainfall modelling is therefore not at the state of development where packages can be taken off the shelf and applied by an uninformed user. However, if used wisely and with due attention to data limitations and performance evaluation, the models represent powerful tools for addressing a range of important issues most notably impacts of climate variability and climate change.

Clearly, rainfall modelling cannot be considered in isolation; joint evaluation of rainfall and rainfall-runoff simulation methods is required. This report includes limited joint testing, with encouraging results. However, it is important that wider testing is undertaken and experience of the application of these methods is developed and is shared.

This project has undertaken development and testing of a range of tools for stochastic generation of rainfall and potential evaporation for use at varying spatial and temporal scales. Of the models developed, software is made available for those models which were considered to be ready for use within the UK given the testing in this report. This brief note provides guidance to the user regarding the software made available through the project. It is hoped that researchers and practitioners will be able to make use of these tools and provide feedback on their experience.

Three software products are to be made available through this project:

- Single-site hourly rainfall model: Random-Parameter Bartlett-Lewis Model
- Multi-site daily rainfall model: Generalized Linear Model (GLM)
- Disaggregation of multi-site daily rainfall to hourly module: GLM with Random-Parameter Bartlett-Lewis Model disaggregation

Single-site hourly rainfall model

The development and testing of the single site model has been detailed within Chapter 2. The Random-Parameter Bartlett-Lewis model with an exponential distribution for cell intensities was chosen for application across the UK based on a series of

individual tests for sites representative of rainfall across the UK. The model can be used to represent rainfall at an individual gauge or a catchment average rainfall.

For individual raingauges (using hourly data), parameter identification has proved a challenge and a new two-stage fitting method has been developed. Overall, the models are able to reproduce a wide range of statistical properties of rainfall, with monthly seasonality, including monthly and annual extremes. However, for some sites extreme value performance is poor. We conclude that the model and optimisation technique are appropriate for use with caution, with the user undertaking careful verification of the simulated outputs. More extensive testing is required on a wider range of sites before the model can be recommended for general application across the UK.

For catchment average hourly rainfall (detailed in Chapter 4), some issues arise where the subdaily model properties are based on an individual subdaily raingauge and the daily properties on a spatial average. These remain to be explored. However, the joint performance analysis (Chapter 5) based on three catchments has shown encouraging performance.

Multi-site daily rainfall model

The generalised linear models detailed in Chapter 6 have been applied to several networks of daily raingauges in the UK and Ireland including applications in flood design studies. They can simulate long sequences of daily rainfall accross a network of sites including effect of spatial location (e.g. topography, rainshadow) and temporal non-stationarity including climate change. Climatic predictors such as the North Atlantic Oscillation can be built in to the model and a time varying probability distribution of extremes can be generated. The models can also be used to infill missing data from a network of gauges.

The report includes a chapter on an application of the software to the Ouse catchment by a non-experienced user and provides some guidance not only to model performance but also to the user's experience of the software. The software is available with user manual.

Disaggregation of multi-site daily rainfall to hourly module

A procedure has been developed within Chapter 8 to generate a spatial distribution of subdaily rainfall for the typical situation where a network of daily raingauges exists and limited subdaily data. Observed daily data can be used directly, or long sequences can be generated using the GLM approach above. The basic concept is simple; a single subdaily timeseries is applied to each location of interest with appropriate linear scaling to match the daily total at that point. Observed subdaily data can be used, or alternatively a modelling procedure based on the Poisson cluster models of the type used in Chapter 2.

The simple concept of spatial transposition of the temporal subdaily sequence has been shown to perform well for the River Lee Catchment in North London. The combined modelling procedure reproduces spatial daily rainfall properties and point

subdaily properties well, but as expected tends to overestimate spatially averaged annual hourly rainfall maxima (due to the coincidence of the subdaily rainfall profiles). Nevertheless, the combined procedure based on observed daily data and a single observed subdaily profile performs well when used as input to a semi-distributed event based rainfall-runoff model. The method appears to be extremely promising, and to have wide applicability.

With respect to the software, a fully integrated procedure is not yet available in suitable form for release. However, the GLM software is available as discussed above and the temporal disaggregation software is available as a package HYETOS (developed outside this programme).

Other procedures developed within this project

The remaining modelling software discussed in this report is not considered to be in a suitable form for general availability at present.

Software and the project website

The software for model fitting and simulation will be available through the project website:
http://www.imperial.ac.uk/ewre/research/currentresearch/hydrology/improvedmethods

Contents

Chapter 1 Introduction

1.1 Introduction

This report details the findings of the DEFRA-funded research project FD2015: Improved methods for National Spatial-Temporal Rainfall and Evaporation Modelling for Broad Scale Modelling (BSM). BSM has been identified by MAFF (now DEFRA) as a priority thematic area and is required to provide a new level of decision-support for strategic assessment of flood risk and flood management, including assessment of impacts of climate and land use change. To achieve this, continuous simulation hydrological modelling is needed. This represents a major development in methodology, which has, in addition to its role in BSM, important benefits for flood design practice.

Traditional hydrograph-based methods of flood risk assessment focus on the evaluation of hydrological response to selected individual rainfall events, and commonly use a simplistic representation of spatial and temporal rainfall. Continuous simulation of the rainfall-runoff process has several potential advantages. For example, the dependence of runoff response on antecedent conditions is explicitly represented (of particular importance for the representation of impacts of climate change), and from a continuous output time series, the frequency of any hydrograph characteristic can be estimated directly. There are of course disadvantages to the continuous simulation approach. It adds considerable complexity to the modelling and simulation task. There is an increased computational burden, but this is tractable with current computer power. Continuous simulation methods will normally involve some extrapolation from the record lengths of fitting data to more extreme rainfall and flow events, and the extent to which they can capture the required extremes is still an open research question. In addition, representation of dry periods is important to establish antecedent conditions for flood runoff generation. Clearly, application of such methods requires understanding of the key properties of rainfall that are important for flood simulation, and it is likely that some compromises between different aspects of performance must be accepted.

Interest in continuous simulation has led to important new developments in rainfall-runoff modelling, for example in the regionalisation of rainfall-runoff models for ungauged catchment application (e.g. Lamb and Kay 2004; Wagener et al. 2004; McIntyre et al. 2006). Research is sufficiently well-developed that UK regionalisation for flood estimation is now feasible, and this is the subject of the companion research project FD2106, undertaken at CEH-Wallingford. However, for continuous simulation on both gauged and ungauged basins there is an associated requirement for time-series of precipitation (and potential evaporation) to provide input to such models. Long sequences (e.g. 1000 years) are required, at the hourly time-scale, which preserve appropriate extreme value properties. In addition, for any application other than very small catchments, representation of spatial rainfall is needed.

This project provides and validates regionally-applicable methods to generate the rainfall and evaporation sequences required for continuous simulation modelling. The development and testing of a suite of approaches to rainfall and evaporation simulation by a joint team at Imperial College London and University College London (ICL/UCL)

forms the bulk of the report. However, within FD2105, software has been provided to CEH-Wallingford for testing with some of the rainfall data and rainfall-runoff models used in FD2106, and this work is reported in Chapters 4 and 5. CEH has also used simulation software for daily spatial rainfall from this project in an unrelated rainfall-runoff modelling study, and this work is included for completeness in Chapter 13. A final element of the project is the provision by the ICL/UCL team of guidance concerning the significance of spatial rainfall for runoff modelling.

1.2 Background

The strategic role of BSM, the need for continuous simulation modelling, and the consequent requirement to simulate long sequences of climatological inputs has been described in Section 1.1 above. Further details are defined in the scoping study of the Broad Scale Modelling Hydrology Programme (Calver and Wheater, 2002).

The proposed research provides the core rainfall modelling elements required to underpin national application of continuous simulation and BSM and builds on and links two major lines of research carried out for MAFF (now DEFRA) by CEH-Wallingford and by the IC/UCL consortium, specifically projects:

- FD0404 Continuous simulation for flood frequency estimation (CEH)
- FD0426/FD2103 Generation of spatially-consistent rainfall data (IC/UCL)

This project aimed to deliver:

1. a national procedure for single-site (or catchment average) continuous rainfall and evaporation modelling;
2. a national procedure to model spatial variability of daily rainfall and to obtain sub-daily spatial rainfall by disaggregation;
3. a methodology to simulate detailed radar rainfall fields and accommodating spatial-temporal heterogeneity;
4. guidance to users concerning the relative importance of spatial rainfall properties and appropriateness of methods as a function of catchment type and scale.

An important element is model validation of the integrated modelling procedure, i.e. testing the combined performance of the rainfall and rainfall-runoff models.

As noted above, this project is closely linked to a parallel project (FD2106), recently completed by CEH-Wallingford. These two projects each provide core elements of a nationally-applicable procedure for continuous simulation modelling, the basic hydrological simulation tools underpinning BSM.

1.2.1 The structure of the work programme

The most basic requirement for continuous simulation modelling is a long time-series of single-site rainfall, for example representing a single raingauge or an areal average. Stochastic models are used, which aim to provide a simplified conceptual representation of the rainfall process using parameters that have physical significance. Storm arrivals are modelled, and then the arrival of (overlapping) cells with random intensities and durations. Following theoretical development of a family of suitable models by members of the team (Cox and Isham, 1988; Rodriguez-Iturbe et al., 1987, 1988), and various subsequent UK and international applications of these methods (see Onof et al., 2000), there was a need to review the performance of model variants for UK data, in particular with respect to parameter estimation and the associated parameter uncertainty and extreme value performance. Methods were also needed to generate long sequences of potential evaporation. Work Package 1 addressed these issues, together with the joint testing of rainfall and rainfall-runoff models.

More generally, there is a need to represent the spatial as well as temporal variability of rainfall, and a range of approaches has been explored. In Work Package 2, an approach to modelling rainfall in continuous space and time has been pursued. The approach is a spatial analogue of the models used in Work Package 1 to represent the temporal process of rainfall at a single site. The models use a hierarchical representation of rainfall, in which the smallest element is a rain cell - a localised area of relatively intense rainfall. It is known that cells tend to cluster within larger-scale structures that we call storms. These storms themselves cluster, to form rain events. At the spatial scale of UK catchments, these correspond to weather systems. We treat the interior of a rain event as a stationary stochastic process. However, there are many different types of event, and the parameters of this process will typically vary between event types. We therefore construct, for each month of the year, a library of model parameters to reflect the prevailing types of weather system during that month and simulate their arrivals. Details may be found in Wheater et al. (2000a), Wheater et al. (2006) and Northrop (1996). Previously, such models have generated rainfall fields that are statistically stationary in space (i.e. no systematic location or orographic effects) and in time (no climatic variability). In Work Package 2 we consider model extension to incorporate spatial non-stationarity. To identify the full model structure, radar data are required. Also in Work Package 2 we report the large effort that has been expended to obtain a consistent long radar time-series for the Chenies radar, and the associated performance of the model in simulating the resulting de-calibrated and re-calibrated data.

The methodology of Work Package 2 relies on the availability of long radar sequences for model development, calibration and testing. Such data are not yet routinely available - nor are the continuous space-time models sufficiently developed for immediate application in engineering practice. Hence in Work Package 3 we develop alternative approaches, based on readily available daily rain gauge data.

Many techniques are available for generating daily rainfall sequences. A powerful and flexible technique, which allows nonstationarities in time and in space to be quantified and incorporated, is that of Generalized Linear Models (GLMs), introduced into the study of daily rainfall by Coe and Stern (1982). The basic idea, which is an extension of linear regression, is to use the values of various predictors to forecast a probability

distribution for the amount of rainfall at a site on a given day. This distribution is specified in two parts - we model the probability of rainfall occurrence separately from the amount of rain if non-zero. Previous days' rainfall amounts can be used as predictors, to account for autocorrelation. Other predictors might include quantities representing regional variation (such as site altitude), seasonal variability, long-term trends and 'external' climatological factors such as the North Atlantic Oscillation (NAO). By defining suitable dependence structures between sites, it is possible to build a multivariate GLM that allows simulation of non-stationary daily rainfall sequences over a network of sites. In this project we have refined and extended the GLM software and report case study applications to two UK data sets.

GLMs allow us to simulate daily precipitation, including spatial dependence, location effects, and climate variability. However, for flood applications in the UK, hourly data are needed. Work Package 3 therefore addresses the problem of spatial-temporal disaggregation from daily to hourly time intervals. A simple approach is developed, whereby models of the form adopted in Work Package 1 are used to disaggregate daily data to hourly at a representative gauge location (for the case that measured sub-daily data are not available). The resulting sub-daily time series is then applied to all sites over the catchment, linearly scaled to reproduce the correct spatially-varying daily totals. This scheme is tested for the River Lee catchment North of London.

In Work Package 4 we address the issue of the significance of spatial rainfall and its representation for flood estimation. A literature review is reported, and a numerical rainfall-runoff modelling study, also based on the Lee catchment, in which alternative rainfall descriptors (i.e. raingauge networks of different density, radar data and disaggregated daily data) are considered and their effects on simulated runoff are evaluated. Finally in this section of the report we report an associated piece of work in which the performance of GLM-simulated daily rainfall as input to a rainfall-runoff model is investigated.

Work Package 5 is concerned with dissemination of the results and tools outlined above.

1.3 Work Programme and Objectives

As noted above, the project is divided into Work Packages (WP) - each of which addresses a particular aspect of the overall modelling objective - the simulation of rainfall and potential evaporation to support continuous simulation modelling.

The objective of each work package as laid out in the project contract is stated below. All objectives have been met with the exception of those signified by bracketing, which, through consultation with Defra have been modified in light of research results or data issues.

Work Package 1: Development of a national capability for single site rainfall and evaporation time-series modelling

- Objective 1a - Intercomparison of rainfall models; development and evaluation of evaporation models, including representation of rainfall and evaporation

seasonal dependence and interactions; joint testing of rainfall, evaporation and rainfall-runoff models;

- Objective 1b - (Generalisation of model parameters for national application. Results to feed into Work Package 5.) As agreed with DEFRA, this was modified in the light of research results and focused on proving a methodology for modelling areal-averaged catchment rainfall, rather than single site.

Work Package 2: Development and testing of non-stationary spatial-temporal rainfall model for medium-term application

- Objective 2a - Model development and testing using radar data; evaluation of fractal and Poisson-based methods;
- Objective 2b - (Integrated performance evaluation of spatial-temporal models and distributed rainfall-runoff models. Results to feed into Work Package 5). As agreed with DEFRA, the extension of the continuous space-time model has not been completed due to data issues. The project reports progress and makes recommendations for the further research needed in this area.

Work Package 3: Development and testing of short-term spatial-temporal modelling, based on daily rainfall modelling and spatial-temporal disaggregation

- Objective 3a - Software development of Generalised Linear Models for national application to model spatially-distributed daily rainfall (with spatial non-stationarity and temporal climate variability)
- Objective 3b - Development and testing of procedure for temporal disaggregation of multisite daily rainfall.
- Objective 3c - Testing the GLM and disaggregation methodologies with distributed rainfall-runoff modelling.

Work Package 4: Effect of spatial structure of (large) catchment-scale rainfall on runoff estimation

- Objective 4 - Data analysis of selected radar rainfall events and long-term raingauge data; simulation studies using distributed rainfall-runoff models for different catchment types and scales; provision of guidance to users as input to Work Package 5.

Work Package 5: Dissemination of continuous rainfall and evaporation simulation procedures

- Objective 5a - To develop procedures from the best of the scientific methods
- Objective 5b - To produce, in consultation, guidance documentation for users
- Objective 5c - To develop, in consultation, dissemination uptake strategies.

1.4 Summary and report map

This chapter has provided an introduction to the modelling tasks undertaken within this project, and the scientific objectives of the task have been defined.

The remainder of the report is broken into a series of four parts, with each part corresponding to a Work Package. These parts have been ordered to present the methods which are considered ready for widespread application first (Part I - single site model, and Part II - the non-stationary multi-site daily model with associated disaggregation to hourly timescale technique), with caveats regarding their application are detailed at the end of each chapter. The spatial-temporal model based on radar data (developed as part of Work Package 2) is detailed in Part III. Part IV details a literature review and experimental results regarding the significance of spatial rainfall for flood estimation.

Each chapter details the methodology undertaken, results of application and recommendations for sub-objectives contained within each part. Conclusions and recomendations are made within Chapter 14

This report, and the methods contained within fulfil the objective 5a and 5b of Work Package 5. Dissemination is underway through a series of presentations at national and international meetings, publication in international refereed journals and through a stakeholder meeting to be held on the 20th Feb 2006. Selected software is available from the authors.

Part I

Development of a national capability for single site rainfall and evaporation time-series modelling

The task addressed here is the development of a national procedure for single-site or catchment average rainfall and evaporation modelling for the continuous simulation of streamflows.

In the literature, a number of single-site models have been developed and validated with a range of rainfall data (Rodriguez-Iturbe et al., 1987, 1988, Onof et al., 2000). The work presented in Chapter 2 identifies the most appropriate model for application in the UK. Because it is not feasible to express the likelihood function in a tractable form, these models have generally been fitted using the Generalized Method Of Moments (GMOM). Appendix C summarises the analytical expressions of the most important statistics characteristic of rainfall. The GMOM is however unable to ensure a good reproduction of extreme values since it is based upon the use of statistics reflecting the average behaviour of rainfall. Moreover, a number of authors have found the parameter estimates it produces to be poorly identified (Velghe et al., 1994; Verhoest et al., 1997). The work presented here therefore addresses the issue of the uncertainty in parameter estimation. It exploits the poor identifiability characterising the GMOM method to develop a two-phase fitting method. This chooses among parameters with non-significantly different performances in terms of the average rainfall characteristics,those which optimally reproduce the extremal behaviour of the observed rainfall.

This methodology is used to compare the performance of a number of models based upon Poisson-cluster processes at three different sites in the UK. A model is chosen on the basis of the GMOM, and the model variant is then selected on the basis of its optimal extreme-value performance. The chosen model is the Random-Parameter Bartlett-Lewis model with an exponential distribution for cell intensities. Its reproduction of extreme-values is significantly improved as a result of the two-phase fitting method, as compared with the performance recorded in the literature (Onof et al., 2000). Model parameter uncertainty results using various record lengths also appear to identify genuine opportunities for parameter bounding, which with further work will aid application of these models in a national setting.

Although the performance is improved with the introduction of the two-phase fitting method, some deficiencies remain in terms of reproduction of extreme statistics. We conclude that the model and optimisation technique are appropriate for use with caution, with the user undertaking careful verification of the simulated outputs. More extensive testing is required on a wider range of sites before the model can be recommended for general application across the UK. It is recommended that alternate means of improving the extreme value performance of the model (such as the use of skewness in fitting) are compared to the newly introduced fitting technique.

In Chapter 3, methods are presented for the modelling of Potential Evaporation (PE). The models developed here are heteroscedastic autoregressions. They allow for the simulation of weekly or daily sequences, and also that of realistic daily sequences consistent with a given weekly sequence. The weekly PE is modelled using joint mean-variance models. Seasonality and linear trends in observed MORECS data are represented successfully by these joint models. Using daily data from a single site, a model is developed which represents daily PE as multivariate normal distribution conditional on given weekly evaporation totals. This can either be used as a disaggregation tool to generate daily PE values from weekly (e.g. MORECS) PE data, or together with the weekly PE generator to enable the simulation of daily PE values.

Chapter 4 presents the results from an extensive program of testing that examines the performance of the single site rainfall models for several UK catchments. Catchment average rainfall data are used, as a precursor to joint testing of the rainfall and rainfall-runoff modelling procedure developed under FD2106.

The results of fitting the single-site model to these data are broadly satisfactory, but there are some notable discrepancies. In particular, some problems in the characterisation of mean rainfall depth and the persistence of wet and dry episodes have been identified. A first analysis found that the original formulation of the model led to the appearance of an unrealistically large number of high magnitude rainfall events. This was resolved by a small modification to the simulation methodology. The modified simulation scheme was tested on three catchments, where the results were much improved. Further testing is required before conclusions regarding the model's performance in different regions in UK can be reached.

Part of the problem of reproducing rainfall characteristics may lie in the way in which the hourly catchment average data series is produced. Hourly catchment average series are obtained by first averaging single-site daily series for all the available gauges in the catchment. Using the hourly profile at a recording raingauge with data of acceptable quality (or a weighted average of such profiles if several such gauges are available), hourly catchment averages are obtained. It is emphasised that the models applied here were designed for use with single-site rainfall (not catchment average rainfall) - with no account taken of catchment statistics being different than those of a single site. It is recommended that the models applied here are further tested regarding the conditions of applicability in terms of catchment size and data availability. Analysis of observed data might allow this correction to be estimated empirically, improving applicability for larger areas.

Chapter 5 completes the investigation into the use of the selected single-site model for the representation of catchment averages. The focus is on the ability to generate realistic streamflows by inputting the generated rainfalls into the continuous rainfall-runoff model developed in FD2106. The analysis has shown that hourly flow data are generally well simulated. However, flow duration curves based on simulated data are generally underestimated, while an analysis of peak-over-threshold (POT) series has shown that flood peaks are generally overestimated, especially for return periods greater than ten years. The underestimation of the flow duration curve is probably linked to underestimation of rainfall persistence within the rainfall model. Consequently, antecedent catchment wetness, conditioned by the presence or absence of antecedent rainfall, may not be adequately simulated and estimates of flow are therefore too low. Some overestimation of flood frequency curves may result from the occasional generation of very large storms by the rainfall model.

We conclude that the models are appropriate to use with caution, but careful assessment on a large number of catchments would be needed before the single-site Random-Parameter Bartlett-Lewis rainfall generator could be used with confidence to provide data for flood risk assessment for ungauged catchments. Any user should validate the model using local data to have sufficient confidence for practical application. More work is recommended to pursue the issue of parameter constraints and non-feasible simulations. There are also ideas for model extension that should be explored to seek improvements in performance.

Chapter 2 Single-site model selection and testing

2.1 Introduction

This chapter aims to recommend a single-site rainfall model for national application, and to provide guidelines on its use. To identify a model for routine national use, the following strategy has been followed:

1. Identify a large number of candidate models on the basis of an extensive literature review. The results of this are summarised in Lekkas (2002).
2. Critically evaluate all of the candidate models and identify a few strong contenders. The criteria used here are that a model should ideally have a long track record of successful and varied application or that, failing this, it should have clearly demonstrated potential to outperform better-established models.
3. Carry out extensive tests of the strong contenders identified in step 2, and make a final recommendation based on the results of these tests.

Section 2.2 provides an overview of the models that have been considered in steps 1 and 2 above.

As well as recommending a model, the chapter discusses issues relating to model fitting. Building upon the two model comparison steps discussed in Lekkas and Onof (2003) (namely the assessment of models in terms of both 'average' and 'extreme' rainfall behaviour), the fitting is done in two stages, as follows:

1. Several sets of parameters are identified that provide good reproduction of 'average' rainfall properties. The numerical optimisation method for this stage replaces that discussed in Lekkas and Onof (2003) as a more efficient tool for the dual purpose of parameter identification and uncertainty assessment. Parameter identifiability is addressed via the method presented in Appendix A
2. Simulations are used to determine the extremal properties of each parameter set identified in stage 1; the parameter set that provides the best reproduction of extremal properties is chosen.

The methodology for the first stage of fitting is reviewed in Section 2.3 of this chapter, and results are presented in Section 2.4. Methodology and results for the second stage are presented in Section 2.5. The results presented in Sections 2.4 and 2.5 are for three gauges which are representative of the rainfall variability over England and Wales. On the basis of these results, recommendations are given in Section 2.6.

2.2 Models considered

This section summarises the models that have been considered in this exercise. Section 2.2.1 provides an overview of the basic model structure; Section 2.2.2 gives

details of the models that have been explored in more detail; and Section 2.2.3 summarises those models from the literature that were discarded after initial consideration (see Section 2.1).

2.2.1 Structure of the models

All of the models considered here are Poisson-based Rectangular Pulse models, which represent rainfall via a continuous-time Poisson process of storm arrivals. In their simplest form, the models represent a storm as a single rainfall cell. The inadequacy of this representation has led to the use of Poisson-cluster processes (Rodriguez-Iturbe et al., 1987). These feature, within each storm, a clustering mechanism that generates a process of cell arrivals. A rectangular pulse of rainfall which has a random depth and a random duration is associated with each cell. Two distinct clustering mechanisms are reported in the literature: the Bartlett-Lewis and the Neyman-Scott processes. In the first, cells arrive according to a renewal process starting at the storm arrival time. This process is truncated by a randomly distributed storm activity duration after which no cells arrive. The Neyman-Scott mechanism generates cells as independent and identically distributed from the storm arrival time. A distribution is chosen for the number of cells per storm.

To complete the description of both clustering mechanisms, it is necessary to specify the associated distributions. In the Bartlett-Lewis Rectangular Pulse model considered below, the cluster duration is exponential and the intercell intervals are exponential so that cells within a storm arrive according to a finite Poisson process. The number of cells per cluster is therefore geometric. In the Neyman-Scott Rectangular Pulse model, the time from the storm origin to each cell origin is exponential and the number of cells per cluster is Poisson. In the remainder of this chapter these two models are referred to, respectively, as the Bartlett-Lewis model and the Neyman-Scott model.

A review of the available single site rainfall models based upon Poisson-cluster processes has been presented in Lekkas (2002). A subset of these models has been selected for comparison here; this subset encompasses most of the advances that have been introduced over the last decade and a half. For each model, analytical expressions for a range of rainfall statistics at any time-scale are given in appendix B. These statistics are, for time-scale h hours, the:

- mean rainfall intensity, $M(h)$.
- variance of rainfall intensity, $V(h)$.
- autocovariance at lag k of the rainfall intensity, $C(k,h)$ (whence the autocorrelations $A(k,h)$ are also available).
- proportion of dry periods, $P_d(h)$.

We now describe each of the selected models in detail.

Table 2.1 Poisson model parameters

Parameter	Definition	Units
λ	Cell arrival rate	hr^{-1}
μ_x	Mean cell depth	mm hr^{-1}
μ_{x^2}	Expected square of the cell depth	mm^2hr^{-2}
η	Cell duration parameter	hr^{-1}

2.2.2 Models considered for detailed investigation

Most of the models considered here are modifications of the Bartlett-Lewis model described above. For example, in Section 2.2.2 we consider a model whose cell arrival structure is identical to that in the Bartlett-Lewis model, but which has more than one type of cell. We also consider different model variants, resulting from different choices for the cell depth distribution. Similar modifications of the Neyman-Scott model can be considered in principle. However, it is generally accepted that the two clustering mechanisms perform comparably in practical terms. Moreover, there are theoretical arguments for their equivalence (Cowpertwait, 1998), and therefore we do not consider these modifications here: rather, we seek to verify the equivalence by comparing the unmodified versions of each model. A simpler Poisson Rectangular Pulse model (hereafter, the Poisson model) will also be used, to provide a benchmark against which to evaluate the improvements brought about by the use of a clustering mechanism.

The remainder of this section provides detailed descriptions of each of the models considered. The notation and parameterisation are the same as those in appendix B. In general, cell depth distributions are parameterised in terms of their first two moments. A more detailed specification is not required for model fitting (see Section 2.4 below), although obviously the precise choice of depth distribution will affect some properties (such as extremes) of the model simulations.

The Poisson model

In the Poisson model, cells arrive according to a Poisson process of rate λ. Cell durations are independent and exponentially distributed with parameter η (so that the mean cell duration is $1/\eta$ time units). The cell depth is distributed with mean μ_x and variance $\mu_{x^2} - \mu_x^2$. The Poisson model therefore has 4 parameters in total, summarised in Table 2.1.

The Bartlett-Lewis model

In the Bartlett-Lewis model, storms arrive according to a Poisson process of rate λ. After each storm arrival, there follows a Poisson process of cell arrivals of rate β. Cells arrive during a time (the duration of storm activity) that is exponentially distributed with parameter γ. As with the Poisson model, cell depths are distributed with mean μ_x and variance $\mu_{x^2} - \mu_x^2$, and durations are independent and exponentially distributed with parameter η. However, since this model is more realistic than the Poisson model we

Table 2.2 Bartlett-Lewis model parameters

Parameter	Definition	Units
λ	Storm arrival rate	hr^{-1}
μ_x	Mean cell depth	mm hr^{-1}
μ_{x^2}	Expected square of the cell depth	mm^2hr^{-2}
η	Cell duration parameter	hr^{-1}
β	Cell arrival rate	hr^{-1}
γ	Storm duration parameter	hr^1

Table 2.3 Neyman-Scott model parameters

Parameter	Definition	Units
λ	Storm arrival rate	hr^{-1}
μ_x	Mean cell depth	mm hr^{-1}
μ_{x^2}	Expected square of the cell depth	mm^2hr^{-2}
η	Cell duration parameter	hr^{-1}
β	Cell arrival parameter	hr^{-1}
μ_c	Mean number of cells per storm	

explore different variants, characterised by the use of the exponential, gamma or Pareto distributions for the cell depth. For the gamma and Pareto versions (which are indistinguishable as far as model fitting is concerned, at least using the methods considered in Section 2.4 below), the model has six parameters; for the exponential version, it has just five since in this case, $\mu_{x^2} = 2\mu_x^2$. The model parameters are summarised in Table 2.2.

The Neyman-Scott model

In the Neyman-Scott model, storms arrive according to a Poisson process of rate λ. The times separating storm from cell arrivals are exponentially distributed with parameter β. The number of cells per storm is Poisson distributed with mean μ_c. Cell durations are exponentially distributed with parameter η. As with the Bartlett-Lewis model, three Neyman-Scott variants are considered arising from different cell depth distributions. This model has either five or six parameters, depending on the variant chosen; these are summarised in Table 2.3.

Random Parameter Bartlett-Lewis model

This is the first of the 'modified' models considered in this chapter. It was introduced by Rodriguez-Iturbe et al. (1988), and is also known in the literature as the 'random η' model and in appendix B, as the 'Linear Random Parameter' model. It is a randomised version of the original Bartlett-Lewis model (Rodriguez-Iturbe et al., 1987). In this

Table 2.4 Random Parameter Bartlett-Lewis model parameters

Parameter	Definition	Units
λ	Storm arrival rate	hr^{-1}
μ_x	Mean cell depth	mm hr^{-1}
μ_{x^2}	Expected square of the cell depth	mm^2hr^{-2}
α	Shape parameter of gamma distribution for η	
ν	Scale parameter of gamma distribution for η	hr
κ	$1/\eta \times \beta$ Cell arrival parameter	
ϕ	$1/\eta \times \gamma$ Storm duration parameter	

randomised version, the cell duration parameter η of the original version becomes a random variable, sampled for each storm from a gamma distribution. To preserve the temporal structure of storms, the cell arrival rate β and storm duration parameter γ of the original model are scaled proportionally to η. The effect is to allow different storms to operate on different timescales.

Although, in the original presentation of this model, an exponential distribution was used for the cell depth distribution, a two-parameter distribution has been found to be preferable for the reproduction of extreme values (Onof and Wheater, 1994). As a result, three versions of this model will be investigated, according to the distribution of cell intensities: the exponential, gamma and Pareto distributions will be considered.

The gamma and Pareto versions of this model have seven parameters each (see Table 2.4). The exponential model only requires six parameters since $\mu_{x^2} = 2\mu_x^2$.

N-Cell Bartlett-Lewis model

Another possible modification to the basic Bartlett-Lewis model is to allow for more than one type of cell. The idea is that any given storm is likely to include precipitation generated by different mechanisms. Cowpertwait (1994) thus presented a 2-cell Neyman-Scott model to represent the co-existence within any given storm of 'heavy' short duration convective cells and 'light' long duration stratiform cells. In this investigation, a Bartlett-Lewis model with two cell types is considered. We shall see further, as a result of the investigation of this model, that including a larger number of cell types would not be warranted.

As for the Random Parameter Bartlett-Lewis model, exponential, Gamma and Pareto distributed cell intensities will be investigated. The gamma and Pareto versions have 10 parameters each (see Table 2.5). The exponential version only requires 8 parameters since $\mu_{x_1^2} = 2\mu_{x_1}^2$ and $\mu_{x_2^2} = 2\mu_{x_2}^2$.

2.2.3 Other models

In addition to those described above, various other modifications have been proposed to the basic Bartlett-Lewis and Neyman-Scott models. In this section we briefly review

Table 2.5 Parameters in the 2-Cell Bartlett Lewis model.

Parameter	Definition	Units
λ	Storm arrival rate	hr^{-1}
μ_{x_1}	Mean cell depth (type 1 cells)	mm hr^{-1}
μ_{x_2}	Mean cell depth (type 2 cells)	mm hr^{-1}
$\mu_{x_1^2}$	Expected square of the cell depth (type 1 cells)	mm^2hr^{-2}
$\mu_{x_2^2}$	Expected square of the cell depth (type 2 cells)	mm^2hr^{-2}
η_1	Cell duration parameter (type 1 cells)	hr^{-1}
η_2	Cell duration parameter (type 2 cells)	hr^{-1}
ψ	Proportion of type 1 cells	
β	Cell arrival rate	hr^{-1}
γ	Storm duration parameter	hr^1

these, and explain why they will not be considered further in this chapter.

Dependent Depth-Duration Bartlett-Lewis model

Observations indicate that intense rainfall is of shorter duration than less intense rainfall. This provides some motivation for exploring a class of models which incorporate an inverse relationship between cell duration and depth. The approach has previously been introduced to both Neyman-Scott (Kakou, 1997) and Bartlett-Lewis (Wheater et al., 2000a) models, with very similar resulting performance.

Unfortunately, the published results in relation to this modelling approach appear to be incorrect. For example, Wheater et al. (2000a, Tables 5 and 6) provide parameter estimates, along with fitted properties, for a Dependent Depth-Duration Bartlett-Lewis model fitted to data from South-West England. However, if the quoted parameter estimates are substituted into the appropriate expressions for fitted properties in Kakou (1998) and Onof (2003), they do not yield the claimed values. We conclude that there were errors in the software used to produce the published results. Attempts to obtain convergence of the objective functions used in the first phase of the fitting algorithm presented below (section 2.3) failed. Since this class of models cannot be considered to have been adequately validated and parameter identification is problematic, it is not a serious contender for national use at this stage.

Extreme Cell Random Bartlett-Lewis model

This model is based upon the work of Cameron et al. (2001). It is a modification of the Random Parameter Bartlett-Lewis model in which, when the cell depth is larger than a certain threshold, it is re-sampled from a Generalised Pareto Distribution. The motivation is to try and model the upper tail of the rainfall intensity distribution directly, to ensure good reproduction of extremes. However, the results presented by Cameron et al. (2001) are questionable insofar as they use data from Plynlimon that are not properly quality controlled. The winter months are characterised by many days where

only a daily total was actually recorded, and the hourly data were derived by division by 24. This model can therefore not be considered to have been adequately validated.

We also note that the model presents certain features which are not attractive:

- The model distinguishes between rainfalls above and below a threshold. However, both rainfall types are lumped together to estimate the Bartlett-Lewis model parameters, including the parameters of the distribution of rainfall intensities below the threshold. Strictly speaking, this is incorrect since the rainfall intensity distribution for rainfalls below the threshold should be obtained by fitting to data from which extremes above the threshold are excluded. This is of course impossible since cell intensities are not observed.
- The model introduces a threshold which is chosen so as to, for instance, improve the reproduction of hourly extremes. But this means that a parameter which in effect depends upon time-scale is introduced to a continuous-time model.

These observations suggest that implementing this Extreme Cell model could only be justified if we observed a systematic underestimation of extreme rainfalls by all other models. We shall therefore not include it in our comparison, but consider it as an option should this comparison identify no satisfactory model.

Quadratic Random Parameter Bartlett-Lewis Model

In the Random Parameter Bartlett-Lewis model of Section 2.2.2, each storm's cell arrival rate β depends linearly upon its cell duration parameter η. Onof (2003) suggested that increased flexibility could be obtained by using a quadratic relationship instead. However, the resulting Quadratic Random Parameter Bartlett-Lewis model is not considered here since certain algebraic expressions for model properties are, as yet, not available for its calibration (Onof, 2003).

Random Parameter Neyman-Scott Model

A randomised version of the Neyman-Scott model, directly analogous to the Random Parameter Bartlett-Lewis model of Section 2.2.2, was introduced by Entekhabi et al. (1989). We have not attempted to fit this model here. This is because, as far as we are aware, for this model no analytical expression is available for the probability that an interval is dry; this property is routinely used in model fitting. In principle, the required probability could be obtained by numerical integration of the corresponding expression for the original Neyman-Scott model. The latter expression is given by Cowpertwait (1991, equation 9b), but is itself difficult to compute. Moreover, expressions for the second-order properties of this model, given by Entekhabi et al. (1989), are in fact approximations that are only valid in certain parts of the parameter space. In our experience, it is not uncommon for a numerical optimisation routine to stray outside these regions during parameter estimation (see Section 2.3 below for details of the estimation procedure); this can lead to unexpected results when fitting.[1]

[1] A similar problem was initially encountered with the Bartlett-Lewis models, for which the expressions given by Rodriguez-Iturbe et al. (1987, 1988) are also approximations. However, in this case more

Table 2.6 Summary of models used in the comparison exercise.

Model	Description	No. of parameters
1	Poisson model with 2-parameter cell depth distribution	4
2	Bartlett-Lewis model with 2-parameter cell depth distribution	6
3	Neyman-Scott model with 2-parameter cell depth distribution and Poisson number of cells per storm	6
4	Random-parameter Bartlett-Lewis model with exponential cell depth distribution	6
5	Random-parameter Bartlett-Lewis model with 2-parameter cell depth distribution (gamma and Pareto versions)	7
6	Bartlett-Lewis model with 2 cell types, each with 2-parameter depth distribution	10

5 km^2 Neyman-Scott Model

It is noted that a daily version of the Neyman-Scott Model applicable over 5 km^2 has recently been developed (Kilsby et al., 2006), and potentially be extended to the hourly level.

2.2.4 Model variants considered

Section 2.2.2 sets out five different models for detailed analysis. As already indicated, for any model we may consider exponential, gamma and Pareto cell depth distributions, leading to three variants for each of the five models, i.e. 15 models in total. To explore all of these model combinations would be prohibitive. We therefore focus on a subset, summarised in Table 2.6, from which conclusions can be drawn regarding the main questions of interest. For example, models 2 and 3 provide a comparisons of the Neyman-Scott and Bartlett-Lewis clustering mechanisms; models 2 and 4–6 provide comparisons of 'original' versus 'modified' models; models 5 and 6 provide a comparison of 'random-parameter' versus 'multiple cell models'; and models 4 and 5 provide a comparison of 'exponential' versus '2-parameter' cell depth distributions.

In Table 2.6, most of the models involve unspecified 2-parameter depth distributions. As indicated in Section 2.2.2, it is not necessary to specify the precise form of these distributions for the work reported in Section 2.4 below. Subsequently however, the choice of distribution becomes important in determining the ability of a model to reproduce extreme rainfall behaviour. In Section 2.5 therefore, two specific versions of model 5 will be considered, based on gamma and Pareto cell depth distributions

accurate approximations have been developed using beta functions (see Onof 2003); it is not clear that the same thing can be done for the Neyman-Scott models.

respectively.

2.3 Model fitting

2.3.1 Methodology

Parameter estimation is difficult for models based on Poisson cluster processes, because the parameters are related rather indirectly to observable properties of rainfall sequences. Fitting techniques may be classified broadly into moment, likelihood and Bayesian methods. The latter two both allow formal comparisons between models, but require the formulation of a likelihood function. This is difficult for the models considered here, because of their complex temporal dependence structure. Attempts have been made by Chandler (1997), who formulated an approximate Gaussian 'spectral likelihood' based on collections of sample Fourier coefficients; and more recently by Northrop (2004), who derived a marginal likelihood for a subset of the parameters in a simple, non-clustered model. However, it has been argued (Rodriguez-Iturbe et al., 1988) that since the models are necessarily over-simplified, inference using likelihoods may not be appropriate in any case.

In the absence of a suitable likelihood function, estimation for these models is usually carried out using a generalised method of moments. This involves choosing parameters that achieve as close as possible a match, according to a weighted least-squares criterion, between the observed and modelled values of selected properties. Specifically, let $\theta = (\theta_1 \ \ldots \ \theta_p)'$ be the parameter vector for the model, let $\mathbf{T} = (T_1 \ \ldots \ T_k)'$ be a vector of summary statistics computed from the data, and denote by $\tau(\theta) = (\tau_1(\theta) \ \ldots \ \tau_k(\theta))'$ the expected value of $\mathbf{T}$ under the model. Then θ may be estimated by minimising an objective function

$$S(\theta|\mathbf{T}) = \sum_{i=1}^{k} w_i [T_i - \tau_i(\theta)]^2 \ . \tag{2.1}$$

for some collection of positive weights $\{w_i : i = 1, \ldots, k\}$.

An overview of moment-based inference is provided in appendix A. A summary of the most important aspects is as follows:

1. The weights $\{w_i\}$ must not be parameter- or data-dependent; otherwise the moment estimators are biased.

2. Under fairly general conditions, the minimiser of (2.1) has approximately a multivariate normal distribution for large samples. This can be used, for example, to construct approximate confidence intervals for the model parameters. The distribution has mean θ_0 (the true parameter vector) and covariance matrix $\mathbf{V} = \mathbf{H}^{-1}\mathbf{J}\mathbf{H}^{-1}$. Here, $\mathbf{H}$ is the expected second derivative of the objective function, estimated using the observed Hessian at the minimum of (2.1). $\mathbf{J}$ is the covariance matrix of the objective function derivatives, estimated as $4\mathbf{W}'\mathbf{C}\mathbf{W}$ where $\mathbf{W}$ is a $k \times p$ matrix with (i, j)th element $w_i \left(\partial\tau_i/\partial\theta_j\right)_{|\theta=\hat{\theta}}$ and $\mathbf{C}$ is an estimate of the covariance matrix of $\mathbf{T}$. If $n > 1$ years of data are

available, fitting properties $\mathbf{T}_1, \ldots, \mathbf{T}_n$ can be computed separately for each year: $\mathbf{T}$ can then be taken as the mean over all years, and $\mathbf{C}$ as $[n(n-1)]^{-1} \sum_{i=1}^{n} (\mathbf{T}_i - \mathbf{T})(\mathbf{T}_i - \mathbf{T})'$. For this to work, however, it is important to use estimators that are, as far as possible, unbiased in small samples, since some components of each $\mathbf{T}_i$, in particular those relating to daily data, will be computed from small numbers of observations.

3. An alternative way to construct confidence intervals uses the objective function itself. Specifically, to calculate a confidence interval for an individual parameter θ_j, it is first necessary to calculate a profile objective function $\tilde{S}(\theta_j)$, obtained for each value of θ_j by minimising (2.1) with respect to all of the remaining parameters. Denote by $\hat{\theta}_j$ the overall estimate of θ_j, corresponding to the unconstrained global minimum of (2.1). Define v_j and h_j^{-1} to be the jth diagonal elements of $\mathbf{V}$ and $\mathbf{H}^{-1}$ respectively, (i.e. the diagonal elements corresponding to the parameter θ_j of interest). Then an approximate confidence interval at a specified level consists of all points θ_j such that

$$\frac{2\left[\tilde{S}(\theta_j) - \tilde{S}(\hat{\theta}_j)\right]}{h_j v_j} \tag{2.2}$$

is less than the appropriate percentile of a chi-squared distribution with 1 degree of freedom. Notice that the threshold defined in this way will typically vary between parameters, since the quantity $h_j v_j$ will usually depend on j.

4. Confidence regions for the entire parameter vector can be constructed directly from the objective function (2.1). An approximate confidence region at a specified level consists of all points θ such that

$$a^{-1}\left\{2\left[S(\theta|\mathbf{T}) - S\left(\hat{\theta}|\mathbf{T}\right)\right] - c\right\} \tag{2.3}$$

is less than the appropriate percentile of a chi-squared distribution with b degrees of freedom. The constants a, b and c are given by

$$a = \frac{|\kappa_3|}{4\kappa_2} \qquad b = \frac{8\kappa_2^3}{\kappa_3^2} \qquad c = \kappa_1 - ab \ ,$$

with $\kappa_r = 2^{r-1}\Gamma(r)\mathrm{tr}\{[\mathbf{VH}]^r\} = 2^{r-1}\Gamma(r)\mathrm{tr}\left\{\left[\mathbf{H}^{-1}\mathbf{J}\right]^r\right\}$.

To fit a model using the method of moments, a choice has to be made regarding the summary statistics to include in the vector $\mathbf{T}$. The choice is limited by the need to compute $\tau(\theta)$, the vector of corresponding properties under the model. At least in the first instance, it is desirable to use properties for which analytical expressions are available — otherwise it is necessary to resort to simulation, which can be computationally expensive. For most of the models under consideration, analytical expressions are available for properties such as the mean, variance, autocorrelation function and proportion of dry intervals at different time scales (see Section 2.2.1). In the literature, these models are often fitted using several of these properties at time scales ranging from 1 to 24 hours. In Section 2.4 below, we explore the effect of including different statistics in the estimation, along with the effect of varying the weights $\{w_i\}$.

It was noted above that the fitting statistics should be unbiased in small samples. Statistics such as the mean, variance and proportion of dry intervals meet this requirement. However, the standard estimator of an autocorrelation coefficient can be seriously biased. Methods for correcting this are given by Kendall and Ord (1990, page 79), for example. In the work reported below, to compute the autocorrelation at lag k for any time series we have used Quenouille's bias-corrected estimator

$$r_Q(k) = 2r(k) - \frac{r_1(k) + r_2(k)}{2} ,$$

where $r(k)$, $r_1(k)$ and $r_2(k)$ are the results of applying the standard estimator to the entire series, the first half and the last half of the series respectively.

Models that incorporate skewness in fitting

It is recognised that there have been recent developments with the Neyman-Scott model which include skewness in fitting (Cowpertwait, 1998; Kilsby et al., 2004). These extensions have not been implemented here, however alternate fitting techniques (detailed within 2.5) are introduced to accommodate closer fit to extreme rainfall characteristics.

2.3.2 Approaches to minimising the objective function

Usually, minimisation of the objective function (2.1) must be done numerically. This can be difficult: experience has shown that it often has local minima, and that poor starting values for an optimisation can result in the failure of standard gradient-based algorithms (e.g. quasi-Newton and conjugate gradient methods). However, given good starting values the use of gradient information usually increases the precision with which a minimum can be identified. Two different minimisation approaches have been explored in the current project, as follows:

- Shuffled Complex Evolutionary Algorithm (Duan et al., 1993). This algorithm is primarily designed to avoid converging to local optima while remaining sufficiently efficient. For an n-parameter problem, complexes of $2n+1$ parameters are independently evolved using a downhill simplex method. The simplex parameter sets are selected on the basis of a probability distribution which gives greater weight to parameter sets that yield an improvement of the objective function. There is a shuffling set in which the evolved complexes are mixed, and the new complexes are then involved in turn, until convergence.
- Use of built-in optimisation routines within the R statistical computing environment (R Development Core Team, 2003). R comes with two nonlinear optimisation commands, `nlm` and `optim`. The former uses a Newton-type algorithm, and the latter offers several different options including the Nelder-Mead simplex method (see, e.g. Gershenfeld 1999), which is robust but slow since it does not use gradient information. After some experimentation, an algorithm has been developed that starts off using Nelder-Mead to identify promising regions of the parameter space, and then uses `nlm` with numerical derivatives to refine the optimisation.

Both of these approaches have been tested extensively, and compared by fitting the same models to the same data. Often, the results obtained were very similar in terms of objective function values. However, the values found by the R routines were invariably slightly smaller than those from the shuffled complex algorithm. This is to be expected, since the R routines use gradient information after the initial search. Further advantages of the R routines include:

1. For the shuffled complex algorithm, bounds must be set on each parameter prior to optimisation. It could be argued that this is not a problem, since some parameter values are physically unreasonable. However, it has been found that when fitting Poisson cluster models using (2.1), the method often yields optimal parameter estimates on a boundary. A consequence of this is that if, say, an upper bound of 100 is routinely set for a particular parameter, in many cases this effectively determines its estimated value. This is slightly unsatisfactory. A more appropriate approach is to find an unconstrained global minimum, and then to use the theory outlined above to determine an 'almost optimal' region of the parameter space, in the sense that the objective function throughout this region does not differ significantly from the global minimum. Other criteria, such as physical realism or reproduction of extremes (see Section 2.5 below) can then be used to select a parameter set from within this region.

2. The gradient-based methods in the R routines yield the Hessian $\mathbf{H}$ as a by-product. This in turn allows the theory outlined above to be implemented.

In the light of these comparisons, all of the results reported below have been obtained using the R routines to minimise (2.1). A summary of the algorithm is as follows:

1. Perform M separate Nelder-Mead optimisations, each beginning from a different initial value of θ. One of these initial values is supplied by the user; the remaining $M-1$ are generated via random perturbations about this.

2. Use a set of heuristics to bound each parameter if possible, based on the best m parameter sets found so far (i.e. those with smallest objective function values). Bounds are introduced or tightened if the best sets are in close agreement, and relaxed if the best sets are close to an existing boundary.

3. Discard all but the m best parameter sets found so far. Take the single best set as a new starting point for future optimisations.

4. Use the `nlm` function to perform $N-m$ additional optimisations, each beginning from a different initial value of θ, generated via random perturbations about the starting point identified in step 3. These optimisations are constrained within the boundaries identified in step 2.

5. Discard any parameter sets for which the objective function exceeds $10S_m$, where S_m is the mth smallest of the objective function values found in step 2.

6. Repeat steps 2 to 5 a fixed number of times.

Step 4 of the algorithm yields N sets of parameters, each with associated objective functions. Some of these are discarded in step 5, so upon completion the routine yields between m and N parameter sets in total. One of these is the overall optimum, and the others usually have very similar objective function values. In the work reported below we took $M = 100$, $m = 20$ and $N = 50$, and repeated steps 2–5 five times.

2.3.3 Identifiability issues

When fitting Poisson cluster models using a method of moments, it is possible for very different parameter sets to yield objective function values that are close to the overall optimum. This has been widely reported in the literature (Velghe et al., 1994; Verhoest et al., 1997; Cameron et al., 2000), as well as in our own experience. When this occurs, it indicates lack of identifiability of the model parameters. Clearly, in general this is undesirable and, if possible, one should choose models whose parameters *are* identifiable on the basis of available data.

To investigate the identifiability of parameters, we use profile objective functions, defined for each parameter as already described in Section 2.3.1. The idea of holding a parameter fixed and optimising over the remaining parameters is similar in spirit to the Generalized Linear Uncertainty Estimation (GLUE) framework (Beven and Binley, 1992; Freer et al., 1996). A plot of the profile objective function can be used, together with thresholds calculated from (2.2), to determine confidence intervals for each parameter.

In statistical terms, lack of identifiability equates to imprecise estimation, which should be reflected in large standard errors for one or more parameters i.e. large diagonal elements in the matrix $\mathbf{V} = \mathbf{H}^{-1}\mathbf{J}\mathbf{H}^{-1}$ where $\mathbf{H}$ is the Hessian of the objective function (see Section 2.3.1 above). Intuitively, it is clear that when the objective function is flat (so that different parameter sets give similar values), some elements of $\mathbf{H}$ will be small and $\mathbf{V}$ will contain large values, as required. Estimated standard errors are provided by the R routines used in this work; however, these do not provide a complete picture of parameter identifiability unless the objective function is reasonably symmetric. We therefore do not consider standard errors further here.

2.4 Assessment of non-extreme performance

In this section we apply the theory outlined in Section 2.3 to the models listed in Table 2.6 (page 17). The aim is to compare the models and to examine the effect of incorporating different fitting statistics into the objective function (2.1).

2.4.1 Data

Three raingauges are used in this comparison. All three raingauges are tipping-bucket gauges with 0.2mm resolution. Two of these have been selected for their location and long records. These are Elmdon (Birmingham) and Heathrow which are typical of Midlands and South-East England rainfall. The final gauge is at Tan Llwyth in

Table 2.7 Data sets used in the comparison

Gauge	Data available for	OS grid reference (East,North)
Heathrow	1/1/1949 – 31/12/1987, 1/9/1988 – 31/12/1992, 1/3/1993 – 31/12/1996, 1/3/1997 – 22/7/2001	(5077, 1767)
Elmdon	1/6/1949 – 30/11/1983, 1/1/1984 – 31/7/1989, 1/9/1989 – 31/12/1996, 1/3/1997 – 30/9/1997	(4176, 2839)
Tan Llwyth	1/5/1975 – 26/7/1997, 14/8/1997 – 16/11/1998, 17/12/1998 – 26/4/1999	(2843, 2876)

Plynlimon, chosen because it represents a very wet climate with an annual rainfall nearly four times as large as that at Heathrow.

These three gauges together are therefore taken to be representative of the variability of rainfall over England and Wales. Consequently, the model comparison carried out on the basis of performance at these three gauges enables the identification of an optimal model for national application.

Table 2.7 shows the data availability for these three sites.

2.4.2 Comparison methodology

As well as comparing the models in Table 2.6, we have explored the effect of using different objective functions of the form (2.1). As described in Section 2.3.1 above, it is common to fit models using some combination of means, variances, autocorrelations and proportions of dry intervals, at time scales from 1 to 24 hours. We have therefore compared results from three different combinations, which are summarised in Table 2.8. Objective function (OF) 1 contains 8 properties in total (note the zero weights attached to the 12-hourly variance, and 6- and 12-hourly ACF and proportion dry) — this may be thought of as a 'basic' set of fitting properties. OF2 and OF3 are 'extended' sets containing 13 properties each. In OF1 and OF2 the weights have been chosen, on the basis of the properties' typical relative magnitudes for UK rainfall, so that each property is expected to contribute roughly equally to the fit (for example, as a rule of thumb the variance of hourly rainfall tends to be around double the mean, and both properties enter quadratically into the objective function, so a weight of 0.25 is attached to the hourly variance). In OF3 the weights for the hourly mean, variance and proportion of dry intervals have been multiplied by 100 so as to force adherence to these properties. This enables us to determine whether it is better to allow all properties to contribute roughly equally to the fit, or whether selected properties should be allowed to dominate. We have not considered the possibility of weighting the fitting properties on the basis of their observed values. This is because data-dependent weights will generally lead to biased parameter estimates, as described in Section 2.3.1.

In assessing the different models and objective functions, the following criteria have been considered:

Table 2.8 Fitting properties and weights for the three objective functions used in the comparison exercise. The weights are the $\{w_i\}$ in equation (2.1).

Fitting Property	Time Scale	Weight		
		OF1	OF2	OF3
Mean	1 hour	1	1	100
Variance of rainfall amount	1 hour	0.25	0.25	25
	6 hours	0.0009	0.0009	0.0009
	12 hours	0	0.0001	0.0001
	24 hours	0.0001	0.0001	0.0001
Lag 1 autocorrelation	1 hour	0.04	0.04	0.04
	6 hours	0	0.1	0.1
	12 hours	0	0.2	0.2
	24 hours	0.25	0.25	0.25
Proportion of dry intervals	1 hour	0.01	0.01	1
	6 hours	0	0.02	0.02
	12 hours	0	0.03	0.03
	24 hours	0.04	0.04	0.04

Reproduction of fitting properties: the ability of a model to reproduce the fitting properties can be summarised using the achieved objective function values. The theory from Section 2.3.1 can be used to judge whether two models differ significantly in this respect.

Identifiability of parameters: as indicated previously, in practical applications it is desirable that model parameters should be well identified on the basis of the data available. For any particular model and objective function, parameter identifiability can be explored by plotting profile objective functions showing the range of parameter values that are broadly consistent with the fitting properties, taking into account the uncertainty in estimation of these fitting properties themselves. This technique can be used to guide the choice of both model and objective function.

Physical realism of fitted models: one of the main justifications for the use of the Poisson cluster models is that they represent, albeit in a simplified stochastic form, the mechanisms of the rainfall process. However, this argument can be convincing only if the model parameters are physically reasonable. For example, the physical entities represented in the models as 'rain cells' typically last from a few minutes to around half an hour (Waymire and Gupta, 1981); mean cell durations markedly outside this range should be regarded as unrealistic.

Sensitivity to the objective function: although it is anticipated that the inclusion of different fitting properties and weights in the objective function will affect the fitted models to some extent, extreme sensitivity to fitting properties is undesirable in a model that is intended for routine application nationally.

Computational cost: since the aim of this work is to provide recommendations for models that can be used routinely in hydrological practice, it is important that the models should not be too expensive to fit.

Table 2.9 Optimal values of the three objective functions, for each model fitted to January data from Heathrow. See Table 2.6 for details of the models, and Table 2.8 for details of the objective functions.

Model	No. parameters	OF1	OF2	OF3
1	4	0.0026	0.0036	0.0043
2	6	1.0×10^{-5}	3.7×10^{-4}	4.5×10^{-4}
3	6	3.1×10^{-5}	4.7×10^{-4}	6.7×10^{-4}
4	6	4.1×10^{-6}	2.4×10^{-4}	2.5×10^{-4}
5	7	3.2×10^{-7}	2.4×10^{-4}	2.5×10^{-4}
6	10	5.4×10^{-8}	1.9×10^{-4}	2.0×10^{-4}

2.4.3 Results

All of the models in Table 2.6 have been fitted, using each of the objective functions defined in Table 2.8, to data from each of the three sites in Table 2.7 using the algorithm described in Section 2.3.2. Models have been fitted for each month of the year at Heathrow, but for January and July only at the other two sites. All optimisation was carried out with respect to the logarithms of the model parameters; this has been found to be both faster and more stable than optimising with respect to the original parameters.

Comparison of models — reproduction of fitting properties

Table 2.9 shows, for each objective function, the minimum values attained by each model for January data from Heathrow. Model 6 (the Bartlett-Lewis model with two cell types) minimises all three objective functions. As a check on the minimisation, it is worth comparing objective function values for pairs of nested models — in such cases, the more complex model should always yield a smaller value. For example, models 5 and 6 are both extended versions of model 2 (the basic Bartlett-Lewis model with a 2-parameter depth distribution); accordingly, they yield smaller objective functions in all cases. Model 5 is also an extended version of model 4, and the results are as expected. Model 6 is not an extension of either model 4 or model 5 so, although in this particular case it yields a lower value for all three objective functions, this is not always guaranteed.

To compare the models in terms of objective function performance, it is useful to calculate the objective function thresholds from (2.3), for the best-fitting model (i.e. model 6 in Table 2.9). In general, the thresholds cannot be used to attach a statistical significance level to the comparison between two models; however, any model with a minimised objective function value above the corresponding threshold is clearly uncompetitive and can be rejected. This provides a useful means of screening out any obvious non-contenders among the models being considered.

Unfortunately, due to numerical instabilities it has not been possible to calculate the thresholds for model 6. The problem is due to ill-conditioning of either $\mathbf{H}$ or $\mathbf{V}$ or both (see Section 2.3.1), and indicates that the model may be overparameterised. However, thresholds can be computed for the next best model (model 5): 95% thresholds are 0.0015, 0.0021 and 0.016 respectively. These thresholds are very

Table 2.10 Optimal values of the three objective functions, for each model fitted to July data from Heathrow. See Table 2.6 for details of the models, and Table 2.8 for details of the objective functions.

Model	No. parameters	OF1	OF2	OF3
1	4	0.0028	0.0059	0.0063
2	6	8.0×10^{-5}	2.2×10^{-4}	4.2×10^{-4}
3	6	1.1×10^{-4}	3.3×10^{-4}	5.9×10^{-4}
4	6	2.7×10^{-5}	1.1×10^{-4}	2.6×10^{-4}
5	7	2.7×10^{-5}	1.1×10^{-4}	1.1×10^{-4}
6	10	4.7×10^{-5}	9.8×10^{-5}	1.0×10^{-4}

similar to those obtained from each of models 2–4, indicating that they are not too dependent on the model used to construct them. On this basis, the Poisson model (model 1) is significantly worse than model 5 according to both OF1 and OF2, and can therefore be rejected at this stage. The fact that it cannot be rejected immediately using OF3 may indicate that this objective function (which is dominated by three fitting properties) does not contain enough information to discriminate between models.

Although the thresholds cannot be used to attach a statistical significance level to comparisons between models, they do at least provide an indication of the range of values that can be regarded as 'almost optimum' after accounting for uncertainty in the estimation of the fitting statistics. In particular, the thresholds computed from model 5 are one or two orders of magnitude larger than the objective function values for the best-fitting models. In this context, the differences between the various models (with the exception of the Poisson) are negligible, and it therefore seems that there is little justification for using models with more than 6 parameters. On this basis, there is a strong argument for using the best-performing of the 6-parameter models, which is model 4 in all cases (the random-parameter Bartlett-Lewis with exponential cell depth distribution). Note also that the original Bartlett-Lewis model (model 2) achieves smaller objective function values in all cases than the original Neyman-Scott (model 3) — this provides at least a tentative indication that the Bartlett-Lewis mechanism may be preferable, although in practical terms the difference may be negligible.

Table 2.10 shows the corresponding results for July data from Heathrow. The story is almost identical to that in January, although here model 5 outperforms model 6 according to OF1. This time, 95% thresholds could be computed for model 6 using OF2 and OF3: accordingly, thresholds for the best-performing models are 0.011, 0.012 and 0.3 for the three objective functions. In this case the Poisson model is not immediately ruled out using any of the objective functions. This reflects the increased variability in fitting statistics in July compared to January, which makes it more difficult to discriminate reliably between models. Nonetheless, the Poisson values differ from the remaining models by orders of magnitude, and once again the differences among the remaining models are negligible in comparison with the uncertainty so that model 4 emerges as the strongest contender. The original Bartlett-Lewis model 2 still outperforms the original Neyman-Scott model 3.

The objective function values give an overall measure of model fit, but do not indicate how well the models reproduce individual properties. To explore the fits in more detail, Tables 2.11 and 2.12 show all of the observed fitting properties for Heathrow, in January and July respectively, along with those obtained from the various models fitted

Table 2.11 Observed rainfall statistics for January at Heathrow, with corresponding properties for models fitted using objective function 1. Bold entries indicate properties used in the fitting.

Property	Time Scale	Observed	Fitted using model					
			1	2	3	4	5	6
Mean	1 hour	**0.070**	**0.070**	**0.070**	**0.071**	**0.069**	**0.070**	**0.070**
Variance	1 hour	**0.098**	**0.110**	**0.098**	**0.105**	**0.099**	**0.099**	**0.098**
	6 hours	**1.536**	**1.752**	**1.511**	**1.531**	**1.515**	**1.519**	**1.530**
	12 hours	3.971	4.002	3.824	3.808	3.873	3.887	3.879
	24 hours	**9.295**	**8.506**	**9.203**	**9.066**	**9.330**	**9.310**	**9.302**
Lag 1 ACF	1 hour	**0.533**	**0.629**	**0.536**	**0.539**	**0.533**	**0.533**	**0.533**
	6 hours	0.293	0.141	0.265	0.241	0.280	0.278	0.268
	12 hours	0.156	0.063	0.204	0.193	0.203	0.199	0.199
	24 hours	**0.122**	**0.030**	**0.122**	**0.122**	**0.121**	**0.122**	**0.122**
Proportion dry	1 hour	**0.892**	**0.928**	**0.872**	**0.859**	**0.875**	**0.893**	**0.891**
	6 hours	0.738	0.792	0.724	0.686	0.738	0.751	0.721
	12 hours	0.614	0.655	0.618	0.584	0.625	0.632	0.610
	24 hours	**0.450**	**0.448**	**0.454**	**0.456**	**0.450**	**0.450**	**0.450**

Table 2.12 Observed rainfall statistics for July at Heathrow, with corresponding properties for models fitted using objective function 1. Bold entries indicate properties used in the fitting.

Property	Time Scale	Observed	Fitted using model					
			1	2	3	4	5	6
Mean	1 hour	**0.065**	**0.065**	**0.067**	**0.068**	**0.065**	**0.065**	**0.065**
Variance	1 hour	**0.304**	**0.323**	**0.311**	**0.313**	**0.307**	**0.306**	**0.306**
	6 hours	**3.906**	**4.207**	**3.705**	**3.723**	**3.761**	**3.780**	**3.810**
	12 hours	9.401	9.186	8.842	8.818	9.005	9.020	9.013
	24 hours	**20.520**	**19.105**	**20.553**	**20.430**	**20.657**	**20.656**	**20.623**
Lag 1 ACF	1 hour	**0.385**	**0.517**	**0.403**	**0.409**	**0.391**	**0.389**	**0.388**
	6 hours	0.223	0.090	0.195	0.186	0.197	0.196	0.814
	12 hours	0.140	0.041	0.161	0.157	0.147	0.142	0.143
	24 hours	**0.103**	**0.020**	**0.101**	**0.103**	**0.102**	**0.101**	**0.102**
Proportion dry	1 hour	**0.939**	**0.964**	**0.906**	**0.893**	**0.952**	**0.968**	**0.920**
	6 hours	0.834	0.875	0.819	0.780	0.857	0.876	0.813
	12 hours	0.750	0.779	0.749	0.717	0.767	0.780	0.735
	24 hours	**0.619**	**0.619**	**0.628**	**0.632**	**0.618**	**0.618**	**0.619**

using OF1. Note that not all of the properties contribute to this objective function; the properties not used are shown in normal typeface, and give some impression of the lack of fit that could be overcome by using OF2 or OF3 in preference to OF1. As far as model intercomparison is concerned, the most obvious feature of the tables is the lack of the fit of the Poisson model in relation to the others; apart from this, there is a tendency for the other models to underestimate the 6-hour lag 1 autocorrelation and overestimate that at 12 hours. The Neyman-Scott model (model 3) underestimates the proportion of dry intervals at subdaily timescales. Otherwise the agreement between the fitted models and the data is excellent in both months, and there is very little in practical terms to distinguish between them. This reinforces the previous argument in support of the best fitting 6-parameter model (model 4), and shows that this model is capable of reproducing features of both winter and summer rainfall.

Similar conclusions hold for all other months at Heathrow, and for the other two sites. Except at Plynlimon, model 4 is invariably the best-performing of the 6-parameter models. Moreover, the basic Bartlett-Lewis model invariably outperforms the basic Neyman-Scott. The exceptions are as follows:

- **Heathrow:** model 2 (non-randomised Bartlett-Lewis with 2-parameter cell depth distribution) slightly outperforms model 4 during February and June for OF1, June and August for OF2, and February, March and August for OF3. The Neyman-Scott model 3 outperforms the corresponding Bartlett-Lewis model 2 only during May, under OF3.

- **Elmdon:** the Neyman-Scott model 3 slightly outperforms the random parameter Bartlett-Lewis model 4 during July for OF3. It also outperforms the corresponding Bartlett-Lewis model 2 in July, under both OF1 and OF3.

- **Plynlimon:** the non-randomised Bartlett-Lewis model 2 is the best-performing of the 6-parameter models in all cases.

Although these exceptions are noted, the differences between the 6-parameter models are always tiny in comparison with the thresholds calculated using (2.3), and can therefore be regarded as negligible in practical terms. Therefore these models should be regarded as essentially equivalent at this stage, although if a single model had to be chosen it would be the random-parameter Bartlett-Lewis model 4, since it usually outperforms its competitors slightly. It is also worth noting that the Neyman-Scott model takes much longer to fit than any of the Bartlett-Lewis models, because it requires numerical integration when calculating the proportion of dry intervals. In view of this, for the remainder of the analyses we will focus on the Bartlett-Lewis clustering mechanism.

Comparison of models and objective functions — parameter identifiability

As usual, the parameters in some of these models are poorly identified. This has been explored using profile objective functions, as described in Section 2.3.3 above. For all parameters, profiles have been computed on a log scale, over a range of values that extends well beyond the limits that may be considered physically realistic.

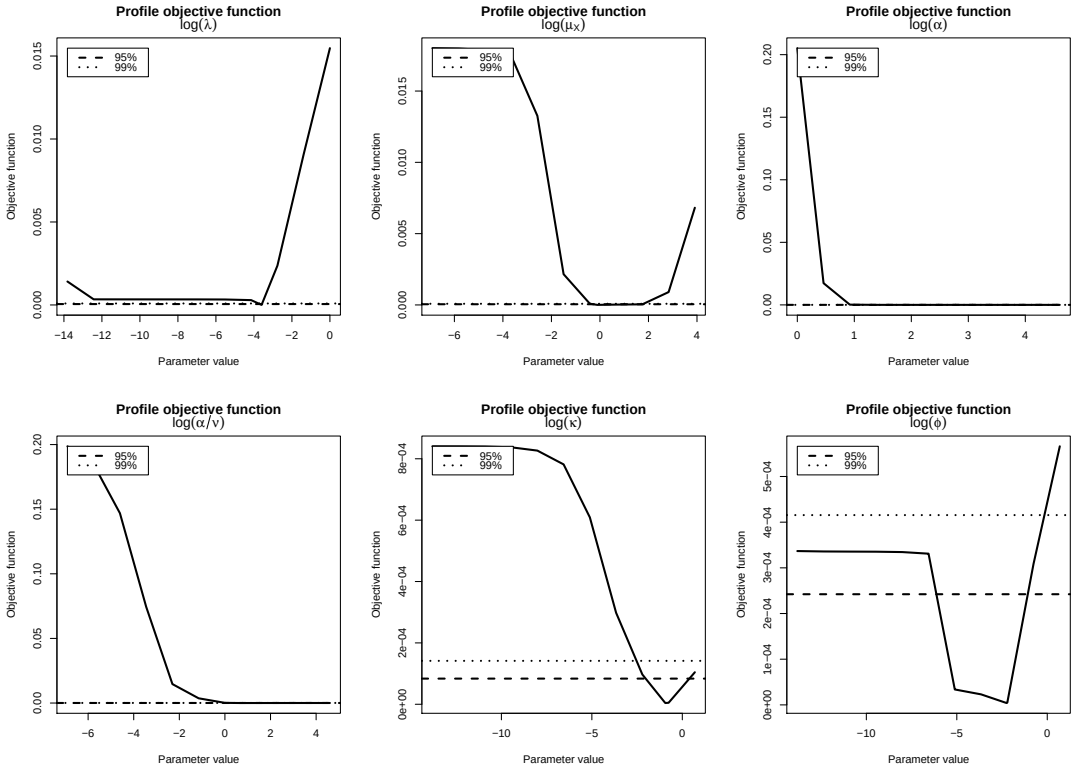


Figure 2.1 Profile objective functions for model 4 fitted to January data from Heathrow, using objective function 1. Horizontal lines define approximate confidence intervals at the levels indicated.

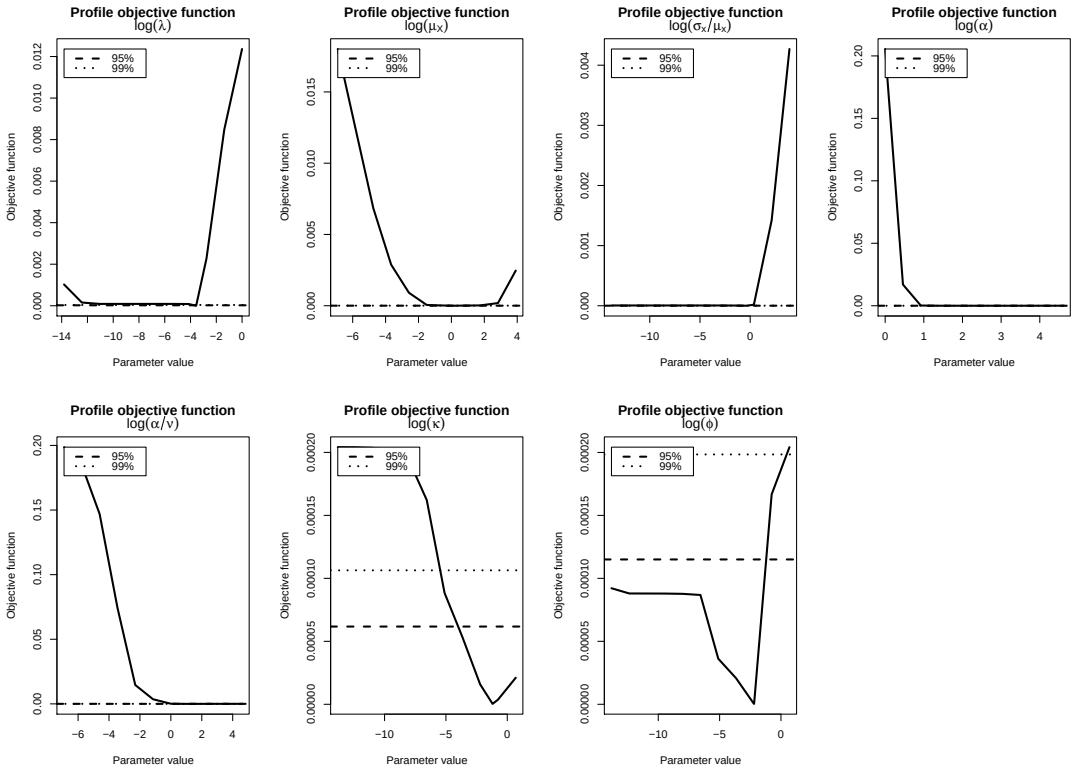


Figure 2.2 Profile objective functions for model 5 fitted to January data from Heathrow, using objective function 1.

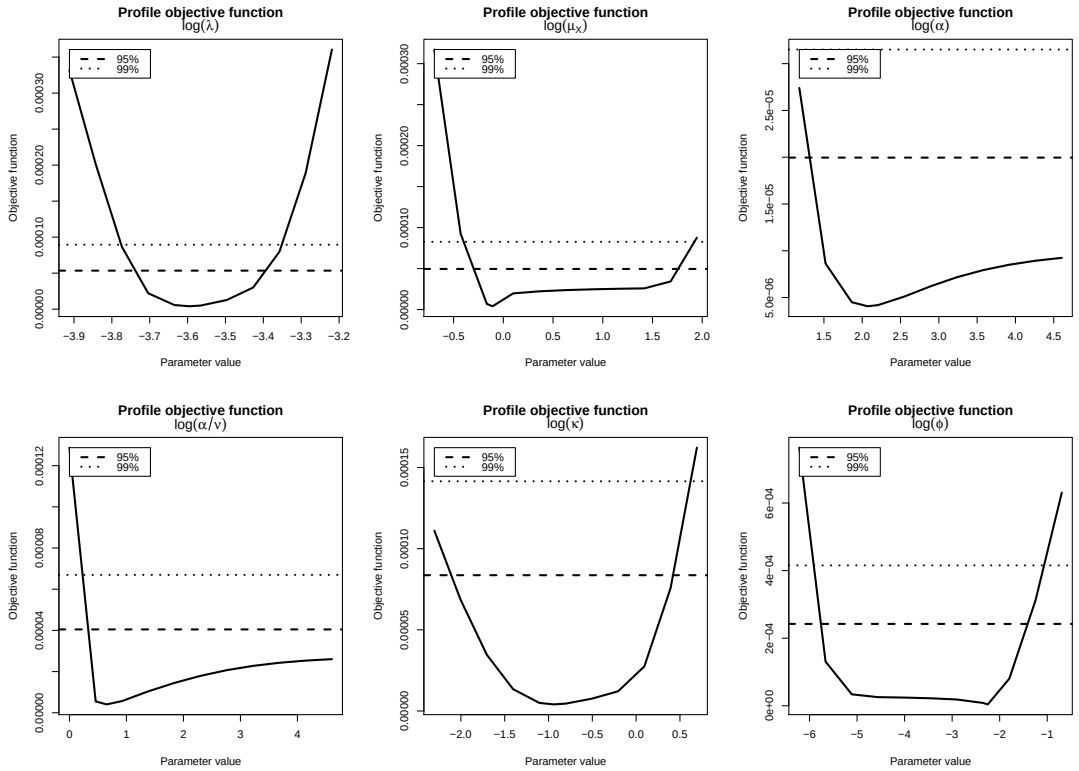


Figure 2.3 Profile objective functions for model 4 fitted to January data from Heathrow, using objective function 1. This is exactly the same as Figure 2.1, but with reduced parameter ranges.

Figures 2.1 and 2.2 show profile objective function plots for models 4 and 5 (random-parameter Bartlett-Lewis with exponential and 2-parameter cell depth distributions respectively), for the January data at Heathrow fitted using objective function 1. On the basis of the 95% objective function thresholds, bounds can be established on λ and μ_X in both cases. Under model 4, bounds can also be established for κ and ϕ, and lower bounds for α and α/ν (to see these more clearly, Figure 2.3 shows the profiles for this model in the neighbourhood of the objective function minimum). Under model 5, bounds can be established for σ_X/μ_X, α and α/ν (it is not clear from the plots, but the profiles do cross the thresholds for these parameters); a lower bound is found for κ and and upper bound for ϕ. Interestingly, the interval for σ_X/μ_X in model 5 runs from 0.5714 to 0.7126 — this does not include the value of 1 corresponding to the exponential model. This suggests that the 2-parameter cell depth distribution yields a significant improvement over the exponential.

Using objective function 2 (see Figures 2.4 and 2.5), the results for model 4 are unchanged. Surprisingly however, the parameters for model 5 become less well identified. For example, no bounds can be established for κ here, and only an upper bound can be established for σ_X/μ_X. This upper bound is 1.8093, which exceeds 1 so that according to this objective function the exponential model is not rejected.

The differences between the results for the two objective functions indicate numerical difficulties with the calculations, since one would expect model parameters to become better identified when more properties are used in the fitting. This serves as a warning against over-interpretation of the plots and approximate confidence intervals.

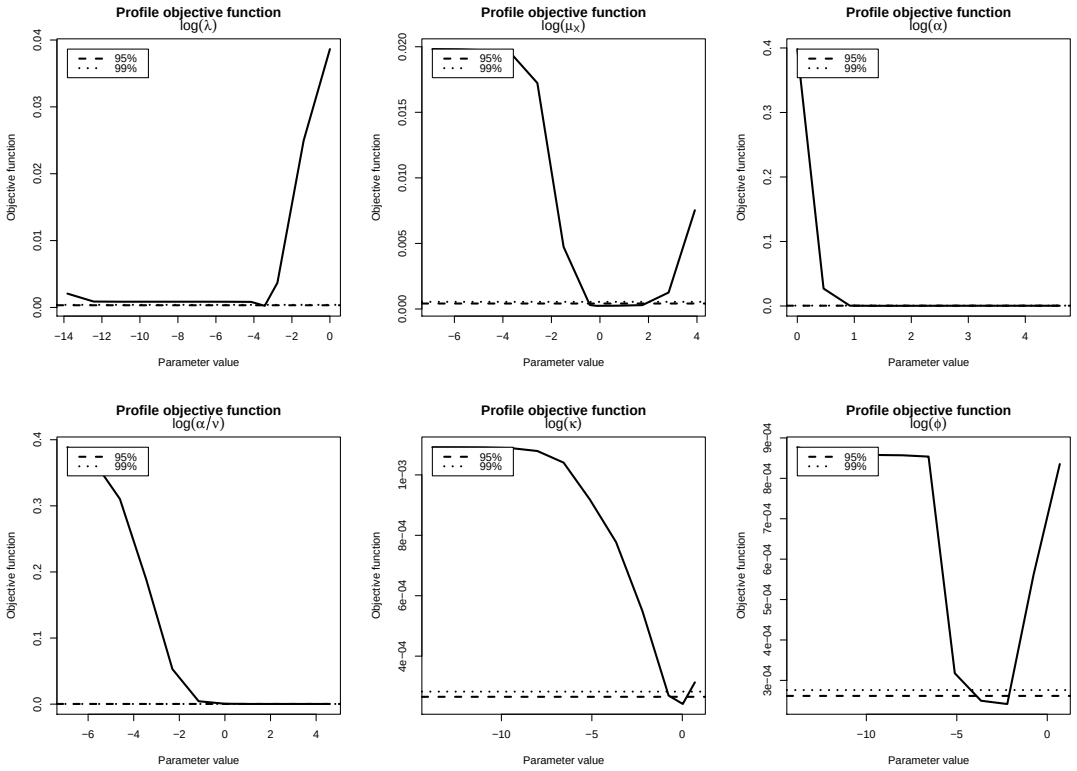


Figure 2.4 Profile objective functions for model 4 fitted to January data from Heathrow, using objective function 2.

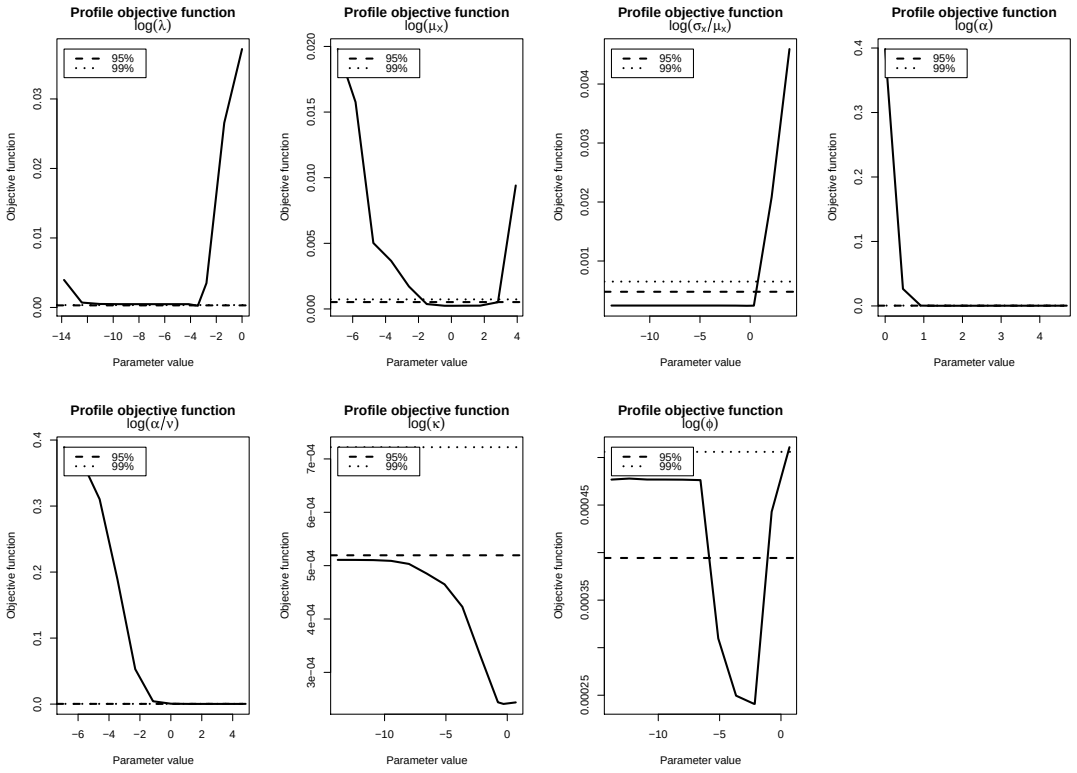


Figure 2.5 Profile objective functions for model 5 fitted to January data from Heathrow, using objective function 2.

Nonetheless, the shapes of these profiles are remarkably consistent across months and sites. In Appendix C for example, Figures C.21–C.24 show the July results corresponding to Figures 2.1–2.5, Figures C.25 and C.26 show the January results from Elmdon (OF1), and Figures C.27 and C.28 show the corresponding results for Plynlimon. In broad terms therefore, the profiles provide an extremely useful means of assessing the relative ease of identifying the various parameters.

There is a tendency for parameters to be less well identified in summer than in winter. For example, at Heathrow in July (Figures C.21–C.24), the only parameters for which bounds can be established in all cases are λ and μ_X, and ϕ is effectively unidentifiable except using OF2 for model 4 (in Figure C.24, the spike in the ϕ profile is probably due to a failure of the numerical minimisation routine at this point). Comparing Figures C.21 and C.23, it appears that here there is some benefit from the inclusion of the additional fitting properties in OF2. Comparing Figures C.23 and C.24, it seems that in July the exponential model is much better identified than the 2-parameter model.

The January results from Elmdon (Figures C.25 and C.26) are almost identical to the corresponding results from Heathrow (Figures 2.1 and 2.2). In this particular case, upper and lower bounds can be established for all of the parameters in the exponential model, and for all parameters except α, α/ν and κ in the 2-parameter model. For these three parameters, lower bounds only can be established.

For Plynlimon, parameters tend to be slightly less well identified than at the other two sites. For example, Figures C.27 and C.28 show that lower bounds cannot be established for λ at this site in January, using OF1. The same is true using OF2, although a lower bound can be established using OF3. The worse performance at this site could be either because its record length (25 years) is shorter than that of the other sites, or because its rainfall regime differs substantially from that at the other sites. This is investigated in more detail below.

Finally, Figures 2.6 and 2.7 show the profiles for model 6 (the 2-cell model) at Heathrow, for both January and July, fitted using OF2. The erratic nature of these plots is an indication of numerical difficulties in fitting the model, which are much worse than for the other models. The January results for this model are some of the best that have been obtained; the July plots are much more typical, with at least two of the parameters (μ_{X_1} and $\mu_{X_1} - \mu_{X_2}$) being completely unidentifiable, and some indication that the profiles for several other parameters (e.g. σ_{X_2}) only cross the objective function thresholds because of a failure in the minimisation algorithm. It is not uncommon, even using the 13 properties in OF2, for 3 or 4 profile parameters to be completely unidentifiable (in some cases, the profiles are almost completely flat). Coupled with the obvious numerical difficulties in the minimisation here, this suggests that such a complex model is not suitable for routine use.

Effect of objective function on parameter estimates

To explore the sensitivity of the estimated parameters to the properties used in the objective function, we now present estimates for model 4 at Heathrow, in January and July, using all three objective functions. The initial model comparisons in Tables 2.9 and 2.10 suggested that, overall, this is the most promising of the 6-parameter models in terms of objective function performance; and the profile plots in the previous section

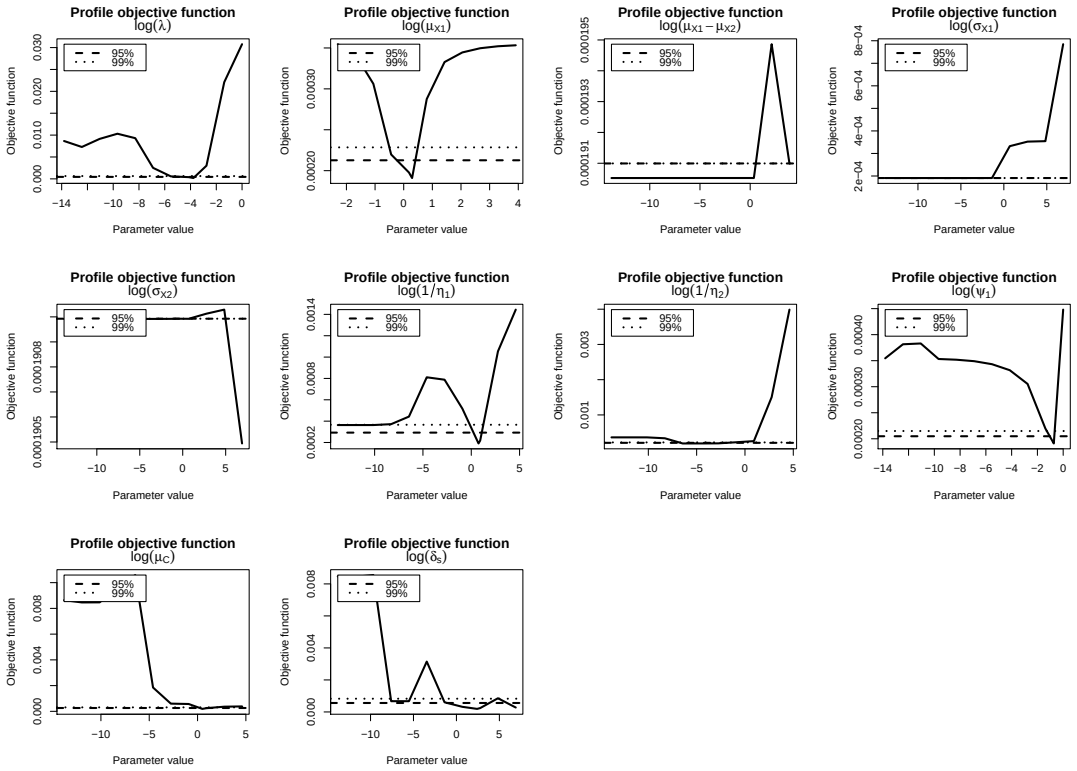


Figure 2.6 Profile objective functions for model 6 fitted to January data from Heathrow, using objective function 2.

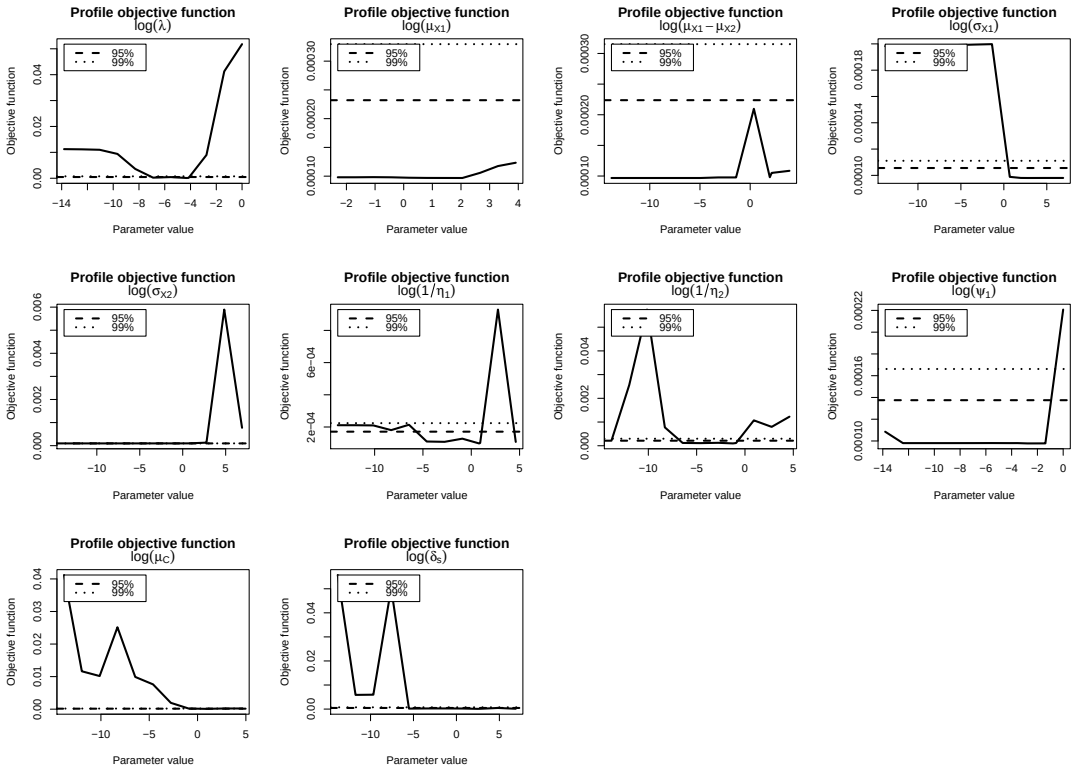


Figure 2.7 Profile objective functions for model 6 fitted to July data from Heathrow, using objective function 2.

Table 2.13 Random-parameter Bartlett-Lewis model with exponential cell depth distribution (model 4 in Table 2.6): parameterisation used for model fitting.

Parameter	Description
λ	Arrival rate of storms
μ_X	Mean cell depth
α	Shape parameter of randomisation distribution
α/ν	Mean of randomisation distribution (the expected cell duration is $\nu/(\alpha-1)$)
κ	Ratio of expected cell duration to expected cell inter-arrival time
ϕ	Ratio of expected cell duration to expected storm duration

Table 2.14 Parameter estimates for model 4 at Heathrow in January, together with approximate 95% confidence intervals derived using profile objective functions.

Parameter	OF1 estimate	OF2 estimate	OF3 estimate
λ	0.0274 (0.0238,0.0336)	0.0323 (0.0291,0.0329)	0.0321 (0.0296,0.0333)
μ_X	0.8978 (0.7431,5.7977)	0.8555 (0.6419,6.5975)	0.8430 (0.6590,1.6547)
α	7.9258 (3.7233,∞)	3.7595 (2.7801,∞)	4.2328 (3.5750,25.7358)
α/ν	1.9225 (1.3907,∞)	5.9016 (3.0796,∞)	4.6448 (2.5834,12.0272)
κ	0.3903 (0.1213,1.5231)	1.0271 (0.5268,1.2808)	0.9092 (0.7501,1.1266)
ϕ	0.1054 (0.0032,0.2436)	0.1069 (0.0199,0.1203)	0.1118 (0.0552,0.1399)

suggest that numerical difficulties can lead to unreliable calculations for more complex models.

The parameterisation used at this stage in the model fitting is slightly different from that given in Table 2.4; this is mainly to ensure stability in the numerical optimisation of the objective function (2.1). The parameters used for fitting are summarised in Table 2.13. The parameter estimates under each of the objective functions, along with 95% confidence intervals based on objective function thresholds, are given in Tables 2.14 and 2.15 for January and July respectively. In these tables, a value of ∞ is reported for the upper end of a confidence interval, if the objective function remained below the 95% threshold at the upper end of the evaluation range. A value of 0 is reported if it remained below the threshold at the lower end of the range (this is distinct from 0.0000, which represents a non-zero value recorded to 4 decimal places). The intervals derived from OF1 and OF2 correspond exactly to the profiles in Figures 2.3, 2.4, C.21 and C.23.

Overall, the parameter estimates are all physically plausible and the three objective functions yield similar values, when set against the context of the uncertainty in all of the estimates. The exceptions to this are α in Table 2.14 (for which the OF3 estimate falls outside the OF1 confidence interval) and κ in the same table (for which the OF1 estimate differs from the other two).

The effect of increasing the number of fitting properties can be seen by comparing the confidence intervals from OF1 with those from OF2. In general, the intervals are comparable, although in both January and July, the intervals for κ and ϕ are clearly narrower under OF2 than under OF1. This suggests that the inclusion of more fitting properties can aid the identification of these parameters. For practical applications

Table 2.15 Parameter estimates for model 4 at Heathrow in July, together with approximate 95% confidence intervals derived using profile objective functions.

Parameter	OF1 estimate	OF2 estimate	OF3 estimate
λ	0.0180 (0.0053,0.0185)	0.0193 (0.0168,0.0194)	0.0193 (0.0173,0.0195)
μ_X	4.4174 (4.1319,19.4399)	4.5977 (4.0207,18.7207)	4.4112 (3.5052,27.7626)
α	3.3580 (2.5090,7.7048)	3.0825 (2.5121,4.2861)	3.1338 (2.8501,4.1746)
α/ν	5.0561 (2.6236,∞)	7.5205 (3.2316,∞)	5.5653 (4.3074,9.5995)
κ	0.1440 (0.0003,∞)	0.2492 (0.0230,∞)	0.1509 (0.0397,0.5669)
ϕ	0.0755 (0,∞)	0.0888 (0.0092,0.7242)	0.0786 (0.0064,0.3614)

however, it may be necessary to weigh the benefits of using extra fitting properties against the cost of obtaining them. For example, in developing a model for national application it may be necessary to regionalise all of the fitting statistics, in which case the costs of regionalising all of the properties used in OF2 would probably outweigh the benefits.

The effect of reweighting the fitting properties can be seen in moving from OF2 to OF3. Here, although the estimates themselves are affected very little by the reweighting, there does seem to be a consistent tendency for the OF3 confidence intervals to be narrower than for OF2. This is most marked for the parameters α and α/ν in January, and α/ν and κ in July, for which the use of OF3 enables upper bounds to be established. This presumably reflects the fact that OF3 allocates most weight to fitting properties that can be estimated reasonably accurately (the mean, 1-hour variance and 1-hour proportion of dry intervals). However, this appears to contradict the finding in Section 2.4.3 where it was not possible to rule out the Poisson model 1 using OF3. Further work would be required to resolve the question of how best to weight the fitting properties.

Effect of record length

In Section 2.4.3, it was noted that model parameters (in particular, λ) are less well identified at Plynlimon than at the other two sites. It is of interest to determine whether this is due to a short record length (there are only 25 years of data at Plynlimon, but around 50 at the other sites) or to a different rainfall regime. To investigate this, models have been fitted to two 25-year subsets of data from Heathrow, covering the periods 1949–1973 and 1977–2001 respectively. Figures C.29 and C.210 show the profile objective functions for model 4, fitted using OF1 to January data from each of these subsets. The corresponding figures for July are given in Figures C.211 and C.212.

Perhaps surprisingly, the 25-year Heathrow results are very similar to those obtained from the entire record (Figures 2.1 and C.21). Tables 2.16 and 2.17 show the parameter estimates and 95% confidence intervals for the three Heathrow datasets, along with Plynlimon, for January and July. It appears from this that the parameters for the two 25-year periods are identified almost as well as those from the entire record. It is also clear that the "difficult" parameters are not the same at the two sites. At Plynlimon for example, there are problems with λ but μ_X is quite well identified. At Heathrow, λ is well identified but the confidence intervals for μ_X are wide. This suggests that the rainfall regime may have an impact upon the identifiability of

Table 2.16 Parameter estimates and approximate 95% confidence intervals for model 4, fitted to January data from Plynlimon, Heathrow and two 25-year sequences from Heathrow.

Parameter	Data set			
	Heathrow			Plynlimon
	1949–2001	1949–1973	1977–2001	
λ	0.03 (0.03,0.03)	0.03 (0.03,0.03)	0.03 (0.01,0.03)	0.02 (0,0.06)
μ_X	0.90 (0.75,5.79)	0.77 (0.63,6.84)	0.98 (0.84,1.06)	1.62 (1.50,3.02)
α	7.93 (3.91,∞)	7.13 (4.78,∞)	8.21 (4.72,∞)	99.62 (98.89, ∞)
α/ν	1.92 (1.59,∞)	3.64 (3.02,∞)	1.59 (1.14,9.09)	0.56 (0.53, 0.88)
κ	0.39 (0.13,1.48)	0.95 (0.11,1.30)	0.21 (0.03,∞)	0.27 (0.05,∞)
ϕ	0.11 (0.00,0.34)	0.11 (0.01,0.12)	0.07 (0,∞)	0.07 (0,0.25)

Table 2.17 Parameter estimates and approximate 95% confidence intervals for model 4, fitted to July data from Plynlimon, Heathrow and two 25-year sequences from Heathrow.

Parameter	Data set			
	Heathrow			Plynlimon
	1949–2001	1949–1963	1987–2001	
λ	0.02 (0.01,0.02)	0.02 (0.02,0.02)	0.02 (0.01,0.02)	0.02 (0,0.02)
μ_X	4.42 (4.13,19.44)	4.23 (1.87,∞)	3.73 (3.00,18.79)	1.69 (1.47,2.89)
α	3.36 (2.51,7.70)	3.54 (3.11,9.47)	3.51 (2.94,6.62)	100.0 (100.0, ∞)
α/ν	5.06 (2.62,∞)	4.61 (3.01,∞)	5.75 (3.02,∞)	0.97 (0.92,2.04)
κ	0.14 (0.00,∞)	0.42 (0.09,∞)	0.10 (0.01,0.76)	0.17 (0.08,0.25)
ϕ	0.08 (0,∞)	0.19 (0.03,∞)	0.04 (0,∞)	0.04 (0,0.06)

individual parameters. It is interesting to note that at Plynlimon, the estimate of α effectively reaches its upper bound (which was set to 100) for both months considered. When α is large, the random-parameter model 4 is effectively the same as the original Bartlett-Lewis model 2 (but with an exponential cell depth distribution). This reinforces the results of Section 2.4.3, where it was found that model 2 outperformed model 4 at Plynlimon.

It is a little surprising that the parameters at Heathrow are identified almost as well with 25 years' data as with 50. To explore this further, models have also been fitted to 15-year sequences of data from Heathrow. In this case, the shapes of the profiles did not change much, but the parameters were much less well identified. This is because the objective function thresholds were higher, reflecting the increased uncertainty due to a smaller sample. The potential effect of this can be seen by looking at the profile for λ in any of the figures in this chapter. It seems almost universally true that the profile for this parameter flattens out close to the optimum; therefore, the parameter is only well identified when the objective function threshold is below the flat region. Our results suggest that there may be a critical record length of around 20 years, below which it becomes extremely difficult to identify some parameters.

2.4.4 Summary of section

On the basis of the results presented here, the following conclusions can be drawn regarding the reproduction of non-extreme rainfall properties:

- There is little to choose between the Bartlett-Lewis and Neyman-Scott clustering mechanisms in terms of performance, although there are some computational advantages to the Bartlett-Lewis if the proportion of dry intervals is to be used for model fitting.

- Little is gained from the use of models with more than 6 parameters; indeed, if models become much more complex than this it becomes very difficult to identify some parameters on the basis of the fitting properties considered here.

- Overall, the best-performing of the 6-parameter models considered here is the random-parameter Bartlett-Lewis model of Section 2.2.2, with an exponential cell depth distribution. However, in practical terms the improvement over other models may be slight as far as non-extreme properties are concerned.

- The inclusion of more fitting properties in the objective function (2.1) can aid the identification of some parameters. However, the benefits from such an exercise must be weighed against the costs. For example, in applications where it is necessary to regionalise the fitting statistics, the extra work required to obtain the 13 properties used in OF2 would probably not be regarded as worthwhile. For national application therefore, it is recommended that models are fitted to the properties used in OF1.

- More work is required to decide how best to weight the fitting properties in the objective function. However, it appears that in general the weights used in OF1 (Table 2.8) can be relied upon to produce reasonable results.

- Some model parameters can be difficult to identify without at least 20 years of hourly observations. Identification problems are worse in summer than in winter.

- Results regarding parameter identifiability using various record lengths appear to offer genuine opportunities for parameter bounding, which will aid application of these models in a national setting. It is recommended that such relationships are further studied to examine such possible parameter bounding and regionalisation opportunities.

In the next section we investigate the models' ability to reproduce extreme rainfall properties. The conclusions above suggest that for this purpose, we should focus on model 4, fitted using OF1. However, since extremal properties are likely to be sensitive to the choice of cell depth distribution, we also consider model 5, which has the additional flexibility of a 2-parameter depth distribution. To proceed further with the model, it is necessary to specify this distribution completely rather than just via its mean and variance. We will therefore examine two versions: the first uses a gamma distribution for cell depths, and the second uses a Pareto. These represent "light-tailed" and "heavy-tailed" cases respectively.

2.5 Extreme value performance of selected models

2.5.1 Methodology

Extreme value performance of the "best" parameters

One of the primary motivations for developing stochastic rainfall generators is the ability to simulate extreme rainfall events which can be used for the purpose of flood design. The models' capacity to reproduce realistic maxima must therefore be assessed.

A first assessment of this capacity is visual. It involves plotting annual rainfall maxima from the observed data and from a model simulation, as in Figures 2.8–2.9. In these plots, the vertical axis represents annual maximum rainfall and the horizontal axis shows the reduced variate

$$y(T) = -\ln(-\ln(1-1/T)) \tag{2.4}$$

of the Gumbel distribution. Here T is a return period, assigned according to the Gringorten plotting position formula:

$$T = \frac{N+0.12}{k-0.44} \tag{2.5}$$

where

N is the number of years of data available

k is the rank of the observation in the ordered set of maxima ($k = 1$ for the largest)

The observed rainfall maxima $\left\{ Y_k^{(h)} : k = 1,\ldots,N \right\}$ at time scale h are plotted in bold with approximate 95% prediction intervals in dotted lines. The fine line represents extremes $\left\{ R_k^{(h)} \right\}$ produced by a long simulation (five times the length of the data set) of model 4 in Table 2.6, using optimal parameters according to OF1 of the previous section.

A number of authors (e.g. Cameron et al. 2000) have noted that mixed results have been obtained from such assessments. Figures 2.8–2.9, as well as corresponding figures for the other two sites examined above, confirm this. Thus, a performance ranging from very good to very poor is obtained in the case of the Random Parameter Bartlett-Lewis model with exponentially distributed cell intensities for instance.

The advantage of the methodology in the previous section is, however, that it allows the identification of near-optimal regions of the parameter space; moreover, the fitting algorithm described in Section 2.3.2 returns several sets of near-optimal parameters whose performance in terms of mean statistics, is not significantly different. For different parameter sets, the extreme-value behaviour could potentially be very variable. This is because the tail of the distribution of rainfall intensity is not uniquely determined by the sole knowledge of mean statistics for a finite number of orders.

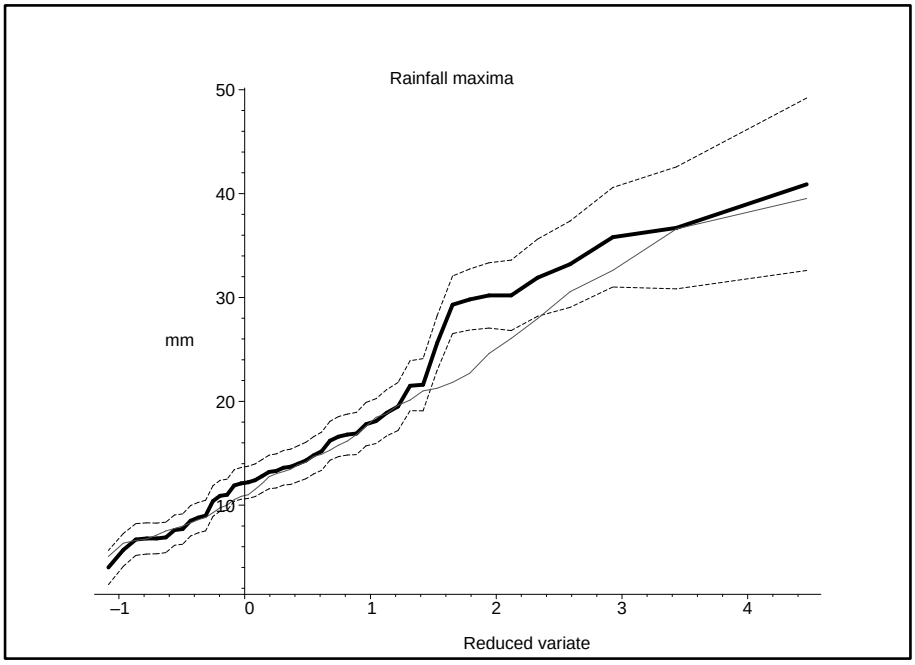


Figure 2.8 Model 4 at Elmdon — July daily extremes for best parameters under OF1

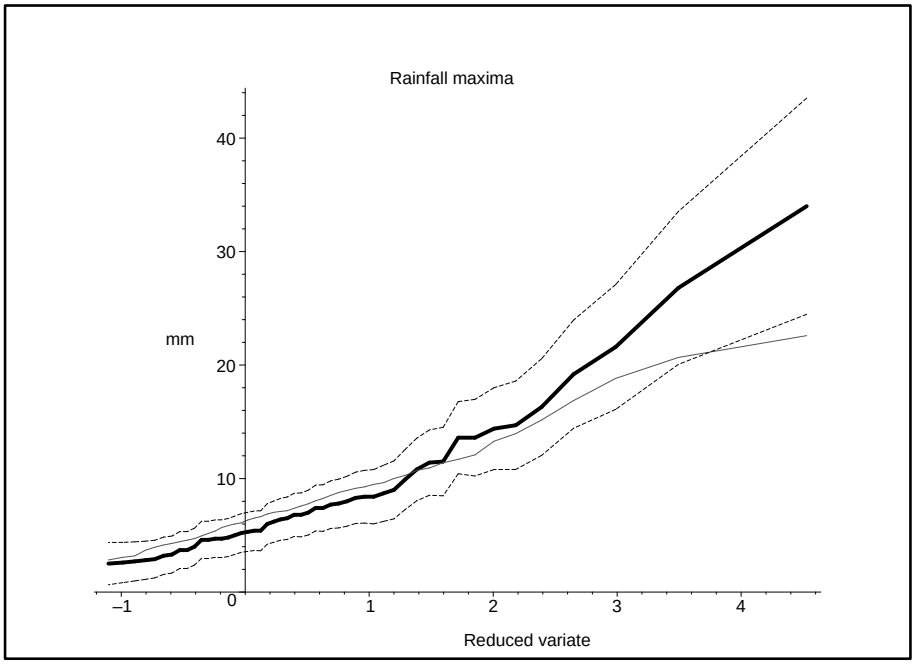


Figure 2.9 Model 4 at Elmdon — January hourly extremes for best parameters under OF1

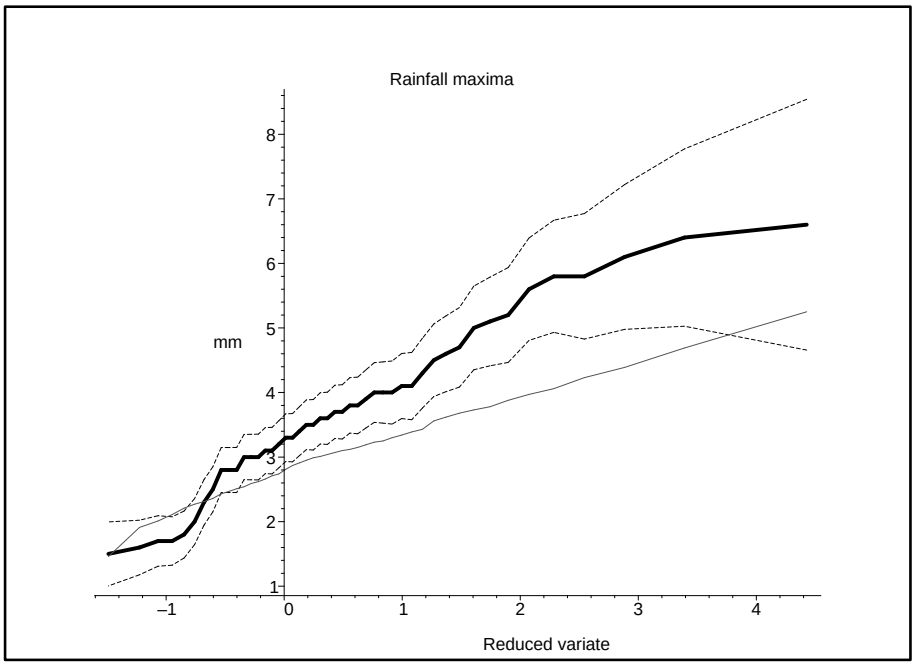


Figure 2.10 Model 4 at Heathrow — July hourly extremes for best parameters under OF1

Two-stage fitting

We therefore propose to add a second phase to the model fitting. This involves comparing the maximum intensities produced by simulations with near optimal parameters, so as to identify the best parameter set in terms of its extremal performance. The method is presented below.

Many studies (as reviewed in Onof et al. 2000) use fits to an Extreme-Value (EV) distribution (such as the Gumbel, the Fréchet or the General Extreme Value distribution) to compare maxima generated by model simulations with the extremes extrapolated using this distribution from the observations. This method has the advantage of enabling comparisons of rainfalls of return periods far larger than the size of the data set. But the comparison is not with the data, but with a model (the EV distribution) fitted to the data. If the model is however to be calibrated to the *observed maxima*, this method is not optimal. We therefore prefer to restrict the comparison to one between observed and simulated maxima. To reduce the variability of the model estimate of the maxima, a long simulation is carried out (here we use a length equal to 5 times that of the observation data set). Details of the simulation are given in appendix C.

For a given time-scale h, the comparison illustrated above (e.g. in figure 2.8) is between two curves. To summarise it, we consider a weighted sum of squares of the differences between observed maxima $Y_k^{(h)}$ and simulated maxima $R_k^{(h)}$:

$$\sum_{k=1}^{k=N} \omega(k,h)(Y_k^{(h)} - R_k^{(h)})^2 \tag{2.6}$$

where $\omega(k,h)$ is the weight to be put on the extremes of rank k at time-scale h.

Since most time-scales of interest are likely to lie between 1 and 24 hours, an objective function can be formed by weighting the sums of squares obtained for these two time-scales, as follows:

$$\Omega \sum_{k=1}^{k=N} \omega(k,1)(Y_k^{(1)} - R_k^{(1)})^2 + (1-\Omega) \sum_{k=1}^{k=N} \omega(k,24)(Y_k^{(24)} - R_k^{(24)})^2 \qquad (2.7)$$

where Ω is the proportional weight to be put on the hourly extremes.

The optimal parameter set for the proposed two-phase fitting method will be that which minimises objective function (2.7).

Choice of weights

A satisfactory set of weights $\omega(k,h)$ will take into account the greater uncertainty involved in estimating maxima of larger return periods. An obvious way of doing this is to use as weight the inverse of the variance of the rank statistic of order k. For a distribution with density function f and c.d.f. F, this variance is (Rice, 1995, p. 376):

$$\text{Var}\left(X\left(k\right)\right) = \frac{k\left(N+1-k\right)}{(N+1)^2(N+2)} \left\{ f\left[F^{-1}\left(\frac{N+1-k}{N+1} \right) \right] \right\}^{-2} \qquad (2.8)$$

(the difference between this and the formula in the cited reference arises because there the observations are ordered from smallest to largest whereas here, they are from largest to smallest). To estimate (2.8), we can assume that the maxima are approximately Gumbel distributed. This is generally a sensible assumption for the range of return periods T obtained for most data sets, i.e. smaller than 100 years, since it is often beyond this range that important divergences between Gumbel and Fréchet fits are observed for rainfall extremes (Onof, 1992). With this assumption, we can obtain the weight for order k:

$$\omega(k,h) = \frac{\alpha(h)^2 \left[\ln\left(\frac{N+1-k}{N+1}\right)\right]^2 (N+2)}{k} \qquad (2.9)$$

where $\alpha(h)$ is the scale parameter of the Gumbel distribution fitted to the extreme intensities at time-scale h hours. This is given by:

$$\alpha(h) = 1.282/s(h) \qquad (2.10)$$

where $s(h)$ is the standard deviation of the observed extreme intensities at time-scale h hours.

The determination of the weight Ω which determines how much relative importance is assigned to the hourly extreme-value performance, ought to be application specific. If the model is used as input to a rainfall-runoff model, the size and responsiveness of the catchment may be determining factors. For small catchments, the emphasis will be almost exclusively upon the finer time-scales.

This weight should however be chosen with care, and the sensitivity of the output to this choice may have to be determined for each data set. For if its value is very different from 0 or 1, it is possible for the selected parameters to perform rather poorly in reproducing both hourly and daily extremes. In the results shown below, Ω is chosen equal to 1.

2.5.2 Monthly performance

In this section, we shall look in turn at each of the three gauges and the three model variants identified in Section 2.4.4, namely the Random Parameter Bartlett-Lewis model with exponential, Gamma and Pareto distributed cell intensities.

The plots shown below are plots of the annual maxima for two months of the year (January and July) for these three model variants. A couple of comments about the plots are required.

- Each line corresponds to a parameter set. The number of parameter sets which have been used for simulation varies between twenty and forty. These parameter sets are found to be near-optimal by the first stage of the fit, using the algorithm described in Section 2.3.2. For each such parameter set, a simulation of length 5 times the length of the data set was performed and the maxima recorded.

- For certain parameter sets, no simulation was carried out, or the simulation was interrupted. The first may have happened because of values of the storm arrival rate that were too small (smaller than 1 in 1000 hours). The second would have been caused by a storm with a number of cells in excess of 2000, or of duration in excess of 1600 hours (which corresponds to a couple of months), or because of a number of storms in excess of 2000 per year. For such parameter sets, the maxima are all set to -1, so that they can be identified in the plots by the presence of horizontal lines below the x-axis.

- The historical maxima are represented by a bold black line and prediction intervals are in fine black dotted lines around it.

- The best performing parameter set for the given time-scale (hourly or daily) is shown as a dashed bold line.

Elmdon

The exponential model performs very well in reproducing the hourly observed maxima for January (Figure 2.11). The optimal simulation is mostly within the 95% prediction interval around the observed maxima for July, but the tendency is clearly to underestimate the observed extremes (Figure 2.12). For daily maxima, the performance is excellent in July but the largest values are underestimated in January (see Figures 2.13 and 2.14).

The gamma distribution does not improve upon the reproduction of the July hourly extremes, (see Figure C.313 in appendix C), and the optimal set of parameters for the Pareto distribution produces maxima that are on the edge of the prediction interval around the observations (Figure C.314). The gamma distribution is, on the other hand, able to produce higher daily extremes in January, thus improving upon the performance of the exponential distribution (Figure 2.15).

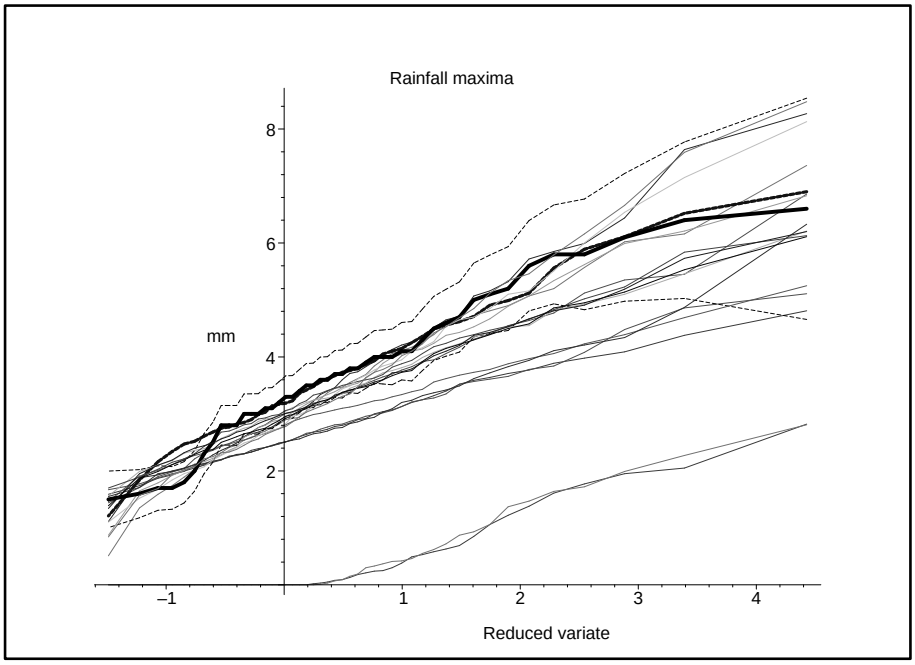


Figure 2.11 Model 4 at Elmdon — January hourly extremes (see details in Section 2.5.2)

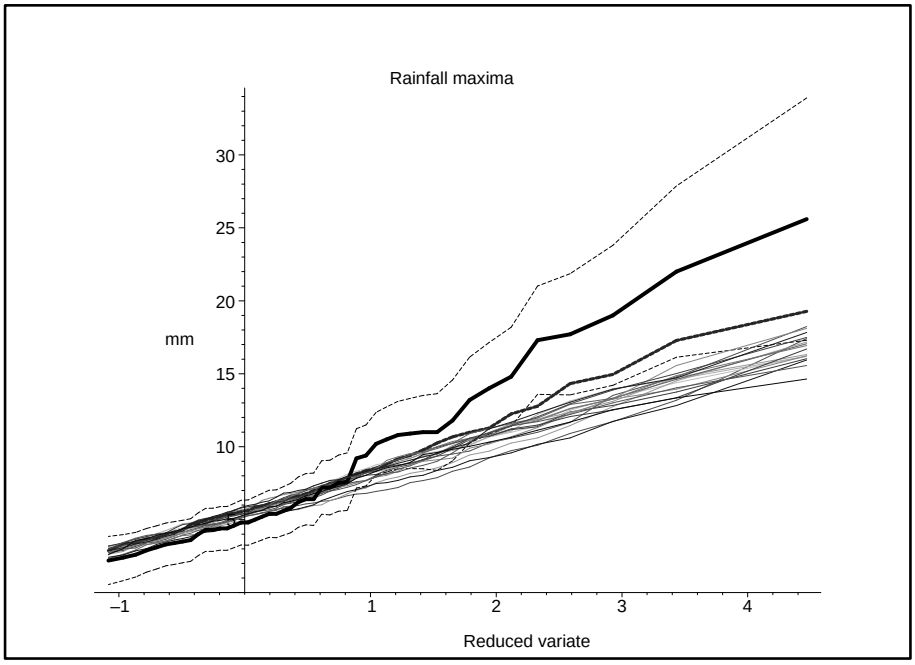


Figure 2.12 Model 4 at Elmdon — July hourly extremes (see details in Section 2.5.2)

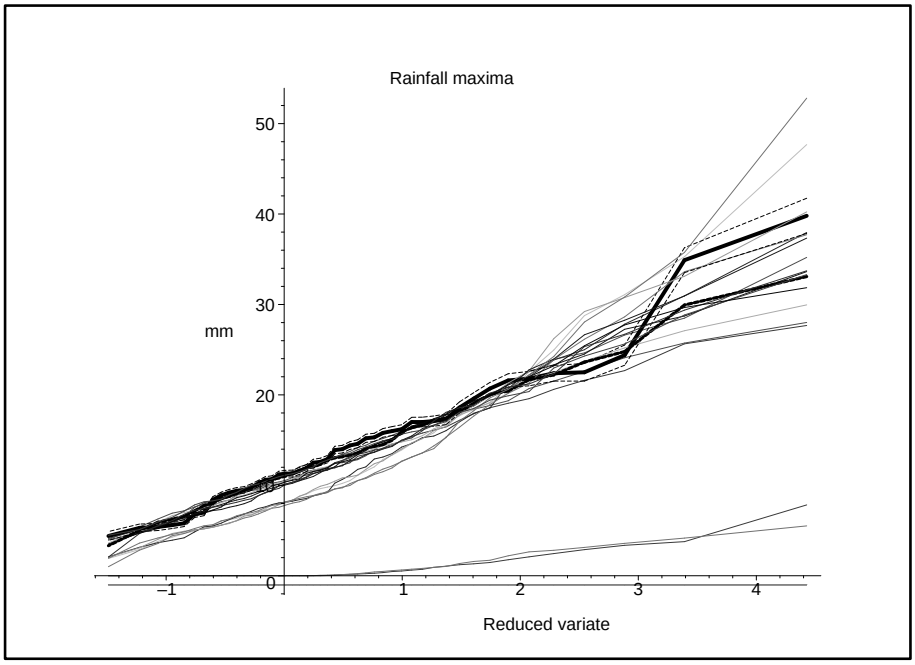


Figure 2.13 Model 4 at Elmdon — January daily extremes (see details in Section 2.5.2)

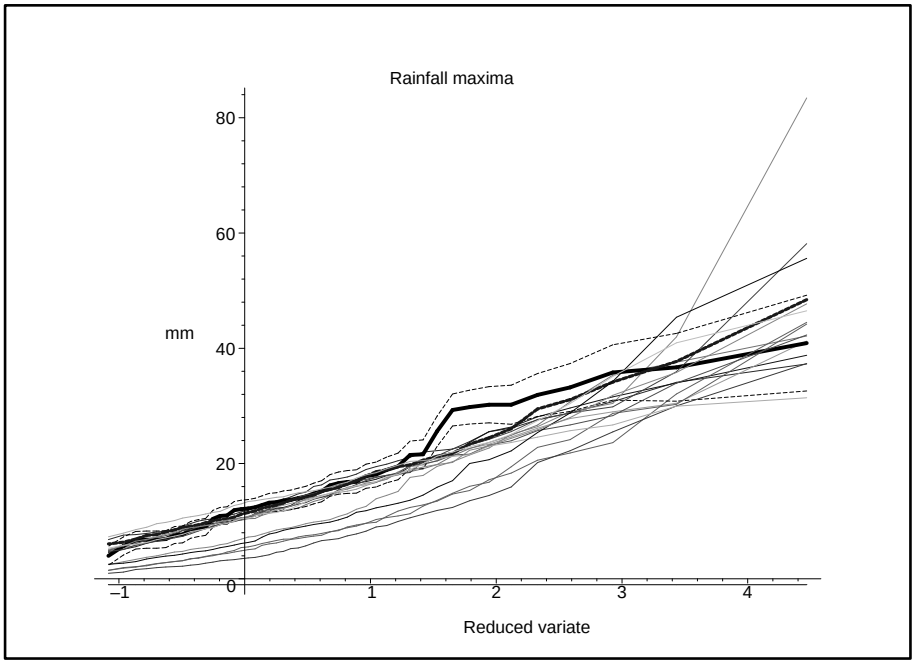


Figure 2.14 Model 4 at Elmdon — July daily extremes (see details in Section 2.5.2)

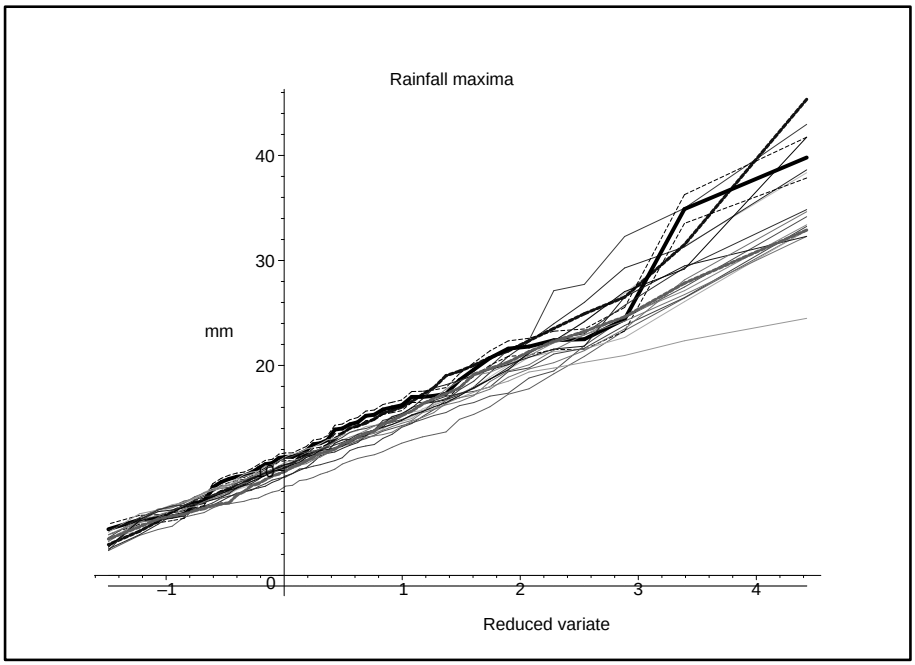


Figure 2.15 Model 5 (gamma version) at Elmdon — January daily extremes (see details in Section 2.5.2)

Plynlimon

The rainfall time-series at Plynlimon is characterised by very extreme maxima, in particular at the hourly time-scale. In appendix C, Figure C.315 shows that the January hourly maxima are well reproduced by the best simulations from the exponential version of the model; in July, however, the simulations underestimate the observed maxima (Figure C.316).

The reproduction of daily extremes is good for both months (e.g., Figure C.317). This shows that the clustering mechanism is able to compensate for the model's underestimatation of the hourly extremes over a range of return periods.

The gamma distribution somewhat improves the July hourly performance, but the January hourly performance is poorer (Figures C.318 and C.319). The characteristic fat tails of the Pareto distribution are able to improve further the reproduction of the July maxima, but again, at the cost of a poorer January hourly performance (Figures C.320 and C.321).

Heathrow

At the hourly time-scale, the exponential model provides a very good reproduction of the January and the July hourly maxima (appendix C, Figures C.322 and C.323). The same is true at the daily time-scale (Figures C.324 and C.325). At this site there are no cases where the addition of a parameter leads to improved performance. For instance, the optimal parameter set identified for January hourly maxima with the gamma model displays a poorer performance than that of the exponential model (Figure C.326).

2.5.3 Annual performance

The results for January and July confirm that no new grounds can be found to prefer a two-parameter distribution to the one-parameter exponential distribution for cell depths. The validation of the model for its overall extreme-value performance will therefore be carried out for the Random Parameter Barlett-Lewis exponential model. In the annual maxima figures shown below, a 95% prediction interval is displayed around the maxima generated by 100 year simulations with the exponential model.

The Elmdon annual extremes show an underestimation of the largest hourly extremes and an overestimation of the largest daily extremes (Figures 2.16 and 2.17).

The picture is very similar at the daily time-scale for the Plynlimon maxima (appendix C, Figure C.427), but the hourly maxima are very well reproduced (2.18). The poor hourly July performance has thus little impact upon the annual results insofar as only one of the historical annual hourly maxima was observed in July.

The annual performance of the exponential model at Heathrow is very similar to that observed at Elmdon. It displays a tendency to underestimate hourly extremes and overestimate daily extremes (see Figures 2.19 and 2.20). Given the excellent reproduction of January and July maxima, this may seem surprising. However, the performance for other months is not as good, as can be seen from the case of June (Figure 2.21).

Thus the variability between gauges of the extreme value performance of the Random Parameter Bartlett-Lewis exponential model is also found displayed between months at a given gauge. This analysis of the annual extremes suggests that, although over a large range of return periods, the Random Parameter Bartlett-Lewis exponential model can reproduce the observed extreme rainfall depths, this no longer holds for return periods longer than 10 years.

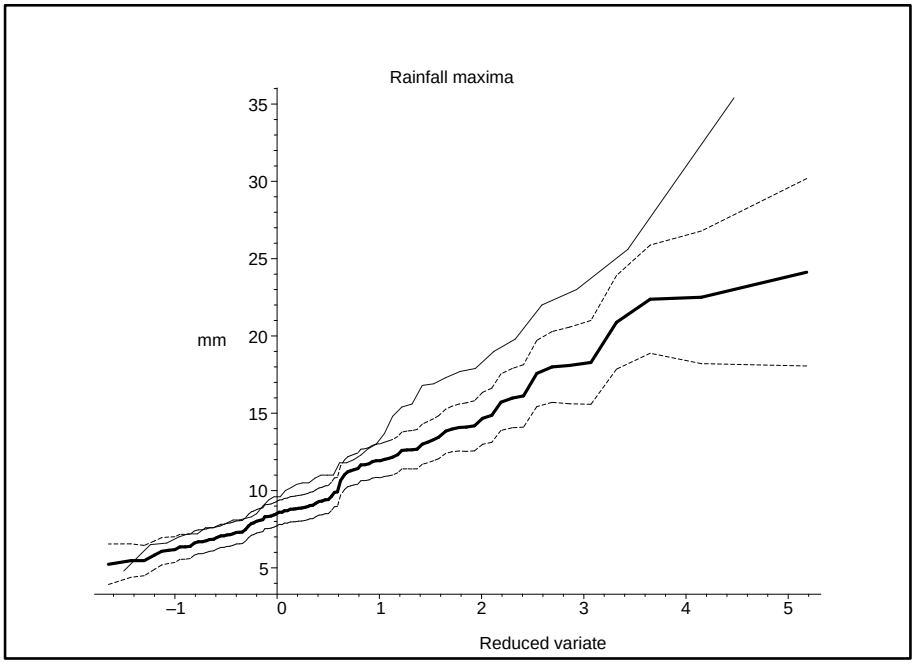


Figure 2.16 Model 4 at Elmdon — Annual hourly extremes. Observed maxima (fine) and maxima from 100 year simulation (bold) are plotted against the reduced variate with 95% prediction intervals (dotted line)

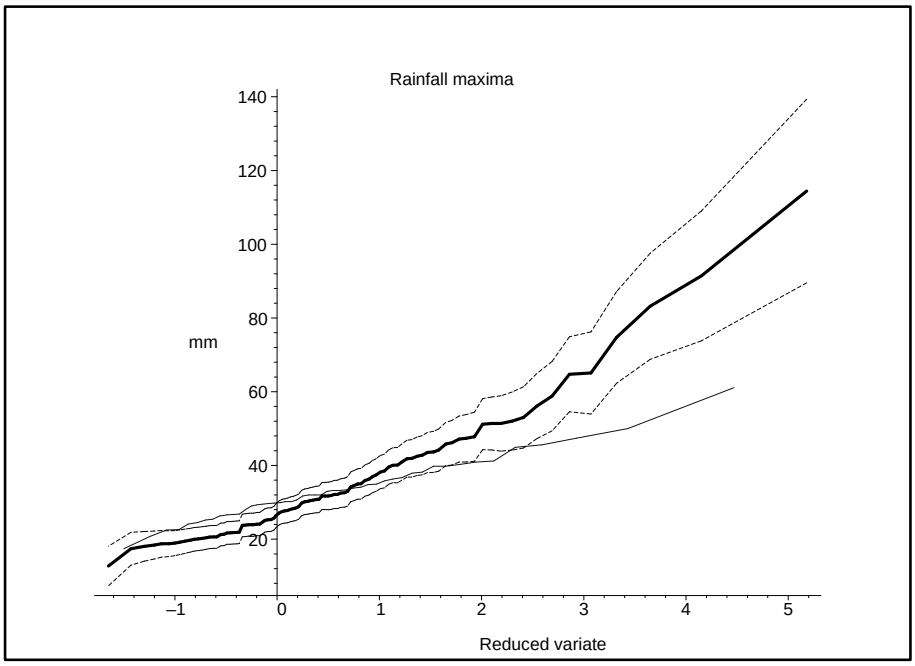


Figure 2.17 Model 4 at Elmdon — Annual daily extremes. See Figure 2.16 for legend.

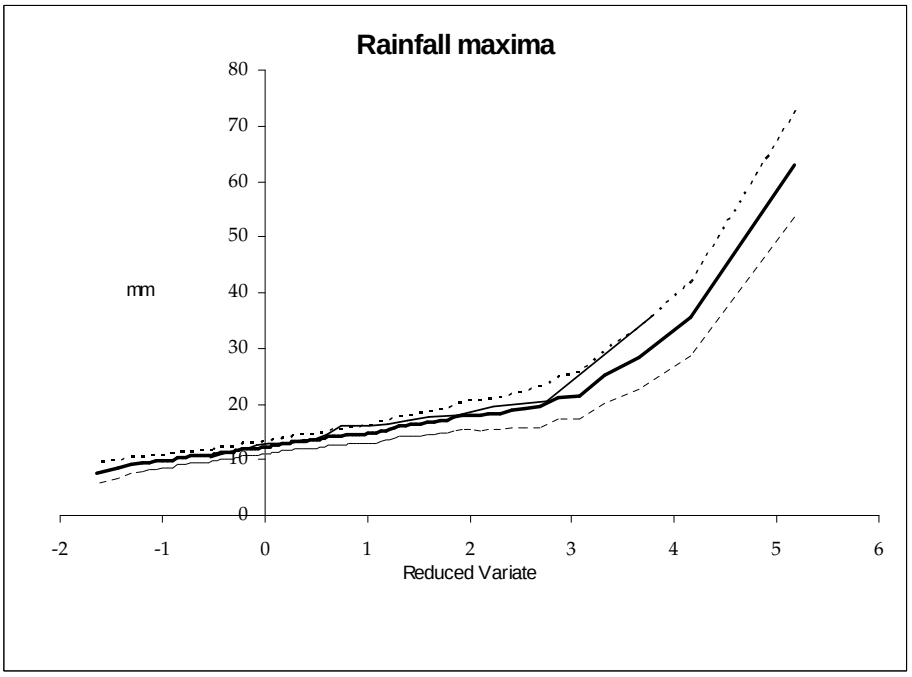


Figure 2.18 Model 4 at Plynlimon — Annual hourly extremes. See Figure 2.16 for legend.

2.6 Conclusions and recommendations

The aim of this work was to recommend a single-site rainfall model for national application, and to provide guidelines regarding its use. Based on the literature review in Lekkas (2002) and the extensive investigation summarised in this chapter, the following recommendations can be made:

- To ensure reproduction of both average and extremal rainfall properties, a two-stage fitting processes is recommended. The first stage aims to generate several parameter sets that are near-optimal in terms of reproduction of average properties; the second identifies which of these sets yields the best reproduction of extremes.
- For the first stage of fitting, an objective function based on 8 properties, corresponding to OF1 of Table 2.8, will generally yield reasonable results.
- Increasing model complexity leads to problems associated with parameter identifiability and numerical instability during optimisation. This provides grounds for preferring a 6-parameter model over more complicated structures.
- Some model parameters can be difficult to identify in the first stage without at least 20 years of hourly observations.
- Problems with parameter identifiability can be reduced, but not altogether eliminated, by increasing the number of properties in the objective function.
- There is little to choose between the Neyman-Scott and the Bartlett-Lewis clustering mechanisms. However, the first is more computationally demanding

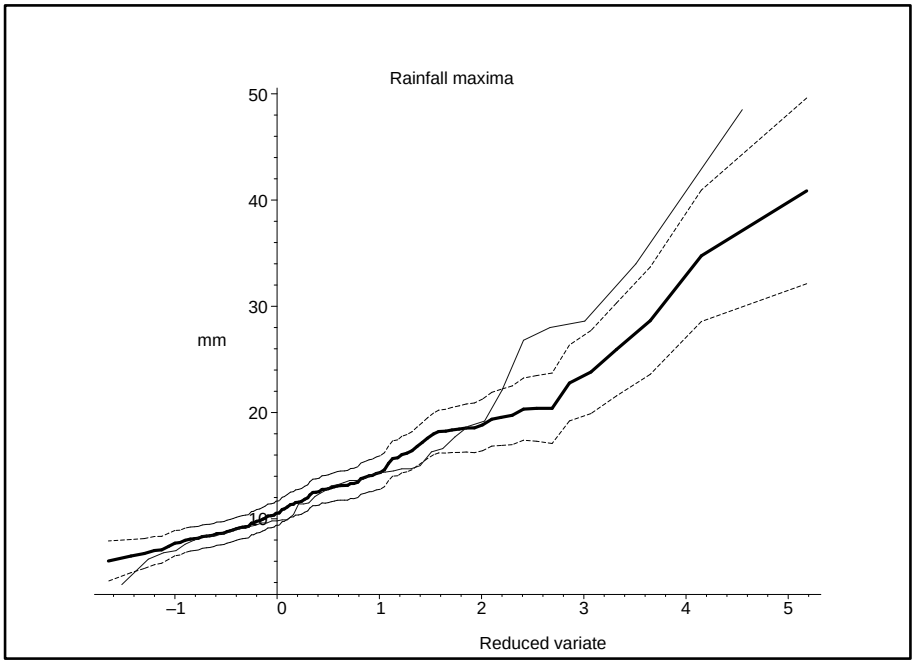


Figure 2.19 Model 4 at Heathrow — Annual hourly extremes. See Figure 2.16 for legend.

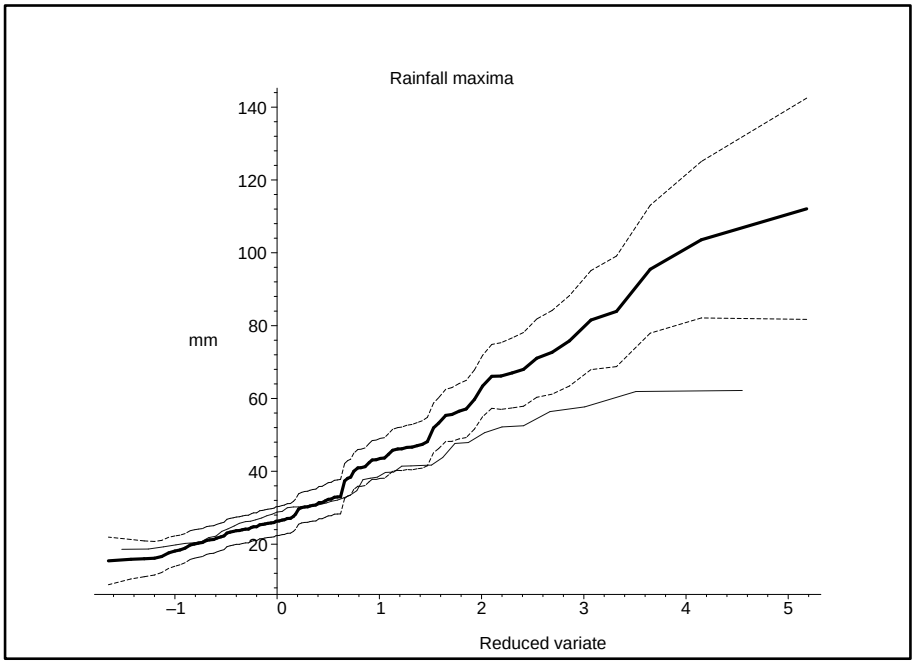


Figure 2.20 Model 4 at Heathrow — Annual daily extremes. See Figure 2.16 for legend.

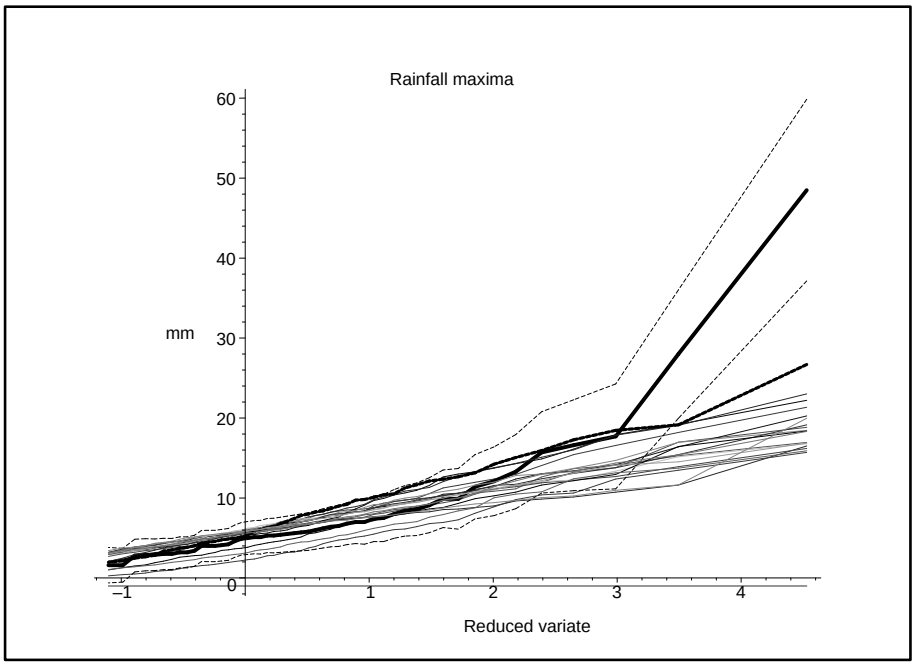


Figure 2.21 Model 4 at Heathrow — June hourly extremes (see details in Section 2.5.2)

since it requires numerical evaluation of the proportion of dry periods at each iteration of the numerical optimisation.

- Results regarding parameter identifiability using various record lengths appear to offer genuine opportunities for parameter bounding, which will aid application of these models in a national setting. It is recommended that such relationships are further studied to examine such possible parameter bounding and regionalisation opportunities.

- In the second stage of fitting, using the objective function (2.7), it is best to choose Ω close to either 0 or 1, so as to enable good reproduction of extremes at one of the two time scales at least. Choices of Ω that give non-negligible weights to both hourly and daily time-scales run the risk of selecting parameter sets that perform poorly at both.

- When compared with the exponential distribution, the use of two-parameter gamma and Pareto cell depth distributions does not systematically improve model performance with respect to extremes. The exponential is therefore preferred. We note however that extremal performance is variable over the year and between gauges, so that the reproduction of the largest extremes cannot be guaranteed.

- The preferred model of choice for national application is the Random Parameter Bartlett-Lewis model with exponential cell depth distribution.

For both phases of model fitting, software has been produced which has been successfully tested on a number of single-site data sets representative of the rainfall variability present within the UK. On the basis of a number of simulations at these

sites, bounds have been set on the distributions which are sampled from, as indicated in appendix C. The chosen models reproduction of extreme-values is significantly improved as a result of the two-phase fitting method, as compared with the performance recorded in the literature.

It is however possible that unacceptably large rainfall depths should be generated for other simulations or other sites. It is therefore important to use the simulated data carefully. In particular, if at any time-scale, rainfall depths are obtained which exceed either the Probable Maximum Precipitation (PMP) or the $10,000$ year rainfall, they should be discarded as outliers.

Although the performance is improved with the two-phase fitting method, some deficiencies remain in terms of reproduction of extreme statistics. We conclude that the model and optimisation technique are appropriate for use with caution, with the user undertaking careful verification of the simulated outputs. More extensive testing is required on a wider range of sites before the model can be recommended for general application across the UK. It is recommended that alternate means of improving the extreme value performance of the model (such as the use of skewness in fitting) are compared to the newly introduced fitting technique.

Chapter 3 Simulation and downscaling models for potential evaporation

3.1 Overview

In simple terms, many hydrological systems can be reduced conceptually to a water balance equation:

$$\text{Precipitation} - \text{Evapotranspiration} = \text{Runoff} - \text{Change in storage} . \qquad (3.1)$$

Hydrological runoff models are designed to solve this equation, for a given precipitation sequence (and possibly the associated evapotranspiration).

There is currently a substantial amount of research interest in the use of runoff models to simulate hydrological systems over long periods of time. The hope is that this use of 'continuous simulation' will yield more insight into the behaviour of a system than simpler approaches based, for example, on analysis of historical events. Continuous simulation does, however, require an ability to simulate realistic precipitation sequences at an appropriate timescale. Consequently there is a vast body of literature devoted to the development of precipitation models for this purpose. In comparison, attempts to simulate evapotranspiration sequences, which may also be required by runoff models, are in their infancy.

The processes governing evapotranspiration are complex, and it can usually be measured only indirectly (Shuttleworth, 1993). For this reason, many runoff models work with *potential evaporation* (PE) rather than actual evapotranspiration. In the UK, since 1978 the primary source of PE data for hydrological applications has been the Meteorological Office Rainfall and Evaporation Calculation System (MORECS). This uses a modified version of the Penman-Monteith equation (Monteith, 1965) to estimate PE on a 40×40km^2 grid covering the whole of the UK. For details, see Thompson et al. (1981). The calculations are based on daily values of meteorological variables, along with information on land usage. However, the output is at a weekly timescale, which is unsatisfactory in those hydrological applications for which daily or subdaily values are required. In dealing with this problem, the pragmatic approach of assuming PE to be constant throughout the week is widespread in current practice (E. Stewart, personal communication). As a zero-th order approximation this is probably reasonable, since evaporation varies much more smoothly in time than does precipitation, for example. However, the effects of ignoring finer-scale variability in PE sequences are not known.

In the work reported in this chapter, models have been developed for both weekly and daily PE sequences. These models are effectively heteroscedastic autoregressions. They have been constructed in such a way as to allow simulation of weekly or daily sequences, and also to simulate realistic daily sequences that are consistent with a given weekly sequence. They are therefore suitable both for use in continuous simulation studies, and for the derivation of realistic daily historical sequences on the basis of weekly MORECS data.

The methodology is illustrated using data from the south of England. The next section provides a summary of these data. Sections 3.3 and 3.4 describe the development of

models for weekly and daily PE, as well as analyses of the relationship between precipitation and PE. In Section 3.4.4, these models are used to develop a downscaling procedure that is appropriate for generating daily sequences subject to fixed weekly totals; and the work is summarised in Section 3.5.

3.2 Data and preliminary analysis

Three separate sources of data have been used in this study:

1. Weekly MORECS PE data (in mm) from a single 40×40 grid-square in Surrey, in Southern England. The data are available from January 1961 to March 1993, and are a subset of the Thames catchment data used in Jolley and Wheater (1996). The MORECS output includes PE calculations for a homogeneous grass surface, and for the estimated real land usage. For current purposes, the choice of PE measure is immaterial; the modelling methodology below can be applied in any situation, and we anticipate that the conclusions are unlikely to be sensitive to land use changes. Here, we focus on the 'grass' data since these are routinely used as a reference land use (Penman, 1948).

2. Daily rainfall and evaporation data (both in mm) from an experimental station at Silwood Park in Berkshire, which is located just outside the MORECS grid square. These run from July 1989 to August 1994. The data were collected as part of a set of lysimeter experiments carried out by Imperial College (Burne et al., 1994). The rainfall data were gathered using a 0.2mm tipping-bucket gauge; the PE calculations used the MORECS form of the Penman-Monteith equation, with a winter wheat crop.

3. Daily rainfall data (in mm) from a gauge located at Rotherfield Park in Surrey, near the centre of the MORECS grid square. The record spans the period 1923–2000; 1557 of the daily values (around 0.5%) are missing. These data are used to investigate the relationship between PE and rainfall.

Figure 3.1 shows the locations of these data sources. The data have been checked for possible errors and outliers; their quality appears reasonable.

To give some idea of the data structure, the boxplots in Figure 3.2 show the monthly distributions of weekly PE and rainfall. The weekly rainfalls here are obtained by aggregating the daily values from Rotherfield Park. It is clear from this figure that the seasonal cycle is much more pronounced in the PE data than in the rainfall. It is also clear that PE is more variable in summer than in winter.

To investigate the possibility of trends in the evaporation data, Figures 3.3 and 3.4 show time series plots of their mean and standard deviation. Separate plots are produced for each month of the year, to allow for the possibility of seasonally-varying trends. In almost all months, there appears to be a long-term increase in the mean series; this is especially noticeable from 1980 onwards. Less structure is apparent in the standard deviations.

As a preliminary indication of the relationship between PE and rainfall, the cross-correlation function between the weekly series was calculated. As expected,

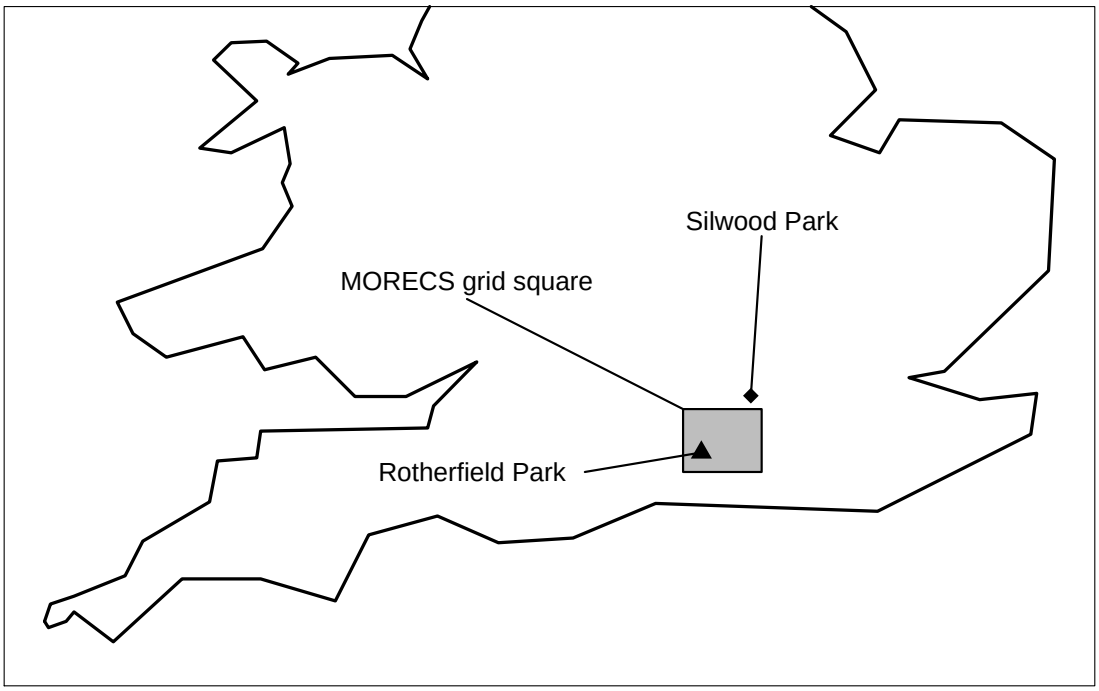


Figure 3.1 Sketch map of the southern UK, showing locations of data sources used for this study.

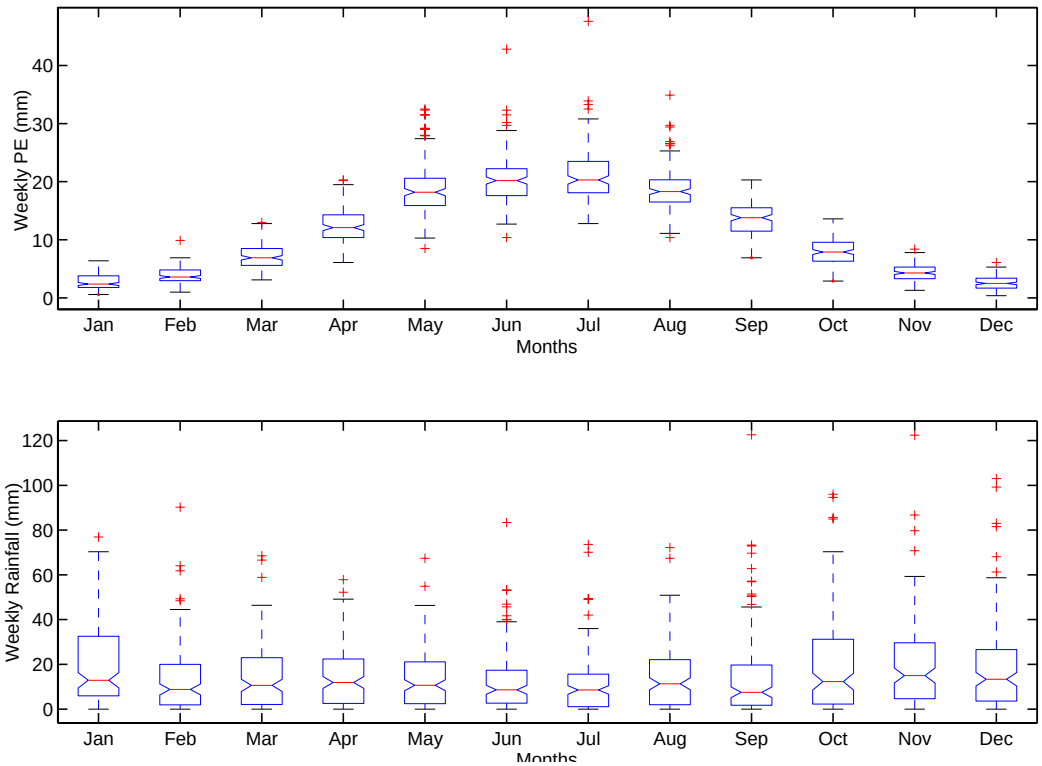


Figure 3.2 Top: monthly distributions of weekly PE, from MORECS. Bottom: monthly distributions of weekly rainfall, from Rotherfield Park gauge.

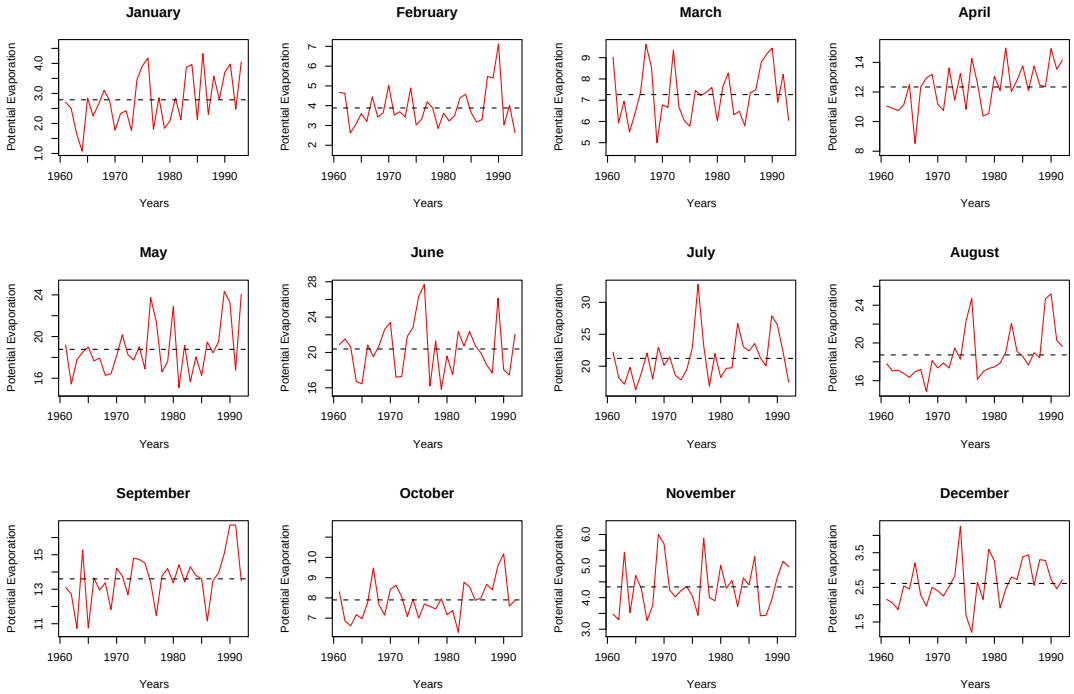


Figure 3.3 Mean weekly potential evaporation (in mm), by month, 1961–1993. Dotted lines represent series means.

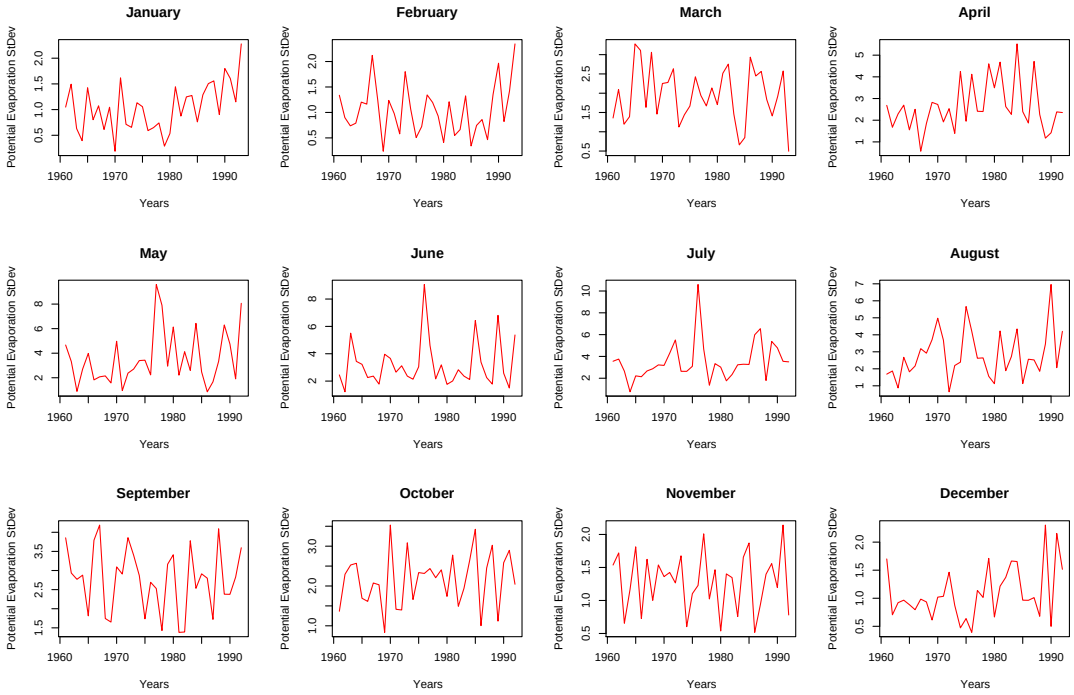


Figure 3.4 Standard deviation of weekly potential evaporation (in mm), by month, 1961–1993.

there is a negative correlation at zero lag (-0.148, decreasing to -0.068 if seasonality is first removed from both series). However, all of the cross-correlations are small, indicating that any relationship is relatively weak.

Similar analyses have been carried out using the limited daily PE data from Silwood Park. The results were similar to those reported above for the weekly MORECS data, and the details are therefore not shown here.

In summary: these preliminary analyses indicate a strong seasonal signal in both the mean and the variance of PE sequences. In addition, any realistic model for the simulation of PE should allow for the possible existence of long-term trends. There is, as expected, a relationship between rainfall and PE, although this relationship appears relatively weak. These are preliminary indications only, however: we now proceed to a formal investigation using statistical models.

3.3 A model for weekly potential evaporation

In this section we focus on the development of a model for weekly PE. As indicated in Section 3.1, the primary motivation for this work is the provision of PE sequences, for use in continuous simulation studies. Strictly speaking, such studies require bivariate sequences of rainfall and evapotranspiration, incorporating a realistic dependence structure between the two variables. The task can be regarded as that of drawing a sample from the joint distribution of two random vectors: $\mathbf{R} = (r_1 \ \ldots \ r_T)'$ representing rainfall and $\mathbf{Y} = (Y_1 \ \ldots \ Y_T)'$ representing PE. Noting that any such joint distribution can be factorised as

$$f_{RY}(\mathbf{y}, \mathbf{r}) = f_{Y|R}(\mathbf{y}|\mathbf{R} = \mathbf{r}) f_R(\mathbf{r}) \ , \tag{3.2}$$

a bivariate sequence can be obtained in two stages:

1. Simulate a rainfall sequence $\mathbf{r}$ from a suitable rainfall model.
2. Simulate a PE sequence $\mathbf{y}$ using a model for $f_{Y|R}(\mathbf{y}|\mathbf{R} = \mathbf{r})$, the conditional distribution of PE given the rainfall sequence $\mathbf{r}$.

When developing models for PE therefore, it is necessary to pay particular attention to its relationship with rainfall. It could be argued that, since PE is calculated deterministically from various meteorological variables, the development of a statistical model for this relationship is unnecessary. However, in many applications the values of the required meteorological variables will not available; a statistical model is therefore a convenient device for generating the required sequences.

3.3.1 Methodology

The PE models developed here are regression-based, and can be written as

$$Y_t = \mathbf{x}_t \beta + \varepsilon_t \qquad (t = 1, \ldots, T) \tag{3.3}$$

where Y_t is the potential evaporation at time t, $\mathbf{x}_t$ is a corresponding vector of covariates, β is a corresponding vector of regression coefficients, and $\varepsilon_1, \ldots, \varepsilon_T$ is an

uncorrelated sequence of zero mean errors. From the analysis in Section 3.2, the covariates should also include some representation of seasonality. Autocorrelation can be incorporated by including lagged PE values as covariates. The preliminary analysis also indicated the possibility of trends in the data; the significance of these can be tested formally within the regression modelling framework, as can the dependence of PE upon rainfall.

The standard theory of linear regression requires that all of the errors $\{\varepsilon_t\}$ have the same variance. It is clear from Figure 3.2 that this assumption does not hold for the PE series — in particular, there is a strong seasonal cycle in the variance. Since there are roughly equal numbers of observations in each month, this heteroscedasticity will have little effect on the standard F test for comparing different models (Montgomery, 1997, Section 3.4). However, it should be reflected in any model designed for the simulation of PE sequences. An elegant solution is possible when the elements of ε are normally distributed. Suppose, for example, that $\varepsilon_t \sim N\left(0, \sigma_t^2\right)$, where the variance σ_t^2 depends on the time of year (and possibly on other factors as well). In this case, $\varepsilon_t^2/\sigma_t^2$ has a chi-squared distribution with 1 degree of freedom, which is the same as a gamma distribution with mean 1 and shape parameter 0.5. From standard properties of the gamma distribution, therefore, ε_t^2 has a gamma distribution with mean σ_t^2 and shape parameter 0.5. This can be represented as a generalized linear model (GLM — see McCullagh and Nelder 1989; Dobson 2001), which allows some function of σ_t^2 to be related to a linear combination of covariates at time t:

$$h\left(\sigma_t^2\right) = \xi_t \gamma \,, \tag{3.4}$$

where ξ_t is another vector of covariates, γ is another vector of parameters and $h(.)$ is a monotonic link function.

Note that (3.3) can equivalently be written as

$$\mu_t = \mathrm{E}\left(Y_t | \mathbf{x}_t\right) = \mathbf{x}_t \beta \,. \tag{3.5}$$

We refer to this as the 'mean model'; (3.4) as the 'variance model', and the model obtained by combining (3.4) and (3.5) as the 'joint model'. A maximum likelihood fit of the joint model can be obtained using an iterative procedure (McCullagh and Nelder, 1989, Section 10.2), alternating between the following two steps until convergence is deemed to have occurred:

1. Fit the mean model (3.5) using weighted least squares, with weights inversely proportional to the variances given by the variance model (for the first iteration, all the variances are assumed equal).

2. (Re-)estimate the parameters of the variance model, by fitting a gamma GLM to the squared residuals from the mean model.

Chandler (2003) gives an extensive example of this technique, in the context of modelling monthly temperatures. This paper also discusses some of the statistical issues that arise when modelling time series using regression methods. The bottom line is that, so long as the covariate vectors $\mathbf{x}_t$ and ξ_t incorporate an adequate representation of time series dependence structure, models can be fitted and compared as if the observations were independent. Comparisons can be based on

log-likelihoods and proportions of variance explained (R^2). R^2 is a familiar measure in regression modelling, but must be applied to the mean and variance components of a model separately — and for the variance component, the concept of 'explained variance in the squared residuals from the mean model' is not a particularly intuitive one. The log-likelihood, on the other hand, accounts for both mean and variance structure simultaneously and therefore provides a single overall measure of model fit. If a model specifies a probability distribution with density $f_t\left(y|\mathcal{H}_t\right)$ for Y_t, conditional on the history $\mathcal{H}_t$ of the process up to time t, then the log-likelihood, given a sequence of observations $y_1,\ldots,y_T$, is

$$L=\sum_{t=1}^{T}\ln f_t\left(y_t|\mathcal{H}_t\right)\ . \tag{3.6}$$

In the current context, f_t is the density of a normal distribution with mean μ_t and variance σ_t^2; hence (3.6) is easily computed for any model.

3.3.2 Dependence of potential evaporation on rainfall

To simplify the modelling task a preliminary linear model, of the form (3.3) and ignoring heteroscedasticity, was used to investigate the importance of the relationship between weekly potential evaporation and rainfall. As indicated above, the changing variance is unlikely to affect the results of such an analysis to any great degree.

Figure 3.2 shows a clear seasonal cycle in the mean PE; hence the first variables included in this model were four trigonometric predictors representing this cycle:

$$\cos\left(\frac{2\pi j\times\text{day of year}}{366}\right)\quad\text{and}\quad\sin\left(\frac{2\pi j\times\text{day of year}}{366}\right)\qquad(j=1,2)\ , \tag{3.7}$$

where 'day of year' here relates to the end of the week under consideration. All four of the estimated regression coefficients were significantly different from zero at the 5% level. The next step was to account for autocorrelation in the PE series, by including lagged values of potential evaporation as extra covariates in the model; it was found that a first-order autoregressive structure describes the temporal dependence adequately.

Having accounted for seasonality and autocorrelation, rainfall predictors were added to the model to determine whether rainfall has a significant effect on evaporation. Various different rainfall parametrisations were investigated. These included the use of weekly rainfall amounts (both for the current week t and the previous week $t-1$) as covariates, as well as the use of separate covariates representing daily rainfalls during weeks t and $t-1$. In all cases, the regression coefficients relating to rainfall were non-significant. Moreover, when comparing the root mean squared error (RMSE) for models with and without rainfall covariates, it was clear that the rainfall information added very little to the models' predictive power. For example, when rainfall was added to a model containing 4 seasonal predictors and a first-order autoregressive term, the RMSE decreased from 2.618mm to 2.614mm. Hence, although rainfall is known to have some role in controlling potential evaporation, for practical purposes this role appears negligible at a weekly timescale. Thus, in developing a joint model for both the mean and variance of potential evaporation, rainfall can be excluded as a potential covariate.

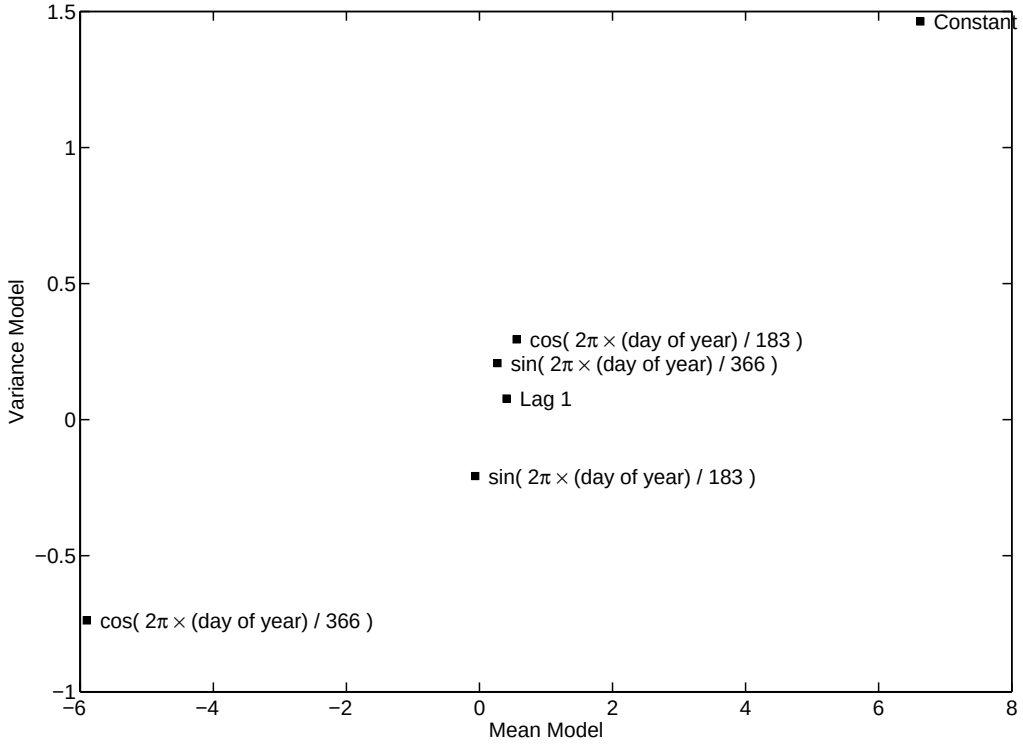


Figure 3.5 Relationship between estimated coefficients in mean and variance models, for joint model containing same covariates in each component.

3.3.3 Development of joint mean-variance model

The above analysis indicates that the mean of potential evaporation can be represented by a linear regression model with at least five predictors (four seasonal components and PE lag 1). We now turn to the development of a joint model for the mean and variance. Such models can be fitted using the iterative scheme described in Section 3.3.1. In the work reported here, convergence was deemed to have occurred when the maximum percentage change in any variance model parameters was less than 0.01.

Based on the preliminary analysis in Section 3.2, it is reasonable in the first instance to consider a variance model in which the standard deviation is linearly dependent on the same predictors as have been identified for the mean model (4 seasonal components and PE lag 1). The relationship with standard deviation, rather than variance, is natural from a consideration of measurement units, and corresponds to the link function $h(x) = x^{1/2}$ in (3.4).

For this initial joint model, the estimated coefficients for both mean and variance components are plotted against each other in Figure 3.5. This shows a clear linear relationship between the two sets of coefficients. Therefore, a simple alternative to modelling the variance as a function of separate predictors may be to represent the standard deviation as a linear function of the mean; this would effectively reduce the number of predictors in the variance model to 1 (besides a constant term).

The analysis in Section 3.2 also indicated the possible existence of linear trends, which may differ according to the time of year, in the weekly PE series. These can be represented via the inclusion of appropriate predictors in the models. Two

representations were investigated here. In the first, a single trend term was included to represent a common linear trend in all months. In the second, 12 separate terms were included, representing different trends in different months.

3.3.4 Results

A variety of different models have been compared, to determine the best balance between model performance and simplicity. The fitted models are summarised in Table 3.1 (standard deviation depending linearly on the mean) and Table 3.2 (standard deviation represented as a linear function of predictors).

Table 3.3 summarises the performance of each model in log-likelihood terms, along with R^2 values for the mean and variance components separately. The table shows that:

- All of the models are, for practical purposes, comparable in terms of R^2 (both for the mean and variance components). The log-likelihoods are much better at distinguishing between the models.
- The models with trend terms (numbers 2, 3, 5 and 6) offer significant improvements over those without (1 and 4). This is evidence that the trends identified in the preliminary analysis are genuine, and that they should therefore be incorporated into any scheme for simulating weekly PE sequences over this period.
- Model 3 leads to a log-likelihood increase of 12.46 over model 2; model 6 leads to an increase of 19.89 over model 5. The statistical significance of these increases can be tested by doubling them, and comparing to chi-squared distributions with 9 and 22 degrees of freedom respectively. The resulting p-values are 0.003 and 0.011 respectively, indicating that the more complicated models are superior in both cases. Hence we may conclude not only that the PE sequence contains a trend, but that the trend is seasonally-dependent.
- A likelihood ratio test of model 6 against model 3 leads to a test statistic of $2 \times (3584.04 - 3554.24) = 59.6$. Compared to a chi-squared distribution with $36 - 18 = 18$ degrees of freedom, this yields a p-value of 2×10^{-6}. This is overwhelming evidence that model 6 fits the observations much better than model 3 — hence the simple mean-variance relationship suggested by Figure 3.5 is not an adequate representation of the variance structure.

Residual analyses for each model are presented in Figures 3.6–3.11. These analyses are all based on the standardised residuals

$$e_t = \frac{y_t - \mu_t}{\sigma_t} \, , \tag{3.8}$$

on the grounds that these residuals will all have a standard normal distribution if the model is correct. For each model, the following plots are produced:

1. A quantile-quantile plot, to test the assumption of normality (if this holds, all the data points should lie approximately on a straight line).

Table 3.1 Coefficient estimates for joint models in which the standard deviation is a linear function of the mean. Model 1: no trend. Model 2: same trend in all months. Model 3: different trends in each month. μ_t and σ_t are linearly related to the predictors as in (3.4) and (3.5), with $h(x) = x^{1/2}$ in (3.4). All models are fitted to 1679 weekly observations.

Predictors in model	Model 1		Model 2		Model 3	
	μ_t	σ_t	μ_t	σ_t	μ_t	σ_t
Constant	7.3089	0.7040	5.7378	0.6757	5.6729	0.6792
$\cos\left(\frac{2\pi \times \text{day of year}}{366}\right)$	-6.5204		-6.7317		-6.6249	
$\sin\left(\frac{2\pi \times \text{day of year}}{366}\right)$	0.1573		0.1200		0.2790	
$\cos\left(\frac{4\pi \times \text{day of year}}{366}\right)$	0.7696		0.7919		1.0362	
$\sin\left(\frac{4\pi \times \text{day of year}}{366}\right)$	-0.0433		-0.0370		-0.0415	
Overall trend			0.0237			
January trend					0.0219	
February trend					0.0210	
March trend					0.0255	
April trend					0.0280	
May trend					0.0378	
June trend					0.0149	
July trend					0.0234	
August trend					0.0346	
September trend					0.0357	
October trend					0.0296	
November trend					0.0254	
December trend					0.0208	
Previous week's value	0.3461		0.3242		0.3097	
Conditional mean		0.1448		0.1458		0.1437

Table 3.2 Coefficient estimates for joint models in which the standard deviation is explicitly represented as a linear function of predictors in the mean model. Model 4: no trend. Model 5: same trend in all months. Model 6: different trends in each month. μ_t and σ_t are linearly related to the predictors as in (3.4) and (3.5), with $h(x) = x^{1/2}$ in (3.4). All models are fitted to 1679 weekly observations.

Predictors in model	Model 4		Model 5		Model 6	
	μ_t	σ_t	μ_t	σ_t	μ_t	σ_t
Constant	7.4674	1.4448	5.9272	0.6746	5.9447	0.7533
$\cos\left(\frac{2\pi \times \text{day of year}}{366}\right)$	-6.6713	-0.7462	-6.8660	-0.8240	-6.9826	-0.7112
$\sin\left(\frac{2\pi \times \text{day of year}}{366}\right)$	0.0934	0.2230	0.0586	0.2129	0.2498	0.6863
$\cos\left(\frac{4\pi \times \text{day of year}}{366}\right)$	0.7845	0.3091	0.8003	0.2859	1.0737	0.0824
$\sin\left(\frac{4\pi \times \text{day of year}}{366}\right)$	-0.0319	-0.2110	-0.0193	-0.1902	-0.0225	-0.3133
Overall trend			0.0232	0.0111		
January trend					0.0226	0.0104
February trend					0.0216	0.0076
March trend					0.0242	0.0006
April trend					0.0261	0.0050
May trend					0.0338	0.0085
June trend					0.0097	0.0155
July trend					0.0167	0.0181
August trend					0.0296	0.0155
September trend					0.0335	0.0148
October trend					0.0294	0.0152
November trend					0.0263	0.0116
December trend					0.0218	0.0111
Previous week's value	0.3323	0.0787	0.3101	0.0692	0.2981	0.0631

Table 3.3 Performance measures for weekly PE models. For the mean components, R^2 is computed from the residuals $\{y_t - \mu_t\}$ in the usual way; for the variance components, it is computed from the variance model residuals $\left\{\sigma_t^2 - (y_t - \mu_t)^2\right\}$.

Model number	R^2 for model components		Total number of parameters	Log-likelihood
	Mean	Variance		
1	0.8806	0.1018	8	-3615.28
2	0.8821	0.1027	9	-3596.50
3	0.8847	0.0926	18	-3584.04
4	0.8803	0.1187	12	-3595.02
5	0.8817	0.1197	14	-3574.13
6	0.8843	0.1092	36	-3554.24

2. A plot of standardised residuals versus fitted values. This is intended to reveal any systematic departures from the assumed linear relationship in the mean model (3.5). Any apparent structure in this plot indicates a problem with the model.

3. Boxplots showing the distributions of standardised residuals by month and by year. These are intended to check whether the models have successfully captured the seasonal structure and trends in the dataset, as well as to indicate possible outliers. Any systematic structure in the location of the boxes indicates a problem with the mean component of a model, and any systematic structure in their heights indicates a problem with the variance.

The residual analyses indicate that in practical terms, there is very little to choose between the models, although trends can be detected in the annual boxplots for models 1 and 4 as expected. All of the models appear to represent seasonality successfully. The quantile-quantile plots do indicate some differences between the models; in particular, models 4, 5 and 6 provide a better fit to the upper tail of the distribution. For each of these models, 7 outlying points remain in this upper tail; however, as these represent about 0.4% of the sample, for many purposes this can be considered adequate. The boxplots indicate that these outliers occurred predominantly in the early summer, during the mid-1970s.

3.3.5 Simplifying the representation of trends

The analyses above indicate that there is a trade-off between model performance and complexity. Formally, the most complicated model (model 6) is the best fit to the data, and hence will produce the most realistic simulated sequences of weekly PE. This model contains a separate trend term for each month of the year, in both mean and variance components. However, in practice it is likely that any seasonal variation in the trends will be smooth, so that we might reasonably expect the trend coefficients themselves to follow a seasonal cycle. This can be represented by including an overall trend term, along with trend/seasonal interaction terms which are formed by multiplying the trend by each of the seasonal terms in the model (see Chandler and Wheater (2002) for examples of interactions in the context of rainfall modelling).

To investigate this, a final model was fitted in which the separate monthly terms in model 6 were replaced by an overall trend and 4 extra predictors defined as

$$\text{Trend} \times \cos\left(\frac{2\pi j \times \text{day of year}}{366}\right) \quad \text{and} \quad \text{Trend} \times \sin\left(\frac{2\pi j \times \text{day of year}}{366}\right) \quad (j = 1, 2)\,. \tag{3.9}$$

This model contains a total of 22 parameters (11 for the mean model and 11 for the variance), and yields a log likelihood of -3559.0. Compared with a log-likelihood of -3554.2 for model 6 in Table 3.2, which contains 14 more parameters, this is not a significant reduction (the p-value for the test is 0.796). We therefore conclude that the interactions offer a substantial simplification of the trend structure, at little cost. Residual plots for the new model are shown in Figure 3.12. There is little visible difference between this and Figure 3.11. In practice, therefore, this model represents a good compromise between complexity and performance.

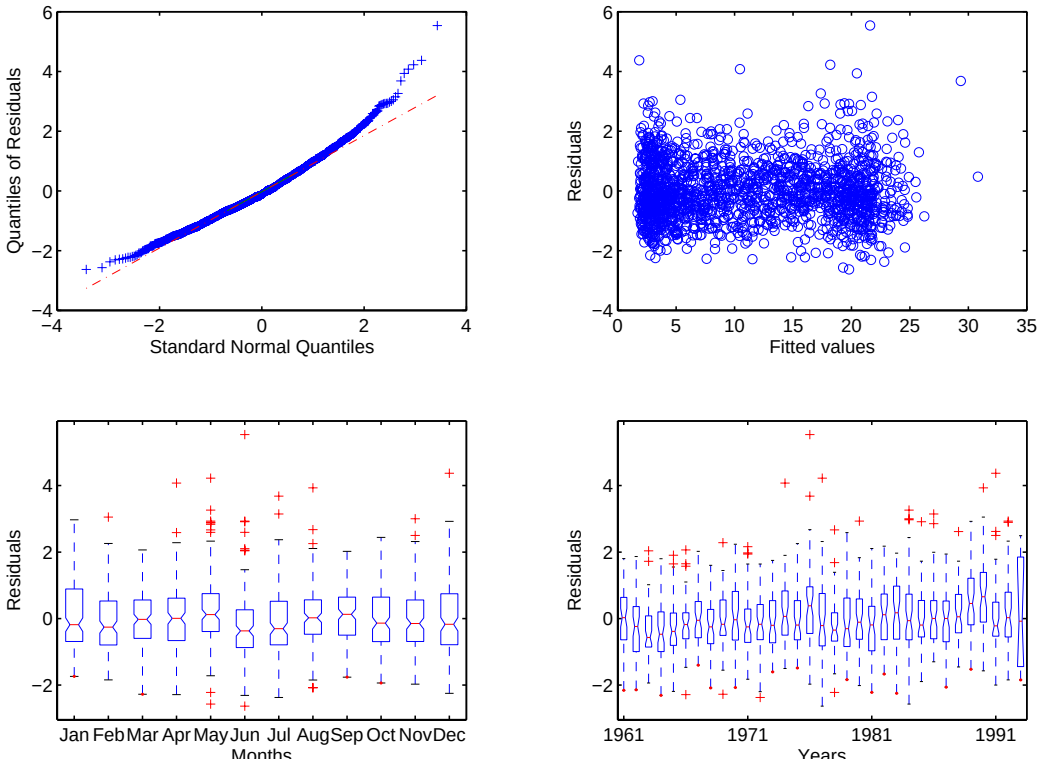


Figure 3.6 Residual analysis for model 1 in Table 3.1. Clockwise from top left: normal quantile-quantile plot of standardised residuals; plot of standardised residuals against fitted values; boxplots showing distributions of standardised residuals by year; boxplots showing distributions of standardised residuals by month.

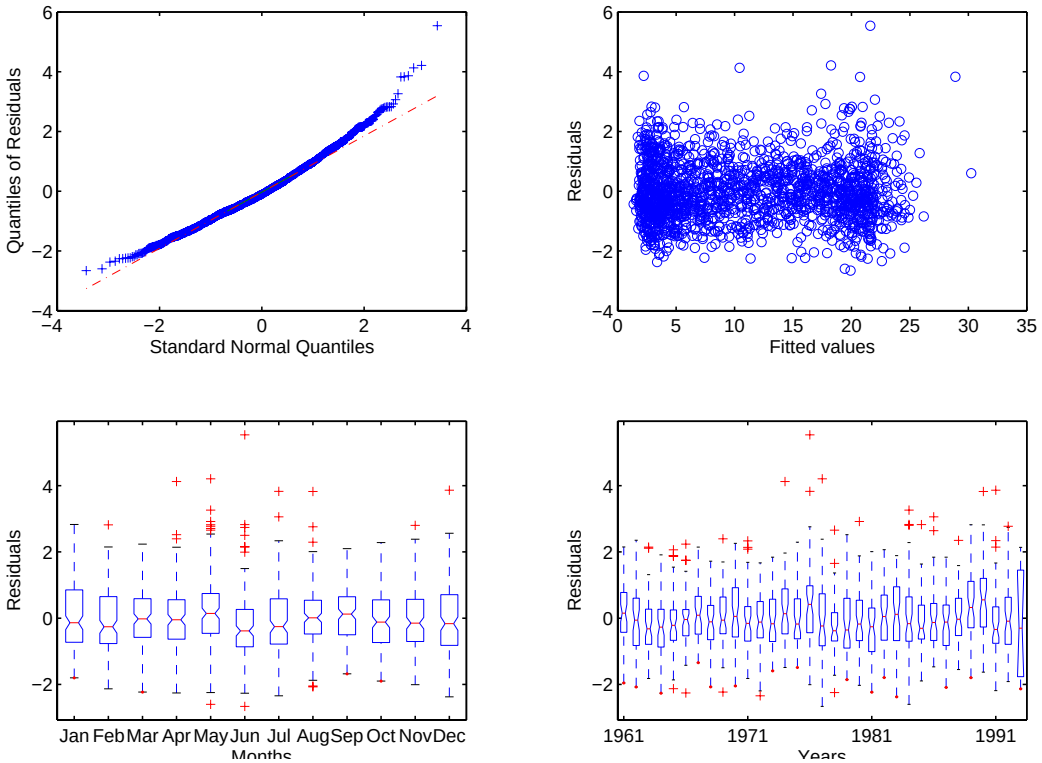


Figure 3.7 Residual analysis for model 2 in Table 3.1. See caption to Figure 3.6 for an explanation of the plots.

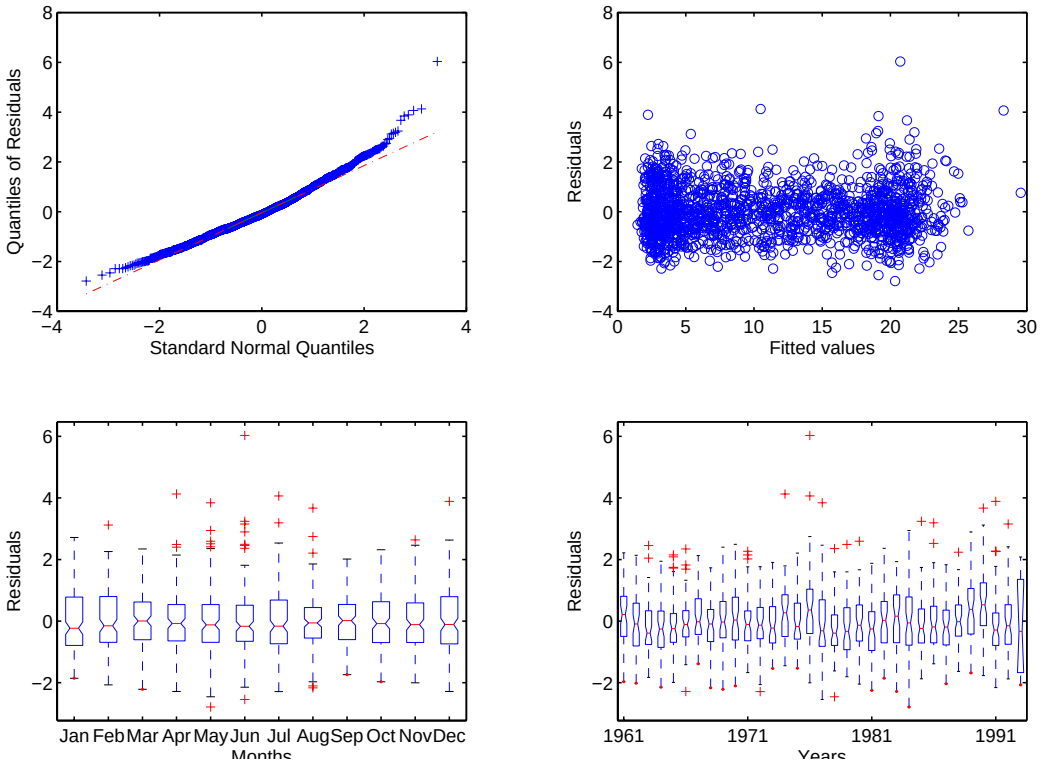


Figure 3.8 Residual analysis for model 3 in Table 3.1. See caption to Figure 3.6 for an explanation of the plots.

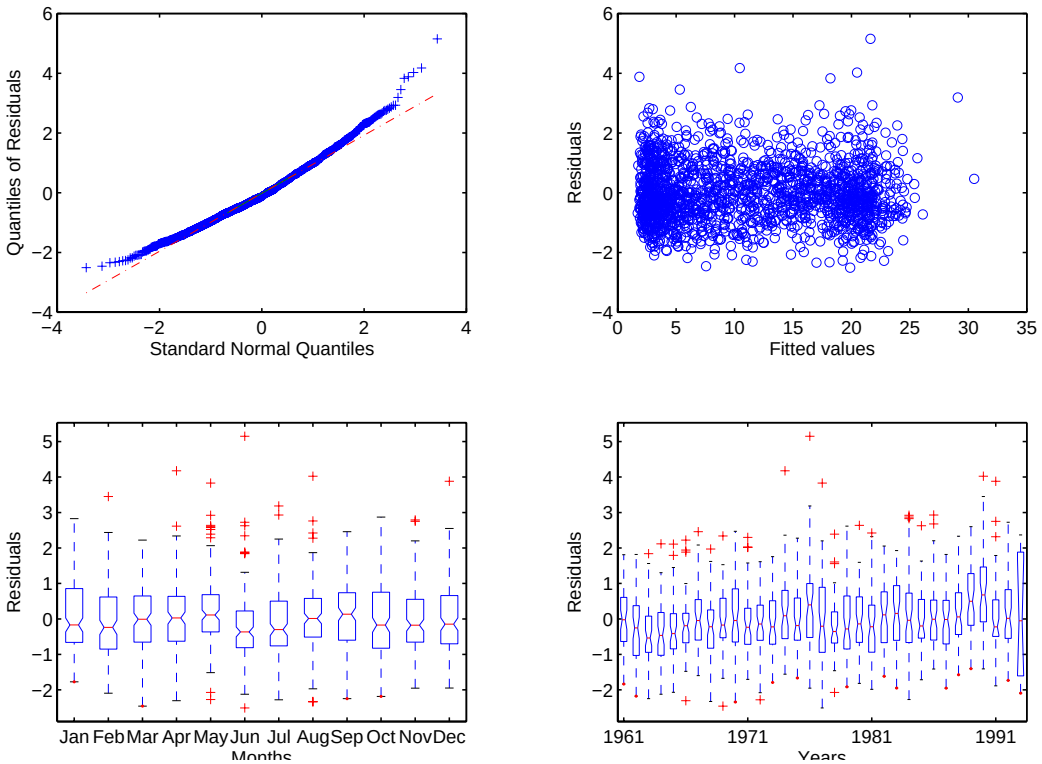


Figure 3.9 Residual analysis for model 4 in Table 3.2. See caption to Figure 3.6 for an explanation of the plots.

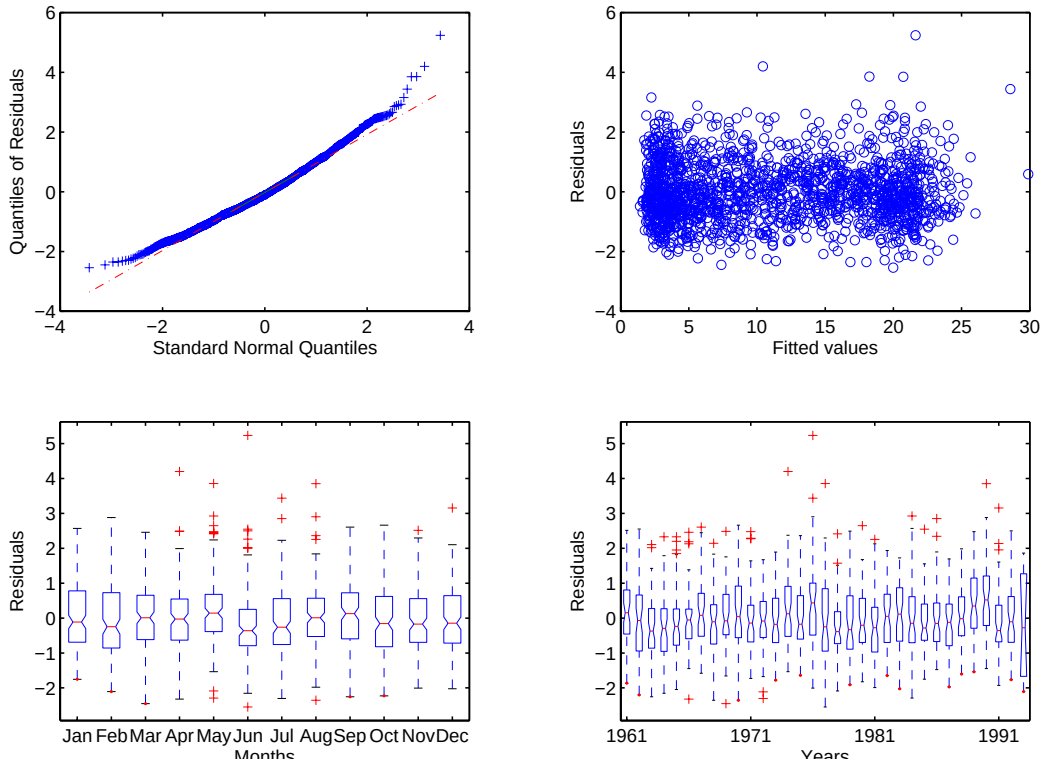


Figure 3.10 Residual analysis for model 5 in Table 3.2. See caption to Figure 3.6 for an explanation of the plots.

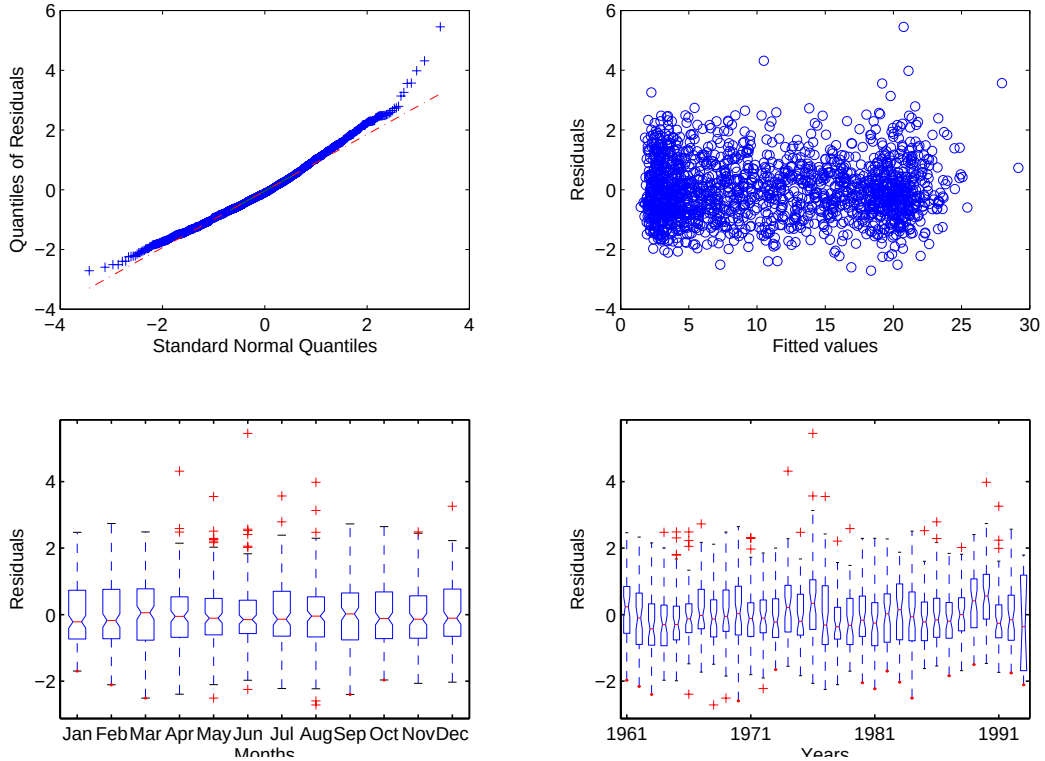


Figure 3.11 Residual analysis for model 6 in Table 3.2. See caption to Figure 3.6 for an explanation of the plots.

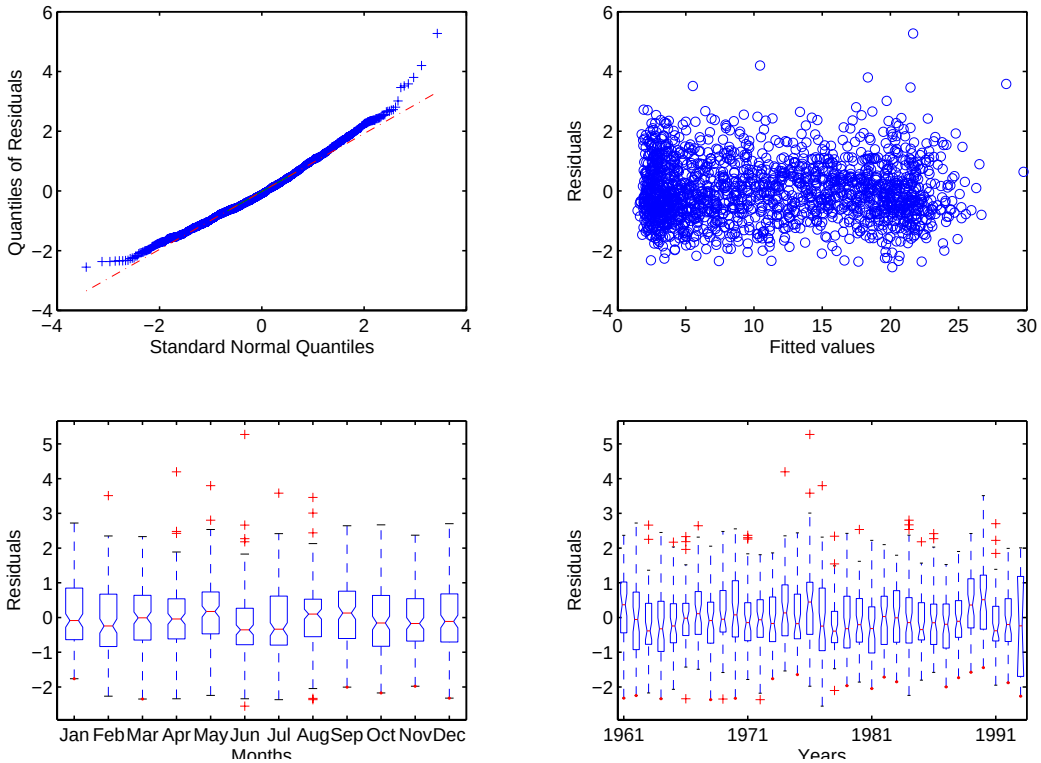


Figure 3.12 Residual analysis for final weekly evaporation model, in which trends are represented via interactions with seasonal components.

Of course, there is no guarantee that exactly the same results will hold for weekly evaporation sequences in very different locations. However, the analyses in this section provide some guidance regarding a general model-building strategy.

3.4 Daily evaporation modelling

3.4.1 Introduction

Models such as those developed in the previous section can be used to simulate potential evaporation sequences at a weekly timescale. But, since runoff models require inputs at a different timescale, in this section daily potential evaporation data will be used to develop a downscaling procedure. Of course, if daily data were widely available then they could be modelled and simulated directly in the same way as weekly data. However, in the UK at least, daily PE data are extremely scarce and are generally insufficient, for example, to identify long-term trends. We therefore propose to make use of limited daily data to study sub-weekly structure, and to use this information to downscale weekly sequences. In this way the dual objectives of downscaling MORECS data and simulating daily PE sequences can both be achieved (daily sequences can be simulated by generating a weekly sequence and then downscaling it).

3.4.2 Daily data analysis

The analysis here is similar to the weekly analysis above — the downscaling procedure developed in Section 3.4.4 below is based on a joint mean-variance model for daily data. However, the procedure also makes the assumption that daily values can be regarded as coming from a multivariate normal distribution. This assumption places some restrictions on the construction of a daily model; the practical consequences of these restrictions therefore need to be investigated.

The data used in this analysis were the 5-year daily rainfall and PE sequences from Silwood Park (see Figure 3.1). As for the weekly analysis, linear regression models were fitted first to give a preliminary indication of structure. The same sequence of models was fitted as for the weekly data in Section 3.3.2. By and large, the conclusions were the same except that this time, a highly significant relationship (with p-value less than 2×10^{-16}) was found between daily rainfall and daily PE (the previous day's rainfall was not significant, however). The resulting linear model contained 7 terms (a constant, 4 terms representing seasonality as in (3.7), rainfall and the previous day's PE). This model yields a root mean squared error of 0.59mm and explains 72.2% of the variance in the daily PE sequence.

The inclusion of rainfall in a daily PE model presents some problems for the downscaling procedure developed below. Thus, although it is statistically significant, it is of interest to determine its real contribution to the model. This can be done by examining the effect of dropping it from the model. In practical terms, this effect is very minor — the root mean squared error rises to 0.61 and the percentage of explained variance drops to 70.4. The need to exclude rainfall for downscaling purposes is therefore not a serious restriction, and it has not been considered further in the daily modelling.

3.4.3 Joint model fitting

As with the weekly analysis, there is a clear seasonal cycle in the variance of daily potential evaporation so that a joint mean-variance model is required. However, with only about 5 years of data, it is not possible to identify any long-term trends, so that there are fewer modelling options in this case — after excluding rainfall from the analysis, the only covariates to be considered are the seasonal terms (3.7), along with previous days' values. Table 3.4 shows three different joint mean-variance models that have been fitted; they differ in their specification of the variance structure. The performance of each model is summarised in Table 3.5, and residual analyses are shown in Figures 3.13–3.15.

Table 3.5 shows that in terms of log-likelihood, model 2 has the best performance and model 3 the worst. However, as far as the other measures are concerned there is very little to distinguish between the models. The same is true of the residual plots, which show reasonable performance for all of the models. There are again a few outliers in the upper tail of the quantile-quantile plot; the boxplots show that these occurred primarily in the winter months, which probably explains why they are associated with small fitted values in the 'residuals vs fitted' plots.

Of the three models presented here, only model 3 implies a multivariate normal

Table 3.4 Coefficient estimates in joint models for daily PE sequence. Model 1: σ_t linearly related to μ_t. Model 2: σ_t linearly related to each of the covariates in the mean model. Model 3: as model 2 but without the previous day's value in the model for σ_t. All models are fitted to 1882 daily observations.

Predictors in model	Model 1		Model 2		Model 3	
	μ_t	σ_t	μ_t	σ_t	μ_t	σ_t
Constant	0.8009	0.0786	0.7916	0.4126	0.7830	0.5237
$\cos\left(\frac{2\pi \times \text{day of year}}{366}\right)$	-0.7692		-0.7606	-0.3140	-0.7516	-0.4126
$\sin\left(\frac{2\pi \times \text{day of year}}{366}\right)$	0.0954		0.0942	0.0619	0.0986	0.0741
$\cos\left(\frac{4\pi \times \text{day of year}}{366}\right)$	0.0986		0.1029	0.0342	0.1076	0.0544
$\sin\left(\frac{4\pi \times \text{day of year}}{366}\right)$	0.0040		0.0048	-0.0080	0.0037	-0.0069
Lag 1	0.3914		0.3981	0.0871	0.4040	
Conditional mean		0.3462				

distribution for collections of daily PE values. The reason is the presence of the previous day's value in the calculation of σ_t in the other two models (in model 1, this is implicit rather than explicit, since σ_t depends on μ_t here, and μ_t depends on the previous day's value). Although the log-likelihoods indicate that this model is significantly worse than either of the other two, the other analyses indicate that this statistical significance is of little practical relevance. We therefore adopt this model, and will use it for downscaling.

3.4.4 Downscaling procedure

The analysis of the previous section indicates that, for practical purposes, any collection of daily PE values can be regarded as being generated from a multivariate normal distribution according to model 3 in Table 3.4. We now develop a methodology that uses properties of this distribution to downscale a given weekly PE sequence to a daily timescale.

The procedure makes use of the following standard result (e.g. Krzanowski 1988, Section 7.2). Suppose a random vector $\mathbf{Y}$ has a multivariate normal distribution with mean vector μ and covariance matrix Σ. Suppose also that $\mathbf{Y}$ can be partitioned into two subvectors $\mathbf{Y}_1$ and $\mathbf{Y}_2$. Then the conditional distribution of $\mathbf{Y}_1$, given $\mathbf{Y}_2 = \mathbf{y}_2$, is also multivariate normal. Specifically,

$$\mathbf{Y}_1|\mathbf{Y}_2 = \mathbf{y}_2 \sim MVN\left(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}\left(\mathbf{y}_2 - \mu_2\right), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}\right) , \tag{3.10}$$

where μ_1 and μ_2 are the unconditional mean vectors of $\mathbf{Y}_1$ and $\mathbf{Y}_2$, Σ_{11} and Σ_{22} are their unconditional covariance matrices and Σ_{12} is the cross-covariance matrix between them. μ_1 and μ_2 can be extracted directly from the relevant elements of μ;

Table 3.5 Performance measures for daily PE models (see Table 3.3 for details of calculations).

Model number	R^2 for model components		Total number of parameters	Log-likelihood
	Mean	Variance		
1	0.7037	0.1920	8	-1121.72
2	0.7038	0.2147	12	-1113.51
3	0.7037	0.2168	11	-1132.72

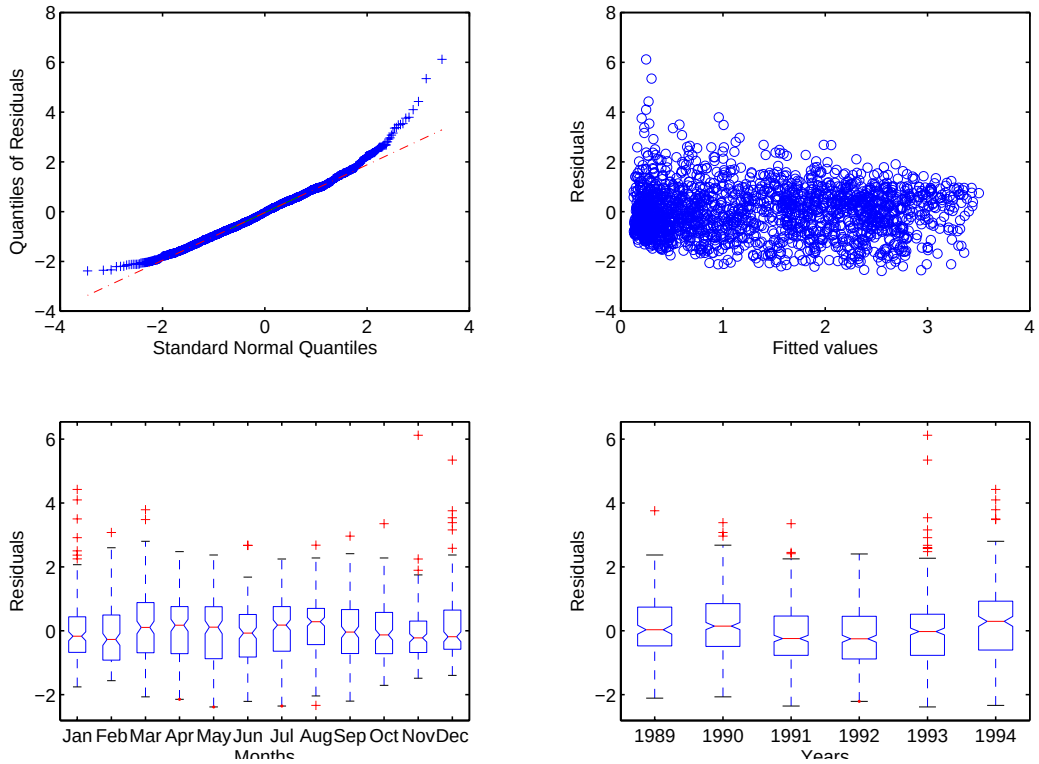


Figure 3.13 Residual analysis for model 1 in Table 3.4. See caption to Figure 3.6 for an explanation of the plots.

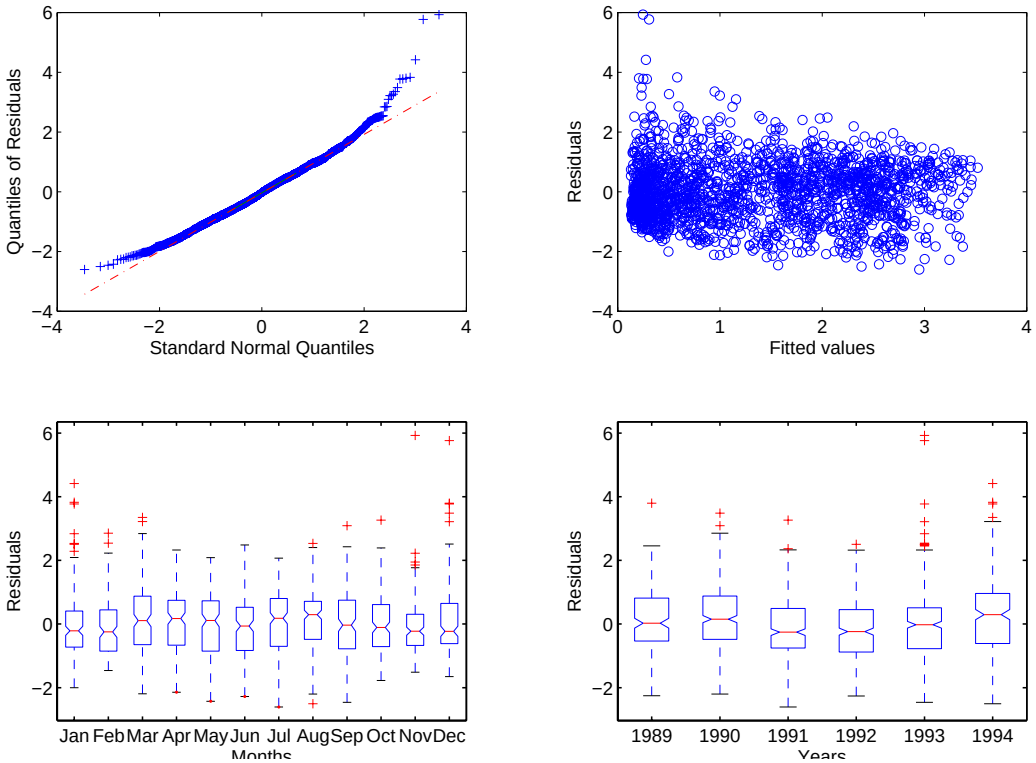


Figure 3.14 Residual analysis for model 2 in Table 3.4. See caption to Figure 3.6 for an explanation of the plots.

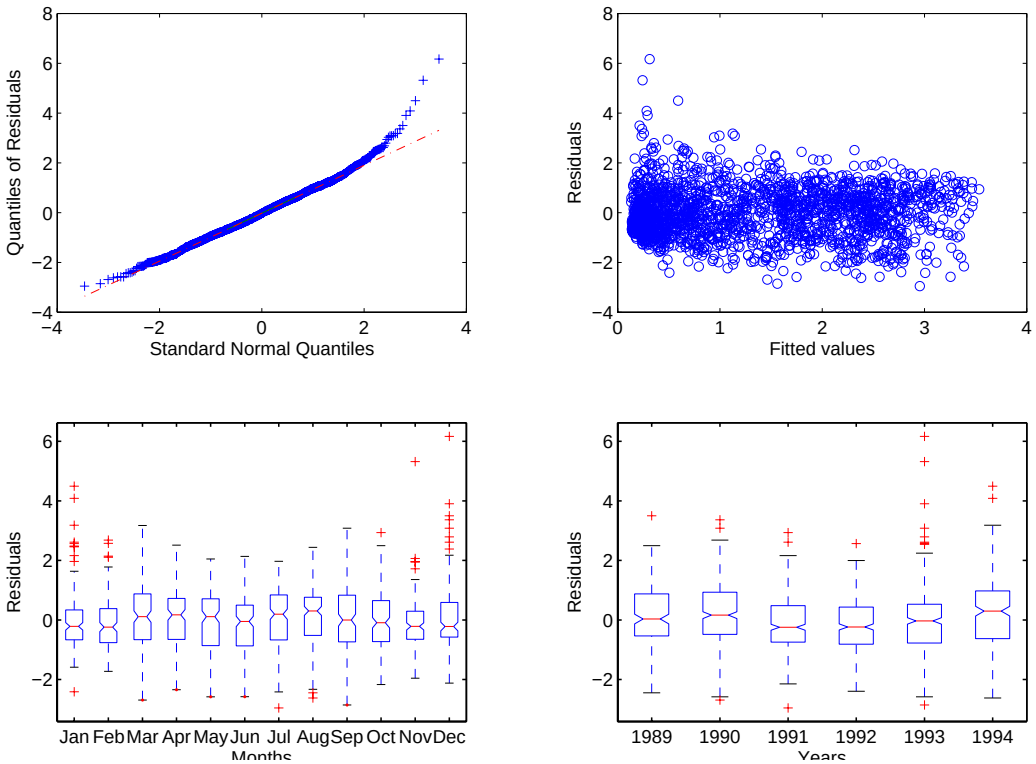


Figure 3.15 Residual analysis for model 3 in Table 3.4. See caption to Figure 3.6 for an explanation of the plots.

Σ_{11}, Σ_{22} and Σ_{12} can be extracted directly from Σ. Therefore, once the mean and the covariance matrix of $\mathbf{Y}$ are known, the conditional distribution (3.10) can be calculated.

The aim of the current exercise is to construct the distribution of seven daily PE values (one for each day of the week), given a sequence of weekly totals (and possibly daily values for previous weeks that have already been generated by downscaling). In the notation of equation (3.10), $\mathbf{Y}_1$ represents the daily values for the current week and $\mathbf{y}_2$ contains the information that will be used to predict $\mathbf{Y}_1$:

1. The last daily value from the previous week, since the daily PE model above involves a first-order autoregressive term;
2. The total evaporation for the current week; and
3. The total evaporation for the next week, since the autoregressive structure of the daily model implies a correlation between the last day of the current week and the first day of the next week.

To calculate the required conditional distribution, let Y_t denote the PE value on day t and let Y_w^+ denote the PE total for week w. Define the vector

$$\mathbf{Y} = \begin{pmatrix} Y_{7w-6} & Y_{7w-5} & \dots & Y_{7w} & Y_{7(w-1)} & Y_w^+ & Y_{w+1}^+ \end{pmatrix}' . \tag{3.11}$$

The first 7 elements are the daily values in week w; the next element is the final day's value from week $w-1$, and the remaining 2 elements are the weekly totals for weeks w and $w+1$. When downscaling for week w, the last 3 elements of $\mathbf{Y}$ will be known (except at the start of a sequence i.e. when $w=1$, in which case $Y_{7(w-1)}$ will generally be unknown and should be excluded from the definition of $\mathbf{Y}$). Based on this information, we wish to calculate the conditional distribution of the subvector

$$\mathbf{Y}_w = \begin{pmatrix} Y_{7w-6} & Y_{7w-5} & \dots & Y_{7w} \end{pmatrix}' . \tag{3.12}$$

Under daily model 3 from the previous section, $\mathbf{Y}$ has a multivariate normal distribution. Hence the conditional distribution of $\mathbf{Y}_w$ can be derived from the mean μ and covariance matrix Σ of $\mathbf{Y}$. Now $\mathbf{Y}$ can be written as

$$\mathbf{Y} = \mathbf{A}\tilde{\mathbf{Y}} , \tag{3.13}$$

where $\tilde{\mathbf{Y}} = \begin{pmatrix} Y_{7(w-1)} & Y_{7w-6} & Y_{7w-5} & \dots & Y_{7(w+1)} \end{pmatrix}'$ is a vector of 15 daily values and $\mathbf{A}$ is the matrix

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} . \tag{3.14}$$

Standard multivariate results then yield

$$\mu = \mathbf{A}\tilde{\mu} \quad \text{and} \quad \Sigma = \mathbf{A}\tilde{\Sigma}\mathbf{A}' \,, \tag{3.15}$$

where $\tilde{\mu}$ and $\tilde{\Sigma}$ are the mean and covariance matrix of $\tilde{\mathbf{Y}}$. These properties of the daily PE sequence can be derived for model 3 of the previous section. It is convenient to reparameterise the model at this point, by combining the sine and cosine terms for each seasonal component into a single cosine term with a phase shift. For example, the mean component of the model can be written as

$$\beta_0 + \sum_{j=1}^{2} \beta_j^* \cos\left[2\pi\left(\omega_j t - \phi_j\right)\right] + \beta_3 y_{t-1} \,, \qquad \text{where } \omega_j = j/366 \tag{3.16}$$

and (from Table 3.4) $\beta_0 = 0.783$, $\beta_1^* = \sqrt{(-0.752)^2 + 0.099^2} = 0.758$, $\beta_2^* = \sqrt{0.108^2 + 0.004^2} = 0.108$, $\phi_1 = (2\pi)^{-1}\arctan(-0.099/0.752) = 0.479$, $\phi_2 = (2\pi)^{-1}\arctan(-0.004/0.108) = 0.005$ and $\beta_3 = 0.404$ (the ϕ calculations here assume that $t = 1$ falls on January 1st). The standard deviation can be rewritten similarly as

$$\sigma_t = \gamma_0 + \sum_{j=1}^{2} \gamma_j \cos\left[2\pi\left(\omega_j t - \psi_j\right)\right] \,, \tag{3.17}$$

where the frequencies ω_1 and ω_2 are the same as those in (3.16). In this new parameterisation, the expected value of Y_t is

$$\mathsf{E}(Y_t) = \frac{\beta_0}{1-\beta_3} + \sum_{j=1}^{2} \beta_j^* \left\{ \frac{\cos\left[2\pi\left(\omega_j t - \phi_j\right)\right] - \beta_3 \cos\left[2\pi\left(\omega_j (t+1) - \phi_j\right)\right]}{1 - 2\beta_3 \cos\left(2\pi\omega_j\right) + \beta_3^2} \right\} \,, \tag{3.18}$$

and the covariance between Y_t and Y_{t-k} is

$$\mathsf{Cov}(Y_t, Y_{t-k}) = \beta_3^k \sum_{p=0}^{6} \gamma_p^* \left\{ \frac{\cos\left[2\pi\left(\omega_p^*(t-k) - \psi_p^*\right)\right] - \beta_3^2 \cos\left[2\pi\left(\omega_p^*(t-k+1) - \psi_p^*\right)\right]}{1 - 2\beta_3^2 \cos\left(2\pi\omega_p^*\right) + \beta_3^4} \right\} \,, \tag{3.19}$$

where

$$\begin{gathered} \omega_0^* = \psi_0^* = 0 \qquad\qquad \gamma_0^* = \gamma_0^2 + \left(\gamma_1^2 + \gamma_2^2\right)/2 \\ \gamma_1^* = 2\gamma_0\gamma_1,\ \omega_1^* = \omega_1,\ \psi_1^* = \psi_1 \qquad \gamma_2^* = 2\gamma_0\gamma_2,\ \omega_2^* = \omega_2,\ \psi_2^* = \psi_2 \\ \gamma_3^* = \gamma_1^2/2,\ \omega_3^* = 2\omega_1,\ \psi_3^* = 2\psi_1 \qquad \gamma_4^* = \gamma_2^2/2,\ \omega_4^* = 2\omega_2,\ \psi_4^* = 2\psi_2 \\ \gamma_5^* = \gamma_1\gamma_2,\ \omega_5^* = \omega_1 + \omega_2,\ \psi_5^* = \psi_1 + \psi_2 \quad \text{and} \quad \gamma_6^* = \gamma_1\gamma_2,\ \omega_6^* = \omega_1 - \omega_2,\ \psi_6^* = \psi_1 - \psi_2 \,. \end{gathered} \tag{3.20}$$

A derivation of these results is given in Appendix D. Note that the variance of Y_t is obtained by putting $k = 0$ in (3.19).

Combining all of these results: in a straightforward series of computations, (3.18) and (3.19) can be used to calculate the mean and covariance matrix of $\tilde{\mathbf{Y}}$, while (3.15)determines the mean and covariance matrix of $\mathbf{Y}$. Hence the required conditional distribution of $\mathbf{Y}_w$ can be derived using result (3.10). Standard algorithms are available for simulating from such a distribution.

The downscaling scheme has been tested using the daily Silwood Park data from Section 3.4.2. The data were first aggregated to a weekly series, which was then

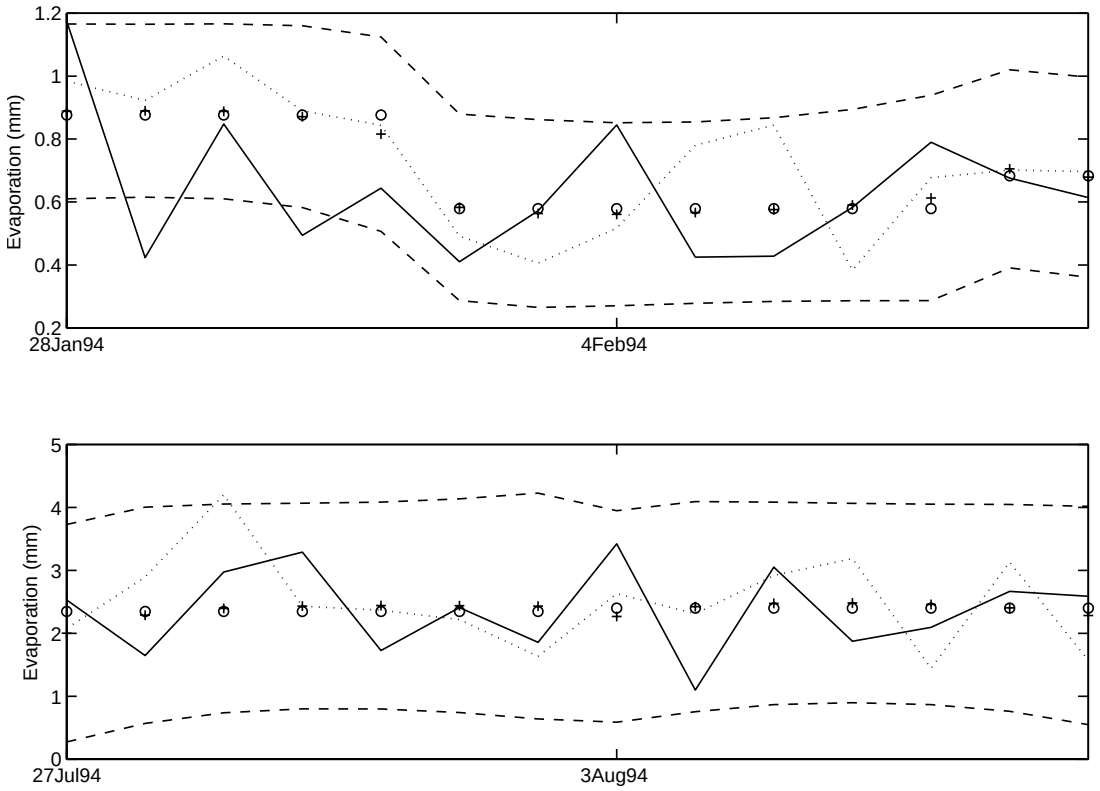


Figure 3.16 Illustration of downscaling scheme performance for two 2-week periods (top: summer, bottom: winter). Dots represent weekly averages, computed from the observed daily sequences (solid lines). The dotted line in each plot represents a single simulated realisation. Crosses represent the mean of the conditional distribution of daily values, and dashed lines are 95% confidence bands obtained from this conditional distribution.

downscaled without reference to the original data. Figure 3.16 shows specimen results for two short periods. The downscaled sequences here represent single simulations — the aim is not to reproduce the observations exactly, but rather to provide sequences that have a realistic level of sub-weekly variability. In these examples, the calculated conditional distributions for each week have almost constant means — which provides some justification for the common practice (see Section 3.1) of assuming evaporation itself to be constant. However, this common practice ignores daily variability and, from this perspective, the simulated sequences in Figure 3.16 are much more realistic. The confidence bands show good agreement between the observed variability and that expected under the downscaling scheme. Note also that the scheme effectively captures seasonal differences in this variability.

These results indicate that the proposed scheme is both viable and successful. It can therefore be used either to downscale historical PE data to a daily timescale or, in conjunction with a weekly simulation model such as that developed in Section 3.3, to simulate daily sequences.

3.5 Conclusions

The aim of this work was to provide simulated sequences of potential evaporation at a daily timescale, to use as inputs in runoff models. The study was divided in three steps:

1. Modelling weekly potential evaporation time series given rainfall;
2. Modelling daily potential evaporation time series given rainfall; and
3. Downscaling weekly potential evaporation to daily values.

Using MORECS data, the modelling of weekly PE was conducted using joint mean-variance models. Both seasonality and linear trends were represented successfully by the joint models. Several model choices were discussed for the dataset studied here. In general, there is a tradeoff between model performance and model complexity. Users can select from these models based on their own needs. Although the methodology has been illustrated using a single dataset, it is completely general and is expected to apply (with suitable modifications to the exact models used) at a wide variety of locations.

A particularly interesting feature to emerge from this work is the existence of significant increasing trends in the studied MORECS evaporation sequence. The magnitude of these trends varies with season. In relative terms, their effect is substantial in winter due to the typically low values of PE during the winter months. In January, for example, the trend corresponds to an increase of around 30% in mean weekly evaporation over the 33-year record; this could have significant hydrological implications. In July however, the increase is only around 3%.

The primary source of potential evaporation data in the UK at present is the MORECS system, which provides weekly values. However, runoff models require inputs at a finer timescale; in current practice, this is often dealt with rather crudely by ignoring subweekly variability. For this reason a new methodology to downscaling weekly to daily values was developed in Section 3.4.4. This procedure uses the theoretical properties of a joint mean-variance model fitted to limited daily data from a single site, to define a multivariate normal distribution for daily potential evaporation given weekly evaporation totals. By sampling from this distribution, it is possible to derive daily potential evaporation values as required by runoff models. The new methodology yields simulated sequences that incorporate a realistic level of sub-weekly variability. It can be used either to downscale historical data or, in conjunction with a weekly simulation model, to generate simulated sequences at a daily timescale. In such applications, the most critical feature of the daily model is the first-order autoregressive coefficient (β_3 in (3.16)), since this essentially determines the smoothness of realisations within a week. The autocorrelation structure of PE sequences is unlikely to vary substantially in space; therefore this model is expected to perform well at other locations, despite being calibrated using a relatively short data sequence. Further work investigating the possibility of regionalisation of this model is therefore recommended, along with comparison of other newly developed meteorologically based techniques (Kilsby et al., 2006).

Chapter 4 Catchment rainfall simulation using the single-site model

4.1 Introduction

This chapter focuses on assessing the performance of the single-site rainfall model described in Chapter 2 in simulating rainfall time series for a number of catchments in Great Britain. The results of the joint testing of the single-site rainfall model and a continuous river flow simulation model developed under Defra/Environment Agency Project FD2106 *National river catchment flood frequency method using continuous simulation* (CEH Wallingford, 2005) are described in Chapter 5.

The analysis described in Chapter 2 of this report identified a single-site rainfall model for routine national use and recommended the Random Parameter Bartlett-Lewis (RPBL) model with exponential cell depth distribution. A two-stage fitting procedure was recommended to ensure reproduction of both average and extremal rainfall properties:

- **Stage 1**: Estimation of near-optimal parameter sets. For each month, several sets of parameters are identified that provide good reproduction of 'average' rainfall properties;
- **Stage 2**: Simulation of rainfall hourly time series. Simulations are used to determine extremal properties for each parameter set identified in Stage 1; the parameter set that provides the best reproduction of extremal properties is chosen and defines the final parameter set for that month.

A set of software tools has been developed to allow the model to be applied in practice.

The aim of the analysis described here was to assess the performance of the RPBL model of single-site rainfall using some of data series compiled during Project FD2106. A review of the individual hourly raingauge records available indicated that they were generally too short and were often affected by missing data. Moreover, the lumped conceptual rainfall-runoff models used within Project FD2106 required catchment average hourly rainfall depths as input. Therefore, the decision was made to test the RPBL single-site rainfall model by fitting to catchment rainfall series for a number of the smallest catchments within the dataset.

4.2 Case study catchments

Initially, eight catchments were selected from the total of 39 catchments with hourly data included in the FD2106 analysis. The catchments were chosen to be representative of a range of rainfall and flood regimes in Great Britain, and at least 15 years of hourly catchment average rainfall data were available in each case. An additional criterion for catchment selection was that the area should not be too large. The locations of the eight catchments are shown in Figure 4.1 and details are given in Table 4.1.

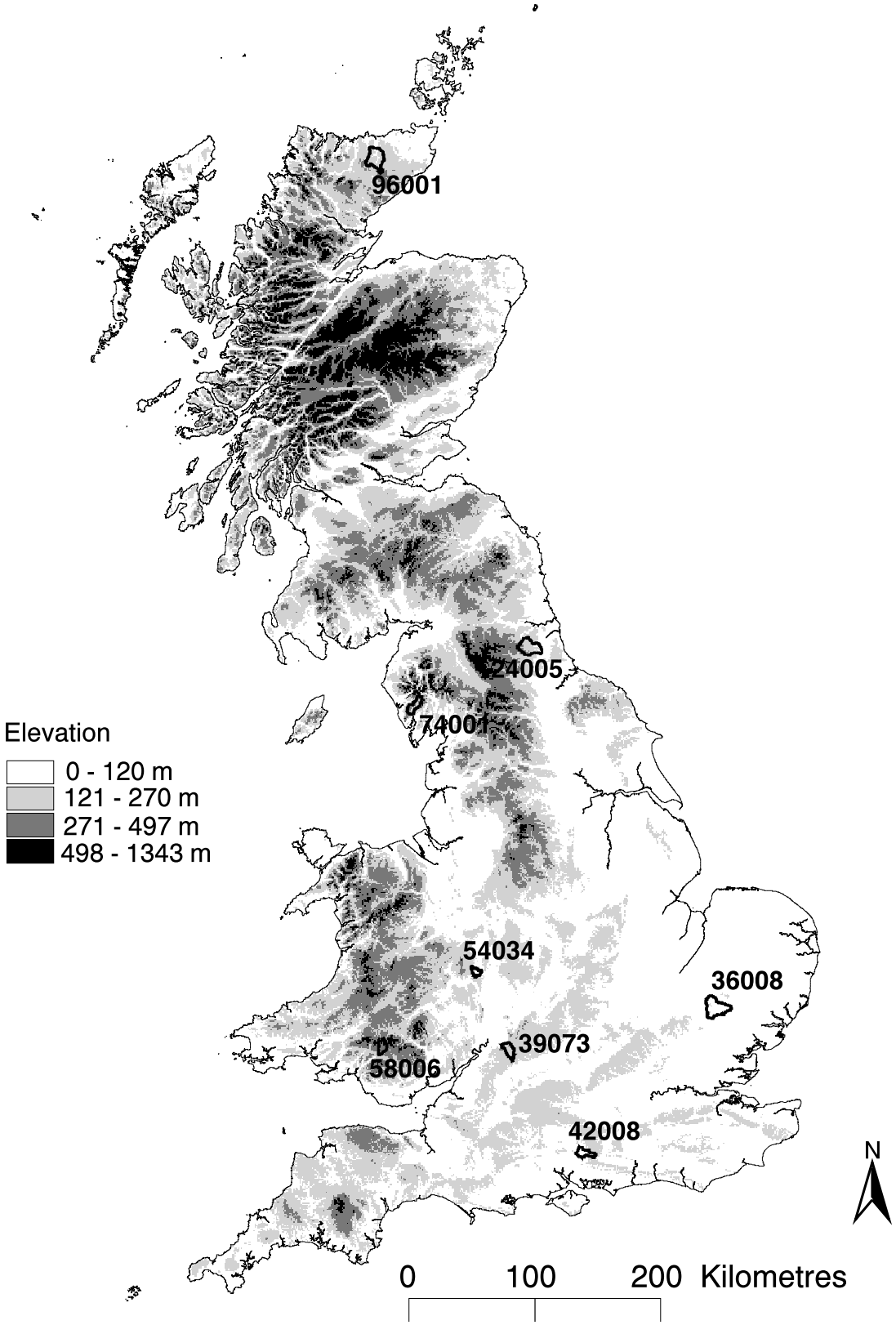


Figure 4.1 Locations of case study catchments.

Table 4.1 Selected catchments and data availability.

No.	River @ site	County	Area km^2	Record	No. Gauges	
					Recording	Daily
24005	Browney @ Burn Hall	Durham	178	1982-2001	2	15
36008	Stour @ Westmill	Suffolk	224	1985-2001	1	19
39073	Churn @ Cirencester	Gloucestershire	84	1987-2001	4	44
42008	Cheriton Stream @ Sewards Bridge	Hampshire	75	1985-2001	1	8
54034	Dowles Brook @ Oak Cottage	Herefordshire	41	1985-2001	1	3
58006	Mellte @ Pontneddfechan	Glamorgan	66	1985-2001	1	4
74001	Duddon @ Duddon Hall	Cumberland	86	1985-2001	1	8
96001	Halladale @ Halladale	Caithness	205	1985-2001	1	2

The hourly catchment average rainfall series were produced using a two-step procedure. First, daily raingauge data within an area encompassing the catchment were averaged using the triangle method (Jones, 1983) to provide a catchment average daily rainfall (CADR) series. The daily catchment average totals were then distributed using the temporal profiles available from the recording raingauge data for days with acceptable data quality. Where data from more than one recording raingauge were available, weights were allocated to each gauge and the profile of weighted sum of the hourly raingauge data was used to distribute the CADR. For days without good quality recording raingauge data, an empirically-derived 24 hr profile based on the average variability method (Pilgrim et al., 1969) was used to disaggregate the daily totals (Calver et al., 2001). The method was designed to compensate, as much as possible, for the effects of topography and the irregular distribution of raingauges, and to take advantage of the denser and more complete network of daily raingauges across the UK (Crooks et al., 2002). The use of this method is therefore justified by the lack of reliable data although its shortcomings should be flagged. The main one is that it alters the variance of the constructed series. This follows from the fact that it applies a single-site profile to a catchment average which ought to have lower variance as a result of spatial averaging.

4.3 Methodology

4.3.1 Rainfall simulation

Hourly catchment average rainfall series were simulated for each catchment following the procedure described in 2 and summarised below:

- For each month, statistical properties of the rainfall series are calculated: mean hourly rainfall intensity; variance of 1, 6, 12 and 24 hr rainfall; lag 1 autocorrelation (1 and 24 hr); proportion of dry intervals (1 and 24 hr).
- Sets of monthly 'near-optimal' parameters that closely reproduce the statistical properties are defined by a generalised method of moments. Because some of these sets may generate unrealistic rain cells (e.g. storm arrival rates too low, or storms with too many cells) which will be discarded in the next step, it is

recommended that a large number of parameter sets (over 80) should be generated.

- For each month and each near-optimal parameter set, rainfall series are simulated and the statistical properties of the extremes are calculated. The final parameter set is the one which provides the best reproduction of 1 hr extreme statistics for that month.
- Long hourly rainfall series are generated using the final parameter set for each month.

Because of the large amount of computing time required for these procedures and the very large number of near-optimal parameter sets to consider (at least 80 sets per month per catchment), it was not feasible to review the graphical outputs available when running the routines (e.g. the shape of the objective functions). The fitting and simulation procedures were automated and the output files of the first stage were used as input to the second stage routines without any further expert selection. Although this automated process filters out many unrealistic parameter sets, this is not a 100% efficient filter. This follows from the fact that there are no hard criteria defining exhaustively the domain of 'realistic' parameters.

4.3.2 Analysis

Two hundred simulations were produced from the final parameter set for each catchment. To avoid bias, the simulations were of the same length as the observed data for each catchment. The performance of the stochastic rainfall model was assessed in a number of ways. The fitted model parameter sets for each month and each catchment were examined to determine whether they reflected the characteristics of observed rainfall series. In addition, graphical comparisons of a number of descriptors of simulated and observed data series were made, and individual outliers were examined.

4.4 Initial results

4.4.1 Fitted parameters of the Random Parameter Bartlett-Lewis (RPBL) model

The parameters of the RPBL are:

λ	Storm arrival rate	hr^{-1}
μ_x	Mean cell depth	mm hr^{-1}
α	Shape parameter of gamma distribution for η	
ν	Scale parameter of gamma distribution for η	hr
κ	$1/\eta \times \beta$ Cell arrival parameter	
ϕ	$1/\eta \times \gamma$ Storm duration parameter.	

In addition, some secondary variables can be derived from the RPBL parameters, that have a more direct physical interpretation:

Expected cell duration (hr)	$\nu/(\alpha-1)$
Expected cell inter-arrival time (hr)	$(\nu/(\alpha-1))/\kappa$
Expected storm duration (hr)	$(\nu/(\alpha-1))/\phi$

where the expectations are here taken over all storms.

A high value of μ_x indicates heavy rainfall, often associated with convective rainfall in Britain. A small value of ϕ represents very short rainfall episodes compared to the whole storm, perhaps typical of frontal rainfall; ϕ close to 1 indicates storms characterised by few rainfall episodes, with few dry episodes in one given storm, often associated with convective storms. A value of κ lower than 1 indicates long dry spells within the storms, interrupted by short rainfall episodes, while when greater than 1, dry episodes are short compared to rainfall events.

For each catchment, a set of parameters was derived for each month and these are given in Table 4.2 together with values of the secondary variables defined above. It can be seen that some of the fitted parameters have values that are physically unrealistic (in bold), for example very large values of μ_x (mean cell depth) or very large or small values of ϕ, resulting in unrealistic expected storm durations. All eight catchments were found to have some months for which at least one of the fitted parameter values was too large or too small. No particular seasonal or geographical patterns in the values were discernible.

The parameter sets shown in Table 4.2 were produced by an unbounded fitting procedure which was considered preferable to a restricted procedure. However, one of the main features of the model is that it represents, in a simplified stochastic form, the mechanisms of the rainfall process, and therefore it is important for values of the fitted parameters to be physically reasonable. Discussions with the project team led to the suggestion that some of the long intense rain cells could be caused by a very small value of η. η is a random variable which is constant throughout the storm and is sampled from the gamma randomisation distribution (shape α and scale ν). It is the inverse of the mean cell duration for that storm (see Chapter 2). The choice of η also determines the mean storm duration, as small values of η will lead to many long storms being simulated. As a result, heavy (long and intense) storms may be produced in the simulations (see Chapter 2). A new version of the evaluation and simulation software was therefore developed. This *restricted eta* version applies a threshold of 0.1 for acceptable values of η.

Because of the long computing time required for the evaluation and simulation and the late stage at which the new version was available, time did not permit a full reanalysis of all eight catchments. Therefore, three catchments were chosen (24005, 42008 and 74001). Details of the initial analysis using the original *unrestricted eta* model for all eight catchments are given in Prudhomme and Stewart (2005).

Table 4.2 RPBL model parameters (fitted using unrestricted eta model)

Catchment	Month	λ (hr^{-1})	μ_x (mm hr^{-1})	α	ν (hr)	κ	ϕ	Expected cell dur (hr)	Expected cell inter-arrival dur (hr)	Expected storm dur (hr)
	1	0.10	0.96	2.12	0.64	0.05	0.12	0.57	11.99	4.77
	2	0.07	1.22	2.05	0.21	0.17	0.06	0.20	1.13	3.56
	3	0.05	1.38	1.85	0.29	0.02	0.02	0.35	19.50	20.11
	4	0.06	1.11	2.41	1.75	0.00	0.02	1.24	**85932.50**	53.21
	5	0.05	1.64	1.99	0.64	0.00	**0.00**	0.65	**54646.90**	1342.39

Table 4.2 (continued from previous page)

Catchment	Month	λ (hr^{-1})	μ_x (mm hr^{-1})	α	ν (hr)	κ	ϕ	Expected cell dur (hr)	Expected cell inter-arrival dur (hr)	Expected storm dur (hr)
24005	6	0.06	1.95	2.29	0.70	0.00	0.15	0.54	**53258.00**	3.63
	7	0.05	4.32	1.81	0.13	0.00	**0.00**	0.16	**17925.80**	729.50
	8	0.05	2.27	2.08	0.23	0.53	0.19	0.21	0.40	1.12
	9	0.05	1.63	2.06	0.46	0.06	0.04	0.43	7.00	10.09
	10	0.09	1.39	1.85	0.45	5.51	**80.57**	0.53	0.10	0.01
	11	0.07	1.01	1.90	0.30	0.67	0.21	0.34	0.50	1.64
	12	0.09	0.70	1.98	0.66	5.13	**4.88**	0.67	0.13	0.14
	1	0.07	1.06	3.01	1.87	0.00	0.00	0.93	**520277.00**	**1609.15**
	2	0.07	**761.70**	1.85	0.00	0.00	**0.00**	**0.00**	0.01	0.33
	3	0.06	1.59	2.16	0.49	0.00	0.00	0.42	**149475.00**	357.77
	4	0.05	1.30	2.49	1.16	0.00	**0.00**	0.78	**110174.00**	**5849.76**
	5	0.05	3.70	1.62	0.07	0.00	**169.49**	0.11	**953.23**	**0.00**
36008	6	0.04	3.61	2.20	0.36	0.00	**148.92**	0.30	**105104.00**	**0.00**
	7	0.04	4.45	2.18	0.29	0.00	0.01	0.25	**114297.00**	28.11
	8	0.03	3.56	2.33	0.16	0.23	0.05	0.12	0.53	2.30
	9	0.05	3.42	1.93	0.18	0.23	0.30	0.19	0.86	0.65
	10	0.04	2.37	2.02	0.52	0.01	0.04	0.50	52.22	13.02
	11	0.07	1.44	2.46	0.92	0.00	**102.24**	0.63	**139229.00**	0.01
	12	0.08	1.43	2.17	0.62	0.00	0.21	0.53	192.70	2.54
	1	0.02	1.01	2.98	1.09	0.21	0.02	0.55	2.69	26.94
	2	0.06	1.29	2.56	1.65	0.00	**105.91**	1.05	**236521.90**	**0.01**
	3	0.05	1.11	2.27	1.40	0.00	0.04	1.10	**9455.16**	27.45
	4	0.03	0.81	4.68	4.91	0.27	0.09	1.33	4.92	14.47
	5	0.03	3.09	1.90	0.30	0.05	0.09	0.33	6.76	3.93
39073	6	0.03	2.67	1.96	0.40	0.00	**50.44**	0.42	**18289.31**	0.01
	7	0.03	2.97	2.96	1.22	0.00	0.00	0.63	**31773.66**	**1044.81**
	8	0.04	3.39	2.11	0.29	0.04	0.06	0.27	7.41	4.83
	9	0.04	1.86	2.29	0.09	0.73	0.03	0.07	0.09	2.04
	10	0.08	2.74	1.91	0.38	0.00	0.04	0.42	**39284.41**	10.94
	11	0.03	1.03	4.00	2.90	0.31	0.12	0.97	3.13	7.97
	12	0.07	1.48	1.90	0.78	0.00	**3.88**	0.87	**77223.97**	0.22
	1	0.02	1.27	39.06	44.50	0.09	0.03	1.17	12.61	42.71
	2	0.05	30.73	2.20	0.00	0.06	**0.00**	**0.00**	0.02	1.14
	3	0.05	2.03	1.92	0.47	0.00	**9.61**	0.51	986.63	0.05
	4	0.03	1.36	39.47	22.20	0.17	0.03	0.58	3.30	16.52
	5	0.02	1.60	43.76	15.17	0.19	0.04	0.35	1.91	9.80
42008	6	0.03	4.40	2.88	0.12	0.24	0.03	0.06	0.27	2.52
	7	0.03	4.56	1.96	0.27	0.00	0.02	0.28	231.61	17.42
	8	0.04	3.12	2.12	0.61	0.00	**0.00**	0.54	**48963.04**	**3649.01**
	9	0.03	1.90	2.82	0.71	0.26	0.07	0.39	1.51	5.40
	10	0.20	3.10	1.71	0.16	0.00	0.07	0.22	**174667.40**	3.03
	11	0.04	1.49	2.85	0.92	0.49	0.15	0.50	1.01	3.41
	12	0.07	1.81	2.19	0.70	0.21	0.26	0.59	2.82	2.27
	1	0.09	2.08	1.82	0.29	2.50E-06	**170.61**	0.36	**1.4E+05**	**0.00**
	2	0.05	16.80	2.01	0.00	0.05	0.00	**0.00**	0.05	1.65
	3	0.06	2.42	2.03	0.28	1.00E-06	**1.90E-05**	0.27	**266350.70**	**14181.76**
	4	0.06	1.26	2.10	0.61	0.20	0.28	0.56	2.77	2.02
	5	0.04	3.30	1.82	0.09	8.50	**6.99**	0.11	0.01	0.02
54034	6	0.05	2.64	2.17	0.58	3.50E-06	0.01	0.50	**1.4E+05**	42.53
	7	0.04	8.07	2.09	0.15	1.60E-06	0.23	0.14	**88109.31**	0.61
	8	0.05	3.11	2.52	0.37	0.17	0.12	0.24	1.39	1.94
	9	0.04	3.85	1.90	0.05	0.11	0.01	**0.05**	0.45	3.96
	10	0.05	7.41	1.99	0.01	0.14	0.01	**0.01**	0.08	1.79
	11	0.03	1.43	2.48	0.74	0.04	0.01	0.50	11.27	49.94
	12	0.03	**7640.32**	21.45	0.00	4.40E-05	**1.80E-07**	**2.53E-05**	0.57	142.30
	1	0.04	1.54	1240.92	1100.17	0.16	0.03	0.89	5.60	30.53
	2	0.06	0.78	9.90	51.37	0.14	**168.87**	5.77	40.04	0.03
	3	0.03	**6642.70**	2.99	5.10E-05	0.00	**3.00E-06**	**2.55E-05**	0.20	8.51
	4	0.02	**462.34**	9.90	0.01	0.00	**2.80E-05**	**0.00**	0.67	22.08
	5	0.01	**588.13**	6.0E+08	2.5E+05	0.00	**1.00E-05**	**0.00**	0.64	40.46
58006	6	0.01	1.17	15.03	8.17	0.29	0.02	0.58	2.03	35.94
	7	0.02	**19470.40**	22099.57	0.49684	2.50E-05	**1.30E-06**	**2.25E-05**	0.91	17.29
	8	0.02	**23296.70**	19.29	0.00036	2.20E-05	**1.10E-06**	**1.96E-05**	0.88	17.81
	9	0.02	1.45	3.61	1.55	0.30	0.02	0.59	1.98	34.85
	10	0.02	4.57	4.92	0.46	0.20	0.01	0.12	0.58	16.43
	11	0.02	1.52	125.71	168.65	0.15	0.02	1.35	8.75	54.96
	12	0.01	**2785.45**	20.64	0.00	0.00	**1.40E-06**	**0.00**	1.09	142.21
	1	0.04	1.96	4.73	5.27	0.05	0.03	1.41	26.56	43.41

Table 4.2 (continued from previous page)

Catchment	Month	λ (hr^{-1})	μ_x (mm hr^{-1})	α	ν (hr)	κ	ϕ	Expected cell dur (hr)	Expected cell inter-arrival dur (hr)	Expected storm dur (hr)
	2	0.11	1.91	3.09	2.63	0.00	**55.09**	1.26	448.92	0.02
	3	0.12	1.96	2.97	2.29	0.00	0.04	1.16	369.74	28.55
	4	0.08	1.57	2.25	1.30	0.05	0.07	1.05	23.04	15.35
	5	0.05	1.85	1.97	0.11	0.94	0.07	0.11	0.12	1.52
74001	6	0.03	2.07	3.65	2.65	0.06	0.03	1.00	15.94	33.74
	7	0.05	1.99	24.33	44.19	1.00E-06	0.00	1.89	**1.9E+06**	723.22
	8	0.06	2.21	5.47	6.99	0.00513	**6.03**	1.57	304.92	0.26
	9	0.13	2.35	2.16	0.82	2.40E-05	0.00	0.71	**30013.75**	**4198.35**
	10	0.21	2.29	2.13	0.79	1.00E-06	**2.70E-05**	0.69	**694052.70**	**25705.66**
	11	0.17	1.85	2.38	1.31	1.20E-06	0.00	0.95	**793285.70**	303.11
	12	0.16	2.29	2.29	1.18	8.20E-06	0.06	0.91	**111521.70**	15.55
	1	0.04	**331.79**	6.65	0.00	0.00	**2.40E-05**	**0.00**	0.67	17.63
	2	0.07	1.72	2.51	0.20	0.14	0.02	0.13	0.96	8.02
	3	0.07	0.94	2.39	0.23	0.39	0.03	0.16	0.41	4.75
	4	0.01	0.90	14.54	11.53	0.07	0.01	0.85	11.50	152.39
	5	0.04	0.96	**9.10E+12**	1.70E+13	8.00E-06	**0.00**	1.83	**227755.00**	**2699.26**
96001	6	0.10	2.30	1.90	0.25	1.50E-05	0.06	0.28	**18091.70**	4.30
	7	0.16	0.42	1.75	0.12	16.99	**2.35**	0.16	0.01	0.07
	8	0.19	2.09	2.02	0.28	1.20E-06	0.00	0.28	**232576.00**	152.26
	9	0.05	1.94	1.86	0.07	0.11	0.01	0.09	0.78	11.94
	10	0.08	14.00	1.87	0.00	0.06	0.00	0.00	0.04	2.04
	11	0.03	0.93	18.47	13.59	0.18	0.02	0.78	4.32	31.62
	12	0.04	**498.69**	3.19	0.00	0.00	**1.90E-05**	0.00	0.56	15.14

4.5 Final results

The model fitting procedure described in Section 4.3.1 was repeated for catchments 24005, 42008 and 74001 using the restricted eta version of the software to select the final parameter sets and simulate the hourly catchment rainfalls. The fitted parameter sets are shown in Table 4.3.

4.5.1 Fitted parameters and secondary variables

Storm arrival rate, λ

Values of λ are less than 0.1 hr^{-1} for catchments 24005 and 42008. Catchment 74001 has values that are slightly higher with a maximum of 0.21 hr^{-1} in October.

Mean cell intensity, μ_x

For all three catchments the value of mean cell intensity, μ_x, lies within reasonable limits. For catchments 24005 and 42008 the maximum value of μ_x occurs in July and is less than 5 mm hr^{-1}. The highest value of 13.1 mm hr^{-1} occurs in December for catchment 74001. In this case, the high value of μ_x is associated with a low value of expected cell duration (0.01 hr), indicating that the model tends to produce rain cells of relatively high intensity and relatively short duration on this catchment in December.

Table 4.3 Final RPBL model parameters for three catchments (fitted using restricted eta model).

Catchment	Month	λ (hr^{-1})	μ_x (mm hr^{-1})	α	ν (hr)	κ	ϕ	Expected cell dur (hr)	Expected cell inter-arrival dur (hr)	Expected storm dur (hr)
	1	0.04	0.88	2.55	0.89	0.06	0.02	0.57	9.80	27.84
	2	0.07	1.22	2.05	0.21	0.18	0.06	0.20	1.12	3.55
	3	0.07	1.28	1.86	0.43	1.27E-06	5.09E-05	0.50	391180.30	4967.99
	4	0.05	1.46	2.18	1.04	1.07E-06	0.03	0.88	825164.87	28.30
	5	0.05	1.64	1.99	0.64	1.19E-05	4.84E-04	0.65	54561.67	1298.57
24005	6	0.06	1.95	2.29	0.70	1.33E-06	2.4E-05	0.54	407700.39	22593.40
	7	0.05	4.32	1.81	0.13	2.06E-06	7.43	0.16	75894.44	0.02
	8	0.05	2.27	2.08	0.23	0.53	0.19	0.21	0.40	1.12
	9	0.05	2.01	1.98	0.16	0.22	0.06	0.16	0.74	2.91
	10	0.05	1.79	1.84	0.24	0.02	0.01	0.28	17.73	37.35
	11	0.07	1.13	1.94	0.42	0.24	0.11	0.45	1.91	4.03
	12	0.10	1.03	1.98	0.61	0.06	0.13	0.62	10.49	4.68
	1	0.04	1.24	3.74	1.72	0.23	0.06	0.63	2.69	11.20
	2	0.07	1.68	2.16	0.83	6.95E-05	171.48	0.71	7131.06	0.00
	3	0.05	2.11	1.90	0.43	6.06E-05	0.01	0.48	4788.45	92.09
	4	0.03	1.33	40.78	21.73	0.20	0.04	0.55	2.76	14.80
	5	0.02	1.60	43.76	15.17	0.19	0.04	0.35	1.91	9.80
42008	6	0.03	4.40	2.88	0.12	0.24	0.03	0.06	0.27	2.52
	7	0.03	4.56	1.96	0.27	1.22E-03	0.02	0.28	235.68	17.46
	8	0.03	2.22	2.62	0.66	0.16	0.05	0.41	2.51	7.69
	9	0.03	1.89	2.85	0.75	0.24	0.07	0.40	1.67	5.76
	10	0.07	3.37	1.89	0.39	3.38E-03	0.05	0.44	128.81	8.05
	11	0.04	1.53	2.85	1.07	0.36	0.14	0.58	1.59	4.04
	12	0.04	1.74	2.49	0.79	0.16	0.06	0.53	3.37	8.80
	1	0.01	2.67	2.48	0.65	0.06	1.75E-03	0.44	7.81	257.30
	2	0.06	1.65	5.22	5.09	0.32	0.23	1.20	3.81	5.13
	3	0.11	1.94	3.34	3.08	3.1E-04	0.12	1.32	4386.96	11.17
	4	0.08	1.57	2.25	1.30	0.05	0.07	1.05	23.02	15.35
	5	0.14	2.15	1.75	0.31	1.16E-05	1.95E-03	0.42	36070.51	220.22
74001	6	0.06	1.97	2.85	2.11	0.01	0.05	1.14	104.73	24.60
	7	0.01	1.97	5.29E+08	8.80E+08	0.04	0.01	1.66	40.57	218.85
	8	0.05	2.01	7.49	8.84	0.20	0.27	1.36	6.92	4.99
	9	0.14	2.35	2.16	0.83	1.96E-05	2.76E-04	0.71	36386.14	2377.23
	10	0.21	2.21	2.14	0.79	0.27	3.26	0.69	2.60	0.21
	11	0.17	1.85	2.37	1.30	4.01E-03	0.42	0.95	236.45	2.24
	12	0.05	13.10	3.93	0.03	0.21	4.83E-03	0.01	0.06	2.46

Parameters of gamma distribution, α and ν

Catchment 42008 shows relatively large values of α and ν in April and May. For catchment 74001, the July values of α and ν are very large.

Ratio of expected cell duration to cell inter-arrival time, κ

Values of κ are mostly less than 0.3 and all three catchments have some values that are close to 0.

Ratio of expected cell duration to expected storm duration, ϕ

Within the framework of the single-site model, a storm is a rainfall event composed of a number of rain cells of various intensities and durations. The storm duration is defined as the time over which the storm is capable of generating cells. This is generally found to be smaller than 1. However, each of the three catchments has one month with a value of ϕ that exceeds 1, the highest being 171.48 for July in catchment 24005. These are cases where most of the rainfall is generated by a single cell, so that a simpler model would probably perform equivalently well.

Expected cell duration

The expected cell duration is the ratio between ν and $\alpha - 1$. For catchments 24005 and 42008 its value is less than 1 hr for all months. For catchment 74001 there are six months when the expected cell duration exceeds 1 hr. The smallest value of 0.01 hr for December is associated with the highest mean cell intensity of 13.10 mm hr^{-1}.

Expected cell inter-arrival time

The expected cell inter-arrival time is the average time between two consecutive rain cells. For all three catchments some extremely large values are shown in Table 4.3. There is no obvious seasonal pattern in the occurrence of such large values.

Expected duration of storm activity

The expected duration of storm activity represents the length of time during which cells can be generated. For catchment 42008, the largest value is 92.09 hr in March, while several values exceeding 200 hr are found for catchment 74001 and maximum mean storm duration for catchment 24005 exceeds 22,000 hr in June.

4.5.2 Summary and frequency statistics

The performance of the RPBL model was assessed by analysing a range of summary statistics of the observed and simulated hourly rainfall series, together with frequency

curves illustrating the reproduction of the general characteristics of the rainfall regime, as well as extreme values. The summary statistics and frequency curves considered are listed in Table 4.4.

The statistics were defined to provide an accurate and comprehensive description of the rainfall regime. A range of accumulation durations was used to capture the properties of different rainfall types: heavy rainfall over short durations tends to be representative of convective events, while longer durations (up to 48 hours) are more likely to reflect the characteristics of frontal events (Hulme and Barrow, 1997). The mean and variance of rainfall depth were evaluated for durations of 1, 2, 3, 4, 6, 8, 12, 18, 24 and 48 hours on an annual basis and also for the months of January, April, July and October, chosen to be broadly representative of the winter, spring, summer and autumn seasons respectively.

Table 4.4 Summary statistics and frequency curves considered.

Summary statistic	**Description**
Mean depth, complete series	Average magnitude of d-duration events
Mean depth, conditional wet ($r > 0.1$ mm)	Average magnitude of rainfall (non-dry) events
Mean depth, conditional very wet ($r > R10$)	Average magnitude of very wet events
Variance, complete series	Variability of d-duration events
Variance, conditional wet ($r > 0.1$ mm)	Variability of rainfall (non-dry) events
Variance, conditional very wet ($r > R10$)	Variability of very wet events
Frequency curve	**Description**
rdc, rainfall depth	Overall rainfall regime
rdc, length of dry episodes	Persistence of dry intervals
rdc, length of very wet ($r > R10$) episodes	Persistence of very wet events
rfc, annual (monthly) frequency curve	Regime of the extreme events

The *rainfall duration curve (rdc)*, an analogy of the flow duration curve which is widely used in the hydrological community to characterise river flow regime (Shaw, 1983), was chosen to illustrate how well the single-site model reproduces the characteristics of the entire rainfall series. The rdc is defined as the cumulative frequency distribution of the complete series and it relates a magnitude (e.g. rainfall depth) to the percentage of time over which that magnitude has been exceeded. The rdc can be calculated from the original rainfall depth series or from derived series. Rainfall duration curves of three data series were analysed: rainfall depth, the length (in hours) of continuous dry period, and the length of periods when rainfall is always above a threshold (very wet periods). The latter two rainfall duration curves are indicators of the persistence of dry and very wet periods, respectively. The chosen threshold was R10 (rainfall depth exceeded 10% of the time) for the very wet periods. The threshold used to distinguish between wet and dry sequences was 0.1 mm.

The *rainfall frequency curve (rfc)* focuses on the largest elements (the maximum observed in each year or each given month), and associates a magnitude with the probability of that value being equalled or exceeded during a specified number of years. This probability is usually expressed in terms of return period (Robson and Reed, 1999). Rainfall frequency distributions are commonly used to characterise the regime of the extremes (Shaw, 1983). Each curve is constructed from a single complete series of rainfall depth.

Rainfall duration curves and frequency curves were derived for observed and

simulated rainfall series for durations of 1, 2, 3, 4, 6, 8, 12, 24 and 48 hours. The rdc of rainfall depth, length of dry period and very wet period were calculated from these accumulated series.

Box-and-whisker plots have been used to present the results of the comparisons between the statistical properties of the observed and simulated data series (Figures 4.2 to 4.23). In each figure, the spread of the 200 simulations is shown by boxes representing the upper and lower quartiles and the median. Whiskers indicate 1.5 times the inter-quartile range and crosses represent outliers (outside the whiskers). Mean values derived from the observed data series are shown on the diagrams as dots. The box-and-whisker plots provide a useful illustration of the level of uncertainty in the stochastic model. The envelope from the 200 simulations (including the outliers) provides approximate 99% prediction intervals.

Mean x hr rainfall depth

Examples of the distributions of observed and simulated mean rainfall depths for a range of durations are shown in Figures 4.2 to 4.4. Since mean rainfall intensity was used in fitting the model parameters, close agreement between the statistics of observed and simulated data is to be expected. The introduction of a threshold in the sampling of η will however tend to lead to a small underestimation of this mean. A much harder test for the model is to reproduce mean rainfall depths calculated from values greater than 0.1 mm (conditional wet), and from values greater than the 10th percentile R10 (conditional very wet).

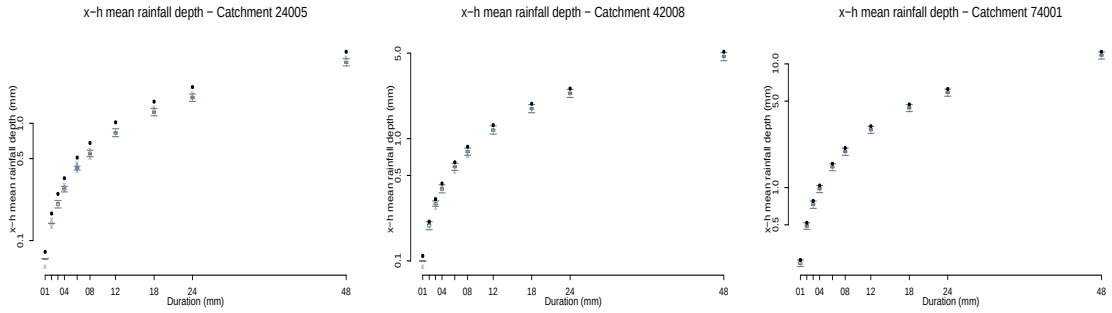


Figure 4.2 Mean rainfall depth (unconditional) - all year (Black dots are observed mean values)

Mean rainfall depth calculated from the complete simulated rainfall series is generally underestimated for all three catchments (Figure 4.2), although the observed mean values lie within the whiskers for catchments 42008 and 74001. Mean rainfall depth for the 1 hr duration is overestimated in two of the three catchments. However, for durations longer than 1 hour, there is a tendency towards increasing underestimation of the mean depth of wet periods (conditional wet) as duration increases (Figure 4.3). The mean depth of very wet periods (conditional very wet) is also underestimated for all three catchments (Figure 4.4).

Mean rainfall depth (unconditional) is reasonably well simulated in winter (January) with observations within the confidence bands for two out of the three catchments for all durations (Figure 4.5). The results for spring (April) show overestimation of the

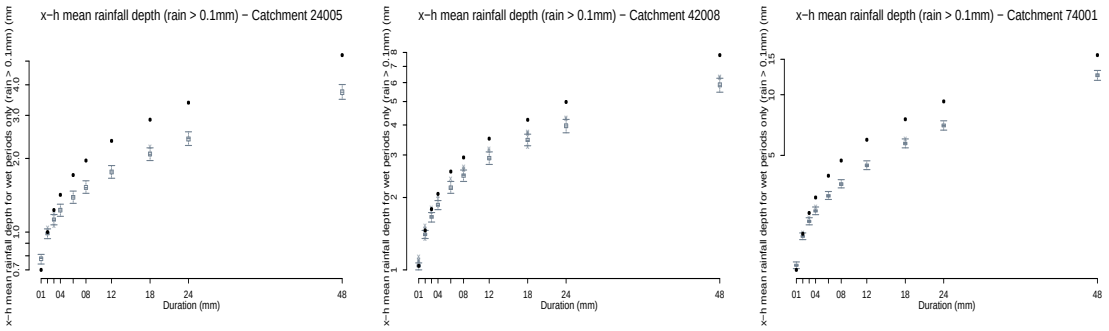


Figure 4.3 Mean rainfall depth (conditional wet) - all year

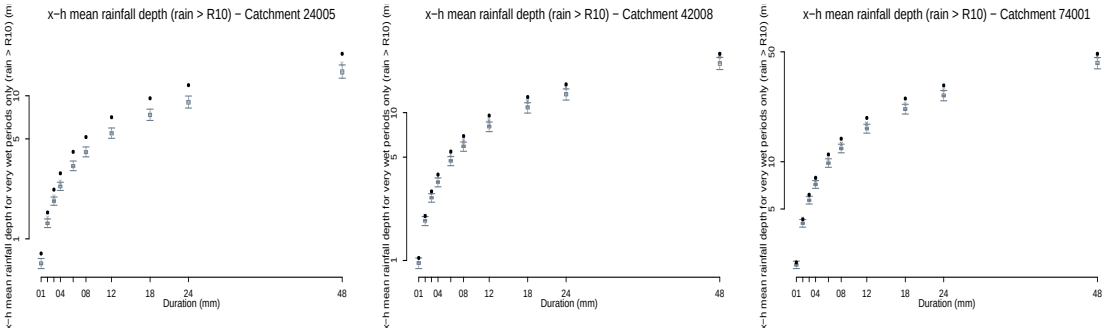


Figure 4.4 Mean rainfall depth (conditional very wet) - all year

unconditional mean depth for catchment 42008 (Figure 4.6), while it is underestimated in summer (Figure 4.7 - July) and autumn (October).

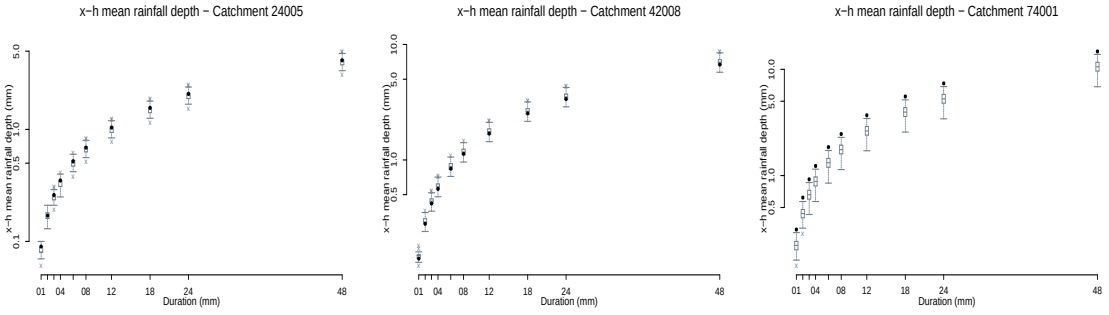


Figure 4.5 Mean rainfall depth (unconditional) - January

In general, the mean rainfall depth during wet and very wet periods is underestimated for all seasons and catchments, except for long durations in catchment 42008 and short durations for catchment 24005 during spring (Figures 4.8 and 4.9).

Variance in x hr rainfall depth

Good reproduction of the variance is a first indication of how well the model captures the variability of the rainfall signal. Annual variance (unconditional and conditional) of rainfall depth of all durations is reasonably well reproduced, with most observed values

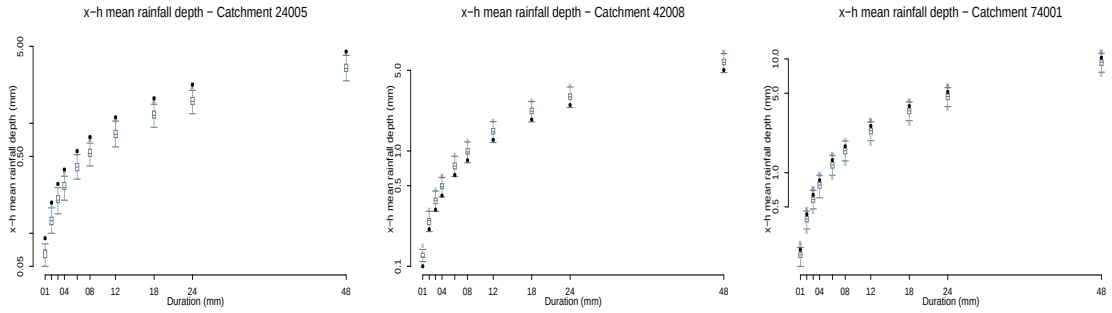


Figure 4.6 Mean rainfall depth (unconditional) - April

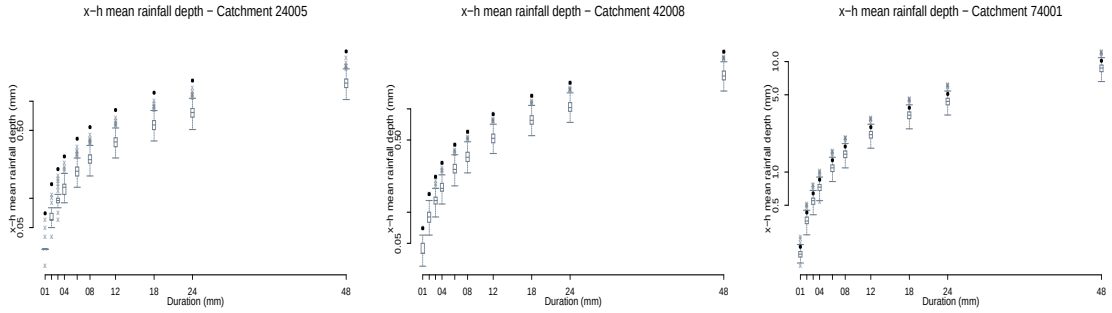


Figure 4.7 Mean rainfall depth (unconditional) - July

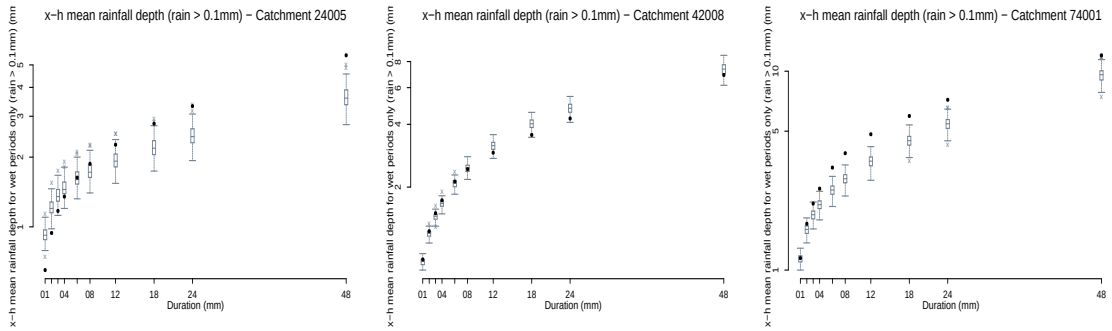


Figure 4.8 Mean rainfall depth (conditional wet) - April

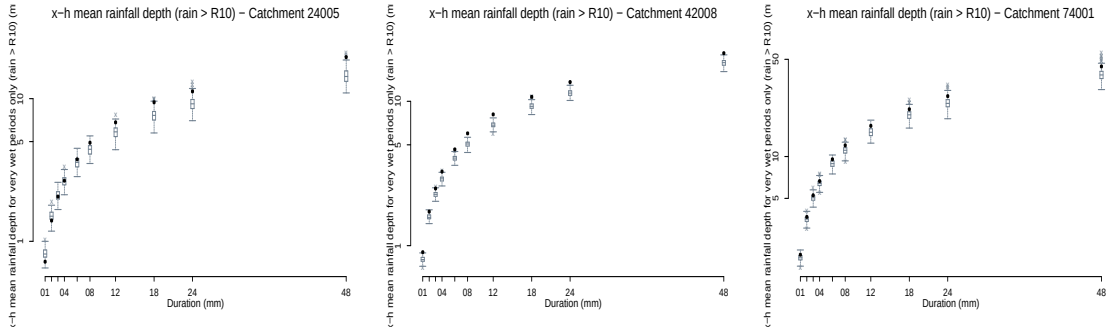


Figure 4.9 Mean rainfall depth (conditional very wet) - April

within (or at the fringe of) the confidence bands (Figures 4.10 to 4.12). For all three catchments, the variance of very wet periods is slightly overestimated. The simulated data are skewed towards 'positive' extremes, with outliers only above the upper confidence band limit (note the log y-axis).

Again, in individual seasons the variance of rainfall depth of the entire series and of wet periods are generally underestimated, while the variance of very wet periods is well estimated in both spring and autumn (Figures 4.13 to 4.16).

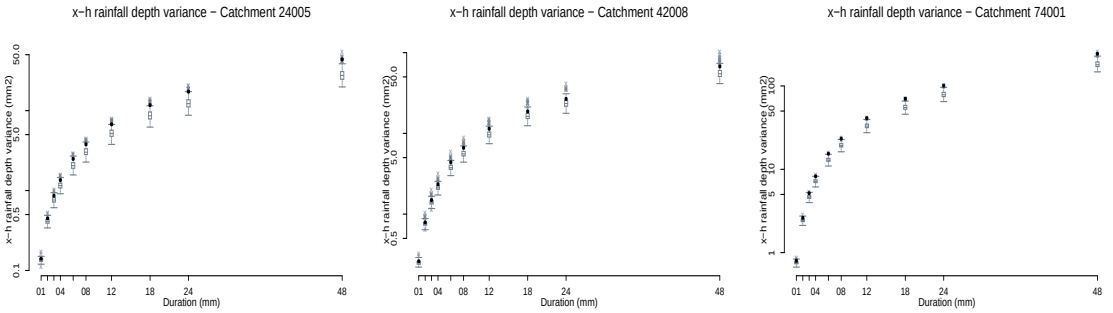


Figure 4.10 Variance of rainfall depth (unconditional) - all year

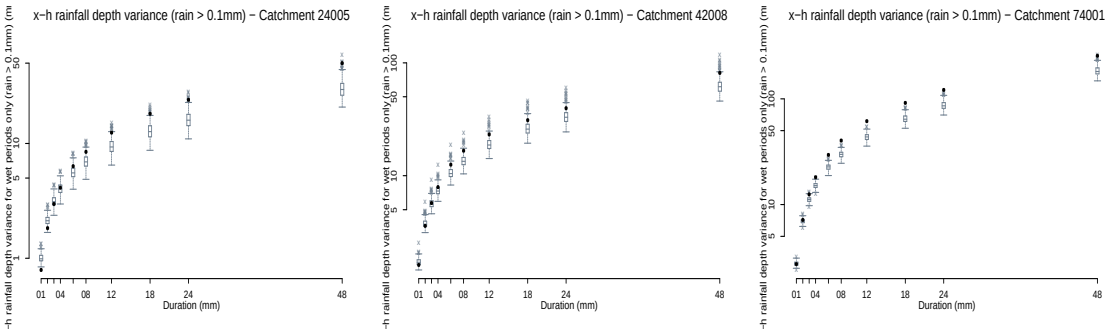


Figure 4.11 Variance of rainfall depth (conditional wet) - all year

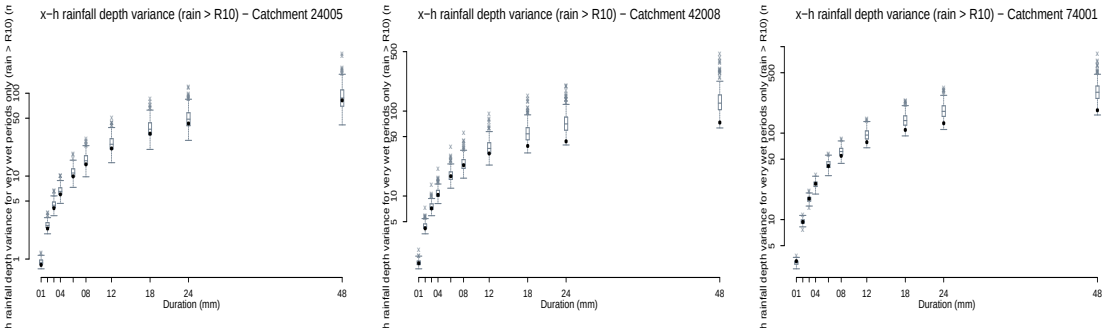


Figure 4.12 Variance of rainfall depth (conditional very wet) - all year

Rainfall duration curves

The rainfall duration curve (rdc) provides an illustration of the frequency distribution of the whole data series. Figure 4.17 shows the 12 hr rdc for the three catchments. The

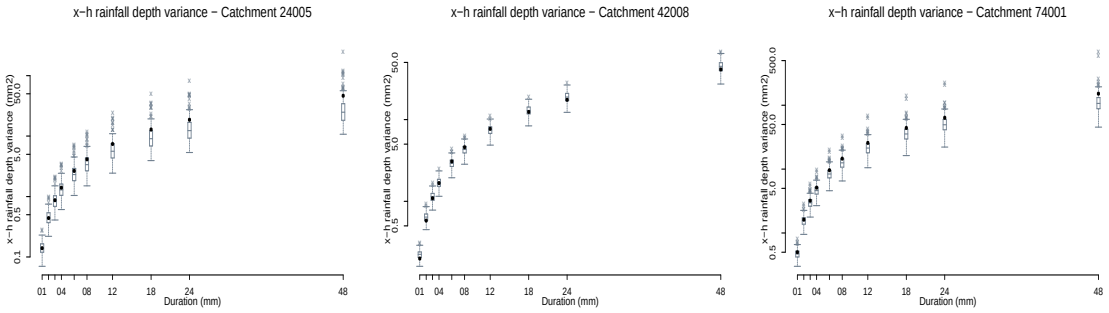


Figure 4.13 Variance of rainfall depth (unconditional) - April

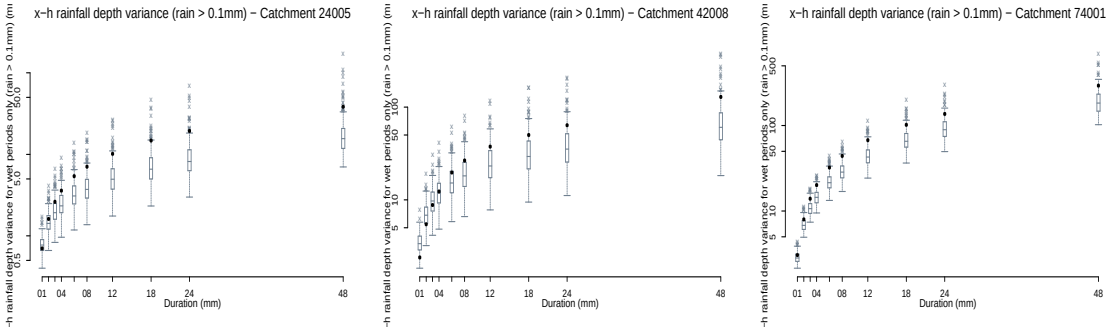


Figure 4.14 Variance of rainfall depth (conditional wet) - October

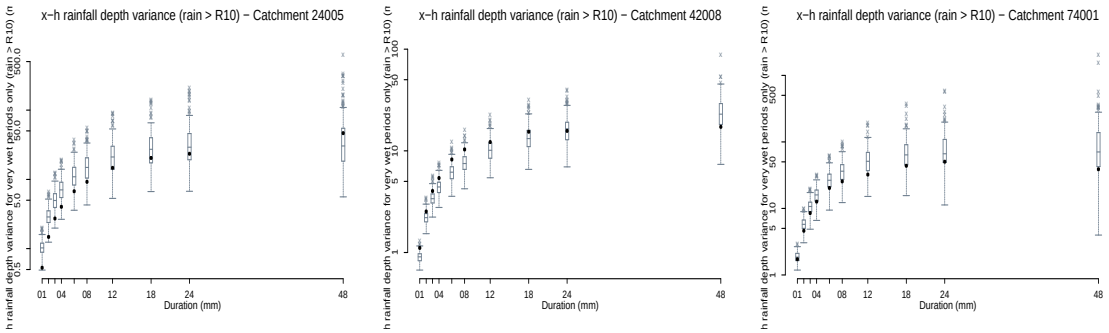


Figure 4.15 Variance of rainfall depth (conditional very wet) - April

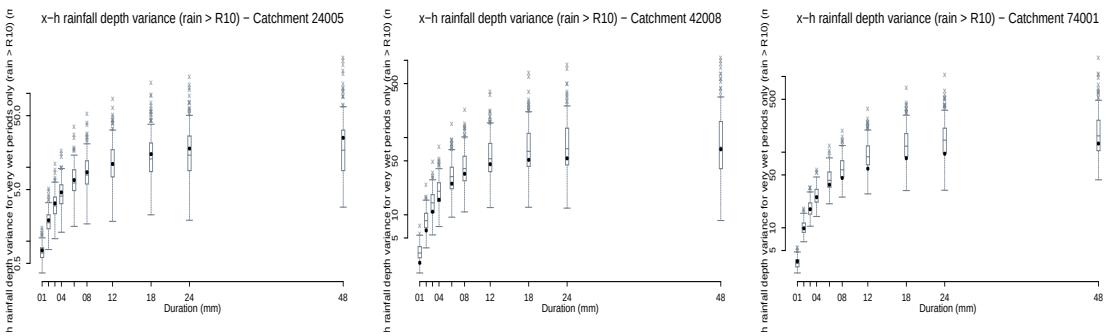


Figure 4.16 Variance of rainfall depth (conditional very wet) - October

frequency distributions of the observed series are well reproduced by the simulations for more frequent events (from about the 20th percentile). However the less frequent events are underestimated for all durations. For all catchments the percentage of time with no rain is well reproduced.

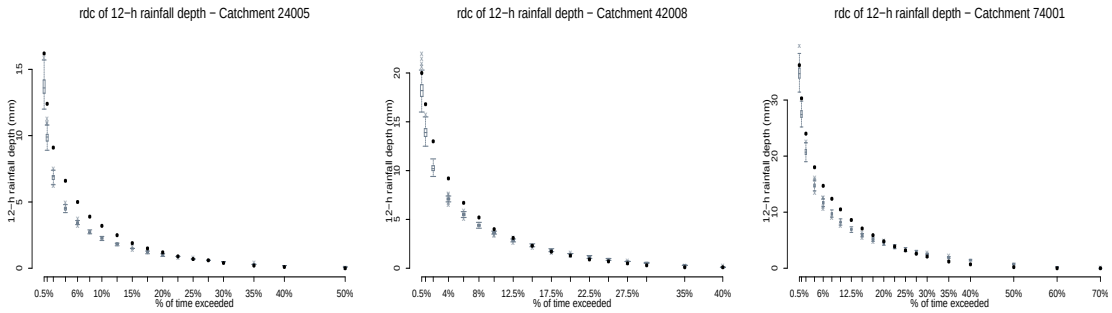


Figure 4.17 Rainfall duration curve of 12 hr rainfall depth

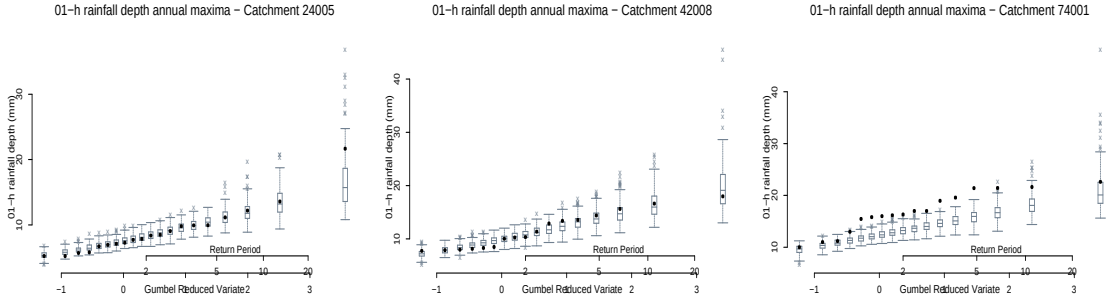


Figure 4.18 Rainfall frequency curve for 1 hr duration - all year

Rainfall persistence

The model's ability to reproduce the persistence shown in the observed data series was assessed from the rdc of the length of dry periods and the rdc of the length of very wet periods. Rainfall depth series of each required duration were derived, and the number of hours when the x-hour rainfall depth was continuously lower than 0.1 mm (dry period) or greater than R10 (very wet period) was calculated. This provided the durations of both dry and very wet periods. Example rdcs are given in Figures 4.18 and 4.19.

For all three catchments, the simulations underestimate the durations of dry spells with the differences being greatest for longer accumulation periods (over 12 hours) and smallest for 1 hr rainfall depths.

The pattern is less clear for very wet spells. Figure 4.19 shows rdcs of very wet periods for the 6 hr duration and some overestimation of rarer events is evident for catchment 42008.

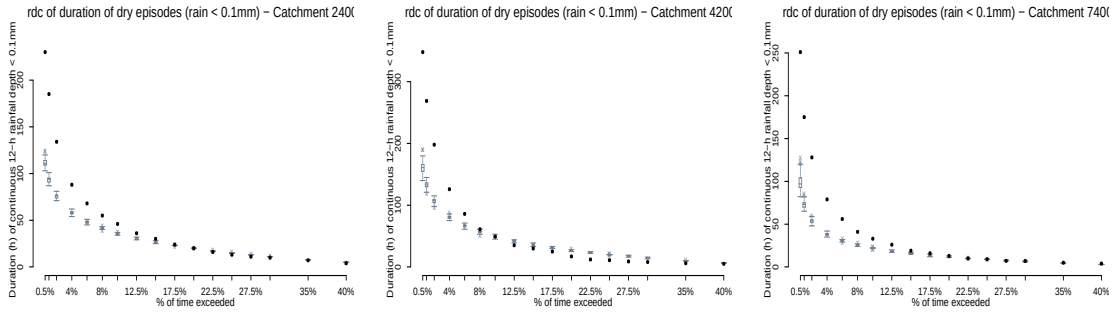


Figure 4.19 Rainfall duration curve of duration of dry episodes

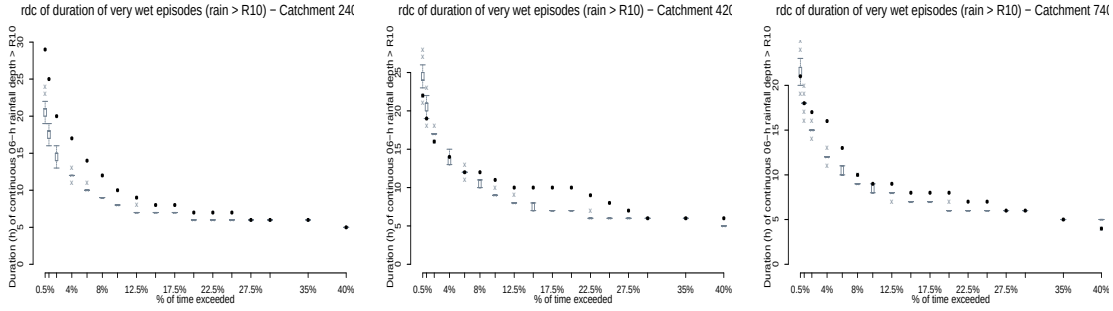


Figure 4.20 Rainfall duration curve of duration of very wet episodes

Rainfall frequency curves

Rainfall frequency curves were derived from annual and monthly maxima abstracted from the observed and simulated data series in order to assess the single-site model's ability to reproduce rainfall extremes. In general, rainfall frequency curves derived from the simulated data correspond well with the observed maxima for all three catchments. The 1 hr annual maxima for catchment 74001 are slightly underestimated by the model, while 4 hr maxima for catchment 42008 are overestimated (Figures 4.20 and 4.21). Frequency curves based on monthly maxima for catchment 74001 are shown in Figures 4.22 and 4.23.

The main reason for introducing a threshold for the value of η in the simulation procedure was that the original model produced a large number of very extreme events for some catchments (Prudhomme and Stewart, 2005). The occurrence of these outliers within the simulations was particularly apparent in spring and summer, especially in the southernmost catchments.

The initial analysis using the unrestricted eta version of the simulation software compared the outliers with point estimates of probable maximum precipitation (PMP) derived from the Flood Studies Report (NERC, 1975), assuming that such estimates might be associated with a return period of the order of 10,000 years. A number of exceedances of the FSR 24 and 48 hr PMP values were found within the simulated data, caused by rain cells of very long duration.

The introduction of the threshold for η has resolved the problem to a large extent, but some values exceeding the FSR 24 and 48 hr PMP estimates were produced by the

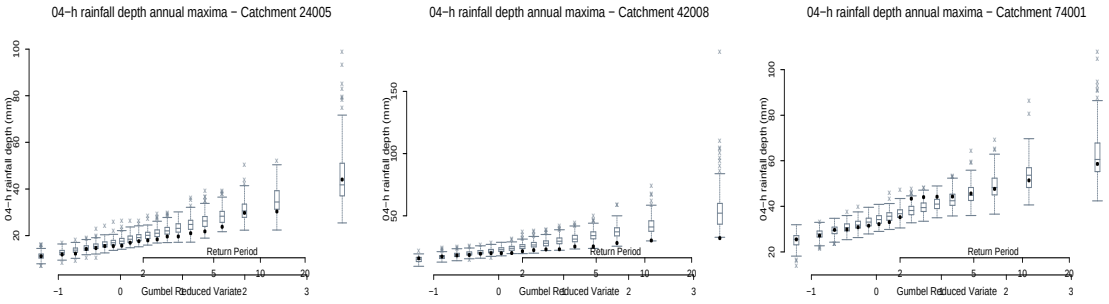


Figure 4.21 Rainfall frequency curve for 4 hr duration - all year

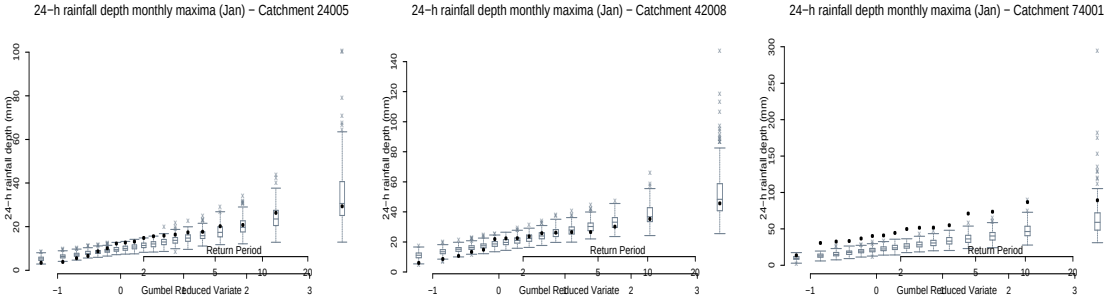


Figure 4.22 Rainfall frequency curve for 24 hr duration - January

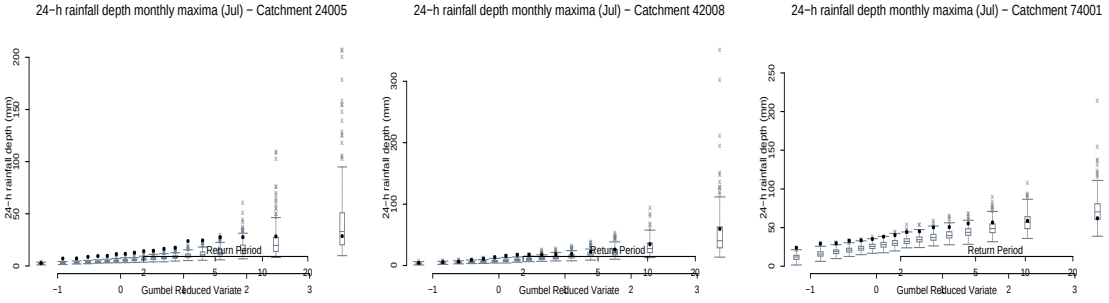


Figure 4.23 Rainfall frequency curve for 24 hr duration - July

Table 4.5 Simulated 24 and 48 hr catchment rainfalls exceeding FSR point rainfall PMP estimates for catchments 24005, 42008 and 74001

	Events > 24 hr PMP			Events > 48 hr PMP		
Catchment	Max. 24 hr depth (mm)	Total event duration (h)	Month of occurrence	Max. 48 hr depth (mm)	Total event duration (h)	Month of occurrence
24005	294	28	Aug	328	53	Apr
42008	303	42	Jul	390	78	Jul
	351	33	Jul	363	73	Dec
	314	32	Oct			
	308	39	Dec			
74001	319	31	Feb	547	87	Mar
	403	37	Oct	474	77	Apr
	334	33	Oct	345	72	Feb
				482	83	Apr
				337	65	Oct
				419	82	Oct
				361	63	Nov
				385	71	Jul
				389	75	Oct

revised simulation procedure as shown in Table 4.5.

It should be pointed out that such exceedances could be eliminated by increasing the threshold used in sampling η. This would however lead to a non-negligible underestimation of the unconditional mean rainfall intensity.

4.5.3 Conclusions

The constraint that the value of η should exceed 0.1 which was introduced in the final version of the simulation software improved the quality of the simulations considerably. In particular, the problem of the generation of extremely large and therefore unrealistic rainfall cells was largely resolved. However, time did not permit a full reanalysis of all the data for the eight catchments originally selected. In addition, the sensitivity of the results to the chosen value of the threshold for η has not been analysed. Further investigation of this aspect of the single-site model is therefore recommended.

4.6 Conclusions

The results of the analysis show the potential of the Random Parameter Bartlett-Lewis (RPBL) model to simulate hourly rainfall series. Some problems in the characterisation of mean rainfall depth and the persistence of wet and dry episodes have been identified. An earlier analysis found that the original formulation of the RPBL model led to the appearance of an unrealistically large number of high magnitude rainfall events. This was resolved by the introduction of a threshold of 0.1 for the value of η (effectively defining the rate at which cells terminate) within the model. The sensitivity of the results to the threshold value has not been tested.

Insufficient time was available to re-analyse all eight catchments in the initial dataset.

Thus the model has been applied to only three hourly rainfall series taken from the FD2106 project archive and further testing is recommended. It is therefore difficult to generalise about the model's ability to simulate rainfall of different types (for example, frontal and convective rainfall) and to draw conclusions about the model's performance in different regions of the UK. Further testing is required before conclusions regarding the model's performance in different regions in UK can be reached.

The data used in the testing of the simulation procedure were catchment average hourly rainfall depths, a choice dictated by the requirement for the joint testing of rainfall and runoff models. Although the catchments chosen were relatively small, the catchment average data may have different statistical properties to the point rainfall series from which they are derived. In particular, the catchment average data would be expected to display less variability than data measured at a single point. Additionally, the number of hourly raingauge records available has an impact upon the hourly variability of the data sets: if one gauge only is available, the data will have the same hourly variance as single-site rainfall, although the daily variance is a catchment average one. This will have a distorting effect upon the optimal parameters obtained, since variances at both time-scales are used in the fit. These issues explain some of the discrepancies noted above. It is emphasised that the models applied here were designed for use with single-site rainfall (not catchment average rainfall) - with no account taken of catchment statistics being different than those of a single site. It is recommended that the models applied here are further tested regarding the conditions of applicability in terms of catchment size and data availability. Analysis of observed data might allow this correction to be estimated empirically, improving applicability for larger areas.

Chapter 5 Joint testing: rainfall simulation and rainfall-runoff modelling

5.1 Introduction

Recent research at CEH Wallingford under companion project FD2106 has developed a new approach to flood frequency estimation based on continuous rainfall-runoff modelling at both gauged and ungauged locations (Wallingford, 2005). This so-called next generation method makes use of advances in hydrological modelling techniques, improvements in computer technology and the increasing availability of rainfall and river flow data of high temporal and spatial resolution. The approach offers potential advantages over statistical and event-based methods since it is able to account directly for the effects of antecedent catchment wetness and river confluences on flood frequency. In addition, the method offers advantages such as the ability to model future changes in the frequency of flooding under alternate climate scenarios. The continuous simulation approach can be used to generate long sequences of synthetic river flow data and thus can be used to supplement gauged records or to simulate data at ungauged sites.

This chapter gives details of the joint testing of the continuous flow simulation methodology with synthetic rainfall series generated by the Random Parameter Bartlett-Lewis (RPBL) model described in Chapters 2 and 4. Further information is given by Prudhomme (2005).

5.2 Methodology

5.2.1 Case study catchments

The FD2106 study used hourly rainfall and flow data from 39 catchments in Great Britain with areas ranging from 0.9 km^2 to just over 500 km^2. The minimum length of record was eight years. Initially, eight catchments were selected for the joint testing of rainfall and flow simulation models, but only three were used in the final analysis due to a modification of the single-site rainfall model (see Chapter 4). Details of the catchments are presented in Table 5.1 and the catchment locations are shown in Figure 4.1.

Although at least 15 years of catchment average hourly rainfall data were available for each site, individual recording raingauge records tended to be much shorter than this. For this reason, and also because catchment average rather than point rainfall was required as the input to the rainfall-runoff models, the RPBL model was fitted to the spatially averaged hourly rainfall series for each catchment. In each case the catchment size was considered to be small enough to justify this decision.

Catchment hourly rainfall series were derived from data for daily and recording raingauges situated within or adjacent to each catchment. The triangle method of Jones (1983) was used to compute catchment average rainfall depths (see Section

Table 5.1 Details of selected catchments.

No.	Catchment	Area km^2	Record Length	No. Gauges	
				Recording	Daily
24005	Browney @ Burn Hall	178	1982-2001	2	15
42008	Cheriton Stream @ Sewards Bridge	75	1985-2001	1	8
74001	Duddon @ Duddon Hall	86	1985-2001	1	8

4.2).

5.2.2 Rainfall-runoff model

Although two conceptual hydrological models were studied in the FD2106 analysis, only one, the Probability Distributed Model, PDM (Moore, 1985) was used in the joint testing exercise. The PDM is based on conceptual water stores. A probability distribution of soil moisture storage capacities is used to evaluate evaporation losses and to determine the time-varying proportion of the catchment which contributes to runoff through 'fast' and 'slow' pathways. A five-parameter version of the PDM had been calibrated for the three catchments selected using an automated sequential procedure (see CEH Wallingford, 2005).

The model was set up to run at an hourly time-step, with rainfall and evaporation series representative of the catchment. Monthly MORECS data were used to determine the potential evaporation (PE) input required by the model. The monthly PE data were downscaled to an hourly time-step using the method applied in the FD2106 analysis. Monthly values were divided equally over the number of days in the month; these were then divided equally over 24 hours to provide hourly estimates. Time did not permit the testing of the alternative method for downscaling monthly PE estimates outlined in Chapter 3.

5.2.3 Rainfall and flow simulation

Hourly catchment average rainfall series were simulated by fitting the restricted eta version of the RPBL model to the data for each of the three catchments. Further details of the method are given in Section 4.3. For each catchment, the final parameter sets shown in Table 4.3 were used to produce 200 simulations of the same length as the data series for that site. The simulations were then used as input data for the PDM, together with the downscaled MORECS potential evaporation data.

5.3 Comparison of observed and simulated flow data

The statistical characteristics of the simulated flow series were compared with those of the observed data using box-and-whisker plots of flow duration curves and flood frequency curves.

5.3.1 Flow duration curves

The flow duration curve shows the percentage of time during which any selected river discharge is equalled or exceeded (Shaw, 1983) and thus characterises the entire flow regime, rather than just the extremes. Flow duration curves for the three catchments analysed are shown in Figure 5.1.

The three catchments show a bias towards underestimation in the reproduction of high flows: the black dots in Figure 5.1. (series derived from continuous flow modelling using observed hourly catchment rainfall) are consistently above the box-and-whisker plots representing the confidence bands of the 200 rainfall simulations. This tendency was also found in the initial results for all eight catchments which were based on the unrestricted eta version of the RPBL rainfall model (see Prudhomme, 2005). Medium to low flows are well simulated for all three catchments.

The higher half of the flow regime (flows higher than the median flow) is underestimated for catchment 24005, while the lower half has the observations within or at the limit of the confidence band. Similar results are shown for catchment 42008.

Catchment 74001 is the only catchment where the simulations overestimate some part of the flow regime (median flow and lower are overestimated). The upper half of the flow regime is either underestimated (flows exceeded less than 20% of the time) or well reproduced (observed percentiles within the simulated confidence band).

5.3.2 Flood frequency curves

Flood frequency curves were constructed using a peak-over-threshold (POT) approach. For each data series, a number, n, of flood peaks were abstracted where n is equal to three times the length of the series in years. The flood peaks were plotted using the Gringorten plotting position. Results for the three catchments are shown in Figure 5.2.

For all three catchments the 'observed' POT series (derived from the PDM with observed rainfall and PE inputs) are within the simulated confidence bands. The simulation procedure performs well for catchments 24005 and 74001, with a tendency to overestimate peak flows at higher return periods. The results for catchment 42008 show underestimation of the observed series at all return periods.

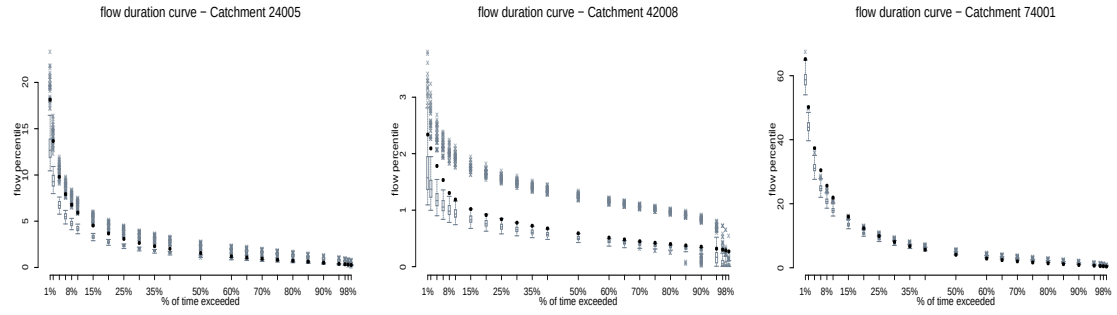


Figure 5.1 Flow duration curves for study catchments. (Black dots show results of rainfall-runoff model using observed catchment rainfall series)

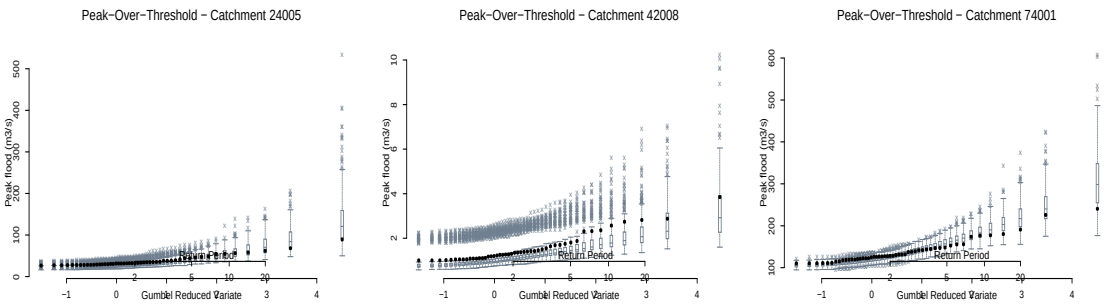


Figure 5.2 Flood frequency curves derived from POT analysis (Black dots show results of rainfall-runoff model using observed catchment rainfall series)

Although the use of the restricted eta version of the RPBL rainfall model has reduced the number of very extreme rainfall events which in turn generate very high peak flows, the simulations for all three catchments contain a number of outliers. Maximum observed and simulated flow values are compared in Table 5.2.

Table 5.2 Maximum 'observed' flow peaks and statistics of simulated flow series.

	24005	42008	74001
Max. 'observed' flow ($m^3\ s^{-1}$)	89.2	3.9	240.4
Upper outlier limit ($m^3\ s^{-1}$)	257.9	6.1	486.7
Upper quartile ($m^3\ s^{-1}$)	159.2	3.9	348.2
Median ($m^3\ s^{-1}$)	120.7	2.9	298.4
Lower quartile ($m^3\ s^{-1}$)	92.7	2.3	254.7
Lower outlier limit ($m^3\ s^{-1}$)	50.2	1.6	176.6
Maximum simulated value ($m^3\ s^{-1}$)	533.1	10.2	607.1
Ratio of simulated maximum to	6.0	2.6	2.5

5.3.3 Inter-annual variability

Although it has not been discussed or investigated here, clustered point process models such as that used here are well known in their relatively poor performance in reproducing variability at high aggregation levels (months and years) and this may be an issue for some applications where antecedent rainfall is important. It is adviseable to check the impacts of such errors on flood frequency simulation when the models are combined with rainfall-runoff models. It is noted that work is underway in project FD2113 to address this issue for the models developed here.

5.4 Discussion and conclusions

Joint testing of the single-site RPBL rainfall model with the PDM rainfall-runoff model for three catchments has shown encouraging results. Further testing of the models on a range of catchment types is recommended before the method can be used with confidence for flood risk assessment at ungauged sites.

The analysis has shown that hourly flow data are generally well simulated. However, flow duration curves based on simulated data are generally underestimated, while an analysis of peak-over-threshold (POT) series has shown that flood peaks are generally overestimated, especially for return periods greater than ten years. The underestimation of the flow duration curve is probably linked to underestimation of rainfall persistence within the RPBL model. Consequently, antecedent catchment wetness, conditioned by the presence or absence of antecedent rainfall, may not be adequately simulated and estimates of flow are therefore too low. The overestimation of flood frequency curves may result from the generation of some very large storms by the rainfall model. The generation of freak storms has been corrected through the use of the restricted eta version of the RPBL model.

We conclude that the models are appropriate to use with caution, but careful assessment on a large number of catchments would be needed before the single-site Random-Parameter Bartlett-Lewis rainfall generator could be used with confidence to provide data for flood risk assessment for ungauged catchments. Any user should validate the model using local data to have sufficient confidence for practical application. More work is recommended to pursue the issue of parameter constraints and non-feasible simulations. There are also ideas for model extension that should be explored to seek improvements in performance.

Part II

Development and testing of short-term spatial-temporal modelling based on daily rainfall modelling and spatial-temporal disaggregation

Part I of this report has been concerned with the development and application of models for the temporal process of rainfall at a single site. The models are constructed in continuous time, enabling their properties to be aggregated into arbitrary discrete time intervals for model fitting and for applications. In principle, these stochastic models can be generalised to simulate spatial-temporal fields of rainfall (such as that presented within Chapter 9). At present, however, such stationary models suffer from some major drawbacks. First they rely on radar data for calibration. Apart from the data quality issues discussed in Appendix F, radar technology is relatively new so that record lengths are limited. Apart from the data quality issues discussed in Appendix F, radar technology is relatively new so that record lengths are limited (the Chenies sequence used above is the longest currently available in the UK). In view of the general consensus that climate varies on decadal or longer timescales (IPCC, 2001, Section 1.2.2), it may be unwise to rely too heavily upon such short runs of data to provide model parameterisations that will hold into the future.

The second drawback of the stochastic models is their stationarity in both space and time. This renders them unsuitable, in the form presented above, for reproducing such hydrologically important features as long term changes in climate or systematic orographic effects.

Although radar and subdaily raingauge data are relatively scarce in the UK at present, long daily raingauge records are much more abundant. This suggests that for short to medium term application, it would be useful to develop models for daily rainfall that can be used to study nonstationarities and to simulate rainfall sequences at a daily timescale. In principle, the output from such models can be downscaled to any desired resolution using some appropriate method; examples of this are presented in Chapter 8. Hourly data may be adequate for most flow simulation purposes.

Many techniques are available for modelling and simulating daily rainfall sequences; some of these are reviewed by Wilks and Wilby (1999), for example. Here, the aim is to produce a methodology that is:

- Conceptually simple
- Suitable for routine application, in terms both of computational demands and of the time taken to fit, validate and simulate models
- Capable of representing realistically, and in reasonable detail, the observed temporal structure of daily rainfall sequences in the UK
- Suitable for the generation of multi-site rainfall sequences, at scales up to a few thousand square kilometres

The final point above ensures that the methodology is suitable for application to all but the largest UK catchments.

This part of the report describes the development and application of models for daily rainfall sequences. Chapter 6 provides an overview of the models themselves; examples of their application to two contrasting data sets are provided in Chapter 7. Chapter 8 then presents and illustrates a methodology for disaggregating a multisite daily rainfall sequence to a subdaily time scale. When combined with the capability for

simulating multi-site daily sequences, this enables (possibly nonstationary) multi-site subdaily sequences to be generated.

Chapter 6 Generalized linear models for daily rainfall

6.1 Introduction to the models

The models developed here are based on Generalized Linear Models (GLMs), which are standard in the statistical literature (McCullagh and Nelder, 1989; Dobson, 2001). The basic idea is an extension of linear regression: we use the values of various predictors (which we call *covariates*) to forecast a probability distribution for each day's rainfall at every site of interest. Possible covariates include previous days' rainfalls (possibly at more than one site), the month of the year, variables representing regional variation (e.g. site altitude) and 'external' climatological factors such as the North Atlantic Oscillation (NAO). Within this framework, temporal and spatial nonstationarities can be accommodated via appropriate choice of covariates. This also allows distributions to be specified at sites for which covariate information is available but rainfall data are not. One application of this is to generate point rainfall sequences at each node of a regular grid, if for example a rainfall-runoff model requires gridded rainfall inputs. This would be an improvement upon the widely-used practice of interpolating the available raingauge observations onto the grid (Jones, 1983; Gannon, 1995), because the properties of interpolated data (particularly their variability) will typically be rather different from those of raingauge sequences. The GLM approach avoids this problem because it provides a means of sampling directly from the distribution of rainfall at each of the grid nodes (if necessary conditioned upon the observed values at raingauges).

Our implementation of GLMs for rainfall modelling is based on that of Coe and Stern (1982) and Stern and Coe (1984). These authors adopted a two-stage approach, as follows:

1. Use logistic regression to model the pattern of wet and dry days. Specifically, denoting by p_{st} the probability of non-zero rainfall at site s on day t, and by $\mathbf{x}_{st}$ a corresponding vector of covariates, the model is given by

$$\ln\left(\frac{p_{st}}{1-p_{st}}\right) = \mathbf{x}'_{st}\beta \tag{6.1}$$

   for a coefficient vector β.

2. Fit gamma distributions to the amount of rain on wet days. If non-zero, the rainfall amount at site s on day t is taken, conditional on a covariate vector ξ_{st}, to have a gamma distribution with mean μ_{st}, where

$$\ln\mu_{st} = \xi'_{st}\gamma \tag{6.2}$$

   for a coefficient vector γ. All gamma distributions are assumed to have a common shape parameter, ν say (if $\nu = 1$ the distributions are exponential). This is equivalent to assuming that, conditional on the covariates, daily rainfall values have a constant coefficient of variation (McCullagh and Nelder, 1989, Chapter 8).

These two models are referred to as ‘occurrence’ and ‘amounts’ models respectively. The right-hand sides of (6.1) and (6.2) are called ‘linear predictors’.

In the GLM framework, models can be fitted and compared using Maximum Likelihood, and tested using a variety of simple but informative checks. This procedure is illustrated in chapter 7 below. Moreover, models can be specified in such a way that the effect of one covariate depends on the values of others — for example, it is known that the NAO affects European climate primarily in the winter months (Hurrell, 1995). In such situations, the covariates involved are said to *interact*. One of the potential advantages of the GLM methodology is that it allows such structures to be incorporated very easily into simulated rainfall sequences.

Further details on the use of GLMs to represent the structure of multi-site rainfall sequences can be found in Chandler and Wheater (2002); a similar approach can be used to study other cllimatological variables such as wind speed (Yan et al., 2002) and temperature (Chandler, 2005).

6.2 Multi-site simulation

As described in the previous section, a GLM enables us to specify probability distributions for daily rainfall at individual sites, conditioned on the values of various covariates. A single-site sequence can then be simulated, given some initial conditions, by sampling a value from the first day’s distribution, using this value to construct a distribution for the second day, sampling a new value for the second day and so on. Typically, the required initial conditions are provided by a few days’ observed data.

For the generation of simultaneous rainfall sequences at several sites, in general it will be necessary to account for the fact sequences at neighbouring sites are not independent. Instead of specifying individual distributions for the next day’s rainfall at each site, it is therefore necessary to specify a joint distribution for at all sites. As in the previous section we proceed in two stages, first defining a joint distribution for the wet-dry pattern of rainfall occurrence, and then the distribution of the rainfall amounts vector at the wet sites. The mean vectors of these joint distributions are specified by the GLMs of the previous section.

The methodology for multi-site simulation is described in detail in Chandler (2002) and in Yang et al. (2006b). Here we present a brief summary.

6.2.1 Rainfall amounts

We deal with the joint distribution of amounts first, since this is easier to specify than that for occurrence. Our approach here was developed in an earlier project for MAFF (Wheater et al., 2000b), and makes use of the observation that if Y is gamma distributed, the distribution of $Y^{1/3}$ is very close to normal. This fact was used by Chandler and Wheater (2002) to check the assumption that wet day rainfall amounts are gamma distributed: specifically, if Y_i is the observed amount for the ith wet day in a database, and μ_i is the modelled mean of the gamma distribution for that case, then

the quantities

$$r_i^{(A)} = (Y_i/\mu_i)^{1/3} \tag{6.3}$$

all share the same normal distribution, approximately. The mean and variance of this normal distribution can be calculated numerically, as described in Section 3.3.2 of Chandler and Wheater (1998). The calculation needs to be carried out only once in any simulation, since the mean and variance depend on ν, the common shape parameter of the gamma distributions, but not upon the means $\{\mu_i\}$.

The quantities $\left\{r_i^{(A)}\right\}$ are called *Anscombe residuals*. Since they are approximately normal, it is natural to describe spatial dependence via a model for their inter-site correlations. One might choose to use a standard geostatistical model to represent these correlations as a function of inter-site distance and direction (Cressie, 1991), or simply to calculate the empirical inter-site correlation matrix from historical data (note, however, that if it is desired to simulate rainfall at locations other than those where raingauge data are available, it will be necessary to use a model to specify the inter-site correlations at ungauged locations).

Simulation of rainfall amounts at 'wet' sites for a particular day therefore proceeds: firstly, by sampling a vector of Anscombe residuals from a multivariate normal distribution with an appropriate mean and covariance structure (standard algorithms exist for this); and secondly by inverting the transformation (6.3) at each site. If any negative values are generated, they are discarded and a new vector is drawn. One potential drawback is that the scheme takes no account of 'dry' sites, since the amounts model is only defined at sites where the rainfall amount is nonzero. Therefore it is not guaranteed to produce small amounts of rain near sites which are dry, although wet sites which are close to each other will tend to have similar rainfall amounts. There are, however, few currently-available simulation methods that overcome this problem, which is discussed in more detail in Yang et al. (2006b).

6.2.2 Rainfall occurrence

In modelling rainfall occurrence, a transformation to marginal normality is not possible and an alternative approach to the modelling of spatial dependence is required. Some possible approaches for achieving this were reviewed in earlier work for MAFF Wheater et al. (2000b, Chapter 4). There, a simple approach was adopted based on a hidden 'weather state' variable, under which each day was classified as either 'wet' or 'dry' over the entire region of interest. On 'wet' days, the probability of rain at each site was increased and on 'dry' days, the probability was decreased, in such a way that the overall probabilities matched those from the GLM.

Unfortunately, experience has shown that this approach is unable to reproduce the very strong dependence in rainfall occurrence that is commonly encountered at small to medium catchment scales in the UK. At these scales, on any given day all sites are usually influenced by the same weather system, so that they tendency to be either mostly wet or mostly dry. The number of wet sites is related to the proportion of an area which experiences rain. Therefore, for hydrological applications it may be important to reproduce accurately the distribution of the number of wet sites, and as part of the current project we have explored the idea of modelling the distribution

directly. This has the advantage of being both conceptually straightforward and computationally feasible. The idea has been used in other applications, notably in the analysis of teratology (developmental toxicity) data — see, for example, Ryan (1995) and references therein.

The approach adopted here is to assume a beta-binomial distribution for the number of wet sites on each day. Let S be the total number of sites, and let $Z_t = \sum_{s=1}^{S} Y_{st}$ be the number of these sites that are wet (here, Y_{st} is a binary variable taking the value 1 if site s is wet on day t and zero otherwise). The probabilities for the beta-binomial distribution are

$$P(Z_t = z) = \binom{S}{z} \frac{\Gamma(\alpha_t + z)\,\Gamma(S + \beta_t - z)\,\Gamma(\alpha_t + \beta_t)}{\Gamma(\alpha_t + \beta_t + S)\,\Gamma(\alpha_t)\,\Gamma(\beta_t)} \tag{6.4}$$

for $z = 0, 1, \ldots, S$. The parameters of the distribution are $\alpha_t \in \mathbb{R}^+$ and $\beta_t \in \mathbb{R}^+$. The mean and variance are

$$\frac{S\alpha_t}{\alpha_t + \beta_t} \quad \text{and} \quad \frac{S\alpha_t\beta_t\,(\alpha_t + \beta_t + S)}{(\alpha_t + \beta_t)^2\,(\alpha_t + \beta_t + 1)}\ , \tag{6.5}$$

respectively. We reparametrise the distribution by setting

$$\theta_t = \frac{\alpha_t}{\alpha_t + \beta_t} \quad \text{and} \quad \phi_t = \alpha_t + \beta_t\ , \tag{6.6}$$

so that

$$\mathrm{E}(Z_t) = S\theta_t \quad \text{and} \quad \mathrm{Var}(Z_t) = \frac{S\theta_t\,(1-\theta_t)\,(\phi_t + S)}{\phi_t + 1}\ . \tag{6.7}$$

Note that the rainfall occurrence model (6.1) allows us to calculate $E(Y_{st}) = p_{st}$, say. Hence $\mathrm{E}(Z_t) = \sum_{s=1}^{S} p_{st}$. This therefore allows us to calculate the value of θ_t directly from the occurrence model:

$$\theta_t = S^{-1}\mathrm{E}(Z_t) = S^{-1}\sum_{s=1}^{S} p_{st}\ . \tag{6.8}$$

The parameter ϕ_t controls the shape of the distribution, as illustrated in Figure 6.1. As $\phi_t \to 0$, the distribution becomes increasingly concentrated around 0 and 1, and as $\phi \to \infty$ the distribution tends to the binomial, with parameters S and θ_t. Since the binomial distribution arises if all the Ys are independent and identically distributed, ϕ can be regarded as an overall summary of spatial dependence, with small values corresponding to strong dependence.

In the work reported here, we have assumed that $\phi_t = \phi$ is constant for all t. Figure 6.1 shows that this is likely to be reasonable in practice: as θ_t varies, the effect is to move along one of the rows of the figure. This provides a plausible means of reproducing typical 'summer' and 'winter' distributions of numbers of wet sites, for example.

When $\phi_t = \phi$ is constant, the beta-binomial model can be fitted easily using a method of moments. Specifically, given data $\{(S_t, Z_t, \theta_t = \sum_s p_{st}/S_t) : t = 1, \ldots, T\}$ we define $R_t^2 = (Z_t - S_t\theta_t)^2 \ / \ [S_t\theta_t\,(1-\theta_t)]$ (note that we write S_t here, because in practice it is unlikely that the number of sites yielding data will be the same for all t). From (6.7), we have

$$\mathrm{E}\left(R_t^2\right) = \frac{\phi + S_t}{\phi + 1}\ ,$$

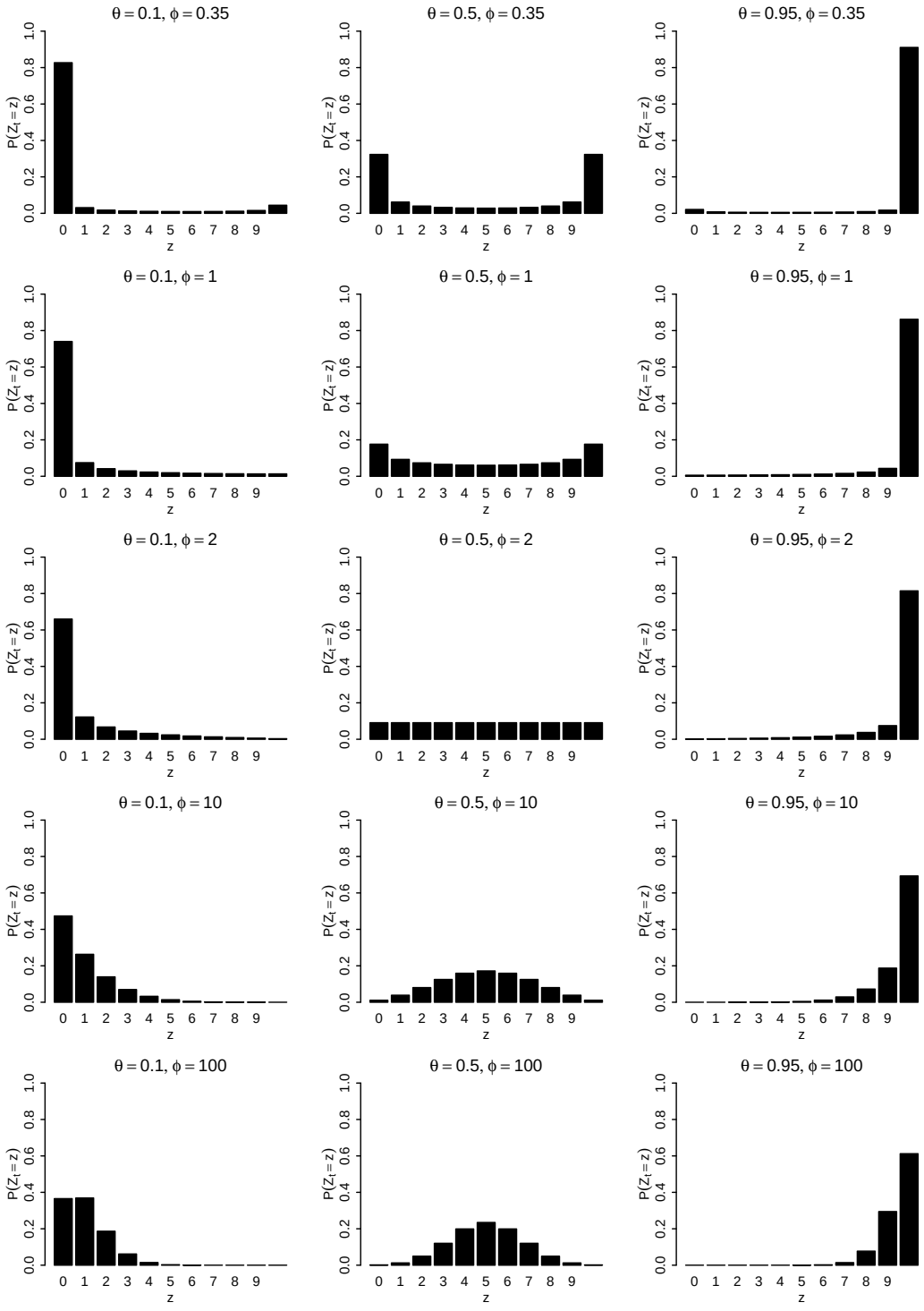


Figure 6.1 Examples of beta-binomial distributions for the number of wet sites (Z) from a total of $S = 10$ sites. The plots show the probabilities $P(Z = z)$ for different values of θ and ϕ.

so that ϕ can be estimated as

$$\hat{\phi} = \frac{\sum_{t=1}^{T}(S_t - 1)}{\sum_{t=1}^{T}\left(R_t^2 - 1\right)} - 1 \ . \tag{6.9}$$

The parametrisation in terms of θ_t and ϕ, and proposal for a moment-based estimate of ϕ, are similar to those given by Williams (1982).

Having specified a plausible model for the distribution of Z_t, a set of dependent rainfall occurrences can be simulated by first sampling the number of wet sites from this distribution, and then allocating the positions of these sites. This needs to be done in such a way as to reproduce correctly the marginal probabilities of rain at each site, according to the rainfall occurrence model. An algorithm for achieving this is described by Chandler (2002).

There are some potential drawbacks with this beta-binomial model. One is that the model may fail if sites are too close together; another is that there is no concept of inter-site distance except as implied by the regional variation of the ps in the occurrence model. More details are given in Yang et al. (2006b). The latter problem is unlikely to be serious at most UK catchment scales where inter-site dependence is uniformly high. As far as the former problem is concerned, our view is that sites would have to be extremely close together (e.g. separated by tens of metres) for the model to provide a gross misrepresentation of local spatial structure, and this is unlikely to cause too many difficulties in practice.

It should be pointed out that the beta-binomial model on its own does not define a unique joint distribution for rainfall occurrence at all sites; hence our simulation algorithm is effectively sampling from only one of many possible joint distributions. We do not regard this as a problem, since the aim is merely to reproduce accurately the distribution of the number of wet sites while preserving the probabilities of rain at each site. Perhaps a more serious problem is that the postulated beta-binomial distribution may not be compatible with the marginal probabilities of rain at each site — obvious examples of incompatibility arise when $P(Z_t = S) > \min_s P(Y_{st} = 1)$ and when $P(Z_t = 0) > \min_s P(Y_{st} = 0)$, for example. In practice however, we have only ever encountered incompatibility in cases where the occurrence model (6.1) generated ps at one or two sites that differed substantially from the majority. This type of behaviour is unrealistic, and is symptomatic of a poor occurrence model. In the work reported in subsequent chapters we have simulated tens of thousands of years of daily rainfall without once encountering the problem.

6.2.3 Imputation

Daily rainfall records often contain missing values, leading to uncertainty regarding the properties of historical rainfall sequences. If we can determine the distribution of these missing values conditional upon the observed values at all sites, then we can simulate from this conditional distribution many times to construct uncertainty envelopes for historical properties of interest. We refer to this process as *imputation*. It can be an extremely helpful aid to the interpretation of historical records.

The spatial dependence structures proposed here are specified in such a way that imputation is straightforward. For the occurence model, the conditional distribution

emerges naturally as a by-product of the algorithm for allocating the locations of wet sites as described in Chandler (2002). For the amounts model, dependence is specified via a multivariate distribution for the Anscombe residuals. If data from some sites are missing, but others are observed, then Anscombe residuals can be computed from the observed sites and the conditional distribution of the missing residuals, which remains multivariate normal, can be calculated (Krzanowski, 1988). Missing residuals can then be simulated from this conditional distribution, and back-transformed to yield imputed rainfalls.

6.3 Practical implementation

The methods described in this chapter are all implemented in the software package GLIMCLIM (Chandler, 2002). This is designed to make the fitting and simulation of the models relatively straightforward without needing to worry about technical details or data manipulation. The package makes use of model definition files to define, for example, the covariates to include in a model and the spatial dependence structure to use used. When models are fitted, the package provides diagnostic checks as well as updated definition files that can subsequently be used as input to a rainfall simulation. This package has been used to produce all of the GLM-based results in this report.

6.4 Summary of chapter

GLMs have been introduced as an intuitive and interpretable means of modelling daily rainfall sequences, either at a single site or at multiple sites. They are able to accommodate nonstationarity, through covariates that change in space or time; and are also able to generate dependent sequences at neighbouring sites. The GLIMCLIM software package (Chandler, 2002) provides a means of fitting and simulating the models, without worrying too much about the technical details or the manipulation of large volumes of data. However, as with all modelling tools, care is required in the implementation of the methodology. Some guidelines are given in the GLIMCLIM user manual (Section 6). Users should also note the following:

1. The models are designed to be fitted to data from a network of daily rain gauges. The amounts models, in particular, assume that the coefficient of variation in rainfall amounts is constant. Our experience is that this is true in general of daily raingauge data in many regions of the world. However, it may not be true for gridded data derived by interpolation from a network of gauges, because the variability at each grid node will depend on the density of gauges in its neighbourhood (among other things). In general therefore, we do not recommend that the models are fitted to gridded data. If gridded rainfall data are required (for example, as input to a rainfall-runoff model), we recommend that GLMs are fitted directly to the available daily raingauge records and then simulated on a regular grid. If necessary, the simulations can be conditioned on all available rain gauge data (this is essentially an imputation exercise, which can be handled by GLIMCLIM). This ensures that the simulated sequences

have, as far as possible, the properties that would have been observed using raingauges at each grid node.

2. The spatial dependence structures available in the GLIMCLIM software (in particular relating to rainfall occurrence) are designed primarily for use when inter-site dependence is uniformly high. This is the case at spatial scales where most sites can reasonably be assumed to be affected by the same weather systems on any day; hence the dependence structures are probably suitable for all but the largest UK catchments. At larger scales (say, above 2000 km^2), the methodology may still produce reasonable results; however, users should check carefully to ensure that the results of any simulation exercise appear reasonable.

3. The beta-binomial model for dependence in rainfall occurrence is unlikely to capture the the strong dependence between sites that are too close together (e.g. tens of metres). It is therefore advisable to ensure a reasonable spacing between sites in any simulation exercise (note, however, that there is no problem with *fitting* the models to data from sites that are very close together).

4. A beta-binomial distribution for the number of wet sites may not always be compatible with the probabilities of rain predicted by the logistic regression model for rainfall occurrence. If the GLIMCLIM software package encounters such incompatibility, it will report a warning and modify the predicted probabilities. However, in practice such incompatibility is caused by predicted probabilities that are unrealistic (usually because the probabilities at one or two sites are very different from the remainder). Warning messages from the software should therefore be taken as an invitation to improve the rainfall occurrence model.

Chapter 7 Case studies in the multi-site modelling of daily rainfall

In this chapter we apply the GLM methodology to two contrasting data sets. Models are fitted to each data set and used to simulate multi-site rainfall sequences in each area. It is shown that the GLMs are able to reproduce a wide variety of statistical properties of observed rainfall sequences, including extremes, over a range of spatial scales. The model for the first data set has also been used, in conjunction with the stochastic spatial-temporal model of Chapter 9, to generate nonstationary rainfall sequences that are continuous in both space and time. The results of this were presented in Chapter 10.

7.1 Data used in the case studies

The first data set used here is from an area roughly $50 \times 40\ km^2$ around the catchment of the river Blackwater in Surrey. The second is from a region of similar size (about 40×60 km^2) in northeast Lancashire. These regions can be regarded as representative of "lowland" and "upland" areas respectively. Daily rainfall data have been obtained from the UK Meteorological Office for both regions; nominally, there are 43 records from the Blackwater region and 41 from northeast Lancashire. The locations of the two study areas, and the gauges, are shown in Figure 7.1.

The data from the Blackwater date from 1904; those from northeast Lancashire date from 1961. In both regions, typically between 10 and 20 gauges to 2000, have been operational at any one time since 1960.

A variety of checks have been carried out on both data sets. These are now summarised briefly. More details can be found in Yang et al. (2006a).

7.1.1 The Blackwater

Although there were nominally 43 distinct records for this region, no data were received for two of the sites. Of the remaining 41 sites, there were five pairs and one triplet of sites in identical locations. For each of these groups, the overlapping portions of the series were found to be identical. Therefore, the longest records in each group have been retained and the remainder discarded. There are 34 unique series left. The locations of the gauges are shown in Figure 7.2, which also shows a relief map of the region. The gauge altitutes range from 10 m to 170 m above sea level.

A preliminary inspection of the data revealed several cases in which a site only had one rainfall value for a particular month; this value was extremely large and was recorded on the last day of the month. The suspicion must be that these values are actually monthly totals (although they are not flagged as such in the data files). Six such values have been discarded from the subsequent analysis.

In addition to the monthly totals, there are many cases in which one day has only one wet site. In some cases the single wet site has a very large amount of rain recorded;

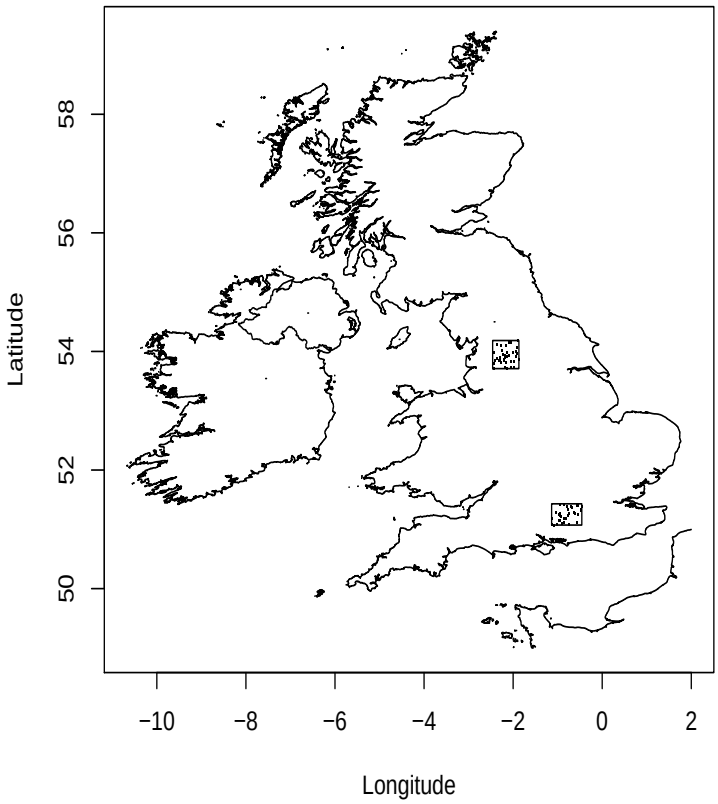


Figure 7.1 Locations of the Blackwater and northeast Lancashire study areas, with positions of raingauges marked within each region.

the most extreme example is a gauge that recorded 75 mm of rain on a day when all other gauges (29 of them) recorded zero. Some of these suspicious cases have been excluded from the following analyses. The criteria for exclusion were:

1. At least 8 gauges were operational on the day in question.
2. Only one gauge reported a non-zero rainfall amount.
3. The non-zero amount exceeded 10 mm

A further difficulty with these data concerns the recording of small rainfall amounts. For the 34 series in the Blackwater, values less then 0.5 mm contribute only about 1% of the total rainfall, but are responsible for almost 20% of the rainfall occurrences. Therefore, when building a model for rainfall occurrence it will be necessary to account carefully for any differences in observational practice concerning the recording of small values. One example of this is that in the mid 1970s, the recording resolution of the gauges changed: prior to about 1975, the smallest value recorded by most gauges is 0.3 mm, and afterwards it is 0.1 mm.

Changes in recording resolution are relatively easy to handle within our modelling framework; we describe how to achieve this below. A more serious problem is that there are substantial spatial inconsistencies in the pattern of rainfall occurrence. To illustrate this, the top left-hand plot in Figure 7.3 compares the cumulative rainfall

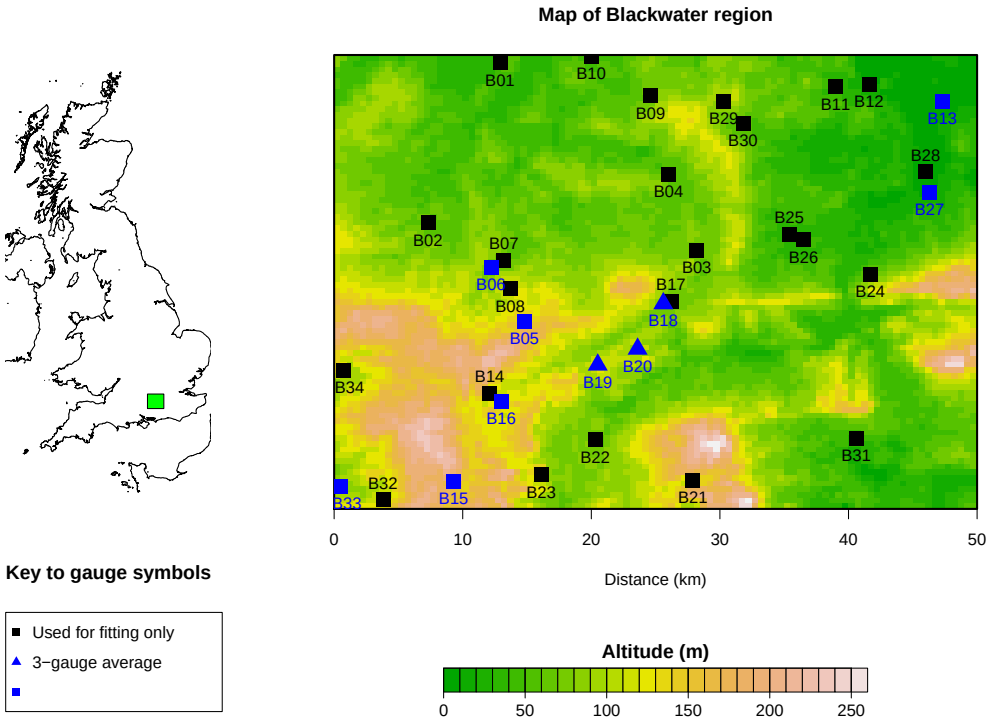


Figure 7.2 Map of the Blackwater area, showing locations of 34 raingauges used in the case study.

occurrence at sites B03 and B17 (see Figure 7.2), from 1961 onwards. Cumulative values are sampled every 10 days. If the rainfall conditions are almost the same for the two sites and the records from the two gauges are both believable, then the cumulative values should be centered along the solid 45°line. However, there is a clear discrepancy between the sites. This discrepancy is less pronounced for rainfall amounts (top right-hand plot), and hence is likely to be caused by different characteristics of small rainfall amounts at the two sites. This conjecture can be verified by setting to zero any daily rainfall below some threshold and replotting. The results, with a threshold of 0.5 mm, are shown in the bottom panels of Figure 7.3. The discrepancies are reduced or almost eliminated.

Another way to investigate spatial inconsistencies in rainfall occurrence is by comparing each gauge with the mean condition of the region. Here we calculate the series of the annual percentage of wet days for each gauge, then calculate the mean series of all the 34 gauges. The starting time for comparison is 1975, when all the gauges have resolution 0.1 mm. The root mean squared error (RMSE) of each series against the mean series is calculated as a measure of discrepancy. The top panel of Figure 7.4 shows all the discrepancies in ascending order. This shows that there is substantial variability about the overall mean; gauges B02, B03 and B31, in particular, are quite different from the rest with RMSEs in excess of 15%. However, for thresholded data (bottom panel), the RMSEs are considerably reduced for almost all sites — particularly B02 and B03. Again, it seems that many of the inter-site differences are associated with small rainfall amounts.

The above analyses suggest two possibilities about the Blackwater dataset:

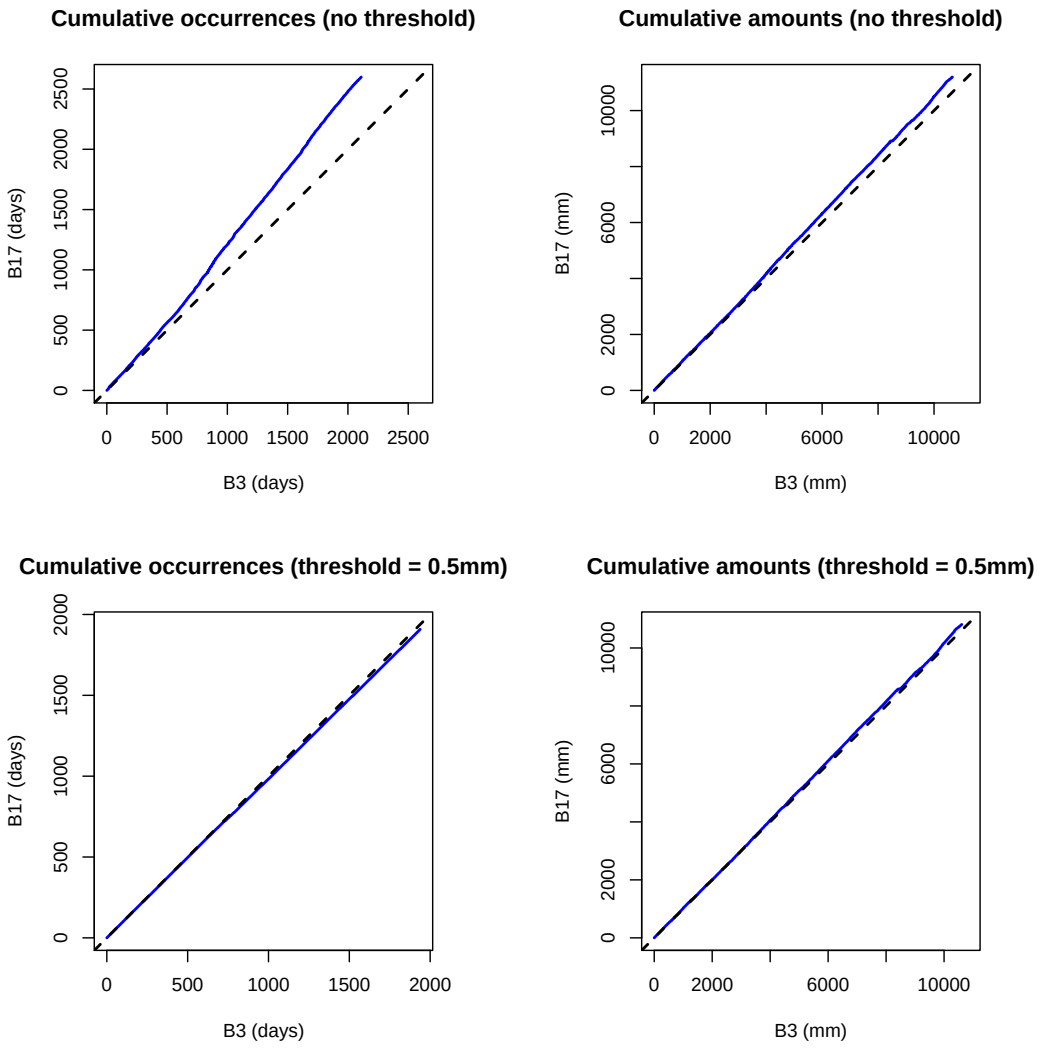


Figure 7.3 Comparison of cumulative rainfall occurrence and amount for neighbouring gauge pair B03 and B17 (see Figure 7.2). The top plots are for the raw data, and the bottom plots for data after applying a threshold at 0.5 mm.

1. The data quality was not well controlled, with human-induced inconsistencies among gauges in different sites, especially when recording small rainfalls; or

2. There may exist different local climate characters in this region. Generally, it is unlikely that a single meso- or large-scale weather system would result in quite different rainfall characters in this small and flat area. However, because the river Blackwater passes through this region, different underlying surface conditions may result in different vapour supplies, and thus unevenly distributed rainfalls under the same weather system. In addition, different underlying surface conditions can also help to form small-scale weather systems and then lead to local rainfall.

However, due to the lack of any other related weather data, particularly the circulation data for this region, we are unable to judge which possibility is true based solely on the rainfall data.

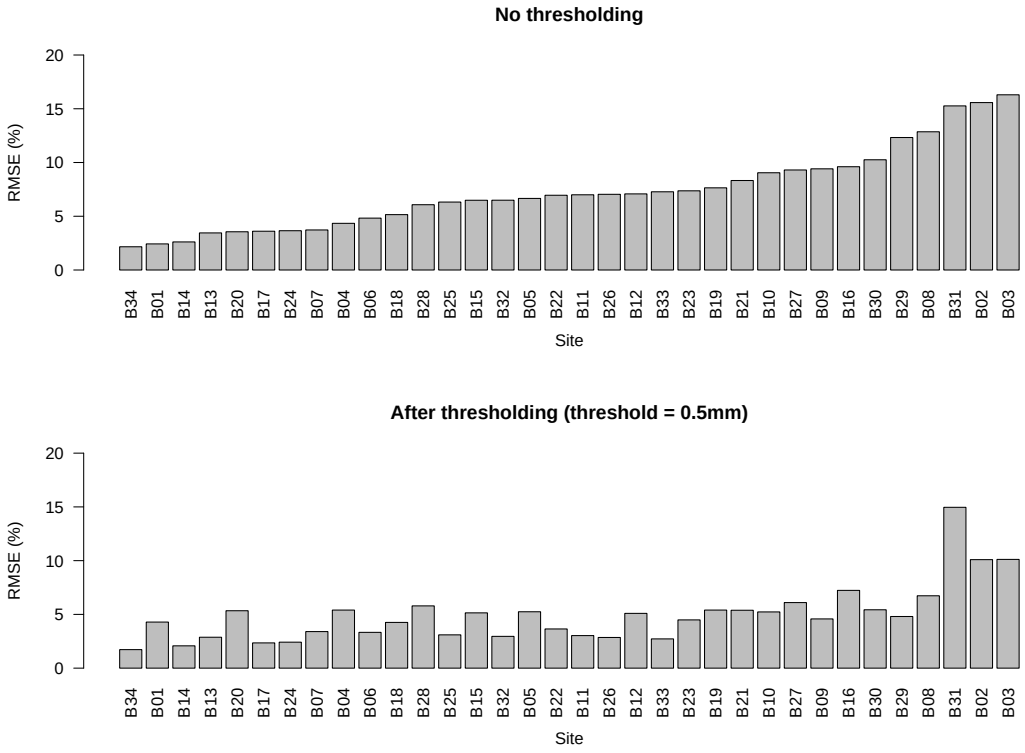


Figure 7.4 Root mean squared errors obtained at each site, when the percentage of wet days each year is predicted using the mean percentage of wet days at all sites. The time period used for the calculations is 1975–2000. Top: before thresholding, bottom: after applying a threshold of 0.5 mm. The sites are listed in the same order, and the vertical scales are the same in both plots.

7.1.2 Northeast Lancashire

In terms of data quality, the features of the northeast Lancashire dataset are very similar to those of the Blackwater. For example, there are two pairs of sites in identical locations and with identical records over the period of overlap. As with the Blackwater, the longest record from each pair has been retained and the others have been discarded. There are 39 unique series left, from sites at altitudes ranging from 83 m to 381 m above sea level. The positions of the gauges are shown in Figure 7.5.

Other features in common with the Blackwater data are as follows:

- The daily data files contain unflagged monthly totals. The problem is more widespread here than in the Blackwater data set — 36 such monthly totals have been identified in the data files and discarded from subsequent analysis.
- The recording resolution changed from 0.3 mm to 0.1 mm in the mid 1970s.
- There are spatial inconsistencies among the gauges, particularly with respect to rainfall occurrence. Once again, these can be resolved to some extent by thresholding the data. However, for this dataset there were still unexplained inconsistencies after thresholding. This may be due to the fact the relief of the area is quite varied, so that local climatic effects are more pronounced.

An additional complication with the northeast Lancashire data is that five of the 39

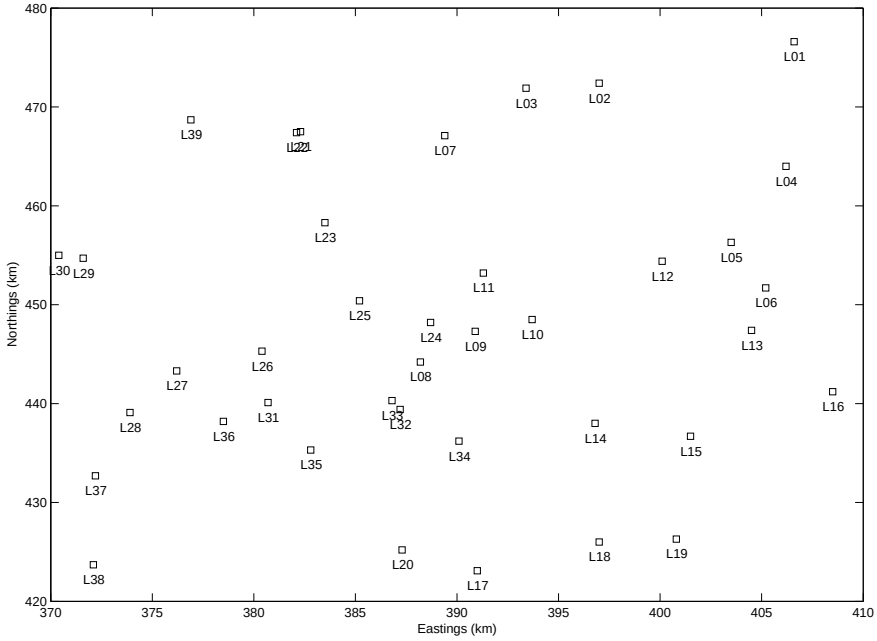


Figure 7.5 Locations of the 39 raingauges in the northeast Lancashire region.

gauges are automatically operated. whereas the remaining 34 are manual. The two recording mechanisms yield rainfall sequences with slightly different properties. Take L21/L22 as an example. L21 is an 0.2 mm tipping-bucket gauge whereas L22 is manually read each day. The two gauges are almost at the same position (see Figure 7.5). However, cumulative plots for both rainfall occurrence and amount (Figure 7.6 (a) and (b)) show apparent biases from the 45°lines. Even after a threshold at 0.5 mm has been applied, these biases still exist (Figure 7.6 (c) and (d)). This suggests that the difference in mechanism may itself contribute to the discrepancy between gauges of different types. Such problems should be taken into account when fitting models to the dataset.

The descriptions above are intended to provide an overview of the kinds of problems that may be encountered when using daily rainfall data from the UK. Some of them, such as changes in recording resolution and different gauge types, can be handled directly within the GLM framework. Thresholding is an effective way to remove spatial inconsistencies, although it cannot be expected to eliminate them completely. Of course, for continuous simulation purposes it is necessary to ensure that thresholding does not affect the important properties of the generated rainfall sequences too much. This will be investigated below.

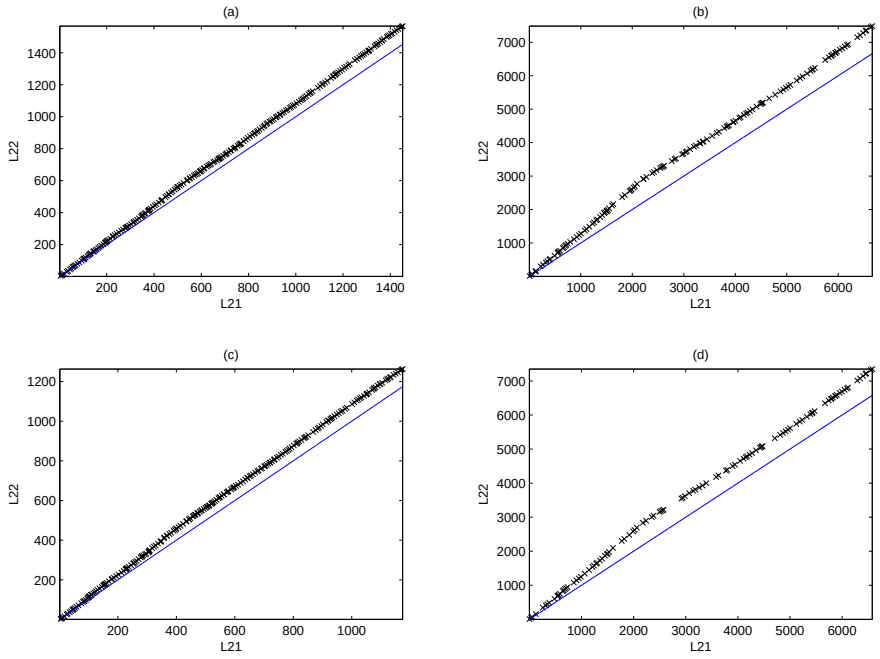


Figure 7.6 Cumulative rainfall occurrence (days) (a) and amount (mm) (b) for gauges L21/L22. (c) and (d) are the counterparts of (a) and (b) when a threshold at 0.5 mm is applied.

7.2 GLMs for the Blackwater data

7.2.1 Fitting to the raw data

When fitting GLMs to rainfall data, it is a good idea to start with a simple model incorporating covariates that are known to affect rainfall (e.g. seasonal effects). Covariates can then be added gradually; at each stage, likelihood ratio tests and informal diagnostics can be used to determine whether or not the extra complexity is justified (Chandler and Wheater, 2002; Chandler, 2005). For the Blackwater data, the covariates considered included Legendre Polynomials and their interactions to represent easting and northing site effects, site altitude, sine and cosine components to represent seasonality, temporal dependence and temporal persistence. In addition, a monthly time series of the North Atlantic Oscillation (NAO) was chosen as a potential external factor, since this is generally regarded as one of the most important indices of large-scale atmospheric circulation affecting European climate (Barnston and Livezey, 1987). The use of polynomials to represent regional affects makes an implicit assumption that these effects are smooth (i.e. that in the long run, a map showing the mean rainfall at each location in the studied region would not show any abrupt discontinuities). This seems a reasonable assumption for practical purposes, although it should be noted that a smooth surface will not be able to handle spatial inconsistencies of the type described in the previous section. Our analysis therefore assumes, in effect, that the spatial inconsistencies are hydrologically irrelevant and can be smoothed over.

Table 7.1 Summary of occurrence model fitted to data from 34 sites in the Blackwater region.

```
Main effect:                                          Coefficient  Std Err
------------                                          -----------  -------
Constant                                               -1.656555   0.0424
Legendre polynomial  1 for Eastings (OS grid, kilomet   0.238056   0.0144
Legendre polynomial  2 for Eastings (OS grid, kilomet   0.021913   0.0126
Legendre polynomial  1 for Northings (OS grid, kilome   0.139888   0.0128
Legendre polynomial  2 for Northings (OS grid, kilome  -0.029776   0.0069
Altitude (decades of metres)                            0.041876   0.0020
Indicator for Data before 1975                         -0.166660   0.0435
February indicator                                     -0.135253   0.0540
North Atlantic Oscillation (Jones et al. index)         0.019972   0.0083
Daily seasonal effect, cosine component                 0.212098   0.0204
Daily seasonal effect, sine component                   0.024927   0.0203
Mean of I(Y[t-1]>0)                                     1.917443   0.0546
Mean of I(Y[t-2]>0)                                     0.610765   0.0540
Mean of I(Y[t-3]>0)                                     0.263004   0.0349
Mean of I(Y[t-k]>0: k=1 to  2)                         -0.525234   0.0755

2-way interactions:                                   Coefficient  Std Err
-------------------                                   -----------  -------
Legendre polynomial  2 for Eastings (OS grid, kilomet  -0.242761   0.0166
with Legendre polynomial  1 for Northings (OS
Legendre polynomial  2 for Eastings (OS grid, kilomet  -0.074485   0.0151
with Legendre polynomial  2 for Northings (OS
Indicator for Data before 1975                          0.489272   0.0616
with Mean of I(Y[t-1]>0)
North Atlantic Oscillation (Jones et al. index)         0.050678   0.0115
with Daily seasonal effect, cosine component

Parameters in non-linear transformations:
-----------------------------------------
Legendre polynomial  1 for Eastings (OS grid, kilomet:
          Lower limit for polynomial rep:  460.0000 (prespecified)
          Upper limit for polynomial rep:  510.0000 (prespecified)
Legendre polynomial  1 for Northings (OS grid, kilome:
          Lower limit for polynomial rep:  130.0000 (prespecified)
          Upper limit for polynomial rep:  170.0000 (prespecified)

Spatial dependence structure:
-----------------------------
Structure used is Beta-Binomial distribution for number of wet sites
          Shape parameter of Beta-Binomi:    0.4600
```

Table 7.1 shows the GLIMCLIM output for an occurrence model fitted to the raw Blackwater data. The occurrence model is a logistic regression as defined by equation (6.1) (page 104). The linear predictor (see page 104) for the model contains 19 terms including the constant term; four of the terms are interactions. Notice the following:

- There is an interaction between NAO and seasonality in the model. This interaction term appears highly significant, when the magnitude of the coefficient (0.051) is compared with its estimated standard error (0.012). This reflects the well-known fact that the NAO affects European climate more strongly in winter than in summer (Hurrell, 1995).
- An indicator variable is used to allow for different resolutions of the data before and after 1975. This takes the value 1 before 1975 (when the recording resolution was 0.3 mm) and zero afterwards. Effectively, this allows the model to shift the probabilities of rain prior to 1975. The negative coefficient suggests reduced rainfall occurrence prior to 1975. This is to be expected, since the coarse recording resolution at this time would have led to more small rainfall amounts being recorded as zero.
- There is an interaction between the pre-1975 indicator and the term `Mean of I(Y[t-1]>0)` (this is the proportion of sites experiencing rain on the previous day). This is because the change in recording resolution leads to a change in the interpretation of a 'wet' day; hence, if the proportion of wet sites is used as a covariate in the model, its effect is expected to change correspondingly. The same reasoning applies to the proportion of wet sites 2 and 3 days previously; however, interactions between these terms and the pre-1975 indicator did not lead to significant improvements in model fit and were therefore omitted in order to keep the model relatively simple.

As in Chandler and Wheater (2002) and Yang et al. (2006b), a simple check of the model structure is to plot mean Pearson residuals by month, site and year. For any GLM, the Pearson residual for the ith case in a database is defined as

$$r_i^{(P)} \propto \frac{Y_i - \mu_i}{\sigma_i} \ , \tag{7.1}$$

where Y_i, μ_i and σ_i are respectively the observed value, the mean of the distribution predicted by the GLM and the standard deviation of this distribution. If the model is adequate, all Pearson residuals should come from distributions with zero mean and the same variance. To illustrate this, the left-hand plot in Figure 7.7 shows the magnitudes of mean Pearson residuals at each site. It is noticeable that at most sites, the mean Pearson residuals differ significantly from zero at the 95% level, although there is no systematic structure. For example, the means at B29 and B30 are both significantly different from zero, but that at B29 is negative whereas that at B30 is positive. It is likely that this is due to the spatial inconsistencies that have already been noted with these data. Note, for example, that both sites B02 and B03 have very large mean residuals — these were identified previously as being different from the remainder of the sites in terms of rainfall occurrence.

An amounts model, of the form (6.2), has also been fitted to this data set. It contains fewer covariates than the occurrence model, mainly owing to simpler representation of

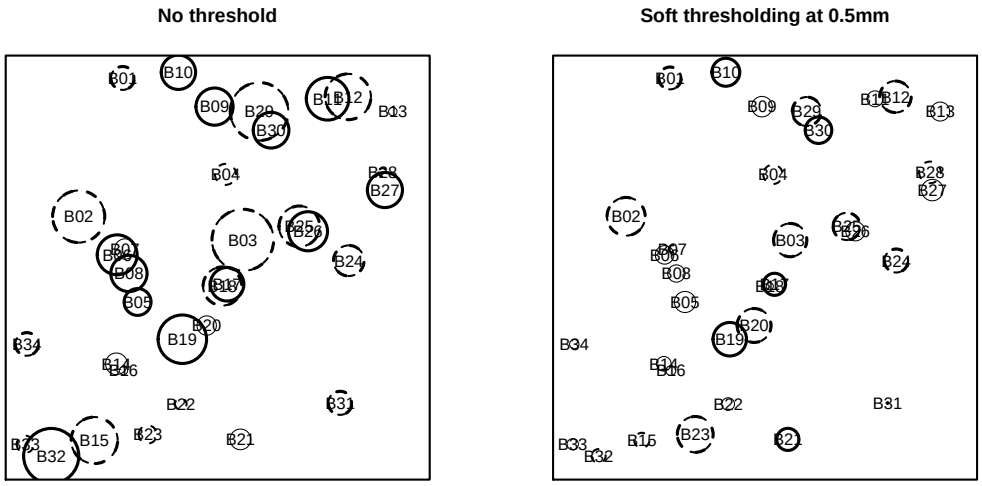


Figure 7.7 Mean Pearson residuals for occurrence model fitted to raw data (left) and data after applying a soft threshold (right). Solid and dashed circles indicate positive and negative mean residuals, respectively. Thick lines indicate mean residuals that differ significantly from zero at the 5% level. Symbol sizes are directly comparable between plots.

regional effects. The mean Pearson residuals at each site are shown in the left-hand plot of Figure 7.8. Again, there are many sites where the mean residual differs significantly from zero, but there is no apparent pattern.

It is clear from the left-hand plots in Figures 7.7 and 7.8 that the fitted GLMs are not able to capture the spatial variability in either rainfall occurrence or amounts. As noted previously, this is not entirely surprising since the representation of regional variation in the models is smooth. In view of the concerns noted earlier regarding the possibility that the spatial inconsistencies may be due to data quality problems, it is worth trying to take account of this when modelling. We have investigated two possible approaches to this, as follows:

1. Fit models to thresholded data.
2. Discard those gauges with the largest discrepancies from the overall mean condition, using the remaining subset of data to fit the model.

Combinations of the two methods are also worth trying. In order to find a simple and reasonable method to make up the spatial inconsistencies and to obtain a better model performance, we tried different combinations of data in fitting the same model, i.e. keeping the set of predictors unchanged for comparison. An extensive comparison exercise was undertaken. The results can be summarised as follows:

1. The estimated coefficients, and residual performance, of the occurrence model did not change much when two or three gauges with large discrepancies were discarded from the analysis.

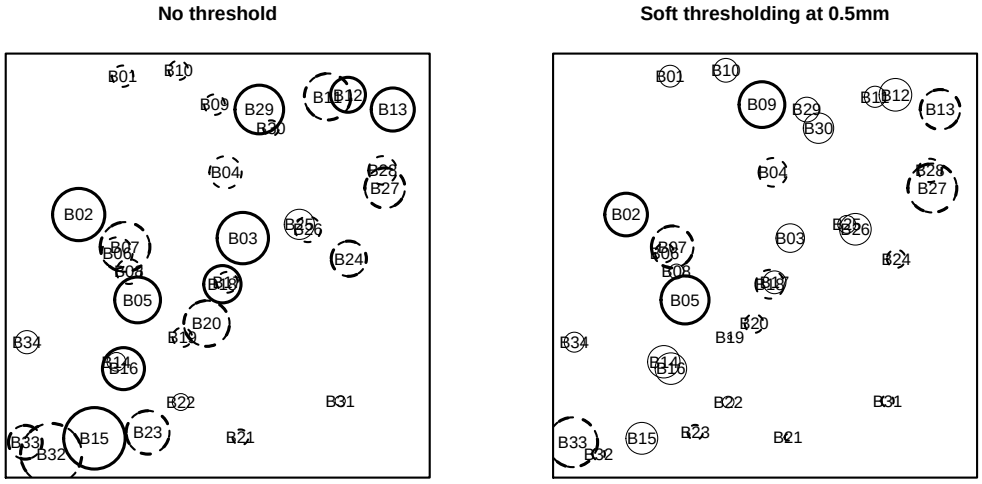


Figure 7.8 Mean Pearson residuals for amounts model fitted to raw data (left) and data after applying a soft threshold (right). Solid and dashed circles indicate positive and negative mean residuals, respectively. Thick lines indicate mean residuals that differ significantly from zero at the 5% level. Symbol sizes are directly comparable between plots.

2. The occurrence model changed quite a lot when fitted to data with a threshold of 0.5 mm applied. Moreover, the mean residuals at each site tended to be much smaller (see the right-hand panel of Figure 7.7).

3. The amounts model was relatively insensitive to both the subset of sites used in the fitting, and to the imposition of the threshold.

The conclusion from all these analysis is that applying a threshold of 0.5 mm for this dataset provides a reasonable resolution of the spatial inconsistency problem. In this case, all the gauges can be included in the dataset, and the model-building task is simplified to some extent because there is no need to try and represent complex site-to-site variation that is mainly associated with very small values. From now on therefore, we focus on models fitted to thresholded data.

7.2.2 Fitting to thresholded data

As far as rainfall occurrence is concerned, the analysis and simulation of thresholded data is straightforward: any non-zero rainfalls below the threshold are set to zero and counted as'dry' days in the logistic regression model (6.1). For rainfall amounts, however, a decision has to be made regarding how to treat the values above the threshold. The GLIMCLIM software offers two alternatives:

'Soft' thresholding: if the original variable of interest is Y and the threshold is τ then models are fitted to Y^*, where $Y^* = 0$ if $Y < \tau$, $Y - \tau$ otherwise.

'Hard' thresholding: if the original variable of interest is Y and the threshold is τ then models are fitted to Y^*, where $Y^* = 0$ if $Y < \tau$, Y otherwise.

The terminology is borrowed from the literature on wavelets.

To determine which of the two thresholding strategies is most appropriate, it is useful to consider the implications for simulation of models fitted to thresholded data:

- For soft thresholding, after simulation the threshold is added back to any nonzero values. Therefore the simulated output will contain no values between zero and the threshold.
- For hard thresholding, no correction is made. In particular, positive values below the threshold are not set to zero, so that the simulated output may contain values between zero and the threshold. This may be seen as a problem. However, from a modelling perspective it is certainly the correct thing to do because, in this case, non-zero values are modelled on the scale of the original data. If the fitted model is reasonable, the problem will be negligible. Conversely, if the problem is non-negligible then the fitted model is unreasonable. If this is the case, 'soft' thresholding should be used instead.

An additional consideration is that the gamma distribution is intended as a model for variables that can take any positive value. Since hard thresholded data do not contain any observations between zero and the threshold, there is a danger of introducing bias when fitting gamma distributions to hard thresholded data. Soft thresholding has therefore been used here throughout. The threshold used is $\tau = 0.5$ mm.

The final occurrence model, fitted to thresholded data, is summarised in Table 7.2. The covariates used are exactly the same as those in Table 7.1 (in fact, the covariates were selected on the basis of thresholded data; for the purposes of this report, the same model was then refitted to raw data purely to demonstrate the effect of thresholding). A comparison of the two tables shows that the coefficients relating to site effects (the Legendre polynomials) are quite different in the two models, as is the estimated shape parameter of the beta-binomial spatial dependence model (see Section 6.2.2). The remaining coefficients are reasonably similar, however. The thresholding has therefore affected the spatial structure of the model (which is to be expected, since it was designed to remove the spatial inconsistencies in the data), but little else.

Table 7.3 gives a summary of the corresponding model for rainfall amounts. This has been built up from a simple model, independently of the occurrence model in Table 7.2. The covariates are quite similar, however. The main differences are in the representation of previous days' rainfalls (the amounts model uses the average rainfall on previous days, whereas the occurrence model used the proportion of wet sites), and in the representation of regional variation which is simpler in the amounts model — in fact, the fitted model indicates a very simple north-south gradient in rainfall amounts here, represented by a quadratic function of northings (note also that site altitude was not found to be significantly associated with rainfall amounts).

As a check on the fitted models, the means and standard deviations of Pearson residuals have been examined by year, month and site. The mean residuals at each site were presented in Figures 7.7 and 7.8: the lack of any systematic structure here

Table 7.2 Summary of occurrence model fitted to thresholded data from 34 sites in the Blackwater region.

```
Main effect:                                          Coefficient  Std Err
------------                                          -----------  -------
Constant                                               -1.667627   0.0409
Legendre polynomial  1 for Eastings (OS grid, kilomet   0.091191   0.0144
Legendre polynomial  2 for Eastings (OS grid, kilomet   0.036959   0.0125
Legendre polynomial  1 for Northings (OS grid, kilome   0.008330   0.0126
Legendre polynomial  2 for Northings (OS grid, kilome   0.006044   0.0071
Altitude (decades of metres)                            0.023006   0.0019
Indicator for Data before 1975                         -0.113156   0.0410
February indicator                                     -0.143939   0.0569
North Atlantic Oscillation (Jones et al. index)         0.016703   0.0084
Daily seasonal effect, cosine component                 0.181196   0.0211
Daily seasonal effect, sine component                   0.024493   0.0207
Mean of I(Y[t-1]>0)                                     1.868315   0.0540
Mean of I(Y[t-2]>0)                                     0.850038   0.0501
Mean of I(Y[t-3]>0)                                     0.263687   0.0346
Mean of I(Y[t-k]>0: k=1 to  2)                         -0.866629   0.0743

2-way interactions:                                   Coefficient  Std Err
-------------------                                   -----------  -------
Legendre polynomial  2 for Eastings (OS grid, kilomet  -0.043513   0.0164
with Legendre polynomial  1 for Northings (OS
Legendre polynomial  2 for Eastings (OS grid, kilomet   0.030764   0.0147
with Legendre polynomial  2 for Northings (OS
Indicator for Data before 1975                          0.203808   0.0619
with Mean of I(Y[t-1]>0)
North Atlantic Oscillation (Jones et al. index)         0.056701   0.0119
with Daily seasonal effect, cosine component

Parameters in non-linear transformations:
-----------------------------------------
Legendre polynomial  1 for Eastings (OS grid, kilomet:
          Lower limit for polynomial rep:  460.0000 (prespecified)
          Upper limit for polynomial rep:  510.0000 (prespecified)
Legendre polynomial  1 for Northings (OS grid, kilome:
          Lower limit for polynomial rep:  130.0000 (prespecified)
          Upper limit for polynomial rep:  170.0000 (prespecified)

Global quantities:
------------------
          'Soft' threshold for +ve value:    0.4500

Spatial dependence structure:
-----------------------------
Structure used is Beta-Binomial distribution for number of wet sites
          Shape parameter of Beta-Binomi:    0.3564
```

Table 7.3 Summary of amounts model fitted to thresholded data from 34 sites in the Blackwater region.

```
Main effect:                                          Coefficient  Std Err
------------                                          -----------  -------
Constant                                                1.376381   0.0202
Legendre polynomial  1 for Northings (OS grid, kilome  -0.176408   0.0074
Legendre polynomial  2 for Northings (OS grid, kilome   0.043523   0.0064
North Atlantic Oscillation (Jones et al. index)        -0.017764   0.0080
Daily seasonal effect, cosine component                -0.151680   0.0268
Daily seasonal effect, sine component                  -0.161528   0.0175
Ln(1+Mean of Y[t-1])                                    0.124163   0.0140
Ln(1+Mean of Y[t-2])                                    0.048250   0.0151
Ln(1+Mean of Y[t-3])                                    0.027155   0.0146
I(Y[t-k]>0: k=1 to  3)                                 -0.063775   0.0268

2-way interactions:                                   Coefficient  Std Err
-------------------                                   -----------  -------
North Atlantic Oscillation (Jones et al. index)         0.027704   0.0111
with Daily seasonal effect, cosine component
Daily seasonal effect, cosine component                 0.055925   0.0180
with Ln(1+Mean of Y[t-1])
Daily seasonal effect, cosine component                 0.097103   0.0192
with Ln(1+Mean of Y[t-2])

Parameters in non-linear transformations:
-----------------------------------------
Legendre polynomial  1 for Northings (OS grid, kilome:
          Lower limit for polynomial rep:  130.0000 (prespecified)
          Upper limit for polynomial rep:  170.0000 (prespecified)

Global quantities:
------------------
          'Soft' threshold for +ve value:    0.4500

Spatial dependence structure:
-----------------------------
Structure used is Constant correlation between all pairs of sites
          Inter-site correlation        :    0.7861
```

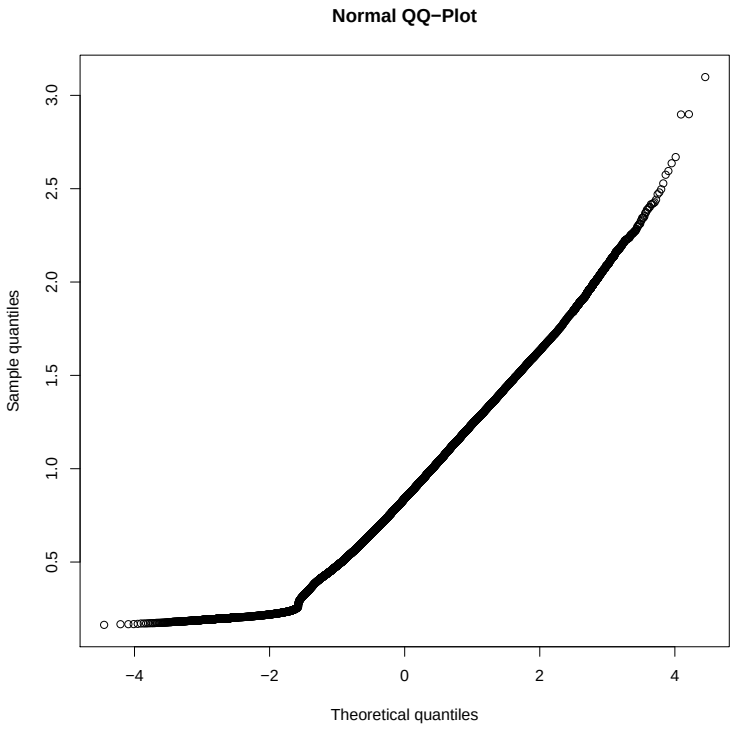


Figure 7.9 Normal probability plot of Anscombe residuals from amounts model fitted to thresholded data from 34 gauges in the Blackwater region.

suggests that the models have done a good job of capturing any systematic regional variation in rainfall over the area. Results by month and site can be found in Figure 2 of Yang et al. (2006b). The only obvious pattern is a tendency for the residual standard deviations to be slightly higher in summer than in winter; this could be rectified, if necessary, by relaxing the assumption of a constant shape parameter in the rainfall amounts models (see Section 6.1). For reasons of brevity, we do not present these plots here: residual plots are, however, presented for the northeast Lancashire study in Section 7.3 below.

As well as checking that the fitted models capture the systematic structure in rainfall sequences, it is necessary to verify the assumption that rainfall amounts are drawn from gamma distributions. As suggested in Chandler and Wheater (2002), this can be achieved using a normal probability plot of the Anscombe residuals defined in equation (6.3) (page 106). This plot is shown in Figure 7.9. Most of the data points fall almost exactly on a straight line, implying that the Anscombe residuals are almost exactly normally distributed over most of their range. The sharp deviation at the left-hand end of the plot is associated with small rainfall values, and is exactly as expected when approximating a positive-valued random variable using a normal distribution (see Chandler and Wheater 2002, for example). This plot therefore confirms that the gamma distribution provides a good fit to the observations.

7.2.3 Simulations

To see if the fitted rainfall models can reproduce the dynamics of the rainfall process, occurrence and amounts models fitted to both thresholded and unthresholded data have been used to simulate rainfall. Note that the overall properties of the resulting sequences are related rather indirectly to the specification of the models themselves (Yang et al., 2006b; Yan et al., 2006); hence it is necessary to carry out checks of this type to ensure that the models are suitable for use in hydrological applications. Various statistics for different site combinations (whole region average, a cluster of some closely located sites and some individual sites) have been calculated to check the model performances. To generate dependent multi-site sequences, the beta-binomial dependence structure (Section 6.2) has been used for rainfall occurrence, and a constant inter-site correlation between Anscombe residuals has been used for amounts. For the occurrence model fitted to thresholded data, the estimated beta-binomial shape parameter is 0.3564 (see Table 7.2). The distributions of numbers of wet sites will therefore look like those in the top row of Figure 6.1 (page 108).

This choice of a constant inter-site residual correlation in rainfall amounts is made primarily for simplicity, and because at this spatial scale all correlations are high (very few of them are below 0.75). If this were deemed to be too simplistic in any particular application, and alternative would be to use the exact residual correlations observed between each pair of sites, or to model the correlations as a function of inter-site distance as described in Section 6.2. We have carried out some exercises comparing the results presented here with those obtained using the observed residual correlation structure. In practical terms, the results were indistinguishable; we conclude that at the spatial scale of the Blackwater region, the precise choice of residual correlation structure is relatively unimportant.

In order to make meaningful comparison with observations, 10 sites were selected that had few missing values over the period 1961–1999. These sites are shown in blue in Figure 7.2. For each fitted model, the missing values were imputed 10 times by the model using conditional simulations to simulate the missing values, i.e., the fitted model generated random numbers conditional on all the observations for the missing values only. Free simulations (simulations for the whole rainfall series over the period unconditional on the given observations) were then carried out 100 times to simulate the modeled dynamics of the rainfall processes. For each realisation, summary statistics (the mean and standard deviation of whole series, proportion of wet days, the mean and standard deviation of wet days, and autocorrelations at lags 1–3) have been calculated for the average of 10 gauges, a cluster of 3 gauges (blue triangles in Figure 7.2) and some individual gauges. The 10 imputations therefore yield 10 values for each summary statistic; the variability of these gives an idea of the uncertainty due to missing data. The 100 free simulations yield a distribution of each summary statistic. The observed values (which should lie somewhere in the envelope from the 10 imputations) can be compared with this distribution — if the models are plausible, the observed values should look like a sample from the simulated distributions. All the calculations were carried out using soft thresholded and unthresholded models respectively for comparison.

Figure 7.10 shows some results for the average of 10 gauges, obtained from models fitted to unthresholded data. Figure 7.11 shows the statistics for a cluster of 3 closely

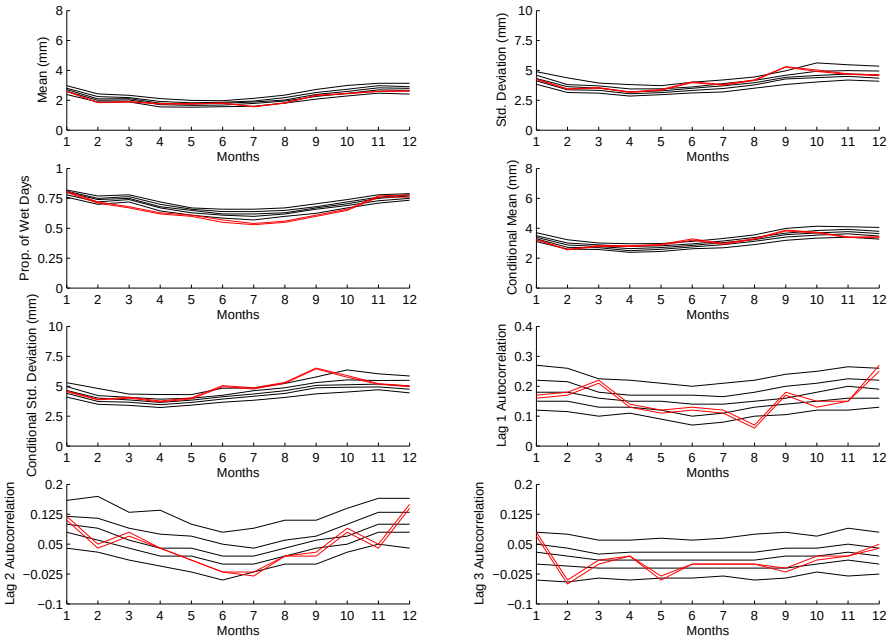


Figure 7.10 Observed statistics (red lines, envelope from 10 imputations) and percentiles (5, 25, 50, 75, 95) for simulated statistics (black lines) from simulations of models fitted to unthresholded data, for a 10-site average. Left to right, top to bottom: mean, standard deviation, proportion of wet days, mean rainfall on wet days, standard deviation on wet days, autocorrelations at lags 1–3.

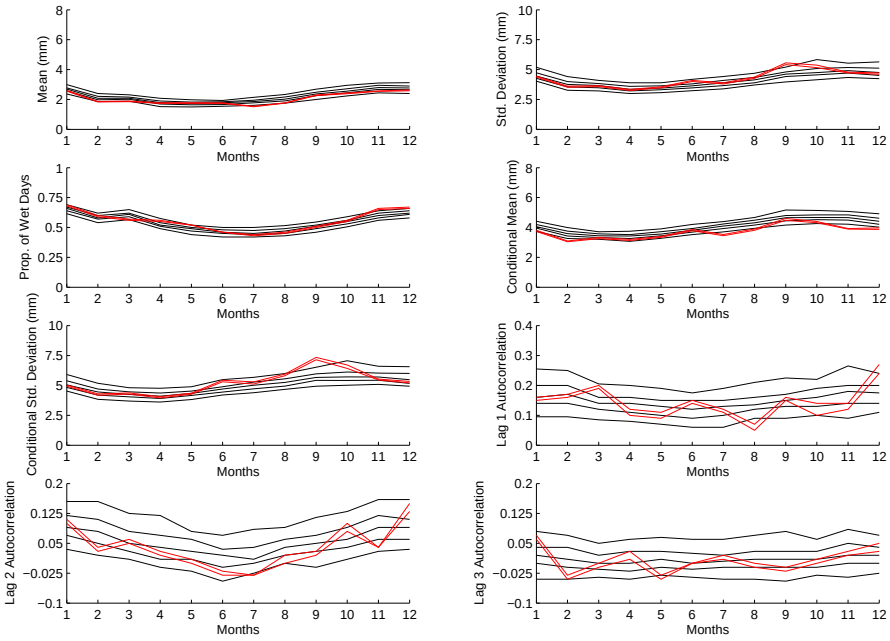


Figure 7.11 Observed statistics (red lines, envelope from 10 imputations) and percentiles (5, 25, 50, 75, 95) for simulated statistics (black lines) from simulations of models fitted to unthresholded data, for a cluster of 3 sites. The statistics are the same as those in Figure 7.10.

located gauges (B23, B25, B27). Similar plots have been produced for individual gauges; the results are very similar, and are not shown here.

Figures 7.12 and 7.13 are comparable with Figures 7.10 and 7.11, but are obtained using models fitted to soft thresholded data. Comparing these with the previous two figures, it is clear that the thresholding has not introduced any significant biases. The proportion of wet days is reduced (as expected), but the effect on the other summary statistics is virtually indistinguishable. It is also clear from both sets of figures that the simulated distributions look reasonable — with few exceptions, the observed statistics fall within the bands of the simulations.

To illustrate the models' ability to reproduce extremes, Figure 7.14 shows the observed time series of annual maxima for the average of 10 gauges, along with the corresponding simulated distributions. Figure 7.15 shows the corresponding plot for models fitted to thresholded data. Again, there is generally good agreement between the observations and the simulated distributions (note that if the simulated distributions are correct, 5% of observations are expected to fall above the 95th percentile, which looks reasonable from the figures); and again, it is clear that the thresholding has not introduced any biases.

Further analyses of GLM simulation performance for the Blackwater region can be found in Yang et al. (2006b), where it is shown that the GLMs provide a good representation of rainfall variability at seasonal time scales (e.g. 3-month rainfall totals). It is also shown the extremal properties of GLM simulations are in good agreement with a conventional extreme value analysis of the data.

Having established that the GLMs provide good reproduction of a variety of rainfall statistics in the Blackwater area, they have also been used in conjunction with the spatial-temporal models of Chapter 9 to generate nonstationary rainfall sequences that are continuous in both space and time. The results of this exercise were discussed in Chapter 10.

7.3 GLMs for the northeast Lancashire data

A similar modelling exercise has been carried out for the contrasting northeast Lancashire data set (recall that this is an upland region with much more topographic variability than the Blackwater). Once again, we have examined spatial inconsistencies by fitting models to both unthresholded and thresholded data. In this case, however, the presence of both automatic and manual gauges (see Section 7.1) is an additional source of potential inconsistency that can be controlled for by introducing an indicator variable for automatic/manual gauges. This indicator takes the value 1 for an automatic gauge and zero otherwise. Its function is very similar to the 'pre-1975' indicator that was used with the Blackwater data to account for the change in recording resolution.

In qualitative terms, the results for this data set were very similar to those for the Blackwater. We therefore do not give so much detail here. In particular, we only report the results of fitting to thresholded data. Once again, soft thresholding was used with a threshold of 0.5 mm.

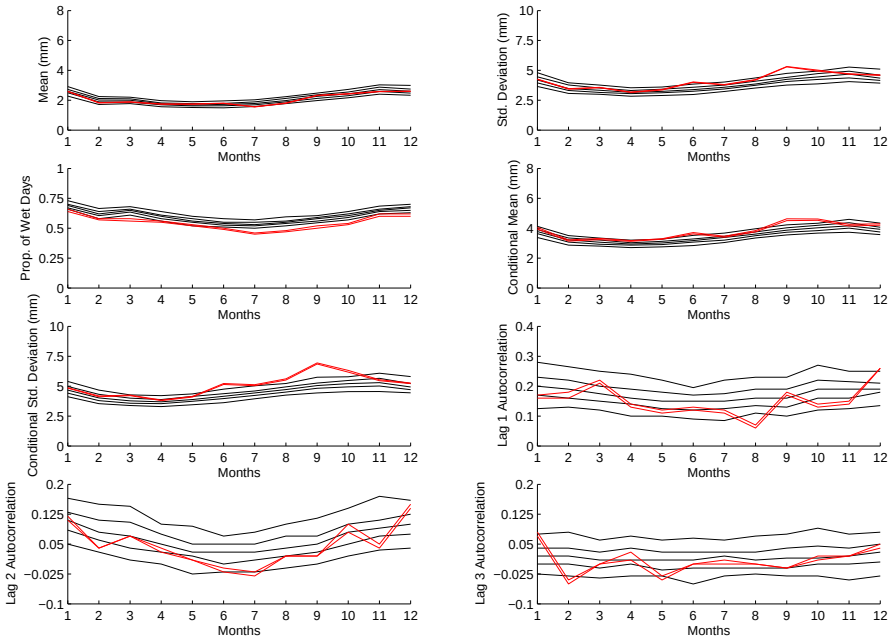


Figure 7.12 Observed statistics (red lines, envelope from 10 imputations) and percentiles (5, 25, 50, 75, 95) for simulated statistics (black lines) from simulations of models fitted to thresholded data, for a 10-site average. The statistics are the same as those in Figure 7.10.

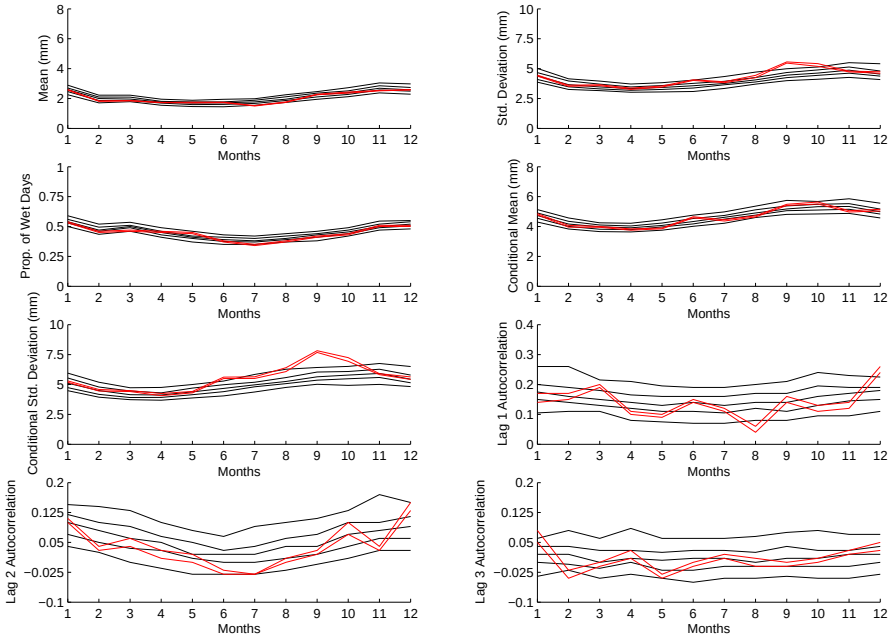


Figure 7.13 Observed statistics (red lines, envelope from 10 imputations) and percentiles (5, 25, 50, 75, 95) for simulated statistics (black lines) from simulations of models fitted to thresholded data, for a cluster of 3 sites. The statistics are the same as those in Figure 7.10.

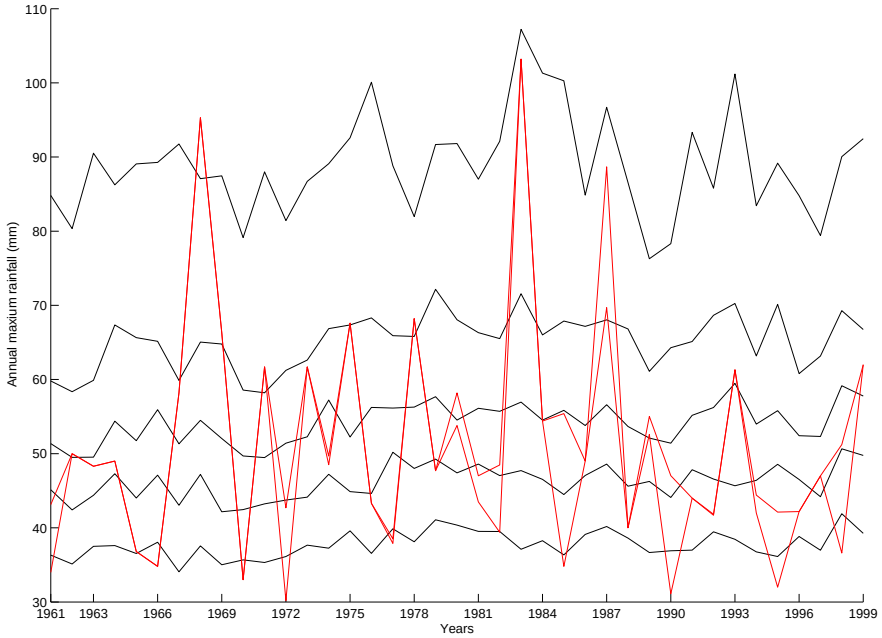


Figure 7.14 Time series of observed annual rainfall maxima (red lines, envelope from 10 imputations) and percentiles from simulated maxima obtained from models fitted to unthresholded data (black lines), for a 10-site average.

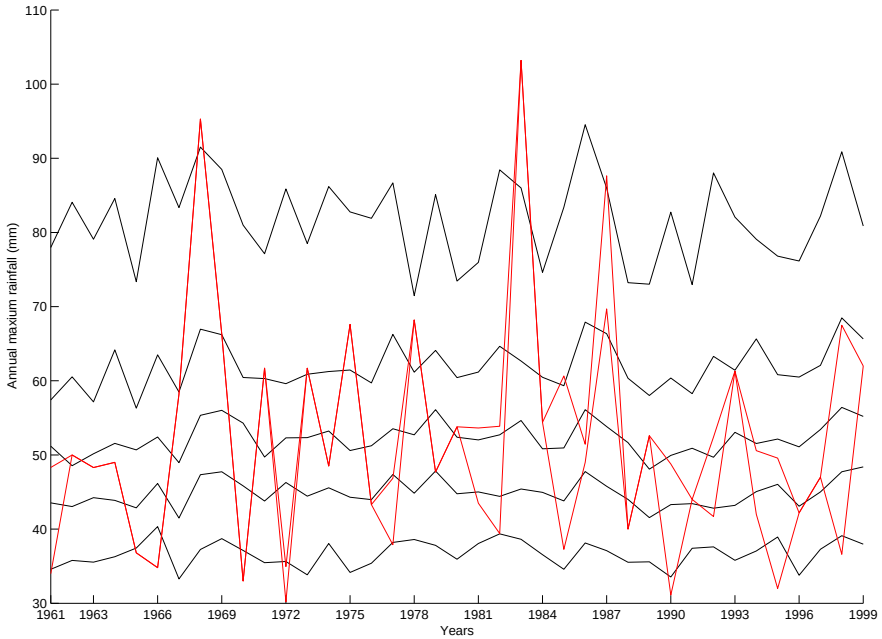


Figure 7.15 As Figure 7.14, but using models fitted to soft-thresholded data.

7.3.1 Fitting to thresholded data

The final models for occurrence and amounts are summarised in Tables 7.4 and 7.5. Covariates chosen include Legendre polynomials and their interactions to represent regional variation, site altitude, an indicator for automatic gauges, a long-term linear trend, sine and cosine components to represent seasonality, and temporal dependence. In addition, two climate indices, North Atlantic Oscillation (NAO) and Arctic Oscillation (AO) were chosen as external factors. In both models, the indicator for automatic gauges is indispensable — the negative coefficients suggest that automatic gauges do not record rain as often as manual ones. This is probably associated with the tipping-bucket mechanism of the automatic gauges.

It is interesting to compare the covariates in these models with those obtained for the Blackwater region (Tables 7.2 and 7.3). Overall, the model structures are very similar — for example, the occurrence models both involve site altitude, a seasonal cycle, an indicator variable for February (which does not fit well into the overall seasonal cycle), the NAO and the proportion of wet sites up to 3 days previously. The main differences are in the representation of regional variability, in particular for the amounts models. The regional variability in amounts for northeast Lancashire is quite complex, requiring cubic polynomials in both eastings and northings, as well as interactions with site altitude (the interpretation is that the effect of altitude varies over the region). This is to be expected, since the topography of the area is more complex than that of the Blackwater.

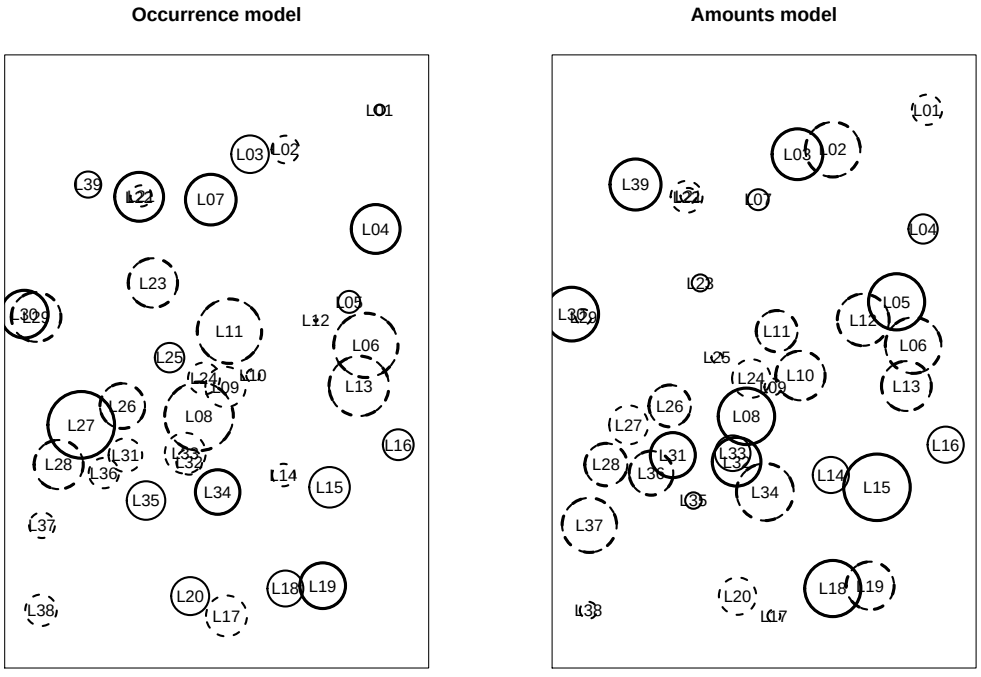


Figure 7.16 Mean Pearson residuals for occurrence model (left) and amounts model (right) fitted to thresholded data from northeast Lancashire. Solid and dashed circles indicate positive and negative mean residuals, respectively. Thick lines indicate mean residuals that differ significantly from zero at the 5% level.

Table 7.4 Summary of occurrence model fitted to thresholded data from 39 sites in northeast Lancashire.

Main effect:	Coefficient	Std Err
Constant	-1.458254	0.0327
Legendre polynomial 1 for Eastings (OS grid, kilomet	-0.088131	0.0084
Legendre polynomial 3 for Eastings (OS grid, kilomet	-0.092157	0.0102
Legendre polynomial 1 for Northings (OS grid, kilome	0.034922	0.0097
Legendre polynomial 3 for Northings (OS grid, kilome	-0.017361	0.0108
Altitude (decades of metres)	0.015828	0.0004
Indicator for automatic gauge	-0.198305	0.0285
North Atlantic Oscillation (Jones et al. index)	0.119389	0.0122
Arctic Oscillation (Thompson et al. index)	-0.148793	0.0225
February indicator	-0.183857	0.0668
Daily seasonal effect, cosine component	0.181198	0.0242
Daily seasonal effect, sine component	-0.058853	0.0241
Mean of I(Y[t-1]>0)	1.880166	0.0417
Mean of I(Y[t-2]>0)	0.343535	0.0453
Mean of I(Y[t-3]>0)	0.238945	0.0423

2-way interactions:	Coefficient	Std Err
Legendre polynomial 1 for Eastings (OS grid, kilomet with Legendre polynomial 3 for Northings (OS	-0.147344	0.0221
Legendre polynomial 3 for Eastings (OS grid, kilomet with Legendre polynomial 3 for Northings (OS	-0.129584	0.0298
North Atlantic Oscillation (Jones et al. index) with Daily seasonal effect, cosine component	0.081505	0.0139

```
Parameters in non-linear transformations:
-----------------------------------------
Legendre polynomial  1 for Eastings (OS grid, kilomet:
          Lower limit for polynomial rep:  370.0000 (prespecified)
          Upper limit for polynomial rep:  410.0000 (prespecified)
Legendre polynomial  1 for Northings (OS grid, kilome:
          Lower limit for polynomial rep:  420.0000 (prespecified)
          Upper limit for polynomial rep:  480.0000 (prespecified)

Global quantities:
------------------
          'Soft' threshold for +ve value:    0.4500

Spatial dependence structure:
-----------------------------
Structure used is Beta-Binomial distribution for number of wet sites
          Shape parameter of Beta-Binomi:    0.4432
```

Table 7.5 Summary of amounts model fitted to thresholded data from 39 sites in northeast Lancashire.

```
Main effect:                                          Coefficient  Std Err
------------                                          -----------  -------
Constant                                                1.327744   0.0237
Legendre polynomial  1 for Eastings (OS grid, kilomet  -0.060686   0.0065
Legendre polynomial  3 for Eastings (OS grid, kilomet  -0.206457   0.0102
Legendre polynomial  1 for Northings (OS grid, kilome   0.256902   0.0223
Legendre polynomial  2 for Northings (OS grid, kilome   0.075928   0.0086
Legendre polynomial  3 for Northings (OS grid, kilome   0.326758   0.0298
Altitude (decades of metres)                            0.007408   0.0004
Indicator for automatic gauge                          -0.073228   0.0209
North Atlantic Oscillation (Jones et al. index)         0.018709   0.0070
Daily seasonal effect, cosine component                -0.118165   0.0326
Daily seasonal effect, sine component                  -0.108276   0.0168
Ln(1+Mean of Y[t-1])                                    0.144700   0.0126
Ln(1+Mean of Y[t-2])                                    0.050361   0.0130
Ln(1+Mean of Y[t-3])                                    0.030747   0.0126

2-way interactions:                                   Coefficient  Std Err
-------------------                                   -----------  -------
Legendre polynomial  1 for Eastings (OS grid, kilomet   0.265727   0.0152
with Legendre polynomial  2 for Northings (OS
Legendre polynomial  3 for Eastings (OS grid, kilomet  -0.289681   0.0269
with Legendre polynomial  2 for Northings (OS
Legendre polynomial  3 for Eastings (OS grid, kilomet   0.178025   0.0226
with Legendre polynomial  3 for Northings (OS
Legendre polynomial  1 for Northings (OS grid, kilome  -0.009238   0.0008
with Altitude (decades of metres)
Legendre polynomial  3 for Northings (OS grid, kilome  -0.013228   0.0013
with Altitude (decades of metres)
Altitude (decades of metres)                            0.000856   0.0004
with Daily seasonal effect, cosine component
North Atlantic Oscillation (Jones et al. index)         0.026544   0.0095
with Daily seasonal effect, cosine component
Daily seasonal effect, cosine component                 0.079181   0.0176
with Ln(1+Mean of Y[t-1])
Daily seasonal effect, cosine component                 0.033685   0.0176
with Ln(1+Mean of Y[t-2])

Parameters in non-linear transformations:
-----------------------------------------
Legendre polynomial  1 for Eastings (OS grid, kilomet:
          Lower limit for polynomial rep:  370.0000 (prespecified)
          Upper limit for polynomial rep:  410.0000 (prespecified)
Legendre polynomial  2 for Northings (OS grid, kilome:
          Lower limit for polynomial rep:  420.0000 (prespecified)
          Upper limit for polynomial rep:  480.0000 (prespecified)

Global quantities:
------------------
          'Soft' threshold for +ve value:    0.4500

Spatial dependence structure:
-----------------------------
Structure used is Constant correlation between all pairs of sites
          Inter-site correlation        :    0.7422
```

Figure 7.16 shows mean Pearson residuals at each site, for both occurrence and amounts models. In both cases, many of the sites have mean residuals that differ significantly from zero at the 5% level (indicated by thick lines on the plots). The performance in this respect is worse than for the Blackwater data, and for the amounts model the results are little better than those for models fitted to unthresholded data (not shown here). However, the results are certainly no worse; and for the occurrence model the thresholding once again yields substantial improvements. Moreover, it is difficult to see any systematic structure in either of the plots in Figure 7.16 (e.g. neighbouring sites L18 and L19 have large residuals of opposite sign in the amounts model). We conclude that the thresholding is still worthwhile, although in this case it does not completely account for the spatial inconsistencies.

To illustrate the use of Pearson residuals to check seasonal structure and long-term trends, Figures 7.17 and 7.18 show the residual means and standard deviations for both occurrence and amounts models, by month and by year. For the occurrence model the performance is excellent — there is no structure in the mean plots, and the standard deviations are almost exactly as expected under the model. For the amounts model, there is no obvious seasonal structure in the residuals, but there is a suggestion of a negative trend in both the means and standard deviations. This could be accounted for, if necessary, by including a linear trend term in the model.

7.3.2 Simulation

As with the Blackwater example, we have tested the models' suitability for hydrological applications by simulation, to see if they can reproduce the dynamics of the rainfall process. The 18 longest series were selected for this exercise. Once again, 10 sets of imputations have been used to examine uncertainty in the historical values of rainfall summary statistics, and 100 free simulations have been used to obtain simulated distributions for the rainfall properties. The properties examined are the same as those for the Blackwater; they have been calculated for an average of all 18 gauges, an average over a cluster of 3 gauges and for some individual gauges.

Results are only presented for models fitted to thresholded data. Figure 7.19 provides a comparison between observed and simulated statistics for the average of all 18 gauges — these statistics can be regarded as representative of the catchment average rainfall series. Figure 7.20 shows the same statistics for a cluster of 3 gauges in the centre of the study region; and Figures 7.21 and 7.22 show the statistics for two individual gauges (L05 and L39).

The plots again show generally good agreement between the observations and simulated distributions, although there are some isolated instances of poor performance (e.g. the standard deviation of wet day rainfall amounts at site L05 in July and August). Overall the performance is encouraging, and we conclude that although there are spatial inconsistencies in the dataset, the fitted models are able to reproduce the dynamics of rainfall in northeast Lancashire. A possible explanation for this is that although the spatial resolution of the dataset is not high enough to reflect the local weather characteristics in the area, the fitted models are still able to capture the large-scale properties.

Further results from this case study are reported in Wheater et al. (2006).

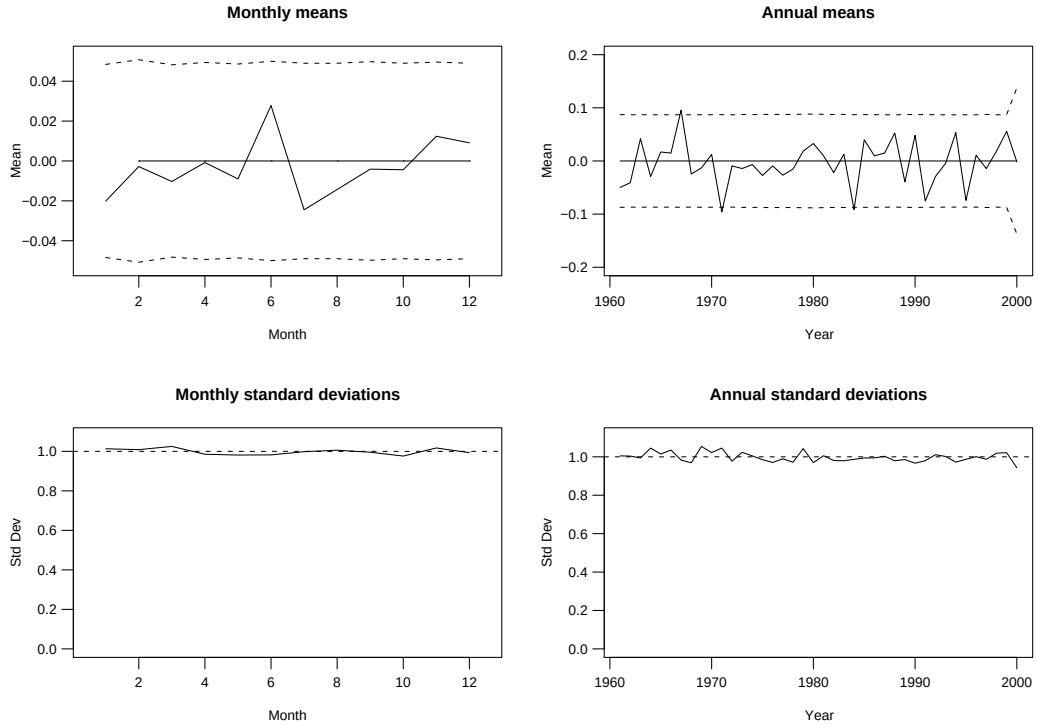


Figure 7.17 Means and standard deviations of Pearson residuals for the northeast Lancashire occurrence model, by month and year. In the top plots, dashed lines show 95% limits, adjusted for intersite dependence as described by Chandler (2002), under the assumption that the model is correct. In the bottom plots, dashed lines show the standard deviations expected under the model.

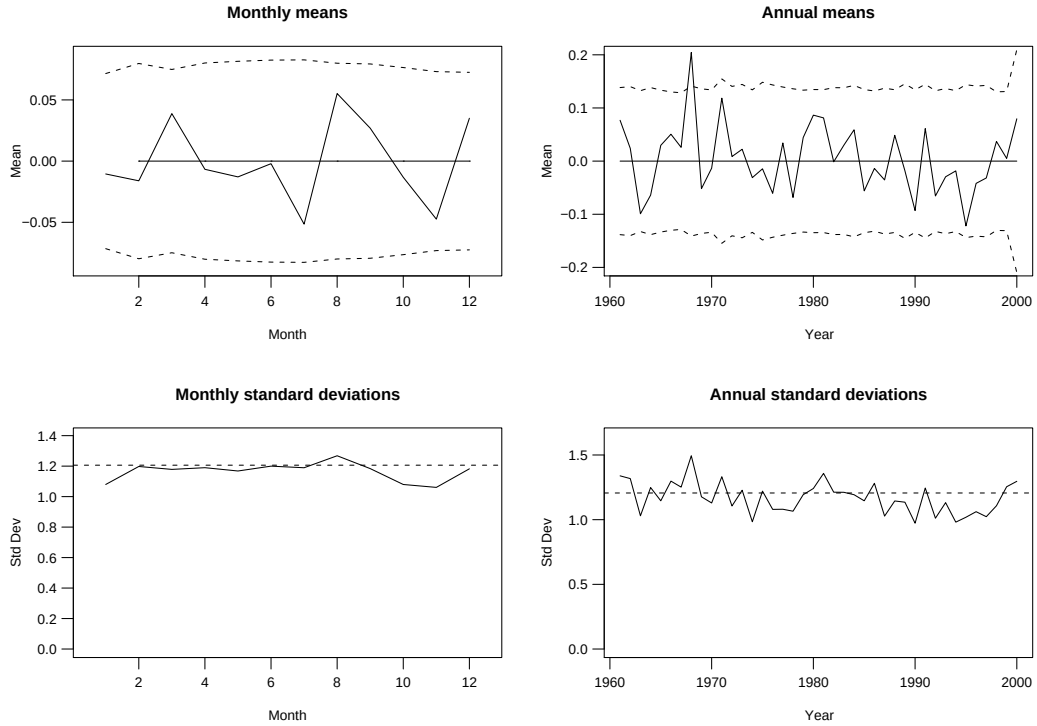


Figure 7.18 Means and standard deviations of Pearson residuals for the northeast Lancashire amounts model, by month and year. See caption to Figure 7.17 for further explanation.

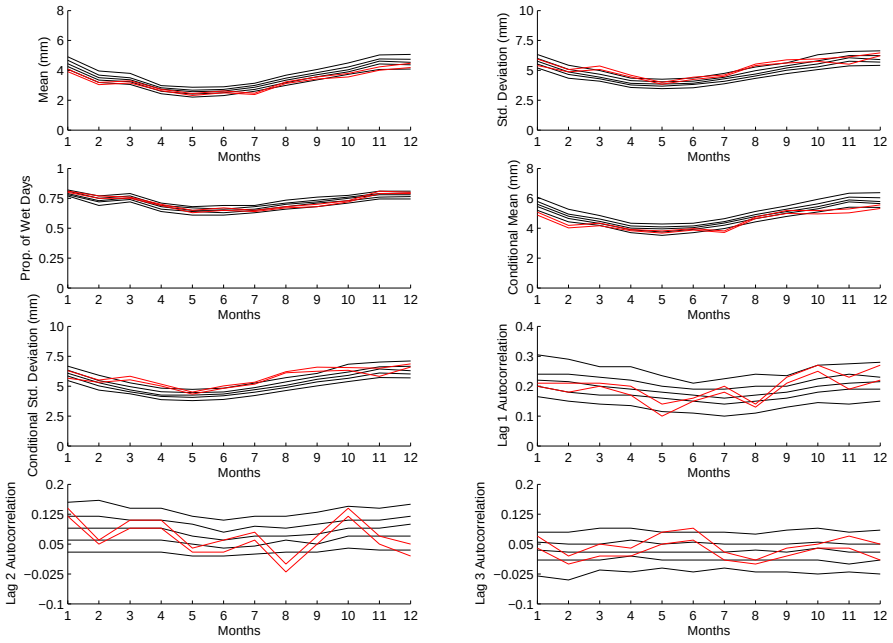


Figure 7.19 Observed statistics (red lines, envelope from 10 imputations) and percentiles (5, 25, 50, 75, 95) from simulations of models fitted to thresholded data, for an 18-site average. The statistics are the same as those in Figure 7.10 (page 128).

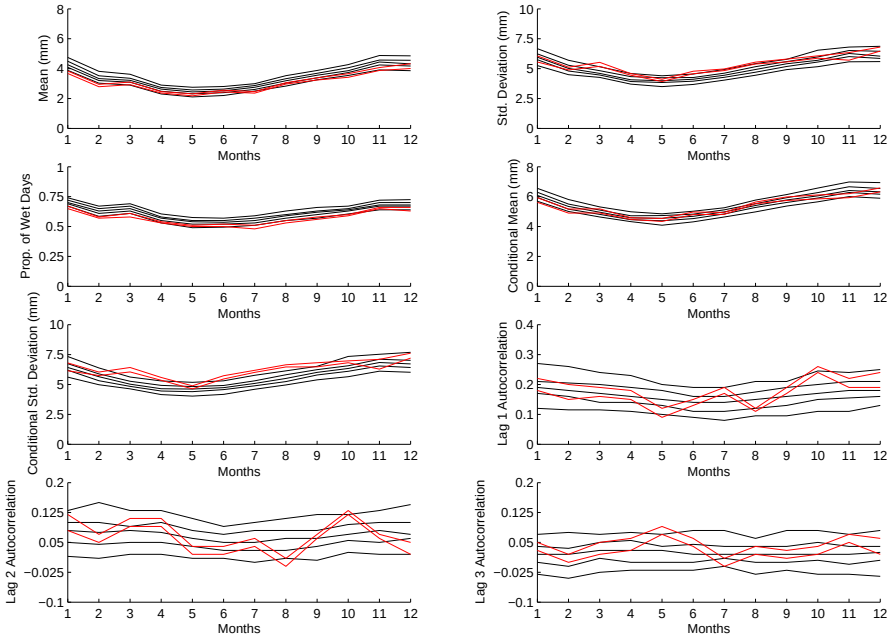


Figure 7.20 Observed statistics (red lines, envelope from 10 imputations) and percentiles (5, 25, 50, 75, 95) from simulations of models fitted to thresholded data, for a 3-site average (sites L05, L09 and L24). The statistics are the same as those in Figure 7.10 (page 128).

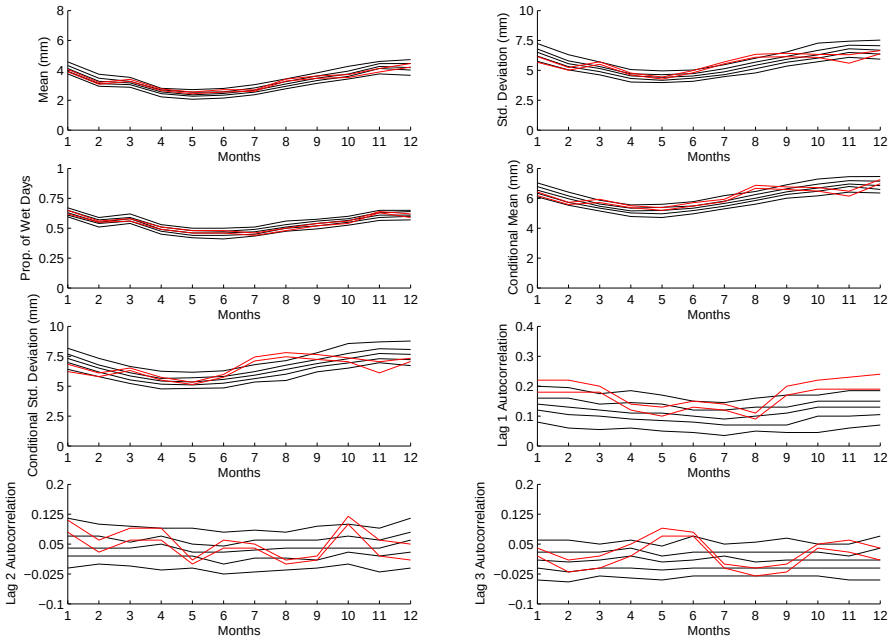


Figure 7.21 Observed statistics (red lines, envelope from 10 imputations) and percentiles (5, 25, 50, 75, 95) from simulations of models fitted to thresholded data, for site L05. The statistics are the same as those in Figure 7.10 (page 128).

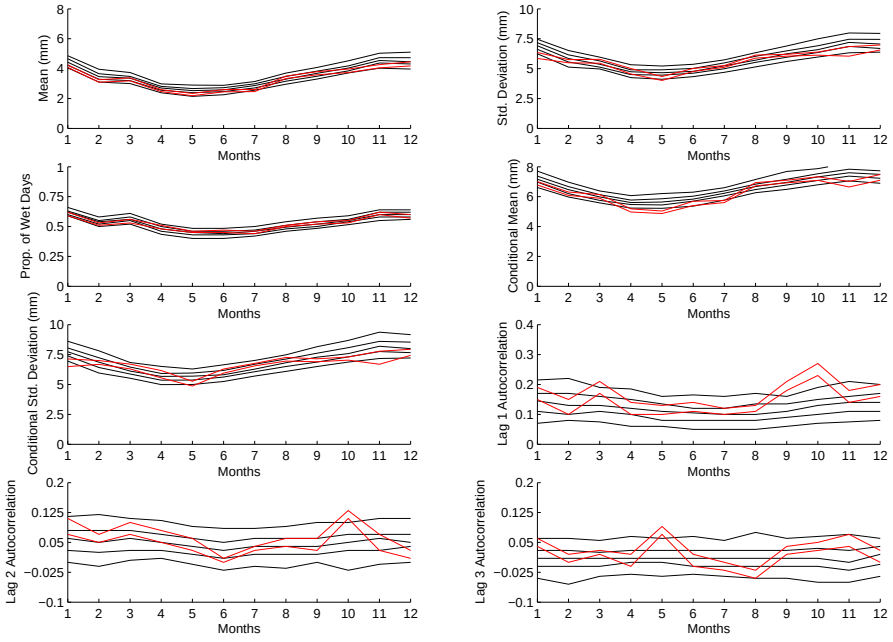


Figure 7.22 Observed statistics (red lines, envelope from 10 imputations) and percentiles (5, 25, 50, 75, 95) from simulations of models fitted to thresholded data, for site L39. The statistics are the same as those in Figure 7.10 (page 128).

7.4 Summary of chapter

This chapter has illustrated the use of GLMs to model and simulate daily rainfall sequences, and has also highlighted some issues that are likely to be encountered when using daily raingauge data from the UK. The results from the two extensive and contrasting case studies were very similar, which gives some confidence that they can be used as a guide to what can be expected when applying the methodology at other locations in the UK.

The main points to emerge from this chapter are as follows:

- Daily rainfall records in the UK should be scrutinised carefully before use. Duplicate records should be removed to avoid biasing any analyses, as should unflagged monthly totals. A further discussion of data quality issues can be found in Yang et al. (2006a).
- After excluding records that are clearly erroneous, daily rainfall records can exhibit apparent spatial inconsistencies. These affect rainfall occurrence in particular, and are in part caused by differences in the characteristics of very small rainfall amounts. The reason for these differences is unknown. Having experimented with a variety of different schemes, and compared simulation results, it seems that thresholding the data using a soft threshold of 0.5 mm can help to remove inconsistencies without affecting the overall rainfall properties to any appreciable degree. Although such thresholding does not always remove the inconsistencies completely, it does seem to reduce them. As far as rainfall occurrence is concerned, this reduction can be substantial. We therefore recommend thresholding daily rainfall in this way as a matter of course, prior to analysis.
- The GLM framework provides a natural and interpretable means of incorporating systematic structures into simulated rainfall sequences. As well as modelling the dynamics of the rainfall process itself, corrections can be made for features such as changes in recording resolution and different gauge types. Such features can, and should, be accommodated via the use of indicator variables.
- In both of the case studies considered here, the fitted models had very similar structures. The main differences were in the representation of regional variation, which is more complex for northeast Lancashire than for the Blackwater. This is to be expected, since northeast Lancashire has a more complex topography. The models for the two regions were built independently of each other. Their similarity therefore suggests that the covariates identified here may be suitable for use in other areas of the UK.
- Pearson residuals can be used to check the systematic structure of the fitted models. The GLIMCLIM software calculates the mean and standard deviation of Pearson residuals by month, site and year; plots of these can be used to highlight any structure that is not captured by the models.
- GLM simulations are able to reproduce a wide variety of properties of observed rainfall sequences, over a range of spatial scales (scales up to around 2500 km^2

— the size of the northeast Lancashire region — have been considered here). These include extremal properties as well as means, variances, proportions of wet days and so on. Further results can be found in Yang et al. (2006b) and Wheater et al. (2006).

- The GLM for the Blackwater data has been combined with the spatial-temporal model described in Chapter 9, to generate nonstationary rainfall simulations in continuous space and time.

Elsewhere in the FD2105 project, GLMs have been fitted to additional data sets. Some results can be found in Chapters 8 and 13. In general, the performance is promising. After extensive testing therefore, we conclude that the methodology is suitable for routine use in the UK.

Chapter 8 Spatial-temporal disaggregation of daily rainfall from a GLM

The GLMs described in the previous chapters can be used to generate multi-site daily rainfall sequences. For many hydrological applications, however, subdaily rainfalls are required. One way to generate these is by combining a GLM with the spatial-temporal stochastic model, as described in Chapter 10. However, it has already been noted that this requires the use of radar data, which are problematic in terms both of quality and availability. As an alternative therefore, in this chapter we explore a simpler method for generating the required subdaily sequences. The method combines the daily output from a GLM with a single-site disaggregation model based on Poisson cluster processes (Koutsoyiannis and Onof, 2001).

8.1 Methodology

The GLM methodology has been discussed extensively in previous chapters. The new feature introduced here is to combine this with the HYETOS single-site disaggregation model (Koutsoyiannis and Onof, 2000, 2001). This single-site generator produces hourly totals that are consistent with the daily amounts. In the implementation described here, the single-site generator is applied at one of the sites of interest, and the resulting sub-daily temporal pattern is then applied uniformly in space (in a method similar to that presented in Woods and Sivapalan (1999)) to daily rainfall simulated from a GLM. The main steps of the algorithm are summarised as follows and can be seen in Figure 8.1:

- Generate a sequence of daily rainfall across a network of sites using a GLM.
- Taking a cluster of wet days, defined as a series of consecutive wet days delimited by at least one dry day, use a single-site model to simulate the hourly rainfall at one gauge, referred to as the master gauge, and disaggregate using an adjusting procedure, so that the hourly totals sum up to the daily totals. The single-site model used in this work is the random-parameter Bartlett-Lewis model described in Chapter 2.
- Use the hourly pattern generated at the master gauge to disaggregate the daily information to sub-daily rainfall at the other gauges. If the master gauge records zero rainfall, the previous day's temporal profile is applied.

The method is assessed using rainfall records from a network of 21 raingauges from the 1400km^2 Lee catchment in the Thames region, UK, for the period 1987-2002. Daily data for the raingauge network and sub-daily data from a master gauge are used to develop the model. The simulated hourly rainfall fields are then compared with the observed sub-daily data from the other gauges to test the model. A map of the Lee catchment together with the raingauge locations is included in Figure 8.2 and Table 8.1. The raingauge elevations vary between 20 and 150m.

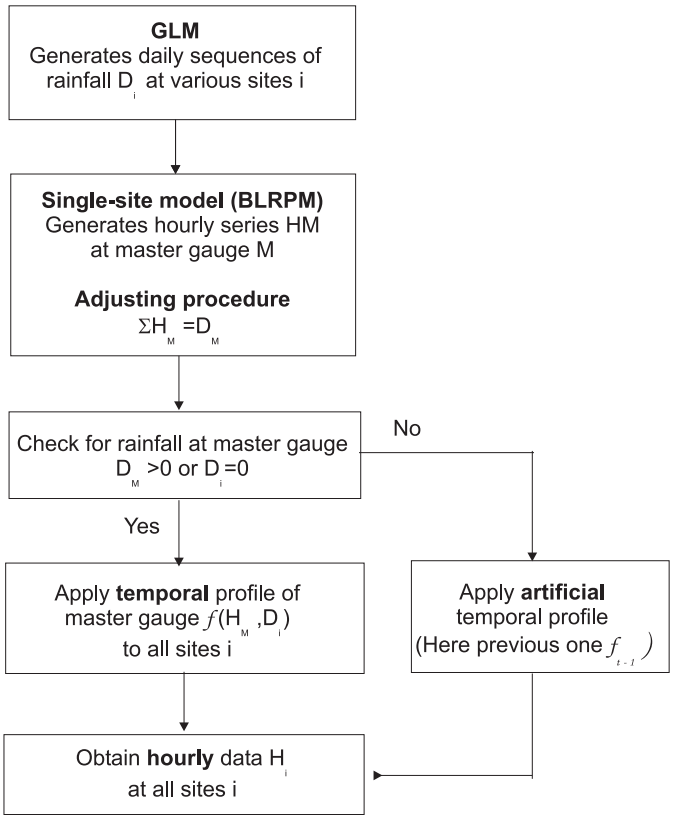


Figure 8.1 Schematic diagram of the proposed disaggregation scheme.

8.2 Simulation of daily rainfall

8.2.1 Model fitting

GLMs have been fitted to daily data from the 21 rain gauges listed in Table 8.1, for the 1987-2002 period. Few gauges have records starting in 1987, and most of them were introduced during 1990; hence most of the comparisons hereafter are run for the 1991-2002 period. A simple model structure is expected when fitting the model to this relatively short time-scale.

The fitted occurrence model is a logistic regression, as defined on page 104. It contains 14 covariates, and is summarised in Table 8.2. Prior to including any systematic regional effects, the bubble map in Figure 8.3 shows the mean Pearson residuals at each site. The plot suggests a possible variation along the x-coordinate which is then introduced in the model by a Fourier representation of site Eastings. Site altitude is also included. The map for the final model in Figure 8.4 indicates that the residuals are more randomly distributed in space, suggesting that the site effects in the model have captured most of the patterns in the area. The model contains a linear trend to improve the performance by year; apart from this, the structure is similar to

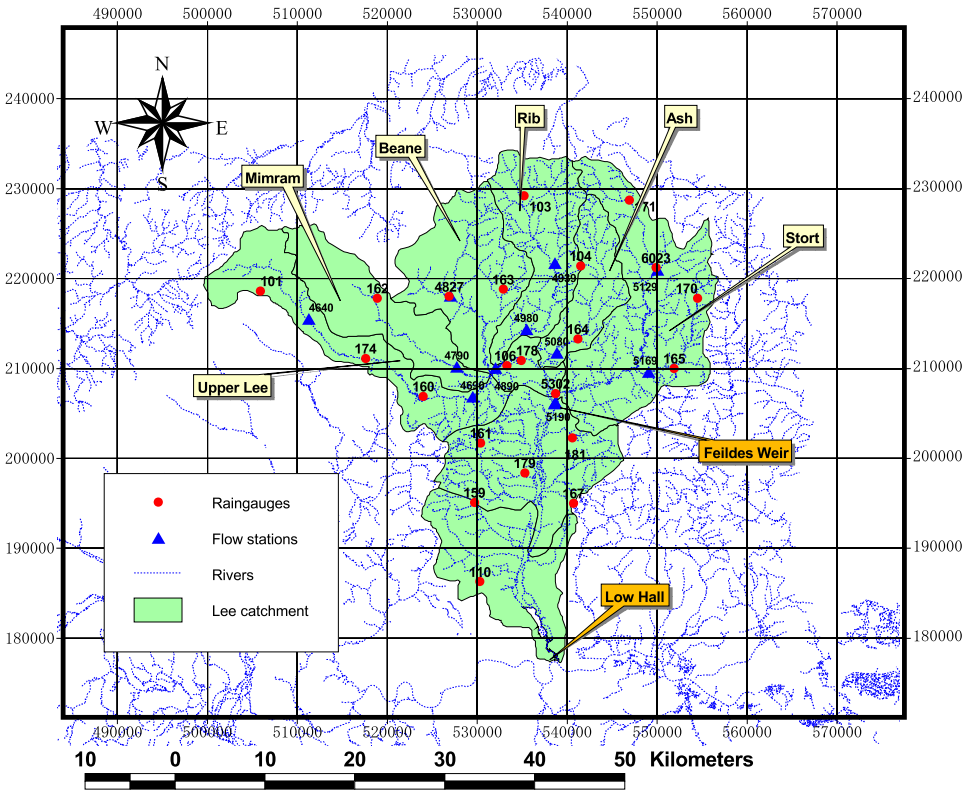


Figure 8.2 Map of the Lee catchment

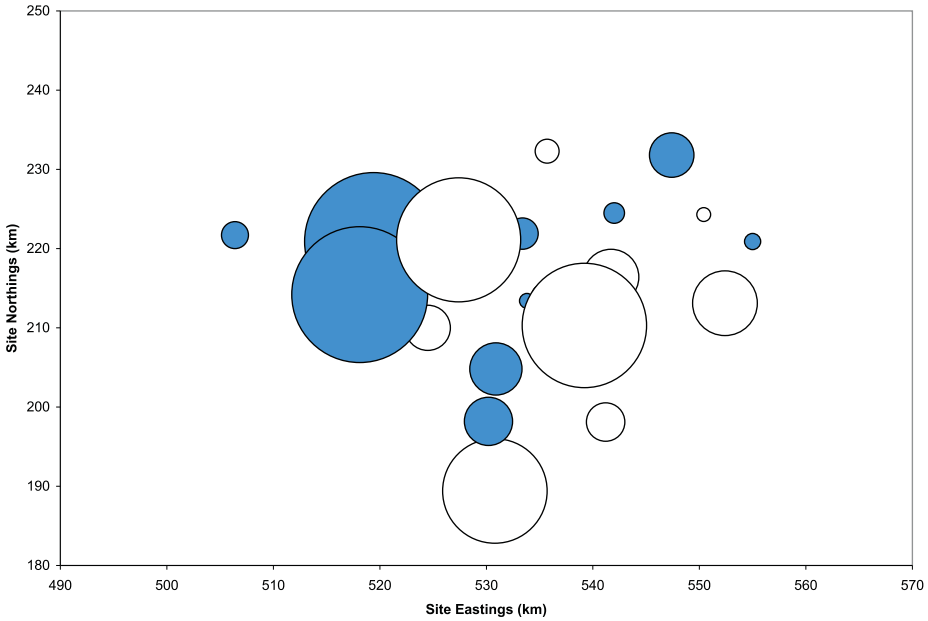


Figure 8.3 Bubble map showing mean Pearson residuals from fitted occurrence model at each site, prior to incorporation of site effects. Positive and negative mean residuals are represented in blue and white respectively.

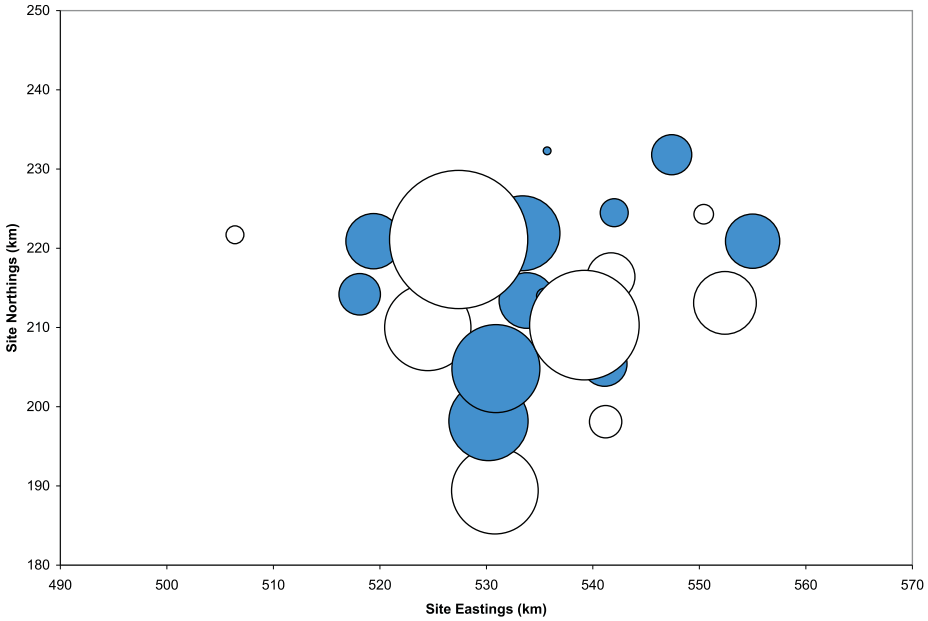


Figure 8.4 Bubble map showing mean Pearson residuals from fitted occurrence model at each site, after incorporation of site effects. Positive and negative mean residuals are represented in blue and white respectively.

Table 8.1 Details of rain gauges in the Lee catchment. The highlighted gauge (number 0163) is used as the master gauge in the analyses that follow.

Site ID	Name	Easting	Northing	Start	End
0101	Runley Wood	506400	221700	1988	2002
0103	Chipping	535700	232300	1988	2002
0104	Braughing Friars	542000	224500	1988	2002
0106	Hertford	533800	213400	1988	1994
0110	Hornsey	530800	189400	1988	2002
0159	Hadley	530200	198200	1990	2002
0160	Mill Green	524500	210000	1989	2002
0161	Darnicle Hill	530900	204800	1990	2002
0162	Whitwell	519400	220900	1989	2002
0163	**Dane End**	**533400**	**221900**	**1989**	**2002**
0164	Widford	541700	216400	1990	2002
0165	Hatfield Heath	552400	213100	1990	2002
0167	Epping	541200	198100	1989	2002
0170	Takeley	555000	220900	1990	2002
0171	Clavering	547400	231800	1990	2002
0174	Wheathampstead	518100	214200	1991	2002
0178	Broadmeads	535400	214000	1994	2002
0181	Nazeing	541100	205400	1997	2002
4827	Bragbury Park	527400	221100	1988	2002
5302	Rye Meads	539200	210300	1990	2002
6023	Stansted Argus	550400	224300	1987	2002

that of the models described in Chapter 7. Note also the use of the beta-binomial scheme for incorporating spatial dependence into rainfall occurrence. The estimated shape parameter (0.3059) for the beta-binomial distributions is very similar to that obtained for the Blackwater data set in Section 7.2.

The fitted amounts model uses gamma distributions, as defined on page 104. It is summarised in Table 8.3. The covariates are similar to those in the occurrence model, but fewer in number. The reduction is mainly due to simpler site effects, implying that rainfall amounts are less affected by spatial variations within the region than rainfall occurrence. This result agrees with that from the Blackwater case study in Section 7.2. Additionally, the amounts model contains only one day's previous rainfall. Prior to including any systematic regional effects, the mean Pearson residuals at each site (Figure 8.5) suggest a possible southwest-northeast gradient, which can be represented as a planar surface. Legendre polynomials of degree 1 in each direction have been used to represent this pattern. A map for the final model is given in Figure 8.6; as with the occurrence model, the residuals now seem randomly distributed in space.

8.2.2 Model checks

As in chapter 7, Pearson residuals have been used to check the models' representation of systematic seasonal and annual structure. Figure 8.7 shows residual plots by month, site and year, for both occurrence and amounts models. Since almost all of the mean residuals lie within the 95% limits, we conclude that the fitted models succeed in reproducing the seasonal and regional structure, along with trends in the

Table 8.2 Covariates used in occurrence model for daily rainfall in the Lee catchment.

Components	Values	
Constant	-2.3966	
Fourier sine component 1 for Eastings	0.0274	1
Fourier cosine component 1 for Eastings	-0.0591	2
Fourier sine component 1 for Eastings	-0.0152	3
Fourier cosine component 1 for Eastings	-0.0521	4
Altitude(km)	0.8810	5
Trend 1 - Linear (0.1 per year)	0.1972	6
Daily seasonal effect, cosine component	0.2738	7
Daily seasonal effect, sine component	-0.0789	8
Mean of I(Y[t-1]>0)	1.6663	9
Mean of I(Y[t-2]>0)	0.7848	10
Mean of I(Y[t-3]>0)	0.2343	11
Mean of I(Y[t-k]>0: k=1 to 2)	-0.7973	12
2-way interaction: covariates 7 and 9	-0.2502	13
2-way interaction: covariates 8 and 9	0.1991	14
'Soft' threshold for +ve values	0.4500	
Shape parameter of beta-binomial distribution	0.3059	

Table 8.3 Covariates used in amounts model for daily rainfall in the Lee catchment.

Components	Values	
Constant	1.3175	
Legendre polynomial 1 for Eastings	-0.0918	1
Legendre polynomial 1 for Northings	-0.0779	2
Daily seasonal effect, cosine component	-0.1720	3
Daily seasonal effect, sine component	-0.1558	4
Mean of I(Y[t-1]>0)	0.2737	5
2-way interaction: covariates 3 and 5	0.1413	6
2-way interaction: covariates 4 and 5	-0.0439	7
'Soft' threshold for +ve values	0.4500	
Dispersion parameter	0.7528	
Observed residual correlation structure		

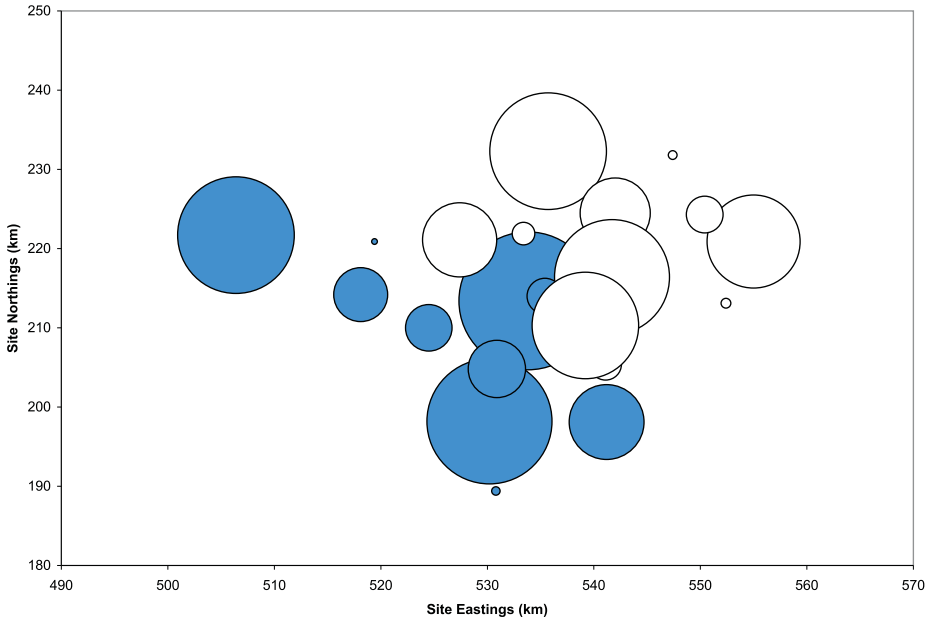


Figure 8.5 Bubble map showing mean Pearson residuals from fitted amounts model at each site, prior to incorporation of site effects. Positive and negative mean residuals are represented in blue and white respectively.

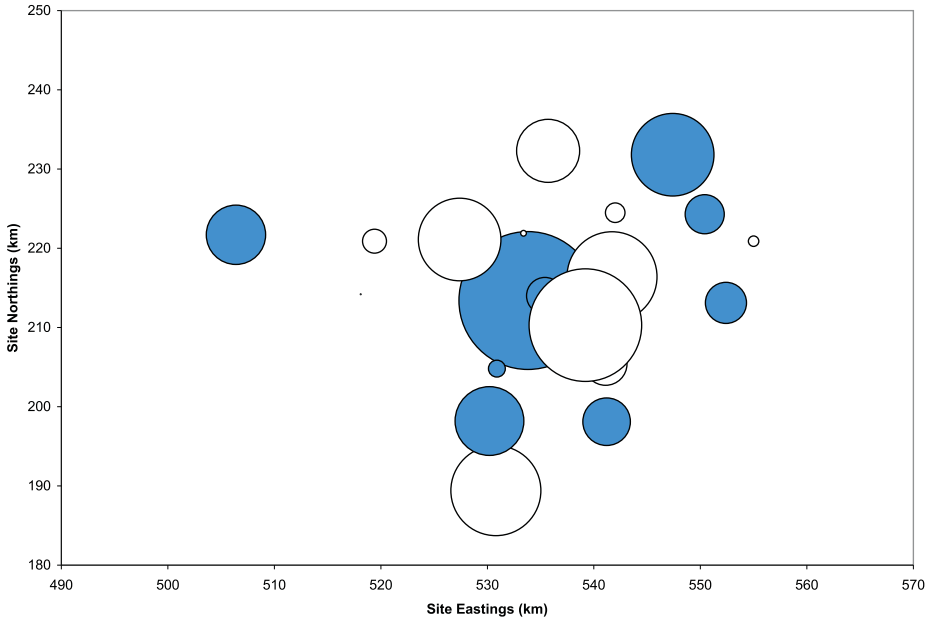


Figure 8.6 Bubble map showing mean Pearson residuals from fitted amounts model at each site, after incorporation of site effects. Positive and negative mean residuals are represented in blue and white respectively.

rainfall sequences.

To check the gamma distributional assumption of the amounts model, a normal probability plot of Anscombe residuals has been produced (see also Section 7.2). The result is shown in Figure 8.8. The shape is almost exactly the same as that for the Blackwater model on page 126 (note that the order of the axes is different on the two plots); hence, once again, we conclude that gamma distributions provide a good fit to the observed rainfall amounts.

8.2.3 Simulation

As in Chapter 7, the calibrated occurrence and amounts models have been tested further by using them to simulate simultaneous rainfall sequences at multiple sites. Conditioned on data from December 1990 to initialise the model, 40 sets of simulations were generated for the 1991-2002 period as these capture most of the sampling variability (an envelope from 40 simulations yields an approximate 95% confidence interval under the model). The performance of the simulation is assessed by comparing the properties of the simulated sequences with the historical series. Summary statistics include the mean, standard deviation, proportion of dry periods, and the lag 1 autocorrelation. At a typical site on the catchment (site 0160, 5 missing days), the results in Figure 8.9 show that the observed series behave as though they were sampled from the simulated distributions. This indicates that the fitted model provides realistic simulations even though the simulations seem a bit high for the proportion of dry periods.

The multi-site properties of the simulations are evaluated by comparing the reproduction of the distribution of wet sites and the inter-site correlations between the GLM simulations and the observed series. Results are displayed for 20 simulations since these are seen as representative of the sampling variability for these properties. Figure 8.10 shows that a spatial dependence structure defined by a beta-binomial distribution reproduces well the distribution of the number of wet sites.

As suggested in Chapter 6, spatial dependence in rainfall amounts is modelled through the inter-site correlations between Anscombe residuals. Figure 8.11 presents some boxplots of pairwise correlations between raingauges versus distance (binned in 5 km intervals), both for observed and simulated at daily and hourly time-steps. There is a clear decay of the correlations with distance here; therefore the observed inter-site correlations have been used to drive the GLM simulations rather than using a constant correlation as in Chapter 7. To illustrate the effect of the choice of correlation structure here, Figure 8.12 shows the results when a constant inter-site correlation structure is used in the simulations.

In both Figures 8.11 and 8.12, the daily correlations from the GLM simulations are underestimated, but those in Figure 8.11 show very similar decaying structure to the observed spatial correlations. When applying a spatially uniform temporal disaggregation, the simulated daily correlation is transferred directly to the hourly level and overestimates the observed hourly correlations. This is one of the limitations of the proposed scheme; the extent to which this is important with respect to the extremal behaviour is evaluated further below.

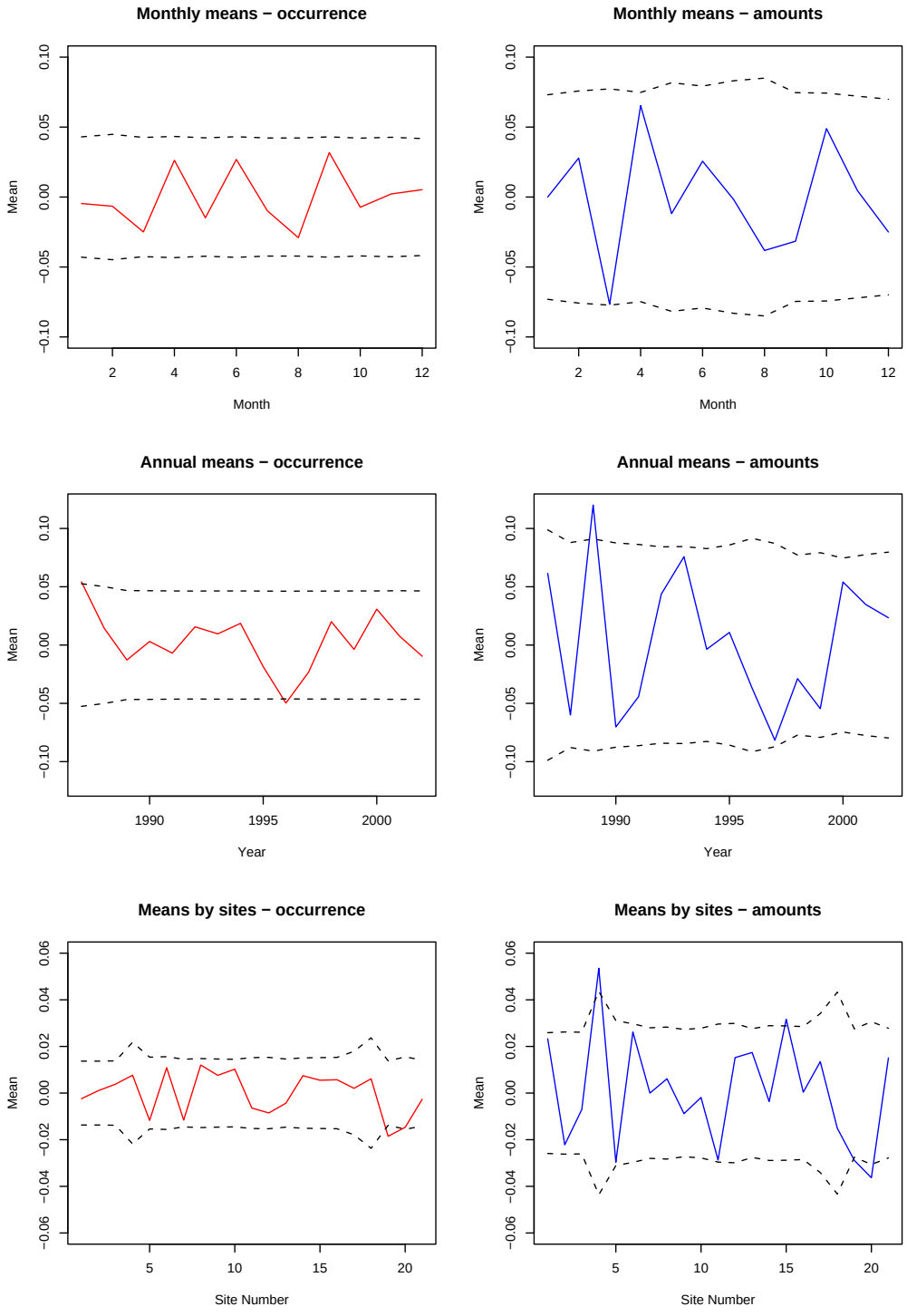


Figure 8.7 Mean Pearson residuals from occurrence and amount models. Dashed lines indicate 95% limits under the model.

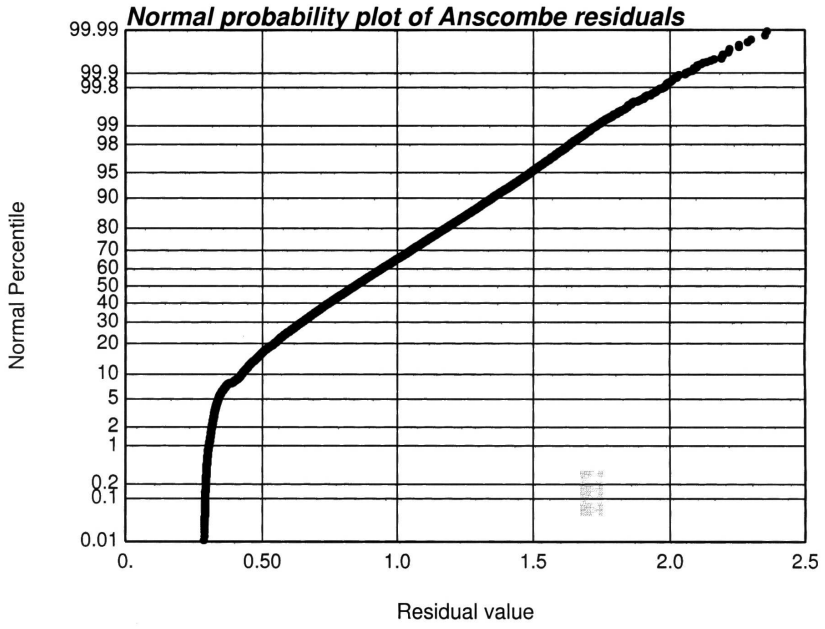


Figure 8.8 Anscombe residuals from amounts model for daily rainfall in the Lee catchment.

8.3 Single-site Disaggregation

The GLM can simulate daily rainfall at various sites, or indeed at any location within the modelled region. At this stage, a model to generate hourly rainfall at a master gauge is required. Here, we build on the work reported in Chapter 2 and use a random parameter Bartlett-Lewis Rectangular Pulse model (BLRPM) with 2-parameter cell intensity distribution. Rather than using the parameterisation given in Table 2.4 (page 14), here we define an alternative parameterisation that is more physically meaningful. The new parameterisation is defined in Table 8.4.

Table 8.4 BLRPM parameterisation used in disaggregation work. The seven model parameters $\lambda, \mu_x, \sigma_x, \delta_c, \sigma_c, \mu_c$ and δ_s are defined in terms of the original parameterisation given in Table 2.4.

Storm arrival rate	λ	(hr^{-1})
Mean cell intensity	μ_x	$(mm.hr^{-1})$
Standard deviation of cell intensity	$\sigma_x = \sqrt{\mu_x{}^2 - \mu_x^2}$	$(mm.hr^{-1})$
Mean cell duration	$\delta_c = \frac{\nu}{\alpha - 1}$	(hr)
Inter-storm standard deviation of cell duration	$\sigma_c = \frac{\nu}{\sqrt{(\alpha-1)^2(\alpha-2)}}$	(hr)
Mean number of cells per storm	$\mu_c = 1 + \frac{\kappa}{\phi}$	
Mean duration of storm activity	$\delta_s = \frac{\nu}{(\alpha-1)\phi}$	(hr)

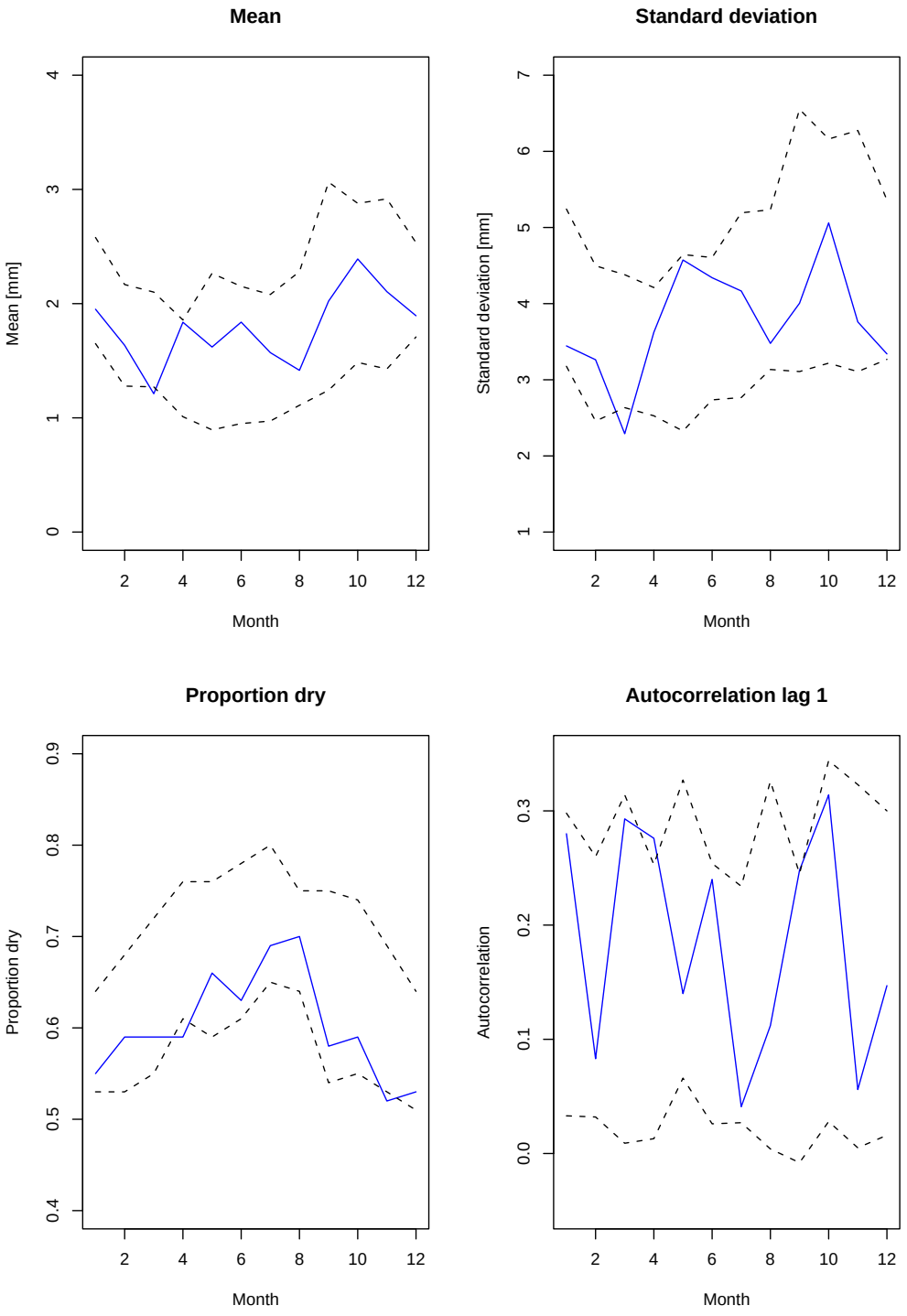


Figure 8.9 Observed and simulated monthly statistics for site 0160 at a daily time-step. In each plot, the blue line is obtained from observations and the dashed lines show the envelope obtained from 40 GLM simulations.

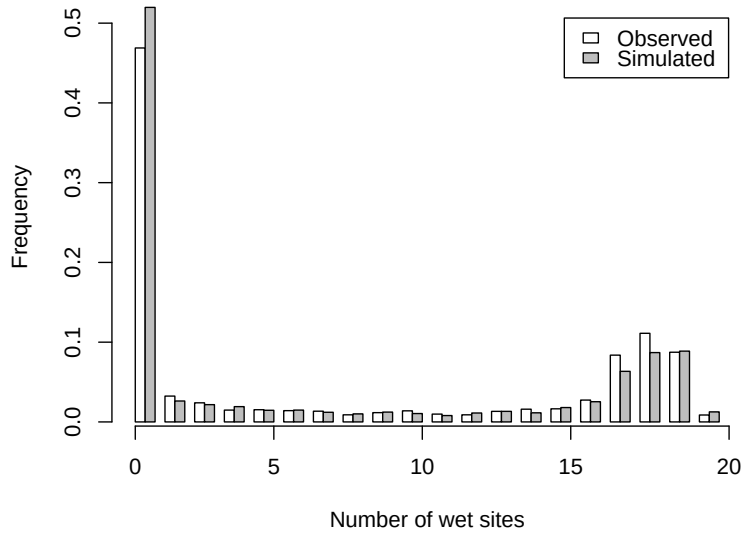


Figure 8.10 Observed distribution of number of wet sites in the LEe catchment, together with distribution simulated by GLMs using a beta-binomial model for spatial dependence in rainfall occurrence.

8.3.1 Fitting of the BLRPM

The BLRPM has been fitted to data from the master gauge (site 0163), using a method of moments as described in Chapter 2. Site 0163 has been chosen as the master gauge because of its central location in the Lee catchment and its continuous record for the period 1990-2002 with few missing values.

Several combinations of fitting statistics have been tested for the moment-based fitting. The statistics that gave best results in terms of minimising the fit between the expected and observed properties are the Mean; 1-hour Variance; 6-hour Variance; 24-hour Variance; 1-hour Autocovariance; 1-hour Proportion of dry periods and 24-hour Proportion of dry periods. In this investigation, we use the inverse of the squared historical values as weights (year average). Minimisation of the objective function (2.1) has been carried out using the Shuffled Complex Evolutionary (SCE-UA) algorithm described in Section 2.3.2, rather than using the R routines described in Chapter 2.

An initial simulation exercise revealed that in a few months of the year, storms with only one cell were generated repeatedly. Therefore, in order to obtain more realistic storms, some boundaries constraining the range of model parameters were defined as shown in Table 8.5. In particular to ensure that storms with more than one cell are produced (i.e. $\mu_c = 1 + \frac{\kappa}{\phi} > 1$), the lower boundary for κ was set to 0.01 so that κ does not become too small, and the upper boundary for ϕ was set to 1 to prevent too large a value of ϕ. An upper boundary on λ was also set so that up to five storms per hour could be generated. With these constraints in place, the estimated BLRPM parameters for each month are given in table 8.6.

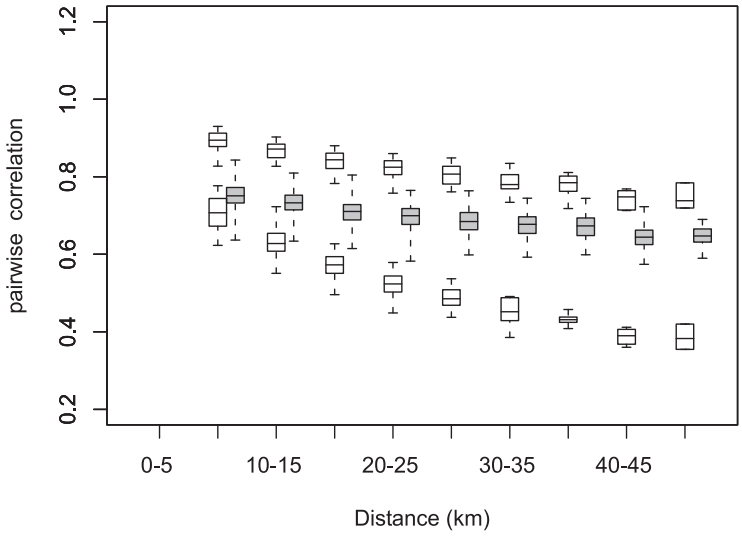


Figure 8.11 Pairwise correlations between rain gauges for the Lee catchment. White box-plots are observed daily (top) and hourly (bottom) correlations, grey box-plots are the results from simulations using the observed inter-site Anscombe residual correlations in rainfall amounts. The box-plots represent the median, 25 and 75% quantiles, the whiskers extend to the extreme values.

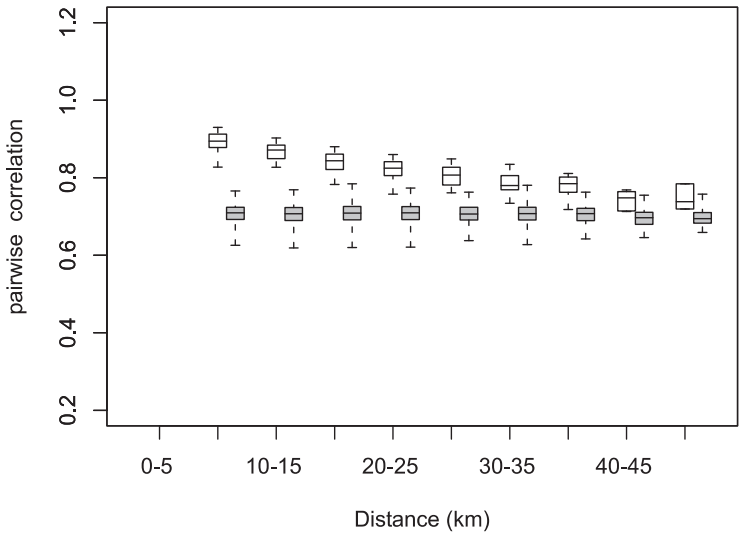


Figure 8.12 Daily pairwise correlation between rain gauges for the Lee catchment. As Figure 8.11, except that the hourly cross-correlations are excluded, and here the simulations used a constant inter-site Anscombe residual correlation in rainfall amounts.

Table 8.5 Constraints imposed on BLRPM parameters for fitting to data from site 0163.

Parameter	Lower bound	Upper bound
λ	0.00001	0.18
μ_x	0.00010	20.00
α	2.00001	20.00
ν	0.00001	20.00
κ	0.01000	10.00
ϕ	0.00001	1.00
σ_x	0.00001	10.00

Table 8.6 Estimated BLRPM parameters for site 0163 (OF: value of the objective function)

Month	λ	μ_x	α	ν	κ	ϕ	σ_x	OF
1	0.03180	1.44480	6.45502	3.48998	0.24764	0.15963	0.66352	6.48E-09
2	0.03088	0.97155	3.70047	1.59769	0.65262	0.21727	0.67948	6.43E-07
3	0.02114	1.25618	4.59485	2.43372	0.22869	0.12865	0.35678	9.44E-07
4	0.02374	1.27010	7.29332	1.26043	0.59964	0.04839	1.34963	2.04E-06
5	0.01829	1.72448	6.38034	0.28931	0.69754	0.02120	4.20670	1.31E-05
6	0.01915	6.58788	3.95294	0.34049	0.14794	0.03947	2.31878	4.53E-06
7	0.01752	1.92690	7.72351	1.20251	1.08564	0.11288	4.45936	7.81E-06
8	0.01666	4.98179	10.60813	1.54685	0.35758	0.10102	3.06124	2.60E-06
9	0.02385	1.14486	10.00757	2.89612	0.88906	0.09982	1.67808	3.83E-06
10	0.02591	1.72476	2.98359	0.22083	0.65200	0.03766	2.21018	2.77E-05
11	0.03227	0.94980	3.74207	0.53414	0.68891	0.04996	1.39419	1.59E-05
12	0.03686	0.95697	4.21056	0.97753	0.74226	0.12601	0.73860	7.46E-11

8.3.2 Disaggregation using HYETOS

Once the single-site model parameters are obtained they can be input to the HYETOS single-site disaggregation model (Koutsoyiannis and Onof, 2000). Given a data series at a daily scale (here from the GLM), several runs from the Bartlett-Lewis model are performed separately for each cluster of wet days. This continues until the sequence of simulated daily depths matches the sequence of daily totals from the GLM within a tolerance distance d, defined as

$$d = \left[\sum_{i=1}^{L} \ln^2 \left(\frac{Z_i + c}{\widetilde{Z}_i + c} \right) \right]^{\frac{1}{2}} , \tag{8.1}$$

where Z_i and $\widetilde{Z}_i$ are respectively the original and simulated daily intensity, L number of wet days in the sequence and c is a small constant (set to 0.1 mm). "Adjustment" is then performed to ensure that the hourly synthetic series generated by the stochastic model is consistent with the given daily series. Procedures are used to allocate the error in the cumulative hourly rainfall over the daily interval. These procedures are accurate in the sense that they preserve explicitly certain statistics. Furthermore, the methodology uses repetitive sampling to improve the approximations of statistics that are not explicitly preserved by the adjusting procedures. Three adjustment procedures were investigated by Koutsoyiannis and Onof (2001), who concluded that a simple proportional adjustment of the hourly values is the most appropriate for the disaggregation problem. The proportional adjusting procedure modifies the initially

generated values $\widetilde{X}$ to get the modified values X according to:

$$X_s = \widetilde{X}_s \left(\frac{Z}{\sum_{j=1}^{24} \widetilde{X}_j} \right), \; s = 1, \ldots, 24 \; , \tag{8.2}$$

where Z is the daily intensity to be disaggregated.

The scheme incorporates four levels of repetition which are entered when d is higher than a certain tolerance.

8.3.3 Preliminary results

To investigate the performance of the single-site disaggregation scheme, the theoretical values of various summary statistics were calculated under the fitted BLRPM (the analytical expressions of the properties can be found in Appendix B) and compared with the values obtained by applying the disaggregation procedure to real daily data from the master gauge for the period 1990-2002. A sensitivity analysis was conducted to assess the impact of the convergence criterion on the iterative sampling procedure. Koutsoyiannis and Onof (2001) recommend a value of 0.1 for the tolerance on d. Various values around 0.1 were tested, but no significant difference could be observed when changing the allowed distance between 0.01 and 0.5, hence the suggested tolerance of 0.1 was kept.

Figure 8.13 gives a comparison of historical, BLRPM fitted, and disaggregated values of some hourly summary statistics at the master gauge, for the period 1990-2002. A good agreement is observed between the historical and the disaggregated series using the previously fitted parameters. An increase in the standard deviation by 0.04 mm on average for the whole year (i.e. around 10%) is however noted and is partly due to the adjusting procedure in the disaggregation algorithm. The optimal set of parameters overestimates the proportion of dry periods in January, June and August by an increment of 0.02 (i.e. around 2%). The temporal autocorrelations are slightly underestimated with a discrepancy of 0.035 between the historical and disaggregated series; about half of this difference is observed between the historical and the expected values. A refinement can be obtained in retaining single-site parameters that fit better the autocorrelations in October and November where the largest discrepancy of 0.1 (i.e. around 20%) is observed between the expected and historical series.

8.4 Multisite Disaggregation

Once the single-site disaggregation model has been used to obtain sub-daily data at the master gauge, a standard profile can be defined as the hourly rainfall at the master gauge divided by the daily total rainfall. This temporal profile can then be applied to any other gauge by multiplying the standard profile by the total rainfall at that gauge:

$$P_{ij} = \frac{x_{i0}}{\sum_{i=1}^{24} x_{i0}} \times \sum_{i=1}^{24} x_{ij} \; , \tag{8.3}$$

where P_{ij} is the profile at hour i for gauge j, x_{i0} is the master gauge rainfall intensity at hour i and x_{ij} is the rainfall intensity at hour i for gauge j. If the master gauge records

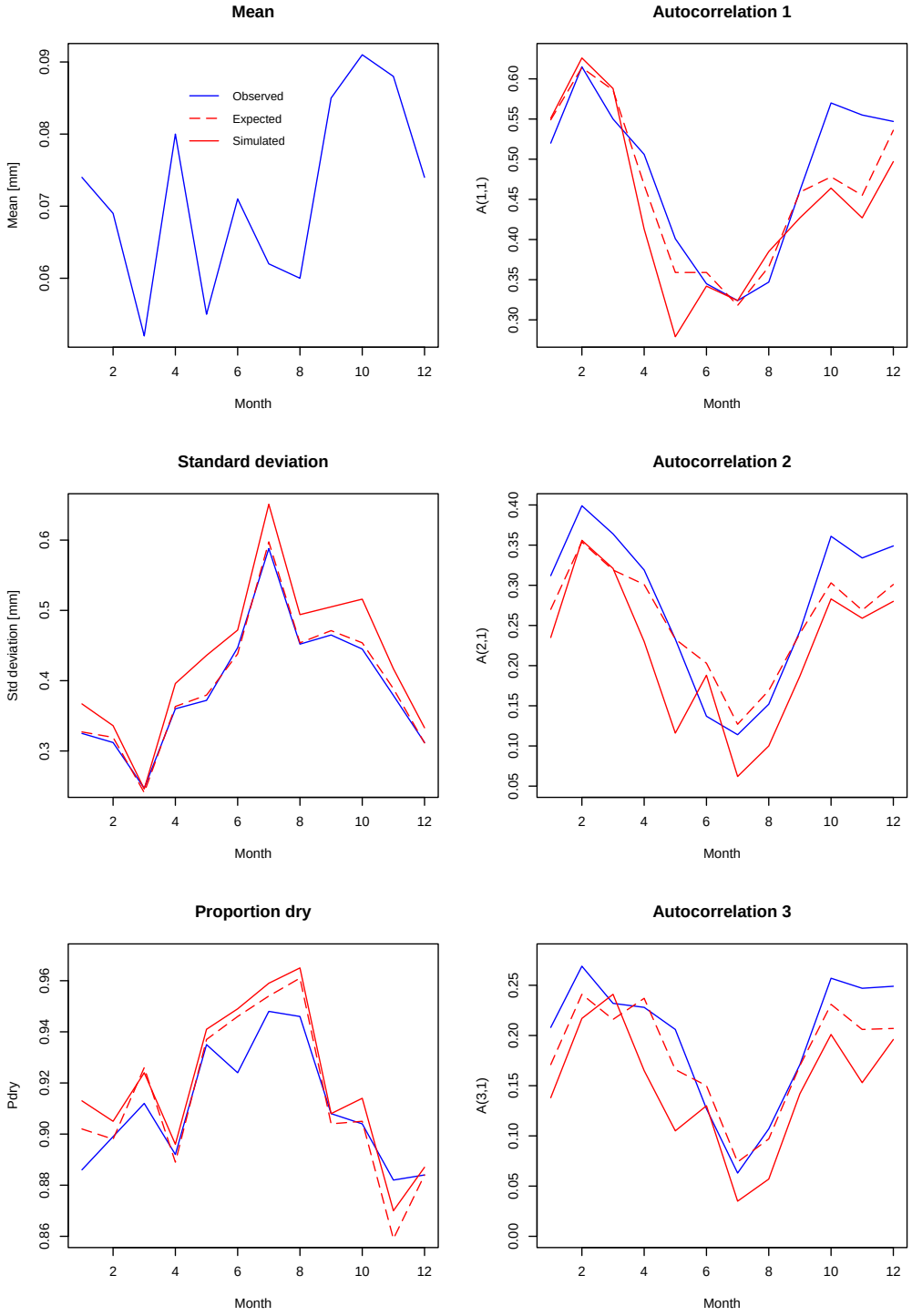


Figure 8.13 Impact of the disaggregation scheme on properties of hourly rainfall at the master gauge. In each plot the blue line represents observed properties; the dashed red line represents the theoretical properties from a BLRPM fitted to the data; and the solid red line represents the properties after applying the HYETOS disaggregation scheme.

zero rainfall while other gauges record non-zero daily intensity, an artificial profile is used to disaggregate at the other gauges. So far the analysis has been carried out using the previously computed profile at a site, but there is scope for refinement of this strategy.

A data analysis of 16 events, taken at different times of the year and representative of most rainfall types, has been undertaken. For the majority of them, the results from applying a spatially-uniform temporal disaggregation are encouraging. Two contrasting events are used to illustrate the method. Using real data, the standard profile of the master gauge (site 0163) is applied to the satellite stations. Figure 8.14 shows the simulated and observed profiles for the event of 3rd February 1990 (daily mean of 13 raingauges: 24mm, range: 17.8-28.6 mm, core duration: 9h). For this event, the modelled data reproduce well the observed series except at site 0101. This site is on the edge of the Lee catchment, and this may indicate the limit of the application of a standard profile. This is representative of usual model performance, however exceptions exist. An example is the intense summer event of 23rd July 1996 (daily mean of 19 raingauges: 20.8 mm, range: 1.4-53.2 mm, core duration: 4h) where the results in Figure 8.15 show some limitations (See sites 0101 and 0110) as the rainfall is spatially highly variable.

8.5 Testing the combined scheme

We have seen that the GLM is able to generate realistic sequences of daily rainfall at various sites, and that the joint Poisson-cluster model together with the adjusting procedure performs well in simulating hourly rainfall at the master gauge. The representative event of 3rd February 1990 illustrates successfully the implementation of a spatially-uniform temporal disaggregation. The next step is to evaluate the temporal disaggregation for the whole period of study. Further, these models need to be tested in combination to assess their simulation performance.

8.5.1 Application of a single temporal profile at all gauges

First, the use of a temporal profile at one gauge to disaggregate at other gauges is tested by applying the observed hourly profile from the master site to observed daily data from other gauges for the entire 1991–2002 period. The properties of the resulting hourly series are compared with those of the observed hourly data. These properties include the mean, standard deviation, proportion of dry periods, autocorrelation (at lag 1, 2 and 3) and skewness. At all sites, a close agreement between the historical and modelled series was found. Figure 8.16 shows some specimen results for site 0160. Compared with the observed values, and on average for the whole year, there is a 0.4% increase in the standard deviation, 3.4% decrease in the lag 1 autocorrelation, 1.1% increase in the skewness and 0.2% decrease in the reproduction of the dry proportion. Hence the difference between the two series of statistics is small and validates the hypothesis of applying a spatially-uniform temporal profile.

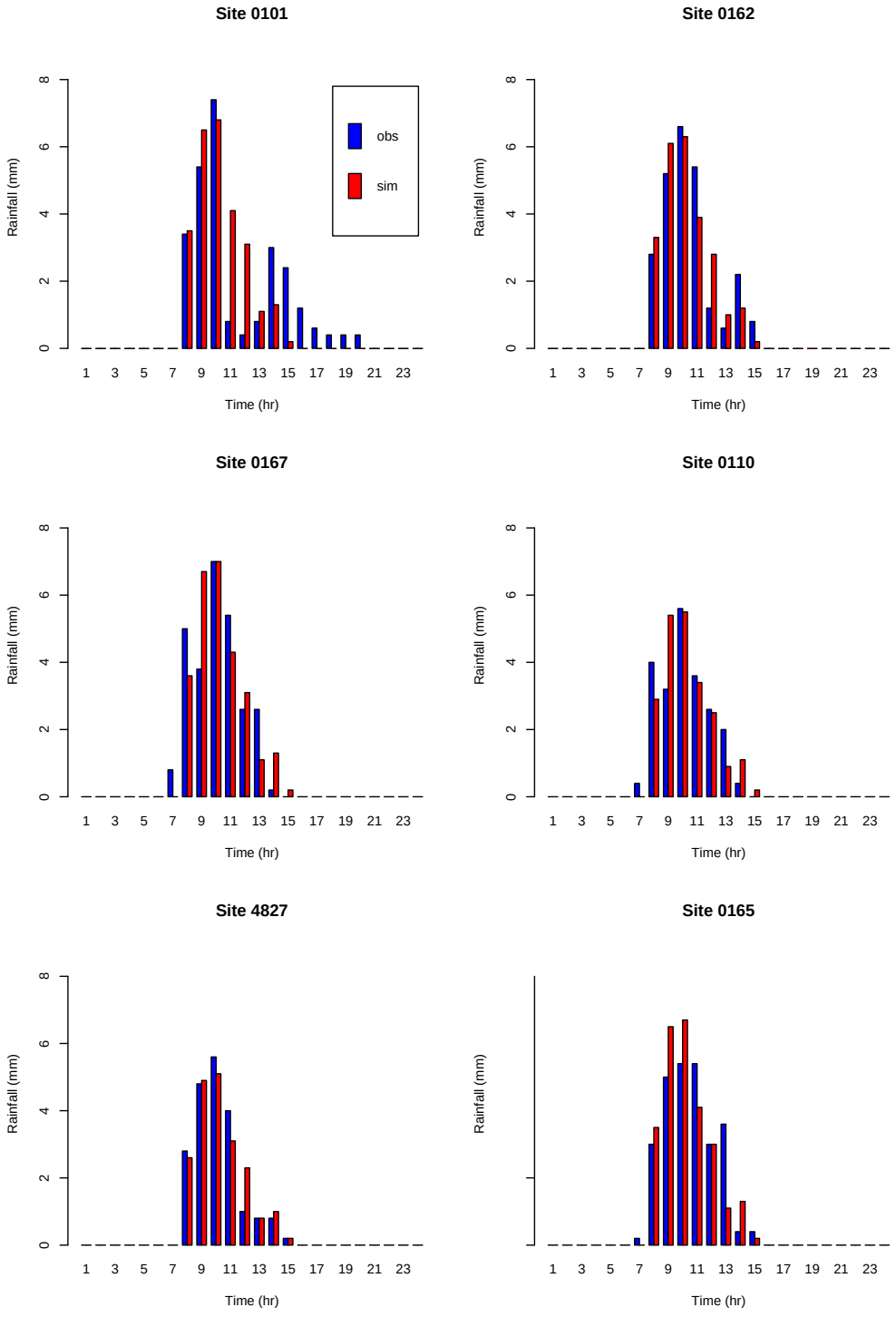


Figure 8.14 Event of 03/02/90 — Applying the observed profile of the master gauge to the satellite gauges

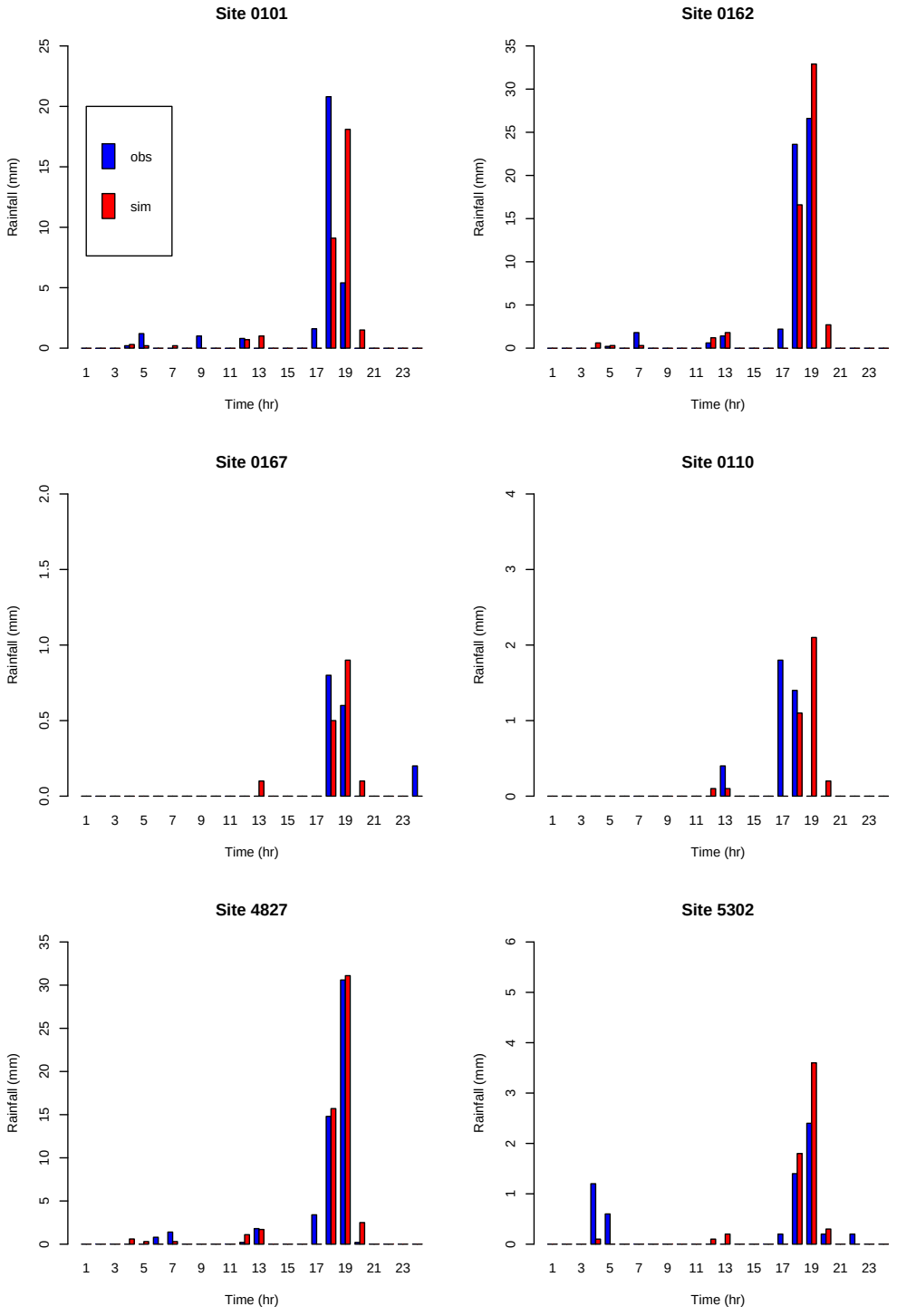


Figure 8.15 Event of 23/07/96 — Applying the observed profile of the master gauge to the satellite gauges

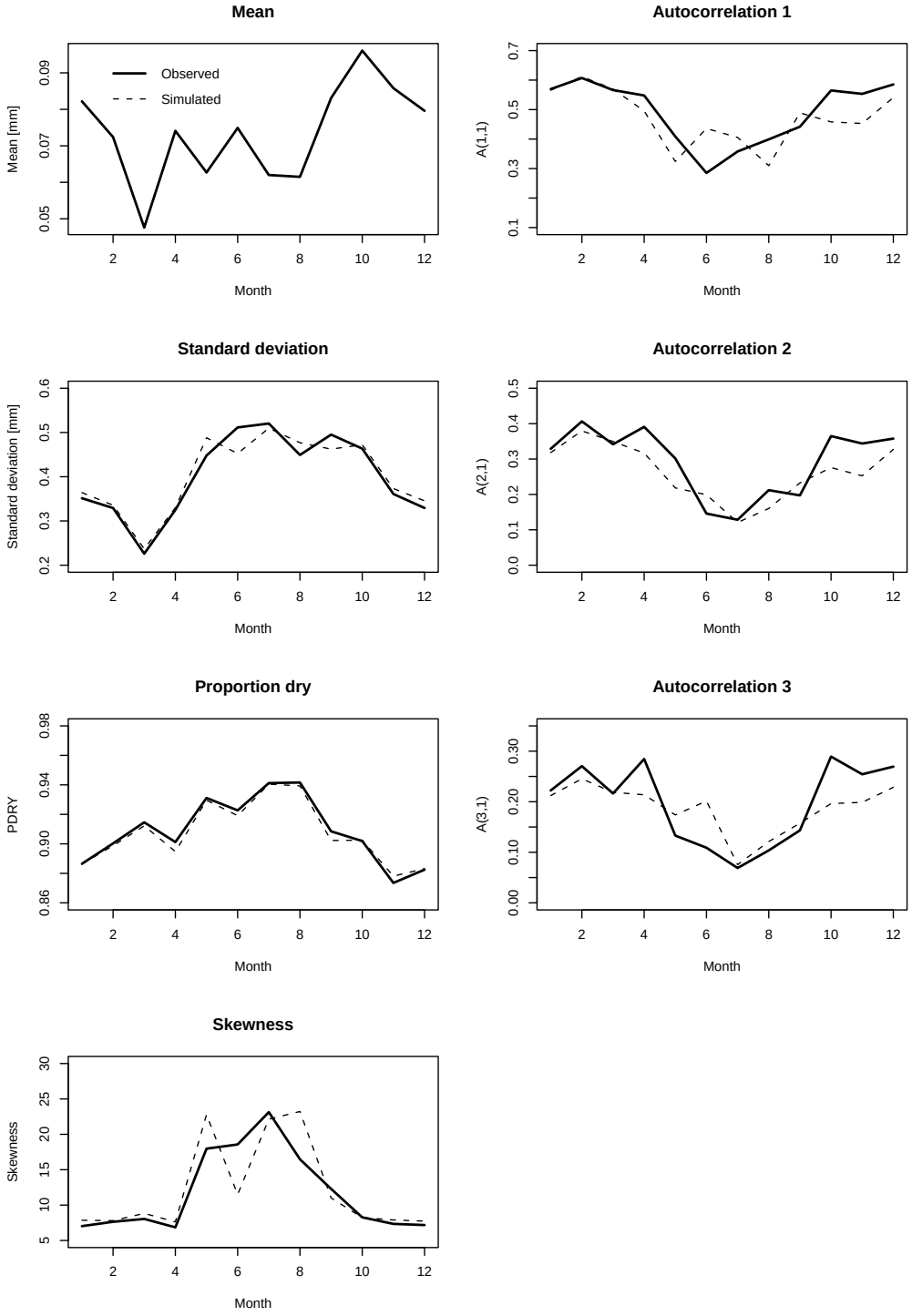


Figure 8.16 Properties of hourly sequences at site 0160, 1991–2002. Solid lines show the observed hourly properties; dashed lines show the properties obtained when the observed daily sequences are disaggregated by applying the observed sub-daily temporal profile from the master gauge 0163.

8.5.2 Effect of single-site disaggregation

Next, the variability introduced by the HYETOS single-site disaggregation scheme is assessed. Using observed daily data, 20 disaggregations are run at the master gauge and each temporal profile is transferred to the other gauges in the catchment. Results show an overall good fit between the modelled and observed series over the 21 sites. In particular the reproduction of the skewness is encouraging for the assessment of the extreme value performance of the model. However, the disaggregation scheme has a tendency to overestimate the standard deviation for all months, as well as the proportion of dry periods for the summer months and to underestimate the autocorrelations. These features were already in evidence when fitting the BLRPM. Figure 8.17 presents the statistics obtained for site 0160. Taking the mean statistics from the 20 disaggregations, here we have a 14% increase in variance, 15% decrease in lag 1 autocorrelation, 5.6% increase in skewness and 1.4% increase in the proportion of dry periods. Hence the disaggregation scheme compares well with the observed statistics even though it introduces some bias.

8.5.3 Effect of GLM simulations

To assess the effect of using GLM simulations in place of actual daily sequences, 20 GLM simulations were generated for the period 1991-2002 and combined with the disaggregation scheme. The results show good agreement between the observed and modelled series, although the overestimation of the standard deviation and proportion of dry period is still present. The envelope from the 20 simulations is represented in Figure 8.18. On average for the 20 simulations, the mean is well represented throughout the year with a 0.3% increase compared to the historical values. We also observe a 24% increase in the standard deviation, 18% decrease in the lag 1 autocorrelation, 1.5% decrease in skewness and 2.6% increase in the proportion of dry periods.

Finally, we explore the effect of using the GLM-disaggregation method at sites not used in the model fitting (Figure 8.18, dotted lines). The GLM model is recalibrated to 19 gauges on the catchment (exclusion of sites 0160 and 0106), daily data are simulated at site 0160 and the combined scheme is run. Similar features in the reproduction of the statistics are observed compared to the previous case. The results are promising and, provided that a site is surrounded by few rain gauges, the combined scheme has the potential to simulate hourly rainfall information at a location where no data are available.

8.5.4 Analysis of extremes

The extremal behaviour of the modelled rainfall is also inspected, since this is of primary importance in applications such as flood management. The annual maxima, both daily and hourly, for the period 1991-2002 are considered and plotted using the Gringorten plotting position method, similarly for the 20 simulations. The relationship

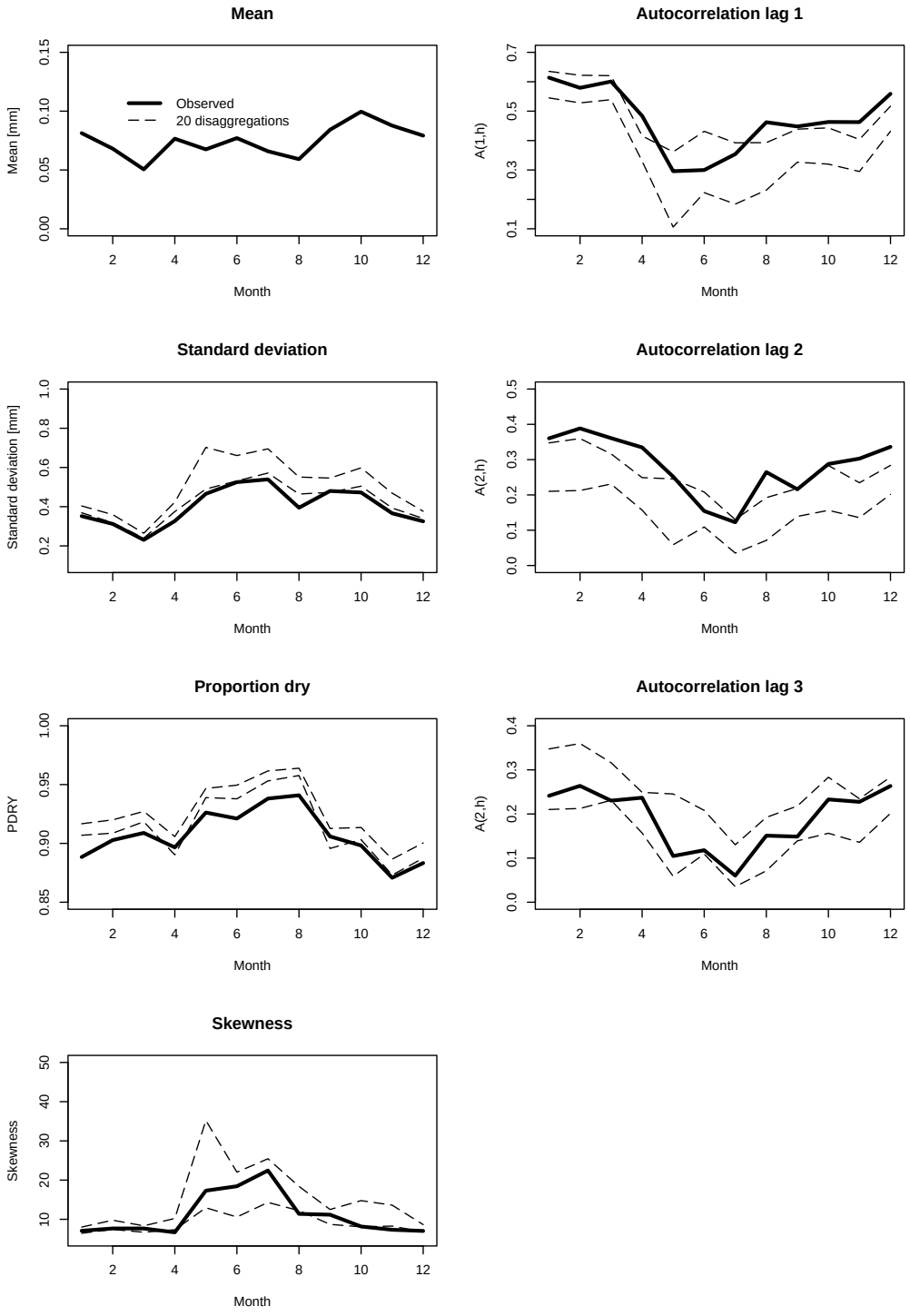


Figure 8.17 Properties of hourly sequences at site 0160, 1991–2002. Solid lines show the observed hourly properties; dashed lines show the envelopes from 20 disaggregations obtained by applying temporal profiles from the master gauge to all sites in the catchment. The profiles were obtained by applying the HYETOS single-site disaggregation technique to the observed daily sequence at the master gauge.

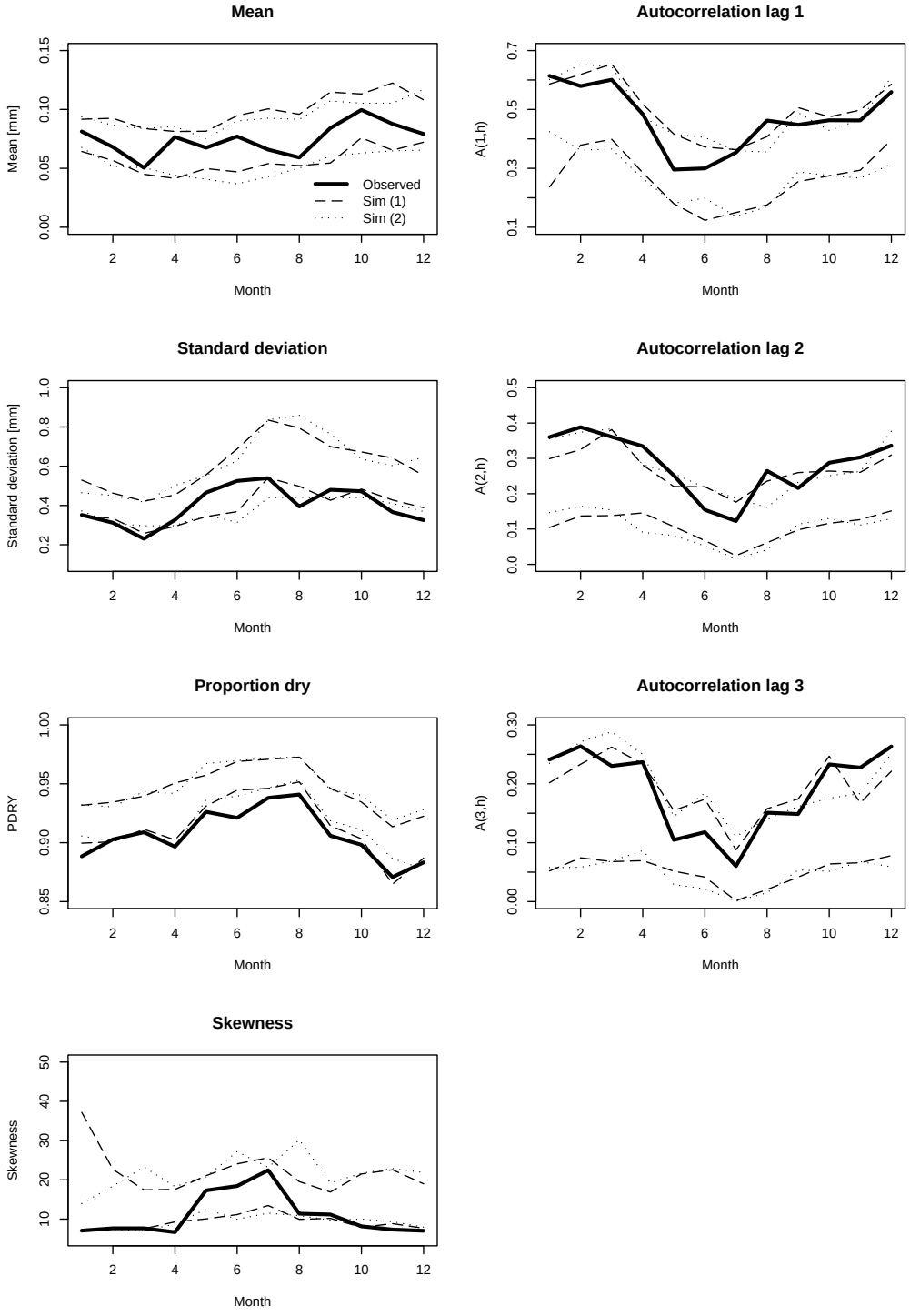


Figure 8.18 Single-site disaggregation at the master gauge conditioned on simulated daily rainfall from GLMs and sub-daily temporal profile applied at a typical site on the catchment. The plot compares the observed series with the envelope of 20 simulations at a gauge used (Sim 1) and not used (Sim 2) in the calibration of the model.

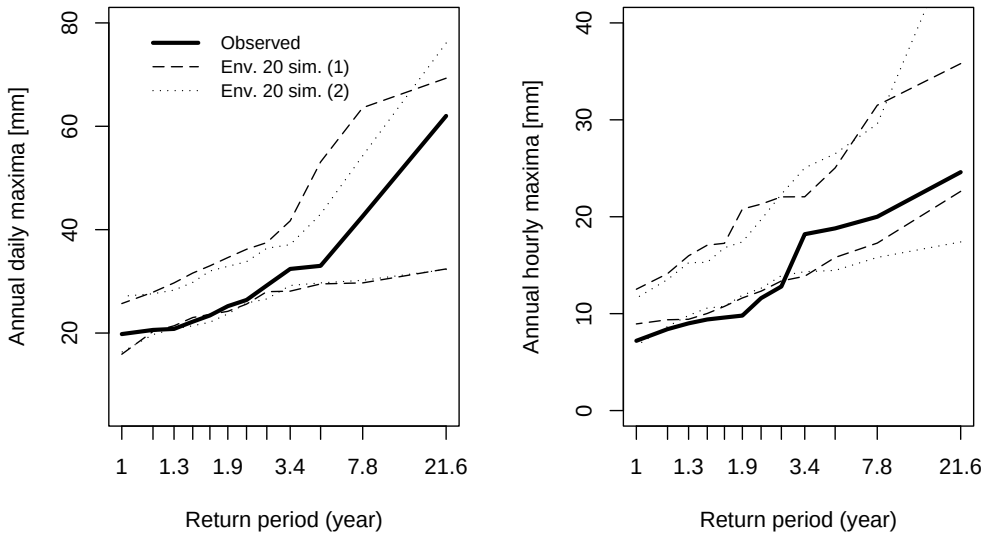


Figure 8.19 Return level plots for daily (left) and hourly (right) annual maxima. The plots compare the observed series with the envelope of 20 simulations at a gauge used (Sim 1) and not used (Sim 2) in the calibration of the combined multisite disaggregation model.

between the the reduced variate y and the return period T is given by

$$y = -\ln\left[-\ln\left(1-\frac{1}{T}\right)\right] . \tag{8.4}$$

Results in Figure 8.19 show a reasonable agreement between the simulations and the historical series at daily level. The results are similar regardless of whether the test gauge was used for model calibration. At an hourly time-step, the historical series is overestimated for return periods less than three years but lies within the envelope of 20 simulations for higher return periods. A closer inspection of the maximum hourly totals reveals that out of 20 12-year simulations, 31 annual maxima are equal to the corresponding daily depths. In one third of the cases, this is due to the single-site model, the other two thirds arise when the master gauge records zero rainfall and the previous profile applied corresponds to a day where the total rainfall falls in an hour. An analysis of the observations reveals two occasions during the 1991–2002 period when an hourly extreme corresponds to the total daily rainfall; but on average the observed daily total is lower than for the simulations (11.3 mm observed compared to 18 mm on average for the simulations). A refinement could be obtained defining a more appropriate artificial profile when the master gauge records zero rainfall.

The overestimation can also be partly explained by the higher proportion of dry periods obtained from the combined scheme, which suggests that more intense short duration rainfall intensities are produced leading to higher hourly maxima. Furthermore, the mixed performance of the BLRPM in reproducing the extreme values has already been discussed in Chapter 2 (see also Onof and Wheater 1993; Onof et al. 2000). The performance of the disaggregation scheme could possibly be improved by using the two-stage fitting procedure suggested previously, to improve the extremal behaviour of the single-site simulations.

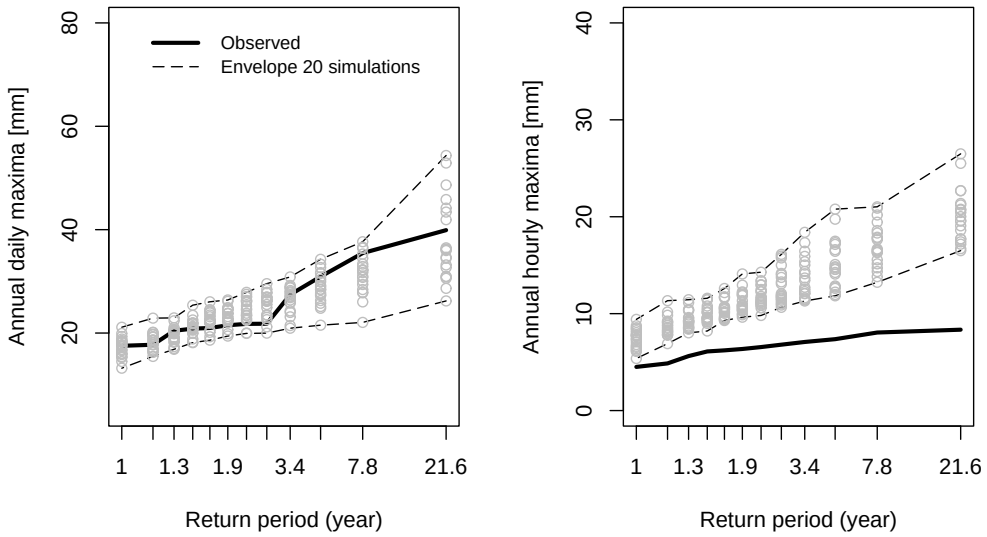


Figure 8.20 Return level plots for daily (left) and hourly (right) annual maxima (gauge average time series).

Examining spatial extremes, return level plots of observed values versus simulated average rainfall from all gauges (Figure 8.20) reveal that the model performs well with respect to the daily extreme values but clearly overestimates the hourly series for all return periods. This is expected since a single temporal profile is applied at all sites and therefore the hourly peak is reproduced at the same time for all gauges. The importance of this will depend on the sensitivity of a catchment to individual hourly rainfall intensities.

8.6 Summary of chapter

A simple disaggregation method has been proposed, which enables multi-site rainfall sequences to be generated at a subdaily time scale. The proposed scheme retains a relatively simple structure, and uses the GLMs of Chapter 6 to include spatial and temporal nonstationarities properties on a daily basis. Hourly data are then obtained using the HYETOS technique (based on the Poisson cluster models of Chapter 2) to obtain a subdaily temporal profile at a master gauge; this profile is then applied to all gauges in a catchment.

The performance of the proposed disaggregation procedure has been assessed in terms of the reproduction of a set of standard statistics and with regard to extremes. Results from a network of 21 rain gauges on the River Lee are promising. Simulated rainfall properties match the empirical statistics to an acceptable level for practical purposes. Although the model overestimates the size of the extreme value properties of rainfall, the larger observed maxima fall within a 95% confidence interval of simulated extremes. Additionally, the combined scheme has the potential to transfer information from sites included in the calibration of the GLM to nearby locations which are not used to calibrate the model.

Further research is required to improve the reproduction of the proportion of dry periods, which in turn will benefit the simulation of extreme values. The performance of the Bartlett-Lewis model is conditional on the finding of suitable parameters to run the model. The first step would be to identify some single-site parameters that provide a better fit to the proportion of dry periods for the months of January, June and August; and also the lag 1 autocorrelation in October and November. A second direction where development can take place involves further investigation on the impact of the adjusting procedure on the overestimation of the standard deviation. Improvement in the reproduction of the proportion of dry periods and the standard deviation would lead to data that are more temporally correlated.

The major limitation of this scheme is the overestimation of the spatial correlation due to a unique temporal profile applied at all gauges. In order to decrease this correlation, the temporal profile could be shifted in time for some of the gauges (either randomly or taking into account the distance away from the master gauge). More generally, when applying the temporal profile, complete spatial dependence between sites is assumed, which does not hold beyond a certain distance. It will be important to identify at what scale allowance for spatial heterogeneity must be made. To address this issue, evaluation of the methodology could be extended to a larger region. Various options could be pursued for more complex procedures. One possible development would be to run independent disaggregation procedures for discrete sub-areas, another would be to build in spatial correlation to a multi-site procedure. More generally, if adequate radar data were available then the continuous space-time models of Chapters 9 and 10 could be used for subdaily data, conditioned on daily GLM simulation.

The scheme proposed here has the advantage of simplicity and can be readily implemented. In the context of flood management, necessary further work includes the testing of the modelled rainfall as input to rainfall-runoff models to assess the impact of the simulated rainfall on runoff generation. A first study on this aspect is presented in Chapter 12 below.

Part III

Development and testing of non-stationary spatial-temporal rainfall model for medium-term application

We turn now, in Part III, to consider rainfall as a fully spatial-temporal process where, for maximum flexibility, both space and time vary continuously. Here, the ultimate goal is to develop a model that can be simulated to produce artificial, high-resolution rainfall data. These data can then be used as input to hydrological rainfall-runoff models for flood risk assessment and design purposes.

In earlier work (DEFRA FD0426) a spatial-temporal rainfall model based on Poisson cluster processes was developed, and some preliminary testing was carried out with a single, very limited, radar data set (from the Wardon Hill radar). This model was stationary both in space and time, apart from allowing basic seasonality. The report on that project noted the need for further development and testing of the model on a more substantial data set. Thus, objectives of the current DEFRA-funded project (FD2105) were to carry out further model testing, and theoretical model development as necessary, and also to develop the model to incorporate temporal and spatial non-stationarities (enabling the representation of possible climate change scenarios, and topographic effects). An additional objective was to review recent developments in fractal methods for rainfall and to compare these with the Poisson-based approach, to ensure that our spatial-temporal model development is based on the most appropriate theory.

A substantial review paper on fractal methods methods for rainfall is attached to this report as Appendix E. The content is somewhat technical and we give here only a brief summary of the main points and conclusions. The basic idea of a multifractal model is that it enables the parsimonious description of rainfall variability over a large range of scales in space and time. Most multifractal fields are based on multiplicative cascade processes, and these may develop through either a continuous or a discrete range of scales. Continuous cascades do not take exactly zero values so that an arbitrary threshold has to be imposed to enable spatial-temporal regions to be defined as ‘dry’. The generating mechanisms of discrete cascades are mathematically simpler and more intuitive, but the resulting fields have a rather artificial rectangular structure.

Much of the work on multifractal models is most applicable to the purely spatial representation of total rainfall over some fixed time period. The contrast between the directionality of time and the symmetry of spatial dimensions means that space-time processes cannot sensibly be represented simply by extending spatial models to include an extra spatial dimension. Advances are being made in research on genuinely spatial-temporal fractal models, but many outstanding modelling issues remain to be resolved. We note also that the simulation of such models is highly computer-intensive.

In general, multifractal models for (usually, spatial) rainfall have been validated in terms of their ability to reproduce the scaling properties of observed rainfall, but their performance with regard to other properties that are crucial for rainfall-runoff, such as wet-dry properties and extremes, have not yet been adequately explored. In addition, there are questions to do with parameter estimation and model identification that need to be addressed. The overall conclusion is that fractal models for spatial-temporal processes are not yet at the stage where they are readily applicable for continuous simulation of rainfall, and that therefore the Poisson-based approach is the most promising for medium-term application.

Given these conclusions of the review of multifractal models (Appendix E), we will

concentrate on the development of Poisson-based nonstationary continuous simulation models in the rest of this part of the report. Thus, in the current project, there has been considerable further testing of the stationary continuous simulation model, together with development of this model to allow for spatial and temporal nonstationarities. The stationary continuous simulation model comprises two main sub-models. The first is a model for the spatial-temporal rainfall field interior to a rain event; details are given in Section 9.2. The second sub-model is for the advection process by which rain events move across a catchment; details can be found in Section 9.3.

Substantial effort has been expended in developing and checking the software used for fitting the model for the rain event interiors, and for the continuous simulation. The stationary version of the model has been fitted to data from the Chenies radar, and extensive model validation has been carried out. The data and their calibration is described in Appendix F, while the results of the continuous simulation will be discussed in Chapter 10. The model is fitted on a month by month basis so that basic seasonality is incorporated, but it is otherwise stationary in space and time.

A nonstationary spatial-temporal model is achieved by combining the stationary model, which provides the required finescale spatial and temporal structure, with a multiplicative rescaling field that drives the spatial and temporal heterogeneities. This extension of the stationary model is described in Section 9.5. The rescaling field is obtained by fitting Generalised Linear Models (see Part II of this report) to daily rain gauge data. A benefit of using gauge data here is that simulations of the final model will be calibrated directly to ground truth, so that careful calibration of the radar data themselves is not essential. Simulation of the nonstationary model enables ensembles of future rainfall over a catchment, allowing for its topographic features, to be explored. The results from the continuous simulation of the nonstationary model will be discussed in Chapter 10.

The overall conclusions of this work are, firstly, that the model for rain event interiors has been extensively validated and provides a good representation of the basic statistical properties of the fine scale spatial and temporal structure of rainfall. Second, methodology to produce a continuous stationary simulation of rainfall across a catchment has been further developed. This has been applied in conjunction with the Chenies radar data and the simulated realisations of the model have been shown to have statistical properties that are generally quite a good representation of those of the empirical data. Third, for the first time, a way has been demonstrated to extend the stationary spatial-temporal model to allow topographic effects and possible climate change scenarios to be included. In the current report, validation of the continuous simulations has focussed on wet/dry properties and first and second-order moments of the spatial-temporal process. In future work, and especially for hydrological applications, it will be important to look at the tail of the intensity distribution and properties of extremes.

The model validation has raised many complex issues, many of which are to do with the radar data themselves and their calibration to reflect ground truth. Ultimately, the results of the continuous simulations can only be as good as the data used to fit the models, and various limitations of currently available radar data are discussed in Appendix F. We have applied the UK standard calibration procedure (Moore et al., 1994) and found that, for the Chenies data, it is possible to improve the agreement

between the calibrated radar values and data from rain gauges under the radar region, by using a modified procedure. As well as improving the agreement between the distributions as a whole, emphasis has been put on reproducing the larger hourly values which are especially important for rainfall-runoff applications. Nevertheless,a particular issue (see Appendix F) is that rather arbitrary choices of constants in the method of calibrating the radar data can lead to calibrated data with different statistical properties (often with reduced temporal autocorrelations) and thus, at least potentially, to differences in the fitted model and its continuous simulation. This sensitivity of statistical properties to the calibration procedure is a source of some concern and means that, at present, models based on radar data must be used with some caution in flood risk assessment. A substantial further study is needed to gain a thorough understanding of all the issues involved. The need for the very time-consuming work (see Appendix F) to process and calibrate the Chenies radar data was not anticipated in the original project inception report. The necessity for its inclusion (with DEFRA approval) meant that although valuable insights have been gained, unfortunately, there has been much less time than expected for development of the spatial-temporal model, and thus various issues remain to be addressed.

At present, a limitation to the use of the continuous simulation model is that the within-event model can only be fitted to data from events that cover the entire radar window for a substantial period of time, and can be clearly distinguished from radar clutter and anomalies. This means that, for example, periods of very light scattered showers or rainfall events that pass over only a part of the window are not fitted. The omission of such events from the continuous simulation may lead to discrepancies between the properties of the simulations and those of the observed data; the significance of these omissions is discussed in Chapter 10. In the comparisons presented there, a threshold coverage of 15% has been applied to all images, with all pixels in images below this coverage being given intensities of zero. Given this threshold, the models generally perform acceptably in reproducing the statistical properties of the observed data. However, further model development is needed if the model is to be able to represent light and spatially and/or temporally intermittent rainfall.

A number of other outstanding issues have been identified both in terms of the identification and sampling of events, and the modelling of the advection process. These are discussed in Chapter 10. In connection with the former, a threshold rule has been applied to the calibrated Chenies radar series for the period 1990–2002 in order to select rainfall events to which the event interiors model can be fitted. This thresholding has resulted in the rather surprising fact that there appear to be an increasing number of rain events over the period, with a jump upwards in the number of wet events around 1994, and an increasing trend thereafter. In part, these observations can be explained by changes in radar recording practice, but inspection of rain gauge data suggests that at least part of the increasing trend during the 1990s is genuine. Such a trend would have various implications for the continuous simulation of rainfall, as discussed in Section F.5. If there are genuine changes in the numbers of events and their characteristics (and to determine long-term patterns far more than 12 years' data are needed), the need for nonstationary continuous simulation, as discussed in Section 9.5, becomes apparent.

In the light of experience with the Chenies radar data, modifications have been made

to some of the procedures described in our earlier work for DEFRA (FD0426). Changes have been made to the way in which models are fitted, as well as to the software implementation of both model fitting and simulation. Considerable time and ingenuity were required to identify the causes of problems and to resolve them. However, as a result, the procedures developed to fit the spatial-temporal models and to produce continuous simulations are now much more robust and firmly based, and can be used with more confidence for further model development.

The extension of the stationary spatial-temporal model to allow the inclusion of topographic effects and possible climate change scenarios is a significant step forward. The results are dependent on the adequacy of the stationary model, and the rescaling method employed involves the implicit assumption that climate change would leave the wet-dry structure of rainfall unchanged. Thus the model will need to be further developed to remove this rather implausible assumption. Nevertheless, the model proposed here is an important first step in demonstrating the feasibility of obtaining nonstationary simulations. For the future, it will be of particular interest to incorporate physically-based climate change scenarios into the rainfall simulations. This can be achieved by conditioning the GLMs on the output of numerical climate models. Research in this area is currently being carried out as part of DEFRA project FD2113.

Chapter 9 Development of a non-stationary spatial-temporal rainfall model

9.1 Introduction

Our basic *stationary* model for spatial-temporal rainfall fields has already been published in a number of places in the literature, and a detailed description can be found in the report on FD0426 (Wheater et al. 2000b) with further developments detailed in Wheater et al. (2006).Thus we give only an outline of the main features here. The aim is to approximate the complex spatial-temporal rainfall process using a simple stochastic mechanism. In particular, we incorporate some physical knowledge about spatial-temporal rainfall structure into a parsimonious stochastic model, parameterized in terms of physically meaningful quantities. This approach can be described as *stochastic-mechanistic*. The model is a spatial analogue of the point process based models, described in Part I of this report, that are used to represent the temporal process of rainfall at a single site. Separate submodels are built for the interior of a rain event (see Section 9.2) and for the timing of a sequence of rain events that will affect a particular catchment (Section 9.3).

The model is based on a stochastic point process, whose points give the spatial-temporal locations of localised areas of relatively intense rainfall that are called *rain cells*. It is assumed that individual rain cells have random durations, spatial extents and intensities. Based on observational studies Austin and Houze (1972),we assume a hierarchical system of rain cells, whereby cells tend to cluster (in space and time) within larger scale structures that we call *storms*. Storms themselves tend to cluster in similar manner to form *rain events*. These rainfall elements can be characterised in terms of their typical spatial extents and durations (see table 9.1). The model has been fitted to rainfall radar data covering an area of approximately 10,000 km^2 with rainfall intensities averaged over pixels of area 4 km^2. In particular, we concentrate on a *fitting window* consisting of a 52 $\times$ 52 array of 2 km $\times$ 2 km pixels. The temporal separation of the radar images is 5 minutes and we use images produced by the lowest radar beam, i.e. the level of the atmosphere nearest the ground. A rain event will typically cover the entire radar window at some point during its lifetime, and our model for event interiors is fitted to the cells and storms that make up the spatial and temporal *interior* of such events, i.e., without including their development and dissipation in time or their periphery in space. Thus only one level of clustering (of cells within storms) is required in that part of the model. In principle, another layer of clustering (of storms within rain events) could be introduced if data were available over an area large enough to contain a reasonable number of rain

Table 9.1 Typical spatial extents and durations of rainfall elements

rainfall element	spatial extent / km^2	duration
rain cell	10–50	up to 40 minutes
storm	100–1000	a few hours
rain event	$>$ 1000	several hours

events. We then use a separate submodel (Section 9.3), to describe the temporal succession of events over the radar window . Once the parameters of the model have been estimated, realisations of rainfall fields can be simulated in continuous space and time. A single parameterisation can produce many different realisations for use in Monte Carlo simulation studies of catchment response.

9.2 Modelling event interiors

9.2.1 Model definition and properties

Our model for rain event interiors was developed by Northrop (1998), building on previous work by Cox and Isham (1988). As described above, each rain event is a superposition of clusters of 'rain cells', each cluster being termed a 'storm'. Each rain event has a random velocity V, and is associated with an underlying random elliptical 'shape' specified by an orientation, Θ, and eccentricity, E. Storm origins, which are unobserved, occur in a Poisson process of rate λ in space-time; given the shape and velocity of an event, its storms are realised independently and identically about their origins. Each storm has a temporal duration (its 'lifetime') that is exponentially distributed with rate γ. During this lifetime, the storm moves with velocity V and gives birth to rain cell origins at a rate β in time, starting with a cell at the temporal storm origin. Thus the temporal structure of the storm is that of a Bartlett-Lewis Poisson cluster process (Cox and Isham, 1980, Section 3.1); the number, C, of cells per storm has a geometric distribution with mean $\mu_C = 1 + \beta/\gamma$. This is illustrated in the left panel of Figure 9.2.1.

In space, each cell origin is independently displaced relative to the moving storm origin, according to a bivariate zero mean Gaussian distribution whose contours share the elliptical shape (Θ, E) of the event. This yields a Neyman-Scott cluster structure (Cox and Isham, 1980, Section 3.1). Storms will tend to have a banded appearance when E is close to 1, but will appear more circular for E near 0. The scale, A_s, of the spatial cell displacements varies independently between storms. The cells themselves move with velocity V and share the elliptical shape (Θ, E), with random semi-major axis length A_c. Each cell has a random temporal duration D, during which it produces rainfall at a constant random intensity X throughout its spatial extent. Thus the cell can be viewed as a cylinder with elliptical cross-section and height X. A_c, D and X are realised independently for all cells within an event.

The right panel of Figure 9.2.1 illustrates the instantaneous spatial structure of a single storm. A rain event is a superposition of many such storms; the total rainfall intensity at any space-time point is the sum of intensities from all cells covering that point. Both storms and cells may (and often will) overlap in space and time. In this regard, it may be helpful to think of overlapping cells as contributions from different levels in the atmosphere.

The assumption that cell intensities are constant throughout their lifetimes and spatial extents is made primarily for mathematical convenience, although other possibilities can be considered. For example, a multiplicative random noise could be applied to give the cell an irregular profile (Rodriguez-Iturbe et al., 1987), or a truncated

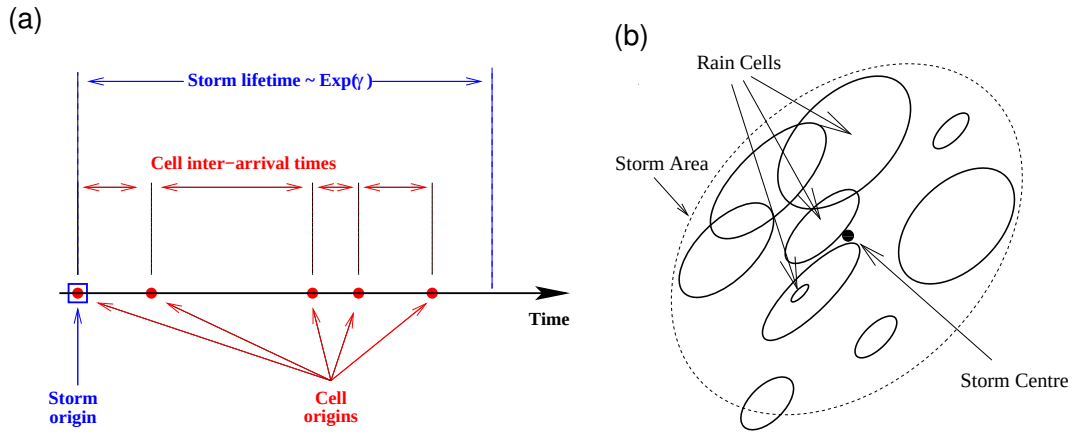


Figure 9.1 Schematic diagram of a storm in the space-time model for the interior of a rain event. (a) Temporal structure: cell origins occur in a Poisson process during the lifetime of a storm, with a cell at the storm origin. (b) Spatial structure: each cell is elliptical, and is displaced from the (moving) storm centre according to a bivariate Gaussian distribution with the same elliptical shape.

Gaussian shape could be used (Northrop and Stone, 2005). Detailed assumptions about the shape of a cell are in any case unlikely to be important except at very short time scales.

According to the model description above, the rainfall intensity at spatial location u and time t can be written as

$$Y(u,t) = \int_{\tau=-\infty}^{t} \int_{w \in \mathbb{R}^2} I(w,\tau;u,t) X(w,\tau) dN(w,\tau) \tag{9.1}$$

where N is the cluster process of cell origins, $X(w,\tau)$ denotes the depth of a cell with origin (w,τ), and $I(w,\tau;u,t)$ is an indicator taking the value 1 if this cell covers u at time t, and zero otherwise. This representation is the key to deriving the statistical properties of the model. For example, the mean rainfall intensity is

$$\mathrm{E}[Y(u,t)] = \lambda \mu_C \mathrm{E}(X)\mathrm{E}(D)\mathrm{E}\left(\pi A_c^2 \sqrt{1-E^2}\right) ,$$

as is intuitively obvious, given that the final factor represents the expected spatial extent of a rain cell.

In general, to derive second-order (covariance) properties of the model when V and E are non-zero, it is helpful to define a linear coordinate transformation $\mathcal{L} : (u,t) \rightarrow (u^*,t)$ such that, in the new space-time coordinate system, cells are circles with radius A_c and the velocity is zero. It is also convenient to assume that the cell duration, D, is exponentially distributed, since in this case the model has a temporal Markov structure. In addition, we assume that the cell scale, A_c, has a gamma distribution; the simpler, exponential form is not appropriate here because the zero mode of the exponential distribution means that large numbers of unrealistically small cells are likely to be generated. Similarly, we assume a gamma distribution for the storm scale, A_s. The second order properties of the model only require the corresponding properties of X, so no particular distributional form need be assumed in this case, although a parameter can be saved by assuming for example that var(X)= E$[(X)]^2$ (as for the exponential distribution).

Within a rain event, all cells have the same velocity, eccentricity and orientation so that, for fitting this part of the model, these can be treated as parameters (v, e, θ respectively) with no distributional forms assumed. In this case, derivation of the autocovariance function is simplified as all contributions come either from single cells overlapping both space-time points, or from distinct cells within a single storm. However, allowance for variation between events will be required subsequently, for simulation of long space-time rainfall sequences. Further details of the model and derivations of its properties are given by Northrop (1998); note however that this paper contains an error in the expression for the space-time covariance function $c_s(u,t)$ for the model with circular cells and zero velocity, when $t > 0$ (Section 3.1.1).

The problem arises in the evaluation of the temporal integrals

$$I = \int_{\tau_1=-\infty}^{0} \mathcal{F}_D(-\tau_1) \int_{\tau_2=-\infty}^{t} \mathcal{F}_D(t-\tau_2)\, \mathrm{e}^{-\gamma|\tau_2-\tau_1|}\, \mathrm{d}\tau_1\, \mathrm{d}\tau_2$$

in Northrop's equation (2.10, page 1882 of the paper), where $\mathcal{F}_D(u) = \mathrm{e}^{-\eta u}$ is the survivor function of the exponential distribution of cell duration. The correct form for I is $[\gamma \mathrm{e}^{-\eta t} - \eta \mathrm{e}^{-\gamma t}]/[\eta(\gamma^2 - \eta^2)]$ rather than $\mathrm{e}^{-\gamma t}/[\eta(\gamma+\eta)]$ as claimed. Thus, the covariance given in Northrop (1998) as $Ae^{-\eta t} + Be^{-\gamma t}$ for appropriate A and B, can be corrected by replacing $\mathrm{e}^{-\gamma t}$ in the second term by $[\gamma \mathrm{e}^{-\eta t} - \eta \mathrm{e}^{-\gamma t}]/(\gamma - \eta)$. We note that the incorrect expression will have affected the model fitting and continuous simulation results given in Wheater et al. 2000b.

9.2.2 Model Fitting

The model described above has been fitted to data recorded by the Chenies radar for the interiors of a large number of observed rain events. These data, and the selection of fitting events that are consistent with the model assumptions as regards the stationarity of rain event interiors, will be described further in Appendix F below. The radar data can be treated as temporally instantaneous but spatially averaged. If the radar pixels have dimensions $h \times h\mathrm{km}^2$, then the data have the form

$$Y_{ij}^{(h)}(t) = \frac{1}{h^2} \int_{(i-1)h}^{ih} \int_{(j-1)h}^{jh} Y(u,t) du_2 du_1 \tag{9.2}$$

where $u = (u_1, u_2)$. The first and second order properties of these variables can be derived straightforwardly from those of $Y(u,t)$.

For any given rain event, the model description involves 4 random variables (X, D, A_c and A_s) as well as the parameters $\lambda, \beta, \gamma, e, \theta$ and the two components of the velocity v. Thus the model has at least 11 parameters. If two-parameter distributions are used for the cell intensity X, the cell scale A_c and the storm scale A_s, the number of parameters rises to 14. In the modelling reported here, an exponential distribution is assumed for cell intensities so that Var(X)=[E(X)]^2. Following exploratory data analysis, the gamma distributions for A_c and A_s are approximated by one-parameter distributions with the scale parameters held fixed. Thus 11 parameters remain to be determined.

For this type of model, likelihood-based fitting is infeasible because the density of the spatially aggregated data is not available in a computationally useful form. Moreover, the model's constant cell intensities impose short term deterministic features that are

not present in real rainfall. Although such features are hydrologically insignificant, they may adversely affect the performance of methods that attempt to maintain fidelity to all aspects of the data (Rodriguez-Iturbe et al., 1988); likelihood based fitting may therefore be undesirable. As an alternative, we choose to use a generalised method of moments (GMM), in which parameters are chosen to provide as close as possible a match between selected sample statistics and the corresponding model properties.

In the work reported here, models are fitted using the mean, $\mathrm{E}(Y^{(h)})$, the variance $\mathrm{var}(Y^{(h)})$, and the space-time autocorrelation function

$$\rho^{(h)}(k,\tau) = \mathrm{Corr}\left(Y_{ij}^{(h)}(t), Y_{i+k_1\, j+k_2}^{(h)}(t+\tau)\right)$$

where $k = (k_1, k_2)$, for various levels of aggregation (h) and spatial and temporal lags, k and τ respectively. Minimisation of the objective function (2.1) is not feasible analytically, so numerical methods must be used. Some reparameterisation helps to improve stability at this point. In particular, β is replaced by the mean number of cells $\mu_C = 1 + \beta/\gamma$; $\mathrm{E}(A_c)$ by the mean cell area $\mu_A = \pi\sqrt{1-e^2}\mathrm{E}\left(A_c^2\right)$; and $\mathrm{E}(A_s)$ by the mean storm area μ_S (defined in terms of the area in which approximately 40% of the rain cell origins are expected to fall).

Fitting is carried out in two stages (see Wheater et al. 2000b for more details). First, the velocity v is determined from the cross-correlation $\rho^{(h)}(k,\tau)$ using the property that $\rho^{(h)}(k,\tau)$ is maximised when $k = v\tau$. Because the arguments of $\rho^{(h)}$ can only take a discrete set of values a centroid estimate is obtained. The velocity is then treated as known in the second stage of fitting.

This second stage exploits the fact that the purely spatial autocorrelation function $\rho^{(h)}(k,0)$ depends strongly on the spatial parameters μ_A, μ_S, e and θ (e and θ determine the shape of the spatial autocorrelation contours, while μ_A and μ_S control their decay rate). Similarly, for given v, the temporal parameters η and γ control the decay of the maximised cross-correlation $\rho^{(h)}(v\tau,\tau)$. The remaining parameters λ, μ_C and $\mathrm{E}(X)$ influence the mean and variance of the rainfall field.

Further details of the iterative process used for parameter estimation will not be given here (the interested reader is referred to Wheater et al. 2000b for further information) and we simply note changes from the procedures described in that report.

- As described above, the expression for the space-time covariance function $c_s(u,t)$ has been corrected.

- The updated programs optimise over the mean cell and storm area (μ_A and μ_S) only at the final step of the iteration. Also, the values of these parameters are no longer estimated by solving equations, as in the original report, but by numerically minimizing the difference between the observed and fitted maximum spatial correlation ($\rho(vt,t)$). This change is necessitated by the correction to the expression for the space-time covariance.

- To improve the numerical optimization of the objective function, initial values for the eccentricity are not allowed to be too close to 1. The new initial value is set as $\min(e_0, 0.8)$, where e_0 is the initial estimate as calculated in the original report.

- Unrealistic parameter sets are excluded. This is not done by constraining the optimization, but rather by checking the optimized values and excluding the sets that do not satisfy the following requirements.
  - Cell duration: $\mu_d \geq 1$ min
  - Storm duration: 5 min $\leq \mu_l \leq 100$ hr
  - Number of cells: $1 < \mu_c < 100000$
  - Storm area: $\mu_s < 10000$ km^2
  - Cell intensity: $10^{-8} \leq \mu_x < 100$ mm/hr

9.3 Durations of events and dry periods

In the previous section, models for the interior of a single rain event have been described. We now consider a model for the sequence of events as they move across the fitting window, focussing first on the times of event arrivals and departures, and subsequently on the shapes of the events as they affect the window. We give only an outline of the model, which has already been described in Wheater et al. 2000b, together with a note of the changes that have been made in the present work.

On the basis of a 4 year sequence of radar data from a station at Wardon Hill in South West England, Wheater et al. (2000b) distinguished two types of event producing non-negligible amounts of rainfall. The first type extended over the fitting window for a period during which the event interiors appeared spatially and temporally stationary, and were therefore suitable for model fitting. By inspection of radar sequences, it was found that the second type of event generally appeared to have a similar structure to the first, but followed a path that only intersected an edge or corner of the window and could not therefore be used for fitting. A semi-Markov model for the sequence of durations of the two types of events, and the dry periods in between them was proposed and fitted.

The duration of an event was defined as the time during which the coverage of the fitting window (i.e. the proportion of the fitting window that is wet) was above 25%. The threshold choice is to an extent subjective, requiring a balance between the non-detection of events if the threshold is too high, and mis-interpretation of radar noise as precipitation if it is too low. Too low a threshold also allows the possibility that a sequence of events following each other rapidly across the window may be mis-identified as a single event, for which the stationarity assumption is likely to be questionable.

For the Chenies radar data (which have rather different characteristics from those from Wardon Hill) we have found that a threshold coverage of 15% for the definition of an event is a satisfactory compromise, and enables a simpler model for event sequences in which only a single type of event is defined. The reasons for this choice are discussed further in Chapter 10. We assume that the durations of the events over the window, and of the 'dry' periods that alternate with them, form independent sequences of independent and identically distributed random variables; thus the sequence of wet and dry period durations forms an alternating renewal process (Cox and Isham, 1980,

Section 3.2). The independence of successive durations has been verified empirically using scatterplots (examples are given in Chapter 10).

In Wheater et al. (2000b) it was demonstrated, using the rather higher threshold of 25%, that positive radar values below the threshold (both rainfall and radar noise) contribute a relatively minor proportion of the total rainfall intensity over the window. Nevertheless, further research is needed to determine the extent of the influence of very light rainfall on the runoff properties of a catchment, and to adjust the model to allow for such rainfall if appropriate.

In Appendix F, it will be seen that there are many breaks in the Chenies radar sequence so that many of the wet and dry interval durations are censored. Breaks of several hours' duration are not uncommon. We have therefore chosen to model both types of interval using Weibull distributions, for which it is straightforward to include censored data in likelihood estimation of the parameters. To incorporate seasonality, separate alternating renewal models are fitted for each month of the year. The renewal model is otherwise temporally stationarity. The fit of the models can be assessed using standard techniques from survival analysis, such as a comparison of empirical and fitted cumulative hazard functions.

9.4 Stationary continuous simulation of a sequence of rain events

For a continuous simulation of rainfall over a catchment, that is stationary in space and time, the alternating renewal model for wet and dry interval durations must be combined with the stochastic model for event interiors. In this report, we will demonstrate the procedure by simulating rainfall over the whole of the fitting window. As described above, events arrive and depart when the radar coverage crosses a threshold which, for the Chenies radar, is set at 15%. The basic idea of the continuous simulation is first to simulate the alternating renewal model to give the sequence of event durations and dry periods; then, independently for each wet period, an appropriate event interior is simulated. By inspection of radar images we find that, at the spatial scales of interest, the edges of weather systems are approximately linear. These edges define, in some sense, the lines that best separate the 'wet' and 'dry' pixels when the event is just entering (or leaving) the window. Thus,for any observed event, edge orientations can be obtained by applying Fisher's Linear Discriminant Function (see, for example, Krzanowski 1988, Section 12.3) to an image near the start or end of the event. An example is given in Figure 9.2. Note that the rules for choosing the image on which the linear discriminant function is calculated have been simplified from those reported in Wheater et al. (2000b).

For continuous simulation over the fitting window, we therefore treat rain events as moving bands of rainfall, defined via the separation and orientations of their leading and trailing edges that best separate the 'wet' and 'dry' pixels when the event is just entering (or leaving) the window. The width of each band is determined so that the rainfall coverage will be above the threshold for the simulated duration of the event. The details of the continuous simulation are rather complex. In particular, in order to deal with edge-effects that would result from an arbitrary finite truncation of the

rainband, the event interior model is simulated within the area of intersection of the moving rain band and a large fixed rectangle (the 'generation region') containing the area of interest, with a toroidal wrap-around.

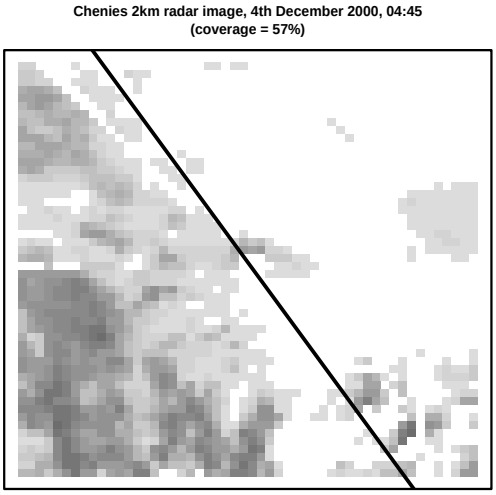


Figure 9.2 The use of Fisher's Linear Discriminant Function to identify the orientation of the leading edge of a rainfall event.The solid line is the linear discriminant; the gradient of this is taken as the orientation of the leading edge.

To simulate an event interior at a particular time in the continuous simulation, a set of parameters is sampled from one of the fitted events for that same month, where different types of event will yield different parameter combinations. However, there is dependence between the model parameters, especially the velocity, and the event duration (in particular, fast moving events are seldom observed to have long durations) and these dependencies should be reflected in simulations. To achieve this, we have constructed a library of the parameters for every fitting event in the Chenies radar record. For each event, the library includes the 11 parameters of the within-event model, the leading and trailing edge orientations and the observed duration of the event. To reproduce the inter-parameter dependence during simulation, a set of parameter estimates is sampled from this library, conditionally on the event duration for that set being close to that simulated from the renewal model.

Further details of the method of continuous simulation will not be given here (the interested reader is referred to Wheater et al. (2000b) for further information) and we simply note changes from the procedures described in that report.

- As noted above, for the Chenies radar data, a threshold of 15% coverage is used to define the arrival and departure of events at the fitting window, rather than the 25% threshold that was appropriate for the Wardon Hill data.

- The rule for selecting the images that are used to estimate the orientation of the leading and trailing edges using Fisher's Linear Discriminant Function has been changed. For Wardon Hill data, the rule was to select the first and last images for which the coverage was above the critical (25%) threshold. It has been found

that using a higher threshold gives a more clearly defined direction, and for the Chenies data we have used a threshold of 60% of the coverage at the start / end of the fitting period (see Chapter 10.1).

- The rules for selecting a parameter set from the library of fitted events from the Chenies radar are different from those applied to the Wardon Hill dataset. For the Chenies data, the new fitting rules (see Section 10.1) require some stationarity criteria to be satisfied for a window of at least one hour. This means that the durations of the fitted events tend to be longer than those of the other wet periods, and therefore the event lengths in the parameter library are not a representative sample from the Weibull distribution fitted to all wet period durations. Thus, for the Chenies data, the parameter set is selected for which the event length is closest to the sampled duration.

- To account for long storms and cells, in the toroidal wrap-around of the simulation rectangle, the maximum number of times a storm or cell can return to the simulation area has been increased to 5, whereas the original program only allowed the storms to return to the simulation area once.

So far, in the work described above, we have developed a model for the continuous simulation of rainfall that is stationary in space and, apart from basic seasonality, in time. In next section, we consider adapting this model to allow for temporal and spatial heterogeneity. The performance of both stationary and nonstationary simulations will be discussed in Chapter 10.

9.5 Continuous simulations that are nonstationary in space and time

So far, in this Chapter, we have described the development of a class of models that can be used for the simulation of a spatial-temporal rainfall. The models have parameters that are easily physically interpretable, and are built in continuous space and time so that their simulation provides high resolution data. However, they are stationary in both space and, apart from basic seasonality, in time. Thus these models do not allow the representation of spatial topographic features that can be hydrologically significant even in small catchments, or possible temporal climate change scenarios.

The Generalised Linear Models (GLMs) that will be described in Part II have the considerable advantage that they can incorporate both spatial and temporal heterogeneity. However, these models are built in discrete space and time (for a set of spatial locations at the daily level). Given appropriate covariates, they can be interpolated to any specific spatial location but they do not have the high temporal resolution that is likely to be needed for some rainfall-runoff situations. An increase of the temporal resolution to, say, the hourly level cannot be easily be achieved, for not only would this impose much greater computational demands but, more importantly, the complexity of the model and the number of covariates involved to represent complicated dependencies at the hourly level would need to increase very substantially.

For input into rainfall-runoff models, it is therefore highly desirable to have the advantages of both types of model, combining high resolution with spatial and temporal nonstationarities and in this section we discuss how this may be achieved. The basic idea is to fit the GLM to rain gauges under the radar, to use simulations of the fitted GLM to determine the statistical properties of nonstationary multi-site daily sequences, and then to condition the continuous simulation of the stochastic model to respect these properties.

Perhaps the easiest way to do this, and the way explored here, is to generate a stationary simulation from the stochastic model and then simply to rescale each image to reflect the spatial structure given by the GLM. More specifically, the GLM is simulated many times to generate the average spatial rainfall at a coarse grid of sites. The number of sites in this grid is somewhat arbitrary, but a guide would be to use roughly the same number as the number of rain gauges used to fit the GLM (the reason for keeping the number of sites for the simulation fairly low is purely to keep the computation time reasonable). Then, by interpolating this coarse grid of values to a higher resolution, a gridded image at the same resolution as the radar data can be obtained for each month in each year of the period of interest, which can then be used for the rescaling. The interpolation is achieved by first regressing the coarse grid values on the site effects that are used as covariates in the GLM, and then using these regressions for interpolation.

Thus the algorithm to obtain the nonstationary simulation is as follows:

1. Fit the GLM to data for a set of rain gauges under the radar.
2. Simulate the fitted GLM many times, to give daily values at a set of sites (e.g. a subset of the centres of the radar pixels) for the time period of interest.
3. For each month in the simulation period, calculate the mean daily rainfall at each site by averaging over simulations.
4. Regress the mean values at each site on the site effects that are covariates in the GLM.
5. Use these regressions to interpolate the average monthly rainfall values, giving an estimated mean field for each month of the simulation period at each point on the high resolution grid.
6. Calculate the ratio of the estimated mean field to the mean for the stochastic model (which is spatially constant). The latter must also be estimated by simulation, since it does not appear feasible to study analytically the contribution of events arriving and departing from the region of interest.
7. Generate a realisation of the stochastic model, and rescale each image using the ratios calculated in step 4.

The results of applying this rescaling algorithm to the stationary continuous simulations will be discussed below in Chapter 10. An additional benefit of this rescaling method is that, since the GLMs are fitted to raingauge data, it automatically guarantees that the final continuous simulation is calibrated to ground truth, even if the

radar data are imperfectly calibrated. This scheme will incorporate the spatial and temporal nonstationarities built in to the GLM; however, the rescaling will not affect the wet and dry space-time regions of the stochastic simulation. At catchment scales this is probably acceptable as far as spatial heterogeneity is concerned. However, it is implausible that, for example, the distribution of dry interval durations in a particular season will remain constant in the presence of appreciable climate change. Therefore a different approach will probably be needed to cope with substantial time trends. One way that that this might be achieved while preserving the subdaily structure of the stochastic model would be to calculate a variety of summary statistics from the GLM simulations, and then to choose stochastic model parameters that reproduce these statistics. There are several possible strategies for this. However, this possibility will not be pursued further in this report.

9.6 Summary

In this Chapter, we have discussed the development of a model for the continuous simulation of spatial-temporal rainfall, taking as our starting point a preliminary model first explored in FD0426. This model comprised two submodels, one for the interiors of rain events, and the other for the advection process determining the passage of rain events over a catchment. In FD0426, only a very short data set of about 4 years from the Wardon Hill radar, was available for model testing. In Chapter 10, results of fitting these submodels to a much more extensive data set from the Chenies radar are described. This testing raised many complex issues. As described in Appendix F, many of these issues were to do with the data themselves and their calibration. Others concerned the methodology used to fit the models and required modifications to the software used. In both cases, considerable time and ingenuity were needed to identify the causes of the problems and to resolve the issues. However, as a result, the procedures developed to fit the spatial-temporal models and to produce continuous simulations are now much more robust and firmly based.

A method has been devised to incorporate spatial and temporal nonstationarities into the continuous simulation of the stationary model, allowing the representation of spatial topography and possible climate change scenarios to be included for the first time. Although this method represents only a first step towards the goal of realistic nonstationary simulations, and will need to be further refined (for example so that the temporal nonstationarities can affect the properties of the wet and dry space-time regions), it is nevertheless a major step forward. For the future, it will be of particular interest to incorporate physically-based climate change scenarios into the rainfall simulations. This can be achieved by conditioning the GLMs on the output of numerical climate models, and research in this area is ongoing.

Chapter 10 Spatial-temporal modelling: fitting and simulation results

In this chapter we report the results from fitting and simulating the spatial-temporal models of Chapter 9 to calibrated data from Chenies. Models are fitted using 5-minute data at $2 \times 2\text{km}^2$ spatial resolution, for the years 1990, 1991 and 1993–2002. Data from before 1990 and during 1992 have not been used, because of quality issues relating to an at-site calibration procedure that renders them effectively unusable (Section F.2.3). The 2km data are available within a circle of radius 76km centred on the radar station; as in Wheater et al. (2000b), we work within the largest square that will fit inside this circle. This square, which we refer to as the *radar window*, has dimensions $104 \times 104\text{km}^2$ (i.e. 52×52 radar pixels).

The stationary space-time modelling scheme consists of two components: the model for event interiors and that for wet and dry period durations. In Sections 10.2 and 10.3 below, these are assessed individually, before giving results for the combined scheme in Section 10.4. Nonstationary simulation is discussed in Section 10.5. Before describing any of these results however, it is necessary to consider how a 'rain event' should be defined for modelling purposes.

10.1 Event definition

To fit the models of Chapter 9 to a sequence of radar data, the following steps are required:

1. Split the radar record into 'wet' and 'dry' periods. 'Wet' periods correspond to rain events in the models.
2. Where possible, fit the event interior model of Section 9.2 to a suitably chosen sequence of data for each 'wet' period. Not all wet periods yield sequences that are suitable for model fitting, for the following reasons:
    - The model is stationary in both space and time. To obtain meaningful parameter estimates therefore, it is necessary to fit to sequences for which the spatial and temporal stationarity conditions are satisfied at least approximately. An event consisting of a single narrow band passing over the radar window could not be regarded as spatially stationary within the window, for example.
    - To obtain reasonably accurate estimates of the rainfall properties used for model fitting, a reasonable number of images is required for which the stationarity criteria are met. Some wet periods are not long enough for the rainfall properties to estimated reliably.
3. Fit Weibull distributions to the durations of wet and dry periods for each month of the year.

Ideally, a visual inspection of radar images would be used to identify 'wet' periods and, for each such period, to identify a suitable sequence for event interior model fitting.

However, with 12 years of 5-minute images available it is simply not practicable to carry out such an inspection; it is therefore necessary to develop automated procedures to carry out these tasks.

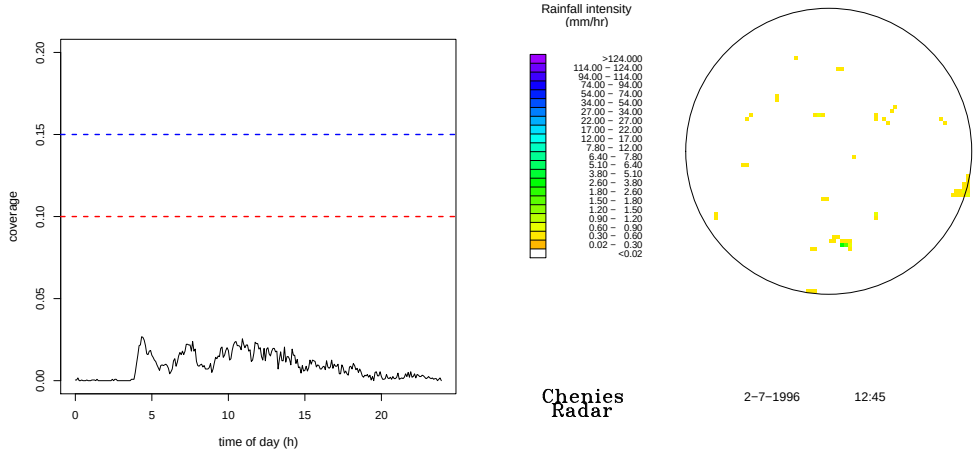
In Wheater et al. (2000b), events in the Wardon Hill radar record were identified using the time series of *coverages* (i.e. the proportions of non-zero pixels in each image). Briefly, an event was defined as starting (arriving) when the coverage crossed a threshold of 25% (chosen to separate real rainfall from radar clutter), and stopping (leaving the radar window) when the coverage crossed this threshold again. Two types of events were distinguished: those that achieved coverages in excess of a second threshold of 50% during their lifetimes (Type I), and those that did not (Type II). The role of the second threshold was to distinguish between events that during their life covered the whole radar window (and hence should meet the spatial stationarity conditions required for model fitting — see above) from those which did not. The event interior model was then fitted to Type I events only, over time periods for which they extended over the whole radar window.

Data sequences from the Chenies and Wardon Hill radars differ in several important respects. In particular, the Chenies data have been pre-processed to remove clutter and anomalies, and subsequently calibrated against observations from gauges (Appendix F). Therefore, different rules for identifying the arrival and departure of events, as well as the fitting periods, have been adopted for Chenies. For example, clutter has been removed from the calibrated Chenies data, so it may be possible to lower the coverage threshold used to define events. The criteria used to determine an appropriate threshold were summarised in Section F.4; further details are given below. Moreover, at Chenies the use of a 50% threshold to identify events suitable for model fitting does not seem particularly effective. A visual inspection reveals that several events with coverages higher than 50% are either bands that cross the radar window, or peripheral events that never affect the entire window. Also, several events with coverage lower than 50% do cover the whole radar window at some stage of their life. The criteria used at Chenies, to determine whether an event is suitable for event interior model fitting, are described in Section F.4.2.

10.1.1 Event arrivals and departures

As noted above, the absence of clutter in the calibrated Chenies data means that event arrivals can be detected reliably using a lower coverage threshold than in Wheater et al. (2000b). We find that a threshold of 8–10% is sufficient to separate rainfall from radar noise (see Figure 10.1, for example). However, lowering the threshold to this extent introduces another potential problem, illustrated in Figures 10.2 and 10.3: if two events follow each other in quick succession so that one enters the radar window just as the other is leaving, the coverage may never fall below 10%. In this case the two events will be treated as a single long wet period. Situations such as that in the left panel of Figure 10.2 are not uncommon, particularly in the winter months. Sequences with large oscillations of the coverage series cannot be regarded as temporally stationary; certainly, the event interior model of Section 9.2 is not designed to reproduce this type of behaviour. However, the interior of each individual rainfall pulse often can be regarded as stationary, as indicated by the images in Figure 10.3; hence it is preferable to regard each pulse as a separate event. This can be

Figure 10.1 (a): Series of low rain coverages associated with radar clutter. The coverage time series is shown as a solid line; the dashed lines correspond to thresholds at 10% and 15%. (b): Example of radar clutter; this image corresponds to one of the time points in (a), and is typical of the appearance of images with very low coverages.



achieved by raising the threshold slightly.

Analysis of the Chenies radar records for 1998 suggests that a threshold of 15% may be adopted as a compromise between retaining as much genuine rainfall as possible, and breaking very long events into smaller stationary ones. Out of 408 rain events identified using a threshold of 10%, 110 have a coverage below 15%. However, most of these are very short: only 28 last for an hour or more. Moreover, only one of these is suitable for model fitting according to the rules described below. The loss of actual rainfall due to raising the threshold (and therefore discarding some short events with very low coverage) is outweighed by the reduction in the number of long events. For example, the initial 10% threshold results in 35 events longer than 12 hours, while this number decreases to 23 when the threshold is 15% (note that this reduction is not achieved by discarding events, but by splitting them up into shorter segments). However, it should be noted that the number of events with extremely long duration (over 24 hours) does not change (6 in both cases). Some of the remaining long events appear temporally stationary, while for others the coverage series has large oscillations but remains above the chosen threshold. In order to split all non-stationary events into shorter stationary ones, the threshold would have to be raised to an unreasonable level, and images with a large area covered by rain would be considered dry.

10.1.2 Fitting periods

To determine whether a particular event is suitable for model fitting, the main criteria are that the event can be regarded as a realisation of a stationary stochastic process over the radar window; and that this realisation is long enough that the summary statistics of that process can be estimated reasonably accurately. As noted above, the rule adopted by Wheater et al. (2000b) (an event is suitable for fitting if its coverage exceeds 50% for at least an hour) seems overly simplistic for the Chenies data. To try

Figure 10.2 Oscillating coverages for two event sequences. In each case the coverage time series is shown as a solid line; the dashed lines correspond to thresholds at 10% and 15%.

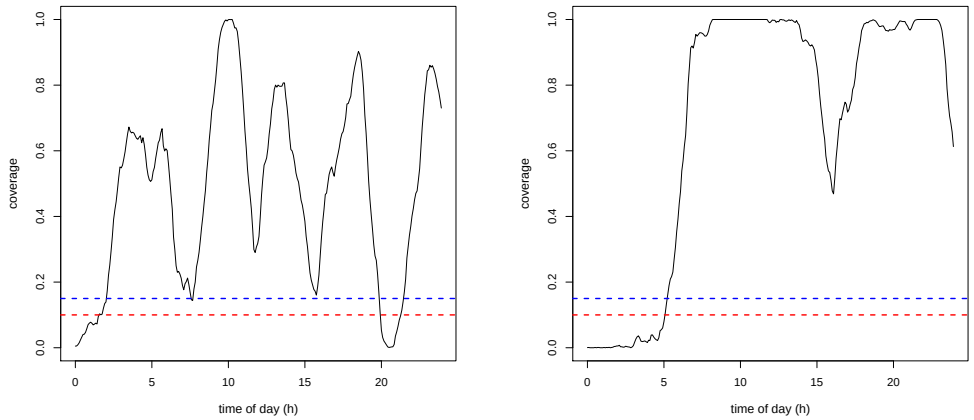


Figure 10.3 An hourly sequence of images from the Chenies radar: a wet spell leaves the radar area while another one arrives. To split this sequence into two separate events on the basis of coverage, a threshold of at least 15% is required.

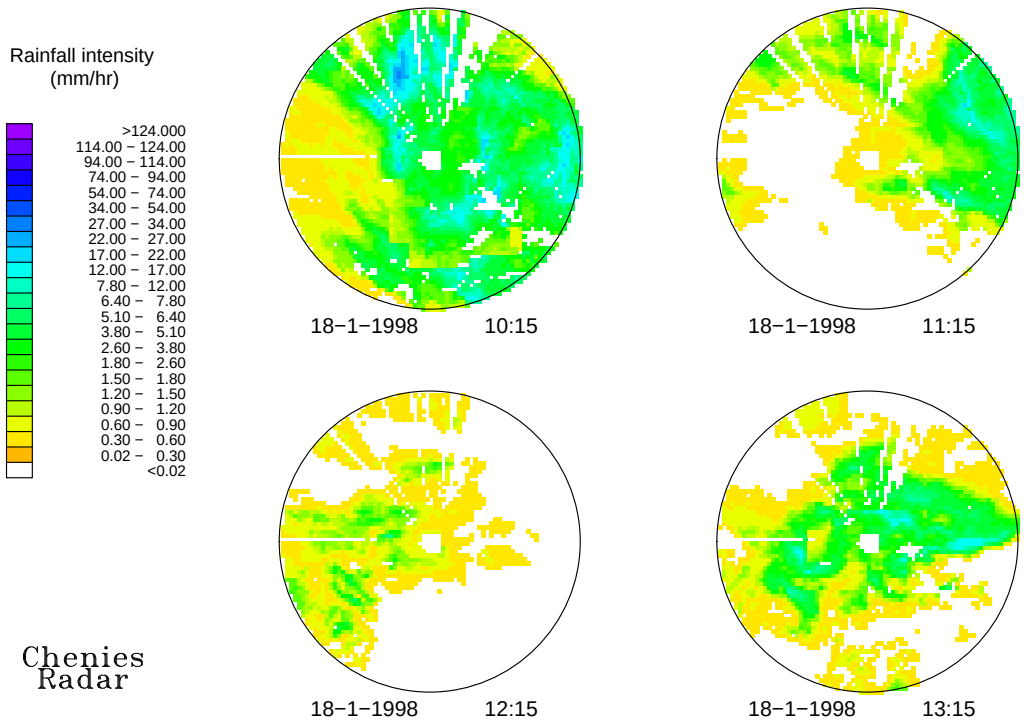


Table 10.1 Criterion for spatial stationarity: an image with a given overall coverage is deemed to be spatially stationary if the coverages for each of four subsquares are within the range indicated.

Overall coverage	Range of subsquare coverages
0.15 – 0.20	(0.05, 0.40)
0.20 – 0.30	(0.12, 0.45)
0.30 – 0.40	(0.18, 0.54)
0.40 – 0.50	(0.25, 0.65)
0.50 – 0.60	(0.35, 0.75)
0.60 – 0.70	(0.46, 0.84)
0.70 – 0.80	(0.58, 0.92)
0.80 – 0.90	(0.68, 1.00)
0.90 – 0.95	(0.78, 1.00)
0.95 – 1.00	(0.82, 1.00)

and improve upon this, several alternative strategies have been investigated. In each case, the classification of events was assessed by comparing with a visual inspection of Chenies events for the year 1998. None of the strategies gave a perfect classification of the events, or always chose the intuitively optimal (from a visual inspection) period for model fitting. However, the following algorithm provides reasonable results:

1. Identify all images within an event (i.e. all images for which the coverage exceeds 15%).

2. To ensure temporal stationarity, eliminate all images that correspond to a rain coverage smaller than p, where p is the solution of

$$p + k\sqrt{p(1-p)} = p_{\max}.$$

   The arbitrary constant k is set to 0.3 and $p_{\max}$ denotes the maximum coverage achieved by the event being considered. This criterion is motivated by the observation that an event is most likely to be stochastically stationary when it achieves its maximum coverage $p_{\max}$, and that a period of temporal stationarity can therefore be defined as the period for which the coverage is close to the maximum. The form $k\sqrt{p(1-p)}$ is motivated by the observation that the variability in coverage depends on p itself — hence more variability is expected (and allowed) in events with a coverage close to 0.5 than in events with coverage close to 1.

3. To ensure spatial stationarity, divide the radar window into four sub-squares, and require that the coverage in each of these sub-squares is ‘close’ to the overall coverage for the event. Various attempts have been made to define ‘close’ for this purpose: the most effective seems to be to define a table of limits depending on the overall coverage. The limits used here are given in Table 10.1.

   Any image that does not meet the spatial stationarity requirement is eliminated from the event.

4. If, after the previous stages, an event retains a consecutive sequence of images lasting at least one hour, this sequence is considered suitable for model fitting.

10.1.3 Implications for simulation

Under the procedures described above, rain events are classified into two types depending on whether or not they are suitable for model fitting. This is the same as in Wheater et al. (2000b). However, the rules used to determine fitting events are now able to recognise a wider variety of events than previously — for example, these rules will recognise an event consisting of widely scattered showers as being suitable for fitting, whereas the previous classification would not. Previously, no attempt was made to simulate rainfall within Type II events, since not enough was known about their characteristics (Wheater et al., 2000b, Section 3.2.3). However, with a lower coverage threshold for defining events, and with an improved algorithm for identifying suitable fitting periods, model parameters are available for a wider population of events than previously. Therefore, in the work reported below we have regarded the fitting events as a representative sample of all events. This eliminates the need to consider the two types of events separately in simulations. However, it should be noted that this could lead to poor simulation performance if the fitting events have substantially different properties from the remainder.

10.2 Event interior model: fitting and performance assessment

The procedure described in Section 10.1.2 above has been applied to the entire calibrated Chenies record from 1990–2002, excluding 1992, to identify events that can be used to fit the event interior model. Of a total of 4388 events in the radar record, just 380 events are identified as suitable for fitting. Note, however, that some of these 'events' are extremely short, with a single image above the 15% coverage threshold. In this section we summarise the fitting results for the 380 selected events, and examine the ability of the model to reproduce the properties of events to which it is fitted.

10.2.1 Model fitting

The model has been fitted to each of the 380 events individually, using a generalised method of moments as described in Section 9.2. The results have been used to construct a library of parameter sets, obtained from events where the fitting was successful. One criterion for judging the success of a fit is that the resulting parameter estimates should be within limits that are deemed to be physically realistic (these limits are also given in Section 9.2). The fitting may fail, therefore, if it yields unrealistic parameter values. This occurred for 83 of the 380 events. For a further two events, the software failed to find suitable initial values for the parameters. In total therefore, the fitting failed for 85 events, leaving a total of 295 parameter sets in the library.
Table 10.2 breaks down the fitting failures by their cause. The six cases where $\mu_d = \mu_l$ involve a division by zero, as the space-time covariance has a $(\mu_d - \mu_l)$ term in the denominator. If this occurs while fitting, the software issues a warning message and deems the fit to have failed. In fact however, in all six cases the values of both μ_d and μ_l were estimated as zero to the available precision — hence it could be argued that in

Table 10.2 Causes of failure when fitting the event interior model

Cause of Failure	Number of Cases
$\mu_d < 1$ min	6
$\mu_l < 5$ min	29
$\mu_l > 100$ hr	12
$\mu_c < 1$	1
$\mu_x \geq 100$ mm/hr	1
$\mu_d < 1$ min & $\mu_c \geq 100000$	21
$\mu_l \geq 100$ hr & $\mu_c \geq 100000$	1
$\mu_x < 10^{-8}$ mm/hr & $\mu_c \geq 100000$	5
$\mu_d = \mu_l$	6
no initial estimate	2
TOTAL	85

practice these events should have been excluded in any case, under either of the criteria '$\mu_d < 1$ min' or '$\mu_l < 5$ min'.

To some extent, the limits used to define 'realistic' parameter values are arbitrary, and if these were changed then some of the 85 events in Table 10.2 could be be added to the library. For example, up to 19 of the 29 events with $\mu_l < 5$ minutes could be included if it was considered reasonable to impose a smaller limit for the mean storm length. The remaining 10 events could not be included, however, since in each case the estimate of μ_l is zero to the available precision.

We have found that in many of the 83 cases where an event yields unrealistic parameter estimates, the properties of the event are reproduced poorly by the event interior model with the estimated parameters. By restricting the parameter values to reasonable intervals, poorly fitting events are therefore excluded automatically from the library. Of course, this does not guarantee that the model is able to reproduce the properties of the events where the fit is deemed to be successful. The only way to determine this is to examine the fit of each event individually, as described below.

10.2.2 Performance assessment

Following Wheater et al. (2000b), we assess the fit of the model to a given event in three ways:

1. By visually comparing the calibrated radar images for the event with images simulated by the model.
2. By examining the ability of the model to reproduce the summary statistics used in the fitting procedure.
3. By examining the ability of the model to reproduce properties that were not used in the fitting procedure.

Clearly, it is not feasible to carry out these checks for all of the events to which models have been fitted. Instead therefore, a small subset of events was identified prior to

Table 10.3 List of Events Selected for Model Performance Evaluation

Fitting Period	Total Duration (Minutes)	Mean Intensity (mm/hr)
03/02/1990 07:40–10:15	550	7.03
19/11/1991 09:05–10:00	1155	5.44
16/06/1993 11:40–12:55	320	3.95
22/02/1994 16:40–17:40	1560	0.47
28/02/2000 02:45–04:00	600	1.09
04/12/2000 07:00–08:10	205	1.68
02/04/2000 19:35–10:40	3060	1.42
16/03/2001 22:50–0035	970	3.25
15/03/2002 11:20–12:15	350	1.41

Table 10.4 Summary statistics for the mean duration and intensity of all 380 events considered suitable for fitting the event interior model.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Duration (minutes):	60	340	515	600	770	3060
Mean intensity (mm):	0.05	0.73	1.23	1.37	1.83	7.03

fitting, and the checks have been carried out where possible on all events from this subset. The events in the subset were selected to encompass a range of different rainfall characteristics. Table 10.3 gives the dates, durations and mean intensities associated with these events. For comparison, Table 10.4 gives summary statistics for the durations and intensities of all events considered for fitting. The quartiles presented here can be used to categorise events as ‘short’ (duration below first quartile), ‘long’ (duration above third quartile), ‘light’ (intensity below first quartile) and ‘intense’ (intensity above third quartile). for example, the event of 16th June 1993 is short and intense; that of 22nd February 1994 is long and light; and that of 19th November 1991 is long and intense. The other selected events represent other combinations (including ‘average’ events such as that of 28th February 2000).

Unfortunately, the fitting procedure failed for two of the selected events in Table 10.3. For the event of 3rd February 1990, the estimate of μ_l was less than 5 min; the event of 16th June 1993 yielded $\mu_d < 1$ min and $\mu_c \geq 100000$. It may be noteworthy that these are both events with high mean intensity (although the fitting was successful for the event of 19th November 1991, which also has a high mean intensity).

For the remaining events, the fits have been checked as indicated above. The first check is a visual comparison of simulated sequences with the original radar data. To illustrate this, Figure 10.4 shows part of the observed sequence from the event of 4th December 2000; Figure 10.5 shows a corresponding specimen simulation from the fitted model. The results are typical: the two sequences share broadly similar features and appear to have similar spatial organisation, although on closer inspection the cellular structure of the simulation is apparent.

Figure 10.4 Part of the observed sequence of calibrated radar images from Chenies, for the event of 4th December 2000. The images show clearly where anomalies have been removed during the calibration process (described in Appendix F).

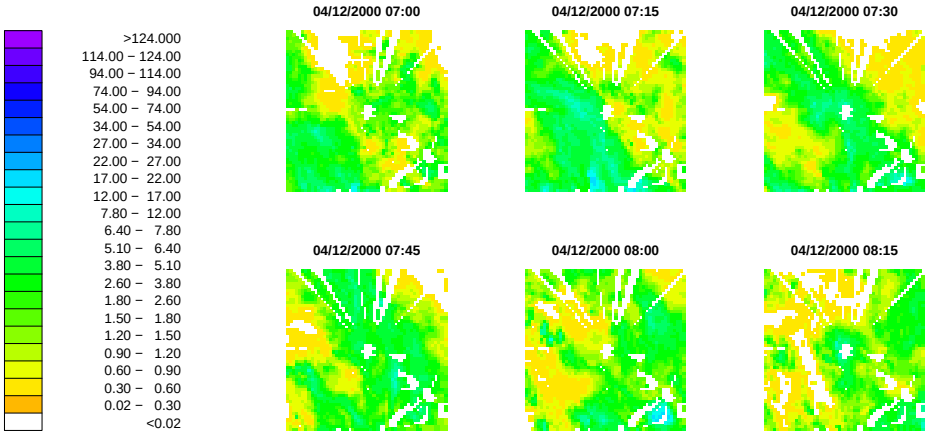
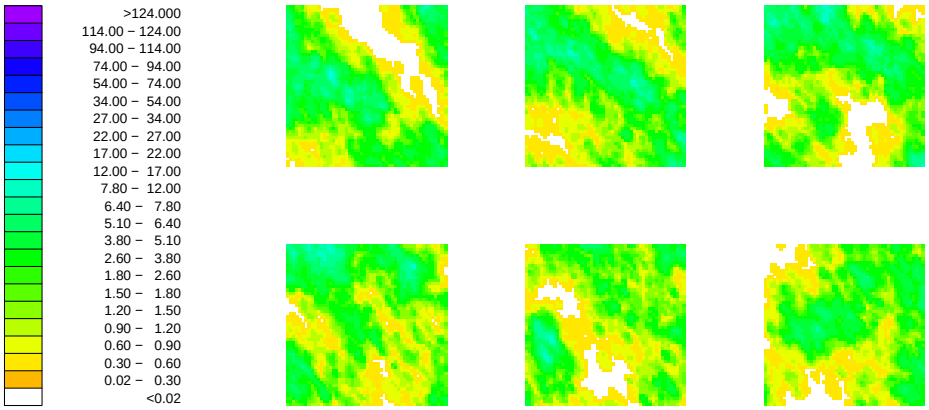


Figure 10.5 Specimen simulation of the model fitted to the event of 4th December 2000. The time separation of the images here is 15 minutes, as in Figure 10.4.



Properties used in the model fitting

The main properties used in fitting the model (apart from the mean, which is reproduced exactly by the fitting procedure) are as follows:

- The variance of spatially aggregated rainfall intensity at scales of 2×2 and 8×8 km^2.
- The space-time autocorrelation function $\rho(\mathbf{v}t, t)$, for $t = 10$ and 20 min. Here, $\mathbf{v}$ is the estimated velocity of the event so that $\rho(\mathbf{v}t, t)$ is effectively the purely temporal autocorrelation in the direction of storm movement. According to the model, $\rho(\mathbf{v}t, t)$ maximises the space-time autocorrelation at temporal lag t. It is therefore referred to as the "maximum spatial autocorrelation" below.
- High spatial autocorrelations (the precise lags used in the fitting depend on the rate of decay of the spatial autocorrelation function in the original data).

These properties should therefore be reproduced almost exactly by the fitted models. It is also expected that the variance of spatially aggregated intensity, the spatial autocorrelation function and the maximum spatial autocorrelation will be reasonably well fitted at other lags, at least up to the limits considered in the fitting. In addition to these properties, the purely temporal autocorrelation function $\rho(\mathbf{0}, t)$ may be used as a check.

Figures 10.6 to 10.12 show the observed and fitted values of these various properties, for each of the events in Table 10.3 where the fitting was successful. For the fitted version of $\rho(vt, t)$, the spatial lag is approximated to nearest number of pixels, for comparison with the observed correlations.

The results of these checks can be summarised as follows:

Variance of spatially aggregated data: for all of the events considered, there is excellent agreement between the observed and fitted variances, at spatial scales up to 40×40km^2 which is well beyond the largest scale (8×8km^2) used in the fitting.

Temporal autocorrelation function: there is good agreement between data and model at lags of 5 and 10 minutes in all cases; for some of the events however, there is a tendency for the fitted model to underestimate the temporal autocorrelations at lags of 15 minutes and above. The exceptions are the event of 4th December 2000, where the agreement between the two autocorrelation functions is excellent at lags up to 30 minutes (the highest lag considered here); and that of 16th March 2001, for which the model overestimates the temporal autocorrelation at lags above 10 minutes. Note that the model can only generate decaying autocorrelation functions, and hence is not capable of reproducing the observed autocorrelation function for the event of 22nd February 1994 (Figure 10.7), which has a minimum at lag 15 minutes. It is also worth noting that the observed autocorrelations are subject to sampling error. Nonetheless, the fact that the model underestimates in five out of seven cases suggests that this is an area for possible future improvement.

Figure 10.6 Observed and model based variance, temporal autocorrelation, maximum spatial correlation and spatial correlation for the event of 19/11/1991

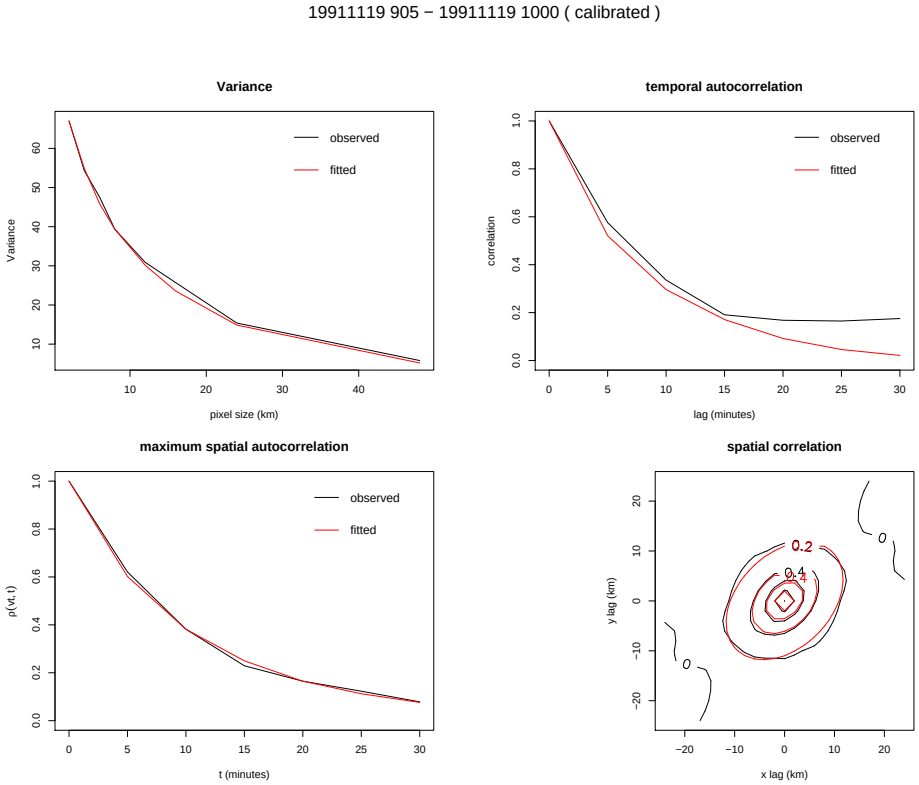


Figure 10.7 Observed and model based variance, temporal autocorrelation, maximum spatial correlation and spatial correlation for the event of 22/02/1994

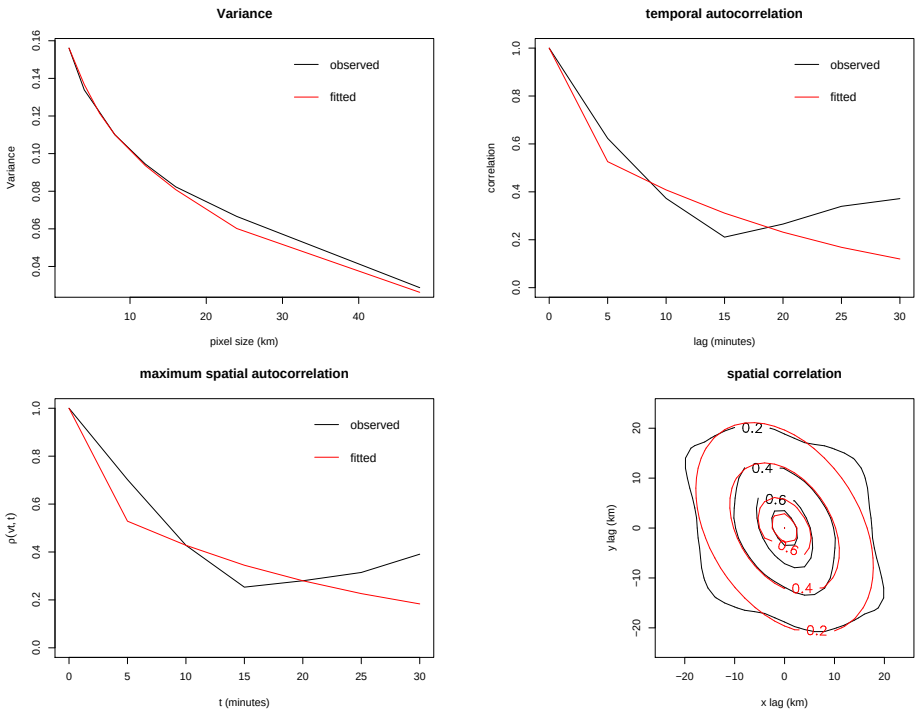


Figure 10.8 Observed and model based variance, temporal autocorrelation, maximum spatial correlation and spatial correlation for the event of 28/02/2000

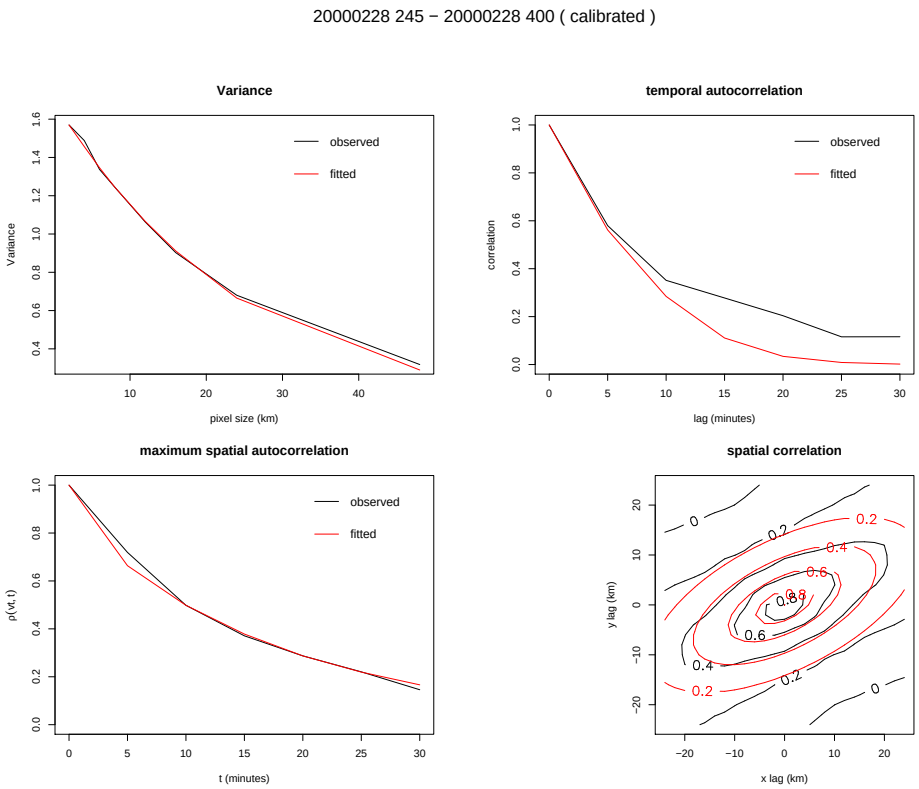


Figure 10.9 Observed and model based variance, temporal autocorrelation, maximum spatial correlation and spatial correlation for the event of 02/04/2000

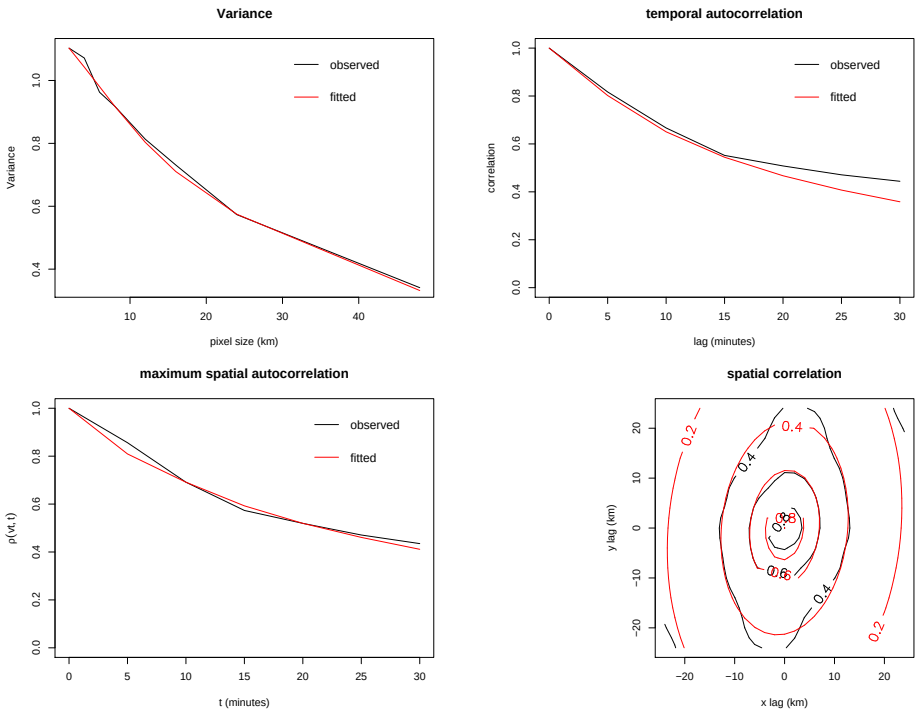


Figure 10.10 Observed and model based variance, temporal autocorrelation, maximum spatial correlation and spatial correlation for the event of 04/12/2000

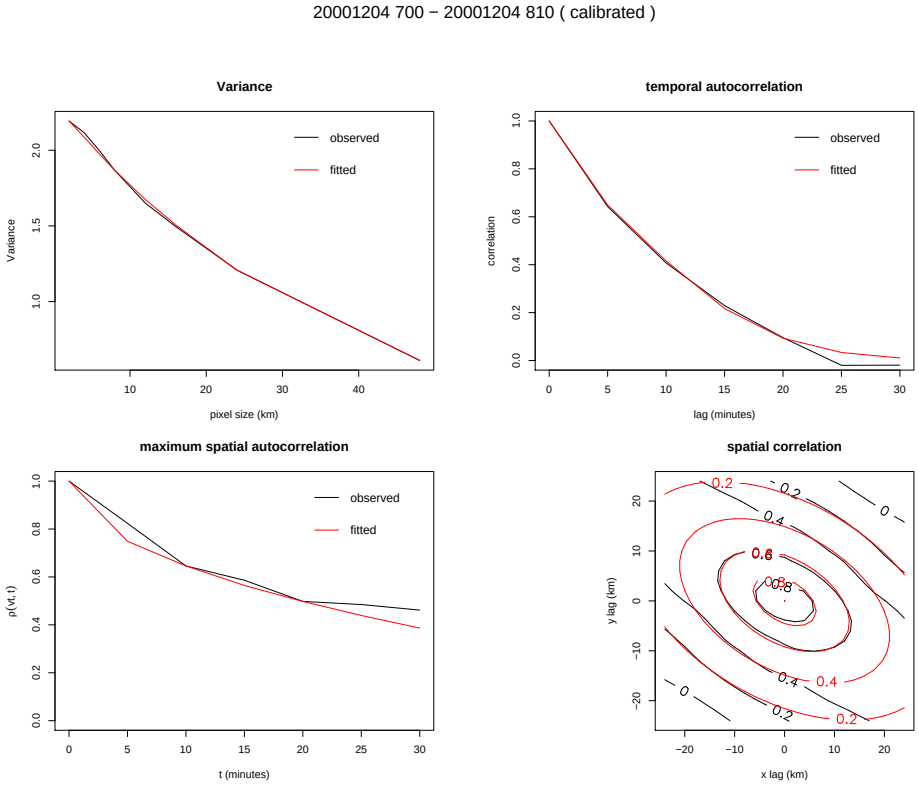


Figure 10.11 Observed and model based variance, temporal autocorrelation, maximum spatial correlation and spatial correlation for the event of 16/03/2001

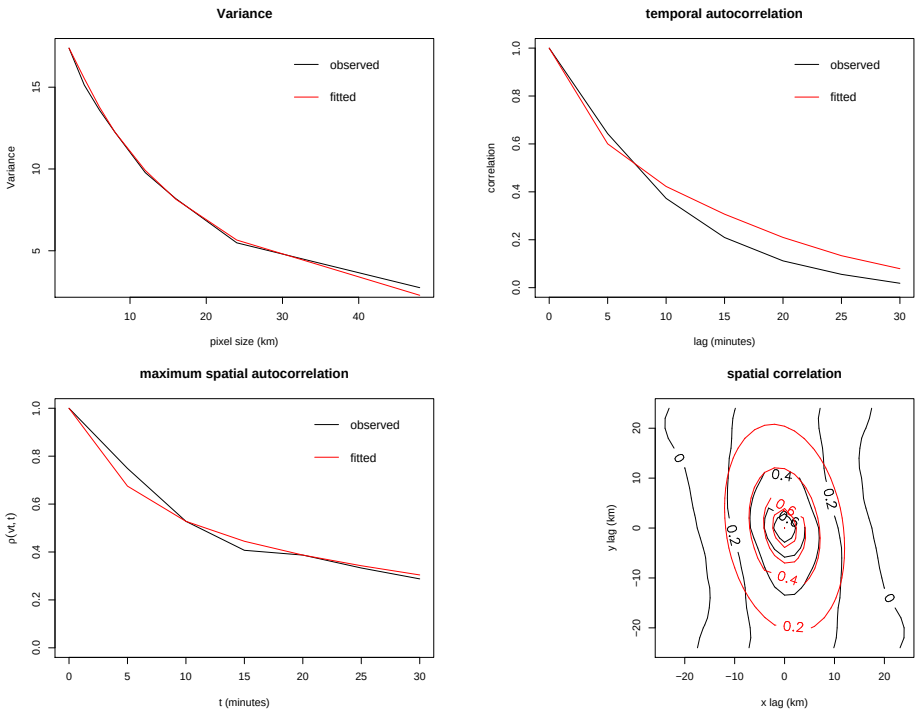
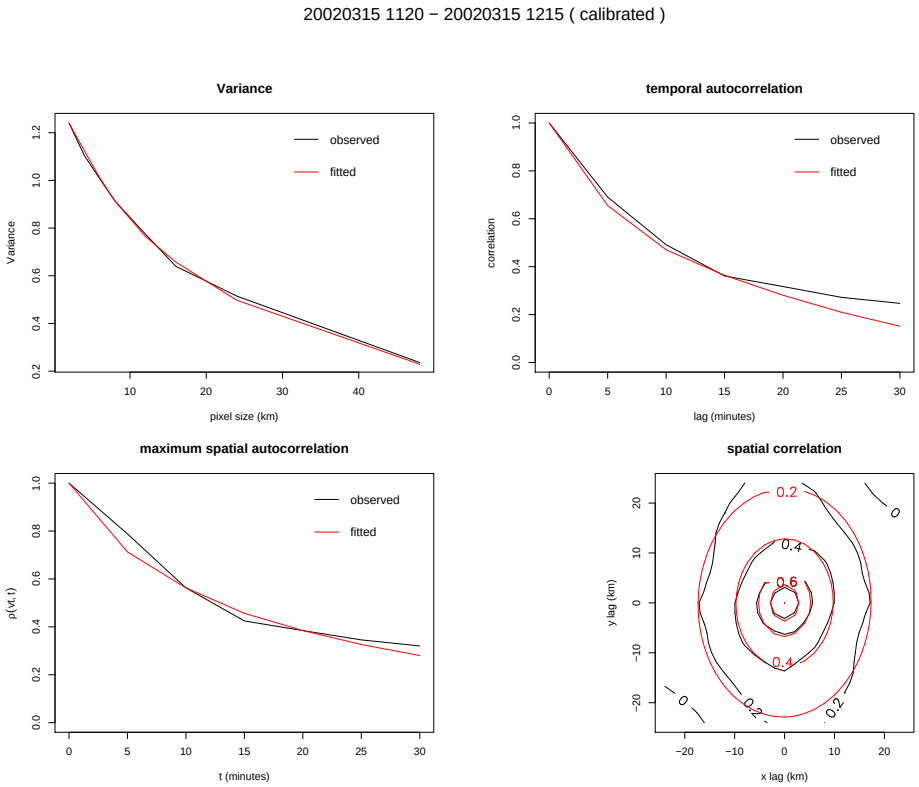


Figure 10.12 Observed and model based variance, temporal autocorrelation, maximum spatial correlation and spatial correlation for the event of 15/03/2002



Maximum spatial autocorrelation: the agreement between observed and fitted values is excellent, except for the event of 22nd February 1994 (again, the model is not capable of reproducing a pattern such as that observed, which could in any case be an artefact of sampling error). Although the fits are constrained to equal the observations at times $t = 10$ and 20 minutes, the plots show that the fitted curves follow the observations closely at other times.

Spatial autocorrelation function: once again, there is excellent agreement between observations and model, at least at lags where the spatial autocorrelations are high. It is interesting to note, however, that for three of the events (28th February 2000, 4th December 2000 and 16th March 2001) the shapes of the observed spatial autocorrelation contours change quite dramatically at higher spatial lags; this is another feature that the model cannot capture. One possible way to accommodate it would be to relax the assumption that storms and cells have the same elliptical shape.

These results suggest that where the fitting procedure has been successful, the model is indeed able to reproduce the properties used in the fitting, as well as other closely related properties. As a first check on the model and the software, this provides reassurance that the model has the required flexibility to cope with a variety of different event types, and that the implementation in the fitting software is correct. The main problem highlighted here is that the model tends to underestimate temporal

autocorrelation at lags above 15 minutes. However, given that temporal autocorrelation is one of the features that is most sensitive to radar data calibration, at this stage there seems little point in trying to improve the model in this respect without first resolving the calibration issues.

Other properties

As well as checking the first and second order properties of the models, it is of interest to examine other properties not used in fitting. Some of these must be determined by simulation because analytic expressions are not available — these include extremes, the probability that a pixel is dry, and the coverage. Properties such as the coverage, the mean intensity and the conditional mean intensity (taken over only those pixels that are wet) are average properties and do not provide information about the spatial organisation of the rainfall field. It is therefore informative to take an increasing sequence of thresholds and, for each, to look at these properties for fields of threshold exceedances. This provides, for example, an assessment of the models' ability to reproduce localised areas of heavy rainfall. To provide a reference point, it may be useful to note for an exponential distribution, the mean exceedance over any threshold is constant.

In this type of analysis, for each event we study properties of both instantaneous and cumulative rainfall fields (the cumulative fields are defined by accumulating rainfall totals at each pixel, throughout the fitting period for the event). To derive the model properties, we simulate from the event interior model over a length of time equal to the period used for model fitting; several simulations are used to illustrate the inherent variability.

Figures 10.13–10.19 show both the instantaneous and cumulative versions of the coverage and mean exceedance over threshold, for the same seven events as in Figures 10.6–10.12. In these figures, the solid lines represent the observations, and each dashed line represents a single simulation from the fitted model. The thresholds considered in each case are up to the 90th percentile of the corresponding intensity distribution, since sampling variability starts to dominate beyond this. The results can be summarised as follows:

Instantaneous coverage: the decay of the coverage with threshold varies enormously between events. Nonetheless, the simulations do a reasonable job of reproducing the observed profiles in all cases. For a couple of the events, the simulations underestimate the observed coverage with a threshold of zero (i.e. the proportion of the area receiving rain); but the discrepancy disappears when the threshold is raised slightly, to 0.5 or 1 mm $\mathrm{hr}^{-}1$; it therefore appears associated primarily with very small rainfall amounts.

Instantaneous mean over threshold: for all but two of the events, the model does a good job of reproducing the observed profile. The problematic events are those of 19th November 1991 (Figure 10.13) and 28th February 2000 (Figure 10.15). For the former, the simulations all show a substantial jump at the left hand end of the plot; this indicates a large number of very small simulated values, as noted already when discussing discrepancies in coverage. Moreover, the

Figure 10.13 Coverage and mean exceedance over threshold, instantaneous and cumulative, for the event of 19/11/1991.

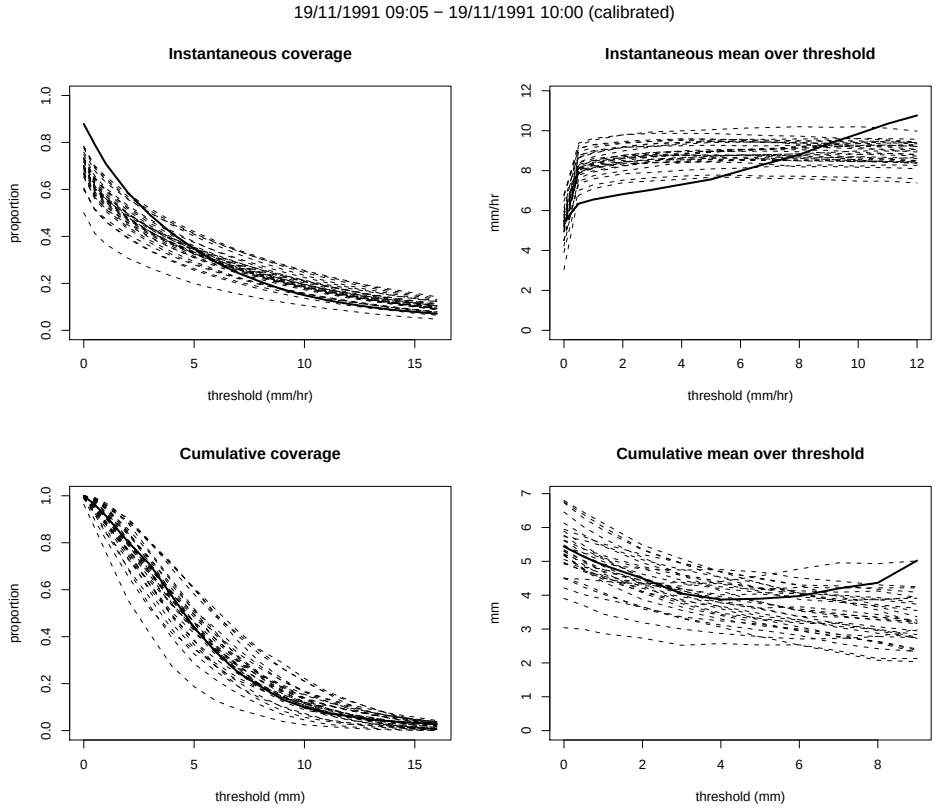


Figure 10.14 Coverage and mean exceedance over threshold, instantaneous and cumulative, for the event of 22/02/1994.

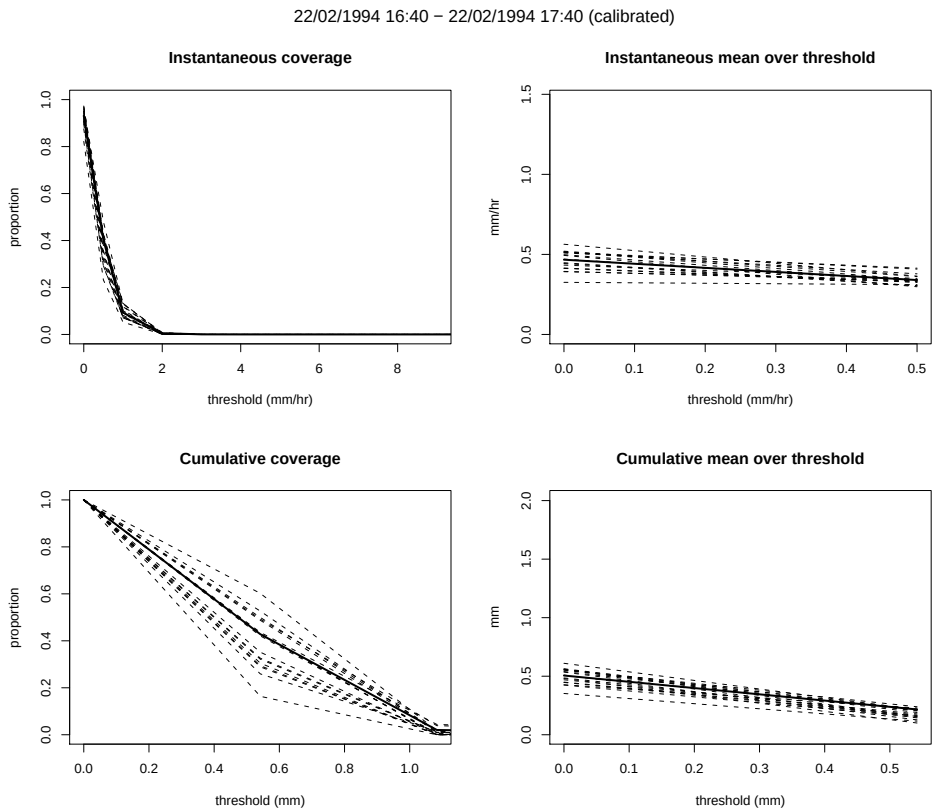


Figure 10.15 Coverage and mean exceedance over threshold, instantaneous and cumulative, for the event of 28/02/2000.

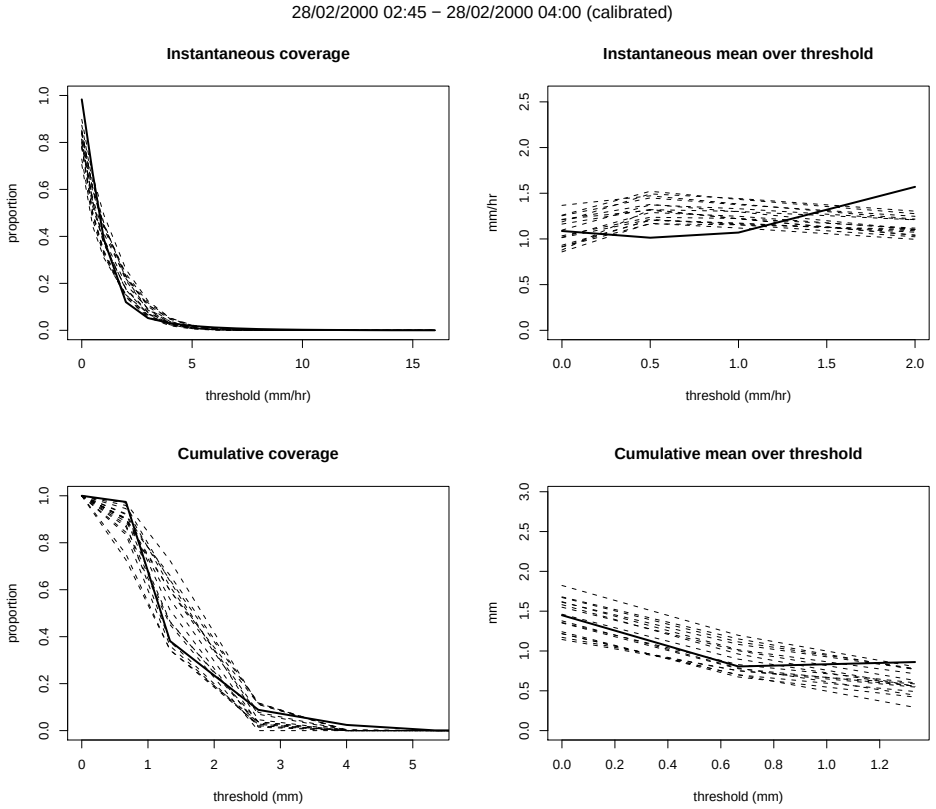


Figure 10.16 Coverage and mean exceedance over threshold, instantaneous and cumulative, for the event of 02/04/2000.

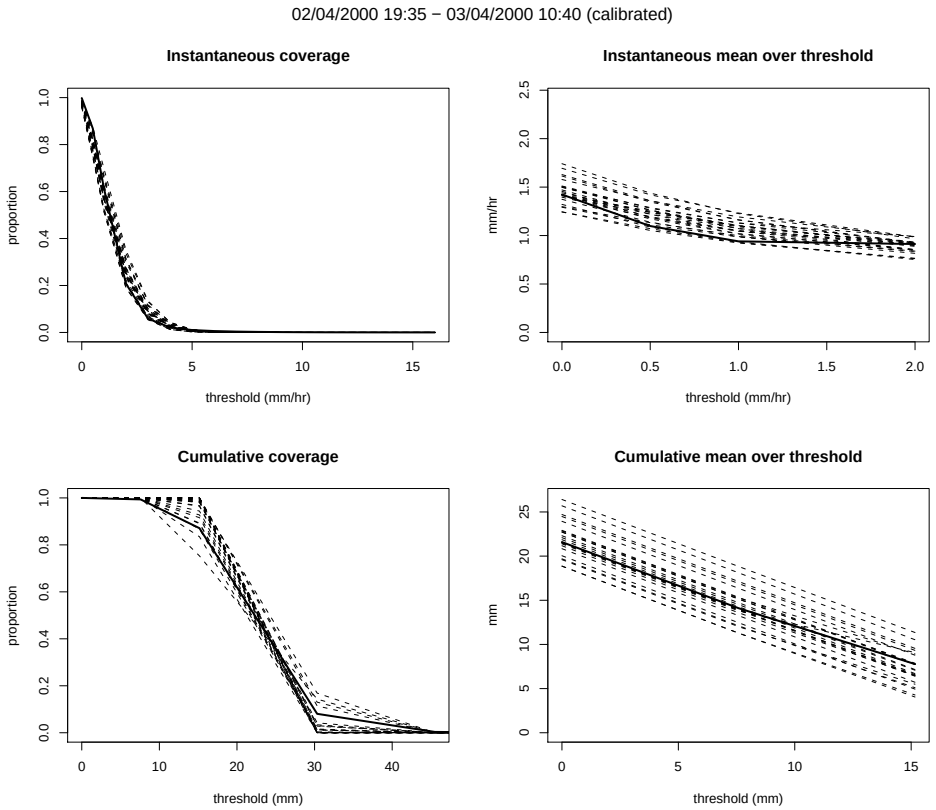


Figure 10.17 Coverage and mean exceedance over threshold, instantaneous and cumulative, for the event of 04/12/2000.

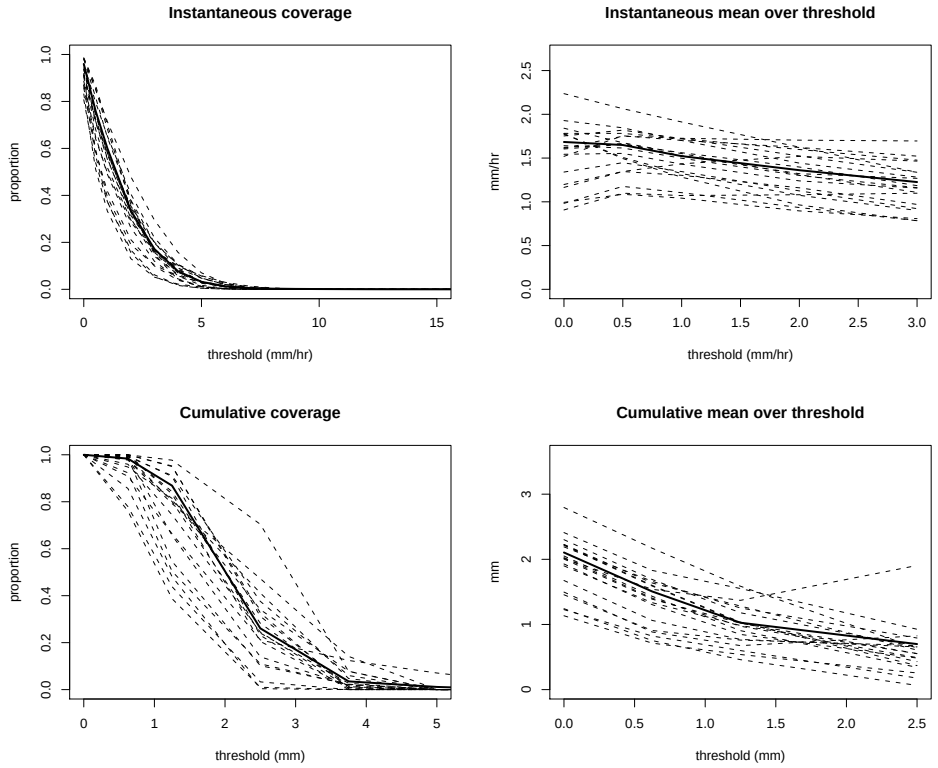


Figure 10.18 Coverage and mean exceedance over threshold, instantaneous and cumulative, for the event of 16/03/2001.

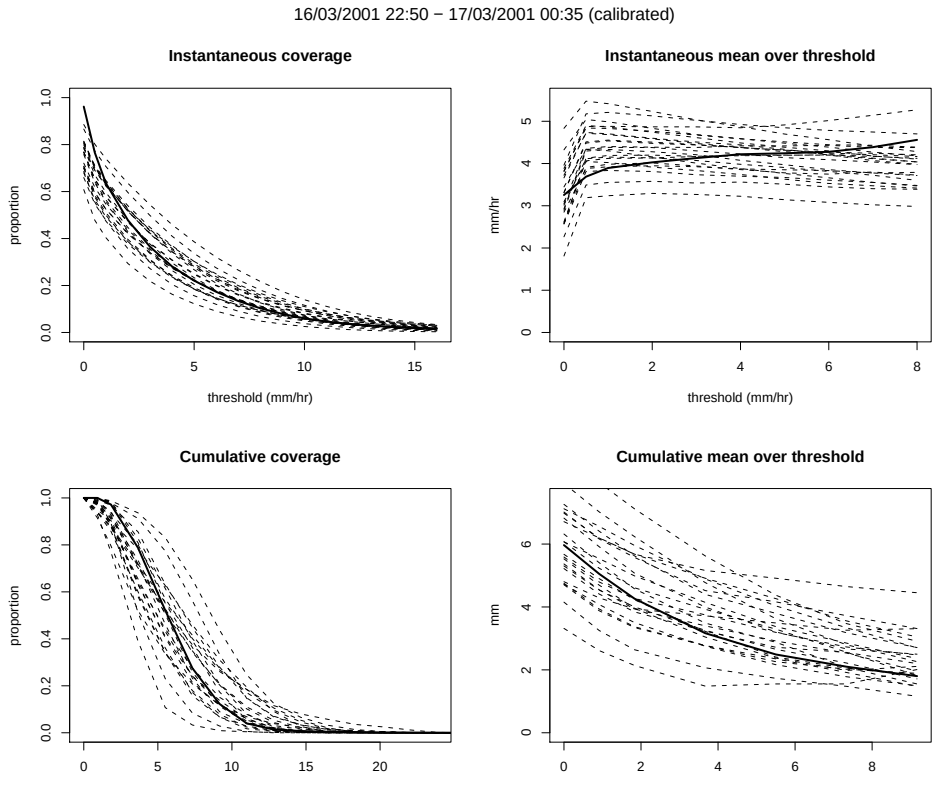
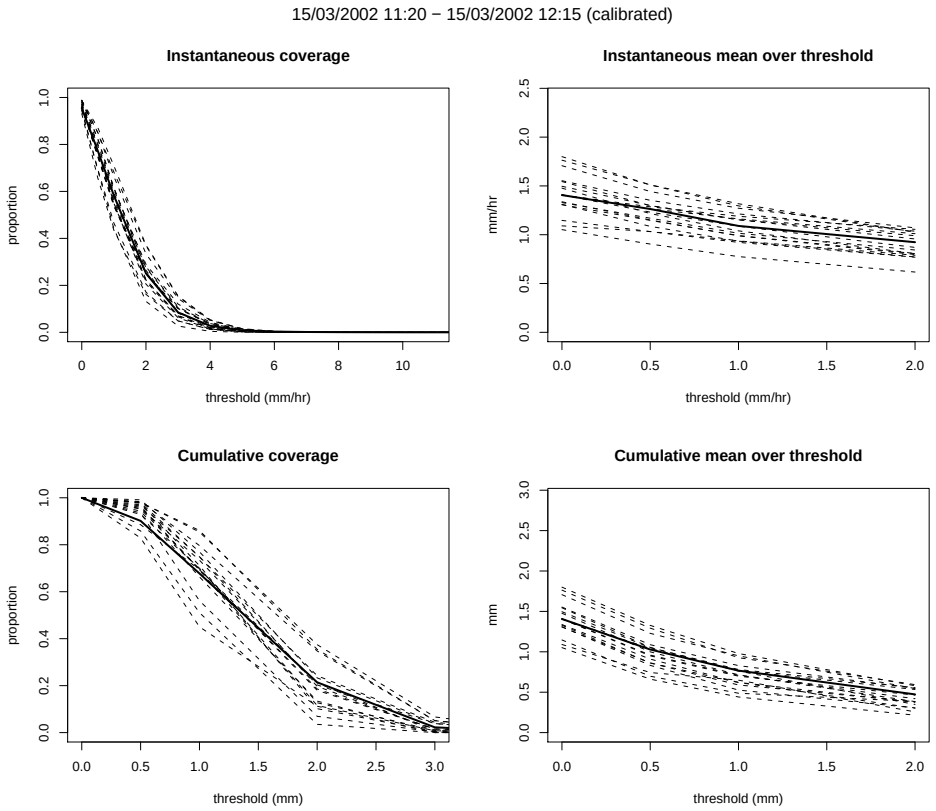


Figure 10.19 Coverage and mean exceedance over threshold, instantaneous and cumulative, for the event of 15/03/2002.



simulated traces for this event are all fairly flat after this jump (as noted above, this is consistent with an exponential distribution), whereas the observed trace rises steadily (indicating a distribution with heavier tails than the exponential).

For the event of 28th February 2000, both observations and simulations show curvature but in opposite directions. In absolute terms the disagreement is not substantial, but the plot does indicate that the model fails to reproduce the highest rainfall intensities for this event.

Cumulative coverage: the agreement between observations and simulations is excellent for all events.

Cumulative mean over threshold: the results are broadly comparable with those for the instantaneous mean over threshold, with good performance for all events except 19th November 1991 and 28th February 2000. A general feature is that the simulations show more variability than for the instantaneous plots.

The analyses in this section represent quite a challenge for the event interior model: although its construction is relatively simple, it appears sufficiently adaptable to reproduce properties relating to the spatial organisation of a variety of contrasting events. However, a matter of some concern is that the worst performance seems to be associated with the intense events that would typically be of most interest in a flood risk assessment. Recall that the fitting failed for two of the three most intense events in

Table 10.3, so that they have not been considered in Figures 10.6–10.19. In these figures, the worst results are for the event of 19th November 1991: this completes the trio. For the three most intense events in the original list therefore, the model performance is inadequate.

10.3 Durations of wet and dry periods

As described in Chapter 9, for continuous simulation purposes it is necessary to specifiy a model for the durations of rain events and the intervening dry periods; continuous simulation then proceeds by generating a sequence of wet-dry durations, and generating a realisation of the event interior model within each 'wet' period in such a way that the simulated coverage remains above the 15% threshold for the correct period of time (Section 9.4).

Our duration model is an alternating renewal process (Cox and Isham, 1980): successive durations are mutually independent and, in each month of the year, separate Weibull distributions are used to model the durations of wet and dry periods. Following Wheater et al. (2000b), the Weibull family has been chosen because it is flexible, and because it enables adjustments to be made for missing data (see below).

10.3.1 Validity of the alternating renewal assumption

To check the assumption of independence between successive events, plots of successive event durations have been produced for each month of the year. The plots produced are as follows:

- Duration of current wet period against duration of next wet period.
- Duration of current wet period against duration of next dry period.
- Duration of current dry period against duration of next wet period.
- Duration of current dry period against duration of next dry period.

Some examples of these plots are shown in Figures 10.20 and 10.21. These are typical: in general there are small positive correlations between successive durations, but these generally do not differ significantly from zero. Moreover, even those that are statistically significant (such as the "dry then wet" correlation in July — see Figure 10.21) are sufficiently small to be of little practical consequence. The highest correlation observed is 0.2, between the lengths of successive dry periods in December. On this basis, it seems that little will be lost by assuming that successive interval lengths are independent.

10.3.2 The problem of missing data

As noted in Appendix F, there are frequent breaks in the Chenies radar record; this means that some event durations cannot be observed fully. Figure 10.22 illustrates the

Figure 10.20 Durations (in hours) of successive wet and dry period durations in January. Logarithmic axis scaling is used on all plots.

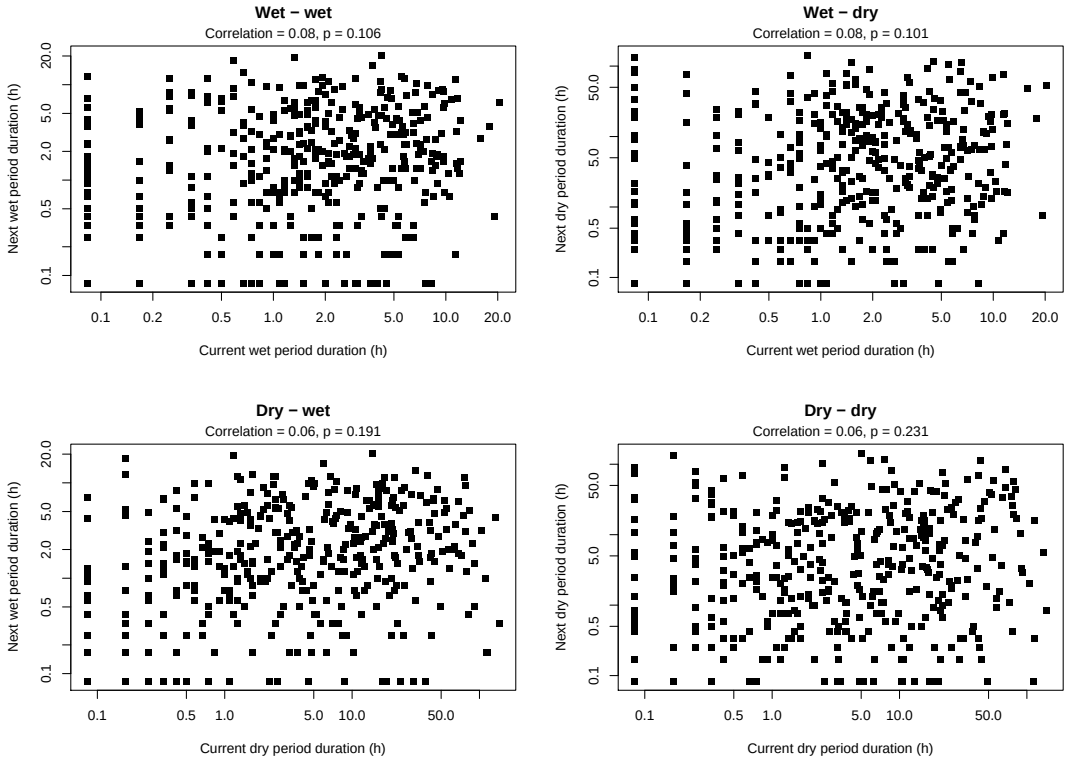


Figure 10.21 Durations (in hours) of successive wet and dry period durations in July. Logarithmic axis scaling is used on all plots.

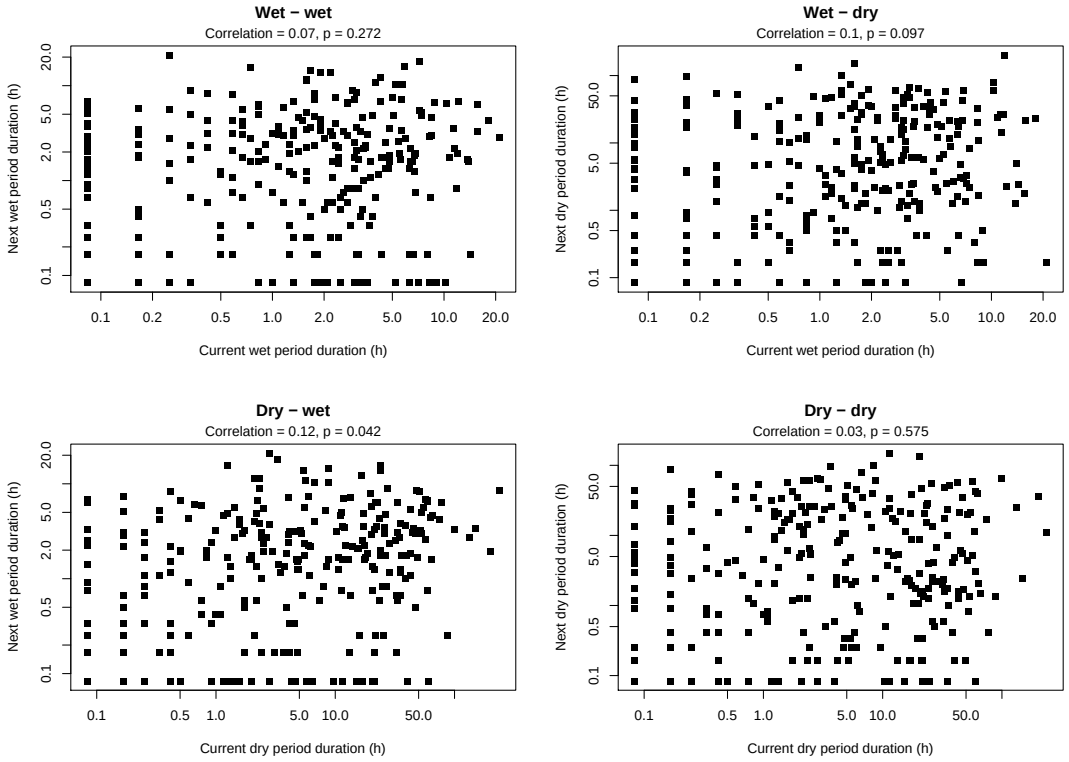
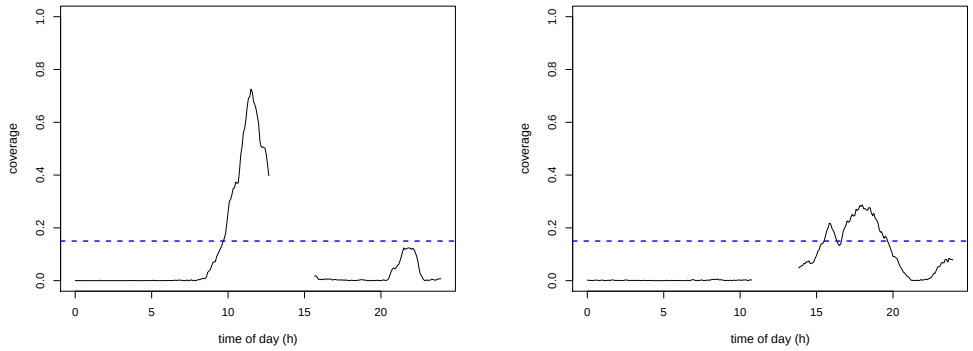


Figure 10.22 Censored coverage series; (left): the system is in different states before and after the missing period; (right): the radar window is dry before and after the missing period.



problem. In the left hand plot, a block of missing data starts midway through an event (when the coverage exceeds the 15% threshold), and ends during a dry period. In this case, it is not possible to determine the duration of either the event itself or the subsequent dry period. In this particular example, since only a couple of hours of data are missing it would be plausible to extrapolate across the missing data and estimate the end time of the event. However, this is not always the case: gaps of several hours are not uncommon, for example. Strictly speaking, all that is known about the duration of the event in this plot is that it lies between some interval — the start time is known, but the end time is somewhere in the period of missing data. An interval whose exact duration is unknown is said to be *censored*.

A scenario such as that in the right hand plot of Figure 10.22 presents more of a challenge. Here, the coverage is below the 15% threshold on both sides of the missing data period. In this case it is not known whether the two segments of the plot correspond to the same dry period, or whether the coverage series has crossed the threshold in the missing patch.

The presence of missing data complicates the fitting of distributions to the wet and dry period durations. Some simplistic options for dealing with the problem are as follows:

- For all censored intervals, take the interval length equal to the duration of the partial interval that is observed (this was the approach taken to produce Figures 10.20 and 10.21). For example, in the left hand panel of Figure 10.22, the event duration would be taken as the time from the start of the event to the last observation before the missing data period. Clearly however, such an approach will lead to systematic underestimation of interval lengths.

- Discard all censored intervals, and fit distributions only to the remainder. It may seem that this will overcome the problem of bias associated with the previous approach; however, this is not the case. The reason is that long intervals are more likely to suffer censoring than short intervals; hence long intervals will be underrepresented in the remaining sample if censored intervals are discarded.

This problem is known as *length-based sampling* (Cox and Isham, 1980).

An alternative is to use methods from survival analysis, which are specifically designed to cope with censored data. One common technique is to fit a distribution using a likelihood function that explicitly accounts for the censored observations. Suppose, for example, that wet period durations are assumed to be drawn from a distribution with probability density function $f(\cdot;\theta)$, where θ is a vector of unknown parameters. A sequence of durations $T_1,\ldots,T_n$ is partially observed: for a subset $\mathcal{U}\subseteq\{1,\ldots,n\}$, the values $\{T_i : i\in\mathcal{U}\}$ are observed exactly, but the remainder $\{T_i : i\in\mathcal{C}\}$ are censored in such a way that for $i\in\mathcal{C}$, all that is known is that $T_i\geq\tau_i$ for some threshold τ_i (here, $\mathcal{U}$ and $\mathcal{C}$ denote the sets of uncensored and censored observations, respectively). Such observations are said to be *right censored*. In this case, it is a standard result (Davison, 2003, Section 5.4.2) that the log likelihood function for θ given the available data is

$$\log L(\theta)=\sum_{i\in\mathcal{U}}\log f(T_i;\theta)+\sum_{i\in\mathcal{C}}\log S(\tau_i;\theta)\ ,$$

where

$$S(t;\theta)=\mathrm{P}(T_i\geq t;\theta)=\int_t^\infty f(u;\theta)\,du \tag{10.1}$$

is the *survivor function* of the distribution. This enables θ to be estimated using maximum likelihood.

To implement this idea, it is convenient to work with a flexible family of distributions for which the survivor function has a simple analytic form. The Weibull family satisfies these requirements, and is therefore commonly used in survival analysis. The Weibull density is

$$f(t)=\frac{c}{\alpha}\left(\frac{t}{\alpha}\right)^{c-1}\exp\left[-\left(\frac{t}{\alpha}\right)^c\right]\qquad(t>0) \tag{10.2}$$

for parameters $\alpha>0, c>0$; the corresponding survivor function is

$$S(t)=\exp\left[-\left(\frac{t}{\alpha}\right)^c\right]. \tag{10.3}$$

The mean and variance of the distribution are

$$\alpha\Gamma\left(1+c^{-1}\right)\qquad\text{and}\qquad\alpha^2\left[\Gamma\left(1+2c^{-1}\right)-\left(\Gamma\left(1+c^{-1}\right)\right)^2\right]$$

respectively, where $\Gamma(\cdot)$ denotes the gamma function. If $c=1$, the Weibull distribution reduces to the exponential.

10.3.3 Fitting results

Weibull distributions have been fitted, using standard survival analysis software, to the wet and dry period durations for each month of the year. Any period overlapping a month end was counted as belonging to the first month affected. In the fitting, any censored duration was considered to be right censored — for example, the event in the left hand plot of Figure 10.22 was regarded as finishing at some time after the start of the missing period, but no attempt was made to incorporate the information that it

finished before around 15:00. In general, this is likely to lead to slight upward biases in all fitted distributions; there is scope for improvement here in the future. A further point is that in situations such as that in the right panel of Figure 10.22, it is not known whether the coverage crossed the threshold during the missing period. In our analyses, we have discarded the second partial period in such situations to avoid the possibility of counting the same interval twice. Again, however, there is room for future improvement.

The fitted distributions are summarised in Tables 10.5 and 10.6, along with information on the numbers of wet and dry periods in each month for the 12 years used in the analysis. To assist in the interpretation of the fitted distributions, their means and standard deviations are also given. Note that the observed means and standard deviations cannot be calculated for comparison, because of the censoring. The tables show that the proportion of censored intervals can be high, especially for dry periods in the summer months. In general it is expected that dry periods will be more heavily censored than wet periods, simply because they tend to be longer. The means of the fitted wet period duration distributions seem relatively stable, at around 3–4 hours, throughout the year. There is, however, a strong seasonal cycle in the fitted dry period distributions, with mean dry period durations exceeding 24 hours in the summer months. It is interesting to note that the estimated shape parameters (c) show very little variation throughout the year, for either wet or dry periods. From the expressions for the mean and variance given above, it can be seen that a constant value of c implies a common coefficient of variation for the duration distributions in all months.

In Appendix F, it was noted that the Chenies radar underwent a number of changes during the 1990s, and that these may have affected the characteristics of the events recorded there. In particular, a modified transmitter was installed around 1996, which led to improved detection of light rain and contributed in part to an increase in the number of recorded events at that time (Section F.4.3). To see if this affected the properties of the duration distributions, a separate analysis was performed in which the Weibull scale parameter, α, was allowed to change in 1996. The analysis was carried out using the `survival` library in R (R Development Core Team, 2003). Likelihood ratio tests indicated that there was no significant change in the scale parameter in any month of the year, for either wet or dry period durations.

10.3.4 Assessing the quality of the fit

To assess whether a particular distribution provides a good fit to data, a standard technique is to use quantile-quantile plots of the fitted and empirical distribution functions. However, it is not straightforward to produce such plots here because the exact values of the censored observations are unknown. To assess the fit of a Weibull distribution to censored data, a useful diagnostic is based on the *cumulative hazard function*

$$H(t) = -\log S(t) \ , \tag{10.4}$$

where $S(t)$ is the survivor function defined at (10.1). In the presence of censoring, $H(t)$ can be estimated empirically using the Nelson-Aalen estimator (Venables and

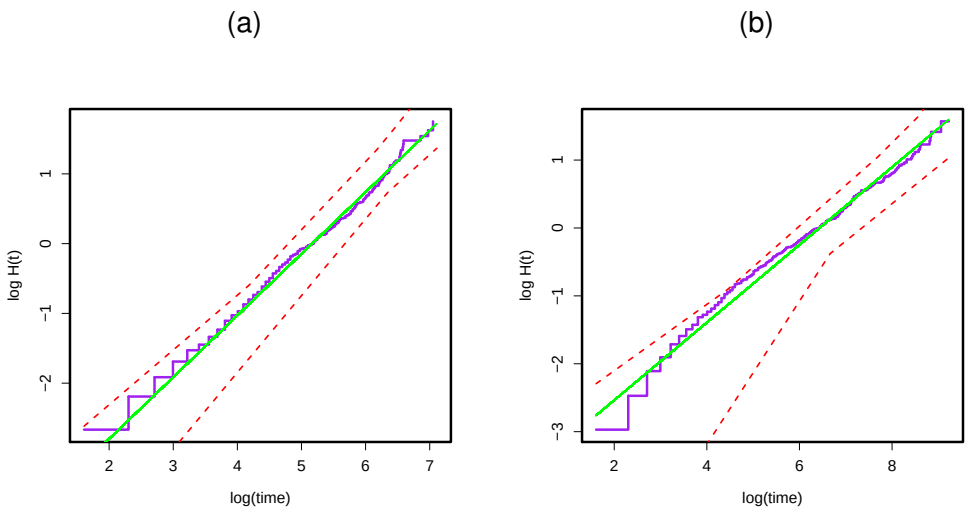
Table 10.5 Parameters and properties of Weibull distributions fitted to wet period durations (in minutes)

Month	No. of wet periods	Percentage censored	$\hat{\alpha}$	$\hat{c}$	Fitted mean	Fitted std dev
1	417	6.2	175.0	0.88	186.0	210.9
2	427	7.0	172.8	0.79	196.8	250.0
3	379	3.4	177.2	0.80	201.0	263.8
4	345	7.5	240.5	0.78	278.4	362.0
5	372	6.4	181.6	0.76	213.5	283.9
6	363	6.6	164.3	0.77	191.7	252.2
7	286	8.7	173.1	0.82	192.4	235.5
8	274	8.8	184.4	0.84	202.3	242.4
9	324	7.4	204.5	0.83	226.5	275.7
10	371	4.9	184.6	0.80	208.7	262.2
11	420	9.3	178.0	0.77	206.8	270.5
12	410	3.4	198.8	0.91	207.9	228.6

Table 10.6 Parameters and properties of Weibull distributions fitted to dry period durations (in minutes)

Month	No. of dry periods	Proportion censored	$\hat{\alpha}$	$\hat{c}$	Fitted mean	Fitted std dev
1	420	13.6	622.4	0.57	999.6	1866.6
2	432	14.4	570.6	0.60	849.7	1479.1
3	383	14.1	605.4	0.54	1072.7	2179.6
4	348	16.7	662.6	0.58	1043.6	1914.8
5	331	18.4	729.3	0.54	1296.9	2643.2
6	365	17.3	701.4	0.54	1214.8	2422.0
7	290	23.8	899.7	0.54	1577.4	3177.4
8	277	22.7	880.6	0.51	1679.1	3619.0
9	328	19.8	779.4	0.53	1425.6	2972.8
10	376	17.0	707.9	0.58	1112.2	2036.2
11	423	17.0	620.4	0.58	979.1	1799.6
12	412	10.7	699.0	0.60	1044.9	1825.6

Figure 10.23 Assessment of Weibull fit: log-log plots of empirical (step) and theoretical (straight line) cumulative hazard functions for wet and dry period durations in January. Dashed lines indicate the envelopes from 10 simulations.



Ripley, 2002, Section 13.1):

$$\hat{H}(t) = \sum_{i:\ i \in \mathcal{U},\ T_i \le t} \frac{1}{N_i}\ , \tag{10.5}$$

where N_i is the number of durations known to be greater than t (this could include both censored and uncensored durations). The sum in (10.5) is over all uncensored observations with durations up to and including t. $\hat{H}(t)$ is a step function with jumps at each of the observed (uncensored) durations.

For the Weibull distribution, from (10.3) the cumulative hazard is

$$H(t) = \left(\frac{t}{\alpha}\right)^c\ ,$$

so that there is a linear relationship between $\log H(t)$ and $\log t$:

$$\log H(t) = c \log t - c \log \alpha\ .$$

This suggests that if the Weibull assumption is correct, a plot of $\log \hat{H}(t)$ against $\log t$ should be approximately linear. Examples, for January data from Chenies, are shown in Figure 10.23. In addition to showing the empirical and fitted hazards, the dashed lines in these plots provide an envelope based on ten simulations to give an idea of the variability expected under the model. Each simulation was subject to the same censoring pattern observed in the actual data: a sequence of wet and dry period durations was generated for the period 1990–2002 according to the fitted alternating renewal model, and the same missing periods were imposed on this sequence as occurred in the radar data. The hazard function was then estimated from the censored simulation using the methodology described above.

The results obtained for all months were similar to those in Figure 10.23. In general, model performance is satisfactory for wet period durations; in some months, however,

there is a tendency to overestimate the lengths of dry periods. This could be due to the fact that all censored intervals have been treated as right-censored in the analyses, as discussed abovce.

10.4 Continuous simulation results for the stationary space-time model

In previous sections, we have discussed the performance of the individual components of the stationary continuous simulation model: the stochastic event interiors model (Section 10.2) and the alternating renewal process for the wet and dry durations advection (Section 10.3). In this section, we now assess the performance of the combined simulation model fitted using the Chenies data (calibrated as described in Section F.2). In particular, we compare basic summary statistics for simulated data sequences with those of observed data after calibration, to see whether the statistical properties of the two sequences are similar. As for the event interiors model discussed above, our focus is on the first and second order properties of the sequences (that is, the mean, standard deviation, spatial and temporal autocorrelation functions) and the proportion of wet pixels.

As in Wheater et al. (2000b), we assess the performance of the spatial-temporal model at various spatial and temporal scales, focussing mainly on the hourly time scale and the 64km^2 spatial scale (ie on blocks of 4×4 pixels). The spatial scales that have been considered are 4km^2, 16km^2, 64km^2, 256km^2, while the temporal scales are 15 minutes, 1 hour and 24 hours. Each time scale has been combined with the basic 64km^2 spatial scale, and each spatial scale has been considered at the hourly time scale. Thus the actual combinations of spatial and temporal scales that have been considered are:

- Hourly & 4km^2
- Hourly & 16km^2
- Hourly & 64km^2
- Hourly & 256km^2
- 15 mins & 64km^2
- Daily & 64km^2.

In the figures that are presented below, we shall present only results for three space-time combinations as exemplars. These are

- Hourly & 4km^2
- Hourly & 64km^2
- Daily & 64km^2.

For the continuous simulations, the realisations of the rainfall intensity are stored in the same format as the observed data, that is as 5 minute images consisting of 52×52 4km^2 square pixels. Measurements at larger spatial and/or temporal scales are simply obtained by averaging over the relevant number of pixels and/or images, and are then available for calculating the summary statistics. For the remainder of this section, when we refer to pixels and images we are referring to those which have already been aggregated to the appropriate space and time resolution.

The spatial-temporal continuous simulation rainfall model is intended to represent the main rain events moving across the catchment/simulation window. No attempt has yet been made explicitly to develop the model to include periods of very light spatially- and/or temporally-intermittent rainfall. Thus, before taking averages of the original 5 minute, 4km^2, images for the empirical data, all of these images having a coverage of less than 15% have been made completely dry (i.e. all pixels within the image have been set to a rainfall intensity of zero). The threshold of 15% has been chosen to correspond to the threshold used to define fitting events (see F.4.2). In addition, this means that any residual effects of rainfall anomalies or clutter remaining after data calibration should not influence the comparisons. This procedure was also carried out on the original 5 minute, 4km^2, simulated images to allow for a more informed comparison. We note that a similar pre-comparison thresholding of data and simulated images was carried out in the comparisons reported by Wheater et al. (2000b) for the Wardon Hill radar, although in that case a much more stringent 25% threshold was used (again in line with the threshold used in that case for model fitting).

Five realisations of the stochastic model have been simulated, each for a period of 12 years. Radar data from Chenies were available for 13 years (for 1990 to 2002) but as explained earlier (Section F.2.3), data from one month in 1992 had been precalibrated on site by the Met. Office and we were unable to remove the effect of this precalibration in order to apply our recommended calibration method. We have therefore omitted all the 1992 data from the model fitting and comparisons between the empirical data and simulated realisations so that all months have the same number (12) of years of data. Summary statistics are then calculated for the empirical data and for each of the five realizations, on a month by month basis. The average of each summary statistic over the five realizations is also obtained. The summary statistics fall into two main categories: unconditional statistics and conditional statistics. The unconditional statistics are calculated over all pixels, whereas the conditional statistics are only calculated over pixels which are ‘wet’ (that is, pixels where the mean rainfall intensity is nonzero). While the unconditional statistics provide us with an overall picture, the conditional statistics allow us to focus specifically on rain events, and thus provide complementary information. The various statistics calculated for each particular month are as follows:

Unconditional Statistics:

(1) Overall Mean — the mean rainfall intensity, in millimetres per hour, taken over all pixels through both space and time.

(2) Overall Standard Deviation — the standard deviation of rainfall intensity taken over all pixels through both space and time.

(3) Within Image Standard Deviation — for each image, the spatial variance of rainfall intensity is calculated over all pixels; the square root of the average

variance over all images then gives an average within-image standard deviation.

(4) Overall Proportion Wet — the proportion of pixels through both space and time that are wet.

(5) Spatial Autocorrelation Function (Spatial ACF) — calculated for 'nearest neighbour' spatial lags (1,0), (0,1), (1,1) and (-1,1) defined as follows. Each aggregated image is made up of a square grid of square pixels so that any given pixel p within the image will generally have 8 nearest neighbour pixels, as shown in the diagram below

(-1,1)	(0,1)	(1,1)
(-1,0)	p	(1,0)
(-1,-1)	(0,-1)	(1,-1)

using notation based on Cartesian co-ordinates to label the neighbouring pixels. It follows that there are four distinct spatial ACF lags to calculate if we just consider neighbouring pixels. To calculate the spatial ACF at (for example) lag (1,1) of a particular image, we use all pairs of pixels within the image having the same relative orientation as pixels p and (1,1), and calculate the spatial ACF for each image that contains some wet pixels. The average of these values is taken to give the overall value of the autocorrelation.

(6) Temporal Autocorrelation Function (Temporal ACF) — calculated at time lags 1, 2 and 3. The temporal ACF at a particular lag is calculated at each particular pixel location in the spatial image. The average of these values is taken to give the overall value of the autocorrelation.

Conditional Statistics:

(1) Conditional Overall Mean — the mean rainfall intensity, in millimetres per hour, taken over all wet pixels through both space and time.

(2) Conditional Overall Standard Deviation — the standard deviation of rainfall intensity taken over all wet pixels through both space and time.

(3) Conditional Within Image Standard Deviation — for each image containing two or more wet pixels, the spatial variance of rainfall intensity is calculated over all wet pixels; the square root of the average variance over all images then gives an average within-image standard deviation.

(4) Proportion Wet Conditional On Wet Images — the proportion of pixels that are wet averaged over all images that contain at least one wet pixel. (Note that this statistic is slightly different from the other conditional statistics, for which only wet pixels are involved in the calculations.)

(5) Conditional Spatial Autocorrelation Function — defined exactly as its unconditional counterpart except that, in calculating the spatial ACF for each image, pairs of pixel values are included in the calculation only if both are non-zero.

(6) Conditional Temporal Autocorrelation Function — defined exactly as its unconditional counterpart except that, in calculating the temporal ACF at a particular pixel location, pairs of values are included in the calculation only if both are non-zero.

The calculations of these statistics are exactly as reported in Wheater et al. (2000b). In particular, when calculating the spatial and temporal autocorrelation functions, we have followed the procedure adopted in that report in that we have used the number of pairwise pixel comparisons as the divisor in the calculations of the autocovariance functions (rather than the total number of pixels in the spatial image or temporal sequence). This is especially important for the empirical data where there are periods of missing images, and enables a more informed comparison between the ACFs at various lags. Discussion regarding the choice of estimator for an autocorrelation function may be found in Priestley (1981, section 5.3.3) and in Kendall and Ord (1990, chapter 6).

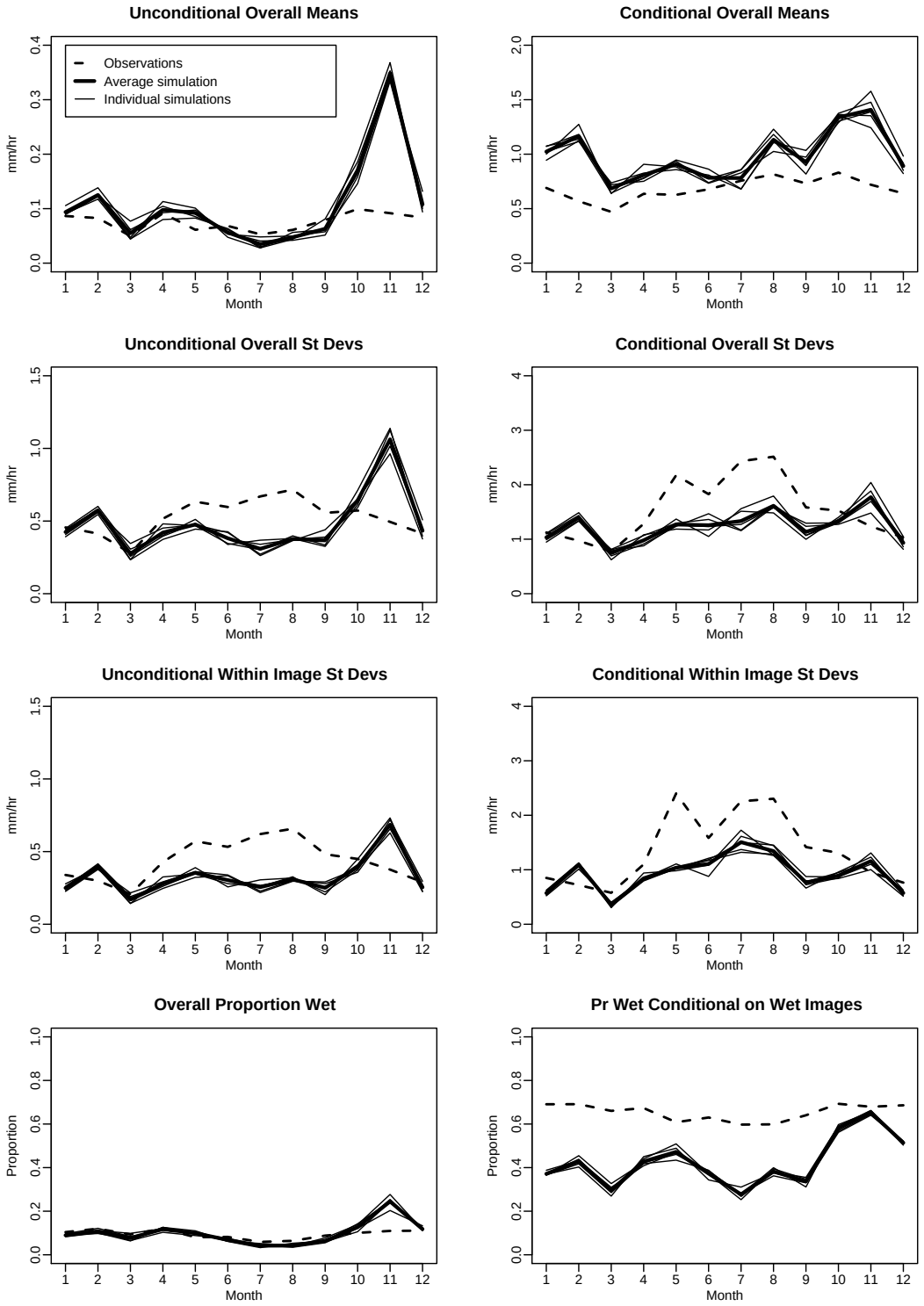
Plots of the summary statistics, as described above, for the five simulations, their average and for the empirical data, are shown in the following figures for the three chosen spatial-temporal scales (hourly & 4km^2, hourly & 64km^2, daily & 64km^2 respectively). Figures 10.24, 10.27 and 10.30 show the mean, standard deviations (overall and within-image) and proportion wet, giving both unconditional and conditional properties. Figures 10.25, 10.28 and 10.31 show the corresponding conditional and unconditional spatial autocorrelation functions for the four nearest neighbour spatial lags specified above, while Figures 10.26, 10.29 and 10.32 show the conditional and unconditional temporal autocorrelation functions for time lags 1,2 and 3.

We consider first the performance at the hourly level, at the finest spatial resolution of 4km^2, remembering that all the original 5 minute, 4km^2 images for both empirical data and simulated realisations have been subject to a threshold of 15% coverage, with pixel intensities within images having coverages below this threshold set to zero.

From Fig. 10.24 we see that, other than in November, the unconditional means of the individual simulated realisations reflect the values of the empirical data. A possible explanation for the overestimation of mean rainfall in November is that, to select a set of parameters for a particular simulated event, the procedure is to choose that set having an event duration that is closest to the duration generated by sampling from the Weibull distribution fitted to event durations in that particular month. Given the relatively small number of parameter sets in the library for each month, it is possible that a particularly heavy event has been chosen unrepresentatively often.

The overall probability of a pixel being wet is very well reproduced by the model (except, again, for November). However, if we focus only on images where there is rain in some part of the image (i.e. there would have been a coverage of at least 15% on the pre-thresholded image), then we see that the model tends to produce too few wet pixels. In compensation, for the mean rainfall calculated conditionally over the wet

Figure 10.24 Summary statistics for simulated and empirical data (hourly & 4 km^2): means, standard deviations and proportion wet.



pixels (for the thresholded images), we see that the fitted model is tending to overestimate this rainfall relative to the empirical data.

The unconditional overall standard deviations tend to be too small in the summer months, possibly reflecting the smaller numbers of fitted parameter sets available in the parameter library in those months for selection in the continuous simulation. Both overall standard deviations and those calculated as average within-image values are very similar. Again, the standard deviations calculated only over wet pixels, show evidence of more variability in the data than in the values simulated from the model in the summer months, with the overall values and those calculated as average within-image values showing similar patterns.

We note that the model of event interiors matches the within-event mean rainfalls exactly (Section 10.2), and reproduces the variances well at a range of spatial scales, so that the discrepancies observed here must be artefacts of the event sampling procedures. One possible cause, as mentioned above, is restricted number of rain events in the radar record that are suitable for fitting the within-event model, and hence the relatively small number of fitted parameter sets in the library that are available for the continuous simulation. A further, possible explanation for the discrepancies noted above is that the current procedure for continuous simulation omits the possibility that major events, similar to those in the fitting set, do not pass centrally over the fitting window but just overlap, say, one corner of the window for a short while. The inclusion of such events would provide greater variability within images as well as affecting the mean values.

The simulated spatial autocorrelations are shown in Figure 10.25. It can be seen that there is generally good agreement with the autocorrelations for the observed data at all four spatial lags, and for both conditional and unconditional autocorrelations. The spatial ACFs for the simulated data tend to show slightly more seasonal structure than is present in the ACFs for the observed data. It is worth noting that for both the average simulation and the empirical data, the spatial ACF is slightly greater at lags (1,0) and (0,1) than it is at lags (1,1) and (-1,1). This is to be expected given that in the spatial ACF at lags (1,0) and (0,1) we are comparing pixels that are closer together than at the other two lags.

The temporal autocorrelations are shown in Figure 10.26. It can be seen that the autocorrelations for the simulations generally reflect well the properties of the empirical data at all lags, and for both unconditional and conditional ACFs. In particular, the seasonal patterns in the ACFs from the simulated data are similar to those observed. The conditional ACFs show slightly closer agreement between observed and simulated functions than the unconditonal ones.

Comparison of figures 10.27–10.29 with figures 10.24–10.26 shows the effect of increasing the spatial scale to 64km^2, with the temporal scale unchanged. The comparisons of means, standard deviations and proportions wet between the empirical and simulated images, are all very similar at the two spatial scales. The standard deviations are generally slightly smaller at the larger spatial scale and the discrepancies between observed and simulated values also tend to be a little smaller. At the larger scale, the probability that a pixel is wet is necessarily higher for both empirical data and model simulations. This is true both for the overall values and for those calculated only for wet images, although the effect in the latter case is much

Figure 10.25 Summary statistics for simulated and empirical data (hourly & 4 km^2): spatial autocorrelations.

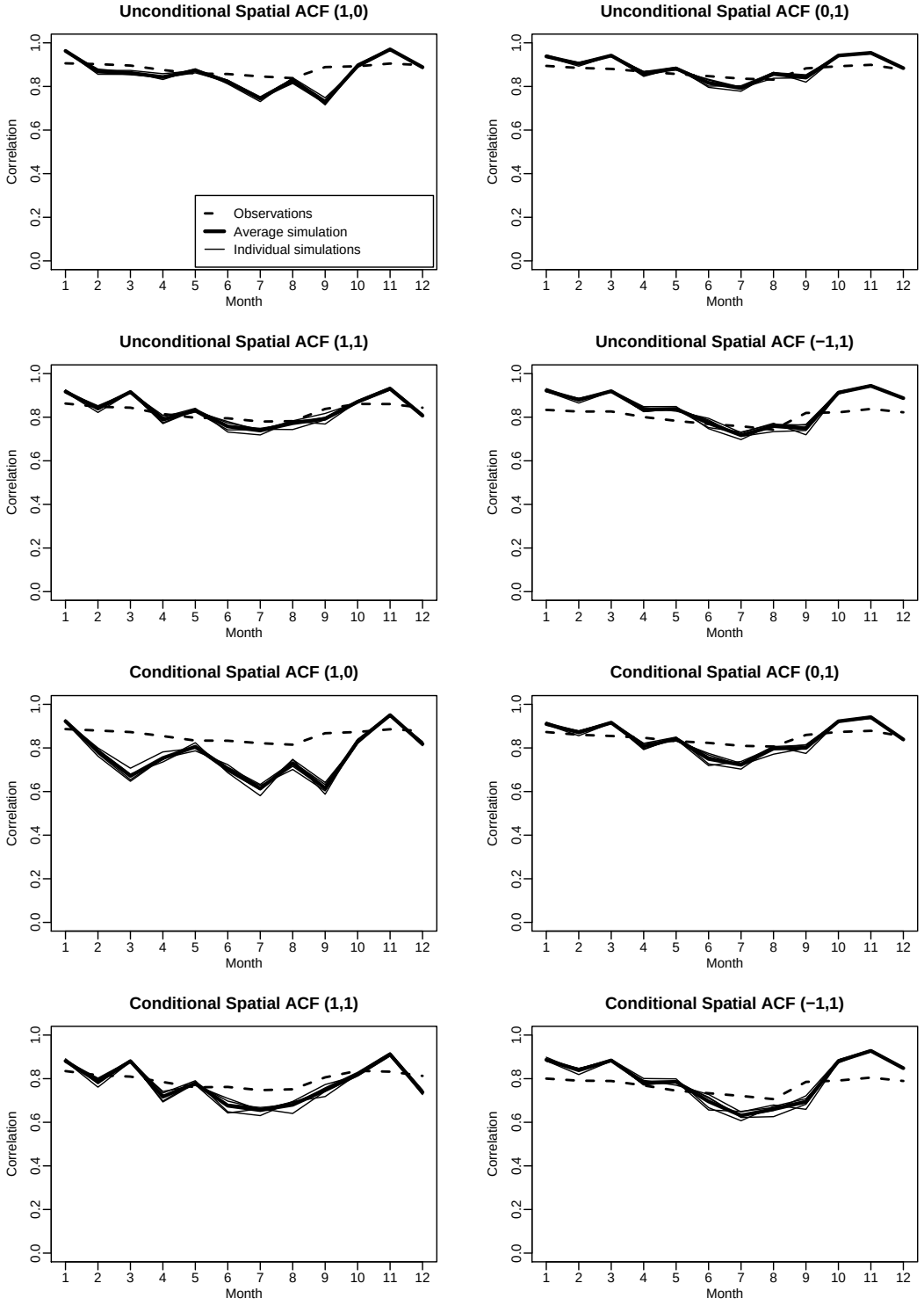


Figure 10.26 Summary statistics for simulated and empirical data (hourly & 4 km^2): temporal autocorrelations.

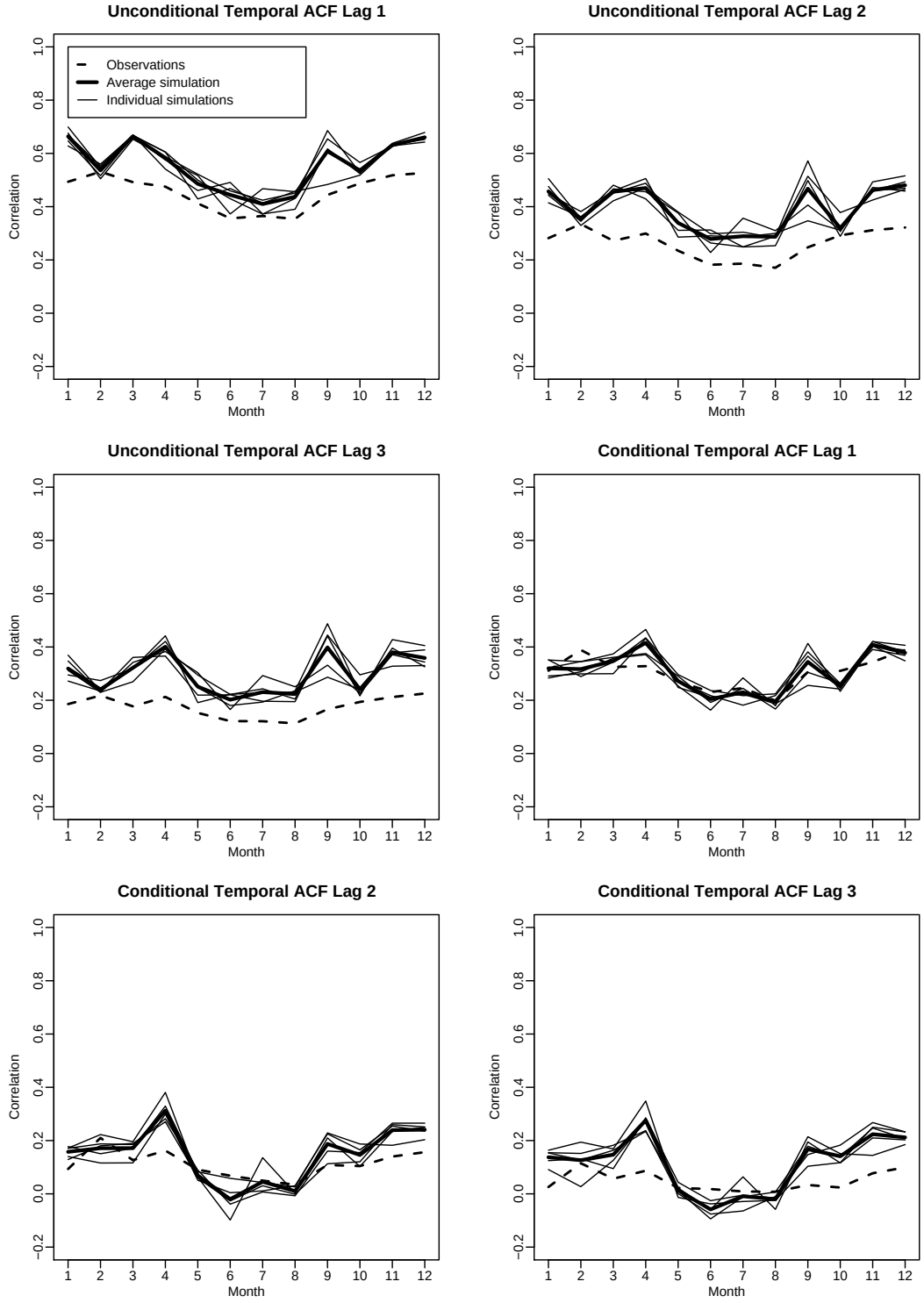


Figure 10.27 Summary statistics for simulated and empirical data (hourly & 64 km^2): means, standard deviations and proportion wet.

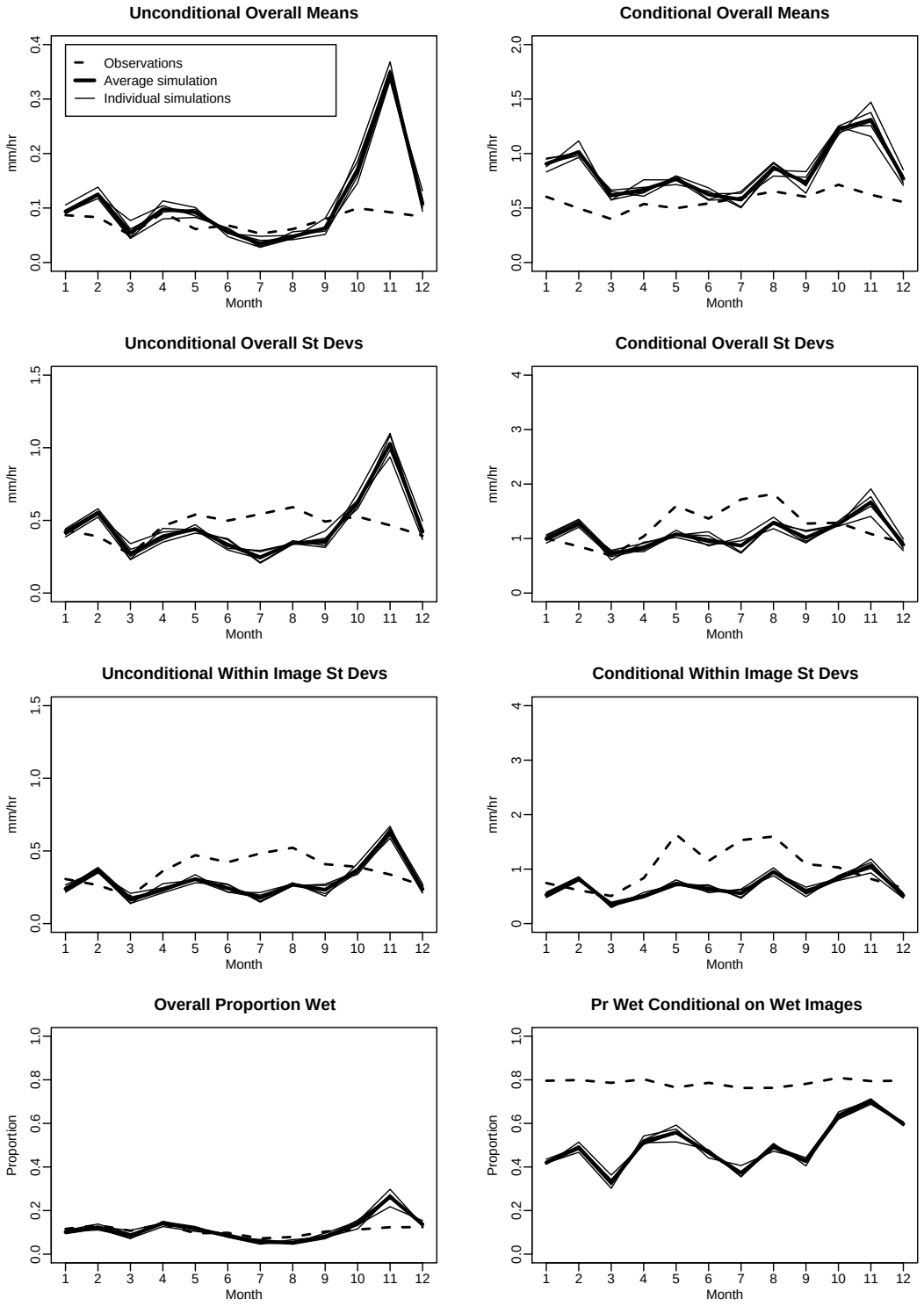


Figure 10.28 Summary statistics for simulated and empirical data (hourly & 64 km^2): spatial autocorrelations.

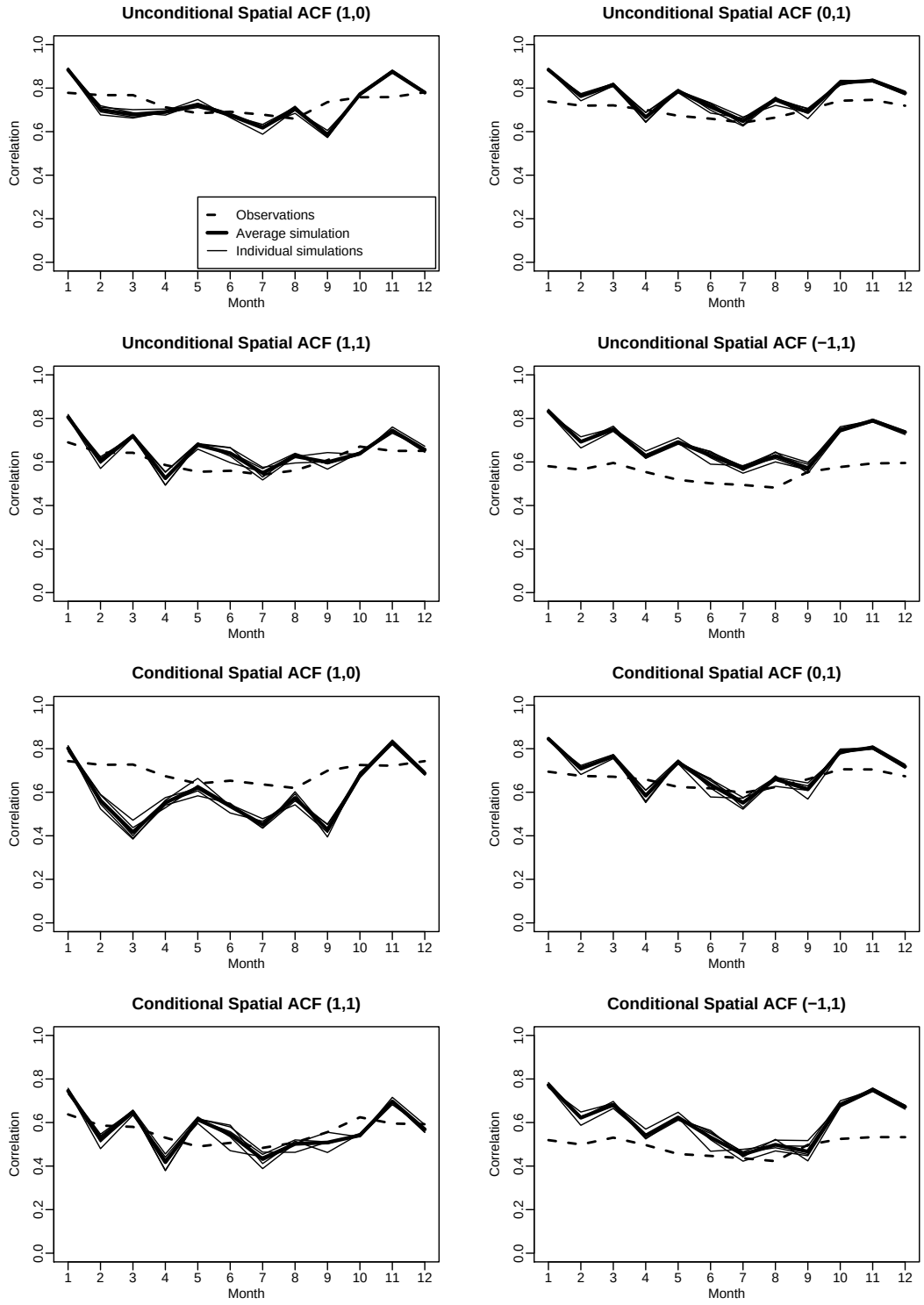
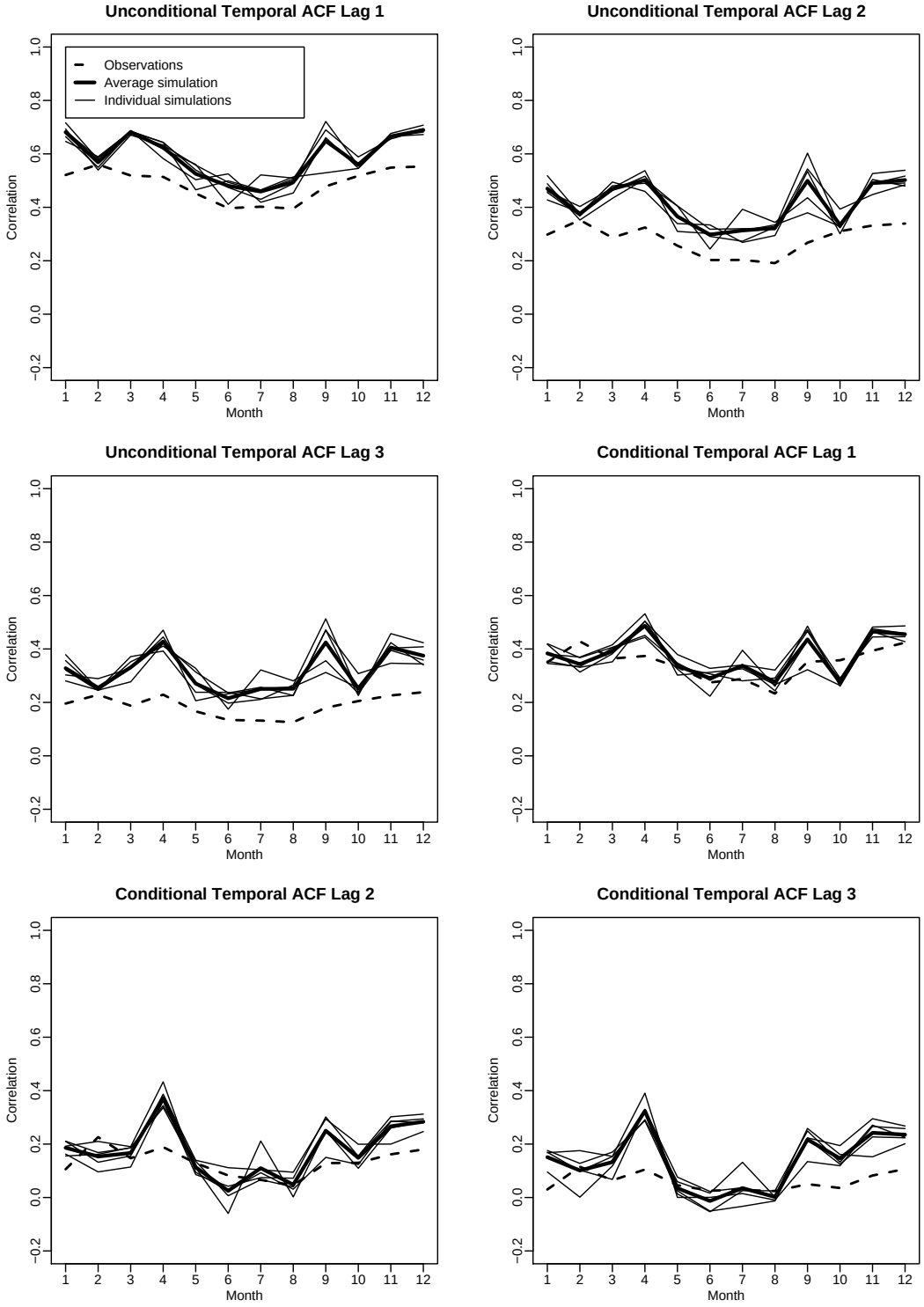


Figure 10.29 Summary statistics for simulated and empirical data (hourly & 64 km^2): temporal autocorrelations.



more apparent in the figures presented here.

As is to be expected, the spatial ACFs are generally a bit lower than before for both data and simulations. In addition, there is rather more discrepancy between the observed and simulated properties at the higher spatial scale, perhaps reflecting a greater emphasis on the finer-scale properties in the fitting process. The temporal ACFs are more or less unchanged.

Comparison of figures 10.30–10.32 with figures 10.27–10.29 shows the effect of increasing the temporal scale to the daily level, while maintaining the spatial scale at the 64km^2. The overall proportions of wet pixels increase considerably with the increased temporal scale, with the simulated values now overestimating the proportions in the data for quite a few months. Similarly, the probability that a pixel is wet, conditionally on the image being wet, is higher at the larger temporal scale. This effect is greater for the simulated images than for the empirical data, so that the discrepancies between two sets of probabilities are reduced at the larger scale. The conditional overall means are much closer together at the daily level, which is to be expected given all the extra pixels that are dry at the hourly level and are now being included in the averages. The standard deviations are, not surprisingly, much smaller at the daily level, with very good agreement (apart from November) between simulated and empirical values.

The spatial autocorrelations are to be expected to increase slightly with increased temporal averaging, reflecting the movement of the rain events over the day. However, the fit of the model is very similar at both temporal levels. At the daily level, there is very little temporal autocorrelation left in either simulated or empirical data, although at this level differences between the individual simulations begin to be noticeable, especially for the conditional moments. Note that the calculations for lag 3 at the daily time scale involve hourly values up to 95 hours apart.

There is another possible cause, in addition to those discussed above, for slight discrepancies between the observed and simulated data. This arises from the algorithm used in the simulation software for ensuring that events exceed the 15% coverage threshold for a specified period of time: when generating a short event with a long mean storm duration parameter, this algorithm may generate rain for a longer period than the specified event duration. This is because the algorithm generates storms until it achieves the coverage threshold of 15%, then continues to simulate storms and eventually deletes as many as necessary to ensure that the coverage falls below the threshold after the required duration. However, the initial storms necessary for the event arrival are never deleted. If a storm longer than the event duration is generated before the coverage threshold is achieved, that storm will continue to produce rain until it dies, so that the simulated duration of the event will be longer than intended. It is, however, not thought that the effect of this is significant.

As noted earlier, we have also looked at results for three other spatial-temporal scales: hourly & 16km^2; hourly & 256km^2; 15 mins & 64km^2. The comparisons for the hourly data at the intermediate, 16km^2 spatial scale and the larger, 256km^2 are in line with those given here for the other two spatial scales. We have only looked at the shorter, 15 minute time scale in conjunction with the 64km^2 spatial scale, but here a notable effect is a considerable improvement in the temporal autocorrelations.

Figure 10.30 Summary statistics for simulated and empirical data (daily & 64 km^2): means, standard deviations and proportion wet.

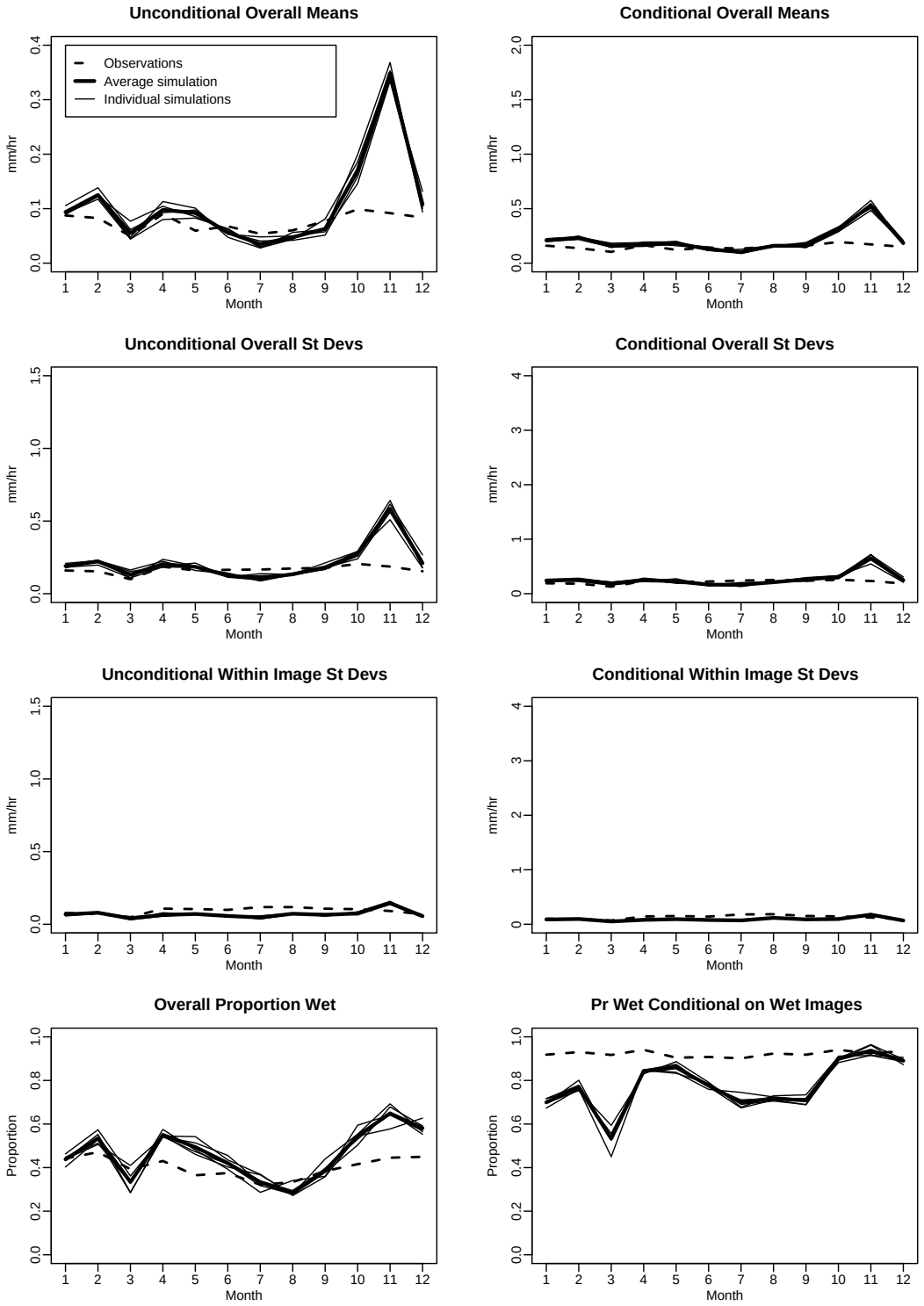


Figure 10.31 Summary statistics for simulated and empirical data (daily & 64 km^2): spatial auto-correlations.

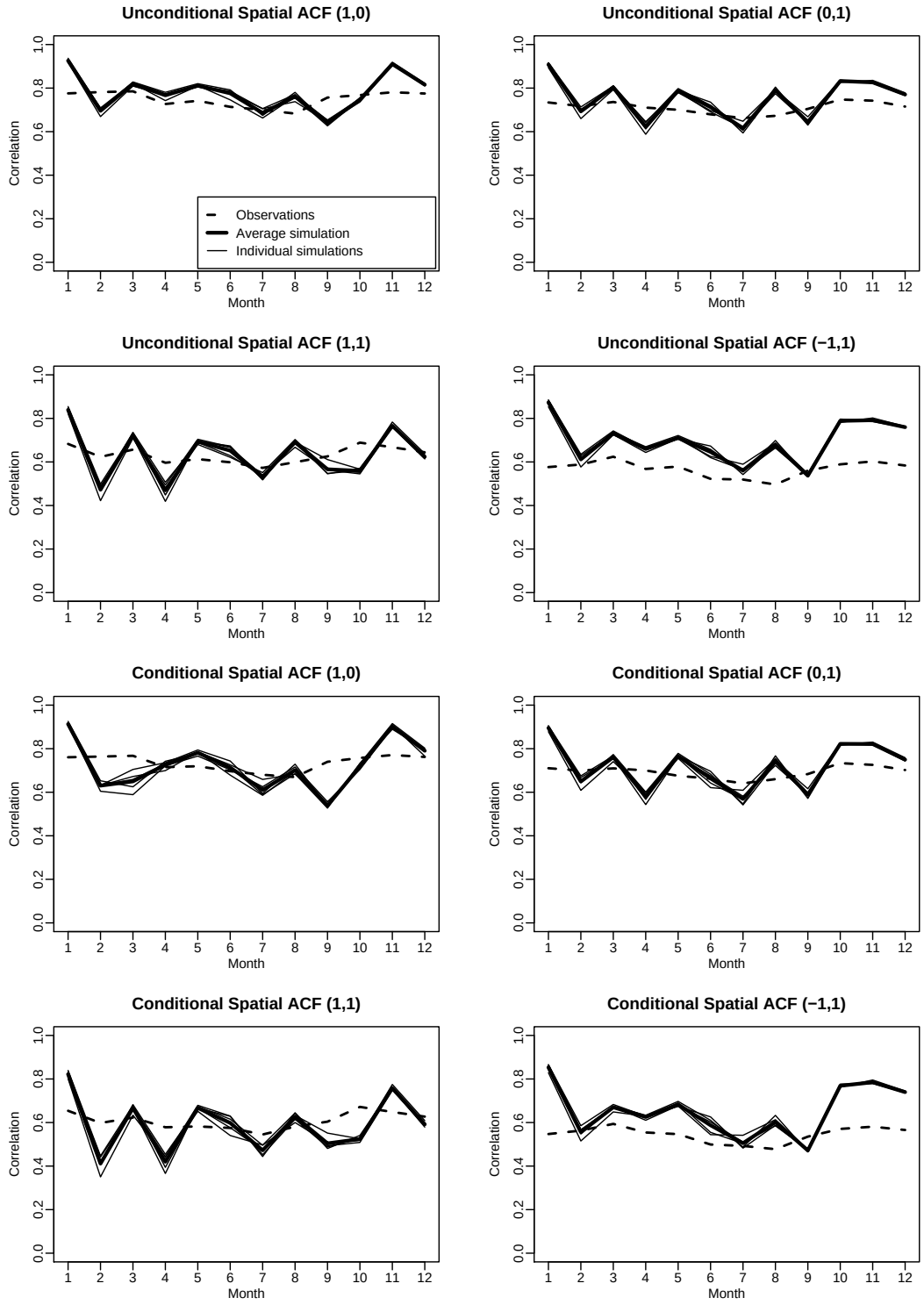
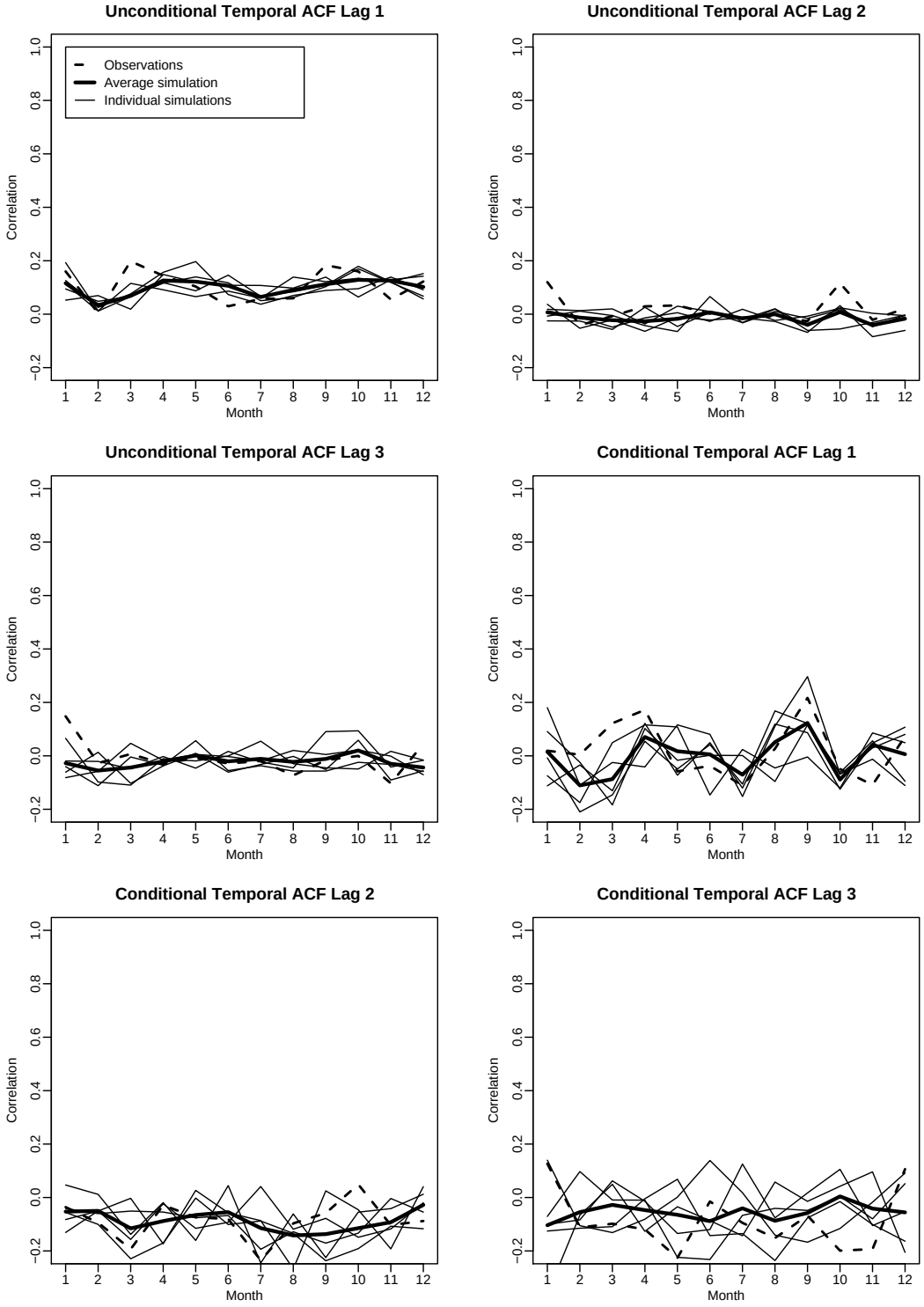


Figure 10.32 Summary statistics for simulated and empirical data (daily & 64 km^2): temporal autocorrelations.



10.5 Continuous simulation results for the non-stationary model

In the previous section, the performance of the continuous rainfall simulations for the Chenies radar region has been assessed, and various discrepancies between the properties of the observed and simulated data have been noted. The model is spatially homogeneous, whereas topological and other geographic effects will influence the empirical process. In addition, in the model, rainfall intensities are constant—and hence perfectly correlated—within rain cells, while real rain cells have spatial autocorrelations below unity. It is therefore possible that some of the observed lack of fit will be reduced by use of the non-stationary continuous simulation in which stationary simulations are rescaled using simulations of GLMs that allow spatial heterogeneity, as described in Section 9.5. Any such effects will be most evident at the larger spatial scales. Another advantage of the rescaling, is that it allows some temporal nonstationarities to be incorporated. Thus, in the rescaling results presented below, the simulation period is for 12 years, from 1 January 1990 to 31 December 2002, but omitting 1992 as discussed in Section 10.4. In addition, the GLMs are fitted using raingauge data, so that the rescaling will have the effect of further recalibrating the simulation results to represent ground truth (although of course the radar data to which the stochastic models have been fitted have already been calibrated against a data from network of gauges).

As for the stationary simulation discussed in Section 10.4, we shall present results for three space-time combinations as exemplars. As before, these combinations are hourly & 4km^2; hourly & 64km^2; daily & 64km^2. Figures 10.33, 10.35 and 10.37 show the mean, standard deviations (overall and within-image) and proportion wet, giving both unconditional and conditional properties. Figures 10.34, 10.36 and 10.38 show the corresponding conditional and unconditional spatial autocorrelation functions for the four nearest neighbour spatial lags specified above. Since the rescaling at a particular pixel is constant over a month, the temporal autocorrelations are unaffected, so these results will not be repeated here. The rescaling has been done by using data from a network of raingauges in the Blackwater catchment (see Sections 7.1 and 7.2). This network extends over a rather smaller region than the fitting window for the Chenies radar, and we have therefore limited our comparisons of rescaled simulations and empirical data to a subset of the fitting window of size 25×20 pixels (50×40 km^2) that corresponds approximately to the region covered by the rain gauge network.

First, we look at the finest resolution, and compare Figures 10.33 and 10.34 with Figures 10.24 and 10.25. The comparisons between results for stationary and nonstationary simulations at the other two resolutions (Figures 10.35 and 10.36 versus Figures 10.27 and 10.28; and Figures 10.37 and 10.38 versus Figures 10.30 and 10.31) are qualitatively the same. As expected, rescaling removes the problem previously observed with the overestimation of the overall mean in November. The various standard deviations are generally somewhat improved, particularly in the summer, where the GLM provides extra information to help to make up for the relative shortage of rain events. The rescaling is multiplicative and therefore has no effect on the proportion of wet pixels either overall or conditionally on a wet image. The differences that can be seen in these plots are explained by the fact that, as noted above, the figures for the rescaled simulations are restricted to a smaller spatial region

Figure 10.33 Summary statistics for rescaled simulations and empirical data (hourly & 4 km^2): means, standard deviations and proportion wet.

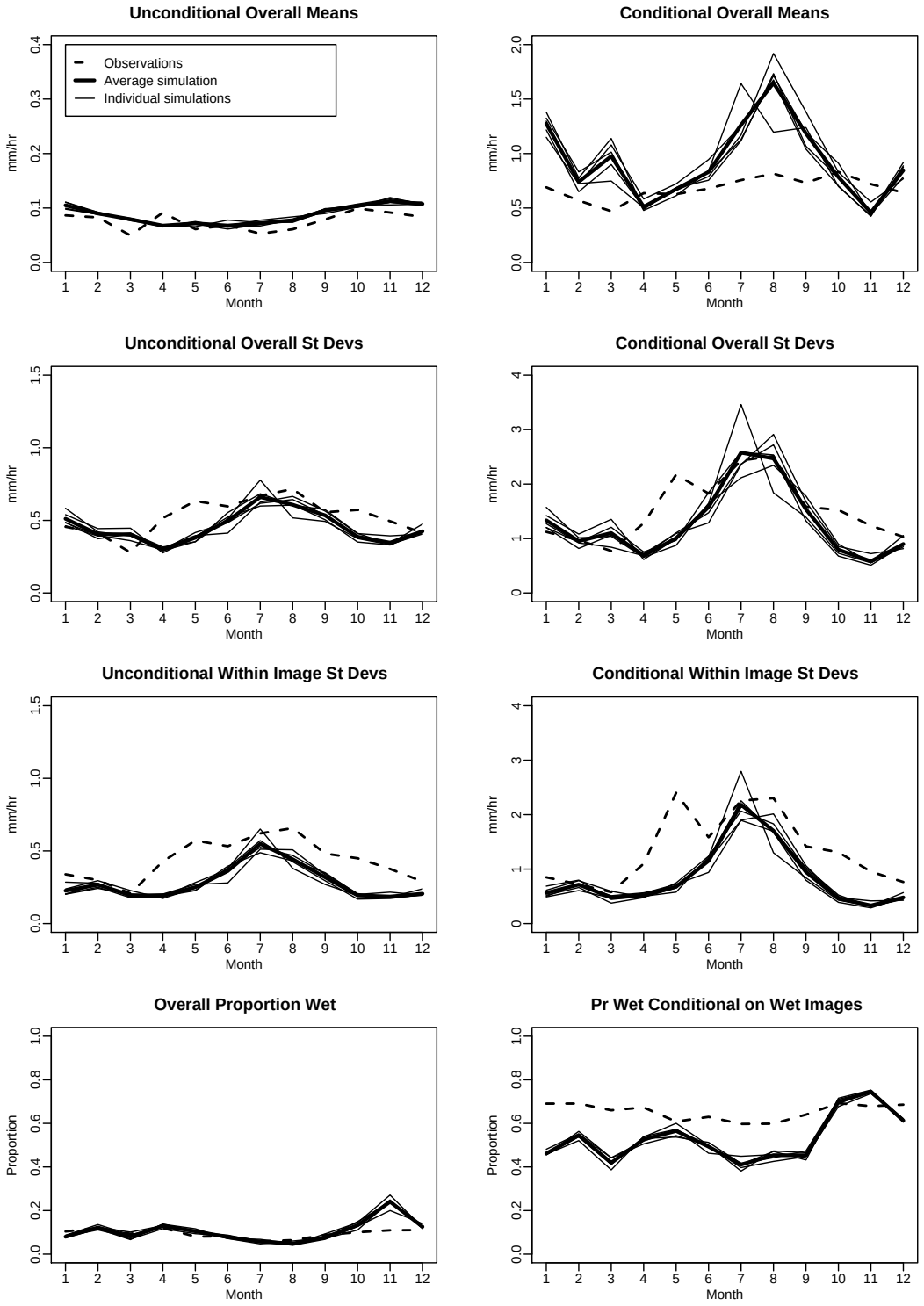
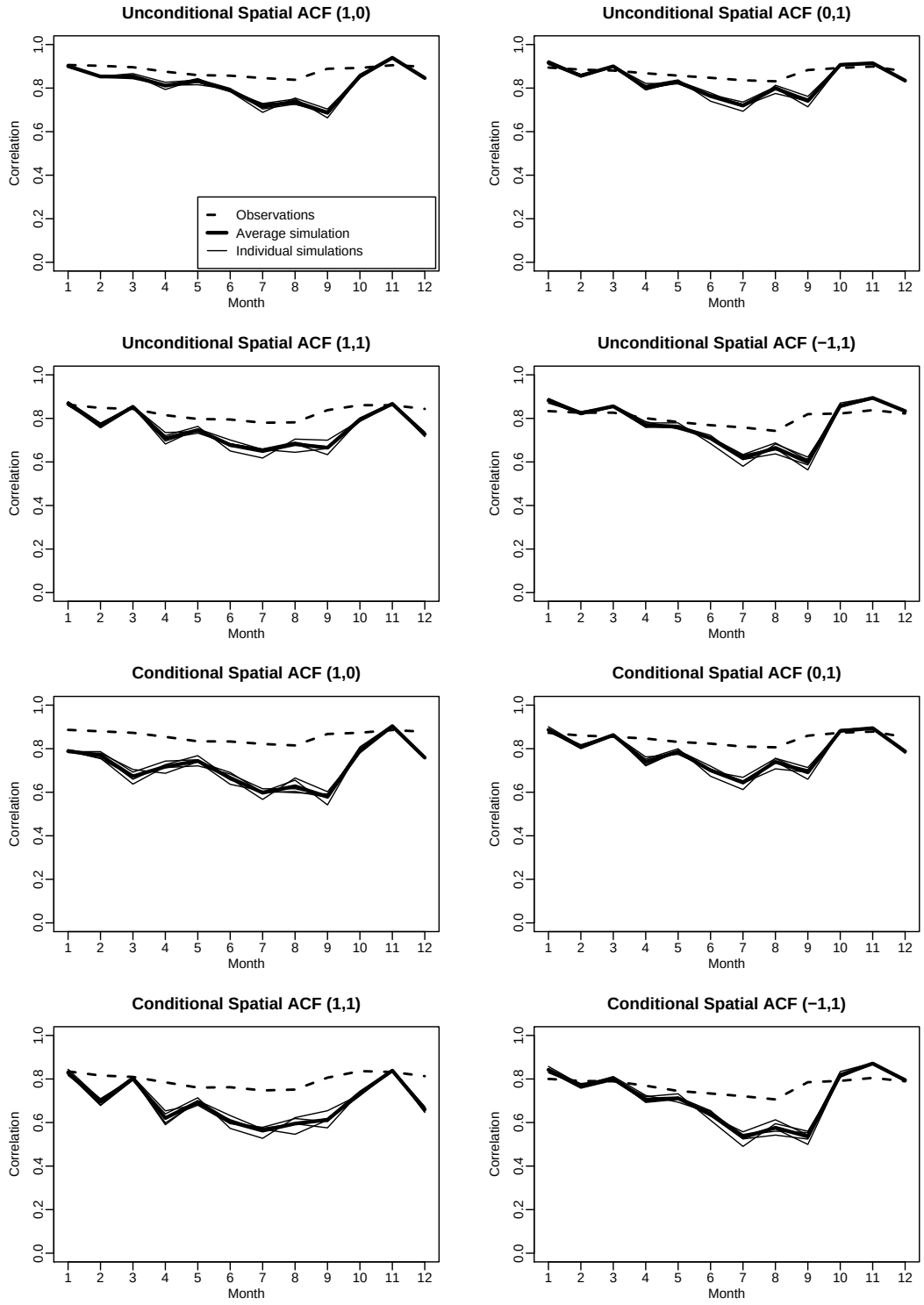


Figure 10.34 Summary statistics for rescaled simulations and empirical data (hourly & 4 km^2): spatial autocorrelations.



than that used for the the stationary simulation.

A comparison of Figures 10.25 and 10.34 shows that the rescaling reduces the simulated spatial autocorrelations, particularly at lags (1,0) and (1,1). This can be explained heuristically by the fact that the rescaling tends to make neighbouring pixels “less similar” than previously. Since the stationary simulations show generally good agreement between the observed and simulated spatial autocorrelations, the rescaling therefore induces underestimation. The primary aim of the simple rescaling scheme adopted here is to reproduce the systematic variation in mean rainfall: it seems that underestimation of the spatial autocorrelations is the price to be paid for this. To overcome the problem it would be necessary to develop a rather more sophisticated scheme for nonstationary simulation. A considerable amount of further work would be required to achieve this.

10.6 Summary of continuous space-time simulation

In this chapter, we have described the performance of the components of the model for the continuous simulation of space-time rainfall fields (that is, the stochastic models for the interiors of rain events and for the arrival and departure of rain events over the simulation window or catchment) together with the results of continuous simulations of the combined model both before and after rescaling to allow for spatial and temporal nonstationarities.

Overall, the event interior model and its fitting have been extensively tested and the results are encouraging. In particular, when model fitting is ‘successful’ (according to the criteria described in Section 10.2), the model performs well in reproducing basic statistical properties, including the proportion of wet pixels, over a range of time and spatial scales. The model does appear to have some difficulties in fitting intense events; further development is required to resolve this. However, at present a bigger problem seems to be that the model can only be fitted to data from events that cover the radar window for a substantial period of time, and can be clearly distinguished from radar noise and clutter. Parameter sets from the fitted events are used as a library from which sets are sampled for the continuous simulation. This means that the simulations have no means of representing rainfall events, for example periods of very light scattered showers, that are not suitable for fitting under the criteria that have been used (in fact, less than 10% of all events were suitable for fitting). Similarly, rainfall events that only pass over a part of the window and for a rather short time will not be fitted. It was found in the previous report (Wheater et al., 2000b) that such events are not generally qualitatively different from fitting events, and so the parameter sets in the fitting library will be representative of these events. However, by its construction, the current simulation program will only generate realised events that move across the whole window. In this respect, the continuous simulations differ significantly from the observed data. To see beyond this when comparing data and simulations, a threshold coverage of 15% has been applied to all images, with all pixels in images below this coverage being given intensities of zero. Given this threshold, the models generally perform acceptably in reproducing the statistical properties of the observed data. However, further model development is needed to enable events that pass over only a part of the simulation window, and rainfall that is light and spatially and/or temporally

Figure 10.35 Summary statistics for rescaled simulations and empirical data (hourly & 64 km^2): means, standard deviations and proportion wet.

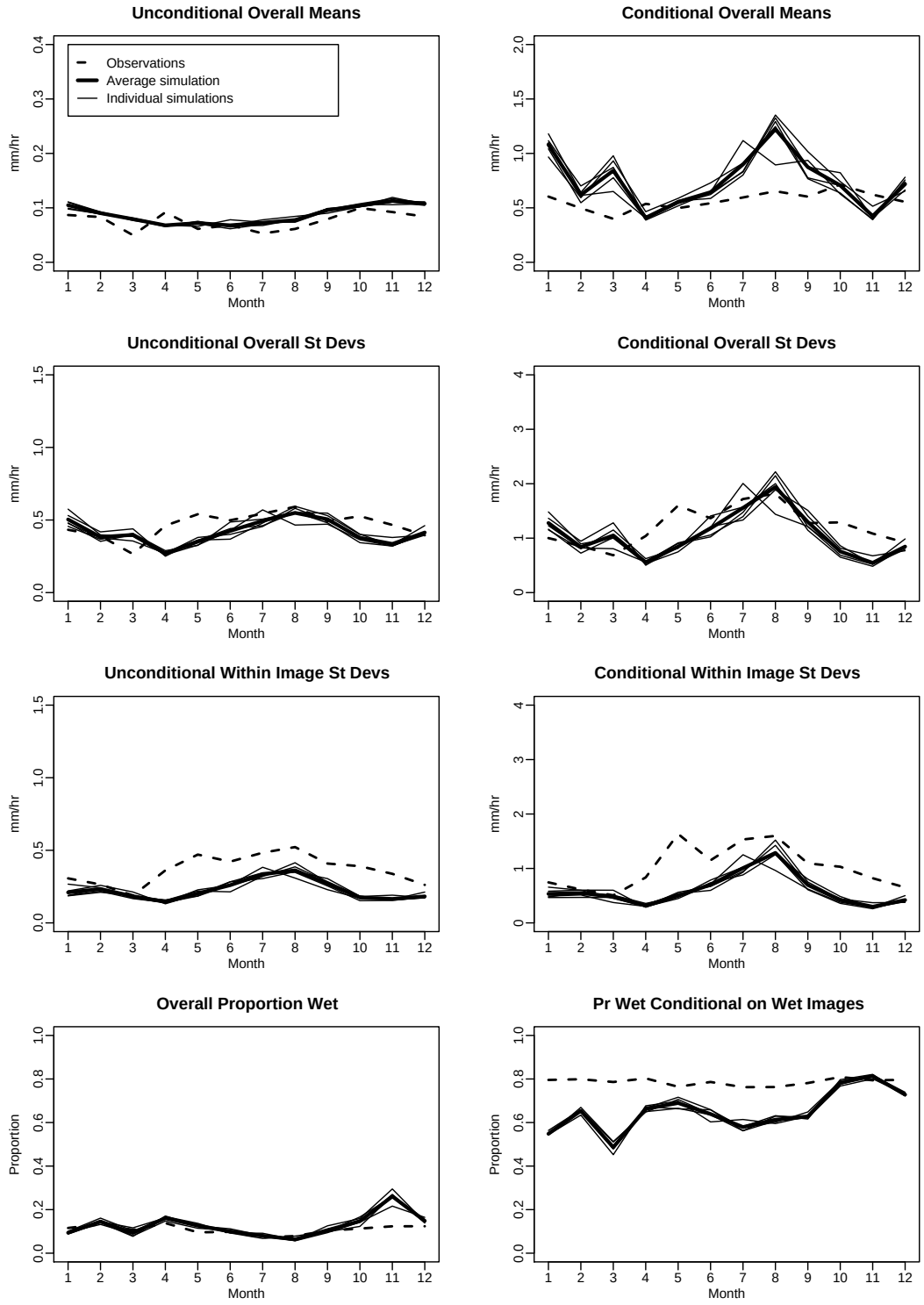


Figure 10.36 Summary statistics for rescaled simulations and empirical data (hourly & 64 km^2): spatial autocorrelations.

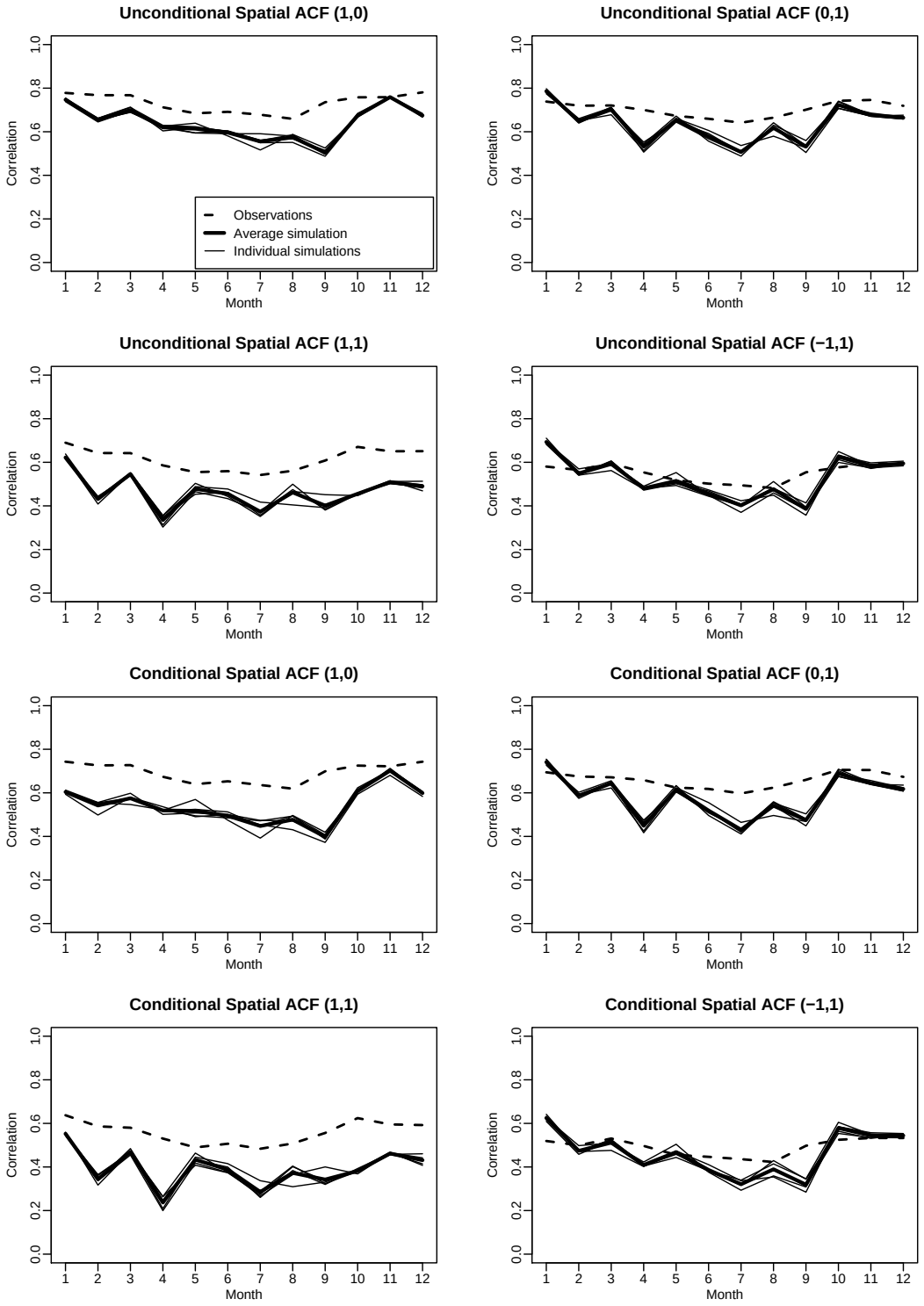


Figure 10.37 Summary statistics for rescaled simulations and empirical data (daily & 64 km^2): means, standard deviations and proportion wet.

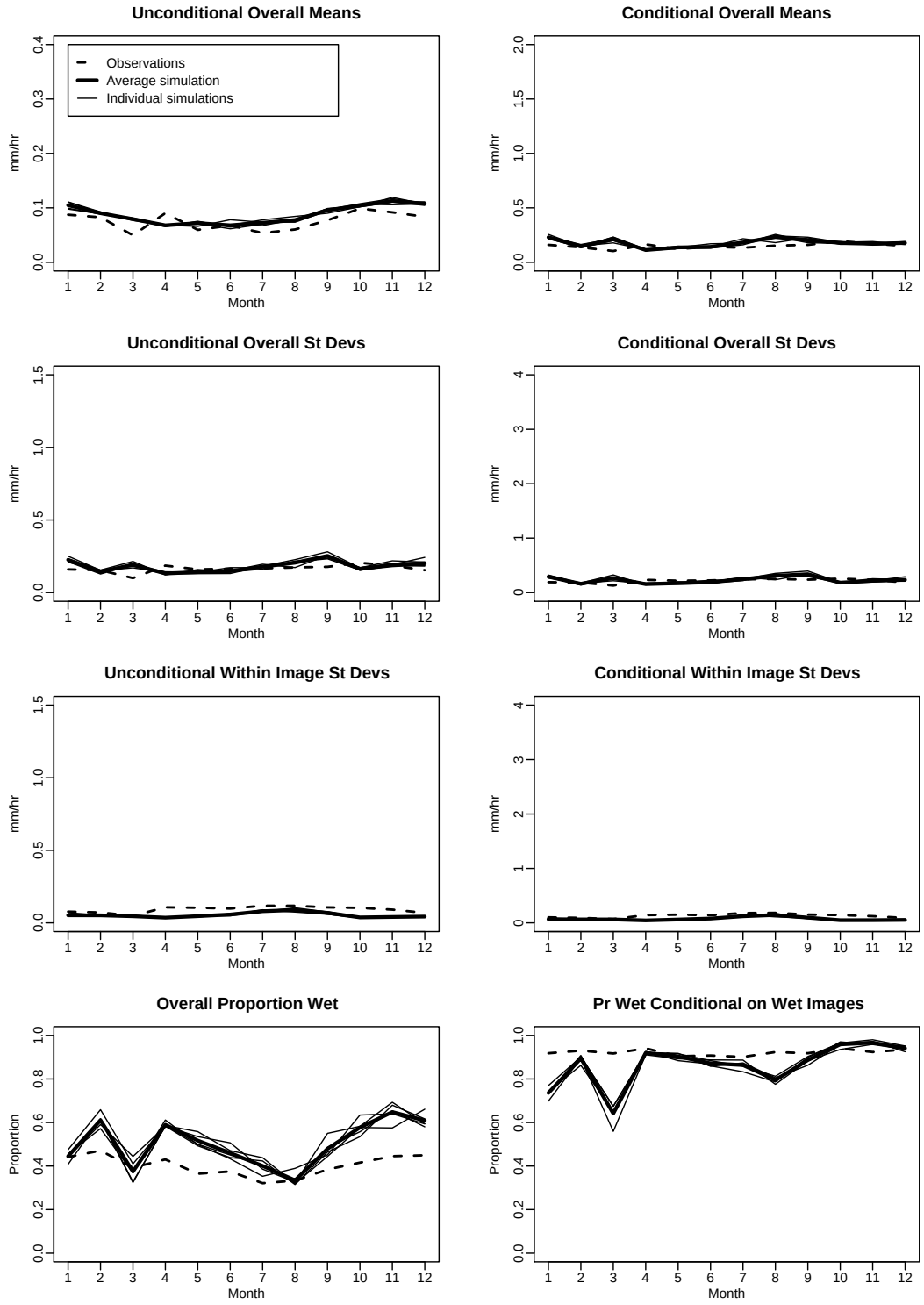
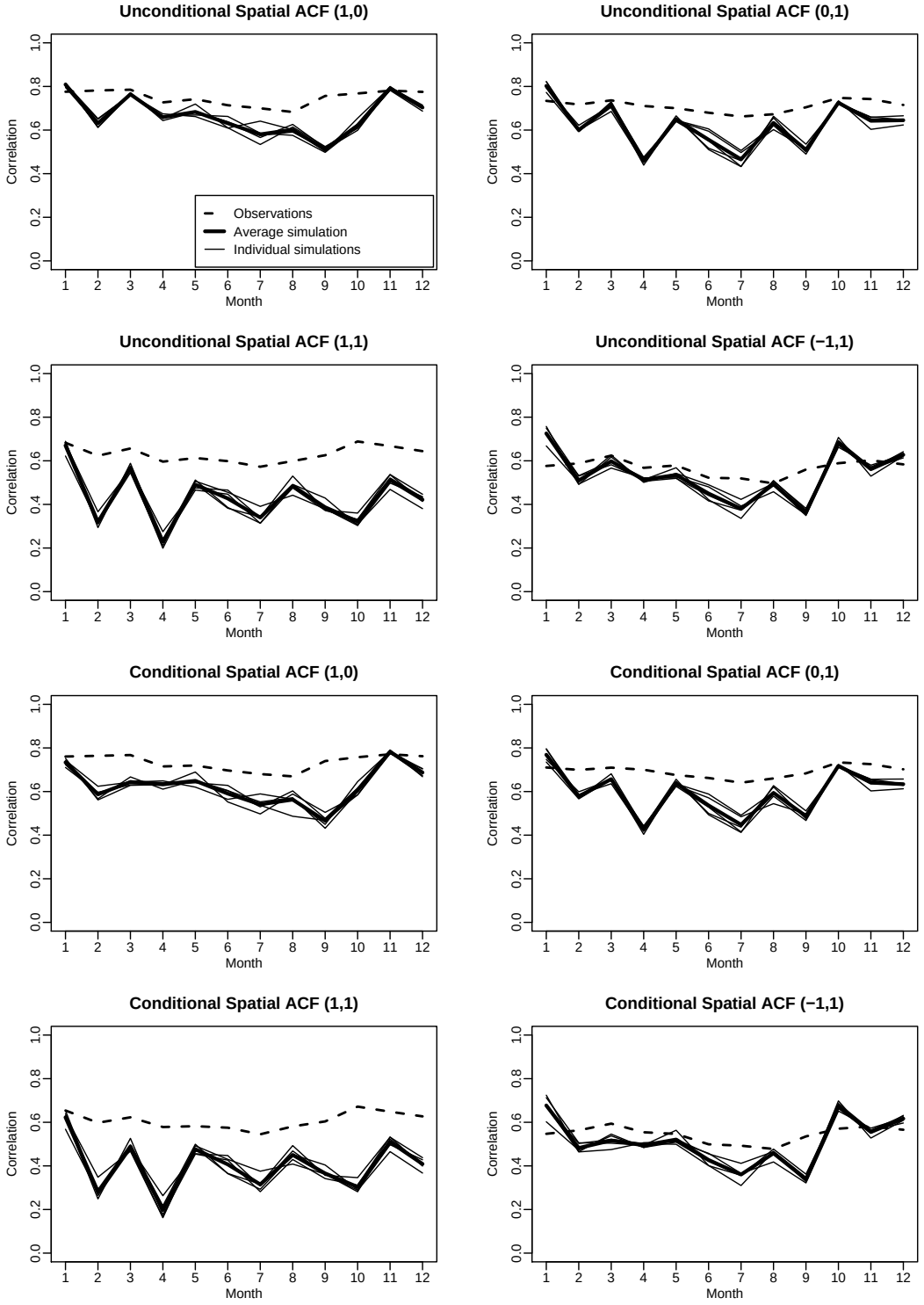


Figure 10.38 Summary statistics for rescaled simulations and empirical data (daily & 64 km^2): spatial autocorrelations.



intermittent to be included in the simulation.

As well as carrying out extensive tests of the methodology for continuous simulation of stationary sequences, we have demonstrated for the first time the potential for generating sequences that are nonstationary in both space and time. The proposed scheme works by rescaling stationary sequences so as to match the nonstationary means derived from raingauge networks. As such, it is conceptually simple and, to some extent at least, removes problems associated with radar calibration since systematic errors resulting from radar over- or underestimation will ultimately be rescaled to match the gauge means. Furthermore, it provides a means of incorporating temporal nonstationarity into simulated sequences (although the rescaling affects only non-zero amounts so that trends in, for example, the proportion of wet pixels cannot be reproduced by this method). As well as the mean, rescaling seems to yield a better match between observed and simulated standard deviations. The price to be paid for these improvements is that the spatial autocorrelation structure is less well reproduced after rescaling. Further work is required to address this.

As described above, the main drawback of the continuous simulation scheme at present is its inability to generate small, scattered rainfall episodes. There are, however, a few other possible reasons for discrepancies between the continuous simulations and the radar data, all of which will need to be examined to improve the simulation performance. They are as follows:

- The characteristics of the radar data themselves change over time. In particular, there is more clutter in the late 1990s and it has been noted (Section F.4) that to allow for this, it may be necessary to raise the coverage threshold used for defining events during this period.
- The use of any fixed coverage threshold to define events is not entirely satisfactory — for example, with the 15% threshold used here there are still several events that last in excess of 24 hours and appear to consist of distinct “pulses” of rain (Section 10.1). It may be that improved performance could be obtained by splitting some of these long events up into shorter ones on the basis of a manual inspection of the data.
- Some of the properties of the radar data themselves are extremely sensitive to decisions made in the calibration process (Section F.3.2). This affects temporal autocorrelations in particular. It is necessary to resolve these calibration issues, in order to be sure that the simulation models are being fitted to realistic rainfall sequences rather than to artefacts of the calibration.
- The event interior model represents rain cells as having a constant intensity over ellipses with regular boundaries, whereas in reality cell boundaries are likely to be less well defined and intensities will vary. This will affect the comparison of simulated and observed rainfall properties at relatively small spatial and temporal scales.
- The sampling of the event interior parameter library may not be representative. It has already been noted that the fitting events may not be a representative subset of all the events to affect the radar window; in addition to this, however, it is necessary to determine whether the procedure for selecting a parameter set

based on the simulated event duration (Chapter 9) is appropriate; it is possible that some parameter sets are being sampled too often.

- When fitting Weibull distributions to the durations of wet and dry intervals, all censored intervals have been treated as right censored, ignoring any information about their maximum durations. This will tend to result in a slight overestimation of all interval lengths; the effect will be greater for dry intervals, however, since these tend to be longer and hence more likely to be censored than wet intervals.

It is unfortunate that within the current project it has not been possible to investigate these issues and, more generally, the extension of the stationary model to represent light and spatially and/or temporally intermittent rainfall, in greater depth. The main reason for this was the unexpectedly very substantial amount of work that had to be done to process and calibrate the Chenies radar data before the work on model development could begin. In consultation with DEFRA, it was decided that it was essential that these data were properly processed and calibrated before the spatial-temporal modelling could usefully proceed. However, this set back progress on model development very considerably and means that further work is needed before continuous simulations from these models will properly reflect the full range of spatial-temporal rainfall variability. We also note that a substantial amount of effort has been devoted to improving the software for fitting and simulation of the continuous space-time simulation models; this software is now, we believe, rather more robust than previously and can be used with much more confidence for further model development.

Part IV

Effect of spatial structure of (large) catchment-scale rainfall on runoff estimation

The report thus far has presented models for single-site rainfall, and alternative procedures for spatial-temporal rainfall. For the practitioner, a central question is:

- To what extent is the representation of spatial rainfall important in rainfall-runoff modelling?

This can be expected to depend on the type of rainfall, the type of catchment, and the spatial scale of the problem (i.e. the catchment area) and, following from this:

- If the spatial properties of rainfall are important, what is necessary to characterise these in terms of both observational data and simulation methods?

To provide guidance, Part IV presents, in Chapter 11, a review of the significance of spatial data for flood runoff generation based on the international research literature. More specific guidance is needed, with particular reference to UK conditions, and Chapter 12 reports on a series of numerical experiments, based on the river Lee, to the north of London. An event-based semi-distributed rainfall-runoff model is fitted to individual sub-catchments, and to the catchment as a whole, and the sensitivity of simulated runoff to different representations of rainfall input is explored. A reference rainfall field is based on the full set of subdaily raingauge data, and compared with descriptors based on reduced numbers of raingauges and on radar. The spatial-temporal disaggregation procedure of Part II (Chapter 8) is also evaluated.

The concluding Chapter of this section (Chapter 13) presents results from an independent case study, by staff of CEH-Wallingford, based on daily rainfall and flow modelling. This presents an application of a distributed rainfall-runoff model using the GLM methods of Part II (Chapter 6) to define input rainfall sequences.

Chapter 11 The significance of spatial data for flood runoff generation

11.1 Introduction

The spatial distribution of rainfall and its temporal evolution are of fundamental importance to hydrological applications. In particular, flood and water resource simulation studies require spatial-temporal rainfall data as input to distributed rainfall-runoff models. Different types of data are available. A raingauge gives reasonably accurate measurements of rainfall at a point, however, a considerable number of raingauges would be required to obtain a detailed representation of the rainfall field structure, and this would prove too costly for routine monitoring. Weather radar can provide a representation of the structure of the rainfall field, but the quantitative measurements are less accurate than what can be obtained by raingauges. Although considerable work has been done over the last three decades in developing procedures to adjust the radar measurement of precipitation, calibration issues remain problematic. In addition, radar data are not always available and existing record lengths are relatively short. An alternative approach is to design synthetic rainfall or develop stochastic models that generate realistic rainfall fields that can be input to rainfall-runoff models.

The explicit representation of rainfall spatial variability in hydrological model input allows derivation of a better understanding of the relationship between runoff response and spatial rainfall patterns. A question that still needs to be answered is the level of complexity required in the representation of spatial rainfall. Singh (1997) showed that the spatial and temporal variability of rainfall can significantly influence the flood hydrograph and shape, but the importance varies as a function of catchment rainfall properties. However no clear overall conclusions emerge from his review. The aim of this chapter is to gain insights into the significance of spatial rainfall for runoff generation for a range of rainfall types, catchment types and catchment scales. A practical objective is to provide recommendations on the spatial resolution of precipitation input needed to model watersheds in medium temperate climatic regions. The review is supplemented by an investigation of the effects of spatial rainfall representation on runoff carried out on the Lee catchment, UK and reported in Chapter 12.

The sensitivity of the hydrological response to the variability of rainfall is usually the result of a nonlinear interaction between the rainfall and some other component of the system (Over, 1995). When studying the runoff response of a given catchment, the modeller has to take into account (a) the storm characteristics, (b) the catchment properties, (c) the effect of scale and (d) the interactions between these three components. Assuming a moderate impact of the model structure, a rainfall-runoff model can be used to test various rainfall representations and the generated streamflow is compared with the observed one. However hydrological models and their parameters are scale-dependent and their use as a tool to determine the impact of spatial rainfall is therefore questionable (Morin et al., 2001; Koren et al., 1999).

We propose to study the impact of these three interactions in association with the

modelling needs for different conditions. On impervious areas, a high proportion of the rain becomes effective and produces runoff whereas on pervious areas, rainfall variations are damped by the integrating reaction of the catchment. The review is structured such that we consider different catchment types ranging from fast to slow flow response. Similarly, the effect of rainfall type is introduced ranging from localised highly variable convective rainfall to relatively more homogeneous large scale frontal structures. The effect of catchment scale is also included. In order to get a coherent view of the controls, three catchments types are selected.

Fast-responding, usually small, urban catchments are presented in Section 11.2. As a natural extension, the impact of highly spatially variable rainfall on fast-responding natural basins is assessed. To illustrate these conditions, the focus is on semi-arid areas characterised by a dominance of Hortonian overland flow and associated with convectively-dominated rainfall. The findings are presented in Section 11.3 for a range of scale, from small ($<$100 km^2), medium (100-2000 km^2) to large ($>$2000 km^2) catchments. The review at this stage does not aim to be comprehensive; articles are selected so that key points are highlighted. A more extensive review for humid temperate regions is undertaken as more complexity is encountered when modelling this type of catchment and because ultimately we want to apply the findings to catchments under these climatic conditions. Section 11.4 presents the findings for the dominantly large-scale frontal rainfall in humid temperate climatic regions, as a function of catchment scale, again ranging from small ($<$100 km^2), medium (100-2000 km^2) to large ($>$2000 km^2) catchments. The main conclusions for urban, semi-arid and natural basins are presented in Section 11.2.3, 11.3.4 and 11.4.4. Overall conclusions and proposed extensions to future work can be found in Section 11.5.

11.2 Urban Catchments

11.2.1 Effect of Storm Movement

Singh (1997) observed that the effect of storm movement is significant for urban watersheds. From his review of the literature, he suggested that on an urban watershed a rain storm moving in the direction of flow produces a higher peak than it would if it were moving in the opposite direction. The effect of storm speed on peak discharge is much less for rapidly moving storms than for storms moving at about the same speed as the flow velocity.

Using synthetic urban catchments, represented by a hydrologic, semi-distributed non-linear rainfall-runoff model, Ngirane-Katashaya and Wheater (1985) showed that the hydrograph is sensitive to the velocity and direction of the storm. Different storms were used, designed with uniform intensity in time and space, passing at 30 degrees intervals over 3 synthetic catchments of 0.31, 2.5 and 20 km^2. Eight storm speeds in the range 0.156-20.0 ms^{-1} were tested under fast and slow responding catchment conditions. Results were compared with equivalent stationary storms. Peak discharge for a downstream-moving storm was greater than for the corresponding upstream-moving storm over the whole range of storm speeds, and with the exception of the 5 ms^{-1} storm speed, a consistent variation with storm direction was observed. The peak discharge owing to the downstream-moving storm increased initially with

increasing storm speed, reached a maximum and then decreased. However, for the upstream-moving storms the peak discharge increased, although at a decreasing rate, throughout the range of storm speeds. Identical patterns were seen for all catchments. However, the variation of magnitude of discharge between different catchment areas was masked by the effect of the response characteristic of the catchment. Discharge rates for the 2.5 km^2 catchment with high response were greater than for the 20 km^2 catchment with low response. The smaller catchments gave similar results, with a maximum increase in peak discharge, in comparison with the one obtained from a uniform rainfall, of 45% and a maximum reduction of 23-24%. For the larger catchments a maximum reduction of 24% was also observed, but the maximum enhancement is slightly less than for the other catchments (41%). For fast storm movement and a rapidly responding catchment, at resonance conditions, the peak discharge downstream was on average 87% greater than upstream. The resonance speed was lower for slow response conditions and for the largest catchment area.

de Lima and Singh (2002) conducted a study for a synthetic impervious plane surface of 100m$\times$1m using a hydraulic model of overland flow, based on the nonlinear kinematic wave. The results for hypothetical storms with varying temporal pattern (uniform, intermediate, advanced and delayed), length, speed (0.5 to 2 ms^{-1}) and direction (upstream and downstream) were compared. They found that peak discharges and hydrograph shapes depend strongly on the storm temporal pattern; two distinct hydrologic responses were observed for storms moving upstream and downstream. Storms moving upstream are normally characterised by hydrographs with early rise, low peak discharge, not so steep rising limb, and long base time. They observed that the sensitivity of runoff to storm patterns decreases at high storm speeds. Finally, when comparing equivalent storms, downstream storm movement presents larger differences in the hydrograph shapes for different rainfall patterns then do upstream movement.

11.2.2 Spatial Rainfall

Aronica and Cannarozzo (2000) developed a conceptual, hydraulic, semi-distributed rainfall-runoff model and assessed the impact on the modelled hydrograph due to changes in the spatial representation of rainfall and in the spatial discretisation of the sub-catchments. The urban, 12.8 ha Parco d'Orleans catchment was used as a case-study. Four events were selected and recorded by two raingauges. Three rainfall scenarios were tested, a uniform rainfall for each subcatchment based on raingauge 1; a uniform rainfall for each subcatchment based on raingauge 2 and a spatially variable hyetograph based on both raingauges. Then using the spatially variable rainfall as input to the model, the catchment was subdivided into 32, 18, and 10 subareas with a mean subarea of 0.40, 0.71 and 1.28 ha respectively. Results showed no influence on the time to peak but both spatial representation of the rainfall and spatial discretisation of the subcatchments had a significant effect on the outlet hydrograph peak (between 0.3 and 44.5% change depending on the rainfall representation) and volume (0.9-19.4% change), especially for high intensity events.

Berne et al. (2004) provided recommendations on the temporal and spatial resolution of rainfall measurements required for urban hydrological applications, based on a high resolution precipitation data set from the HYDROMET Integrated Radar Experiment

98 (HIRE 98). Rainfall data came from a network of 25 rain gauges and from 2 radar networks covering 6 sub-basins of 38 ha to 105 km^2 around Marseilles, France. The three most intense events during autumn 98 in terms of rainfall intensity and total rain amount were selected as representative of the Mediterranean precipitation.

The characteristic time of the catchment, taken as the lag time, i.e. the time difference between the centroid of the catchment-averaged rainfall and the centroid of the modelled hydrograph, was calculated for each of the six catchments and plotted versus surface area (in ha). A power law relation between the lag time t_c (min) and surface area S (ha) was established. The required time resolution of rainfall Δt (defined as $t_c/4$) was estimated as a function of the surface area of the catchments according to:

$$\Delta t = 0.75 S^{0.3} \tag{11.1}$$

Using geostatistics, a simple relation between the range of the variogram of the spatial rainfall field (r km) and the time steps (Δt min) was fitted (both using raingauges and radar data at 1, 2, 3, 6, 12 min resolutions). Similarly, the required space resolution of rainfall ($r/3$) for Mediterranean regions was estimated as:

$$\Delta r = 1.5\sqrt{\Delta t} \tag{11.2}$$

where r the spatial resolution. They concluded that hydrological applications for urban catchments in the Mediterranean region of the order of 1000 ha require a temporal resolution of about 6 min and a spatial resolution of about 3.7 km. Modelling urban catchments of the order of 100 ha requires a resolution of about 3 min and 2 km.

Jensen and Pedersen (2005) examined the variability in accumulated rainfall within a 500m grid single radar pixel (as part of the Local Area Weather Radar, 15 km south of Arhus, Denmark) using nine high-resolution rain gauges. The spatial variability of rainfall was measured by the coefficient of variation. Results were presented for 7 events over a 65 days period. The mean of the rainfall events varied between 1.2 and 33.8 mm, the coefficient of variation ranged from 13.9 to 80.4%. The magnitude of variation in the accumulated rainfall at the nine points within this small area was found to be high with a variation from the mean ranging from 4 to 200%. This confirms that a high density raingauge network is required for runoff modelling of fast-responding urban catchments.

11.2.3 Conclusions

The studies show that the runoff response on urban catchments is sensitive to storm variability in space and time and is also affected by storm movement. More variability in the runoff hydrograph is introduced when the storm moves in the direction of flow and at low speed. Hydrological modelling requires fine spatial and temporal rainfall resolution. Smaller catchments require a higher resolution (Berne et al., 2004) and are more sensitive to storm movement (Ngirane-Katashaya and Wheater, 1985).

11.3 Semi-arid regions with convective rainfall

This section concerns semi-arid regions characterised by Hortonian overland flow and affected by convective storm systems; i.e. combining rapid runoff response with high

spatial variability of precipitation.

A large number of investigations were carried out on the semi-arid experimental Walnut Gulch watershed in Arizona across a range of scale from 4.4 ha to 150 km^2. The summer weather of this region is strongly affected by the North American monsoon, which results in frequent air mass thunderstorms that are highly convective, intense, localised and of short duration. This is complemented with studies in semi-arid areas also subject to convective thunderstorms and conducted in other part of the US and in Israel. An attempt to summarise the main findings is undertaken below.

The last study is conducted in Mexico City (Arnaud et al., 2002). This is not exactly a semi-arid area, possessing an average annual rainfall between 600 and 1000 mm. However it has pronounced dry and wet seasons associated with heavy convective rainfall, and the catchment is fast-responding. The findings are therefore presented here to illustrate effects at the large catchment scale.

11.3.1 Small Catchments ($<$ 100 km^2)

Faurès et al. (1995) found that accurately estimating the spatial variability of rainfall in small catchments submitted to convective thunderstorms is of fundamental importance for runoff simulation. The study was conducted on a 4.4 ha catchment as part of the 150 km^2 semi-arid Walnut Gulch experimental watershed using a distributed model applied to 8 events. Results show that simulation errors are halved as the number of raingauges is increased from one to five, and that at a small scale of 4.4 ha, at least 4 raingauges are required to model appropriately the runoff.

Lopes (1996) examined the effect of uncertainty in spatial estimates of rainfall in predictions of runoff volume, peak runoff and sediment yield using a distributed, event-based runoff-erosion model on a 6.73 km^2 catchment as part of Walnut Gulch. The best areal rainfall estimate was given by an average of the 10 gauges available on the catchment and served to calibrate the model. Three storm rainfall events, representative of the localised, convective rainfall in the area were selected. This includes one large event (areal mean of 32.01 mm with a standard deviation of 10.84 mm) and two medium-sized events (areal mean of 24.51 mm with a standard deviation of 8.71 mm, and areal mean of 23.45 mm with a standard deviation of 6.11 mm). Results demonstrated that the density of the raingauge network, the spatial arrangement of individual raingauges, and the spatio-temporal characteristics of storm events had a significant effect on catchment response prediction. The relative error of prediction with reference to the case of 10 rain gauges was computed. Considering the largest storm, when the rain gauge density was decreased (randomly) from 10 to 6 rain gauges, prediction errors of the order of 30% were observed. As the network density was further reduced to 1 rain gauge, prediction errors increased to about 130%. However, errors in areal estimates of rainfall at the catchment scale increased by only 30% when computed from one raingauge. These results emphasise that the catchment response to rainfall input is nonlinear. Overall it seems that a network density of 5 raingauges ensures good modelling at the 6.7km ^{2} scale.

Morin et al. (2001) used radar data to determine a characteristic temporal scale for the hydrological response of a basin. The response time scale is defined as the time

resolution at which the radar data should be aggregated so that the rainfall hyetograph is most similar to the pattern of the measured outlet runoff hydrograph and represents the integrative properties of the drainage basin in transforming rainfall into runoff. One approach to study this scale is to test the effect of spatial and temporal variability of rainfall and basin properties on the streamflow using a specific rainfall-runoff model. However hydrological models and their parameters are scale-dependent and their use as a tool for determining characteristic scale is therefore questionable. Radar data with spatial resolution of 0.5 to 4 km^2 and a temporal resolution of 5 minutes were used together with discharge data from 4 small basins (10-100 km^2) in Israel. Spatially, the radar rainfall was averaged over the entire basin so that the temporal scale could be studied independently of the spatial scale. The response time scale was objectively identified using a heuristic method based on matching rainfall and runoff peaks and finding the minimum number of unmatched peaks for different time resolutions. They found a similarity between radar rainfall and outlet runoff at a specific temporal scale which was stable over several storms and therefore could be considered as an intrinsic property of the watershed. Relatively short time scales were found when modelling arid areas, especially for the urban and arid catchments (15-30 min) while for the rural catchments larger time scale were identified (90-180 min). Hence, the study confirms that arid catchments are fast-responding and this suggests that the runoff response is highly affected by the storm variability.

Ogden and Julien (1993) explored two-dimensional watershed sensitivity to the spatial and temporal variability of rainfall using a physically-based rainfall-runoff model. The rainfall-runoff model was applied to two watersheds in Idaho and Colorado, covering areas of 32 km^2 and 121 km^2, respectively. Spatially-variable precipitation fields of identical statistics were generated using a rainfall model (Rodriguez-Iturbe and Eagleson, 1987). A total of 50 spatially-variable rainfall fields were applied to each basin using a Monte-Carlo methodology. Defining and as rainfall duration and time to equilibrium, respectively, they found that spatial variability dominates at large scale when , while the temporal variability dominates at small scale when.

Ogden and Julien (1994) examined the sensitivity of a physically-based rainfall-runoff model to the spatial resolution of radar rainfall input data with consideration of length scales. Based on a 2 hour convective rainfall event recorded by the NEXRAD radar, a set of 50 and 40 independent simulations at 1, 2, 3, 4, 6, 8 km resolution were analysed on two semi-arid watersheds of 32 and 121 km^2 (model grid size of 125 and 200 m respectively). The results from simulations with the finest rainfall resolution (1km) were the best available outflow calculations and were used as reference outflow hydrographs. They defined two parameters:

$$\frac{L_R}{L_S} \tag{11.3}$$

which is an indication of ”storm smearing”.

$$\frac{L_R}{L_W} \tag{11.4}$$

which is an indication of ”watershed smearing”. L_S is the correlation length of the rainfall, L_R the grid size of the rainfall data, L_W the characteristic length of the watershed taken as the square root of the area. They considered the standard

deviation of peak discharge, runoff volume and rainfall volume versus the two parameters defined above. Results from simulations without infiltration showed storm smearing occurring as $L_R \rightarrow L_S$. Watershed smearing caused more significant deviations from simulations using the finest resolution data when L_R/L_W exceeded 0.4. The deviations are larger in the case of the smallest basin. The first set of simulations assumed impervious basins (no infiltration) and every storm produced runoff. In the case of impervious basins, for dry conditions, not all rainfall subsets generated runoff (11 and 18 events out of 40 and 50 storms for the 32 and 121 km^2 basins, respectively). Results with infiltration revealed that excess rainfall volumes decrease with increasing L_R/L_W, i.e. due solely to rainfall data aggregation.

11.3.2 Medium Size Catchments (100-2000 km^2)

Michaud and Sorooshian (1994) examined the effect of rainfall sampling errors on simulation of flash floods using a distributed model on the 150 km^2 Walnut Gulch watershed. They studied 10 storms triggering large runoff events, using rainfall fields based on kriging interpolation from 58 raingauges, and a subset of 8 raingauges. Rainfall-sampling errors and their effects on hydrologic simulations were inferred by comparing the peak runoff from the eight-gauge and areal cases (defined below) to the peak runoff from the 58-gauge case. Total errors represent the sum of rainfall-sampling errors and errors between the peak runoff from the 58-gauge case to observed peak runoff. Errors are expressed as percentage bias of the mean observed peak. Results showed that inadequate raingauge density (1 gauge per 20 km^2 compared to 1 gauge per 2.5 km^2) gave total errors in simulated peak of 13.7% of the observed peak flow. Rainfall sampling errors accounted for approximately -10% of errors between observed and simulated peaks. Degradation of the rainfall field from the 58 raingauges and spatial averaging over a 4-km grid resulted in total errors of -24% of the observed peak. Rainfall sampling errors accounted for -47% of errors between observed and simulated peaks. The 4-km pixels were too coarse to preserve the small cells of high rainfall intensity. Although the watershed average rainfall volume was preserved, the distribution of rainfall intensity was modified in a manner that reduced Hortonian runoff. Decreasing the temporal resolution from 2 min to 1 hour also resulted in a marked reduction of simulated runoff. The authors concluded that at least 2-km spatial resolution is required for 50-500 km^2 semi-arid watersheds submitted to convective thunderstorms.

11.3.3 Large Catchments ($>$ 2000 km^2)

Arnaud et al. (2002) studied the sensitivity of distributed hydrological models to different patterns that represent the spatial distribution of rainfall: spatially averaged (i.e. synthetic hyetographs obtained by weighted average of the different observations) or rainfall field (hyetographs associated with different spatial observations). The rainfall data came from a network of 44 rain gauges that covers about 2000 km^2 around Mexico City. The 50 most significant events were selected for the period 1988-1998 and events were classified as extreme or more frequent events. The mean duration was approximately 10 hours and the maximum point depth reached 170 mm for a 126-hour storm with cumulative basin precipitation of 81 mm. Each event was

separated by at least 1 hour dry at all stations. Four fictitious basins ranging from 20 to 1500 km^2 were defined. The tests were undertaken for a distributed model allowing three different rainfall-runoff processes and different basin size to avoid conclusions based on runoff-generation mechanisms and catchment size. For the 50 events, the relative difference between the empirical distribution of the peak flow and runoff volumes was calculated according to the two patterns of rainfall input (UR for uniform basin rainfall and NUR for non uniform rainfall field).

They concluded that runoff volumes and peak flows can vary considerably between spatially uniform and spatially distributed rainfall patterns. Differences in flows range from 10 to 80% depending on the type of runoff production model used, the size of the basin and the return period of the event. The differences were generally moderate for extreme events (10 to 30%) and more significant for more frequent events (20 to 80%). A possible explanation is that highest flows are generally obtained for rain events that affect the whole basin and have lower spatial heterogeneity of cumulative rainfall. This seems to be a particular feature of Mexico City. Higher relative errors were observed for the largest basins in comparison to the smallest ones, which suggest that the runoff sensitivity to rainfall spatial variability increases with catchment scale, again a particular feature of Mexico City.

11.3.4 Conclusions

The studies conducted in semi-arid environments associated with localised convective rainstorms confirm the importance of spatial rainfall for runoff generation. Studies conducted on the Walnut Gulch watershed in Arizona across a range of size reveal that spatial rainfall is important at all scales and increases as the scale decreases. This is also confirmed by Ogden and Julien (1994) who observe larger deviations due to rainfall aggregation in the case of the smallest basin. It seems that the basin response is more sensitive to larger events.

A density of 4 raingauges is required to model adequately a 4.4 ha catchment (Faurès et al., 1995), at least 5 raingauges are necessary at the 6.7 km^2 scale (Lopes, 1996). Michaud and Sorooshian (1994) use rainfall input from 58 raingauges at the 150 km^2 scale and suggest 2 km resolution to model 50-500 km^2 watersheds (Michaud and Sorooshian, 1994). Ogden and Julien (1994) recommend a spatial resolution less than 0.4 the square root of the watershed area. (i.e., 1 km resolution for a 10 km^2 watershed, 4 km resolution for a 100 km^2 watershed). The basin response is sensitive to the magnitude of events (Michaud and Sorooshian, 1994; Syed et al., 2003).

Conversely, in the study conducted at the large catchment scale in Mexico City, higher differences in the flow were observed for more frequent events compared to extreme events. Higher relative errors were observed for the largest basins in comparison to the smallest ones, which suggest that the sensitivity of the runoff on spatial variability of rainfall increases with catchment size. This seems to be a particular feature of Mexico City and may highlight the complexity introduced when mixed weather and catchment conditions are considered

11.4 Temperate Climate

In the temperate mid-latitudes, frontal low-pressure systems dominate in winter, and rainfall tends to be higher. During the summer, the subtropical high-pressure anticyclones expand to influence the weather patterns. Consequently, precipitation is somewhat lower, and most that falls is convective in nature, arising from the uplift air during elevated daytime heating and the generation of heavy short-lived showers and thunderstorms (EAE, 2005).

11.4.1 Small Catchments ($<$ 100 km^2)

Shah et al. (1996b) developed a stochastic rainfall field model and studied the effects of spatial variability of rainfall on catchment response on the 10.5 km^2 Upper Wye catchment, UK, using the Systme Hydrologique Europeen (SHE), a physically based, distributed rainfall-runoff model. Conditioned on ten storms from three raingauges, the rainfall model, which uses the Turning Band Method and a fractionally differenced line process, generates Gaussian random fields with a specified space-time correlation structure. The model has 7 parameters, the rate of decay of the space-time correlation function of the Gaussian field which is assumed isotropic and defined by the lag-one correlation space-time coefficient, the mean and standard deviation of the transformed Gaussian field, two parameters controlling the duration and the spread of the storm, and two components for the storm velocity (see details Shah et al. (1996a)).

Antecedent soil moisture conditions are described by the depth of the phreatic surface component of the rainfall-runoff model. Provided that there is at least one gauge available, results revealed that under ”wet” conditions, good prediction of runoff could be obtained with a spatially averaged rainfall input. However, for ”dry” conditions, the errors were considerably higher if spatially averaged rainfall is used. In comparison to the reference flow (driven by the full rainfall field, i.e. an equivalent of 169 raingauges), the mean absolute error in peak flow and volume were of the order of 6 and 3% for wet conditions and 14 and 8% for dry conditions. They related this to the interaction between the spatial variability in rainfall and the spatial distribution of soil moisture. Results also showed the nonlinearity of the runoff response with higher errors in runoff compared to errors in rainfall depth. By decreasing the lag-one space-time correlation parameter of the model (i.e. assuming more spatially variable rainfall regime or larger catchment areas with the same number of raingauges), they observed that the errors increase with decreasing correlation and are again larger for dry conditions. They also found that larger errors in runoff prediction arise when applying a lumped model to the catchment compared to averaging the rainfall (the R^2 statistics decreased from 0.88 for wet conditions to 0.66 for dry conditions).

Corral et al. (2000) presented a semi-distributed version of TOPMODEL and applied it to Ample, a mediterranean 48 km^2 basin in France. 30 min radar rainfall estimates averaged on 1 km^2 grid served as input to the hydrological model at the same spatial resolution. Based on a set of 13 events, they compared the performance of the semi-distributed model with two lumped models: SCS and the original TOPMODEL run using spatially uniform rainfall. They found that the semi-distributed model, which is able to incorporate the spatial variability of rainfall, outperformed the lumped models

in terms of reproduction of the peak flow, recession curve and Nash-Sutcliffe efficiency (NSE). The NSE for the semi-distributed model was 0.90 in optimisation mode compared to 0.80 for the lumped models, in validation mode a NSE of 0.72 was compared to 0.55. However it is not possible to distinguish the effect of spatial rainfall from the effect of the model discretisation.

Obled et al. (1994) used a semi-distributed rainfall-runoff model on a 71 km^2 rural catchment in the south-east of France containing a network of 21 rain gauges. They studied the sensitivity of this hydrological model to spatial rainfall patterns. Three spatial representations of rainfall were tested: spatially uniform rainfall from a network of 5 (PU5) and 21 (PU21) rain gauges, and a spatially variable rainfall field based on the 21 (PS21) rain gauges, distributed over subcatchment areas of 6-8 km^2. Calibration and validation of the model was undertaken, based on a set of 19 events, for each rainfall representation. The NSE increased from 0.76 for PU5 to 0.80 for PU21, suggesting that a better estimate of the volume of rainfall (a difference of 3% was observed) improves the model performance. The results also showed that the optimised parameters vary quite strongly according to the type of uniform rainfall used. However when using input from the spatially variable rainfall PS21, the NSE was 0.78 compared to 0.76 with PU5 and shows lower performance than PU21. This is because the main peaks were better represented than with PU5 but the model also responded to a rainfall event that the catchment ignored or dampened. This evidenced the sensitivity of the particular rainfall-runoff model rather than the sensitivity of the basin to spatial rainfall. These secondary peaks were not dampened enough by the isochrone routing of the model to fit the observed discharge. With the isochrone assumption, the routing of the runoff generated over contributing areas has been ignored and all the surface water is immediately available to the closest channel reach. An extension of the model was implemented to allow for a specific hillslope unit hydrograph. It preserved the isochrone concept but allowed some spreading in time by assuming that a transfer function, working at the land-unit scale, routes the water to the river. This was defined in terms of a parameterised Gamma function to route the surface water down to the channel only. Compared to the previous case, the NSE increased from 0.78 to 0.82. Hence the results with distributed information remain similar to the case of PU21. Further minor improvement from PU5 to PU21 was observed, which suggests that a spatial average of 5 raingauges is appropriate to model this 71km^2 catchment. They concluded that the spatial variability of rainfall, although important, was not sufficiently organised in time and space to overcome the effects of smoothing and dampening of the catchment. The authors reasoned that the dampening behaviour may be due to the mechanism of runoff generation. In this case, it is of the Dunne type where most of the water infiltrates and local variations in input are smoothed as the water is stored and delayed within the soil. They suggest that the catchment is characterised by small contributing areas which makes the catchment less sensitive to rainfall patterns. Woods and Sivapalan (1999) reached similar conclusions in a study conducted in New-Zealand (see Section 4.2).

11.4.2 Medium to Large Catchments (100-2000 km^2)

A number of errors can be observed in radar measurements and long records of continuous data are still lacking. Nevertheless, provided that a simple correction is

applied, radar data represent an important source of information, which allows the continuous spatial distribution to be studied. An example of this is presented below.

Borga (2002) analysed the impact of errors in radar rainfall estimates on rainfall-runoff modelling using a lumped, continuous simulation, model on the 135 km^2 Brue catchment in Southwest England. Adjustment of the radar estimates include correction for range-related bias due to vertical profile of reflectivity (VPR) (Cluckie et al., 2000); orographic enhancement of precipitation (Kitchen et al., 1994); anomalous propagation; uncertainty in Z-R calibration (Illingworth et al., 2000); mean field bias adjustment against raingauges (Borga et al., 2002).

5 minute radar rainfall estimates were available at 4 elevation scans, together with a high density network of 49 raingauges, for a 2.5 year period. Spatially-uniform areal rainfall, given by an average of the raingauges, was used for the calibration (Dec 97 -July 99) and validation (Nov 93 - Aug 96) of the model. Good agreement in terms of Mean Relative Error (MRE) and Nash-Sutcliffe efficiency (NSE) between the calibration (0.043 and 0.82 for MRE and NSE, respectively) and validation (0.054 and 0.83) period were obtained. Next, the model was run over the validation period using both unadjusted and adjusted radar rainfall estimates, and estimates from the four elevation scans.

Results showed that adjustment of the radar rainfall significantly improved the model performance. In comparison to the reference flow driven by the raingauges, the NSE for the adjusted radar data ranged from 0.92 at 0.5° scan (0.82 for unadjusted data) to 0.84 at 2.5° scan (0.58 for unadjusted data). In comparison to the observed flow, the NSE ranged from 0.75 at 0.5° (0.69 for unadjusted data, 0.83 for the raingauge-driven simulation) to 0.5 at the higher elevation scan of 2.5°. The low value of NSE at the higher scan may be attributed to an effect of residual bright band remaining after VPR adjustment. It may also be expected that the use of radar rainfall estimates for calibration purposes could force model parameters to compensate for the biases and sampling characteristics of radar areal precipitation estimates, thus leading to different simulation results during validation.

Bell and Moore (2000) investigated the sensitivity of catchment runoff models, lumped and distributed, to rainfall at a variety of spatial scales using data from a dense raingauge network and weather radar on the 135 km^2 Brue catchment. 49 raingauges were available, together with radar data at 2 and 5 km grid resolution. The sensitivity of the catchment runoff models was evaluated for two representative periods of stratiform and convective rainfall events that triggered peak flows of about 15 m^3s^{-1} and 10 m^3s^{-1}, respectively.

The distributed model was calibrated twice, using catchment average rainfall from the raingauges and radar data. To determine the effect of rainfall variability on flow response at the catchment scale, rainfall estimates for each of the 28, 2 km pixel containing raingauges were used as alternative inputs to rainfall-runoff models to obtain an ensemble of 28 hydrographs. It was found that the distributed model is sensitive to the location within the catchment of the grid-square raingauge data. As expected, the greater spatial variability in rainfall for the predominantly convective period is reflected in the wide variation in modelled hydrographs. The generated peak flow ranged from 1 to 45 m^3s^{-1} for the convective event compared to 7-22 m^3s^{-1} for stratiform rainfall.

Sensitivity to rainfall (both radar and raingauge data) and distributed model resolution was explored at 2, 5 and 10 km grid resolution. Model recalibration returned the model performance to that obtained from catchment average rainfall data. Best performance was obtained using a lower resolution of rainfall data and model (5 and 10 km than for the 2 km scale). The improvement in performance of the distributed model over the lumped one was marginal in case of the stratiform event (R^2 of 0.922 compared to 0.911) but enhanced for the convective event (R^2 of 0.953 compared to 0.909).

Using raingauge data, sensitivity to rainfall resolutions at 2, 4, 6, 8, 10 and 16 km was explored with the 2 km grid distributed model. Results show that provided the model is recalibrated for each rainfall resolution, equivalent performance is observed for the convective event. However if the model is not recalibrated, the simulation accuracy decreases with the rainfall resolution and drops to 0.4 above 10 km. For the more spatially homogeneous stratiform event, there is little change in simulation accuracy with resolution of rainfall data.

In conclusion, the runoff response is sensitive to the type of rainfall event and the particular model used. The modelling of the Brue catchment benefits from using a distributed approach over a lumped one. Results show that recalibration of the model parameters compensate for bias in rainfall. Using a model discretisation of 2 km, the adequate modelling of the runoff response during convective events requires a rainfall resolution of 2 km.

Chaubey et al. (1999) used a distributed-model on a 148 km^2 watershed in Oklahoma and a network of 17 rain gauges, to simulate surface runoff, sediment nutrients and pesticide transport. The model was calibrated on five events using rainfall input obtained from the Thiessen Polygon method. The four calibrated model parameters served as reference. Applying one rain gauge at a time uniformly across the watershed, results showed that a large uncertainty in model parameters is introduced due to the spatial variability of rainfall. The smallest parameter uncertainty resulted from the rainfall that was the most homogeneous in nature. The correlation analysis between the parameters and the input rainfall revealed that for a particular event, the model parameters are highly correlated. One question that still needs to be answered is the number of raingauges required to estimate appropriately the model parameters.

In the context of real time flood forecasting, Maskey et al. (2004) designed a method based on fuzzy set theory combined with a genetic algorithm to propagate and assess rainfall uncertainty in rainfall-runoff modelling. It used the temporal disaggregation of the precipitation over a period into sub-periods and allowed the uncertainty in both the temporal and spatial distributions of the rainfall to be accounted for. The method consisted of division of the temporal period over which the accumulated rainfall is known into a fixed number of sub-periods and random disaggregation of the accumulated sum over sub-periods. The disaggregated rainfall sums up the accumulated rainfall. The disaggregated precipitation signal is then used as input to the rainfall-runoff routing model. Further details and implementation of the methodology can be found in Maskey et al. (2004). The results, based on 60 hours forecast on a 1744 km^2 catchment in Poland with 9 sub-catchments ranging in size from 64 to 280 km^2, showed that the output uncertainty due to the uncertain temporal distribution of precipitation dominates the uncertainty due to the magnitude of the precipitation (a maximum difference of 80% in forecast discharge is observed when introducing temporal uncertainty). Marginal differences were observed when varying

the temporal pattern over sub-basins, which accounts for the spatial variations of the rainfall field, and when disaggregating into 6 sub-periods instead of 3. Woods and Sivapalan (1999) proposed an analytical method to identify the importance of different components of the hydrological cycle during storm events in a humid temperate catchment of 420 km^2 in New Zealand. In particular, the relative importance of the spatial or temporal distribution of rainfall, and the interaction between the rainfall variability and the flow routing can be assessed. For a single event, they assume that the space-time variability in rainfall and runoff generation can be separated into independent multiplicative time and space components, according to:

$$P(x,y,t) = P_{x,y}(t)P_t(x,y) \tag{11.5}$$

where $P(x,y,t)$ is the space-time rainfall field, $P_{x,y}(t)$ is the time-series of spatially averaged rainfall rates for the study-area delimited by x and y, $P_t(x,y)$ is the dimensionless rainfall space pattern for the duration of the storm event. A similar expression for the runoff generation function is assumed.

Implicitly, in Equation 11.5, a spatially-uniform temporal profile conditioned on storm-averaged totals at each pixel is assumed, which may be a suitable approximation in practice (see Chapter 12) for predominantly frontal rainfall in humid areas. However this scheme may not be appropriate for rapidly moving storms, more spatially variable (and intermittent) rainfall in summer or larger catchments.

The model of rainfall, runoff generation, and routing can be summarised as follows: rain falls on the catchment and either becomes rainfall excess through the action of a runoff generation process or is stored. Rainfall excess is routed to the base of hillslopes (where it enters the channel) via a time-invariant hillslope unit hydrograph, which is assumed to be the same throughout the catchment. Hillslope discharges are then routed through the channel network to a river basin outlet (where they become catchment runoff) by a time-invariant channel network unit hydrograph with constant velocity. Only the runoff generation mechanism in this theory can provide a nonlinear hydrological response. This simple model has the advantage of accounting for the main interactions between rainfall and catchment properties. Apart from the storm movement, storm variability is directly incorporated. Catchment properties are included in the mechanism of runoff generation, the hillslope routing and the channel routing. The spatial organisation of rainfall is taken into account via the distribution of rainfall-excess with flow distance to the outlet. The effect of scale can be studied by applying the methodology at various scales.

To treat these interactions analytically, two sets of equations are derived and the method is illustrated on a 10-hour storm using hourly radar rainfall data. The first set deals with understanding controls on rainfall excess and runoff for a single catchment. For instance, expressions for the mean and variance of runoff duration can be derived. For this event, the mean catchment runoff time (i.e. a characteristic time of the catchment defined here as the lag time between the the start of the rainfall event and the centroid of the runoff hydrograph) is dominated by the storm duration and the flow distance, whereas the storm duration, temporal variability and flow distance govern the variance of runoff duration. In both cases spatial variability has little importance. The authors argue that the extreme values of rainfall excess occur at flow distances representing only a small fraction of the catchment area.

The second set of equations shows the effects of between-catchment variability on

variability between hydrographs. To test these expressions, sub-catchments were created by applying thresholds of 0.0625, 0.2, 0.625, 2, 6.25, 20 and 62.5 km^2. Results show that for small catchments, the mean runoff time is dominated by the storm duration, and for larger catchments the mean travel time in the network gradually becomes more important. The variance of runoff time is dominated by the duration of the rain event and the temporal variability of rainfall excess; for larger catchments the variability of network travel time exerts greater influence. Considering variability among catchment of similar size, results suggest, for both the mean and variance of runoff time, that differences in the temporal spread of rainfall excess dominate between-catchment variability for small catchments, and differences in the distribution of flow distance become more important. Although the example is purely illustrative, it seems that the apparently substantial spatial variations in rainfall fail to exert any significant influence on the timing of catchment response. For this event, at each flow distance, there is a wide range of rainfall depths so the distance-averaged rainfall excess is relatively uniform (see also Naden (1992); Obled et al. (1994)).

Carpenter et al. (2001) analysed the sensitivity of a distributed hydrologic model to change in parameter and radar rainfall input. Rainfall estimates came from the NEXRAD radar (4 km^2) and were applied to 4 basins ranging in size from 285 to 2483 km^2 in Oklahoma and Arkansas. Mean areal precipitation over 80 km^2 was used as input to the rainfall-runoff model for a set of 13 events. They allowed for a 50% change in precipitation and model parameters. Results show that the relative variability in flow increases as basin size decreases and the authors conclude that the sensitivity of flow statistics to parameters and rainfall-input is scale-dependent. For the larger sub-catchments ($>$800 km^2), the 5-95% uncertainty range in simulated flow is between 10-15% of the median of simulated flows for a 50% radar rainfall error magnitude. When independent parametric and radar rainfall uncertainties are combined, the flow uncertainties remain in the range of 25-40% of median flow for the larger basins. Increasing the scale of model spatial discretisation from 80 to 160 km^2 increases the range to 38-100% of median flow, while the median flow itself varied substantially. The analyses show that in 60% of the events studied, the spatially lumped model response driven by raingauge data was outside the 5-95% uncertainty boundaries of the flow simulated by the spatially-distributed model indicating that a distributed approach is preferred in this case. In conclusion, the modelling of the runoff response benefits from a distributed approach even though the sensitivity of flow statistics to parameters and rainfall-input is scale-dependent. It seems that for basins larger than 800 km^2 , a mean areal rainfall over 80 km^2 (i.e. 9 km grid) is adequate. For the smallest basin of 280 km^2, larger variability in flow is observed when uncertainties in the rainfall and parameters are introduced.

In the US, the Distributed Model Intercomparison Project (DMIP) was defined as a general comparison between 12 distributed rainfall-runoff models with reference to the Sacramento Soil Moisture Accounting Model (SAC-SMA), a lumped model used for operational river forecasting. Three main objectives were defined. First, to assess the benefit of using distributed modelling in an operational forecasting environment. Second, to investigate the dominant form of rainfall spatial variability and to refine the understanding of the effect of rainfall spatial variability on simulated basin outlet hydrograph; and finally, to evaluate the performance of using distributed model to produce hydrologic simulations at interior points (Smith et al., 2003).

7 years of hourly, 4 km resolution data from the Next Generation Radar (NEXRAD) were made available to each participant. The study was conducted on one small basin of 65 km2 and 7 medium to large basins ranging in size from 285 km^2 to 2485 km^2 in Oklahoma, Arkansas and Missouri. Comparison from DMIP simulations to observed streamflow and to simulation from the SAC-SMA model was based on 12 statistics (see Appendix A (Smith et al., 2003)). These include measures of relative improvement to evaluate the gain in simulation accuracy realised by using distributed versus lumped models and calibrated versus uncalibrated models. Uncalibrated models were derived with parameters that were estimated without the benefit of using the available discharge time-series. For statistical analysis, 16 to 24 events per basin were selected. Model performance was studied across a range of event sizes: all largest events, several moderately sized storms and few small storms.

Results (Reed et al., 2004) showed improvement in average model performance after calibration for distributed modelling. Although for the majority of catchment configurations, the lumped model performed better than distributed models, some distributed models showed comparable results to lumped models in many basins and clear improvements in one or more basins. Noteworthy improvements in predicting flood peaks were demonstrated in a basin distinguishable from other basins studied in its shape, orientation, and soil characteristics. Greater uncertainties in modelling the smallest basin were observed. We focused on two articles from the DMIP project which are introduced below.

Smith et al. (2004) observed that catchments characterised by marked spatial variability in precipitation, and less of a filtering effect of the rainfall input signal, present improved outlet simulations from distributed versus lumped models. Both a lumped and distributed model were used on three catchments in Oklahoma: Eldon (795 km^2), Blue (1233 km^2) and Watts (1645 km^2). They defined two indices to measure the spatial variability of rainfall for specific events. The locational index I_L describes the variability of rainfall in comparison with channel flow distance and investigates the spatial organisation of rainfall (i.e. where the heavy rainfall lies in the basin), the general variability index I_σ quantifies the intrastorm variability (including terms for the standard deviation of the hourly gridded radar rainfall estimate covering the entire basin). These indexes, in particular I_L , express similar concepts to those developed by Woods and Sivapalan. An index of outflow hydrograph variability is defined in terms of filtering performed on the input signal. They seek the use of wavelet transforms to derive a composite measure of rainfall signal dampening for an event and used the Mexican Hat wavelet to analyse the hourly streamflow. Consequently, an index R_d of outlet hydrograph variability in terms of the amount of filtering or dampening on the input rainfall signal was defined. Next, an analysis to assess the benefit of using distributed modelling was carried out by plotting the modified correlation coefficient R_{mod} statistics (correlation coefficient between the simulated and observed flow time-series weighted by the standard deviations of the simulated and observed hydrographs) against I_σ and I_L .

Results show that for the Blue basin, the distributed model on average provides improvements over the entire range of I_σ values (R_{mod} in [0-0.4]). The distributed model shows more gains as the I_L statistic varies from a value of 1 in both directions. The rainfall for the Blue is the most spatially variable with I_L in the range [0.5-13]. These results suggest that the spatial variability of rainfall is large enough so that

distributed modelling is necessary to improve the results. Besides, the Blue basin is characterised by low levels of filtering of the input signal with a narrow range of dampening ratio R_d in [2-25] (compared to [2-47] for the Watts). The results show that distributed modelling presents improvement for about half of the events for Eldon and only marginal improvement for Watts. Overall, distributed modelling does not provide improvements for the most dampened events. These events have I_L value around 1.1 and I_σ values in the range 3-5, indicating that the distance-averaged rainfall and the rainfall itself are fairly uniform. In this case, a lumped model performs well.

In conclusion, this article proposes indices which give inferences about whether spatial variability of rainfall overcomes the filtering of the basin so that distributed modelling is required. Knowledge of spatial rainfall is important when (a) the rainfall is variable, (b) the catchment is responsive and (c) the distance-averaged rainfall excess is also relatively variable. This approach is analogous to the analytical framework proposed by Woods and Sivapalan (1999), in which the relative importance of the spatial and temporal distribution of rainfall and the interaction between the rainfall variability and the flow routing can be assessed. It would be interesting to apply the methodology of Woods and Sivapalan (1999) and check whether the space-time covariance between rainfall and runoff generation is negligible in the Watts basin and significant in the Blue basin.

Ajami et al. (2004) compared lumped and semi-distributed versions of the SAC-SMA model based on NEXRAD radar data, also for the 1645 km^2 Watts basin in Arkansas. For the semi-distributed model, they developed a multi-step automatic calibration scheme using the Shuffled Complex Evolution optimization algorithm (Duan et al., 1992) for model calibration and use rainfall averaging over 9 sub-basins between 20 and 500 km^2. For this homogeneous basin, analysis of the overall performance at the outlet showed that there is no improvement when moving from the manually calibrated lumped model (DMIP reference) to a semi-distributed structure. NSE is 0.77 and 0.91 for the reference lumped model in calibration and validation, respectively, compared to 0.68 and 0.78 for the semi-distributed model. A closer inspection indicates a better performance for the distributed model during specific periods, especially during the summer and early fall, when the basin response is dominated by baseflow and also provides better simulations for high flows. In terms of rainfall spatial variability, this suggests that knowledge of spatial rainfall during convective events in summer and extreme events is required in order to improve the model performance. Also this indicates that the usual measures of efficiency (such as NSE and R^2), while giving a general indication of the model performance, may not be able to capture the details of the fit.

Next the results were compared for three model calibration strategies defined as follows: lumped, semi-lumped and semi-distributed. The calibration results showed that moving from a lumped calibration strategy to a semi-distributed approach while keeping the model parameters constant over each sub-catchment (semi-lumped), improved the simulation results. However, varying the parameters between sub-basins (semi-distributed) did not further improve the simulations results, either at the outlet or at an interior point. Hence when a semi-distributed model structure is adopted, information of spatial rainfall (averaged over each sub-basin) improves the model performance.

11.4.3 Large Catchments ($>$ 2000 km^2)

Andréassian et al. (2001) analysed the link between the performance of a rainfall-runoff model and the representativeness of the areal rainfall estimate used to run the model. To this end, they defined two indexes to compare the rainfall estimates given by different raingauge combinations to the reference rainfall (defined below). The GORE index compares the sum of squared errors in the rainfall estimate to the temporal variance of the reference areal precipitation. The BALANCE index assesses the quality of rainfall total depth by quantifying the overestimation or underestimation of the reference rainfall by a given raingauge subset.

These indices were compared with model efficiency (Nash-Sutcliffe criterion NSE) and model parameters for three French catchments of increasing size 71 (same catchment studied by Obled et al. (1994)), 1120 and 10700 km^2, and different climate conditions (mediterranean climate for the smallest basin and temperate oceanic climate for the two larger catchments); and three continuous, lumped, daily rainfall-runoff models. The reference areal rainfall estimate was taken as the spatial average of all (or the best positioned) rain gauges. 33, 8 and 20 rain gauges were used for reference areal rainfall estimation for the 10700, 1120 and 71 km^2 catchments, respectively. The sensitivity of model performance to the rain gauge network quality was assessed by plotting the NSE criteria against the GORE or BALANCE index. Results show that there is an overall increase in model efficiency with the GORE index. For all three watersheds, GORE values above 0.9 lead to maximum NSE criteria above 80%. Similarly, the model performance is maximum (NSE $>$80%) for BALANCE values between 0.9 and 1.1, i.e. when the estimate of total precipitation is within 10% the reference value. However this feature is necessary but not sufficient to achieve good performance (cases where the balance index is 1 and the corresponding NSE value is around 50%). The models proved to be robust and reliable due to their ability to cope with imperfect rainfall input estimates. For instance, a single raingauge for the 10700 km^2 basin, with a GORE index of 0.74, achieves a NSE criterion of 81%. Bias in basin rainfall estimates may, at least partially, be compensated by model parameter adjustment. However, the conclusions may be affected by the lumped nature of the model, which is less sensitive to the spatial variability of rainfall. Further the three models showed improved performance and reduced variability of efficiency with more accurate rainfall input. The smaller watersheds require more precise areal rainfall estimates (a higher concentration of rain gauges) to ensure good modelling results. Poor modelling results are obtained when a single gauge is used to model the 71 km^2 watershed even though it provides a greater density per unit area than the 33 and 8 raingauges used for the 10700 and the 1120 km^2 basins. In this study, the catchment is considered as a lumped unit even at the 10700 km^2 scale. This article gives an indication of the raingauge network density required to model the streamflow at the outlet for a range of scale. They used 20 raingauges at the 71 km^2 scale, it seems that 8 and 33 raingauges are suitable to model the 1120 km^2 (i.e. 1 gauge per 140 km^2), and 10700 km^2 (i.e. 1 gauge per 324 km^2) catchment, respectively. Combined with the study by Obled et al. (1994) a density of five raingauges may be appropriate to model the 71 km^2 catchment. Overall the findings suggest that the importance of spatial rainfall decreases with scale. As in Smith et al. (2004), the rainfall variability is describes in terms of two indexes accounting for spatial and temporal variability.

In her study of the Thames catchment, Naden (1992) observed that the slow component of the hillslope response was dominant over the channel routing response and masked any impact of the spatial variability of rainfall. It seemed that the hillslope dominance was due to the large proportion of limestone and chalk in this catchment. Runoff moving through the catchment had a much longer residence time in the hillslope component than in the channel system. Naden (1992) obtained very similar catchment response functions for seven events despite the observed spatial variability in precipitation. Hence, she concluded that a semi-distributed modeling approach over a lumped model was not justified when applied to this 7000 km^2 basin. However her methodology that incorporates the spatial distribution of rainfall in the form of a weighted network with function into a unit hydrograph type flood estimation procedure showed improved performance compare to the traditional unit hydrograph method used in the UK Flood Studies Report (NERC, 1975). In comparison with the observed flow, an efficiency of fit of about 80% on average is obtained, compared with 60% with the traditional method. The weighted width functions could be expressed in terms of the distribution of (relative) rainfall excess with flow distance and in terms of the probability density function for flow distance in the channel network to the catchment outlet, two components developed by Woods and Sivapalan (1999).

(Koren et al., 1999) examined the scale dependencies of hydrologic models to spatial variability of precipitation using four lumped rainfall-runoff models. The 4km grid, hourly rainfall estimates for the three year period May 93 - July 96 from the NEXRAD radar were used in this study. The test area was the 256×256 km Arkansas-Red river basin. Results from different resolutions of spatial averaging of hourly data (4, 8, 16, 32, 64, 128 and 256 km grid) over sub-basins defined by the grid scale, revealed that all models were scale-dependent but the level of dependency varied with the type of rainfall-runoff partitioning mechanisms. It was found that each model produced less surface runoff with increasing scale. Surface runoff is gradually reduced within the range of 20% (relative to cumulated surface runoff at the finest scale) up to the scale of 32 km. At higher scales, there is a faster reduction in surface runoff with a wide range of variability from one model to the other (between 25 and 100% depending on the model). They deduced that the main factor governing runoff reduction at the smaller scales is rainfall variability over a rainfall averaging area. At the largest scales, the rainfall coverage becomes the major factor controlling runoff reduction. In this case, the event covered about 50% of the area at the largest scale. Total runoff is less sensitive to scale than surface runoff (reduction at largest scale is 10-25%). Defining the runoff coefficient as the ratio of accumulated runoff to accumulated precipitation, it is observed that the runoff coefficient is rather stable at the beginning of the event when initial loss are satisfied. Once initial losses have been satisfied, the runoff coefficient varies significantly over different scales with the highest values at the finest scale. Hence, when assessing the runoff response to uncertainty in precipitation, bias is introduced due to the runoff partition mechanism. Infiltration-excess type models were the most sensitive whereas saturation excess models were less scale-dependent. They observed that scale dependency was reduced when the spatial variability of rainfall was taken into account in the rainfall-runoff model structure. This paper shows that there is a need to improve the hydrological lumped model structure.

In the context of nonlinear dynamical systems (Wiggins, 1999) applied to nonparametric streamflow prediction (Porporato and Ridolfi, 2001), Dodov and

Foufoula-Georgiou (2005) propose a methodology which combines the spatial and temporal distribution of rainfall and basin geomorphology into nonlinear analyses of streamflow dynamics. The methodology is based on optimal phase-space reconstruction of the coupled rainfall-runoff process. The (multi-variate) time-series of rainfall and the corresponding (univariate) time-series of streamflow are independently reconstructed in their optimal phase-space by using Takens (1981) delay time method so that two state vectors or rainfall and runoff states are obtained.

A brief summary of the methodology is given here, further details can be found in Dodov and Foufoula-Georgiou (2005). Catchment geomorphology is accounted for by two criteria, firstly by spatially aggregating the rainfall information according to the (sub)catchment topology. Various aggregation levels of rainfall have been investigated: pixel size, catchment scale, and sub-catchment scale. A geomorphologically derived response function (Unit Hydrograph) is used to assign weights to the elements of the rainfall state-vector according to their contribution to the discharge. The streamflow of a reference state can be predicted by coupling the rainfall and runoff information in a distance space. In this distance space, the reference state represents the origin, the x-axis corresponds to streamflow distances (i.e. the Euclidean distance between the reference streamflow and all other observed streamflow states) and the y-axis corresponds to rainfall distances. The similarity between the reference rainfall and all other rainfall states is evaluated by means of root mean square error, multiscale root mean square error and Hausdorff distance (see Appendix A-C in Dodov and Foufoula-Georgiou (2005) at pixel level, weighted (using the geomorphologically derived response function mentioned above) and non-weighted Euclidean distances otherwise. First, the rainfall distances are plotted against the corresponding streamflow distances. Then a kernel function is defined so that it captures a neighbourhood of points around the reference state. Finally, the average of the points within the neighbourhood gives the streamflow prediction at the time horizon defined.

The methodology is applied to two basins in Missouri and Oklahoma with contrasting geology: deep limestone aquifer underlying the Gasconade River basin (7352.4 km^2) and in absence of dominating aquifer in the case of the Chikaskia River basin (4813.7 km^2). This study was based on 6-hour streamflow time-series and radar data available for a period of 5 years.

Here is a summary of the findings:

- Including information about rainfall in the phase-space approach improves the streamflow prediction by up to 18% for the Chiskakia basin and 6% for the Gasconade catchment, compared to prediction based on streamflow only.
- Including information of rainfall spatial distribution increases the prediction accuracy (up to 12% and 5% for the Chikaskia and Gasconade basins, respectively), compared to the results based on averaged rainfall over the catchment.
- Best performances are obtained using a crude catchment discretisation of 3 or 10 subcatchments. Despite the damping effect of the catchment and the uncertainty in the data, the results suggest that rainfall averaging at the subcatchment scale captures enough information on rainfall spatial and temporal variability to predict the runoff response.

- Including information about (sub)catchment's response based on a geomorphologically-derived unit hydrograph, additionally improves the accuracy of streamflow prediction by 2-5%.
- The difference in prediction accuracy due to the information about the spatial distribution of rainfall depends on the geological characteristics of the basin. More improvement is achieved in the case of the Chikaskia River (10-15%), while in the case of the Gasconade River the prediction accuracy is significantly affected by the smoothing effect of the underlying limestone aquifer (improvement of 3-8%).

11.4.4 Conclusions

Contrasting observations arise from the studies conducted in temperate areas. The catchment response is a trade-off between the impact of spatial variability of rainfall and the smoothing effect due to the heterogeneity of the catchment. The importance of spatial rainfall depends on how variable the rainfall is and whether there is enough variability to overcome the damping and filtering effect of the basins (Smith et al., 2004; Obled et al., 1994).

The rainfall variability depends on the type of event. Bell and Moore (2000) observed lower variability in rainfall and runoff response during a stratiform event as opposed to a convective event in a study on the 135 km^2 Brue catchment in the UK. On the 1645 km^2 Watts basin in North America, Ajami et al. (2004) concluded that knowledge of spatial rainfall is important when modelling extreme events and convective events in summer.

When there is not enough variability in rainfall to overcome the damping effect of the catchment, detailed knowledge of rainfall spatial variability is not required to model the catchment response. Rather reliable information of catchment averaged rainfall is important (Woods and Sivapalan, 1999; Smith et al., 2004; Andréassian et al., 2001; Naden, 1992). For instance, Smith et al. (2004) observed that improvement in model efficiency using a distributed model is observed when spatial variability of rainfall is included to model the 795 km^2 Blue catchment whereas a lumped approach using catchment-averaged rainfall is suitable to model the 1645 km^2 Watts basin. The Blue basin is characterised by variable precipitation and low level of basin filtering whereas the Watts show higher level of filtering and more uniform precipitation.

The flow routing and the distribution of rainfall-excess with flow distance control the catchment response. Woods and Sivapalan (1999) proposed an analytical method to identify the importance of different components of the hydrological cycle during storm events in a 420 km^2 catchment in New Zealand. The method is illustrated on a 10-hour storm using hourly radar rainfall data at 1 km grid resolution and a lumped rainfall-runoff model. They conclude that if the space and time covariance between rainfall and runoff generation are negligible then methods for estimating catchment average rainfall and runoff generation may be more important than methods for estimating their variability. The apparently substantial spatial variations in rainfall do not seem to have any significant influence on the timing of catchment response. For the small sub-catchments the space pattern of excess rainfall is unimportant because it only influenced channel routing, and the timescale of channel routing in those

catchments is small compared to storm duration. For larger catchments, spatial variability of rainfall has little influence because rainfall does not vary significantly with flow distance in spite of its spatial variability in (x,y) coordinates. For this event, at each flow distance, there is a wide range of rainfall depths so the distance-averaged rainfall excess is relatively uniform.

Obled et al. (1994) reached similar conclusions. In a study conducted on a 71 km^2 catchment in France, they found that a spatially-uniform rainfall from five raingauges was enough to estimate the streamflow hydrograph because of the large dampening behaviour of the basin. In this area, the runoff generation mechanism is predominantly of the Dunne type, the water infiltrates and local variation of the rainfall input is smoothed and delayed within the soil.

The underlying geology is important and can mask the impact of spatial rainfall as in the case of Naden (1992) where the runoff had a much longer residence time in the hillslope component due to the large proportion of limestone and chalk in the 7000 km^2 Thames catchment. She concludes that a lumped approach is appropriate to model the 7000 km^2 Thames catchment. In the context of nonlinear analyses of streamflow dynamics, Dodov and Foufoula-Georgiou (2005) find that including information of rainfall spatial distribution improves the prediction performance of streamflow but the improvement is less in the case of a 7300 km^2 catchment in central North America with presence of limestone aquifer (3-8%).

The antecedent catchment condition controls the catchment response to the spatial variability of rainfall. Based on results from a 10 km^2 catchment in the UK, Shah et al. (1996b) observed that under dry conditions, higher error prediction of runoff (14 and 8% in peak flow and volume) are obtained with a spatially averaged rainfall input compared to the case with wet conditions (6 and 3%).

The importance of spatial rainfall on runoff generation is also a function of catchment scale. More precise areal rainfall estimates are required at small scales (Shah et al., 1996b; Andréassian et al., 2001; Carpenter et al., 2001). Obled et al. (1994) recommend 5 raingauges for a 71 km^2 basin in the South of France. The study from Bell and Moore (2000) suggest a 2 km grid rainfall resolution to model the 132 km^2 Brue catchment. As the scale increases, in the case where knowledge of spatial rainfall improves the runoff prediction, a crude resolution of rainfall is adequate. Smith et al. (2004) use 4km rainfall grid resolution to model a 800 km^2 basin. Carpenter et al. (2001) suggest a rainfall resolution of 80 km^2 (i.e. about 9km grid) for catchment scales greater than 800 km^2 and recommend a smaller one for the 280 km^2 catchment. At the larger scale, Dodov and Foufoula-Georgiou (2005) suggest that aggregation of rainfall at the sub-catchment scale captures enough spatial information that improves the prediction performance of streamflow by about 10-15% and 3-8% for two catchments of 4800 and 7300 km^2, respectively. They use catchment-averaged rainfall over subcatchments of 530 km^2 (9 subdivisions scheme) and 1600 km^2 (3 subdivisions scheme) for the 4800 km^2 basin, and subcatchments of 735 km^2 (10 subdivisions scheme) and 2450 km^2 (3 subdivisions scheme) for the 7300 km^2 basin.

Further complexity in unscrambling the effect of spatial rainfall on runoff is added when using a rainfall-runoff model. Obled et al. (1994) observed the sensitivity of a particular model to the spatial variability of precipitation, and not the sensitivity of the actual basin. Koren et al. (1999) concluded that lumped hydrologic models are

scale-dependent and that scale-dependency decreases when spatial variation of rainfall is accounted for in the model structure. Using semi-distributed model Chaubey et al. (1999) showed that large uncertainties in model parameters are introduced due to the spatial variability of rainfall. Shah et al. (1996b) found that larger errors in runoff prediction arise when a lumped instead of a distributed model is applied on the catchment, compared to averaging the rainfall. One reason for moving from the lumped to the distributed approach is the potential to simulate streamflow at interior points of the catchment.

To conclude, the runoff response is the result of a complex interaction between the rainfall spatial variability and the heterogeneous catchment properties such as the flow routing, the soil moisture and the geology. The spatial organisation of rainfall (i.e. the distribution of rainfall excess with flow distance) also controls the runoff response. The importance of spatial rainfall decreases with catchment scale. The relative importance of spatial rainfall is dependent upon the modelling strategy, hence the impact of the model should be quantified and kept to a minimum. For instance a small catchment discretisation is preferable to avoid smearing of rainfall information due to spatial averaging. The impact or rainfall errors only can be assessed by defining a reference rainfall and comparing the errors between the reference flow and the flow triggered by an alternative rainfall scenario. Modelling errors can be estimated by computing the difference between the reference and the observed flow. Appropriate calibration is required to keep them to a minimum.

11.5 Conclusions and Future Work

A review of the literature on the importance of spatial rainfall for runoff generation has been conducted for three conditions: (a) urban areas where most rainfall events contribute to runoff, (b) semi-arid areas associated with localised thunderstorms and (c) temperate regions where storms are the results of frontal and convective activities, or a combination of both, and the catchment is more heterogeneous in nature.

Across all catchment types, the review stresses the nonlinearity of the runoff response to rainfall input (Lopes, 1996; Shah et al., 1996b) and highlights three main interactions: (a) rainfall, (b) catchment properties and (c) scale. The effect of rainfall includes variability in space and time, storm movement and spatial organisation. The responding or damping effect of the basin is an intrinsic property of the catchment and can be characterised by the lag time (sometimes referred to as the time of concentration, characteristic time scale, mean runoff time, catchment response time). It is a measure of the catchment response time and expresses the average travel time from all points on the catchment to the outlet (Lekkas and Onof, 2005). As the scale increases, the importance of spatial rainfall decreases and there is a transfer from spatial variability of rainfall to catchment response time as the dominant factor governing the runoff generation. Hence these three interactions control principally the runoff response; however their relative contributions differ depending on the type of catchment considered.

Small urban catchments are fast-responding and sensitive to storm movement, storm variability and event magnitude. Their hydrological modelling requires fine spatial and temporal rainfall resolution of the order of minutes and kilometers. Since the damage

as a result of flooding can be important in urban areas, there is a strong incentive to use radar data or develop disaggregation techniques (Onof et al., 2005) to meet the requirement of urban hydrological modelling. Berne et al. (2004) show that finer spatial and temporal resolutions are required at smaller scale.

In semi-arid areas dominated by convective thunderstorms, the catchment response is also very rapid and a detailed representation of the rainfall distribution is necessary to reproduce the outflow accurately. The findings from this review highlight the importance of spatial rainfall at the small scale and suggest a spatial resolution of 1 km below 50 km^2, 2 km between 50 and 500 km^2. This is a hard recommendation to implement in practice as common operational networks are usually not as extensively instrumented as the catchments reviewed (Reefsgaard and Knudsen, 1996).

Contrasting observations arise from the studies in sub-humid areas where more complex interactions between spatial rainfall and catchment characteristics exist. A key paper is presented by (Woods and Sivapalan, 1999). Their analytical framework includes the main components of the hydrological cycle and quantifies the effects of rainfall and catchment response. The distribution of rainfall-excess with flow distance to the outlet was found to be a dominant factor influencing the runoff response. This finding was also incorporated by other authors (Smith et al., 2004; Naden, 1992; Dodov and Foufoula-Georgiou, 2005). In the event studied, although substantial variation in rainfall distribution was observed, the distribution of rainfall-excess with flow distance was relatively uniform so that knowledge of catchment-averaged rainfall was sufficient to model the runoff response. Obled et al. (1994); Naden (1992) and Smith et al. (2004) reached similar conclusions. Woods and Sivapalan (1999) conclude that the mean runoff time is dominated by storm duration at small scale and for larger catchments the mean travel time gradually becomes more important. Smith et al. (2004) suggest that when catchments are characterised by marked spatial variability in precipitation and less of a filtering effect of the rainfall input signal, they present improved outlet simulations from distributed versus lumped models. In this case, spatially-variable rainfall input is required but a detailed spatial representation is not necessary. The raingauge network density proposed by Obled et al. (1994) and Andréassian et al. (2001) seems adequate to model the behaviour of humid catchments, i.e. 5, 8 and 33 raingauges at scale of 71, 1120 and 10700 km^2. Daily data are adequate for lumped modelling, when distributed (or semi-distributed) modelling is undertaken, subdaily data are recommended.

In rainfall-runoff modelling, a simulated hydrograph is derived from a transformation of the rainfall input via the use of parameters that are the results of a calibration procedure at a fixed spatial scale. The modelling strategy adds another level of complexity and can mask the effect of spatial rainfall, therefore its impact should be quantified independently from the effect of rainfall (see discussion in Section 11.4.4). The main advantage of watershed modelling is that it provides a framework to use rainfall patterns to test hypothesis about the behaviour of the catchment.

This review has identified a number of open problems in the importance of spatial rainfall for runoff modelling in temperate humid areas. It is not clear whether an explicit rainfall representation is required. The findings are usually subjective and there is a need to quantify the effect of rainfall, catchment response and catchment scale independently. It is also difficult to distinguish between the effect of spatial rainfall and the impact of the model spatial discretisation.

In Chapter 12, an investigation on the Lee catchment, in the Thames region is undertaken in the context of flood management in the UK, to explore some of the issues highlighted above. This analysis has three goals: (a) to test whether spatial rainfall is important to model medium to large scale catchment in humid areas, (b) assuming that it is important, to investigate the rainfall spatial resolution required at catchment scale, for a range of scales and catchment types (c) to identify the modelling needs.

The requirements of a rainfall-runoff model for this analysis are as follows: (a) a realistic model incorporating nonlinear routing, (b) the capability of accepting a large number of distributed rainfall input data to assess the importance of rainfall representation, (c) an event-based structure to analyse the impact of the type of rainfall event, (d) a small number of parameters to facilitate parameter identification of catchment response.

The event-based, semi-distributed hydrological model RORB is selected to carry out the investigation. The study is based on 15 years of radar data together with 16 supporting raingauges and 12 flow stations from the 1000 km^2 Upper Lee catchment, UK.

The experimental design is constructed so that spatial rainfall is tested according to rainfall representation and the type of rainfall events. To address this issue, selected events from the period 1987-2002 from 2km grid radar data, subdaily data from a network of 16 raingauges and subsets of the raingauges are used. In the case where only daily data are available, a spatially uniform temporal disaggregation scheme (see Chapter 8) conditioned on observed daily data is also tested. The aim is to quantify the effect of rainfall using an appropriate measure of rainfall spatial variability among rainfall representation.

The Lee is divided into six subcatchments, representing the main tributaries. Each subcatchment is first modelled independently and the simulated runoff is then routed into the total runoff generated at the watershed outlet. Hence the relationship between the spatial structure of observed rainfall and runoff production for a range of catchments can be analysed. To investigate the sensitivity to catchment type, the existing (sub)catchments are then artificially turned into fast-responding urban basins. The catchment response time is taken as an independent variable of catchment response.

This modelling strategy allows analysing the spatial structure of catchment-scale rainfall at scales ranging from 80 to 1000 km^2.

This investigation should provide specific guidance concerning the importance of spatial rainfall for flood estimation and the associated data and modelling needs.

Chapter 12 The significance of spatial rainfall representation for flood runoff estimation: a numerical evaluation on the Lee catchment

12.1 Introduction

The review in Chapter 11 has identified a number of open problems concerning the importance of spatial rainfall for runoff modelling in temperate humid areas. The literature is inconclusive and there is a need to quantify the effects of rainfall type, catchment response and catchment scale independently. It is also difficult to distinguish between the effect of spatial rainfall and the impact of the model spatial discretisation. In this chapter, an investigation based on the Lee catchment in the Thames region is undertaken, in the context of flood management in the UK, to explore some of the issues highlighted above. This analysis has three goals: (a) to test whether spatial rainfall is important to model medium to large scale catchments in humid areas, (b) assuming that it is important, to investigate the rainfall spatial resolution required at catchment scale, for a range of scales and catchment types, (c) to identify the modelling needs.

Event-based, semi-distributed hydrological rainfall-runoff modelling of the Upper Lee catchment is undertaken to analyse the spatial structure of catchment-scale rainfall at scales ranging from 80 to 1000 km^2. Spatial interpolation of the point raingauge data is used to calibrate the model. Events are selected to represent extreme and more typical frontal and convective rainfall. Simulated runoff, driven by radar data and subsets of the raingauges, is compared with the reference streamflow generated from spatially variable rainfall, as defined by the full raingauge network. To investigate the sensitivity to catchment type, the existing (sub)catchments are then artificially turned into fast-responding urban basins. Finally the rainfall simulation method developed in Chapter 8 is tested as input to the rainfall-runoff model. A measure of rainfall spatial variability is introduced. The catchment characteristic time is taken as an independent variable of catchment response. The impact of scale is also assessed.

Section 12.2 introduces the Lee Catchment and the data used in this study. Section 12.3 describes the rainfall-runoff model. The modelling strategy is discussed in Section 12.4. Calibration of the Lee catchment is presented in Section 12.5. The indicators of rainfall variability and basin damping are presented in Section 12.6. Results using alternative descriptors of observed rainfall on existing catchments and artificially urbanised basins can be found in Section 12.7 and 12.8. Performance of the developed rainfall disagregation scheme is presented in Section 12.9. Discussion and recommendations conclude the chapter in Section 12.10.

12.2 Case-Study area

12.2.1 The Lee catchment

The River Lee drains a relatively large catchment of 1416 km^2 and is the principal tributary of the River Thames within Greater London, UK. The catchment is characterised as humid-temperate with a mean annual rainfall of 632 mm. With a maximum length (east-west) and width (north-south) of 56 km, its elevation ranges from 20 to 250 m. The River Lee essentially flows in a southerly direction from its source in Bedfordshire through North London before joining the River Thames on the east side of London, near Bow Creek. It is convenient to divide the catchment into two parts: The Upper Catchment comprising the part of the basin upstream of Feildes Weir and the Lower Lee, defining the area below it. A map of the Lee catchment is provided in Figure 12.1.

The solid geology of the Upper Catchment is relatively simple and can be divided along a NE-SW line (See Figure 12.2), being largely Chalk of the Cretaceous period, overlain in the south and east by London Clay (Sumbler, 1996). Throughout much of the basin, the superficial deposits which overlay the solid geology complicate the picture. These include the predominant boulder clay of central and east Hertfordshire, the Clay-with-flints of much of west Hertfordshire, and the gravels of the river valleys. The Chalk is the most important aquifer of the region. The Chalk has complex hydraulic properties. It has a porous matrix providing groundwater storage but, because the pores and their interconnections are so small, the matrix permeability and thus the speed with which water can be taken up or released, is low. Chalk has also a second porosity, due to the presence of joints and fissures. These fissures can provide rapid groundwater transport and contain the most mobile water. Because of these properties, the level of the water table changes rapidly in response to winter rains or summer droughts (Sumbler, 1996). The Upper Catchment to Feildes Weir is mainly rural, characterised by arable farming. The area has seen significant growth in housing since the 1950s, with urban areas covering 15% (See Table 12.3) of the total Upper Catchment area. From West to East, the main tributaries of the upper catchment are the Upper Lee, the Mimram, the Beane, the Ash, the Rib and the Stort. The Upper Lee, a chalk catchment, is substantially urbanised up to Luton Hoo and is joined by the Mimram at Hertford. The Mimram drains a mostly rural area and is predominantly chalk, characterised by low natural runoff. The Beane joins the Middle Lee just downstream in the centre of Hertford. Low runoff from chalk areas is augmented by runoff from overlying boulder clay and by urban runoff from the town of Stevenage. The Rib catchment is underlain by chalk, but with extensive overlying deposits of boulder clay, runoff is moderately high, albeit characterised by a marked lag. The Ash catchment is almost entirely rural and has similar runoff characteristics to the Rib. Finally, the Stort joins the Middle Lee immediately upstream of Feildes Weir. Runoff rates from the Stort vary considerably, being high on the clay and much lower on the chalk. The Stort presents a mainly rural upper catchment, but includes substantial urban development in the valley. It encompasses the towns of Stansted Mountfichet, Bishop's Stortford, Sawbridgeworth and Harlow. The flow regime is dominated by the influence of the Stort Navigation. The hydrogeological characteristics, main components of drift and level of urbanisation can be found in

Tables 12.1, 12.2 and 12.3.

Table 12.1 % Distribution of Hydrogeological characteristics of the Upper Catchment of the Lee. (Source: http://www.nerc-wallingford.ac.uk/ih/nrfa/spatialinfo/Index/indexEAThames.html)

	High permeability (fissured)	Very low permeability	Mixed permeability
Upper Lee	90.1	6.0	3.9
Mimram	97.0	-	3.0
Beane	96.1	-	3.9
Rib	93.9	-	6.1
Ash	63.9	10.5	25.5
Stort	24.1	70.8	5.1
Feildes Weir	71.4	22.1	6.4

Table 12.2 % Distribution of the main components of drift (Source: http://www.nerc-wallingford.ac.uk/ih/nrfa/spatialinfo/Index/indexEAThames.html)

	Sand and gravel	Boulder clay and morainic drift	Clay with flints
Upper Lee	22.3	13.5	9.8
Mimram	14.9	9.3	35.6
Beane	14.9	41.8	7.5
Rib	12.5	81.3	-
Ash	14.9	77.2	-
Stort	18.4	76.2	-
Feildes Weir	18.4	50.8	7.2

Table 12.3 Urbanisation (Source: http://www.nerc-wallingford.ac.uk/ih/nrfa/spatialinfo/Index/indexEAThames.html)

	Built-up areas (%)	Main towns
Upper Lee	34.7	Luton, Hatfield
Mimram	12.7	Welwyn
Beane	12.7	Stevenage
Rib	3.5	-
Ash	3.5	-
Stort	12.9	Bishops, Harlow
Feildes Weir	14.9	Hertford

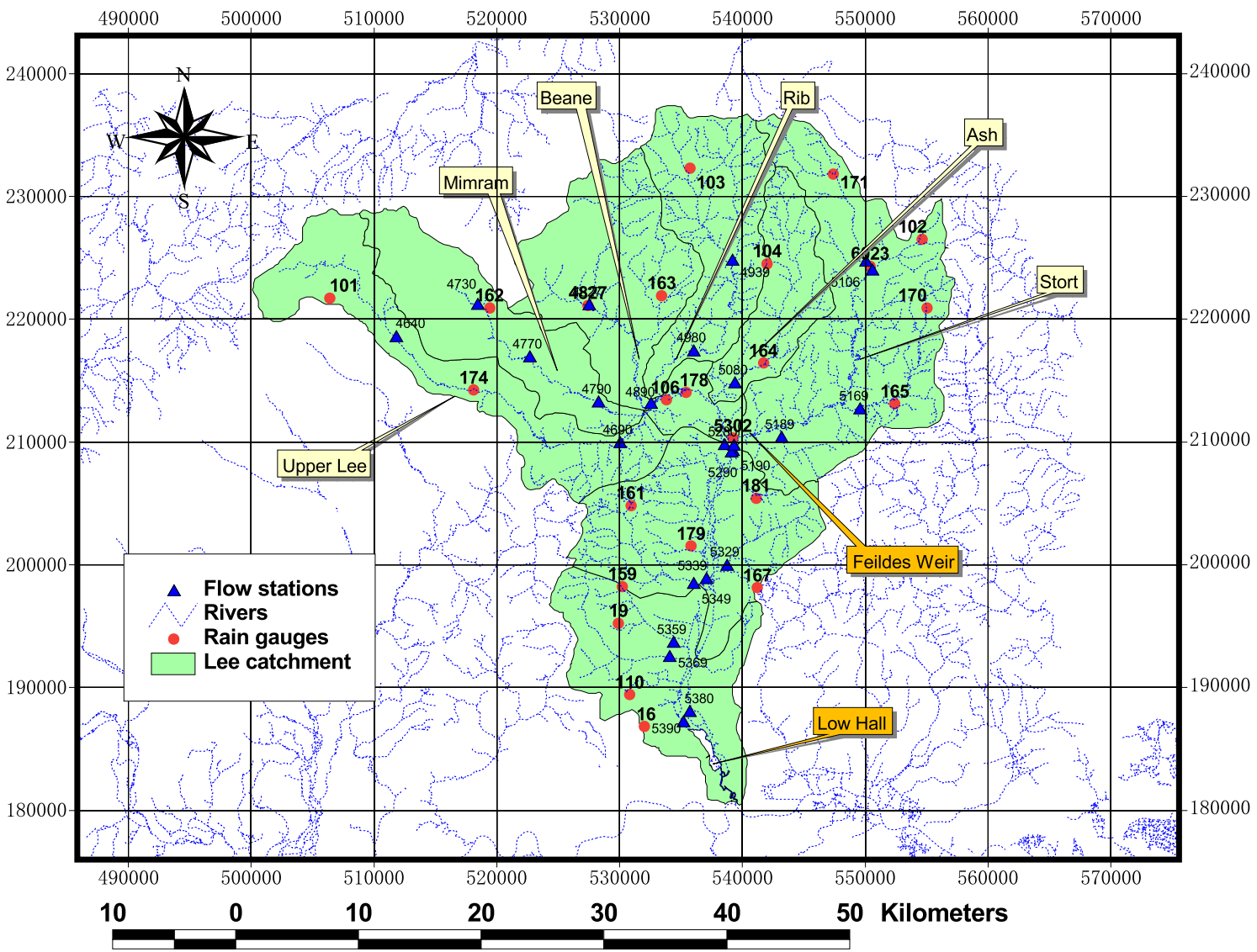


Figure 12.1 Map of the Lee catchment

12.2.2 Database

Radar Archive

A long record of continuous data (1987-2002) from the Chenies radar in the Thames region together with a network of 122 supporting raingauges have been obtained from the Environmental Agency of England and Wales. The radar visualisation has been achieved at 2 and 5 km grid resolution. The Lee catchment lies within the radar umbrella and the 2km data, aggregated at an hourly time-step, are used in this study. These records represent rainfall intensities averaged over 2km by 2km National Grid squares covering a 152 by 152 km square centred at the Chenies radar. The raingauge data are recorded every 15 minutes and represent rainfall amount. The display of the radar archive revealed a number of anomalies and a calibration procedure has been developed to improve the accuracy of the radar rainfall estimates (Bellone, 2003).

Raingauge records

Hourly data from a network of 17 raingauges, the characteristics of which are given in Table 12.4, are used as input to the semi-distributed rainfall-runoff model. Their location on the Lee catchment can be found in Figure 12.1. Due to the proximity of sites 106 and 178 (1.7 km distance) and the fact that their period of record is complementary, they are taken as a single gauge for the rainfall-runoff investigation.

Table 12.4 Raingauge records for the Upper Catchment of the Lee

Site ID	Name	Easting m	Northing m	Altitude m	Start	End
0101	Runley Wood	506400	221700	137	1988	2002
0103	Chipping	535700	232300	104	1988	2002
0104	Braughing Friars	542000	224500	116	1988	2002
0106	Hertford	533800	213400	37	1988	1994
0160	Mill Green	524500	210000	68	1989	2002
0161	Darnicle Hill	530900	204800	73	1990	2002
0162	Whitwell	519400	220900	90	1989	2002
0163	Dane End	533400	221900	70	1989	2002
0164	Widford	541700	216400	60	1990	2002
0165	Hatfield Heath	552400	213100	67	1990	2002
0170	Takeley	555000	220900	94	1990	2002
0171	Clavering	547400	231800	84	1990	2002
0174	Wheathampstead	518100	214200	86	1991	2002
0178	Broadmeads	535400	214000	35	1994	2002
4827	Bragbury Park	527400	221100	69	1988	2002
5302	Rye Meads	539200	210300	30	1990	2002
6023	Stansted Argus	550400	224300	70	1987	2002

Flow data

Hourly flow data from 12 stations on the Upper Lee catchment are used in this investigation The location of the flow gauging stations can be found in Figure 12.1 and

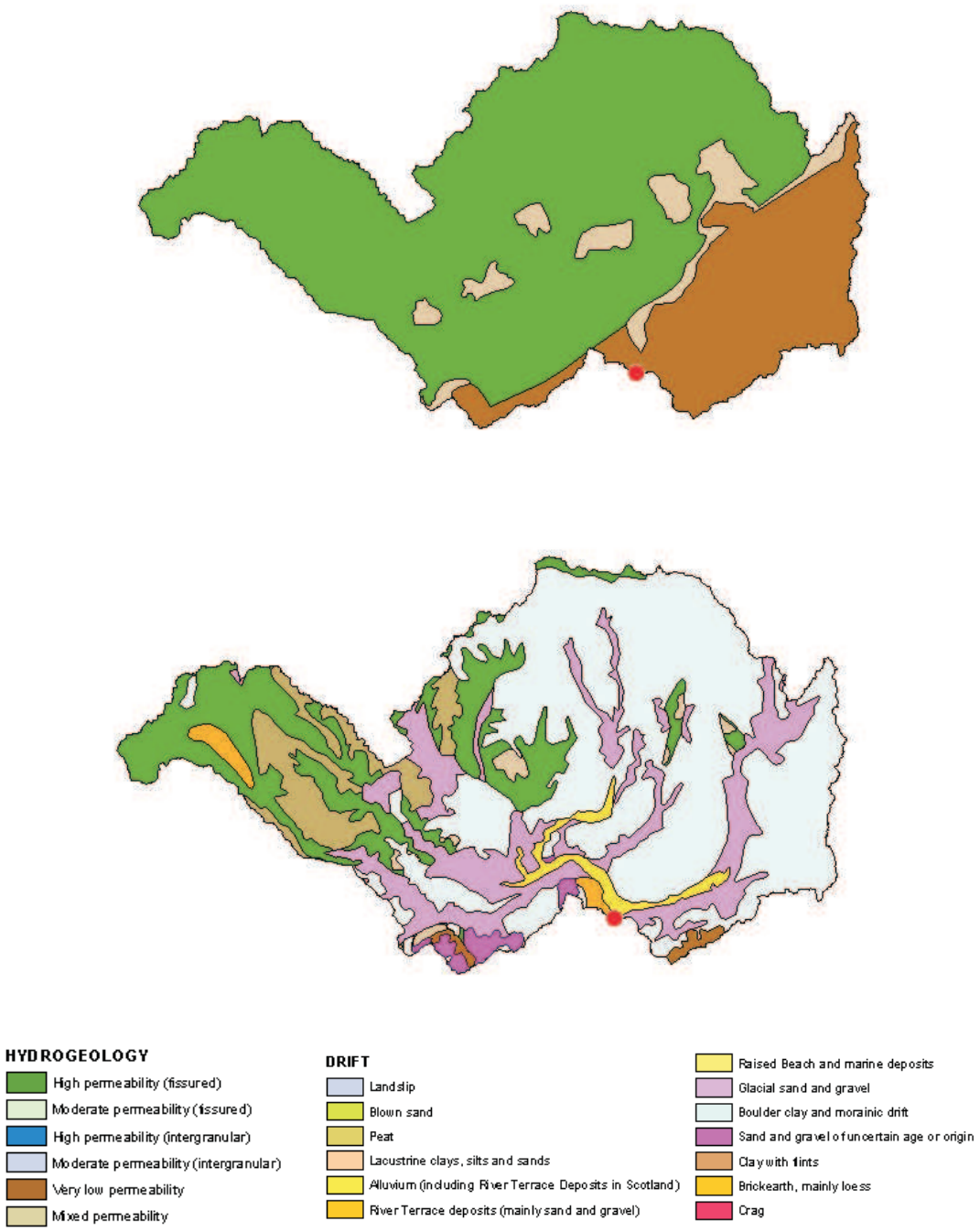


Figure 12.2 Hydrogeological characteristics of the Upper Catchment of the Lee (top map). The lower map shows the drift deposit on top of the hydrogeological properties.(Source:http://www.nwl.ac.uk/ih/nrfa/spatialinfo/Geology/geology038001.html)

their characteristics are given in Table 12.5. The (hourly) temporal autocorrelation function of the streamflow time-series, plotted in Figure 12.3, reveals different behaviour according to the geology of the catchment. The two predominantly chalk subcatchments, the Upper Lee and especially the Mimram show a high persistence in the flow compared to the watersheds with a higher proportion of clay.

Table 12.5 Flow gauging data for the Upper Catchment of the Lee

Station	Location	River	Grid reference	Start	End
4640	Luton Hoo	Upper Lee	TL11791860	1987	2002
4690	Water Hall	Upper Lee	TL29990995	1987	2002
4790	Panshanger	Mimram	TL28241328	1987	2002
4827	Stevenage (Bragbury)	Beane	TL27492116	1987	2002
4890	Hertford (Hartham)	Beane	TL32511314	1987	2002
4939	Griggs bridge	Rib	TL39142480	1987	2002
4980	Wadesmill	Rib	TL3597144	1987	2002
5080	Mardock	Ash	TL39361484	1987	2002
5129	Gypsy Lane	Stort	TL50532403	1987	2002
5169	Sheering	Stort	TL49531270	1987	2002
5190	Glen Faba	Stort	TL39180931	1988	2002
5290	Feildes Weir	Middle Lee	TL39070920	1987	2002

12.3 RORB Rainfall-runoff model

A semi-distributed rainfall-runoff modelling approach is applied to account for the catchment heterogeneity. The semi-distributed model permits to explore spatial variations of rainfall without entering the level of complexity of fully distributed models. Preference is given to an event-based modelling procedure for its simplicity as it excludes components describing losses such as evapotranspiration and soil infiltration. The RORB model is selected as it satisfies these requirements. It has only a small number of model parameters (3 or 4 depending on the type of run) to be identified and has been extensively used in practice to represent the hydrological routing of this particular area (Lotufo Conejo, 1979; Flynn & Rothwell, 1991b).

RORB is a hydrologic simulation model designed to estimate the flood hydrograph by routing rainfall excess through the catchment. Its general concept is illustrated in Figure 12.4. From the total rainfall over a subcatchment, losses are deducted through the loss model in such a way that the volume of the generated hyetograph of rainfall excess equals the volume of the observed surface runoff hydrograph (hydrograph less baseflow).

The user can choose two loss models:

- Model 0 is based on an initial loss which must be satisfied before rainfall excess is generated, followed by a constant proportional rate of loss and hence of runoff (runoff coefficient)

- Model 1 is composed of an initial loss followed by a constant loss rate

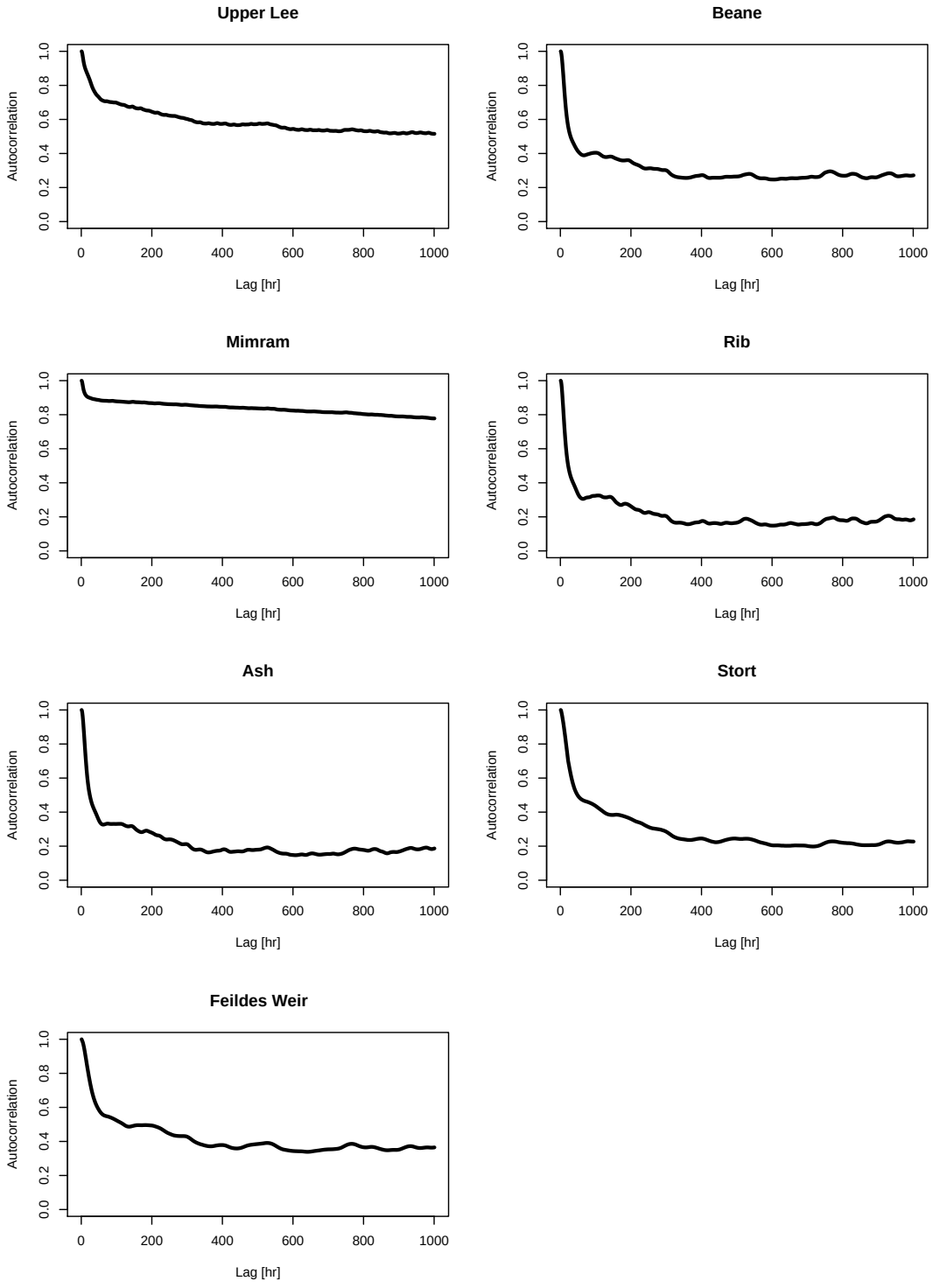
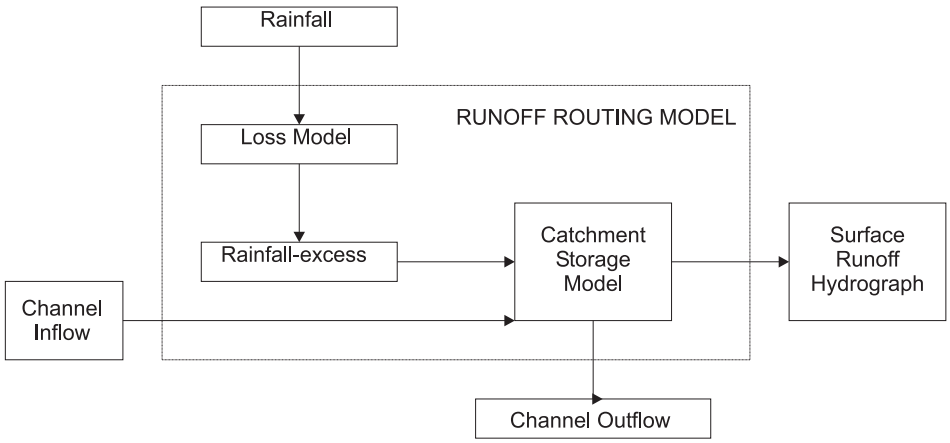


Figure 12.3 Autocorrelation function of the streamflow times-series

Figure 12.4 RORB structural diagram of the model (Source:Laurenson and Mein (1988))



In this study, Model 0 is selected. The excess hyetograph is converted to an input hydrograph which is then routed through the subcatchment storage model and the result is a surface runoff hydrograph at the outlet of the subcatchment. The model can include several subareas and allows for different rainfall depths and patterns over the catchment. Areal variability of loss parameters over sub-areas is taken into account via the fraction imperviousness for urban areas or runoff capacity index for rural areas. The parameters for each sub-area are then calculated as follows:

$$IL_i = (1 - F_i)IL_{perv} \tag{12.1}$$

$$C_i = F_i C_{imp} + (1 - F_i)C_{perv} \tag{12.2}$$

$$CL_i = (1 - F_i)CL_{perv} \tag{12.3}$$

where IL, C and CL are respectively the initial loss, runoff coefficient and continuing loss rate, F is the fraction imperviousness (or relative runoff capacity), and subscripts i, imp, and $perv$ indicate the ith sub-area, the impervious area and the pervious area. C_{imp} is set to 0.9 by the program (Laurenson and Mein, 1988). The storage-discharge relation is represented by a non-linear element:

$$S = kQ^m \tag{12.4}$$

where S represents the storage in the subcatchment, Q is the discharge leaving the subcatchment and k and m are constant parameters defining the functional relationship between storage and outflow. The second equation used is the continuity equation expressed in finite difference form:

$$\frac{I_t + I_{t+1}}{2} - \frac{Q_t + Q_{t+1}}{2} = \frac{S_{t+1} - S_t}{\Delta t} \tag{12.5}$$

where I relates to the ordinate of the hydrograph obtained from the rainfall excess, Q to the ordinate of the outflow hydrograph, S to the model storage, t and t+1 are time indices referring the variables in the equation to successive time steps and Δt is the time interval used. Equations 12.4 and 12.5 require the use of an iterative technique to

solve the non-linear system (the Newton Raphson method is used in the model which is described in the appendix of the user manual (Laurenson and Mein, 1988)).

k is the product of two factors:

$$k = k_c \times k_r \tag{12.6}$$

k_r is a dimensionless ratio called the relative delay time for a given model storage, determined from physical characteristics of the stream channel. k_c is an empirical parameter, which is calibrated by fitting the model to recorded rainfall and streamflow data. For a given storage i, the relative delay time is calculated as

$$k_{ri} = A\frac{L_i}{d_{av}} \tag{12.7}$$

where L_i is the length of reach represented by storage i (km), d_{av} is the average flow distance in the channel network of subarea inflows (km) and A is a factor depending on the type of reach. For instance, A is 1 when the reach is in natural condition or can be computed as:

$$A = \frac{1}{9S^{0.5}} \tag{12.8}$$

for lined or piped reaches (typically applied to urban catchments), where S is the slope of the channel reach. Hence the program has four main parameters to be calibrated: the initial loss (IL), the runoff coefficient (or constant loss rate) and the parameters m and k_c. The parameters m and k_c influence the shape of the hydrograph, regulating the attenuation and delay caused to the input hydrograph by the storage of water in the model storage. The runoff coefficient defines the volume of water in the input hydrograph.

The model provides three types of run: FIT, TEST and DESIGN, the characteristics of which are given in Figure 12.5. On gauged catchments, a FIT run is used first with data for one or more of the available recorded floods to evaluate the model parameters. When the initial loss has been entered as input by the user, the program calculates the pervious area runoff coefficient or continuing loss rate by iteratively achieving a volume balance of rainfall-excess with measured surface runoff. The TEST run is used next with data for the remainder of the recorded floods to test the model. Assuming that the results of the TEST run are satisfactory, a DESIGN run is then used to predict the catchment behaviour under various hypothetical design conditions. Model parameters should all be entered as input in a DESIGN run.

In order to run the model, the following steps are taken:

- The catchment is divided into sub-areas to provide for areal variation of rainfall and losses and the fact that runoff from different parts of the catchment travels different distances to the outlet. Rainfall and losses are averaged over each sub-area.
- Normal (model) storages are placed in each river reach located between the centroid of the subcatchment and the outlet. Special storages such as storage reservoirs, lakes and retarding basins, can also be defined. They are modeled differently from the normal channel reach storage.
- To derive the surface runoff hydrographs for input to the model the baseflow contribution needs to be removed from the observed hydrographs

TYPE OF STUDY	ITEM		TYPE OF RUN		
			FIT	TEST	DESIGN
CATCH-MENT STUDIES & FLOOD ROUTING STUDIES	HYDROGRAPH		Known	Known	Unknown
	MODEL PARAMETERS		Unknown	Known	Known or assumed
	PURPOSE OF RUN		To determine model parameters	To test model	To determine design hydrograph
	OUTPUT LABELED:		FIT RUN	TEST RUN	DESIGN RUN
	SPECIAL STORAGES		Existing only	Existing only	Existing or to be designed
CATCH-MENT STUDIES ONLY	TRIAL ROUTING PARAMETERS		Suggested by program	Not suggested	Suggested, but k_c only if $m = 0.8$
	LOSS PARA-METER	GAUGING STN.	Calculated	Calculated	Input
		NO GAUGING STATION	Input from terminal	Input from terminal	Input from terminal

Figure 12.5 Types of run. (Source: Laurenson and Mein (1988))

- The model calibration requires the determination of the initial loss (IL), m, k_c and depending on the type of run, the runoff coefficient (RC). m, k_c are constant parameters, IL and RC vary depending on the event.

12.4 Modelling Strategy

12.4.1 Model Configuration

Previous studies were undertaken using the RORB model on the Lee catchment. Lotufo Conejo (1979) designed a configuration for the whole Lee catchment. The aim was to develop a flood prediction method to route flood flows in the River Lee from Feildes Weir to Low Hall (Wheater and Lotufo Conejo, 1979). RORB was used to estimate inputs to the Lower Lee, both from the Upper Catchment tributaries and from the subcatchments below Feildes Weir. Flynn & Rothwell (1991b) focused on the upper part of the catchment and modelled various tributaries independently (Hawnt, 1987b,a). Their aim was to integrate all tributary RORB models in order to cover the whole Upper Lee and estimate design floods at Feildes Weir (Flynn & Rothwell, 1991b). Ideally, previous and new RORB models for all tributaries would be combined to create a single model. However, a large number of sub-areas needs to be modelled and this would require too much time due to the computer memory constraints within RORB. The existing RORB models are run independently for each tributary to provide tributary inflows to a new model combining all flow to Feildes Weir (Middle Lee model).

Due to data inaccuracy of the Stort at Glen Faba, the Middle Lee model incorporates the Upper Lee, the Mimram, the Beane, the Rib and the Ash but excludes the Stort contribution, which is modelled independently (Flynn & Rothwell, 1991a). The Upper Catchment of the Lee is heterogeneous and a single model would result in a loss of accuracy at the subcatchment scale and extra "special storages" would be needed to locally modify the global routing parameters "m" and "k_c". Following Flynn & Rothwell (1991b), the Lee catchment is subdivided into 6 subcatchments, of zones assumed internally homogeneous, and the rainfall-runoff model is applied to each of these regions. The flow from each tributary, including the Stort configuration, is then added and routed to the outlet of the catchment at Feildes Weir. This option allows for spatial variation of the model parameters. Further it was perceived essential to keep the same model discretisation at subcatchment and catchment scale. The impact of spatial rainfall can be studied as a function of subcatchment type, subcatchment scale (80-280 km^2) and catchment scale (1000 km^2). One drawback is that it is not easy to run the combined model under design conditions. Design runs have to be made for each tributary model, predicted flood hydrographs abstracted from output files and then combined with the Middle Lee model datafile as concentrated inflows.

12.4.2 Catchment Subdivisions

In order to minimise the impact of model discretisation, a relatively fine subdivision of the catchment is required. Apart for the Mimram, each subcatchment is subdivided into subareas of about 10 km^2 based on the stream network and the drainage divides of the catchment. The very low response of the Mimram, which is baseflow dominated due to the chalk geology, does not justify detailed subdivision (Flynn & Rothwell, 1991b). This size was chosen on the basis of previous studies (Hawnt, 1987b; Flynn & Rothwell, 1991b). This scale is small enough so that the impact of various rainfall representations leads to variation in sub-areal rainfall estimates and to satisfactory computations. Moreover, since for this catchment a raingauge influences on average 60 km^2 a smaller subdivision does not seem justified. For calibration purposes, spatially variable rainfall from the full raingauge network is used as input and this representation is referred to as the reference rainfall. Since a relatively dense network of representative raingauges is available, a simple method to derive the sub-areal precipitation is selected. Spatial interpolation of the point raingauges using the Thiessen Polygons method is used to estimate the areal rainfall. One of the drawbacks is the inflexibility of this method as a new Thiessen network needs to be constructed each time there is a change in the raingauge network, such as when data are missing from one of the gauges. The temporal pattern applied to each subarea is given by the nearest raingauge. The simulated flow driven by the reference rainfall is referred to as the reference flow.

12.4.3 Event Selection

The calibration of the rainfall-runoff model is based on events that span the whole dataset. The events are selected so that they are preceded by periods when there is no rainfall intensity recorded at all gauges for several hours (usually about 24 hours) and their characteristics in terms of event duration, arithmetic average of available

raingauge time-series (Mean rainfall depth), number of available gauges (No. RG) return period (T) determined using the methodology in Faulkner (1999), and coefficient of variation of rainfall (CV) are given in Table 12.6. ”Common” refers to events with return period less than a month (NERC, 1999).

Table 12.6 Events used for calibration

Event	Rainfall Duration (hr)	Mean rainfall depth (mm)	No. RG	T (Year)	CV(%)
11-Nov-91	19	22	15	1.2	13.6
18-Nov-91	20	22.4	15	1.2	13.4
08-Jan-92	25	11.6	16	common	49.1
28-May-92	56	55.2	16	7.5	33.9
13-Aug-92	13	20.9	15	1.4	13.5
19-Oct-92	34	30.7	16	1.56	38.1
09-Apr-93	21	14.4	16	common	12.2
12-Aug-93	10	16.1	16	1.1	26.6
11-Oct-93	65	62.7	16	11.2	16.9
03-Feb-94	8	12.1	16	1.01	16.5
15-Sep-94	24	15.4	16	common	28.6
07-Dec-94	47	21.9	16	1	27.9
16-May-95	25	17.1	16	common	11.1
22-Dec-95	45	27.4	16	1.08	14.6
08-Jan-96	19	20.3	16	1.09	14.3
12-Feb-96	41	18	16	common	18.9
23-Jul-96	4	20.6	16	4.6	81
22-Aug-96	18	17.3	16	1.01	23.1
24-Feb-97	39	21.2	16	1	22.6
14-Apr-98	33	26.3	15	1.17	16.3
02-Apr-00	49	35.7	14	1.69	38.6
27-May-00	20	22.2	15	1.19	12.8
10-Nov-00	32	13.8	14	common	19.6
03-Jan-01	44	20	16	common	20
26-Jan-01	14	14	16	1.02	28.6
21-Oct-01	24	40.7	15	5.8	32.7
30-Jul-02	30	21.4	16	1.03	42.7
09-Sep-02	16	20.5	16	1.19	39.3

12.4.4 Baseflow Separation

Direct runoff is required as input to RORB and as a first approach a simple straight line method is applied (Chow et al., 1988) to each flow event to remove the baseflow component. The basic principles are illustrated in Figure 12.6 where AB and DE represent the baseflow recession, BC the rising limb, CD the falling limb. A straight line from the point at which surface runoff begins (B) to the intersection with the recession limb (E) is drawn to produce the direct runoff storm hydrograph. The point of inflection from the logarithm of the discharge time-series determines the point (E) at which the straight line and the hydrograph recession limb intersect.

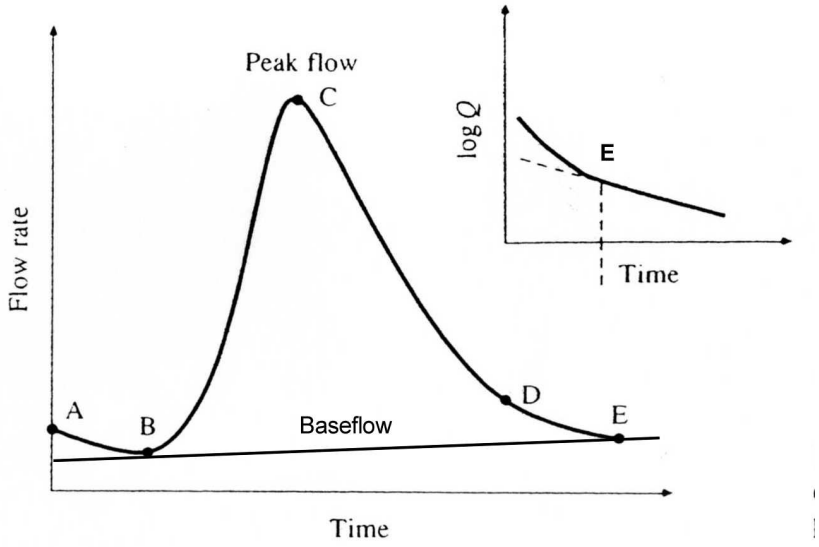


Figure 12.6 Baseflow separation (Adapted from Chow et al. (1988))

12.4.5 Performance criteria

Model parameter calibration and validation is based on the reproduction of the peak outflow (Q_p) and the time to peak (T_p), which is defined as the lag time between the start time and the peak time of the runoff hydrograph. This is complemented by a non-dimensional form of the time-series least squares error, the well-known Nash-Sutcliffe efficiency (NSE), which is also sensitive to hydrograph peak error.

$$NSE_{obs} = 1 - \frac{\sum_{i=1}^{M}(C_i - O_i)^2}{\sum_{i=1}^{M}(O_i - \overline{O})^2} \tag{12.9}$$

where C_i and O_i are the calculated and observed discharge at hour i, $\overline{O}$ the mean discharge over the event duration, M the number of time steps.

12.4.6 Calibration Strategy

The aim of the calibration is to identify model parameters which provide a suitable basis for analysis of model sensitivity to precipitation inputs. Manual calibration is undertaken, varying the model parameters m, k_c and IL for a range of events to achieve a close match between the observed and calculated hydrograph. The procedure is as follows:

- Several events covering a wide range of peak discharge and taken at different time of the year are selected (typically four to six events per subcatchment).
- In the calibration or FIT runs, for each event, a selected range of m, typically $\{0.6, 0.7, 0.8, 0.9\}$, is used. For each m value, k_c is varied until the best fit in

terms of peak flow and time to peak is obtained. Decreasing kc increases the hydrograph peak and decreases the lag, while increasing k_c does the opposite (Laurenson and Mein, 1988). Finally, the initial loss is adjusted to improve the fit. The parameter set maximising the NSE_{obs} criteria is selected for each calibrated parameter set.

- In the validation or TEST runs, the calibrated parameter sets are tested on a few (typically two to four) other events to determine an *a priori* optimum set.
- A sensitivity analysis is conducted, applying each calibrated parameter set to all events in the FIT runs and computing the resulting NSE_{obs} criteria. This indicates a standard parameter set reproducing best all events and either confirms the results from the TEST run, or highlights the need to use an alternative set in addition to the standard one to fit specific types of flow.

12.5 Calibration and Validation

Example datafiles were provided for the Upper Lee, the Mimram, the Ash, the Stort and the Middle Lee. Documentation on the Mimram, the Ash and the Middle Lee is given in Flynn & Rothwell (1991b). Documentation on the Stort can be found in Flynn & Rothwell (1991a). Some description of the model structure of the Beane and the Rib can be found in Hawnt (1987b) and Hawnt (1987a). The Mimram, the Ash and the Stort were calibrated by Quasem (2004).

In the first instance, seven RORB models are calibrated independently and the results for the individual subcatchments are presented in Section 12.5.1. The calibration of the Feildes Weir outlet using observed tributary inflows is referred to as the Middle Lee model. Second the integrated Lee calibration is presented by choosing a set of events common for all tributaries in Section 12.5.2. The integrated modelling of the Upper Catchment of the Lee is referred to as the Feildes model.

12.5.1 Modelling the individual subcatchments

A map showing the discretisation of each subcatchment is included in Figure 12.7. A summary of the subcatchment configurations can be found in Table 12.7. This refers to the number of subareas, the number of raingauges used for each tributary, the intermediate and main flow gauging stations (FGS), the type of reaches, the number of special storages and whether a coefficient of imperviousness is included in the modelling configuration. For a particular subcatchment, it was not possible to determine a unique parameter set. The fitted k_c generally increased with the size of the flow event. A summary of the calibrated model parameters $\{m, k_c\}$ for each subcatchment is presented in Table 12.8. ”Standard” refers to the default parameter set fitting most events, ”Low” and ”Large” refer to the parameter set required to represent low or large flows determined by a peak flow above the threshold given in Table 12.8. Even though the initial loss is defined as an event parameter, a suggested value fitting most events on a particular subcatchment is included in Table 12.8. As shown in Table 12.9, a good fit was achieved for all tributaries with NSE_{obs} on average

for all subcatchments and events of 0.88 and 0.89 in calibration and validation mode, respectively.

Table 12.7 Individual sub-catchment calibration summary table

Catchment	Upper Lee	Mimram	Ash	Beane	Rib	Stort	Middle Lee
Area (km^2)	156	132	79	176	134	278	78
No. Sub-areas	16	5	8	21	11	32	15
Range (km^2)	1.5-19.2	18.5-31	6.8-13.3	2.5-14.8	6.8-19.7	2-18.6	2.2-10
No. RG	5	5	4	5	6	7	6
Centroid RG	174	162	104	163	104	170	106/178
Inter. FGS	1	0	0	1	1	2	6 (inflow)
Main FGS	Water Hall	Panshanger	Mardock	Hartham	Wadesmill	Glen Faba	Feildes Weir
Reaches							
-Natural	21	9	14	33	18	50	15
-Drowned	9						1
N. Special							
Storage	4	0	0	0	0	4	0
Coeff. imp.	Yes	No	No	Yes	No	No	Yes

Table 12.8 Calibrated parameters

	Upper Lee	Mimram	Beane	Rib	Ash	Stort	Middle Lee
Low							
m - k_c	-	-	0.8 -7	-	0.8-7	0.8 - 17	0.6 - 7
Flow			<1 m^3/s		<1 m^3/s	<3 m^3/s	<10 m^3/s
Standard							
m - k_c	0.8 - 21	0.8 - 7	0.8 - 21	0.8 - 16	0.8 - 15	0.8 - 37	0.6 - 21
Large							
m - k_c	0.8 - 33	0.8 - 17	-	0.8 - 21	0.8 - 21	-	0.6 - 35
Flow	>5 m^3/s	>3 m^3/s		>15 m^3/s	>7 m^3/s		>50 m^3/s
IL (mm)	3	0	5	9	7	0	0

Table 12.9 Calibration results

Upper Lee		Beane		Rib		Ash		Mimram		Stort		Middle Lee	
Event	NSE_{obs}	Event	NSE_{obs}	Event	NSE_{obs}	Event	NSE_{obs}	Event	NSE_{obs}	Event	NSE_{obs}	Event	NSE_{obs}
Calibration													
09-Apr-93	0.78	09-Apr-93	0.97	08-Jan-92	0.96	19-Oct-92	0.95	28-Apr-92	0.96	18-Nov-91	0.87	08-Jan-96	0.95
03-Feb-94	0.92	21-Oct-01	0.96	09-Apr-93	0.62	11-Oct-93	0.94	12-Aug-93	0.92	19-Oct-92	0.97	22-Oct-01	0.94
21-Oct-01	0.84	08-Jan-96	0.97	03-Feb-94	0.72	07-Dec-94	0.86	14-Sep-94	0.82	09-Apr-93	0.77	03-Feb-94	0.87
08-Jan-96	0.94	16-May-95	0.78	21-Oct-01	0.95	22-Dec-95	0.92	23-Jul-96	0.97	22-Aug-96	0.79	30-Jul-02	0.92
23-Jul-96	0.93					12-Feb-96	0.96	11-Oct-93	0.64			14-Apr-98	0.92
						16-May-95	0.83						
Validation													
14-Apr-98	0.95	14-Aout-98	0.95	08-Jan-96	0.74	26-Jan-01	0.95	02-Apr-00	0.90	09-Sep-02	0.80	13-Aug-92	0.87
16-May-95	0.95	27-May-00	0.94	01-Apr-93	0.88	08-Jan-96	0.97	03-Jan-01	0.81	11-Oct-93	0.75	18-Nov-91	0.80
10-Nov-00	0.92	30-Jul-02	0.92	27-May-00	0.95	30-Jul-02	0.73			08-Jan-96	0.87	11-Oct-93	0.98
11-Oct-93	0.92			11-Oct-93	0.95					03-Feb-94	0.88		

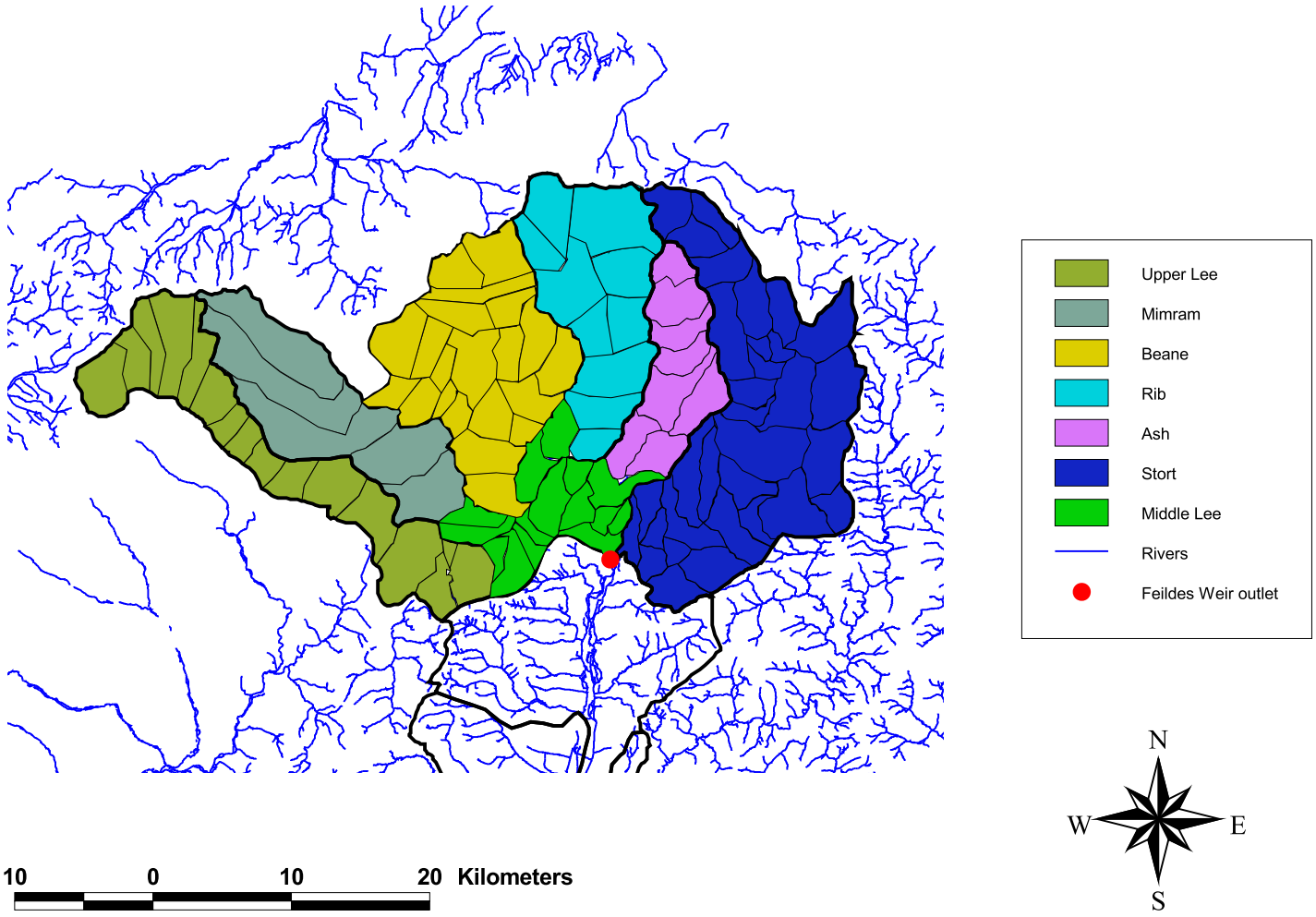


Figure 12.7 Subcatchment subdivisions

12.5.2 Modelling the Integrated Lee Catchment

Following validation of the individual tributaries, the subcatchments are incorporated into an integrated modelling strategy to achieve the catchment hydrological routing. As mentioned in Section 4.1, each subcatchment is first modelled independently and the simulated runoff is then used as inflow to the Middle Lee model and routed into the total runoff generated at Feildes Weir to represent the behaviour of the basin at the 1040 km^2 catchment scale (Feildes Model). A common set of events is defined in order to test the combined calibration at Feildes Weir. This task is more difficult to undertake as radar images, rainfall and discharge data are not always available at all 16 raingauges and 7 main flow stations.

Selected events

Due to data availability and the manual calibration procedure, the study is limited to five individual storm events. In the context of flood management, these events are of medium to large size. They are taken at different times of the year, representative of frontal and convective activities and associated with high, medium and low flows. Hence the largest rainfall event in the database, 11th October 1993, leading to the highest peak plow in the database is selected to represent an extreme event. This is complemented with two medium, more frequent events in winter, on 03rd February 94 and 08th January 1996, and two frequent events in Summer on 16th May 1995 and 30th July 2002. Summer events are difficult to calibrate since, although rainfall may be large, they generate low flow values. Hence the above events are selected so that a minimum peak flow of 5 m^3/s is recorded at Feildes Weir (including baseflow). The coefficient of variation of rainfall is large for the event of 30th July 2002. This event is characterised by a high temporal and spatial variability as illustrated in Figure 12.8, representing two contrasting temporal profiles at two raingauges relatively close in space (10 km apart).

Table 12.10 Characteristics of selected events

Events	Duration hr	Mean mm	CV. %	Direct runoff cumecs
11 Oct 1993	65	62.7	16.9	109.1
03 Feb 1994	8	12.1	16.5	28.8
08 Jan 1996	19	20.3	14.3	30.7
16 May 1995	25	17.1	11.1	4.3
30 Jul 2002	30	21.4	42.7	4.2

The range in runoff coefficient calibrated to the five selected events is given in Table 12.11. The runoff coefficient range given for the Feildes model is applied to the Middle Lee area only (i.e. 78 km^2). For most subcatchments, it can be seen that the runoff coefficient is higher in winter than in summer, which reflects the wetter/drier antecedent catchment condition in winter/summer. Compared to the other subcatchments, the Mimram is characterised by a low value of the runoff coefficient, which is in agreement with the observed damped behaviour of this basin.

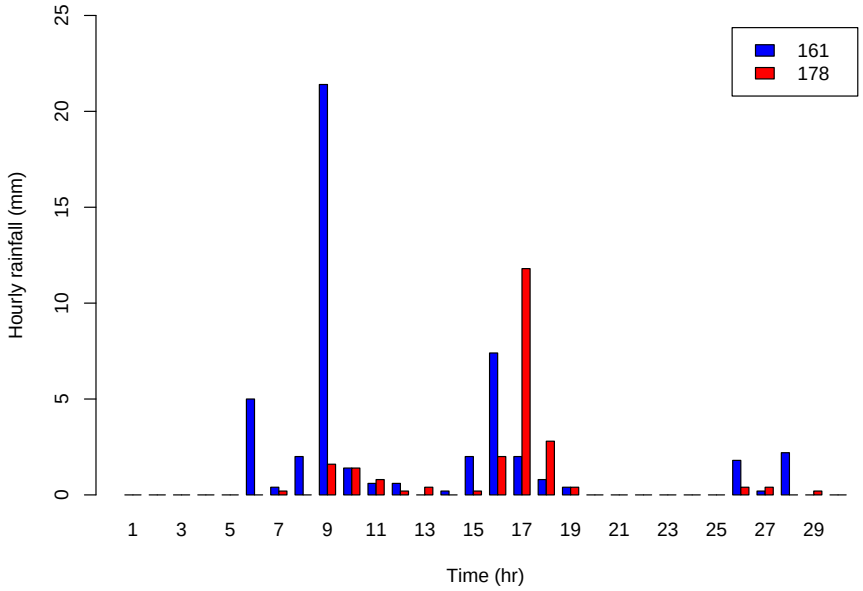


Figure 12.8 Variability of the rainfall of 30/07/02 at two sites

Table 12.11 Runoff coefficient (RC) calibrated for the 5 selected events

	Upper Lee	Mimram	Beane	Rib	Ash	Stort	Feildes
RC winter	0.14-0.23	0.03-0.04	0.09-0.29	0.32-0.74	0.21-0.39	0.33-0.43	0.15-0.69
RC summer	0.07-0.08	0.01-0.02	0.02-0.03	0.01-0.02	0.01	0.02-0.03	0.03

Model Performance

Using the calibrated parameters, the results in terms of NSE_{obs} for the 5 selected events are displayed in Table 12.12. A reasonable fit is achieved with NSE_{obs} of 0.77 on average for all subcatchments and events. The performance at interior points is lower in the case of extreme and summer events. The observed and reference flow at each (sub)catchment outlet are plotted in Figure 12.9 to 12.13.

Table 12.12 Calibration results in terms of NSE_{obs}

	Upper Lee	Mimram	Beane	Rib	Ash	Stort	Feildes
11 Oct 1993	0.92	0.55	0.54	0.96	0.94	0.75	0.84
03 Feb 1994	0.92	0.65	0.83	0.86	0.84	0.88	0.95
08 Jan 1996	0.94	0.64	0.97	0.85	0.97	0.88	0.87
16 May 1995	0.73	0.61	0.78	0.78	0.22	0.59	0.92
30 Jul 2002	0.61	0.84	0.92	0.51	0.73	0.23	0.88

Hence, calibration of the individual subcatchment has been achieved and the integrated modelling to Feildes Weir has been validated on five events. These events constitute a reference set to which rainfall representations based on radar data, subsets of the raingauges and the disaggregation scheme presented in Chapter 8 will be tested to assess the significance of spatial data for runoff generation. The findings are presented in Sections 12.7, 12.8 and 12.9.

12.6 Measure of Rainfall Variability and Basin Damping

12.6.1 Rainfall Representation

As mentioned previously, the reference rainfall consists of spatially variable areal estimates determined from the Thiessen polygons method. The temporal pattern applied to each sub-area is given by the nearest raingauge. In this study, in order to analyse the effects of various rainfall representations, the subcatchments are also run with the following rainfall inputs:

- Raingauge network:
  - Single raingauge with uniform rainfall over the subcatchment. The raingauge which is nearest to the centroid of the subcatchment is selected (See 'Centroid RG' in Table 12.7). Hence at the catchment scale, the flow at Feildes Weir is driven by seven raingauges (SG7).
  - Single raingauge with uniform rainfall over the catchment. The raingauge which is nearest to the centroid of the Feildes Weir catchment is selected (SG1).
- Spatially variable rainfall from radar data calibrated using the network of supporting raingauges (Bellone, 2003). The areal rainfall on each sub-area is

estimated from the average of the pixels included in the subarea. The temporal pattern applied on each subarea is given by the hourly averaged radar rainfall estimate for the sub-area.

- Spatially variable rainfall from spatially-uniform temporal disaggregation (presented in Chapter 8) conditioned on real daily data at the raingauges location. The temporal profile of the raingauge in the centroid of the catchment is selected.

SG7 and SG1 represents a degradation of spatial information in comparison to the reference rainfall and we expect lower model performance. The use of radar data represents an improvement in terms of spatial information but since the model is calibrated using the reference rainfall, similar model performance is expected. The disaggregation scheme preserves the areal rainfall variability in comparison to the reference rainfall but tests the hypothesis of a single temporal profile applied at all sites.

12.6.2 Spatial Deviation Index

The Spatial Deviation Index (SDI) (Quasem, 2004) is defined for each storm and (sub)catchment as a measure of spatial variability over a (sub)catchment. It measures the deviation between the sub-areal reference rainfall to the (sub)catchment averaged areal rainfall and is defined as follows:

$$SDI = \frac{1}{N}\sum_{i=1}^{N}\frac{|P_{ri} - P_T|}{P_T} \tag{12.10}$$

where P_{ri} represents the reference areal precipitation (i.e. as defined by the full raingauge network) for sub-area i in mm, P_T the (sub)catchment Thiessen averaged areal precipitation in mm, and N the number of sub-areas for a given (sub)catchment. A Reference Spatial Deviation Index (SDIR) compares the rainfall spatial deviation between the sub-areal estimates from the tested rainfall representation to the reference sub-areal precipitation estimates.

$$SDIR = \frac{1}{N}\sum_{i=1}^{N}\frac{|P_i - P_{ri}|}{P_{ri}} \tag{12.11}$$

where P_i represents the areal precipitation from either radar data or subsets of the gauges for sub-area i in mm, P_{ri} the reference areal precipitation in mm for sub-area i, and N the number of sub-areas.

For each subcatchment, the events defined in Table 12.10 are used to test the model sensitivity to the rainfall representation defined above. The model is run using the parameters selected from the calibration stage and the results in terms of SDI (%) and SDIR (%) are presented in Tables 12.13 and 12.14.

Results in Table 12.13 show that across all events and all subcatchments, the SDI ranges between 2.3% and 25.3%. As expected, larger discrepancies are introduced for the more variable event of 30th July 2002.

Table 12.13 SDI(%)

	Upper Lee	Mimram	Beane	Rib	Ash	Stort	Feildes
11 Oct 1993	11.3	5.5	10.0	7.9	2.3	9.1	8.1
03 Feb 1994	6.6	5.2	8.3	9.7	7.2	5.9	7.3
08 Jan 1996	6.9	8.3	8.6	13.1	4.6	10.6	8.3
16 May 1995	4.4	4.1	5.2	3.2	2.6	3.8	4.6
30 Jul 2002	19.6	11.6	11.2	11.7	25.3	25	19

Table 12.14 SDIR (%)

	Upper Lee	Mimram	Beane	Rib	Ash	Stort	Feildes
11-Oct-93							
Radar	11.5	18.6	11.6	18.6	10.8	18.9	15.9
SG7	18.0	9.3	12.9	10.0	2.2	12.5	10.9
SG1	14.6	17.5	18.4	8.7	4.7	9.1	10.8
03-Feb-94							
Radar	7.8	11.3	7.9	8.2	6.8	5.0	6.8
SG7	4.0	5.5	12.5	15.6	9.2	5.0	8.3
SG1	8.5	6.3	22.8	33.5	24.8	22.9	19.3
08-Jan-96							
Radar	7.4	2.8	7.2	10.4	4.2	10.3	8.2
SG7	9.0	10.6	9.1	20.7	4.8	21.9	12.9
SG1	7.2	10.4	9.3	26.4	11.3	19.2	13.0
16-May-95							
Radar	10.8	15.6	7.8	12.7	8.2	10.4	9.8
SG7	8.8	9.2	5.2	4.0	3.3	4.1	6.1
SG1	22.6	13.4	30.1	33.5	36.5	33.7	27.4
30-Jul-02							
Radar	30.0	34.6	28.5	22.0	16.2	27.1	25.0
SG7	39.9	15.4	20.8	11.8	17.3	32.8	26.3
SG1	21.2	21.2	27.1	19.0	24.8	42.3	29.2

Results from Table 12.14 show that SDIR ranges between 2% and 42%. Radar data lead to the largest deviation in areal rainfall, compared to SG7 and SG1 for the event of 11th October 1993 but in general, a closer agreement is seen between the reference rainfall and radar data, and the discrepancies are worse when SG1 is used. The event of 30th July 2002 lead to the highest deviations of the order of 30% for all rainfall representations.

12.6.3 Catchment Response Time

Indexes have been derived to measure the runoff variability in relation to rainfall (Woods and Sivapalan, 1999; Ogden and Julien, 1993; Smith et al., 2004; Morin et al., 2001). A common descriptor is the lag time which represents an intrinsic property of the catchment. It is a measure of the catchment response time and expresses the average travel time from all points on the catchment to the outlet. Following NERC (1975), the time from the centroid of total rainfall to the peak flow is taken as a measure of the lag time. In case of multipeaked events, the definition is widened to be the time from the centroid of total rainfall to the "centroid of peaks". For each subcatchment, a rainfall time-series is obtained as the arithmetic average of the raingauges. The lag time between the centroid of the averaged rainfall time-series and the reference flow is calculated. It varies between events according to the spatial and temporal distribution of the rainfall and to the antecedent condition of the catchment. Therefore, for each subcatchment, an average of the lag time over the five events is taken as the catchment response time.

12.7 Rural Catchment Results

In order to assess which level of spatial representation of rainfall is required to adequately reproduce the streamflow at the outlet, radar data, SG7 and SG1 are used as input to the rainfall-runoff model with the parameters calibrated to the reference rainfall. The simulated runoff is compared to the reference streamflow as illustrated in Figure 12.9 to 12.13. A modified definition of NSE is introduced at this stage to measure the performance of the simulated runoff in comparison to the reference flow according to:

$$NSE_{ref} = 1 - \frac{\sum_{i=1}^{M}(C_i - R_i)^2}{\sum_{i=1}^{M}(R_i - \overline{R})^2} \tag{12.12}$$

where C_i and R_i are the calculated and reference discharge at hour i. The results in terms of NSE_{ref} and NSE_{obs} can be found in Table 12.16 and Table 12.17. On average for all subcatchments and events, NSE_{ref} values of 0.95, 0.91 and 0.80 are observed for Radar, SG7 and SG1, respectively. In terms of NSE_{obs}, values of 0.73, 0.73 and 0.60 are observed for Radar, SG7 and SG1, respectively. It is difficult to interpret the results in terms of NSE_{obs} since the calibration is not perfect. The average value of NSE_{obs} between the reference and observed flow is 0.77 and the difference to a value of 1 (perfect fit) should be seen as a model error. The results in terms of NSE_{ref} express the errors introduced by the rainfall representation.

Radar data always lead to high performance at both catchment and subcatchment

scale ($NSE_{ref} > 0.95$). Considering the variation over all events, lower performance is observed for the extreme event on the Rib and the Ash (NSE_{ref} of 0.86 and 0.83, respectively) and during summer events (NSE_{ref} of 0.78 on the Upper Lee on 16th May 1995 and NSE_{ref} of 0.74 and 0.73 on the Upper Lee and the Beane on 30th July 2002). This suggests that the rainfall information captured by the radar data differ from that of the network of 16 raingauges in the case of the extreme and summer events studied. SG7 also generally gives high values of NSE_{ref} at catchment scale, with somewhat lower performance on the Rib on 3rd February 1994 (NSE_{ref} of 0.88), the Rib and the Stort on 8th January 1996 (NSE_{ref} of 0.82 and 0.72, respectively), the Upper Lee on 16th May 1995 (NSE_{ref} of 0.72), the Upper Lee and the Beane on 30th July 2002 (NSE_{ref} of 0 and 0.75, respectively). The performance is always high for the Mimram, the Ash and the Beane. This suggests that one gauge per triburaty may be adequate for the Mimram, a medium size Chalk basin and the Ash, an entirely rural small basin. A network of seven raingauges may be suitable to model the more heterogeneous large scale catchment (i.e. the flow at the catchment outlet only, no interior point).

SG1 gives lower results for all events at both catchment and subcatchment scale, except for the Mimram. Hence this confirms that the Mimram does not show much sensitivity to spatial rainfall and can be modelled using rainfall from one raingauge.

The graphs in Figures 12.14 to 12.18 display the percentage change in peak discharge (Q_p), peak time (T_p) and runoff volume (V) in comparison to the reference flow. The catchments are ranked by size. T_p is usually well reproduced (less than 8% error on average) but depending on the rainfall representation, larger errors are introduced with a percentage change ranging between [-40,80], [-20,60] for Q_p and V, respectively. Results in Table 12.15 show that the relative error in Q_p T_p and V is less with radar data and increases with SG7 and SG1. However there is no obvious trend between the percentage change in Q_p, T_p, V and the scale of the catchment. In conclusion, radar data lead to high values of NSE_{ref} at all scales and to the smallest variations in Q_p, T_p, V. SG7 lead to high NSE_{ref} values at the catchment scale but occasionally lower performance at subcatchment scale. This representation does not seem to affect the average T_p but compared to radar data, increases the errors in Q_p and V. Lower performance in terms of NSE_{ref}, Q_p, T_p and V are observed at subcatchment and catchment scales with SG1. Lower performance are observed for the extreme event and the summer events. The relative importance of spatial rainfall, basin damping and basin scale is assessed in the following sections.

Table 12.15 Percentage change in peak flow, time to peak and hydrograph volume

	Radar	SG7	SG1
Q_p (%)	12.6	15.3	28.1
T_p (%)	5.6	5.8	8.0
V (%)	7.4	12.4	21.6

Table 12.16 NSE$_{ref}$ rural catchments

	Upper Lee	Mimram	Beane	Rib	Ash	Stort	Feildes
11-Oct-93							
Radar	0.98	0.97	0.99	0.86	0.83	0.94	0.97
SG7	0.91	1	0.96	0.95	0.97	0.96	0.98
SG1	0.97	0.89	0.96	0.96	0.96	0.95	0.96
03-Feb-94							
Radar	0.92	1	0.99	0.99	0.98	1	0.99
SG7	0.93	0.99	0.93	0.88	0.99	1	0.99
SG1	0.97	1	0.73	0.36	0.74	0.89	0.86
08-Jan-96							
Radar	0.99	1	1	0.98	0.99	0.99	1
SG7	0.97	0.99	0.98	0.82	1	0.72	0.98
SG1	0.98	0.96	0.99	0.69	0.99	0.79	0.92
16-May-95							
Radar	0.90	0.97	0.94	0.97	0.99	0.99	1
SG7	0.82	0.98	0.98	0.99	1	0.99	0.94
SG1	0.42	0.92	0.54	0.69	0.54	0.74	0.65
30-Jul-02							
Radar	0.74	0.98	0.73	0.97	0.98	0.9	0.95
SG7	-0.06	0.97	0.75	0.96	0.97	0.78	0.98
SG1	0.35	0.93	0.53	0.99	0.77	0.61	0.85

Table 12.17 NSE$_{obs}$ rural catchments

	Upper Lee	Mimram	Beane	Rib	Ash	Stort	Feildes
11-Oct-93							
Radar	0.93	0.71	0.55	0.83	0.68	0.55	0.78
SG7	0.80	0.58	0.65	0.97	0.92	0.72	0.90
SG1	0.90	0.74	0.51	0.97	0.97	0.87	0.92
03-Feb-94							
Radar	0.87	0.65	0.78	0.88	0.74	0.89	0.92
SG7	0.90	0.58	0.78	0.62	0.81	0.89	0.92
SG1	0.90	0.61	0.57	<0	0.08	0.73	0.75
08-Jan-96							
Radar	0.91	0.63	0.96	0.82	0.97	0.87	0.86
SG7	0.94	0.71	0.93	0.68	0.98	0.57	0.91
SG1	0.94	0.41	0.98	0.36	0.95	0.76	0.83
16-May-95							
Radar	0.60	0.61	0.58	0.66	0.34	0.55	0.90
SG7	0.51	0.64	0.75	0.71	0.07	0.58	0.83
SG1	0.23	0.64	<0	0.83	<0	0.52	0.60
30-Jul-02							
Radar	0.75	0.85	0.54	0.55	0.74	<0	0.86
SG7	0.76	0.77	0.44	0.64	0.75	0.40	0.88
SG1	0.76	0.74	0.10	0.58	0.46	<0	0.78

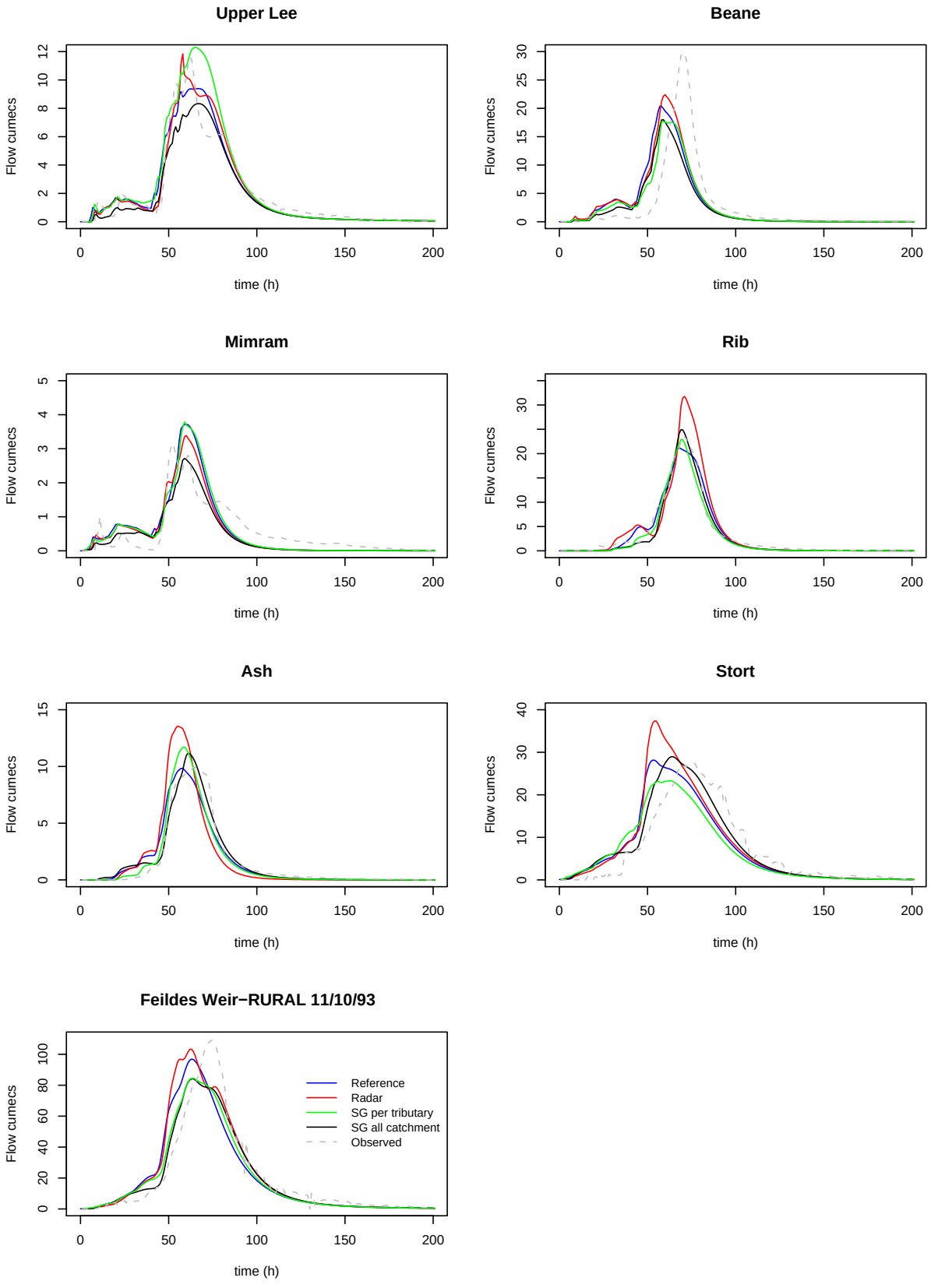


Figure 12.9 Observed, reference and simulated flows at existing catchment outlets on 11/10/93

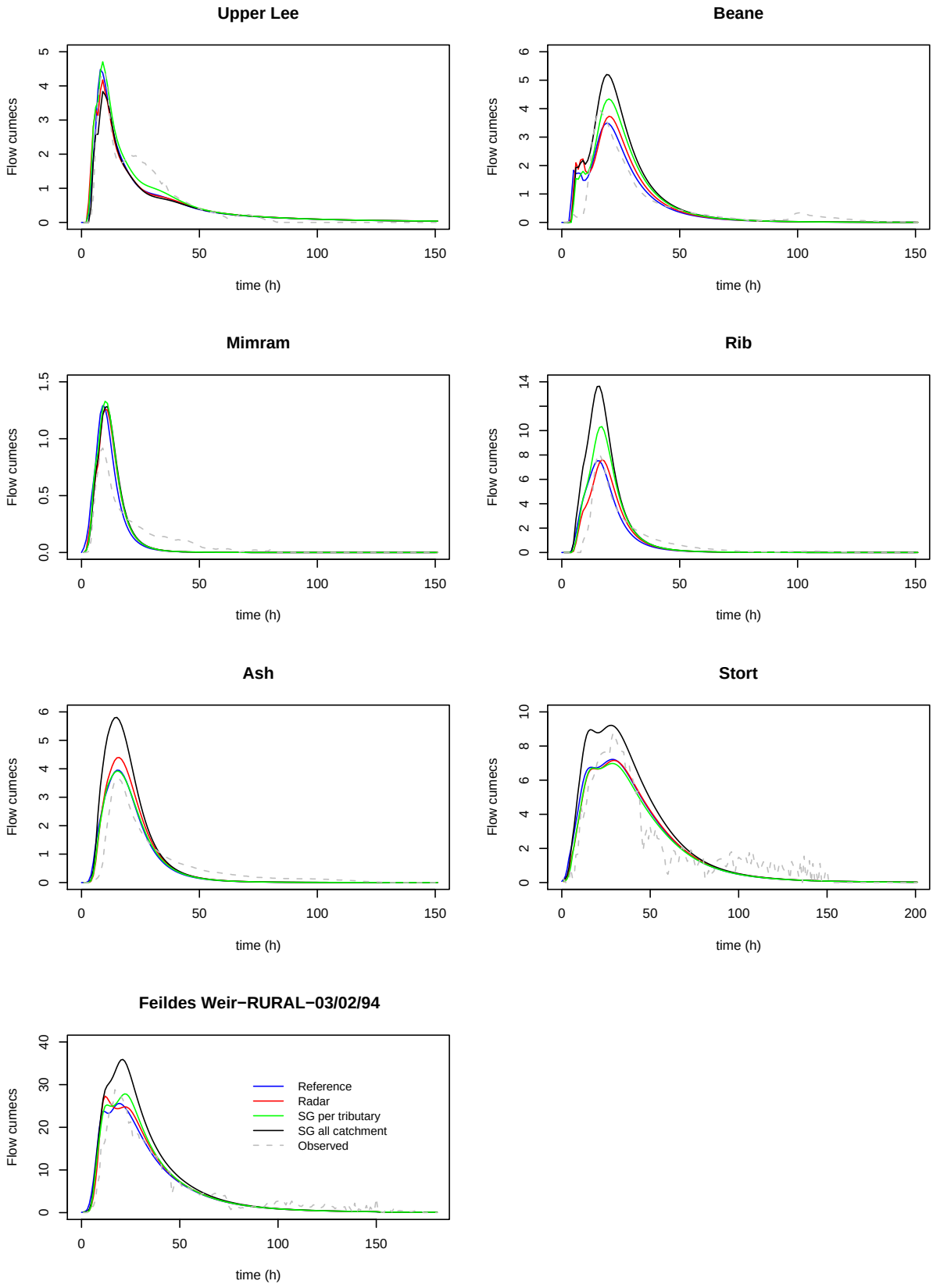


Figure 12.10 Observed, reference and simulated flows at existing catchment outlets on 03/02/94

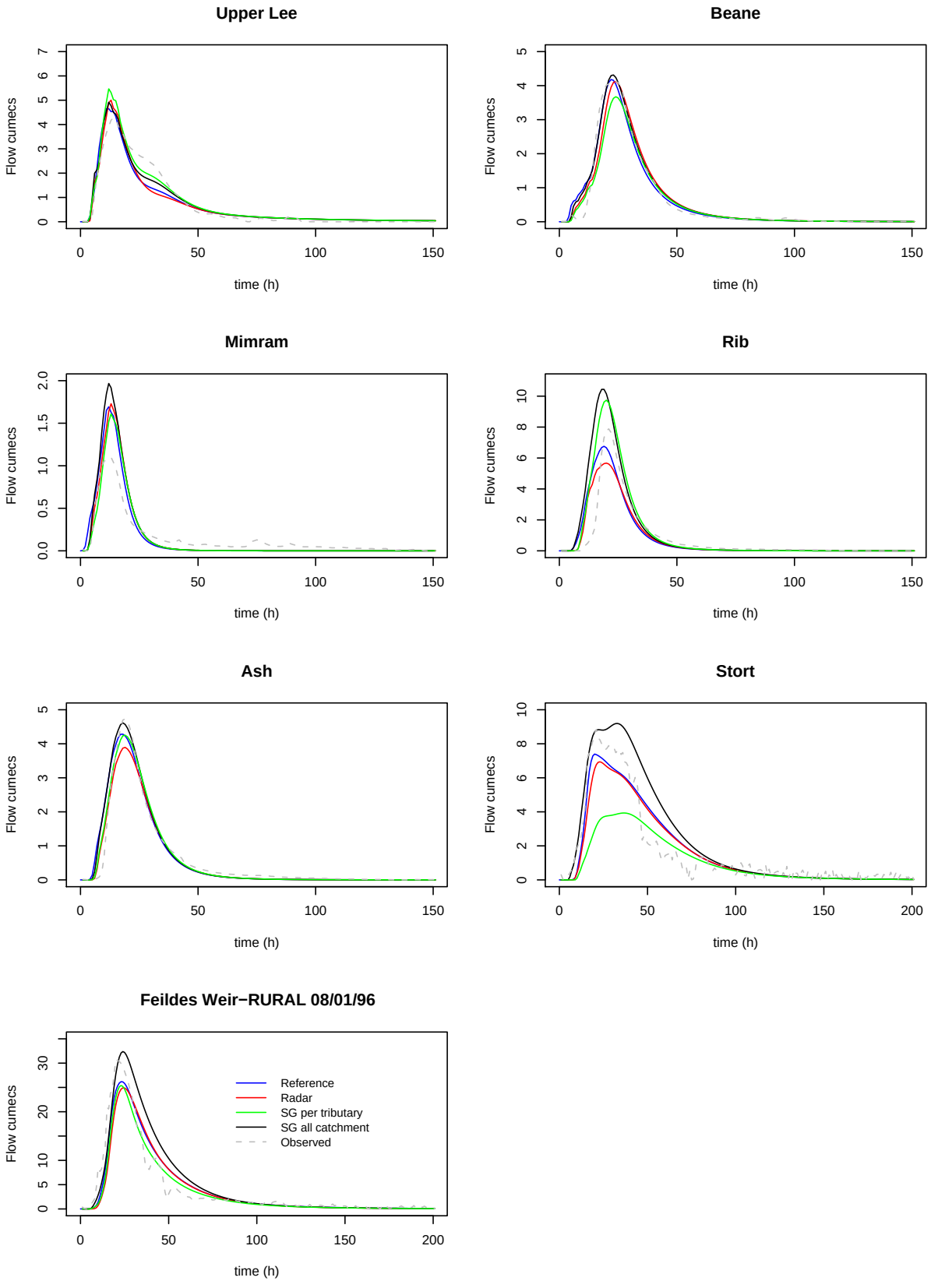


Figure 12.11 Observed, reference and simulated flows at existing catchment outlets on 08/01/96

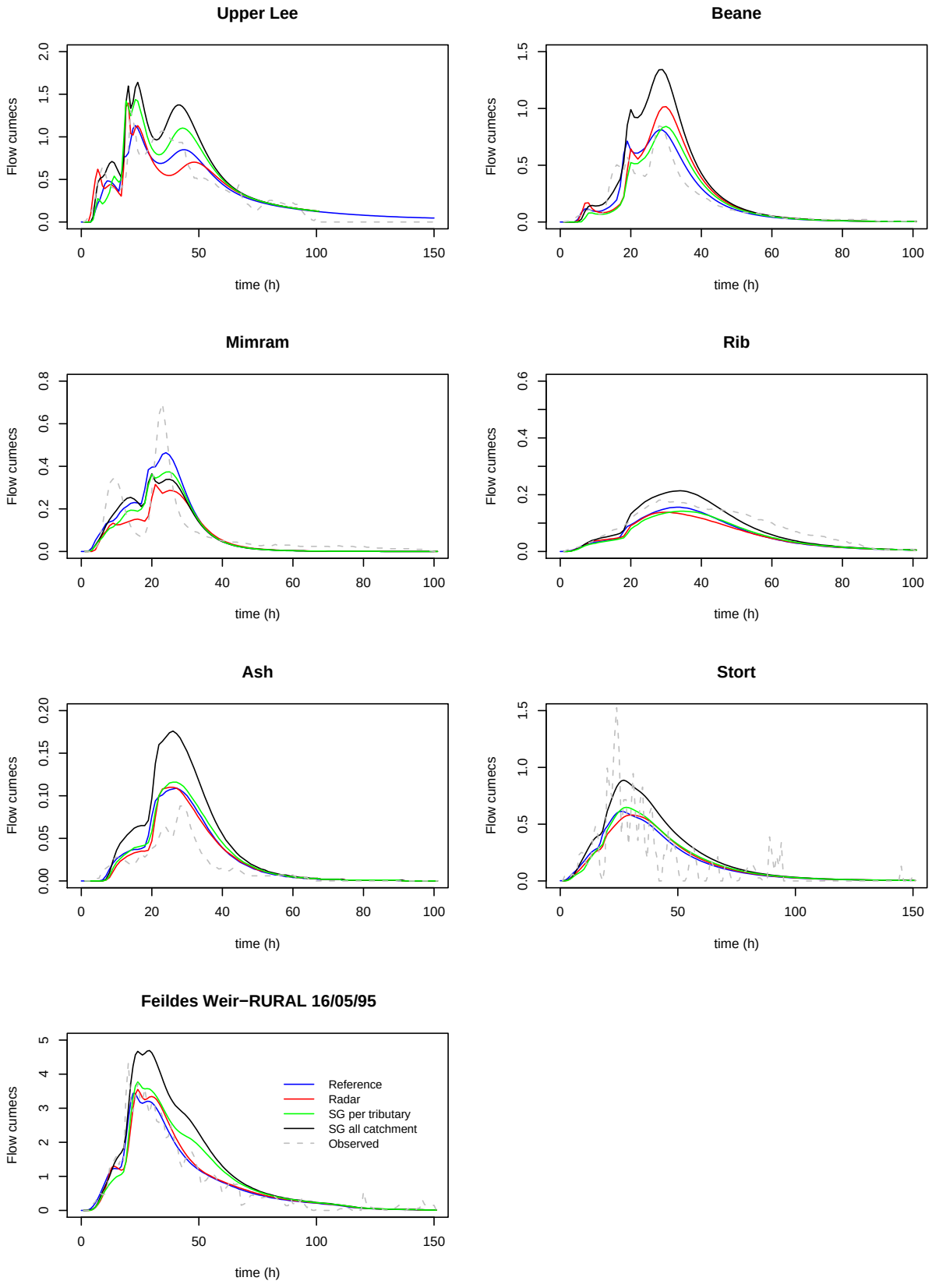


Figure 12.12 Observed, reference and simulated flows at existing catchment outlets on 16/05/95

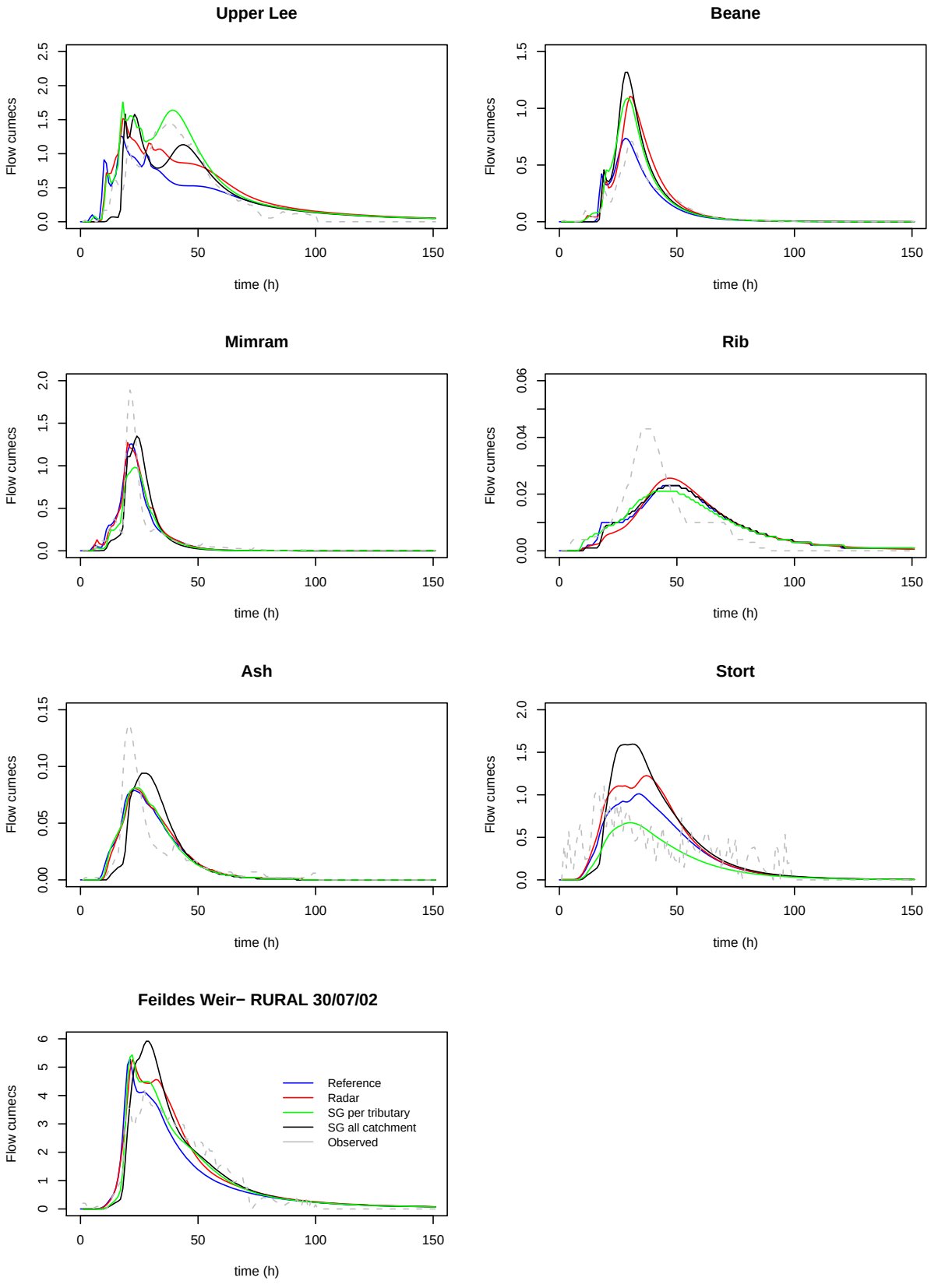


Figure 12.13 Observed, reference and simulated flows at existing catchment outlets on 30/07/02

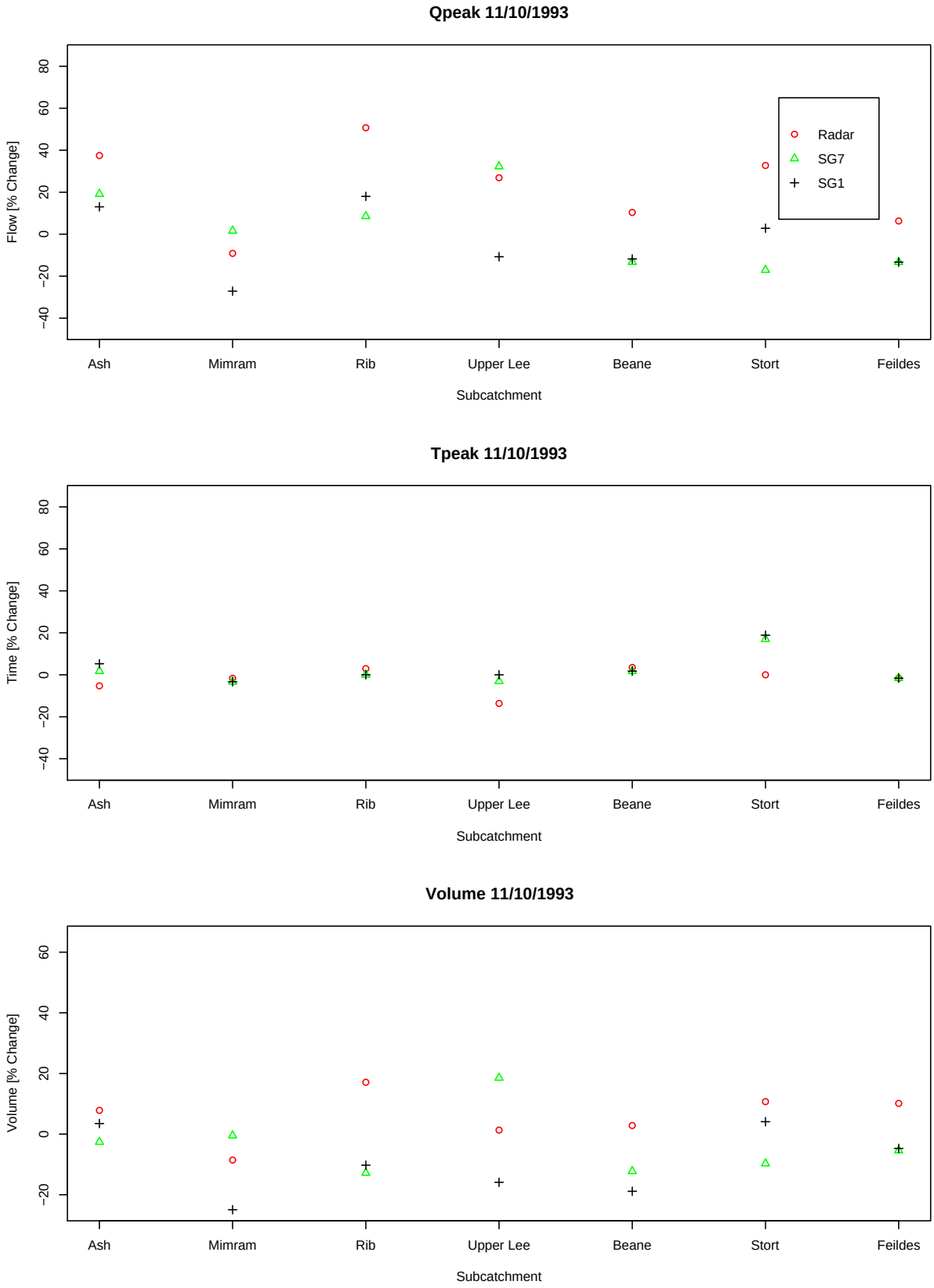


Figure 12.14 Percentage change in peak flow, time to peak and volume in comparison to the reference flow for the event of 11/10/93

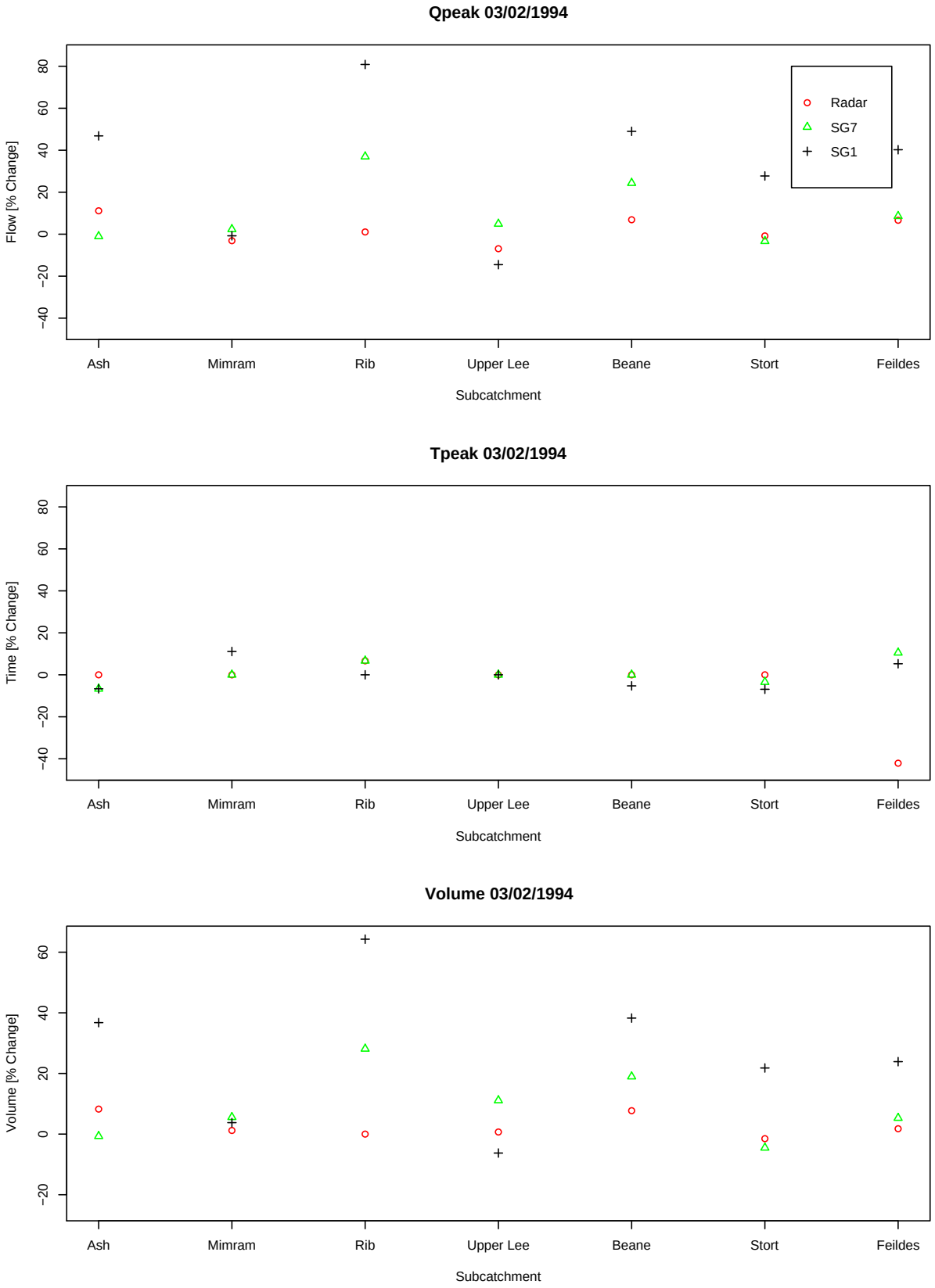


Figure 12.15 Percentage change in peak flow, time to peak and volume in comparison to the reference flow for the event of 03/02/94

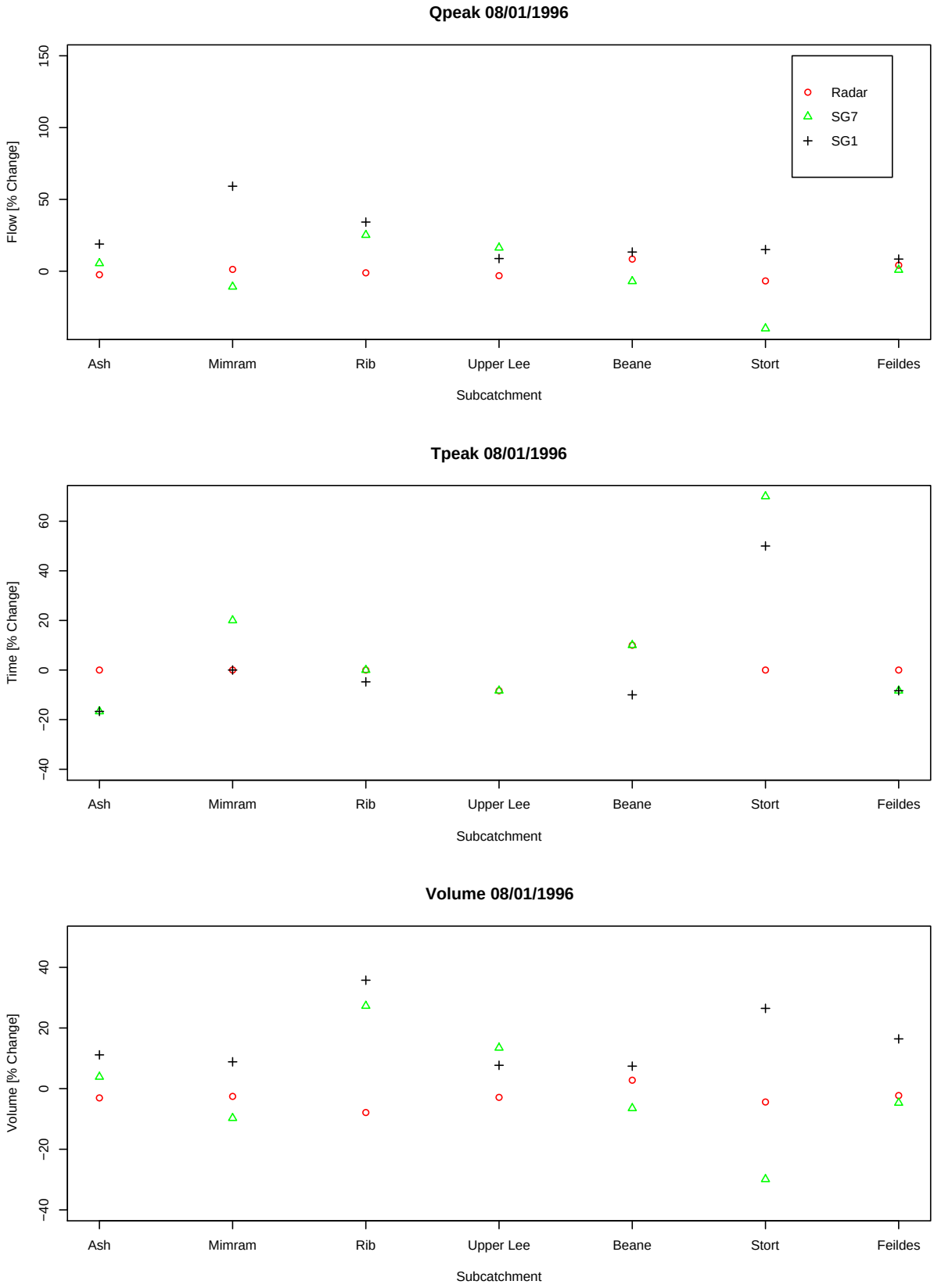


Figure 12.16 Percentage change in peak flow, time to peak and volume in comparison to the reference flow for the event of 08/01/96

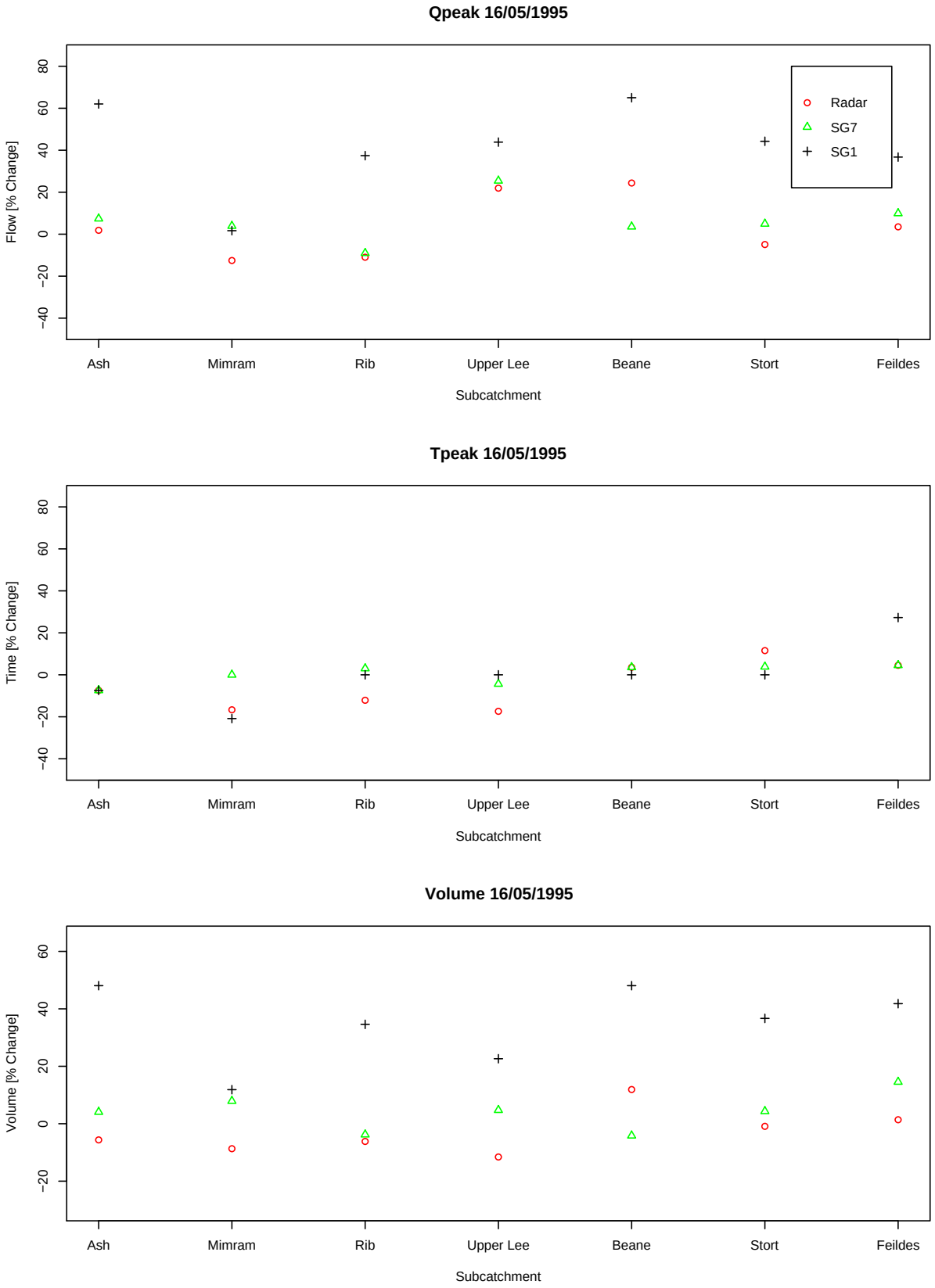


Figure 12.17 Percentage change in peak flow, time to peak and volume in comparison to the reference flow for the event of 16/05/95

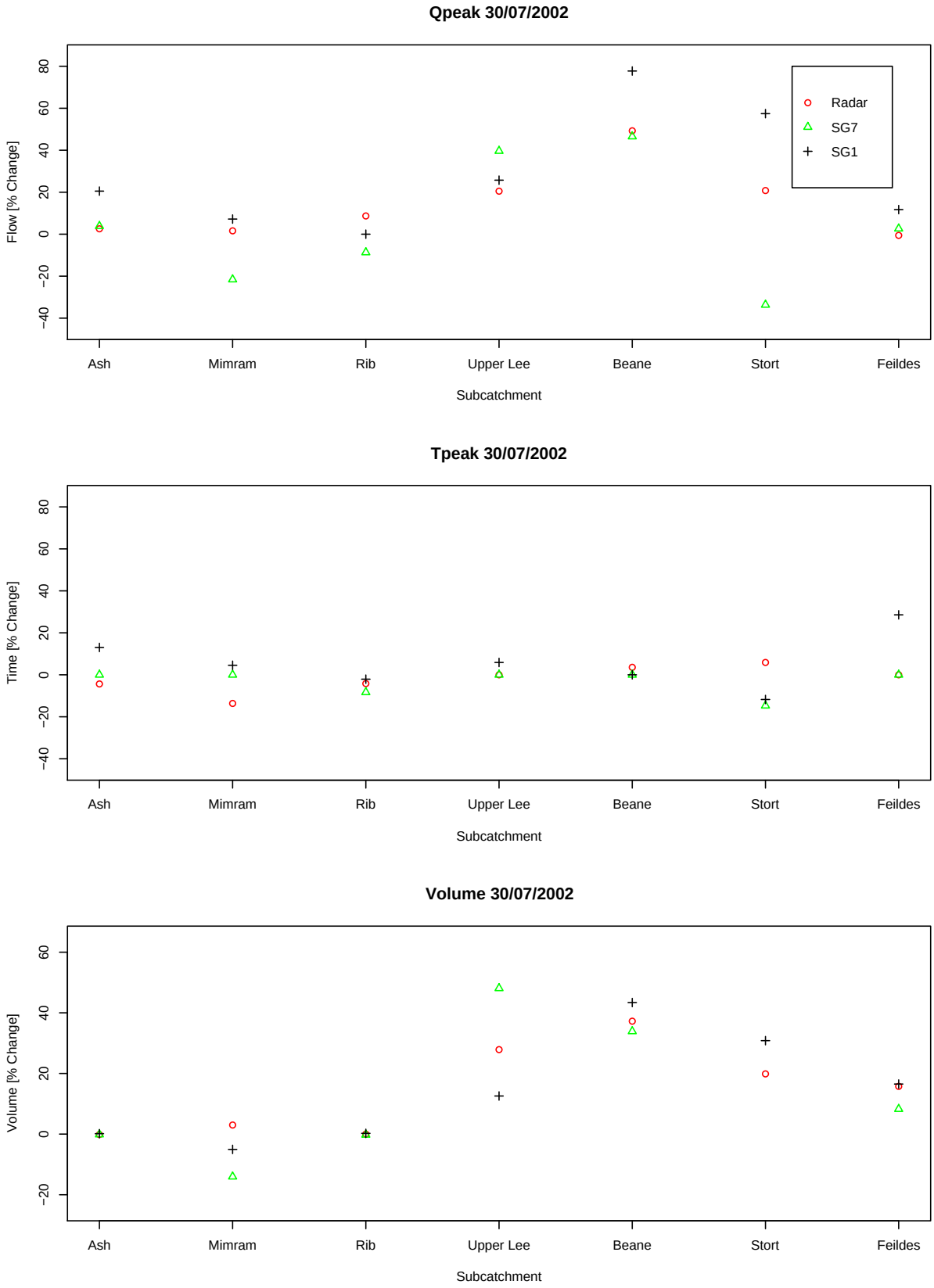


Figure 12.18 Percentage change in peak flow, time to peak and volume in comparison to the reference flow for the event of 30/07/02

12.7.1 Impact of spatial variability of rainfall

In this section, the impact of spatial variability of rainfall, as measured by SDIR, on the model performance is evaluated. Figure 12.19 plots the model performance in terms of NSE_{ref} against SDIR for all five events and all basins. A correlation between the increase in SDIR and the decrease in NSE_{ref} is observed. The top graph illustrates the impact of the three rainfall representations on model performance. Up to about 10% change in SDIR, the NSE_{ref} remains close to 1 for all rainfall scenarios. As SDIR increases, the NSE_{ref} remains high for radar data whereas it decreases for SDIR greater than 15% for SG7 and decreases more sharply for SDIR greater than 10% for SG1. The bottom graph lumps all rainfall representations but makes a distinction between winter and summer events. It can be seen that the NSE_{ref} performance is high (above 0.9) for winter events up to SDIR of 20% and remains above 0.7 up to SDIR of 30%. For summer events, the results are less consistent with a larger spread of NSE_{ref} values for SDIR greater than 10%. The graph in Figure 12.20 displays the change in peak flow against SDIR. Overall for all three rainfall representations, the errors in peak flow increase as the spatial variability of rainfall increases.

Hence, precipitation dominates the hydrological response and its patterns are dependent on the type of rainfall representations and storms. A SDIR greater than 10% affects the peak flow, above 20% it reduces the NSE_{ref} below 0.80.

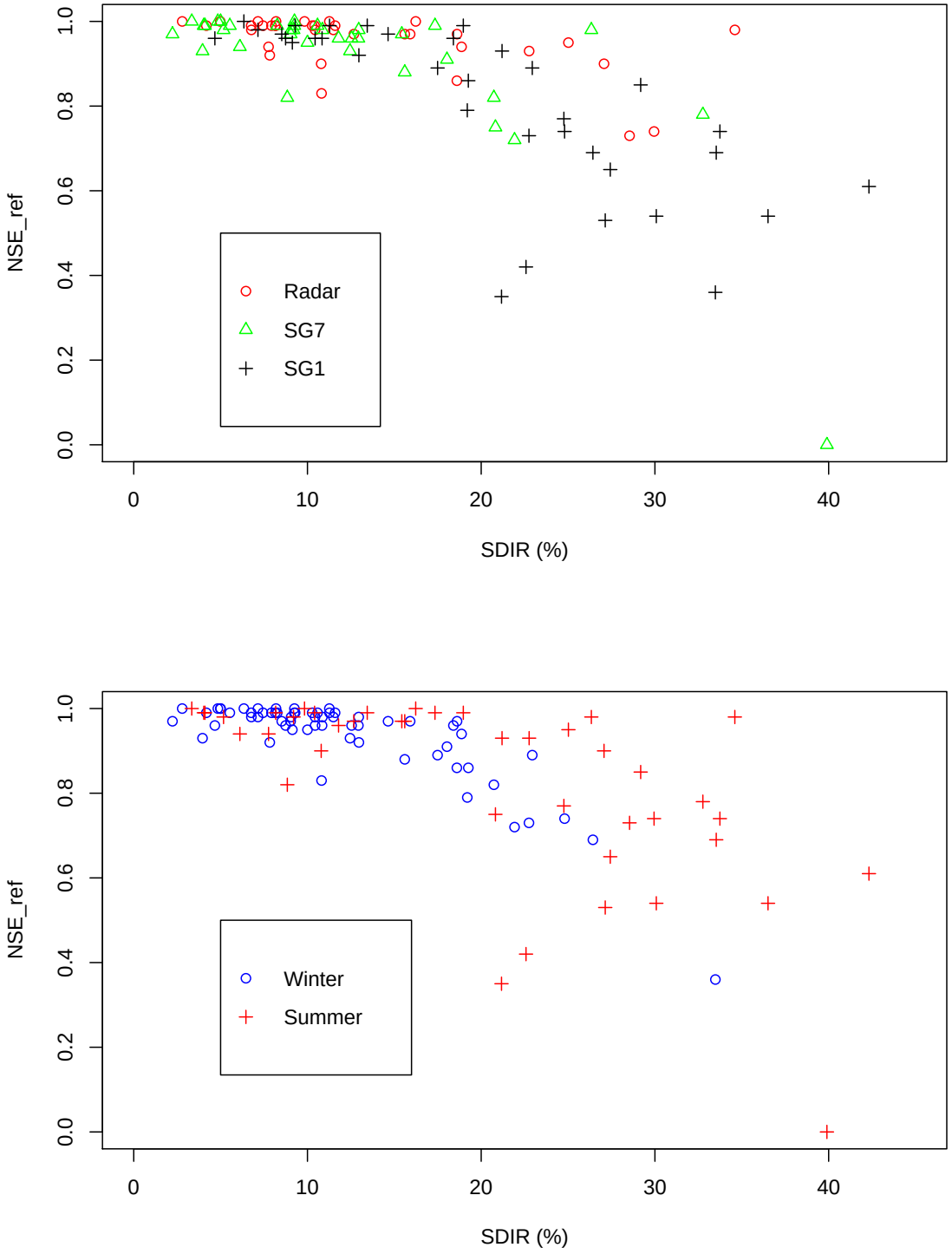
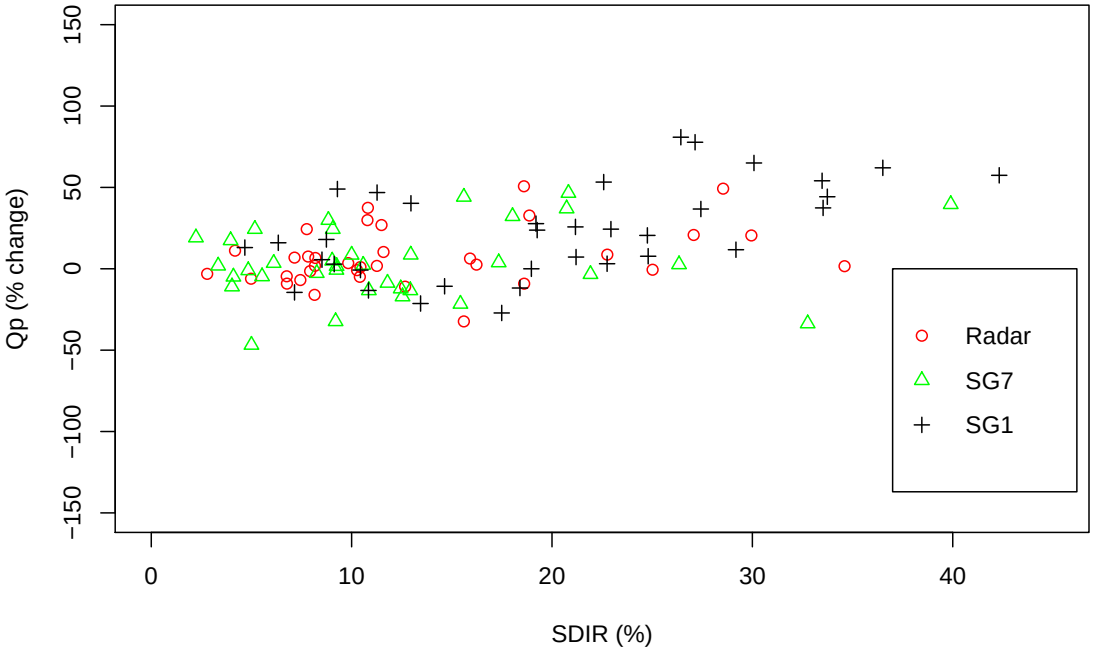


Figure 12.19 Impact of rainfall spatial variability on NSE_{ref}. Influence of rainfall representation (Top) and type of rainfall events (Bottom).



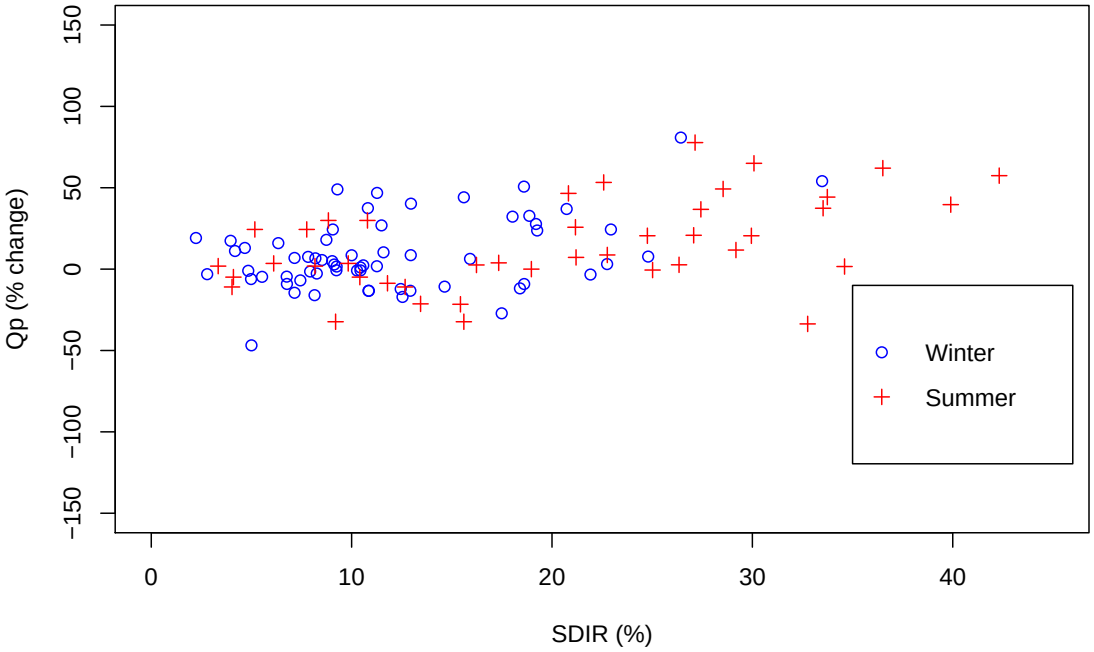


Figure 12.20 Impact of rainfall spatial variability on peak flow. Influence of rainfall representation (Top) and type of rainfall events (Bottom).

Table 12.18 Mean Lag Time (t_c) for existing catchments

	Mimram	Upper Lee	Beane	Ash	Feildes	Stort	Rib
t_c (h)	7.9	9.8	11.2	11.3	16.4	17.1	17.5
Area (km^2)	132	156	176	79	1033	278	134

12.7.2 Influence of catchment response time

The response time of each (sub)catchment is given in Table 12.18 and is plotted against catchment area in Figure 12.21. If the basins were homogeneous, we would expect the catchment response time to increase with the scale of the basin. For instance, Berne et al. (2004) fitted a power law relationship between the lag time and surface area for 6 urban sub-basins of 38 ha to 105 km^2 in the Southwest of France. However in this case, the basins are distinct enough and no obvious trend between the response time and area is evidenced. Hypotheses to explain the drivers influencing the catchment response time are presented below. The reader is also referred to Table 12.1, 12.2 and 12.3, which present the hydrogeological characteristics of the basins.

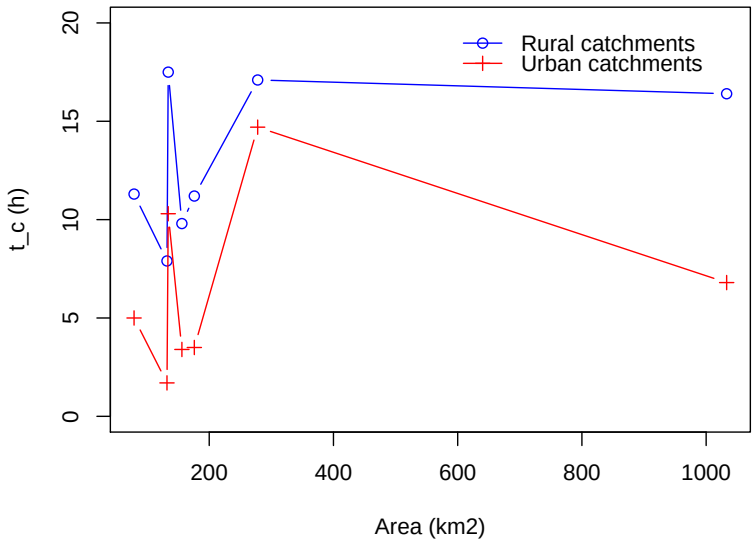


Figure 12.21 Catchment response time versus area

The Mimram (132km^2) is a predominantly rural area. It is an elongated chalk catchment with small outlying patches of London Clay bedrock with boulder clay drift deposits, however these are too small to have any significant influence on the runoff (Hawnt, 1987b). About 13% of the area is developed with the major conurbation of Welwyn Garden City. Even though this is not described as a dense urban area, most of the direct runoff is generated from Welwyn Garden City. From this subcatchment, a fast-response (t_c of 7.9 hours) is expected since the distance from Welwyn Garden City to the outlet is small (about 7 km) but damped due to the influence of the chalk geology. The Upper Lee (156 km^2) has similar geological characteristics to the Mimram (predominantly chalk). However because the urban extent is higher (35%),

more runoff is generated in comparison to the Mimram with the majority of the runoff coming from Luton. A relatively fast response (t_c of 9.8 hours) is expected from this catchment because of the contribution of urban runoff, but slower than the Mimram since the town of Luton is further upstream from the outlet (23 km). The Beane (176 km^2) is mainly rural with 13% of urban extent around the town of Stevenage. The catchment is comprised of chalk bedrock overlaid by superficial deposits of boulder clay. A slower response (t_c 11.2 hour) than that of the Upper Lee is expected since the urbanisation is less, but the larger extent of the impermeable clay layer may generate a higher amount of runoff

The 79 km^2 Ash catchment is characterised as rural (3.5% urban extent). It is predominantly chalk with the south eastern watershed boundary marking the interface with London Clay deposits (Hawnt, 1987a). It is overlaid by a particularly impermeable layer of boulder clay which generates significant runoff. Due to the small catchment size, a relatively fast-response (t_c of 11.3 hours) is expected from this catchment (but not as fast-responding as the above catchments with higher urban extent). A high response is expected due to the proportion of runoff generated from the clay contribution (i.e. not much water is absorbed into the chalk) .

The response at Feildes Weir outlet (1033 km^2) may be slightly faster than some of the more rural catchments (See the Rib below) as there is a higher area covered by urbanisation (15%) which would generate more runoff. However due to the larger size and the heterogeneity of the catchment, a relatively low response is expected.

The Stort catchment is more complex geologically. The underlying, or solid geology is composed mainly of London Clay with chalk in the north-west part of the watershed and is overlaid by boulder clay drift deposits. 13% of the watershed is urbanised. The catchment area of the Stort is 278 km^2 which makes it the second largest of the catchments mentioned which increases the response time. It is more urban than the Rib so that might explain its quicker response. It may be slightly less than that at Feildes Weir due to the influence of man made constructions (reservoir and navigation system). There is likely to be runoff in the south of the catchment where the chalk is confined but also in the north because of drift deposits.

The Rib (134 km^2) is a mixed chalk and boulder clay catchment, predominantly rural (3.5% urban extent). The impermeable boulder clay has a marked influence on the runoff characteristics and generates a significant amount of runoff. There is a major baseflow component to the runoff hydrograph due to percolation of effective rainfall falling directly on exposed (or relatively close to the surface) chalk surfaces. Direct runoff from upstream clay areas can infiltrate, for example when flowing over previously dry stream beds at the start of heavy rainfall (Hawnt, 1987a), (i.e. the water is absorbed, there is no runoff along the dry stream beds, initially, before a stream starts to flow). These location effects introduce a marked lag in the runoff response and high rainfall losses. Hence a slow response occurs (t_c of 17.5 hours) and a high response due to the extent of the impermeable boulder clay deposit.

Figure 12.22 displays the results in terms of NSE_{ref} (radar data, SG7 and SG1) for the five selected events and for each subcatchment ranked according to their characteristic time. The top graph presents the results for each rainfall representation whereas the bottom graph lumps all rainfall representations. In total, there are 15 observations per subcatchment. The boxplot represent the median, upper and lower

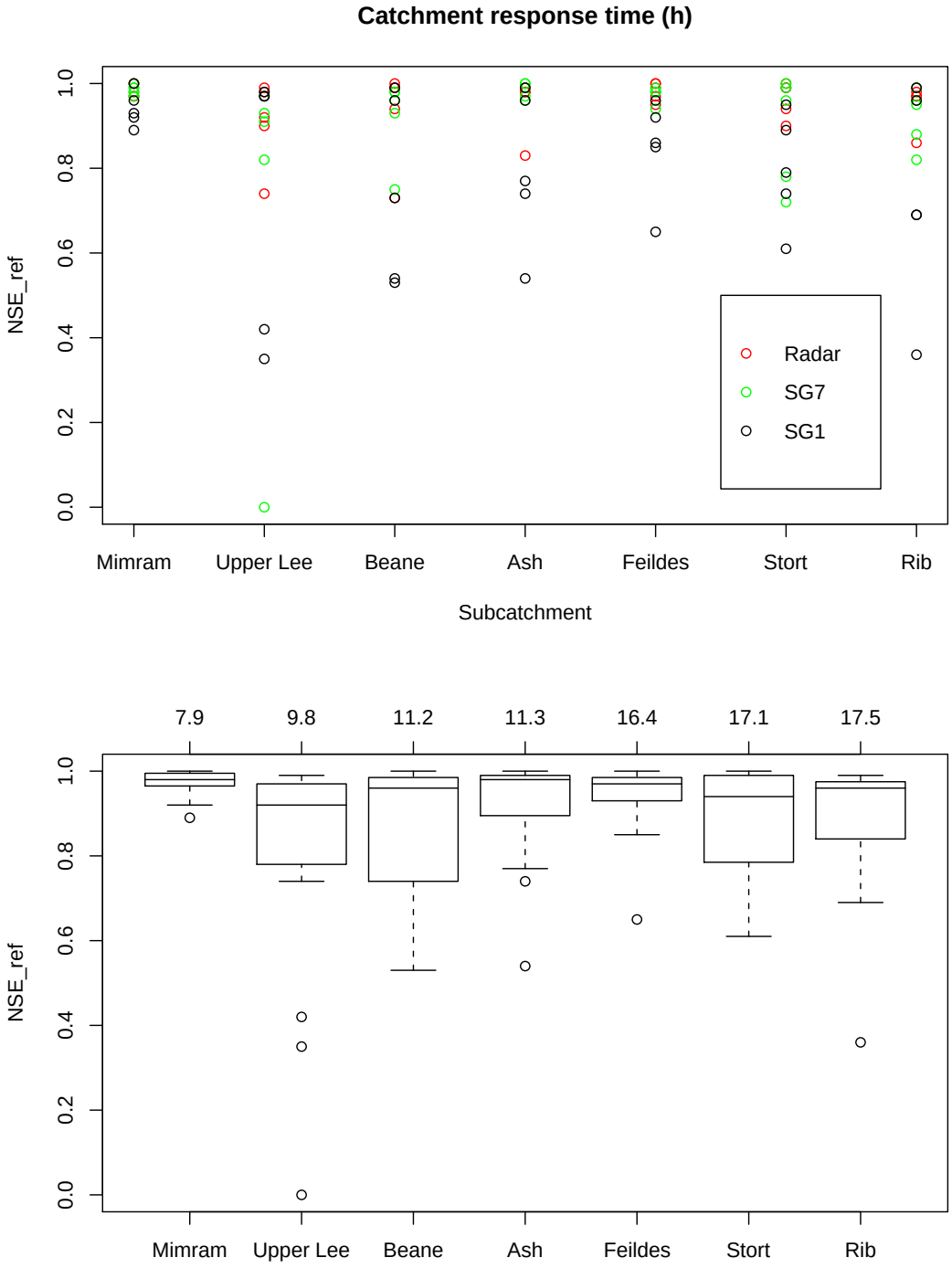


Figure 12.22 Boxplots of NSE criteria for each subcatchment ranked according to their catchment response time (expressed in hour on the upper axis)

quartile, the whiskers extend to the NSE value within 1.5 times the interquartile range from the box. As highlighted in the review of the literature, contrasting observations arise from the studies in humid areas. No clear pattern emerges. It seems that the Mimram and the Lee catchment to Feildes Weir are less sensitive to spatial rainfall with NSE values close to 1, which highlights the damping effect due to the chalk geology of the Mimram and the integrating effect at the large catchment scale for the Lee. This confirms that detailed knowledge of spatial rainfall is not required for basins characterised by a high level of dampening of the rainfall input signal (Obled et al., 1994; Smith et al., 2004; Woods and Sivapalan, 1999; Naden, 1992). In conclusion, this case-study indicates that the dominant controls on the catchment response time are:

- The extent of runoff contributing areas (i.e. the extent of urbanisation and boulder clay drift)
- The location of these (urban) areas, with flow distance to the outlet
- The scale of the catchment

Attention should be focused on these interactions when modelling sub-humid basins in temperate regions. Altitude may be of importance in explaining runoff in areas with more relief.

Woods and Sivapalan (1999) proposed an analytical framework to assess the relative importance of these interactions. They defined the mean catchment runoff time (i.e. a characteristic catchment response time defined as the lag time between the beginning of the storm event to the centre of mass of the hydrograph) as folows:

$$E(T_q) = E(T_r) + E(T_h) + E(T_n) \tag{12.13}$$

where $E(T_q)$ is the mean catchment runoff times, $E(T_r)$ is the mean holding time for rainfall excess (i.e. the holding time for runoff to be generated), $E(T_h)$ is the mean holding time of hillslope travel and $E(T_n)$ is the mean holding time of network travel. An estimate of these holding times would give an insight of the dominant controls on the mean runoff time.

12.7.3 The effect of scale

Results displayed in Figure 12.23 are similar to those presented in Figure 12.22, but with the subcatchments ranked by size (upper axis in km^2). The graph is difficult to interpret, it may be that the sensitivity of the basin to variation in spatial rainfall increases up to the scale of the Beane ($\simeq$ 175 km^2), then decreases with scale as the catchment presents more spatial heterogeneity. At the catchment scale (1040 km^2), the integrating effects of the catchment result in less sensitivity to variation in spatial rainfall with NSE$_{ref}$ values mostly comprised between 0.8 and 1. However lower performance is reached with SG1 in the case of the convective summer event (NSE$_{ref}$=0.65), which indicates that knowledge of spatial rainfall is important at the catchment scale even though a detailed representation is not necessary. For instance, Dodov and Foufoula-Georgiou (2005) use catchment-averaged rainfall over subcatchments of 530 km^2 (9 subdivisions scheme) and 1600 km^2 (3 subdivisions scheme) for a 4800 km^2 basin, and subcatchments of 735 km^2 (10 subdivisions scheme) and 2450 km^2 (3 subdivisions scheme) for a 7300 km^2 basin in central North America. In our case, if the interest is in modelling the catchment outlet only, a network of 7 raingauges seems adequate. This is in agreement with the findings of Andréassian et al. (2001) who recommend a density of 8 raingauges at the 1120 km^2 scale (using lumped modelling).

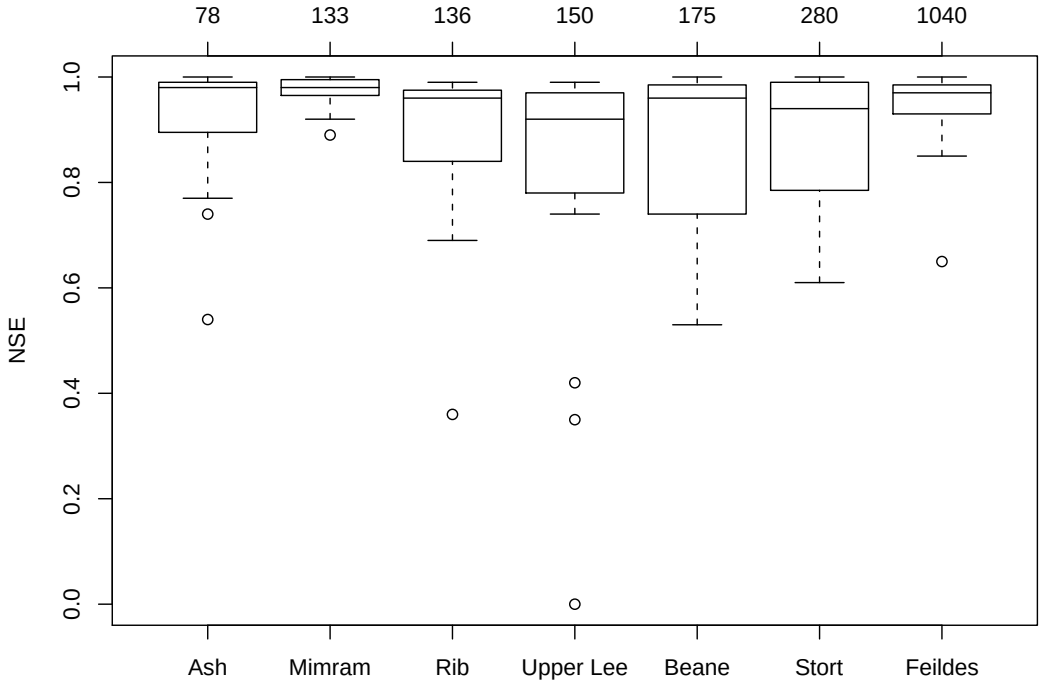


Figure 12.23 Boxplots of NSE$_{ref}$ criteria for each subcatchment ranked according to their size

12.7.4 Rainfall variability, catchment characteristics and model performance

As the model performance is most sensitive with SG1, this rainfall representation is selected to emphasise and assess the relative importance of SDIR, catchment scale and catchment response time. Figure 12.24 represents NSE_{ref} versus SDIR for SG1. The top plot highlights the impact of catchment scale with light to dark colours representing small to large catchments. Similarly, the bottom plot highlights the impact of catchment response time with light to dark colours representing slow to fast catchments. These plots do not indicate any obvious relationships between catchment scale and response time on NSE_{ref}, however the relative importance of SDIR is confirmed with lower NSE_{ref} values as SDIR increases. As seen in Section 12.7.1, SDIR is a function of rainfall representation and storm type.

12.7.5 Conclusion

Results on existing, predominantly rural, catchments reveal that the basins are sensitive to the type of rainfall events, rainfall representation, catchment scale and catchment type. The findings show a clear relationship between the increase in spatial variability of rainfall and the decrease in model performance. More variability is introduced when extreme or summer events occur. Spatially variable rainfall and radar data lead to similar flow response whereas larger discrepancies are introduced when using input data from a subset of the raingauges. Results are worse when using a single gauge to model the whole catchment.

A higher amount of runoff is generated on urban and clay areas whereas the chalk is baseflow dominated. Urban areas speed up the response. Rural basins are more sensitive to the antecedent catchment conditions and generate lower flows in summer than basins incorporating zones of urban development (see the Rib and Ash response on 16th May 1995 and 30th July 2002). Less sensitivity is observed at the catchment scale. Thus the integrating effect of the catchment is due to the extent of runoff contributing areas, their location with respect to the catchment outlet and the catchment scale. These interactions are incorporated in the catchment characteristic time, however it was not possible to assess their impact independently.

12.8 Urban Catchment Results

The effect of urbanisation on runoff is due to (i) increase in impervious area which increases the amount of runoff, and (ii) hydraulic improvement of flow paths (e.g. roofs, roadways, gutters, pipes, channel) speeds up the response. RORB handles (i) by changing the percent impervious. The value input for each sub-area is the proportion of impervious surface which connects in some way to the drainage system. The hydraulic improvement (ii) is dealt with using the Reach Type Flag (Laurenson and Mein, 1988) as in Equation 12.7. The subcatchments are turned artificially into completely urban basins by applying a fraction imperviousness of 30% to all subareas, and with the exception of the Stort, by changing the reach type from Natural to Lined

(or Piped) and a slope of 10% is applied according to Equation 12.8. The same rainfall representations are tested and the simulated runoff is compared to the reference flow. Table 12.20 presents the results in terms of NSE_{ref} for urban catchments. Figures 12.25 to 12.29 show that urban catchments are faster-responding than the existing natural basins. The hydrograph is characterised by a narrow shape and a high peak flow. Further, the response on urban catchments is sensitive to the temporal distribution of rainfall, with bursts of rainfall reflected in the hydrograph shape. The average NSE_{ref} for all catchments and events is 0.95, 0.92 and 0.73 for radar, SG7 and SG1, respectively. In comparison to rural basins, the performance is maintained for radar data and SG7 but decreases with SG1. As mentioned by Young (2005), the NSE statistic is biased towards the fit of high flows and tends to increase as the length of simulation period decreases. However these results serve as a first point of comparison and are complemented by the percentage change in Q_p, T_p, V in Table 12.19. In comparison to rural catchments, the error in Q_p and T_p decreases for radar data (9.7 and 5.4% for urban catchments compared to 12.6 and 5.6% for rural catchments) and increases for SG7 (16.5 and 9.8% for urban catchments compared to 15.3 and 5.8% for rural catchments) and SG1 (38.5 and 12.7% for urban catchments compared to 28.1 and 8% for rural catchments). However errors in hydrograph volume increase slightly for radar data (8% for urban catchments compared to 7.4% for rural catchments) an decreases for SG7 (11.1% for urban catchments compared to 12.4% for rural catchments) and SG1 (18.4% for urban catchments compared to 21.6% for rural catchments). Figure 12.30 to Figure 12.34 display the percentage change in Q_p, T_p and V for all rainfall representations and all subcatchments ranked by size. The event of 30th July 2002 leads to more variation in the catchment response with a percentage peak change up to 120% compared to 80% for rural catchments.

Table 12.19 Percentage change of peak flow, time to peak and hydrograph volume

	Radar	SG7	SG1
Q_p (%)	9.7	16.5	38.5
T_p (%)	5.4	9.8	12.7
V (%)	8.0	11.1	18.4

Table 12.20 NSE_{ref} Urban catchments

	Upper Lee	Mimram	Beane	Rib	Ash	Stort	Feildes
11-Oct-93							
Radar	0.97	0.92	0.98	0.81	0.95	0.97	0.98
SG7	0.93	0.95	0.97	0.90	0.89	0.98	0.99
SG1	0.95	0.93	0.89	0.92	0.92	0.88	0.98
03-Feb-94							
Radar	0.97	0.97	0.99	1	1	1	1
SG7	0.95	1	0.96	0.94	0.98	1	0.99
SG1	0.98	0.98	0.83	0.62	0.86	0.86	0.93
08-Jan-96							
Radar	0.99	0.99	0.92	0.99	1	1	1
SG7	0.95	0.79	0.93	0.91	0.99	0.99	0.99
SG1	0.98	0.85	0.98	0.75	0.92	0.89	0.96
16-May-95							
Radar	0.98	0.99	0.98	0.93	0.99	0.97	0.99
SG7	0.93	0.99	0.96	0.97	0.96	0.94	1
SG1	0.82	0.77	0.69	0.66	0.40	0.81	0.83
30-Jul-02							
Radar	0.91	0.96	0.59	0.80	0.98	0.90	0.94
SG7	0.47	0.72	0.76	0.68	0.82	0.95	0.94
SG1	0.16	0.72	0.01	0.04	0	0.15	0.46

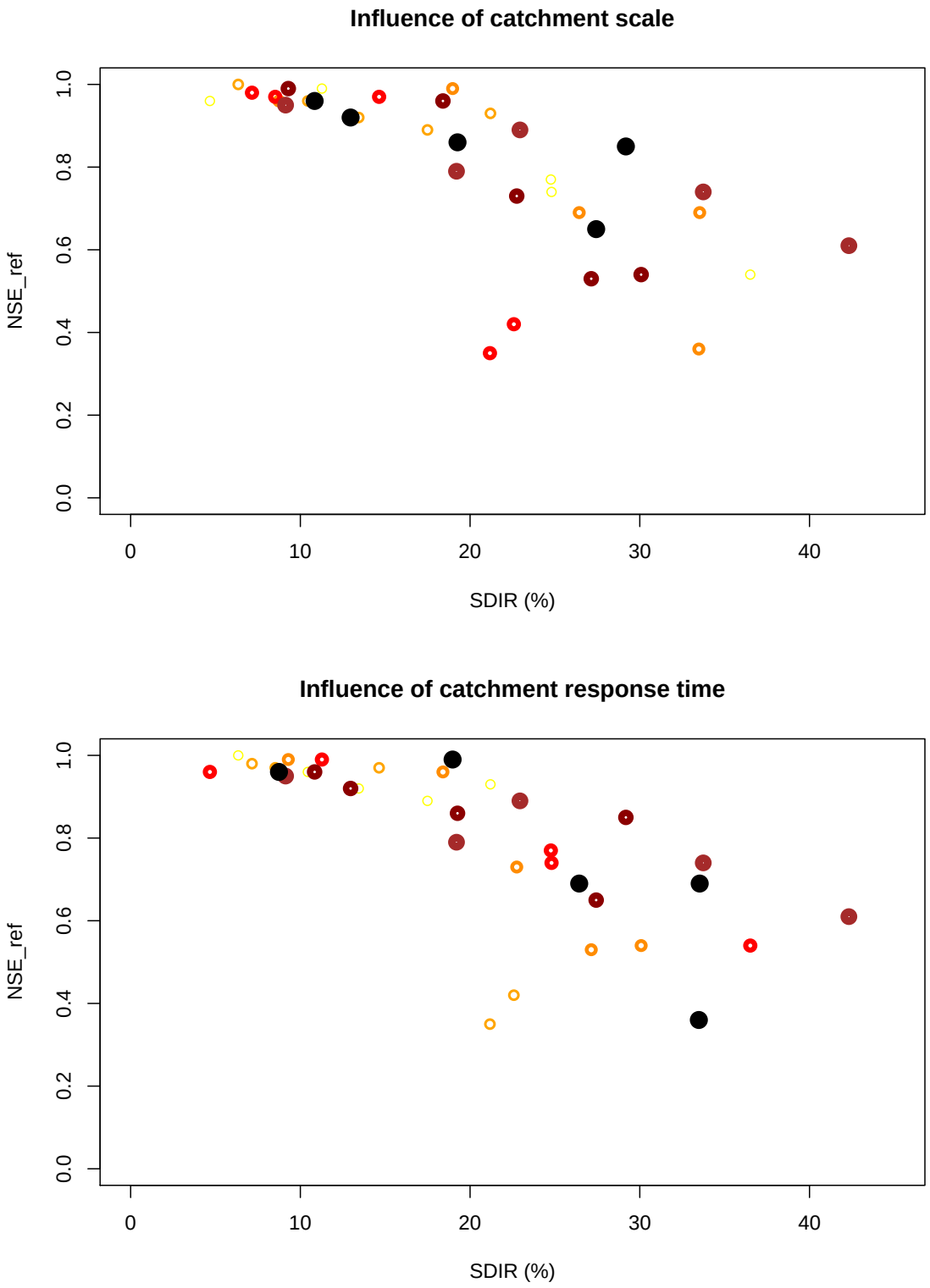


Figure 12.24 SG1: Model performance vs Rainfall variability highlighting the effect of Scale (Top) and Response time (Bottom)

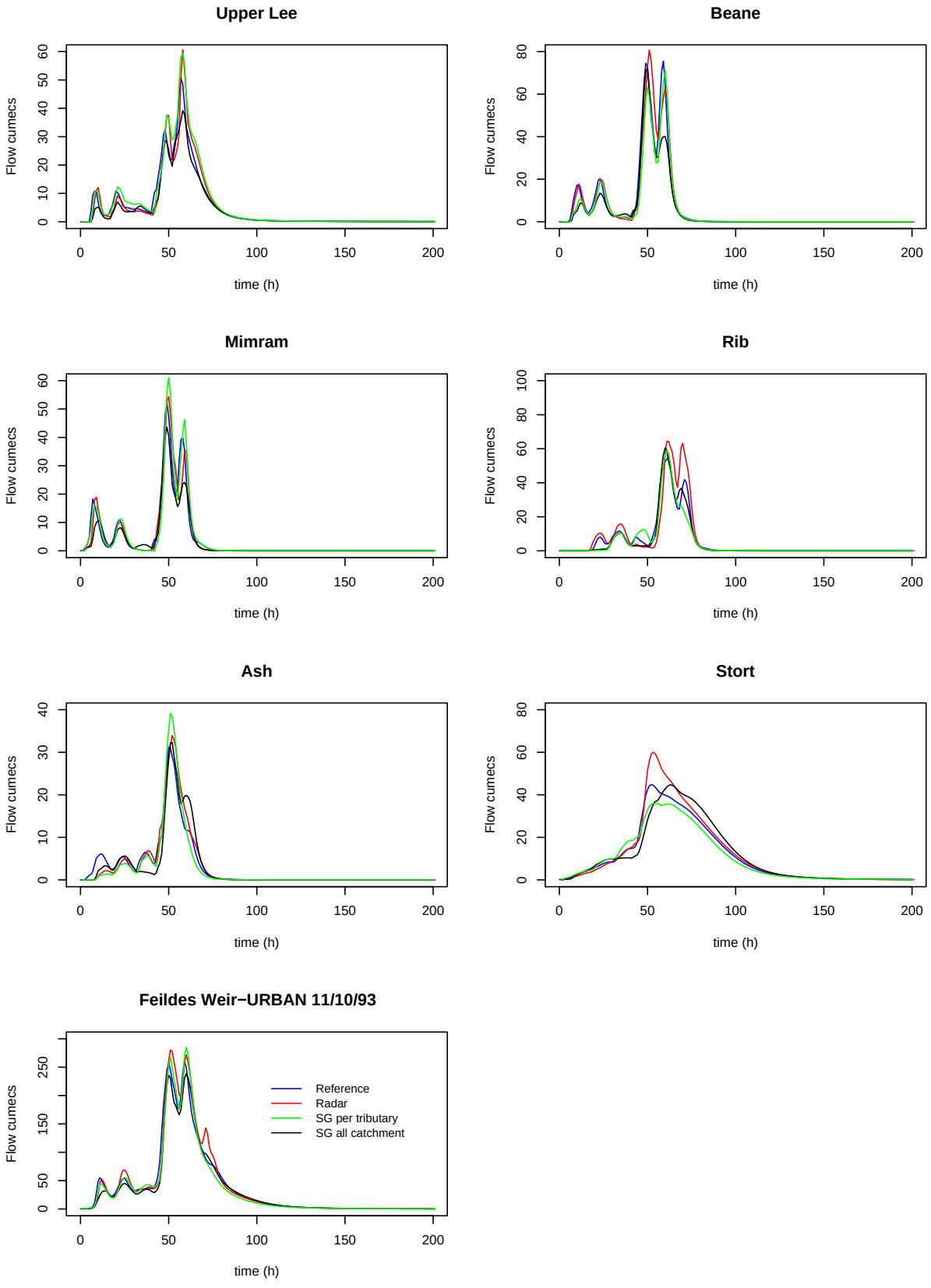


Figure 12.25 Reference and simulated flow at urban catchment outlets on 11/10/93

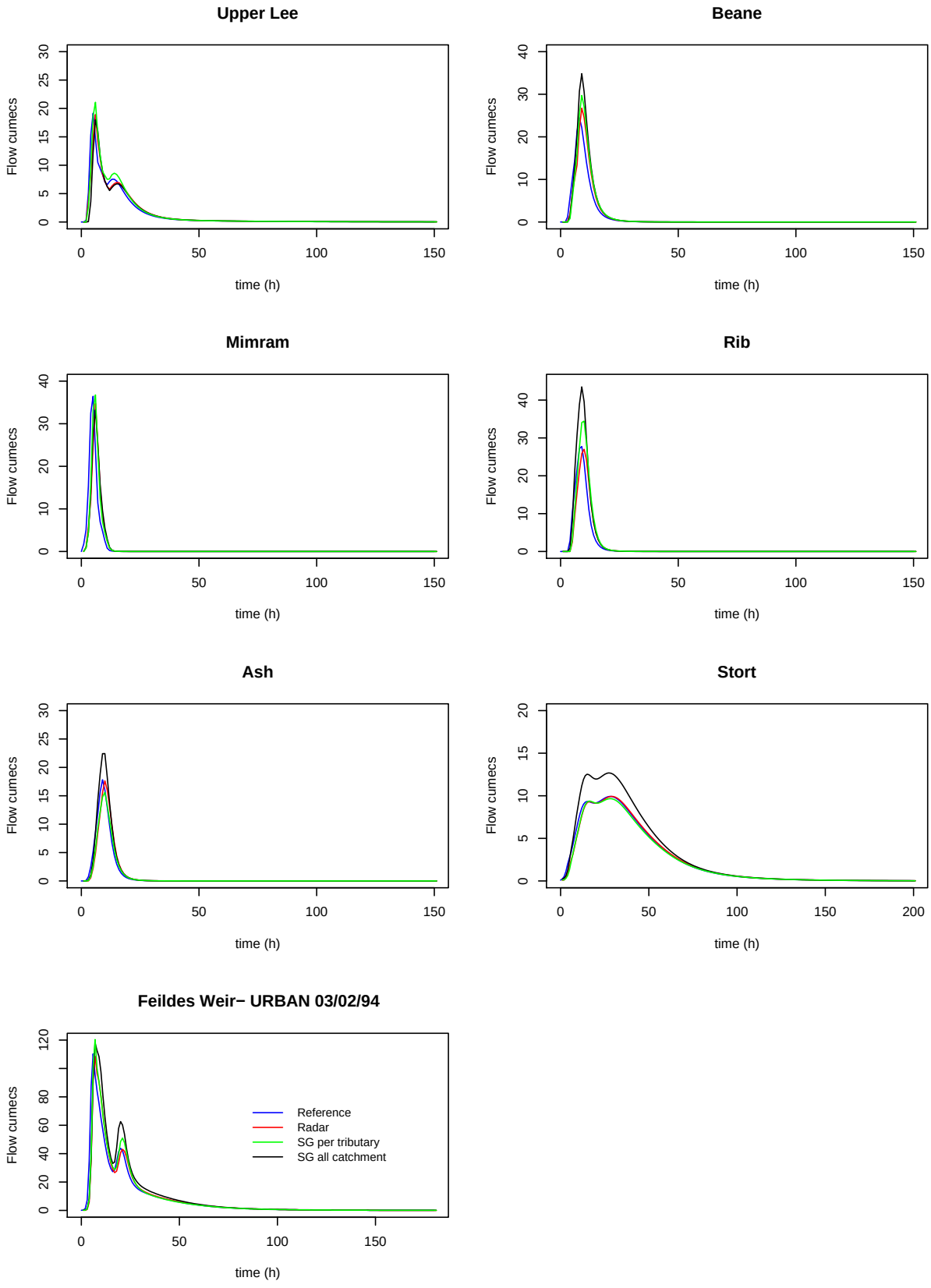


Figure 12.26 Reference and simulated flow at urban catchment outlets on 03/02/94

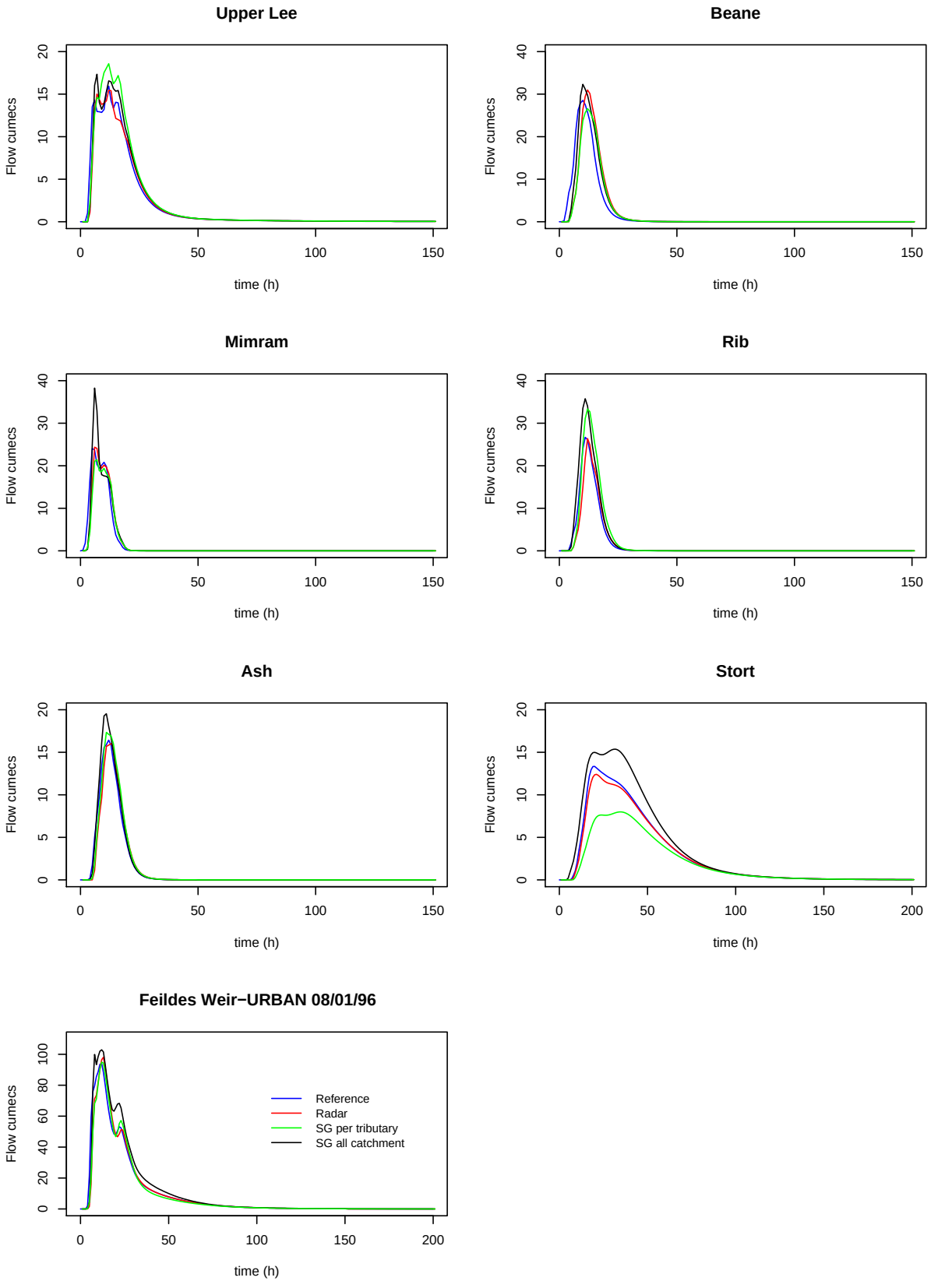


Figure 12.27 Reference and simulated flow at urban catchment outlets on 08/01/96

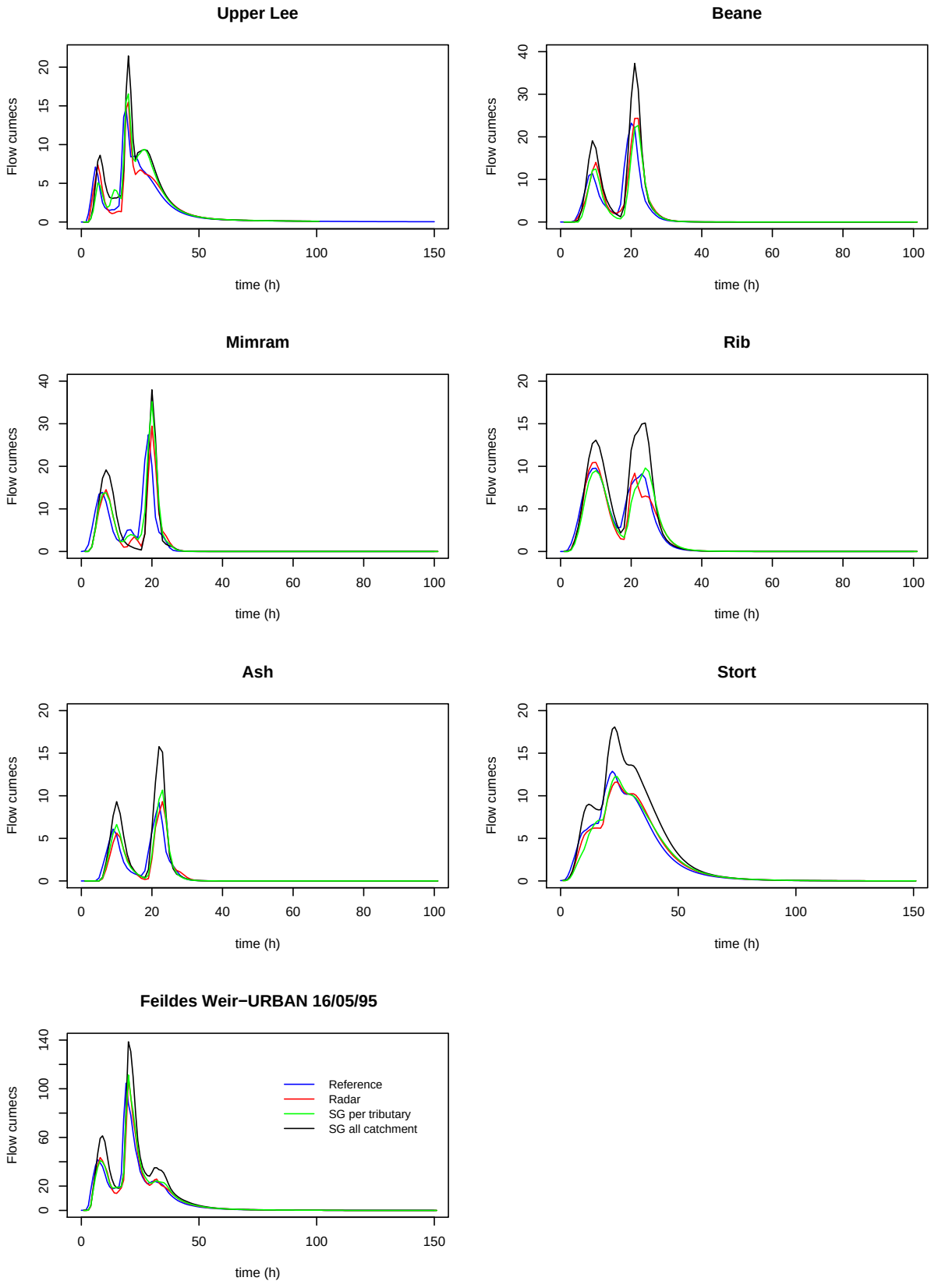


Figure 12.28 Reference and simulated flow at urban catchment outlets on 16/05/95

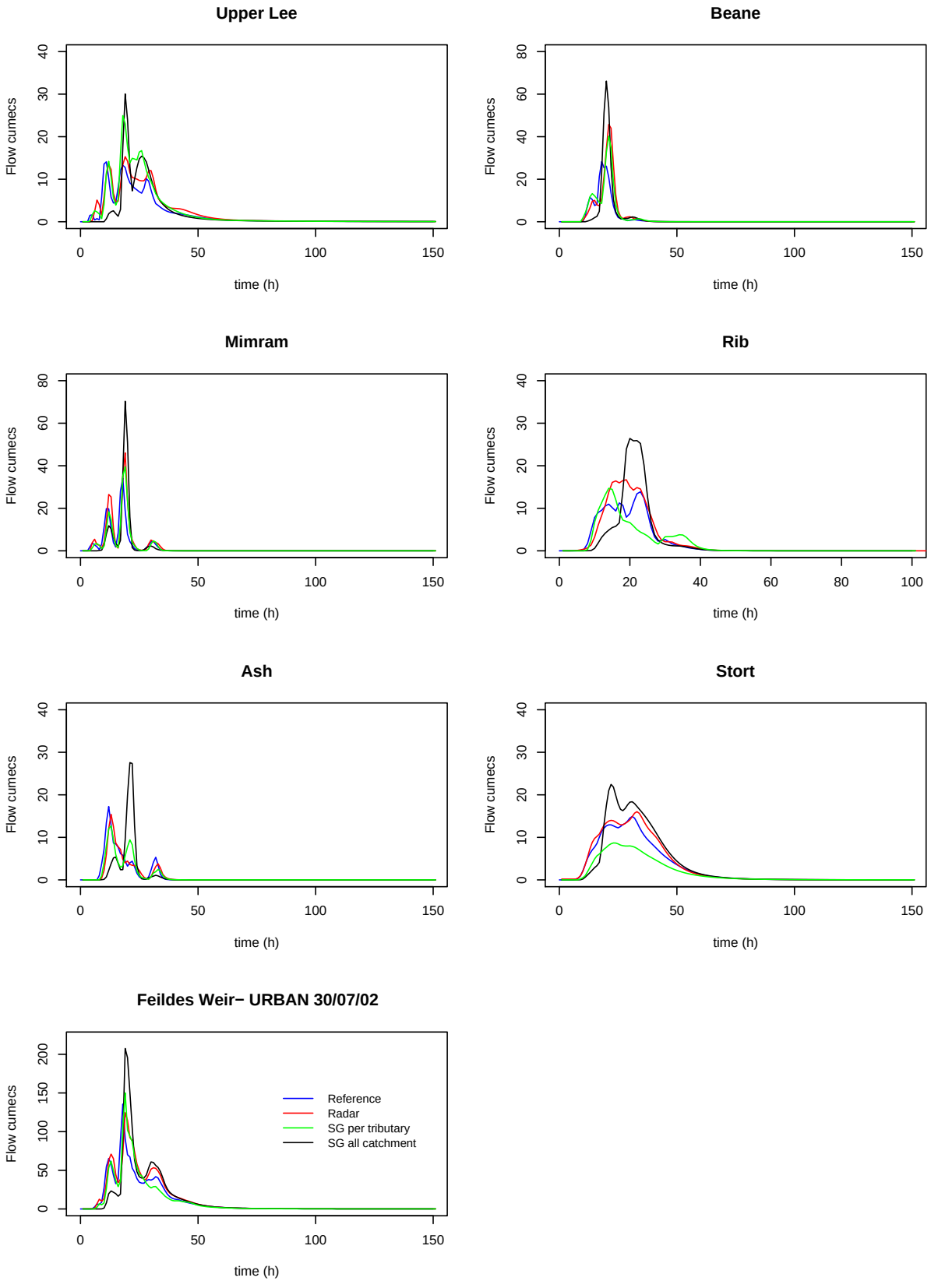


Figure 12.29 Reference and simulated flow at urban catchment outlets on 30/07/02

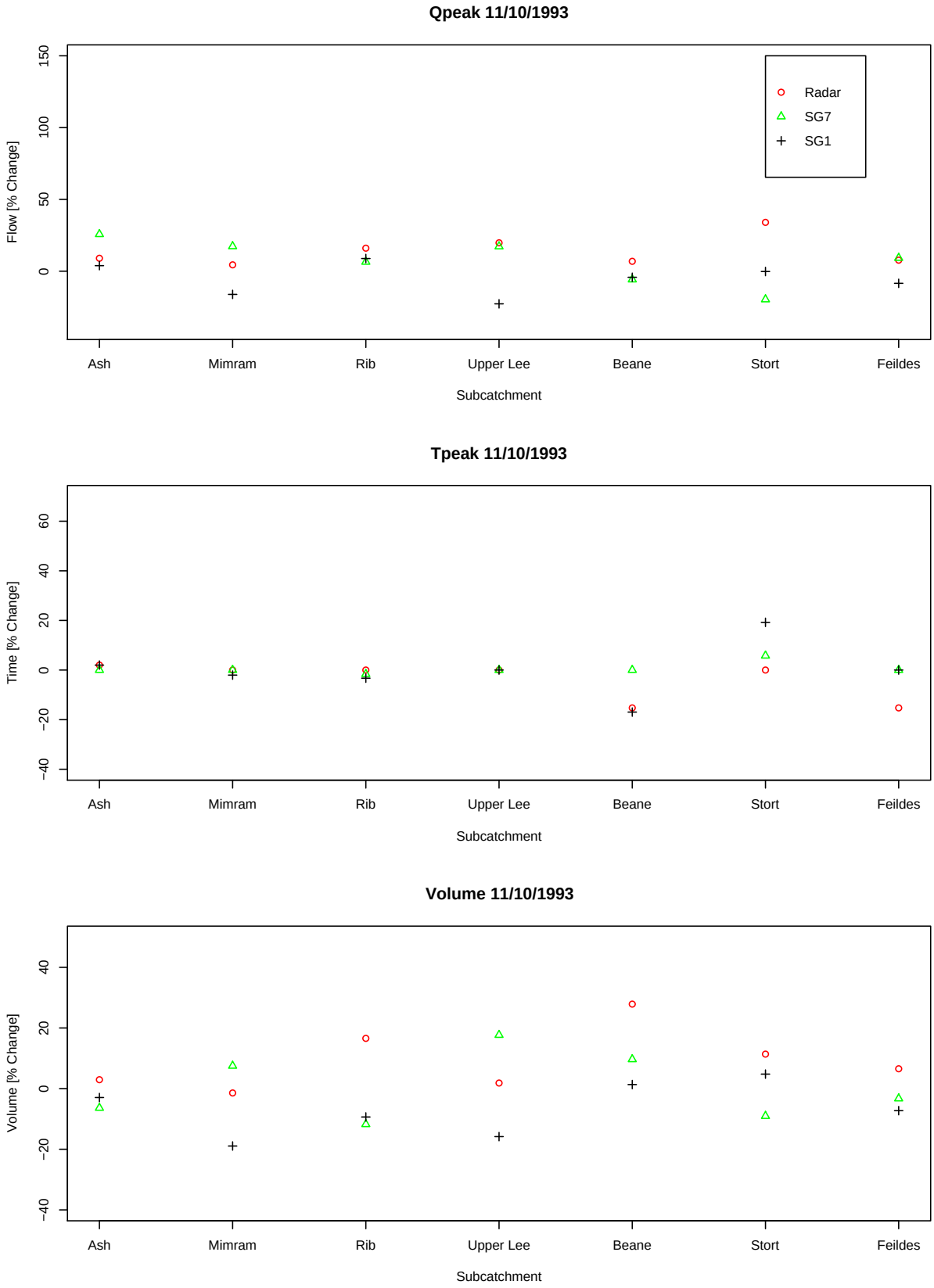


Figure 12.30 Relative change (%) in Peak Flow, Time to Peak and Volume in comparison to the reference flow for the event of 11/10/93

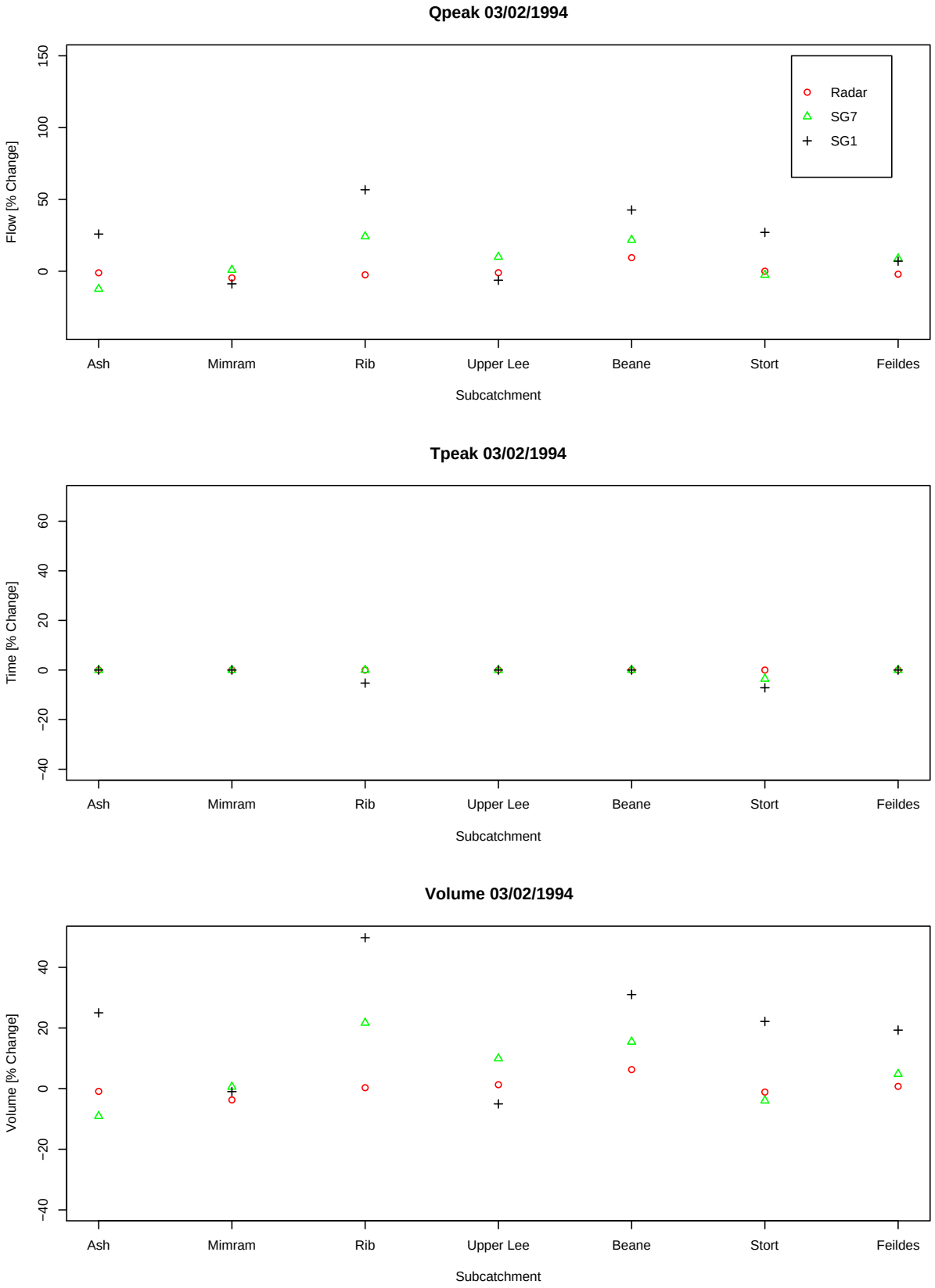


Figure 12.31 Relative change (%) in Peak Flow, Time to Peak and Volume in comparison to the reference flow for the event of 03/02/94

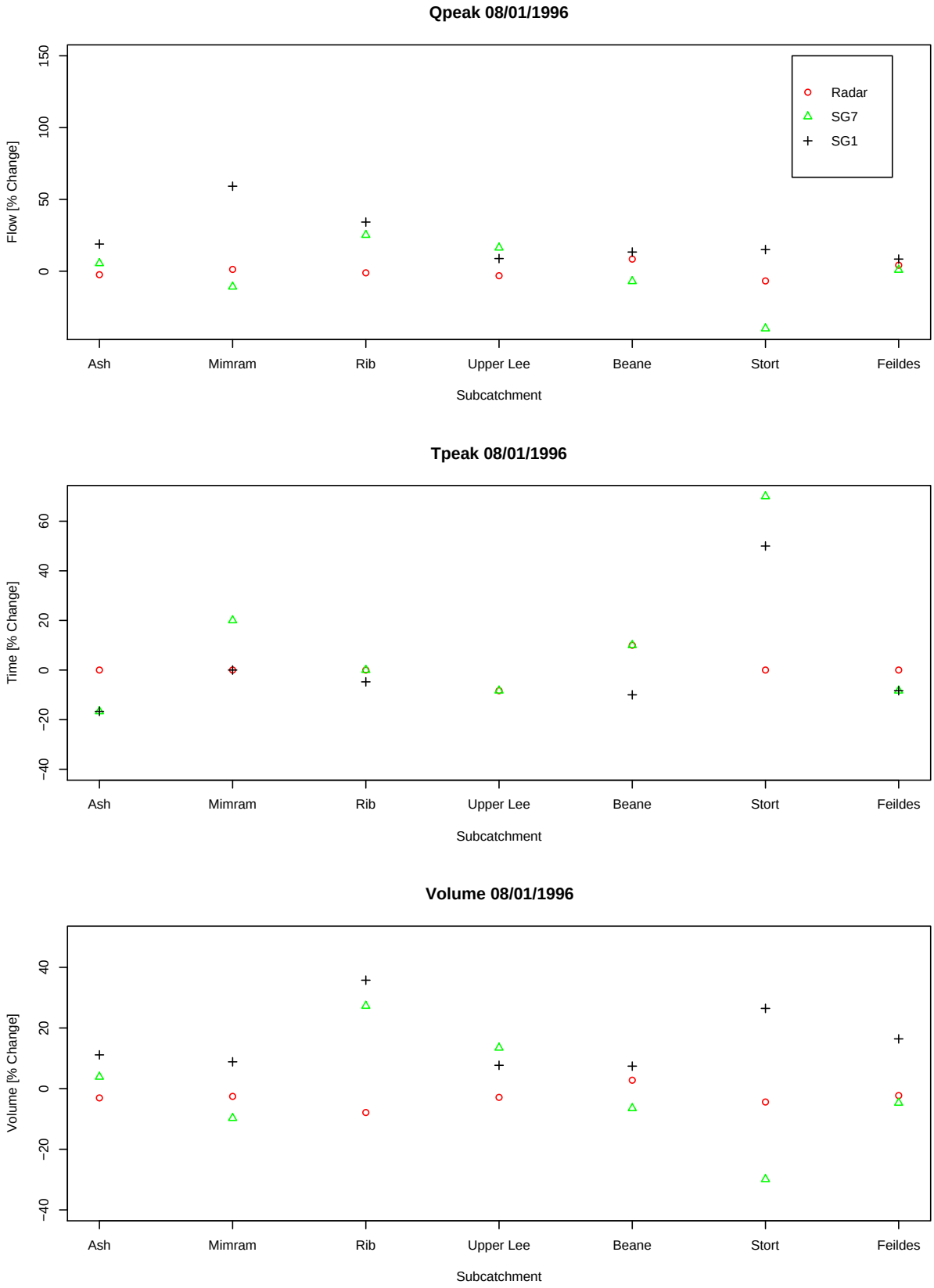


Figure 12.32 Relative change (%) in Peak Flow, Time to Peak and Volume in comparison to the reference flow for the event of 08/01/96

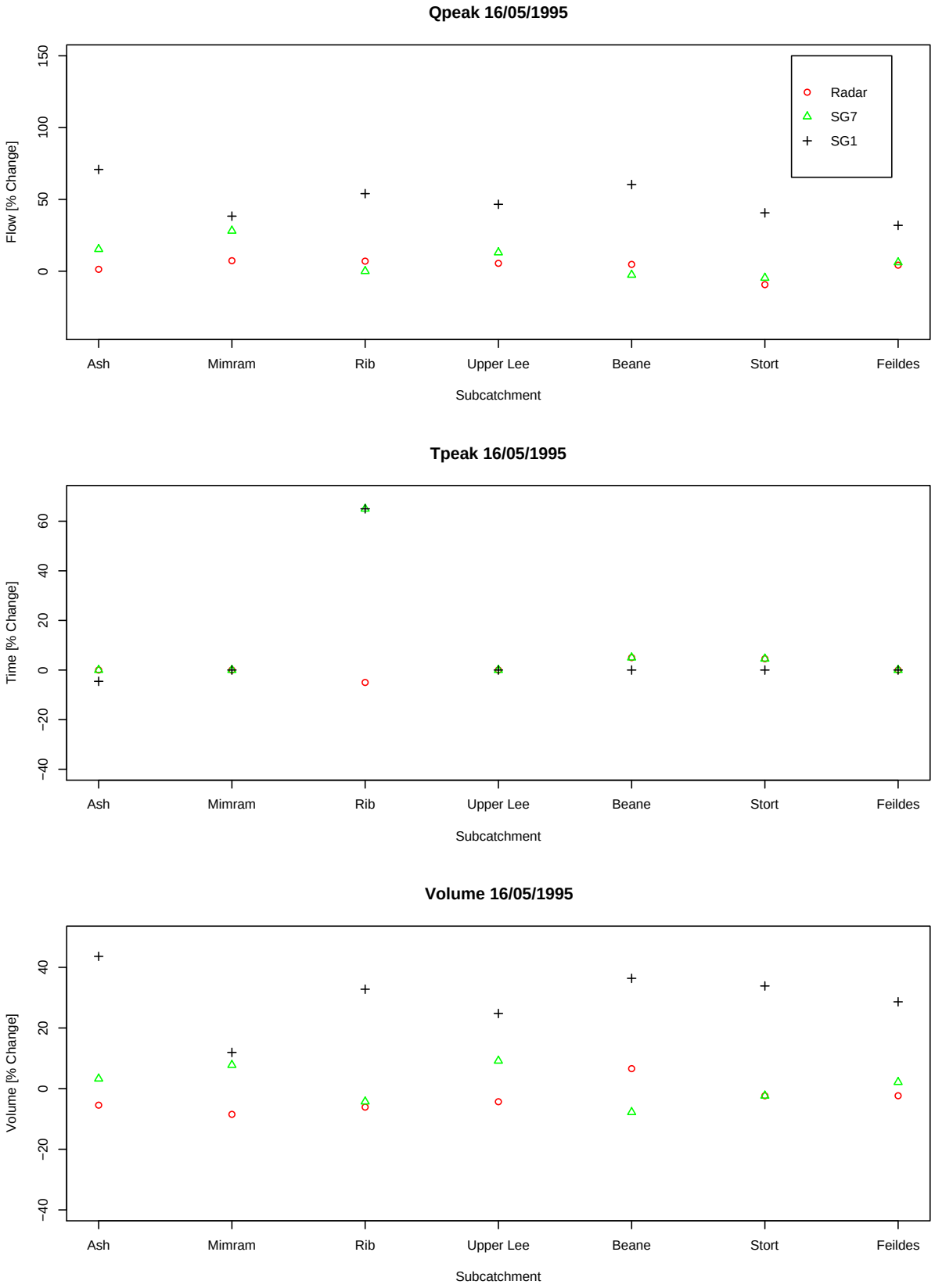


Figure 12.33 Relative change (%) in Peak Flow, Time to Peak and Volume in comparison to the reference flow for the event of 16/05/95

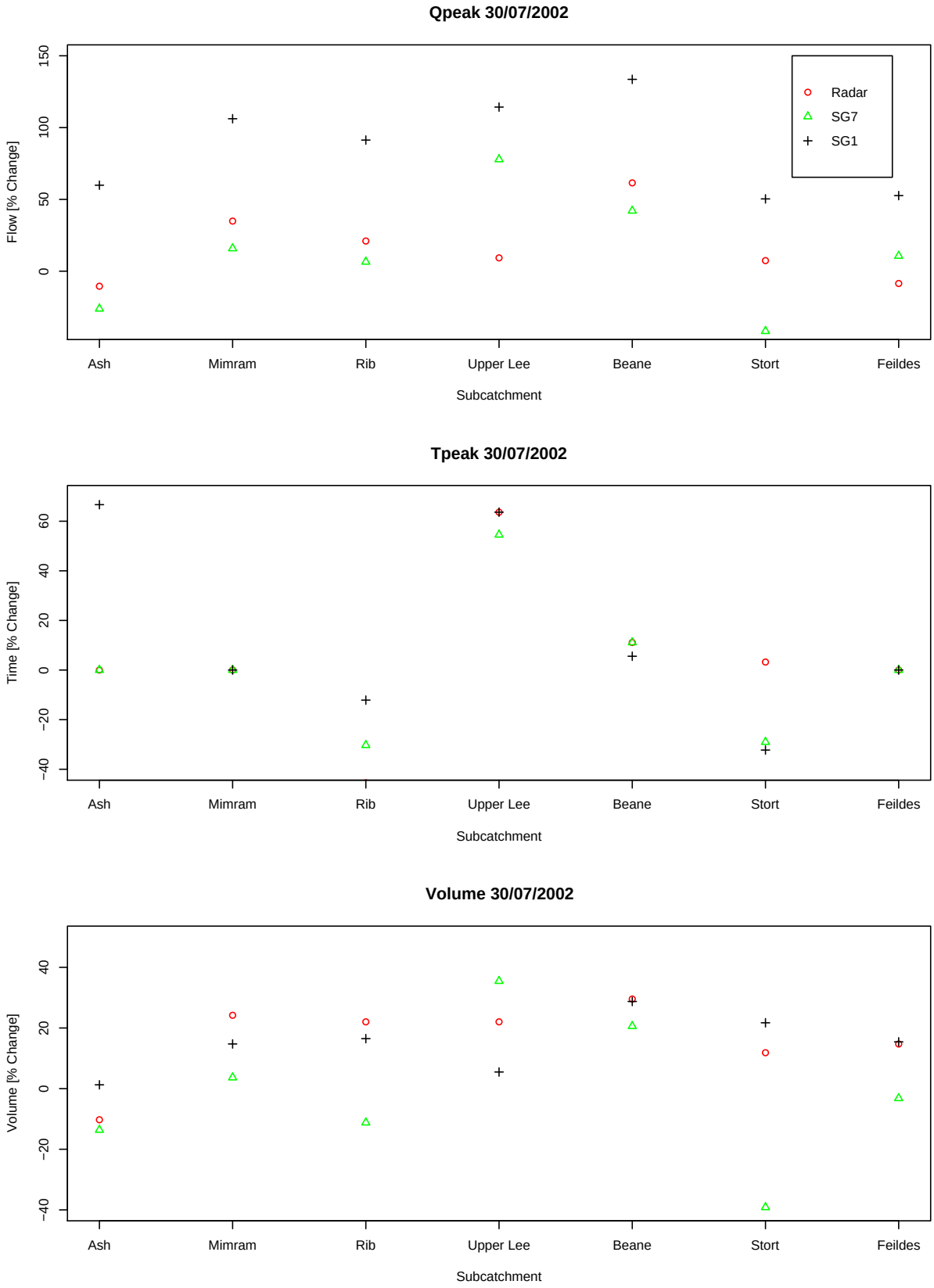


Figure 12.34 Relative change (%) in Peak Flow, Time to Peak and Volume in comparison to the reference flow for the event of 30/07/02

12.8.1 Impact of spatial variability of rainfall

The graph in Figure 12.35 presents a similar trend as in the case of existing predominantly rural catchments with SG1 leading to lower performance compared to the other rainfall representations. Summer events are characterised by lower NSE_{ref} values, expecially for SG7 and SG1.

12.8.2 Influence of catchment response time and catchment scale

Table 12.37 presents the response time of the urbanised catchments. The response time of the existing and urbanised basins are plotted against catchment area in Figure 12.36. It can be seen that the impact of urbanisation decreases the catchment response time by about 50% on average across all basins.

Table 12.21 Mean Lag Time (t_c) for urban catchments

	Mimram	Upper Lee	Beane	Ash	Feildes	Stort	Rib
t_c (h)	1.7	3.4	3.5	5	6.8	14.7	10.3

The graph in Figure 12.37 compares the model performance in terms of NSE_{ref} for the rural and urban subcatchments. The basins are ranked according to the rural catchment response time. The impact of urbanisation has increased the interquartile range in NSE_{ref} for the Mimram, the Ash and the Rib (i.e. the Chalk and the two entirely rural basins) whereas the range has decreased for the upper Lee, the Beane, the Stort and the upper catchment to Feildes Weir (i.e. the more heterogeneous basins including a mixture of Clay and Chalk and zones of urban developments). Similarly, Figure 12.38 compares the performance of rural and urban catchments but ranked by size. The results confirm that urbanisation has an impact on the model performance at sub-basin scale. At catchment scale, the NSE_{ref} remains high using SG7 and radar data. Lower performance is observed using SG1 during summer events.

12.8.3 Conclusion

In conclusion, results show that radar data lead to similar performance on rural and urban catchments. SG7 applied on urbanised basins leads to similar NSE_{ref} values as in the case of rural basins but to larger errors in Q_p and T_p. The discrepancies are enhanced for SG1, especially in the case of summer events. The effect of urbanisation decreases the lag time by 50% on average for all subcatchments. However the impacts of the lag time and the catchment scale on the runoff response were not clearly evidenced. Only at the catchment scale, less sensitivity to spatial rainfall (using radar data and SG7) is observed.

12.9 Disaggregation Scheme Results

In this section, a spatially-uniform temporal disaggregation of observed daily rainfall to hourly data is tested. The sub-daily temporal profile of the raingauge closest to the

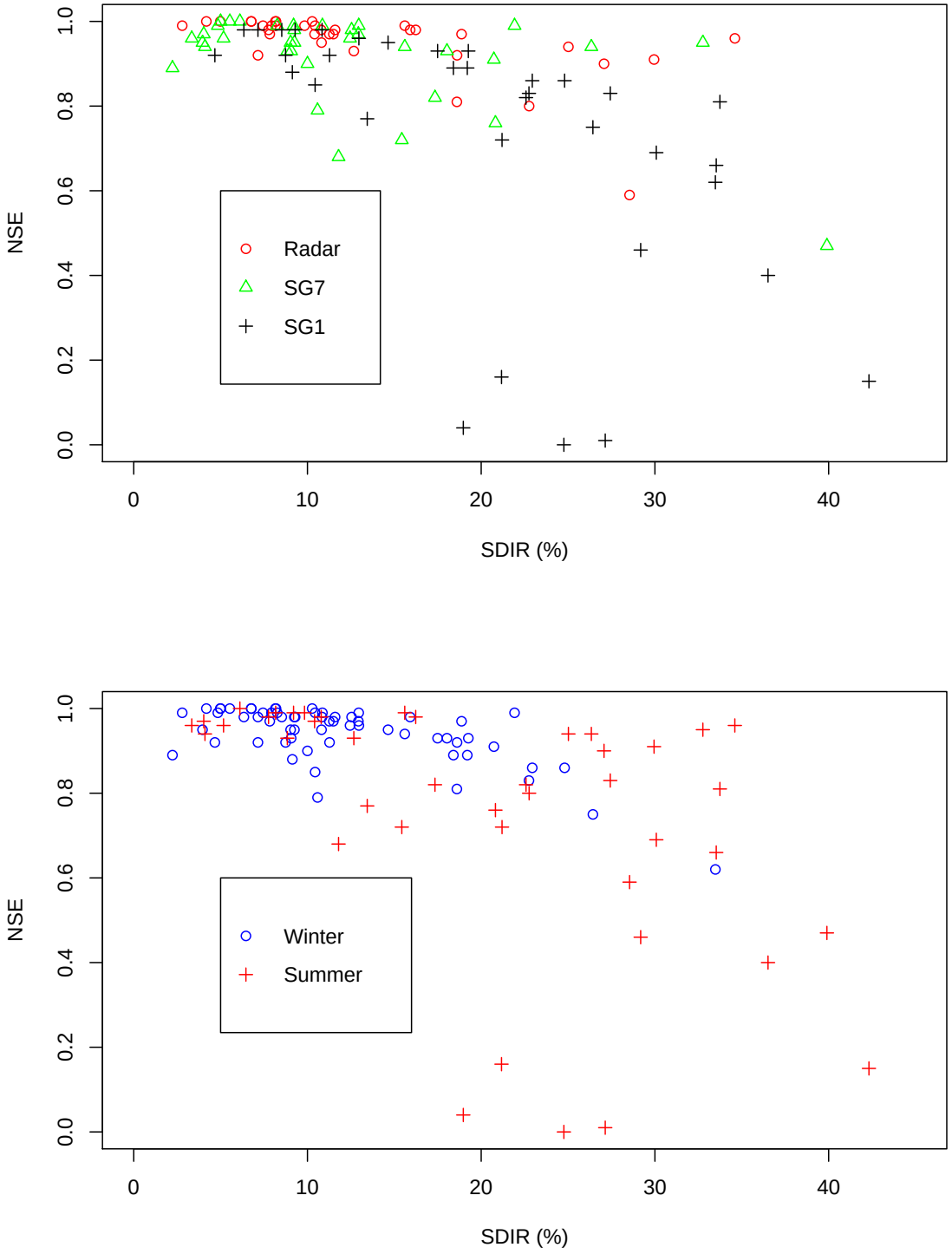


Figure 12.35 Impact of rainfall spatial variability on NSE. Influence of rainfall representation (Top) and type of rainfall events (Bottom).

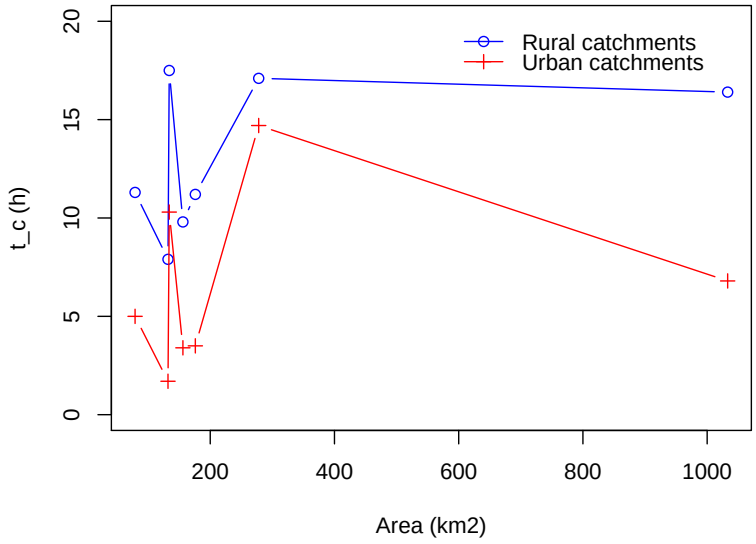


Figure 12.36 Response time of existing and urbanised basins against catchment area

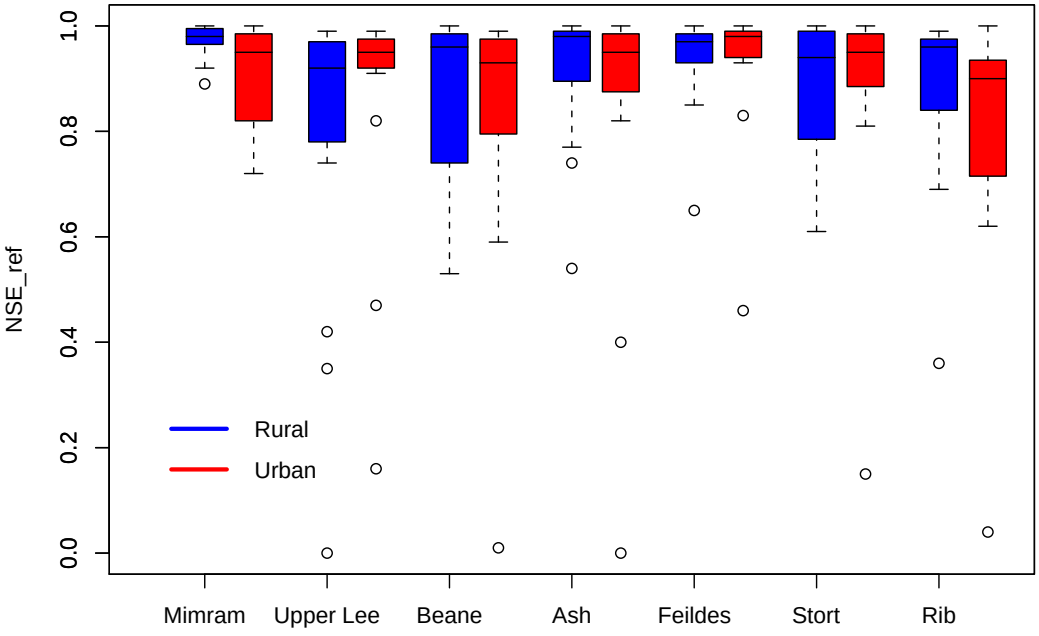


Figure 12.37 Model performance for rural and urban catchments

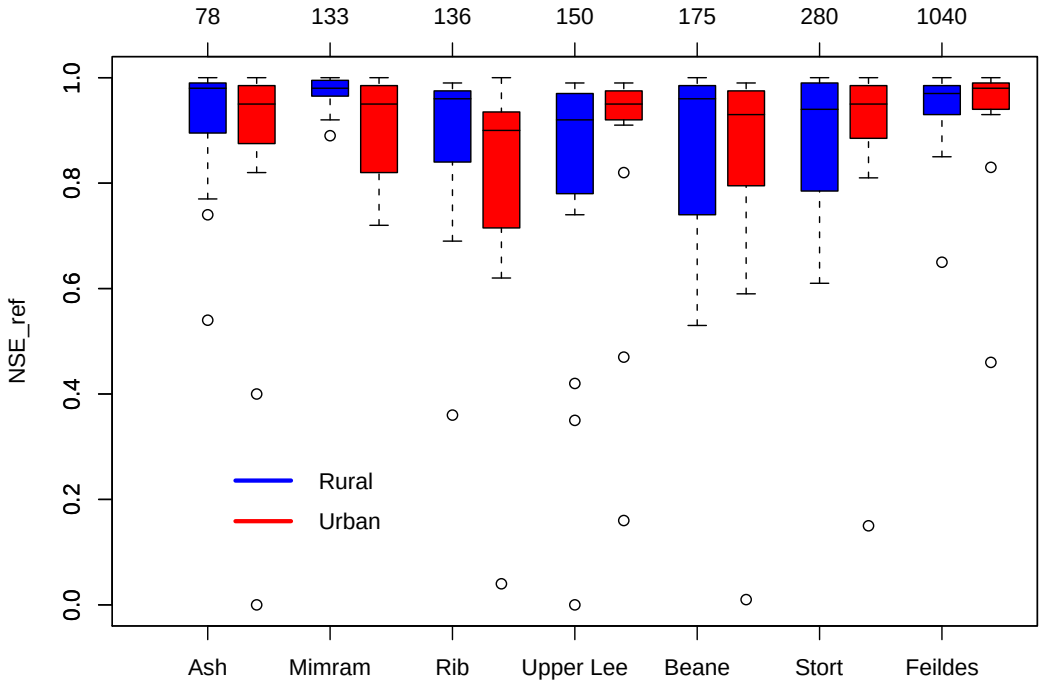


Figure 12.38 Boxplot of NSE criteria for each urban subcatchment ranked according to catchment scale

centroid of the catchment (raingauge 163) is applied to disaggregate the daily values at the other sites. The comparison between the simulated and reference flows are presented in Figures 12.39 to 12.43 for rural catchments and in Figures 12.44 to 12.48 for urban catchments.

12.9.1 Rural Catchments

Table 12.22 NSE_{ref} for the disaggregation scheme

	Upper Lee	Mimram	Beane	Rib	Ash	Stort	Feildes
11-Oct-93	0.99	0.98	0.98	0.99	0.98	0.98	0.97
03-Feb-94	0.97	0.99	1	1	0.98	1	1
08-Jan-96	0.98	0.99	1	1	1	0.96	0.99
16-May-95	0.97	1	1	1	1	1	0.99
30-Jul-02	0.92	0.96	0.95	0.99	0.94	0.98	0.96

Using this disaggregated scheme, the areal rainfall variations are preserved, but a unique temporal profile of rainfall is applied to all subareas. Table 12.22 displays the NSE_{ref} results between the reference and simulated flow. Overall a close fit to the reference streamflow is obtained (NSE_{ref} about 1), the performance is slightly lower for the convective summer event of 30th July 2002 but still remains high ($NSE_{ref} > 0.92$). A shift in the timing of the hydrograph is observed for the extreme event of 11th October 1993 and the convective event of 30th July 2002 as illustrated in Figure 12.39

Table 12.23 Summary performance on rural catchments

	Radar	SG7	SG1	Dis
NSE_{ref}	0.95	0.91	0.80	0.98
Q_p (%)	12.6	15.3	28.1	6.6
T_p (%)	5.6	5.8	8.0	6.5
V (%)	7.4	12.4	21.6	4

and Figure 12.43. Table 12.23 summarises the model performance obtained for each rainfall scenarios. Overall the model performance in terms of NSE_{ref} is higher using the disaggregated scheme (0.98), and decreases with radar data (0.95), SG7 (0.91) and SG1 (0.80). Since the areal variation are preserved using the disaggregation scheme (Dis), this rainfall scenario leads to smallest variations in Q_p and V but the percentage change in T_p is greater than when using radar data or SG7 (6.5% compared to 5.6 and 5.8%).

Since there is a trade-off between the rainfall variability and the integrating effect of the catchment, a simple disaggregation procedure can be applied with good modelling performance. This test also reveals that the impact of spatial variability of rainfall dominates the impact of temporal variability when modelling runoff on these catchments.

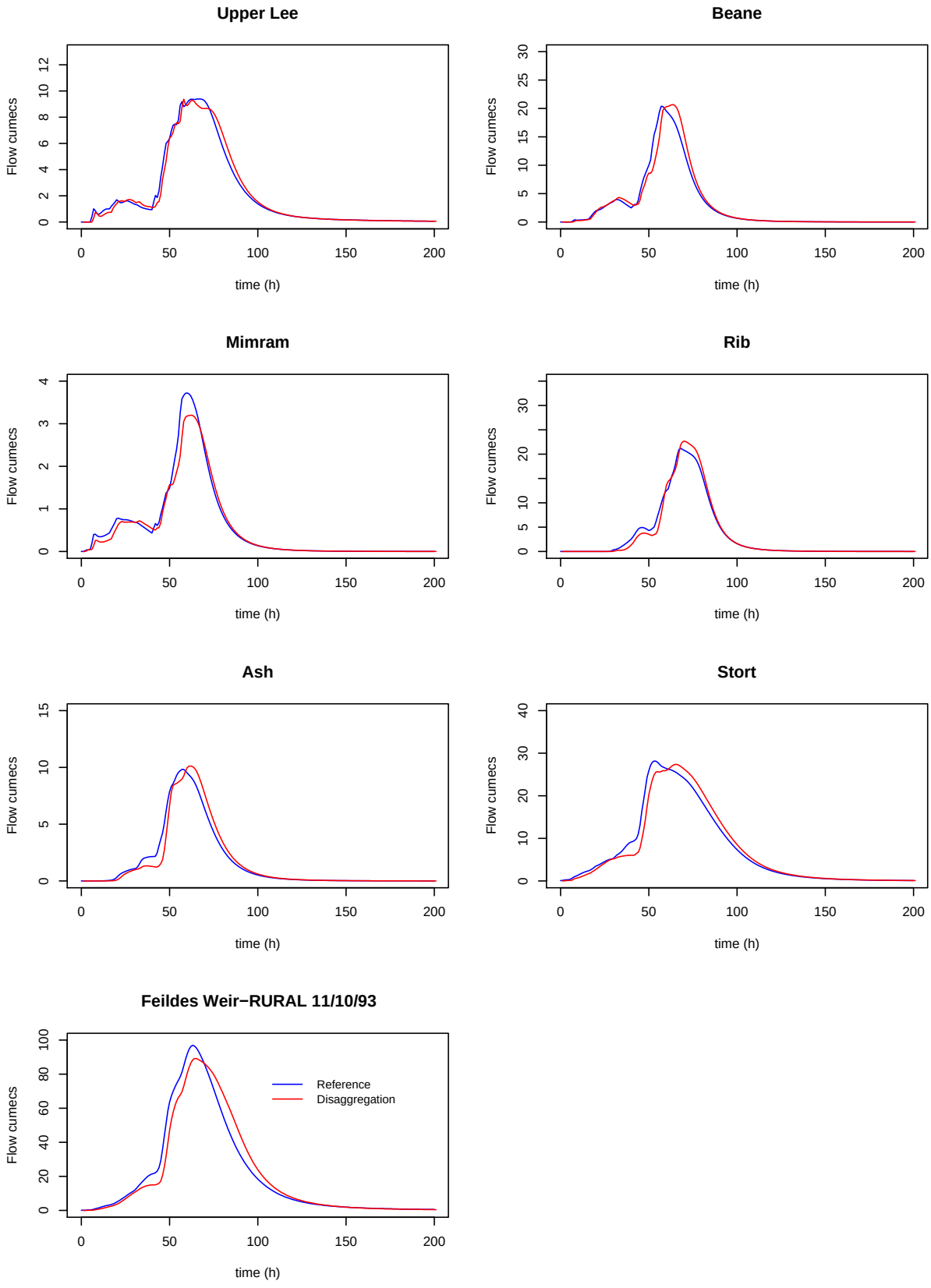


Figure 12.39 Simulated flow using disaggregated rainfall for the event of 11/10/93

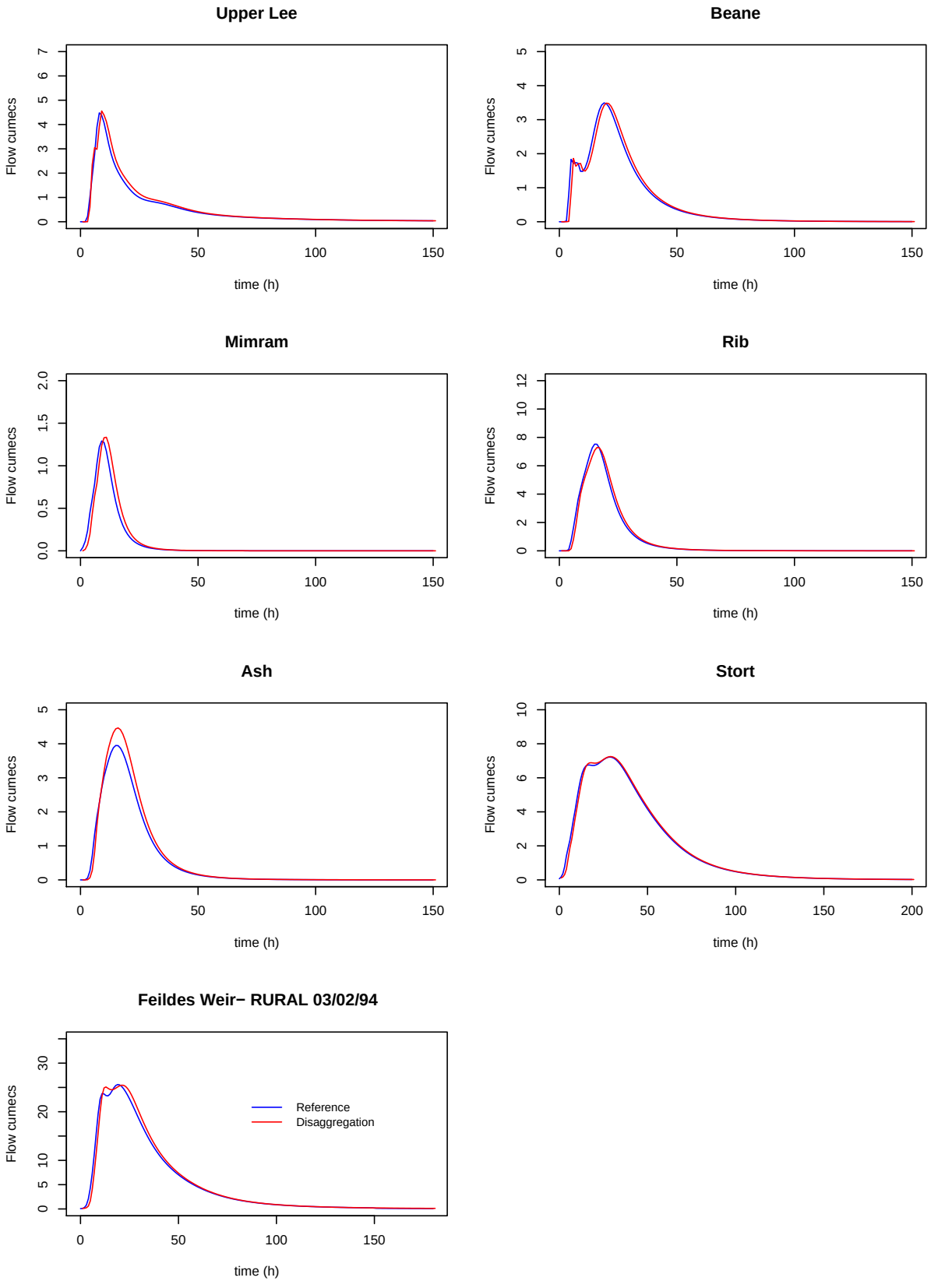


Figure 12.40 Simulated flow using disaggregated rainfall for the event of 03/02/94

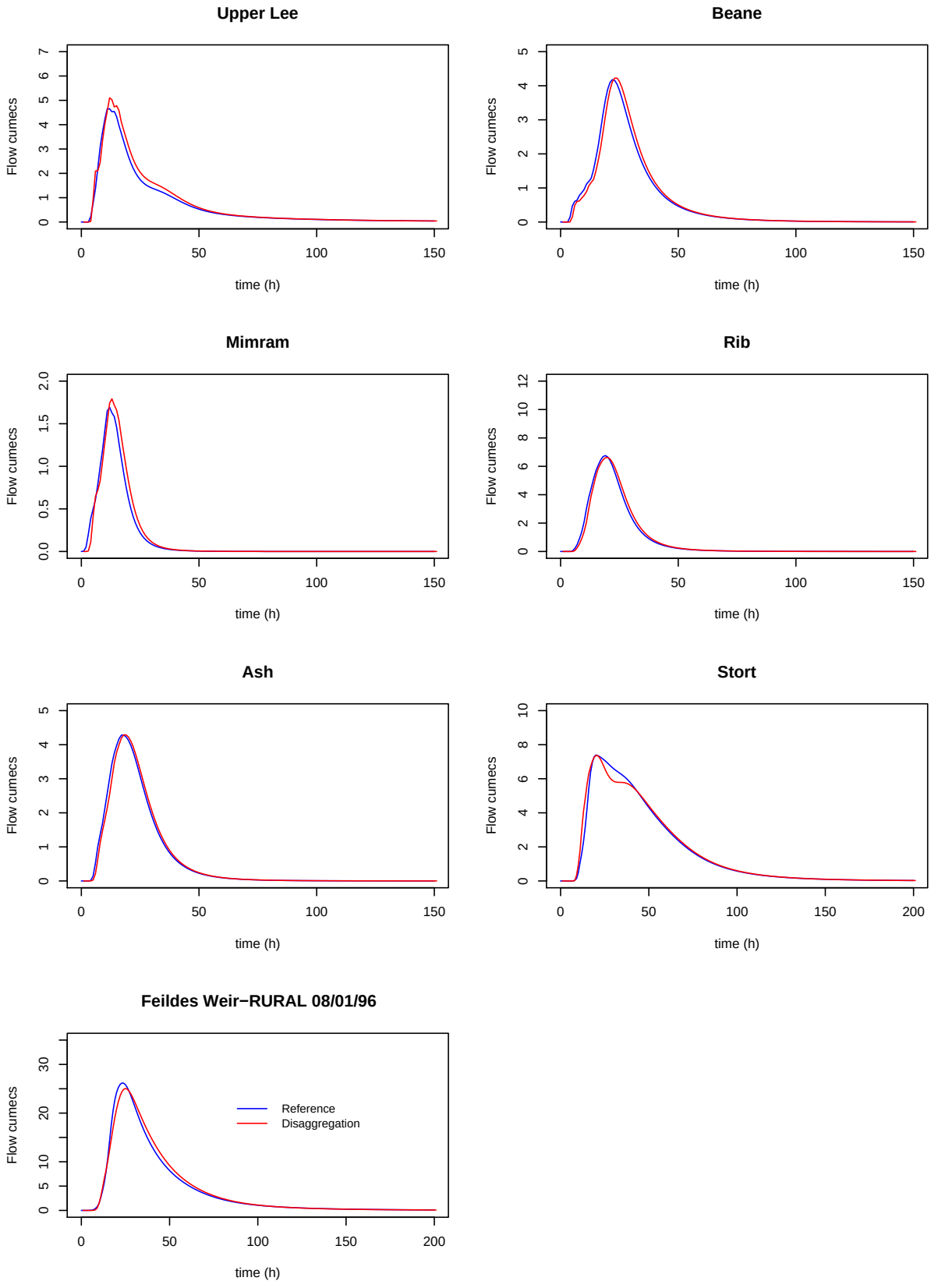


Figure 12.41 Simulated flow using disaggregated rainfall for the event of 08/01/96

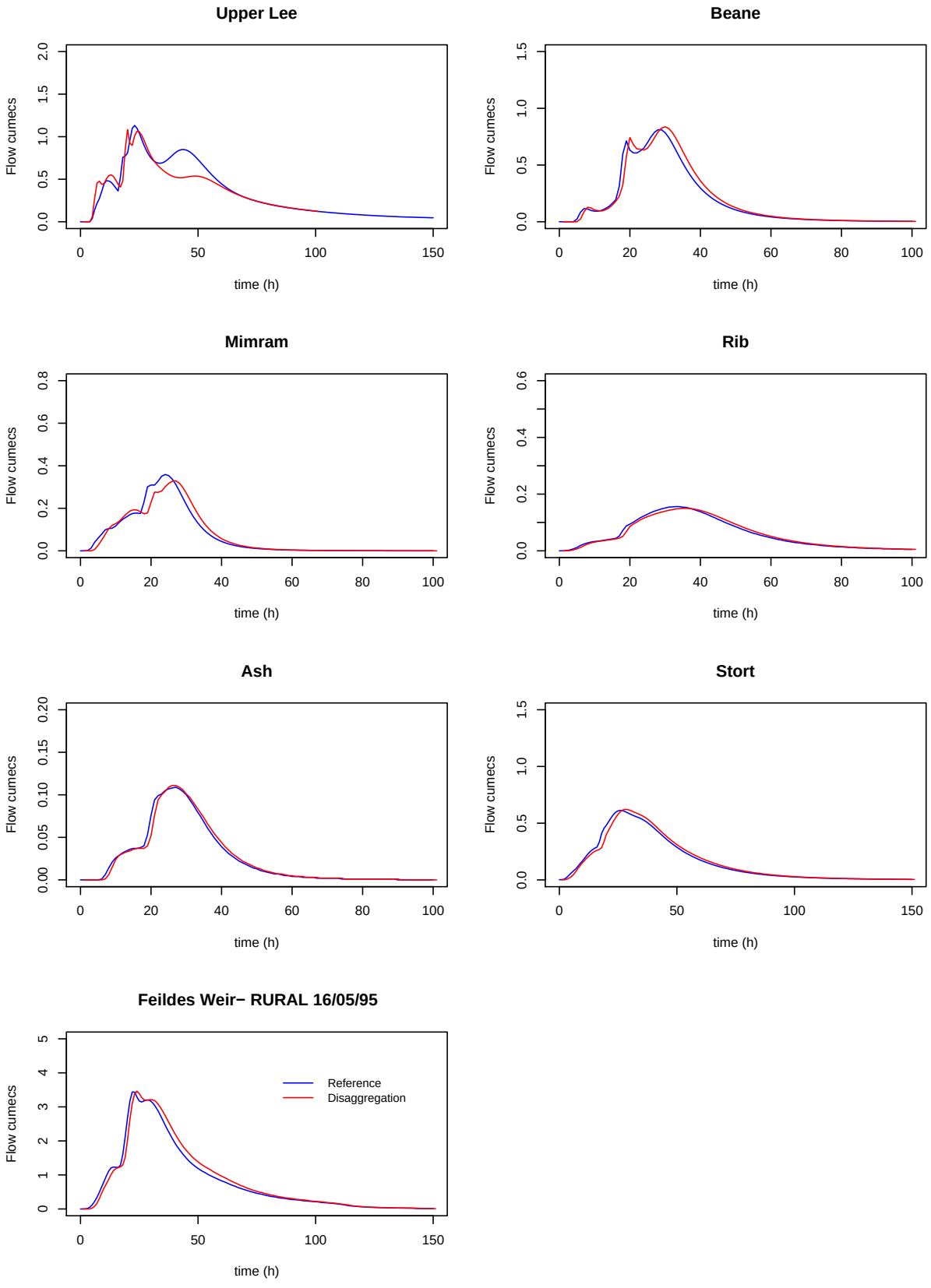


Figure 12.42 Simulated flow using disaggregated rainfall for the event of 16/05/95

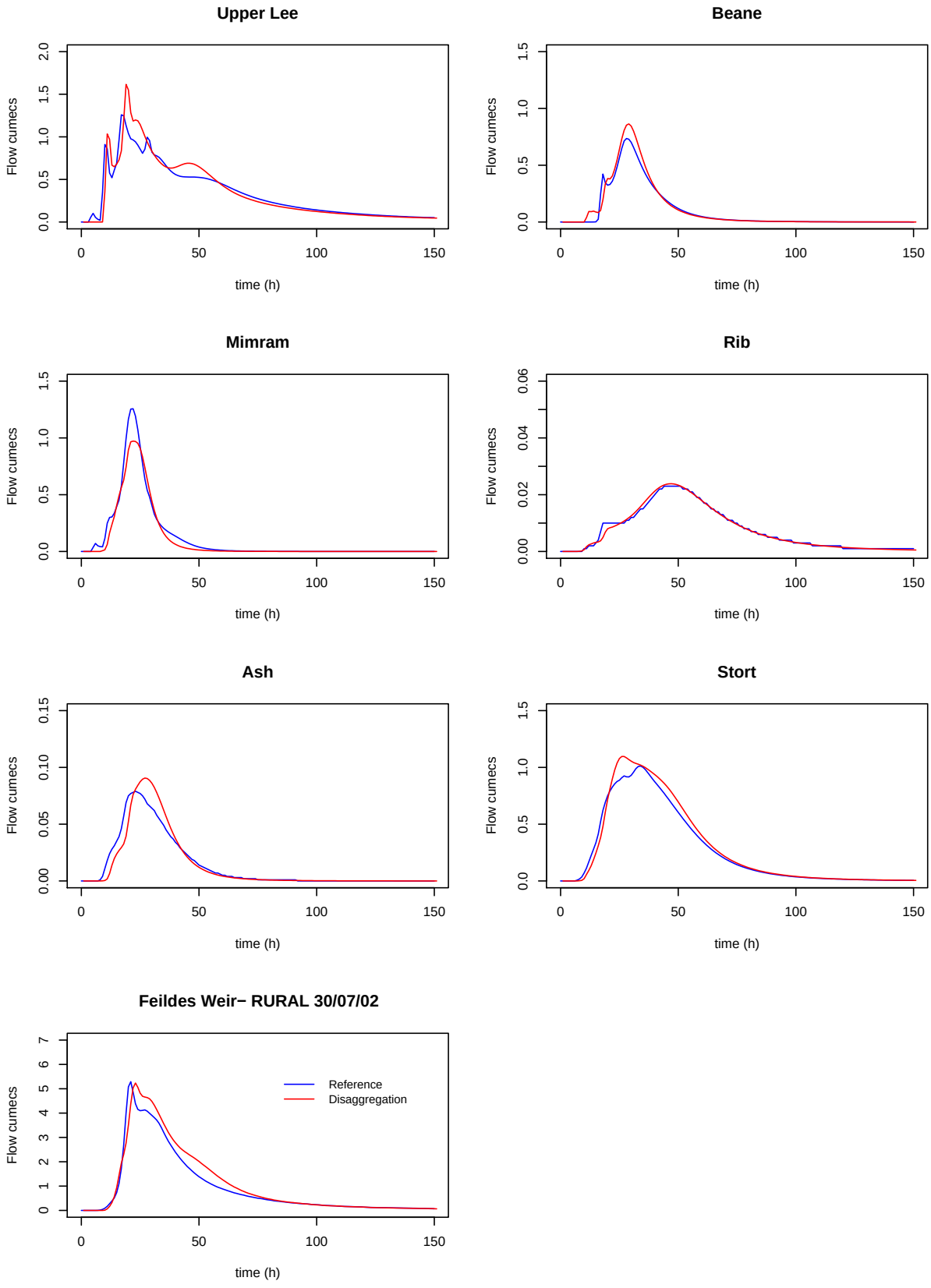


Figure 12.43 Simulated flow using disaggregated rainfall for the event of 30/07/02

12.9.2 Urban Catchments

As seen in Table 12.24, the model performance remains high for the artificially urbanised basins with NSE_{ref} of 0.94 on average for all events and (sub)catchments. Examining the results for the individual subcatchments, results show that the model performance is lower for the Mimram and the Ash on 30th July 2002 (NSE_{ref} of 0.71 and 0.33, respectively). Lower performance at the catchment scale is also observed (NSE_{ref} of 0.88). This indicates the limit of the disaggregation scheme which is not appropriate for temporally variable rainfall as in the case of localised convective storms.

Table 12.24 NSE_{ref} for the disaggregation scheme

	Upper Lee	Mimram	Beane	Rib	Ash	Stort	Feildes
11-Oct-93	0.98	0.94	0.99	0.93	0.90	0.97	0.97
03-Feb-94	0.98	0.98	1	0.99	0.99	1	1
08-Jan-96	0.98	0.97	0.93	1	1	0.97	0.99
16-May-95	0.93	0.83	0.96	0.96	0.97	0.99	0.96
30-Jul-02	0.92	0.71	0.93	0.88	0.33	0.98	0.88

A summary of the model performance characteristics obtained on average for all catchments and events and for all rainfall representations on urbanised catchments is presented in Table 12.25. Higher NSE_{ref} performance (0.95) and lower errors in Q_p (9.7%) and T_p (5.4%) are obtained with radar data. SG7 leads to high NSE_{ref} (0.92) but higher errors in Q_p (16.5%) and T_p (9.8%). The disaggregation scheme achieves the second best performance in terms of NSE_{ref} (0.94). It leads to similar change in Q_p (9%) as in the case of radar data but to larger errors in T_p (15.5%). Lowest model performance characteristics are obtained with SG1.

Table 12.25 Summary performance on urban catchments

	Radar	SG7	SG1	Dis
NSE_{ref}	0.95	0.92	0.73	0.94
Q_p (%)	9.7	16.5	38.5	9
T_p (%)	5.4	9.8	12.7	15.5
V (%)	8	11.1	18.4	4.1

12.9.3 Conclusion

Overall excellent performance of the disaggregation scheme is observed for all types of events and (sub)catchments when the basins are in natural condition (NSE_{ref} of 0.98 on average). Some degradation in performance occurs although the high performance is maintained (NSE_{ref} of 0.94 on average) for the artificially urbanised basins. When there is a lack of available rainfall data, this simple disaggregation scheme conditioned on real or simulated daily rainfall can provide a useful alternative.

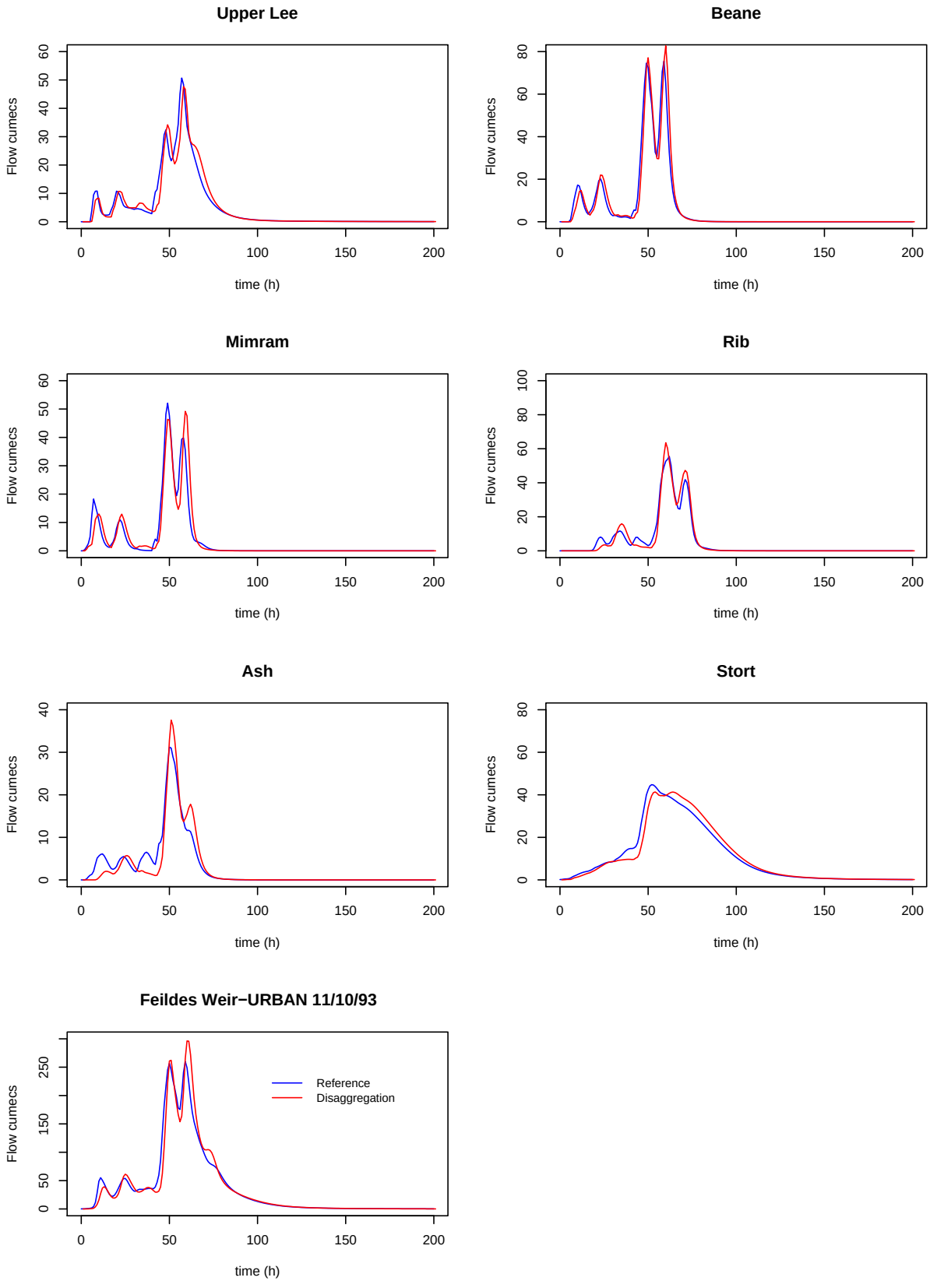


Figure 12.44 Simulated flow for the event of 11/10/93

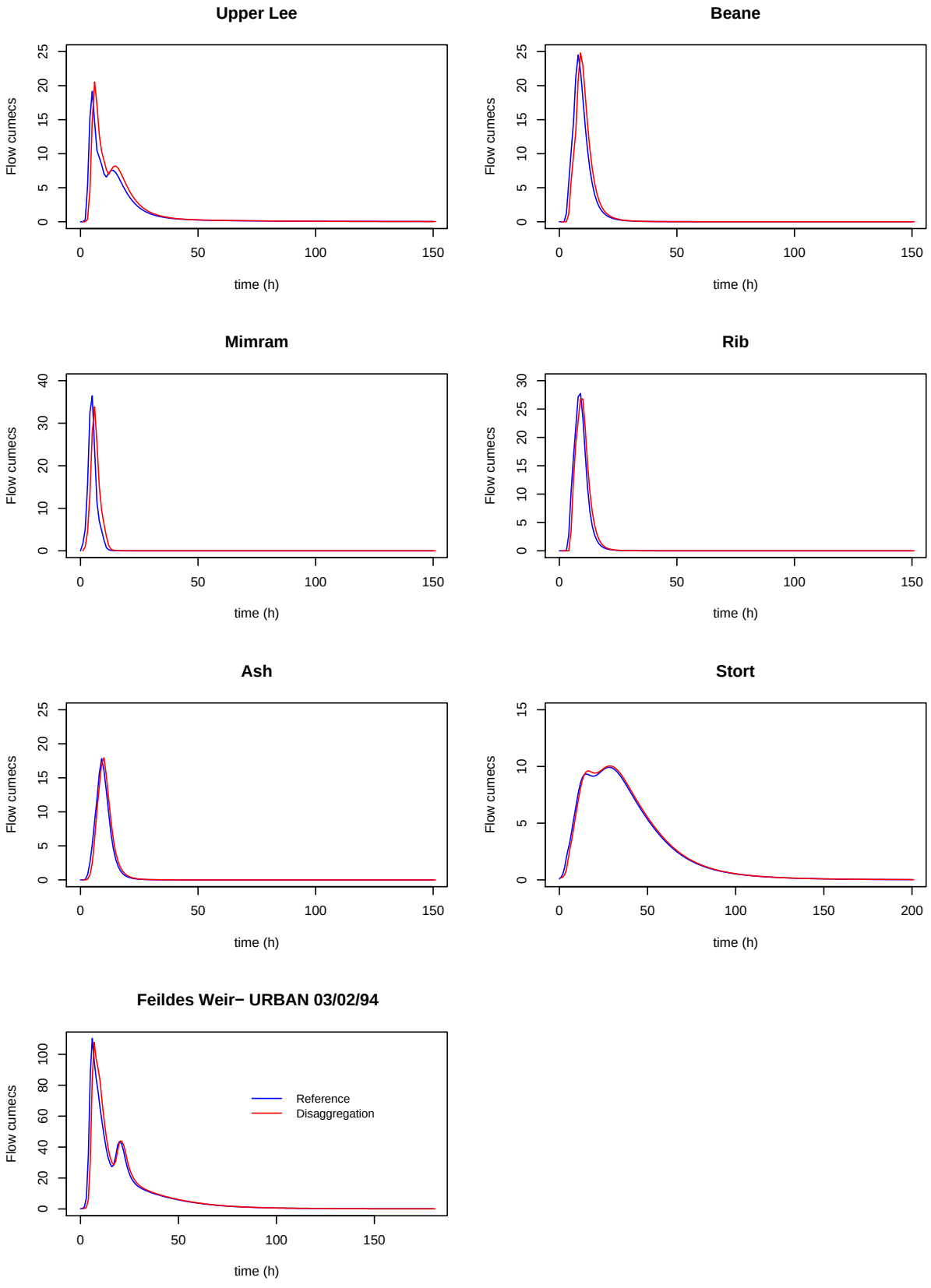


Figure 12.45 Simulated flow for the event of 03/02/94

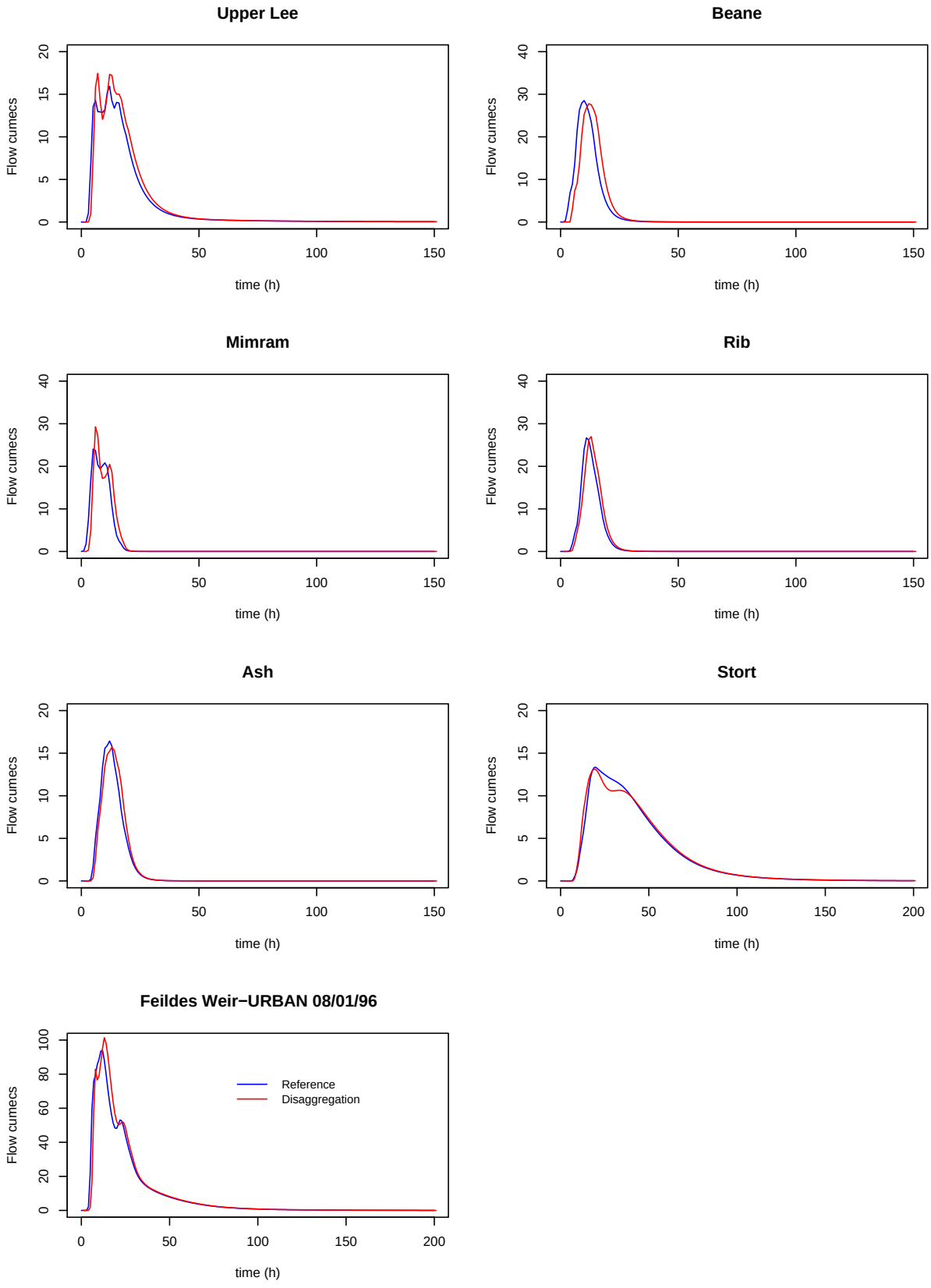


Figure 12.46 Simulated flow for the event of 08/01/96

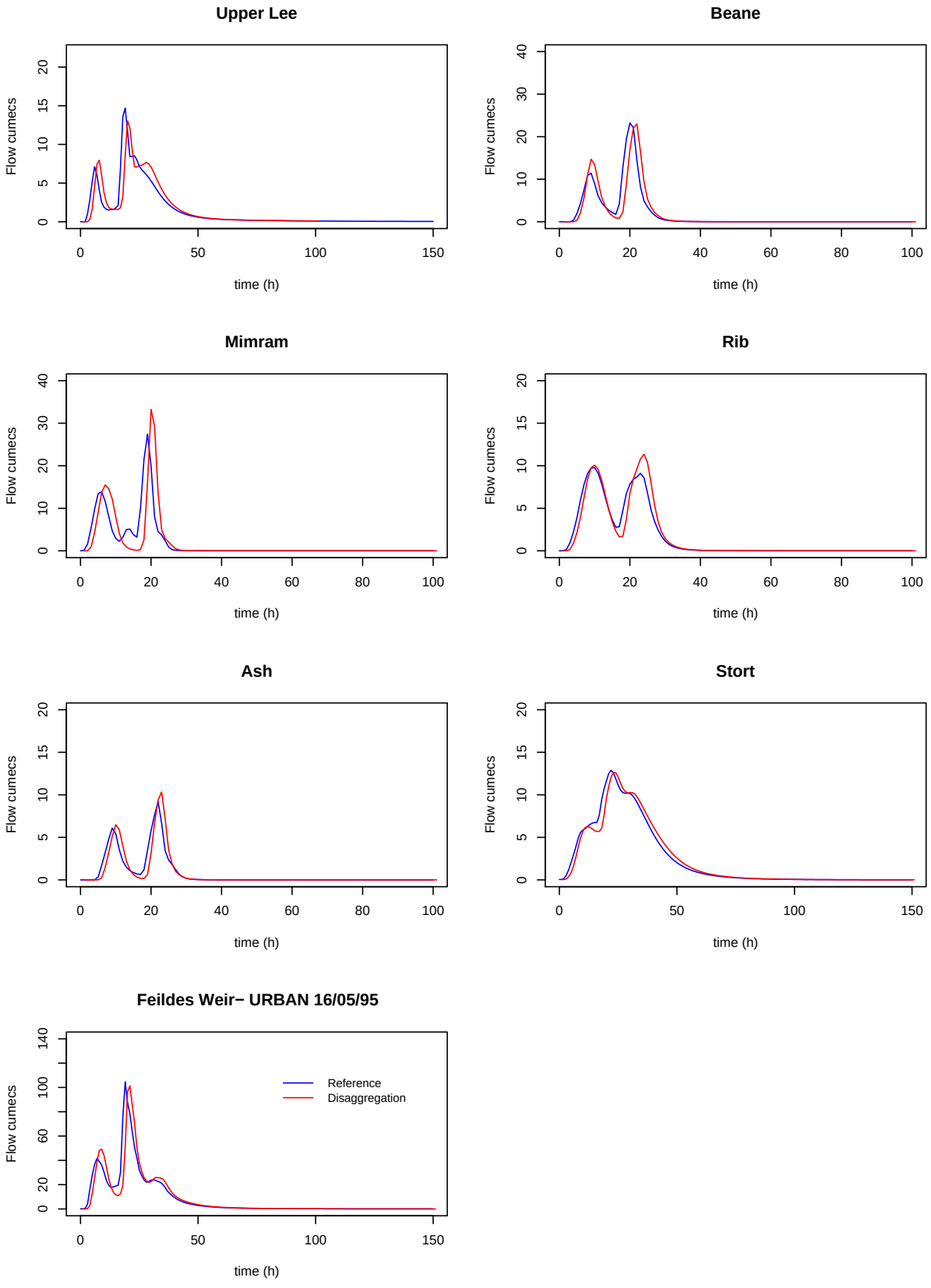


Figure 12.47 Simulated flow for the event of 16/05/95

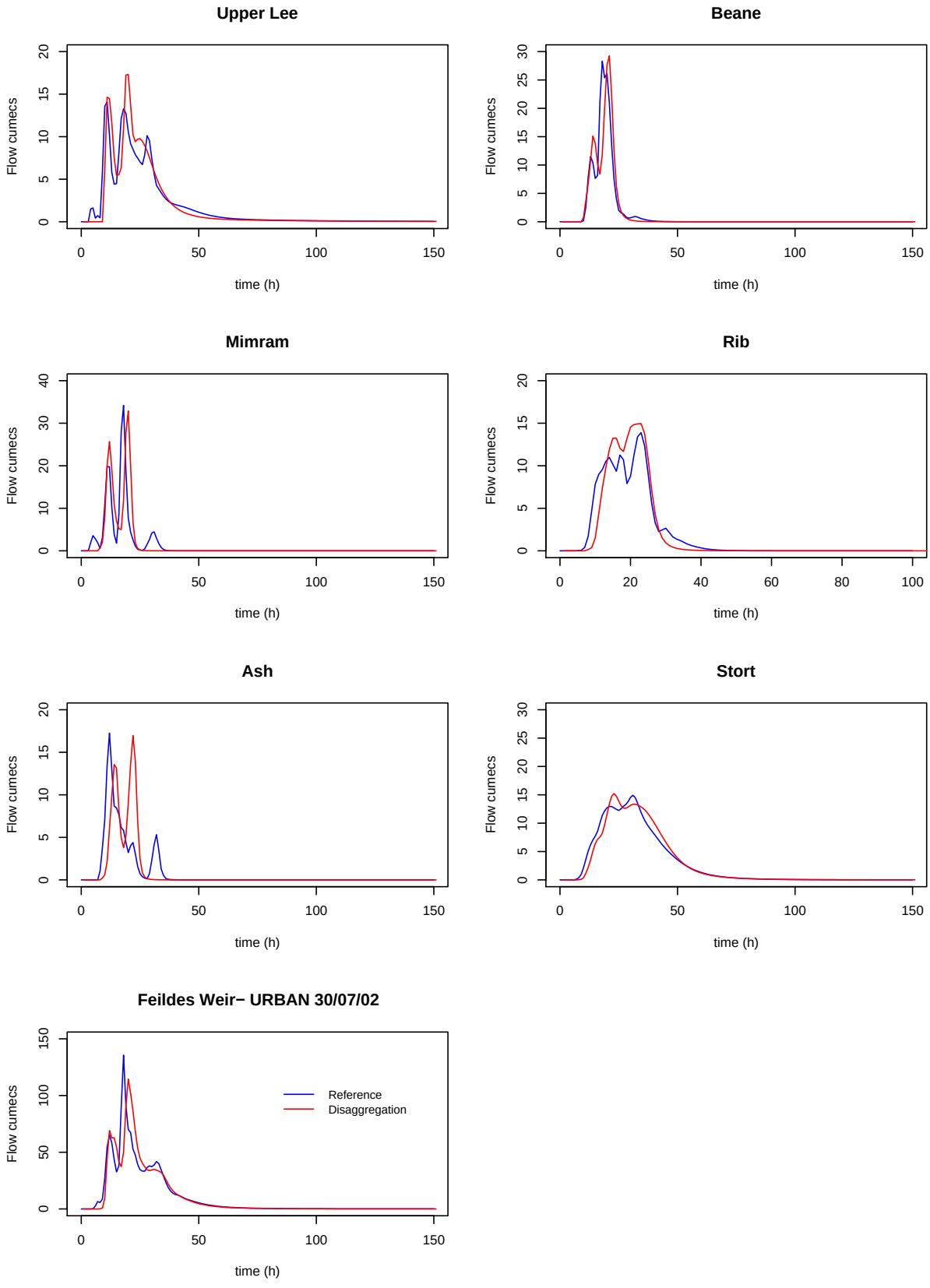


Figure 12.48 Simulated flow for the event of 30/07/02

12.10 Conclusion and Discussion

Event-based, semi-distributed hydrological rainfall-runoff modelling of the Upper Lee catchment has been undertaken. The Lee is divided into six subcatchments, representing the main tributaries. Each subcatchment is first modelled independently and the simulated runoff is then routed into the total runoff generated at the watershed outlet. This strategy allows for spatial variation of both rainfall and model parameters.

Based on a library of storms representative of extreme and more common frontal and convective rainfall, the simulated runoff, driven by radar data, one gauge per tributary and a single gauge representing the whole catchment, is compared to the reference streamflow generated from the spatially variable rainfall as defined by the full raingauge network. The results of this study show that the spatial variation of rainfall is important for runoff production at catchment scale, for a range of scales and catchment types.

The catchment response is a complex relationship between the spatially variable input and the watershed heterogeneity. An index (SDIR) has been defined to quantify the spatial areal deviation in precipitation to the reference rainfall and a clear relationship between spatial rainfall and model performance has been identified. For a range of representative events, spatial deviations in rainfall of up to 40% are observed depending on the type of data available and this is reflected in the simulated hydrograph. At the subcatchment scale, spatial deviations in rainfall above 10% lead to variations greater than 20% in the peak flow and 10% in the volume of the hydrograph. Spatial deviation in rainfall above 20% lead to NSE performance below 0.8. Results show that more variability in the rainfall distribution is introduced when extreme or summer events occur. The findings of this study stress the need for spatial data. Spatially variable rainfall and radar data lead to similar flow responses whereas larger discrepancies are introduced when using input data from a subset of the raingauges. Using radar data, results in terms of NSE are robust and remain high (usually above 0.9). Using one gauge per tributary achieves high NSE value at the catchment scale but can lead to low NSE values at the subcatchment scale. When a unique rain gauge is used to model the catchment response, the results worsen at both subcatchment and catchment scales and also affect the reproduction of the peak hydrograph time. Since the model has been calibrated using spatially-variable rainfall defined by the full raingauge network, the performance of radar data over the spatially-variable rainfall could not be assessed. However results for the largest summer event suggest that large spatial deviation in rainfall (about 30%) can be observed between the radar data and the full raingauge network.

The damping behaviour of the basin has been characterised by the catchment response time. The extent of runoff contributing areas, their location with flow distance to the outlet and the size of the catchments have been identified as the dominant controls on runoff response and should be assessed prior modelling. Further work is required to identify their importance quantitatively. An explicit representation of these interactions is possible within an analytical framework as proposed by Woods and Sivapalan (1999).

It was found that clay and/or partly urban subcatchments are more sensitive to the spatial distribution of rainfall, whereas the chalk rural subcatchment, the Mimram,

exhibits a damping behaviour of the rainfall input. At a catchment scale of 1000 km^2, the impact of spatial variation of rainfall on runoff is less than at the subcatchment scale of 80 to 280 km^2. This underlines the integrating effect of the catchment as the size increases. However the catchment modelling still benefits from spatial information in the case of convective summer events. Entirely rural catchments (Ash and Rib) are most sensitive to the antecedent catchment conditions with insignificant flow (below 0.5 m^3s^{-1}) generated in summer.

When the existing subcatchments are artificially turned into fast-responding urban basins, a decrease in the catchment response time (about 40% on average for all subcatments) is observed and the discrepancies in terms of reproduction of peak flow and runoff volume are enhanced for the most variable events. A spatially uniform temporal disaggregation scheme conditioned on observed daily data is tested as input to the rainfall-runoff model. Using this disaggregated scheme, the areal rainfall variations are preserved, but a unique temporal profile of rainfall is applied to all subareas. A close fit (NSE close to 1) between the simulated flow and the reference flow is obtained for the original subcatchments. It may be that the damping effect of the catchment compensates for the lack of rainfall temporal variability. When using the disaggregated scheme on artificial urban catchments, the peak flow performance is somewhat decreased (but is comparable to radar data) and larger errors in the timing of the peak are observed (15.5%). This confirms that urban catchments requires a fine spatial and temporal resolution to adequately reproduce the streamflow.

Hence the importance of spatial rainfall for runoff generation for medium to large catchments in temperate regions is confirmed. There is therefore a need to represent the rainfall spatial distribution via the use of radar data or a dense enough raingauges network. A network of 16 raingauges seem appropriate at 1000 km^2 scale, between 4 and 7 raingauges are required at 80-280 km^2. This recommendation may not be sufficient for highly convective events in summer, however these events usually generates low flows ($<$5 m^3s^{-1} at Feildes Weir), which is not of primary concern for flood management purposes. For the Mimram, a spatially averaged rainfall at the subcatchment scale is sufficient due to the damping behaviour of the chalk geology. This recommendation would not hold as the urbanisation increases. In order to reduce the impact of modelling discretisation, a small catchment subdivision (10 km^2) is applied. Therefore no spatial rainfall information is lost when using the raingauge network and radar pixel aggregation is done at small scale (average of 2-3 pixels per subarea). Further research to vary the subdivision scale and assess its impact on the flow could be undertaken. A subjective manual calibration procedure has been carried out, however the comparison between simulated runoff response is made to the reference flow so that only the effect of spatial rainfall is assessed. Ideally, the whole catchment to Feildes Weir would have been modelled directly and the hydrograph at interior point generated to evaluate the impact of scale. This option has not been followed for two reasons, a) too large a number of subareas would have been generated and leading to computational problems, b) the subcatchments have very different behaviours that justify a spatial variation of model parameters. Improvement in the modelling strategy would include automatic calibration, incorporation of a larger number of subdivisions in the model structure and the possibility to vary model parameters at subcatchment scale automatically. A simple spatial interpolation scheme based on Thiessen polygon method is used here to calculate the areal

rainfall. Methods that take into account the distance between the estimation point and the observation stations such as inverse distance weighting and multi-quadratics (Hwang et al., 2005) may be more suitable to model summer events. Possible extension of this work involves the use of continuous simulation to have a better understanding of the effect on antecedent conditions.

Chapter 13 Use of a GLM to simulate daily rainfall for input to a rainfall-runoff model

13.1 Introduction

This chapter describes the use of a GLM (Chapter 6) to simulate spatially correlated daily rainfall, and the subsequent use of this simulated rainfall as input to a rainfall-runoff model to model flow for a medium sized catchment (3315 km^2) in North East England. The objective was to test the use of the GLM and the impact of the simulated rainfall on modelled flows, with the performance of the GLM assessed by analyses of both the generated rainfall and flow series.

13.2 Methodology

The application of a GLM was undertaken using the GLIMCLIM software package (Chandler, 2002) following the guidance provided in the User Manual, and the method followed within Chapter 8. A GLM is fitted to a set of observed point rainfall series using an occurrence model to predict the pattern of wet and dry days, an amounts model to determine the rainfall on wet days and the application of spatial dependence to provide spatial correlation between the point values.

The catchment chosen for testing the use of a GLM with a rainfall-runoff model was the Ouse to Skelton, a few kilometers upstream of York, in the north-east of England. The catchment drains the east side of the Pennines with three main contributing sub-catchments of, from North to South, the Swale, Ure and Nidd each with its headwaters in upland moors draining down to low-lying agricultural land. Average annual rainfall ranges from 2000 mm in the headwaters to 625 mm over the lower catchment. The catchment has been previously modelled using a semi-distributed continuous simulation model, CLASSIC (Crooks, 2002), which uses grid based inputs of climatic data, rainfall and potential evapotranspiration (PE), to simulate mean daily flow. The catchment is large enough for a daily time step to be appropriate for modelling flood discharges while a grid square size of 10 km (100 km^2) was selected to allow for spatial variation of climate, soils and land use.

13.3 Model formulation

The rainfall data used to calibrate the GLM were mean daily rainfall for 47 grid squares covering the Ouse catchment determined using an automated version (Gannon, 1995) of the CEH triangle method (Jones, 1983). It should be noted that the GLM was developed for use with point rainfall data which have a greater variability than areally averaged data. However, the grid-averaged rainfall data were readily available and provided long term data sets without missing values. Using grid averaged data also meant that rainfall series simulated with the GLM could be used to drive CLASSIC, with resulting modelled flows directly comparable with those from previous work. Data

for a 20 year calibration period from 1980 to 1999 were used to fit a GLM for the Ouse catchment. Each grid square was defined by the grid reference of the mid-point and the mean altitude of the whole square, determined from a DTM (Digital Terrain Model). The 47 squares were divided into seven regions, shown in Figure 13.1, roughly corresponding to the sub-catchments defined by gauging stations, see Table 13.1.

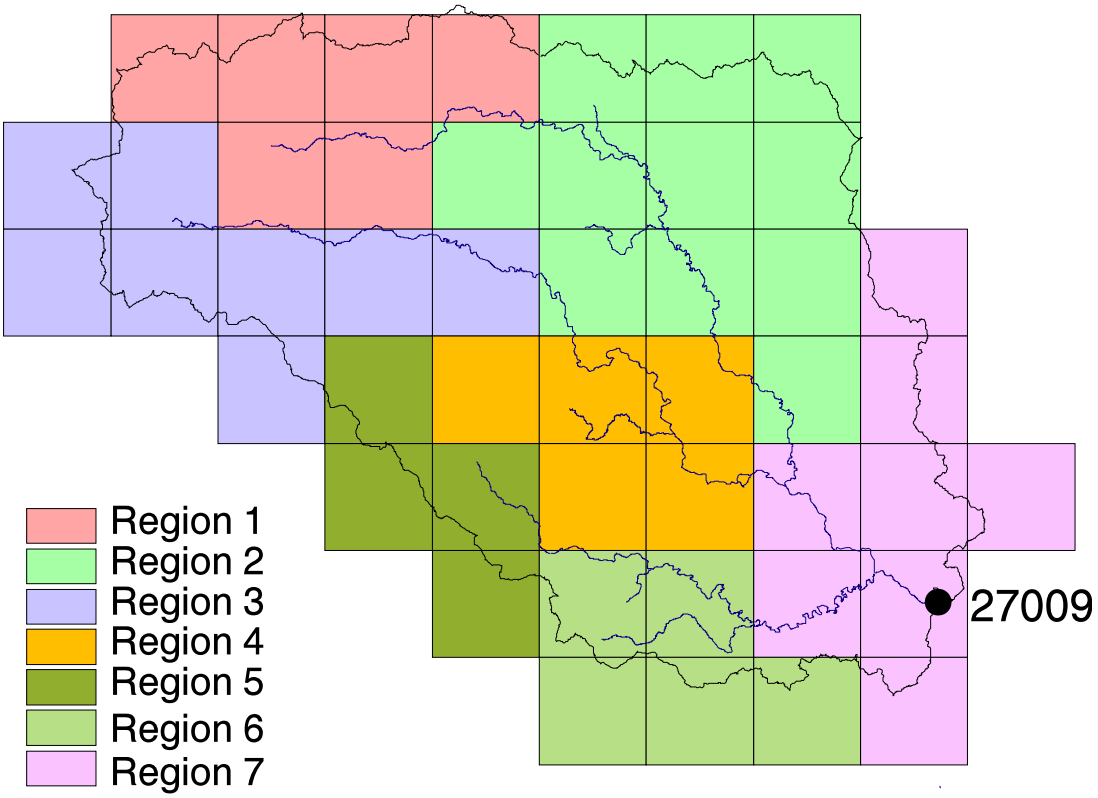


Figure 13.1 Map of the Ouse catchment to Skelton (27009) showing the 10 km grid squares and the seven regions.

Table 13.1 Regional division of Ouse catchment to Skelton

Region	Name	Number of grid squares
0	Whole catchment	47
1	Upper Swale	6
2	Lower Swale	11
3	Upper Ure	8
4	Lower Ure	5
5	Upper Nidd	4
6	Lower Nidd	5
7	Ouse	8

An occurrence model with 15 predictors was constructed with values for the covariates given in Table 13.2. Regional effects were allowed for using Fourier transforms of both the Eastings and Northings, with the main variation in an East-West direction. Seasonality was represented using sine and cosine components with temporal dependence included via indicators of rainfall on the previous 5 days. The 'binary

weather state' model was used to specify the spatial dependence structure with the value of parameter 1 as given in Table 13.2. Other components investigated included persistence and interactions between seasonality, regional effects and temporal dependence but did not result in a significant improvement in the fit of the model.

Table 13.2 Covariates used in Occurrence model

Component	Value	Covariate number
Constant	-1.0637	1
Altitude (m/100)	-0.0253	2
Fourier sine component 1 for Eastings	0.3126	3
Fourier cosine component 1 for Eastings	-0.0163	4
Fourier sine component 2 for Eastings	0.0451	5
Fourier cosine component 2 for Eastings	0.0188	6
Fourier sine component 1 for Northings	-0.0242	7
Fourier cosine component 1 for Northings	-0.0453	8
Daily seasonal effect, cosine component	0.2048	9
Daily seasonal effect, cosine component	-0.0217	10
Mean of I(Y[t-1] $>$ 0)	2.2930	11
Mean of I(Y[t-2] $>$ 0)	0.1167	12
Mean of I(Y[t-3] $>$ 0)	0.1333	13
Mean of I(Y[t-4] $>$ 0)	0.2211	14
Mean of I(Y[t-5] $>$ 0)	0.1320	15
Trace threshold	0.1000	
Parameter 1 in spatial dependence model	4.1538	

The amounts model was constructed in a similar way with the final model having 17 predictors to allow for seasonality, regional effects (again most dominant in an East-West direction), temporal dependence and interactions between regional effects and seasonality. The values for the covariates are given in Table 13.3. The observed Anscombe residual correlation matrix was used to define the spatial dependence; inter-site correlations range between 0.36 and 0.97 so it was thought that use of an average correlation was not appropriate.

The mean Pearson residuals provide a check on the fit of the models for monthly and annual rainfall sequences and regional variation. These are given in Figure 13.2 where the dotted lines show the 95% limits (1.96 standard error of mean) and the red line the mean residuals. Both the monthly and annual residuals indicate that the GLM is fitting the temporal variation, but the spatial variation, represented by the site residuals is not well represented. To provide a further indication of the spatial distribution of the residuals they have been plotted against Easting and Northing for both the occurrence and amounts models, as shown in Figure 13.3. It is possible that patterns in the residuals, for example for Northings 4.85 and 4.95, could be reduced or removed by inclusion of further 'site' parameters. However, as the objective was to test the methodology rather than develop a 'final version', the model was used to generate rainfall data.

Table 13.3 Covariates used in Amounts model

Component	Value	Covariate number
Constant	0.7370	1
Altitude (m/100)	0.0573	2
Legendre polynomial 1 for Eastings	-0.4218	3
Legendre polynomial 1 for Northings	-0.0843	4
Legendre polynomial 2 for Eastings	0.2614	5
Legendre polynomial 2 for Northings	-0.1130	6
Legendre polynomial 3 for Eastings	0.0512	7
Legendre polynomial 3 for Northings	0.0806	8
Daily seasonal effect, cosine component	0.2025	9
Daily seasonal effect, sine component	-0.2022	10
Mean of I(Y[t-1] $>$ 0)	0.5021	11
Mean of I(Y[t-2] $>$ 0)	0.0817	12
Mean of I(Y[t-3] $>$ 0)	0.1423	13
2-way interaction: covariates 3 & 9	-0.3541	14
2-way interaction: covariates 3 & 10	-0.0129	15
2-way interaction: covariates 9 & 11	0.2877	16
2-way interaction: covariates 10 & 11	0.1479	17
Trace threshold	0.1000	
Dispersion parameter	0.6323	

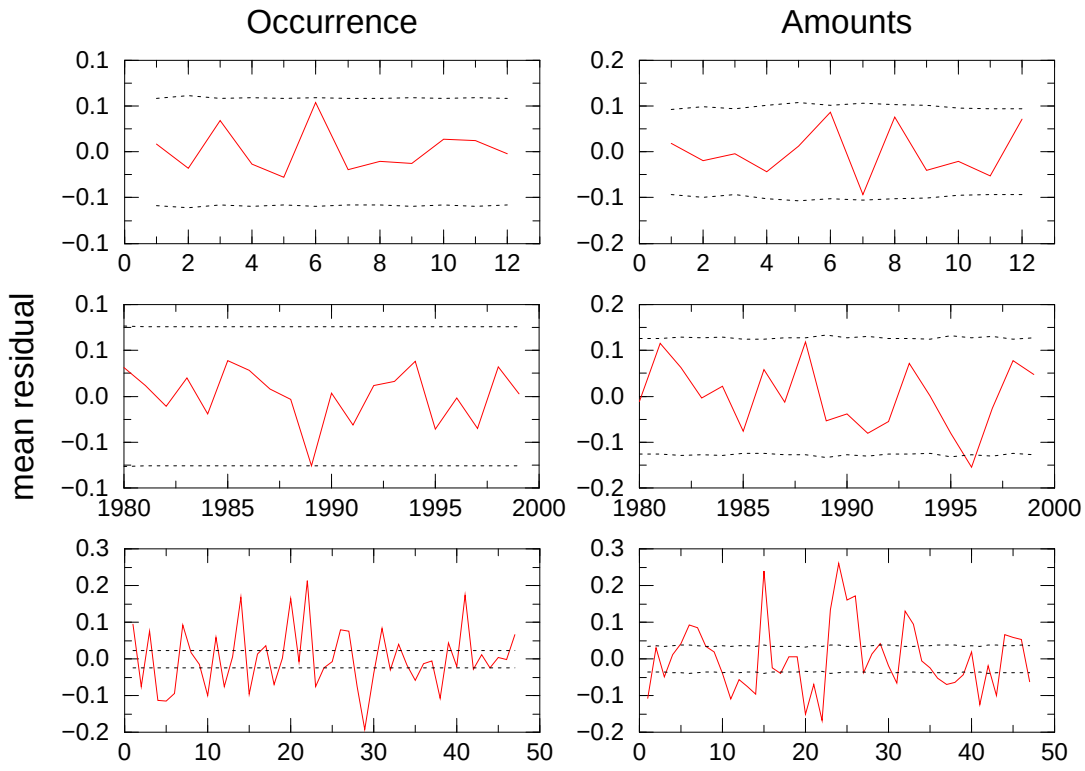


Figure 13.2 Mean Pearson residuals from occurrence and amounts models: top - monthly means, middle - annual means, bottom - means of sites. Red line mean residuals, black dotted lines 95% confidence limits

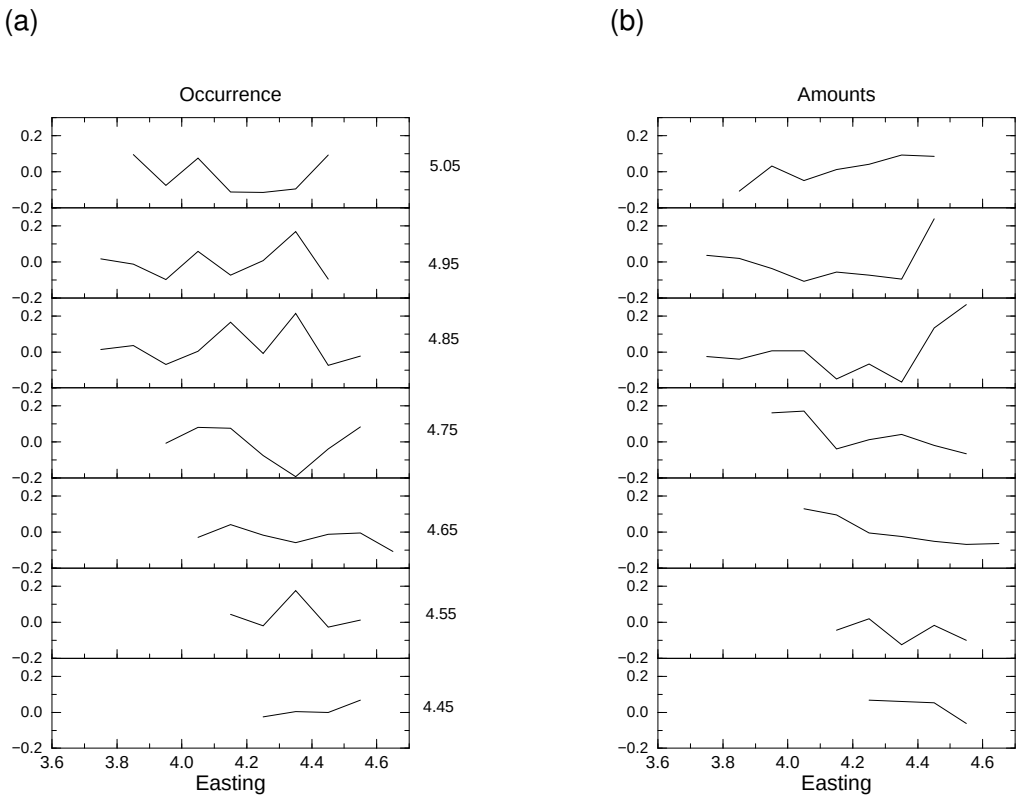


Figure 13.3 Mean Pearson residuals from (a) occurrence and (b) amounts models plotted against Easting. The seven graphs for each model are for different Northings as given by the number in the middle.

13.4 Rainfall simulation

The occurrence and amounts models, together with the observed residual correlation matrix, were used to simulate ten sets of 30 years of daily rainfall for the 47 grid squares using data for 1969 to initialise the simulations. These sets of daily rainfall data were used as input to the rainfall-runoff model, CLASSIC, using monthly PE data for 1970 to 1999 for each grid square, determined from MORECS values (Thompson et al., 1981).

13.5 Results

Peaks-over-threshold (POT) analyses of the generated rainfall and flow series were performed to compare, firstly, the extremes within the generated rainfall values with 30 years of observed rainfall data (1970 to 1999) and, secondly, flows simulated from observed rainfall with those from the generated rainfall. For the comparison of rainfall extremes the analyses were undertaken seasonally (Winter - DJF, Spring - MAM, Summer - JJA, Autumn - SON), using catchment average rainfall from one of the simulated rainfall series for 1 day, Figure 13.4(a), and cumulated over 5 days, Figure 13.4(b), the latter being a typical duration for generation of floods on a catchment the size of the Ouse. Figures 13.4(a) and 13.4(b) show that winter extreme rainfall is well simulated by the GLM but in the other seasons the GLM is underestimating both 1-day

and 5-day cumulated rainfall. This underestimation is probably a result of both the spatial dependence of the rainfall as well as the temporal distribution for each grid square.

POT analysis of the simulated flow series are given in Figure 13.5, where fitted flood frequency curves for all 10 generated rainfall series are shown by the red dotted lines with the red triangles showing the modelled peaks from the series with highest and lowest maximum peak. Flood peaks modelled from observed rainfall are shown by the black circles and the fitted frequency curve by the solid black line. Figure 13.5 shows that the flood peaks generated by the GLM rainfall series compare well with those modelled from 30 years of observed rainfall. This is not surprising as the flood events occur predominantly in the winter for which, as shown in Figures 13.4(a) and (b), the generated extreme rainfalls compare well with the observed. The magnitude of the flood peaks is greatly influenced by the spatial dependence of the rainfall, as using generated gridded rainfall series for the catchment with no spatial correlation has shown (Reynard et al., 2004).

As a further comparison between flows modelled from observed and simulated rainfall, Figure 13.6 shows flow duration curves which provide a comparison across the complete range of flows. Flows modelled from observed rainfall are shown by the black line, those from the 10 GLM rainfall series by the red dotted lines. Apart from the most extreme flows the generated rainfall series simulate higher catchment runoff than that simulated with observed rainfall. The spatial distribution of the generated rainfall was investigated by comparing the mean monthly observed rainfall for the whole catchment and for each region (see Table 13.1) with the envelope obtained from the rainfall simulations. Figure 13.7 shows that for the headwater catchments, regions 1, 3 and 5 there is a strong seasonal pattern to the monthly means and the observed means mostly fall within the GLM envelope. However, for the lower catchment areas, particularly region 7, the seasonal pattern is not as strong and the GLM is overestimating the rainfall in all months. For region 0, the whole catchment, the observed mean monthly rainfall lies at, or just below, the GLM envelope. For all regions the monthly mean for June is not well reproduced by the fitted GLM.

13.6 Notes on use of GLIMCLIM

The use and specification of the GLM was undertaken by CEH as a test of using the GLIMCLIM software. Some observations are provided here regarding the ease of use of this software.

The main work in fitting a GLM is building the occurrence and amounts models by selecting/finding the most appropriate (statistically significant) covariates. The number of possibilities from combinations of components for site, year, month and day effects with interactions, transformations and spatial dependence, defined in Tables 2 to 8 of Chandler (2002), is vast. For the first time user the obvious way to begin to build a model is to start with the covariates used in the examples and then experiment. However, as this is potentially a very time consuming exercise to find the 'best' model, more guidance on when to use (or what to look for in the output that would indicate use of) month/day effects, transformations Chandler (2002, Table 5) etc would be very helpful and aid in avoiding a pure trial and error approach. In particular, further

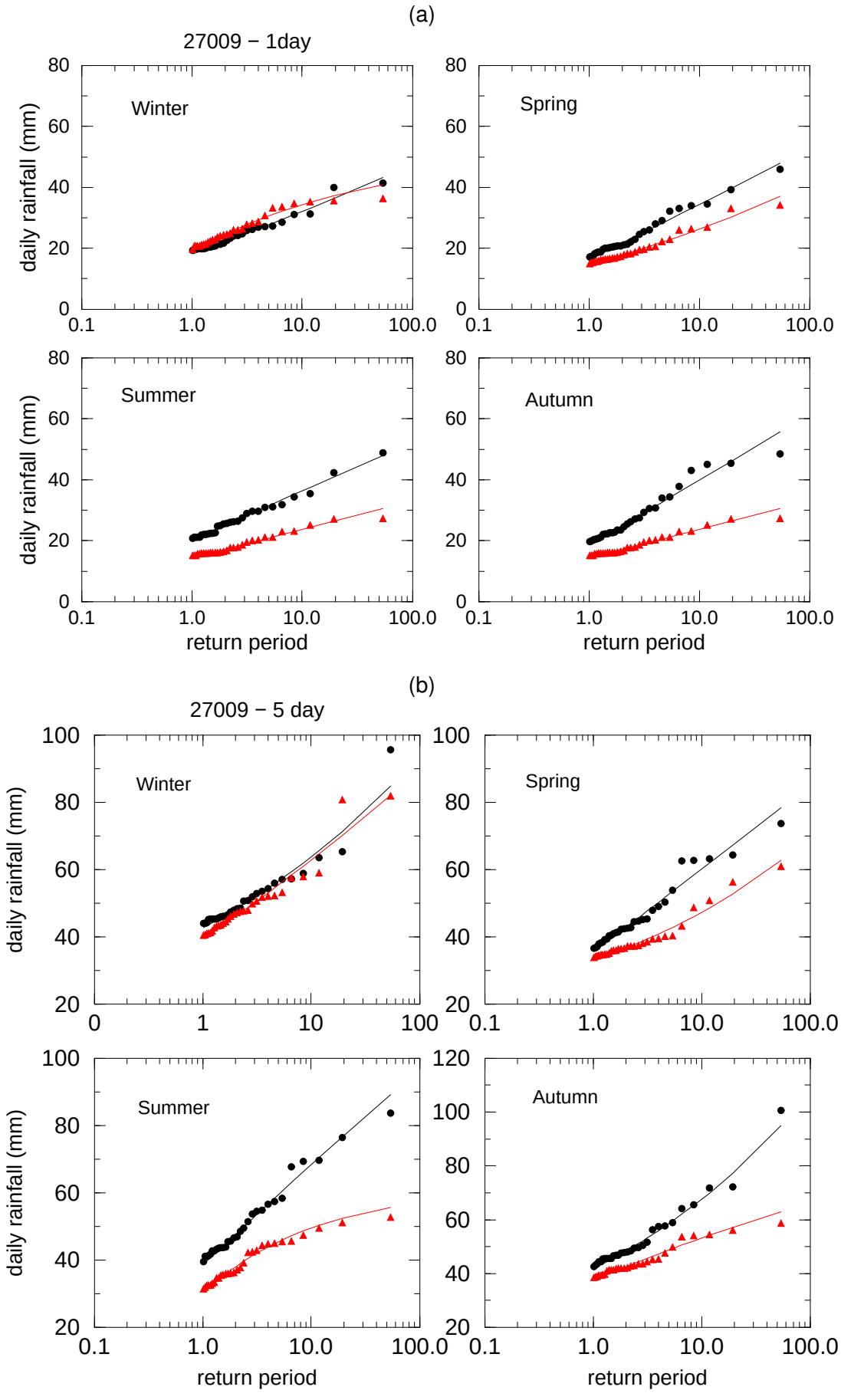


Figure 13.4 Seasonal frequency analysis for 30 years of mean catchment (a) 1 day and (b) 5 day observed (black line and circles), and generated rainfall (red line and triangles)

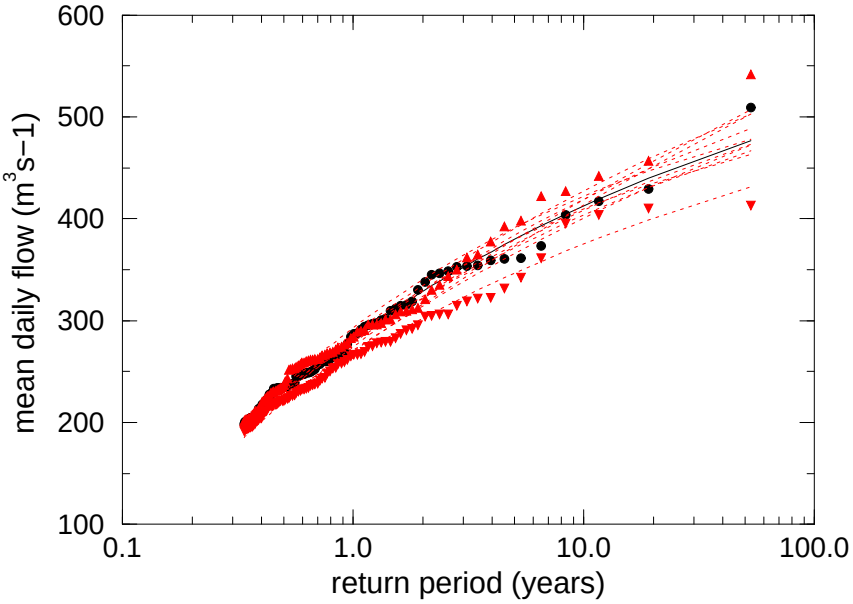


Figure 13.5 Flood frequency curves for the Ouse at Skelton: modelled flows from observed rainfall (black line and circles), modelled flows from 10 GLM simulated rainfall series (red dotted lines), flood peaks from GLM rainfall series generating highest and lowest maximum peaks (red triangles)

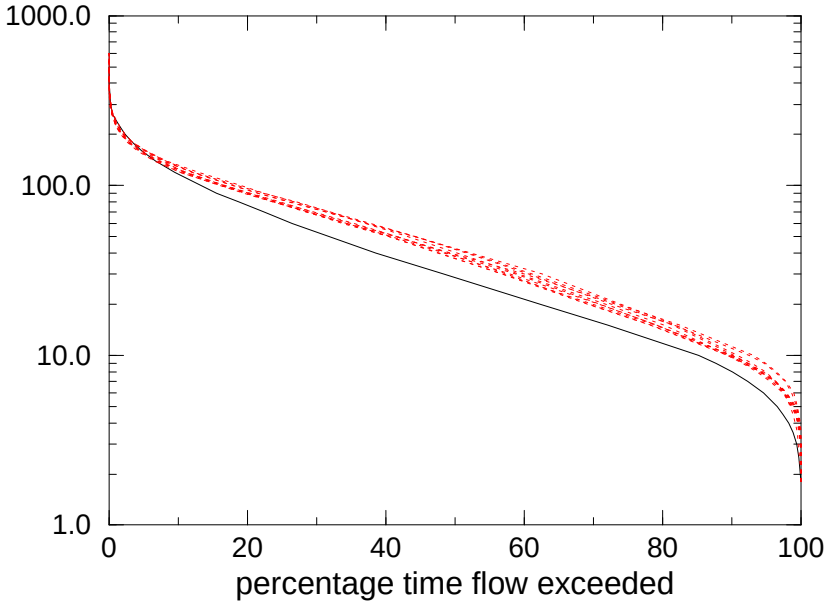


Figure 13.6 Flow duration curves for the Ouse at Skelton: modelled flow from observed rainfall (black line), modelled flows form 10 GLM simulated rainfall series (red dotted lines)

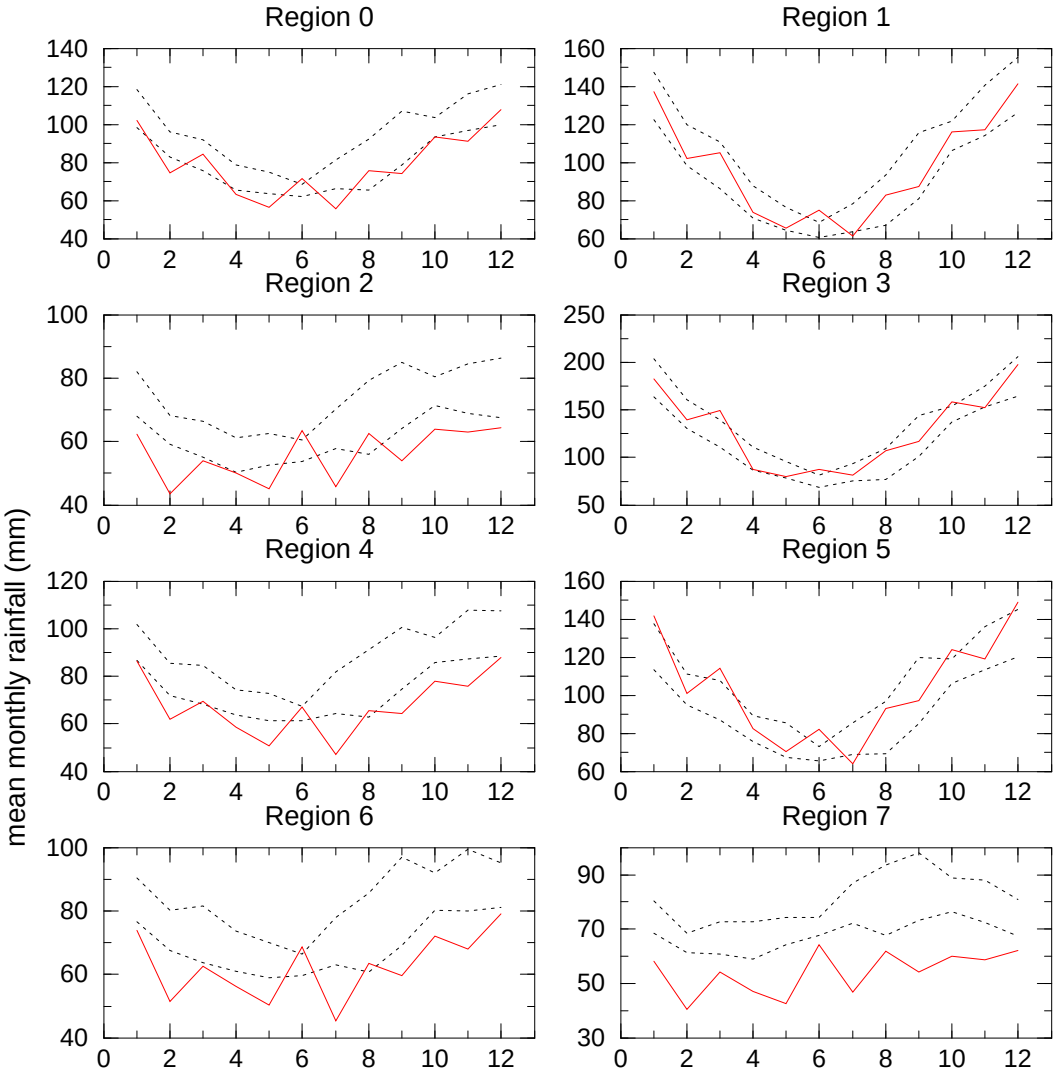


Figure 13.7 Mean monthly rainfall by regions. Red line - observed rainfall 1970-1999, black dotted lines - envelope from 10 GLM 30 year simulations.

explanation for site effects Chandler (2002, Table 4) as to the characteristics that would most benefit from the transformations or Fourier series and Legendre polynomial representations would allow a more informed approach to model building. The provision of code for plotting bubblemaps, as for plotting the Anscombe residuals, would be useful. For simulation, further explanation of the implications of using the two spatial dependence structures in the occurrence model is required.

The GLM was used to model rainfall over a larger area (3315 km^2) and with a greater rainfall gradient than that included in the development of the software, for which the spatial dependence structures were primarily designed for small areas. It is noted that this study took the simulation methodology almost to the limits of what had been intended but there were never any problems in applying the software to the Ouse catchment and the resulting model simulated realistic daily rainfall sequences. Information on spatial limits and rainfall variability used in development and testing could be provided in the User Guide with guidance on usage beyond the tested range.

The software and associated user documentation (Guide, README and output files) were generally straightforward to follow and error messages, usually caused by creating invalid model structures, were normally self explanatory. It was a rewarding process to build the two rainfall models and use them in conjunction with a semi-distributed rainfall-runoff model to generate realistic flow hydrographs. Additionally, it provided the opportunity to understand more about the rainfall characteristics of the modelled catchment.

13.7 Conclusions and chapter summary

The GLIMCLIM software was used to generate spatially gridded data over the Ouse catchment. The study, undertaken by CEH served as a test for use of the GLM, along with testing the application jointly with a spatially distributed rainfall-runoff model.

In terms of reproduction of rainfall characteristics it was found for the model specified that the temporal pattern of the generated rainfall (specifically monthly means) is good for the upper catchments where the rainfall is heavier. Rainfall for the lower catchment is overestimated in all months. Extreme rainfalls, in terms of 1-day and 5-day catchment averages, are well simulated in the winter but underestimated in all other seasons. No check was made of rainfall extremes at single sites (grid squares). It is noted that the GLM was applied to spatially averaged data (over a grid), with irregular number and location of point data within each grid. The GLM was designed for point data, and hence the reproduction of observed statistics can in some part be attributed to the spatial averaging process.

As a note on use of the software, whilst it was found specification of the model was easy, it was found that constructively interpreting the output from GLIMCLIM while fitting the occurrence and amounts models was less easy.

Regarding the flood frequency curves generated, it was found that the flood frequency determined from flows simulated from the GLM series compares well with that simulated from observed rainfall, as flood events are predominantly caused by heavy rainfall during the winter over the upper catchments. Flows, apart from those during flood events, are too high as for four of the regions the rainfall is overestimated. If

rainfall from the GLM were to be used for further study, the occurrence and amounts should be revised to try and improve, at least, the spatial distribution of the rainfall (reduce the mean site residuals).

Chapter 14 Summary and Conclusions

This report has addressed the need to provide continuous inputs of rainfall and potential evaporation to rainfall-runoff models to support continuous simulation for flood risk assessment. A range of methods has been developed and tested. Many of the modelling tools are appropriate for immediate practical application (provided that the appropriate 'health warnings' are heeded). Others need further developments, in data availability, model development and testing; for these we report progress, comment on the potential of the methods, and provide guidance for future work. For clarity, the Conclusions are presented sequentially for the individual work packages (Parts I to IV above).

In Part I we consider the stochastic generation of single site rainfall and potential evaporation time-series. The rainfall models are first tested using long raingauge data records representative of different parts of the UK. This is followed by testing with rainfall-runoff models using shorter data sets to represent catchment average rainfall.

Part II provides methods based on more generally available data sets, namely spatial networks of daily raingauge data. Generalised Linear Models, an extension of regression models, are used to model the distributions of rainday occurrence and the daily rainfall depths, and can represent both spatial non-stationarity and climate variability. Indeed, atmospheric indices, such as the North Atlantic Oscillation, can be used as predictors. To produce subdaily rainfall inputs from the daily rainfall spatial fields, disaggregation is needed, and a simple scheme is developed and tested whereby the models of Part I are used to disaggregate daily to hourly rainfall at an index site, and the temporal pattern is applied to the spatial network, scaled to preserve the spatially-varying daily falls.

In Part III we consider a more general approach to spatial-temporal rainfall modelling, which represents rainfall in continuous space and time. Previous models of this type have assumed statistical stationarity (i.e. topography and location effects were not represented, nor climate variability), and their development was based on short radar sequences. Here we derive a long radar sequence for model testing and explore a new approach to represent spatial non-stationarity.

Finally, Part IV considers the importance of spatial rainfall for flood estimation. A literature review considers the significance of spatial rainfall representation as a function of rainfall type, catchment type and scale. This is complemented by a rainfall-runoff simulation study, based on the River Lee, North of London, in which different representations of spatial rainfall, using raingauges, radar, and the disaggregation scheme of Part II, are used as inputs to a distributed hydrological model. On the basis of the review and these numerical results, the significance of spatial representation is evaluated.

14.1 Part I: Single site modelling of rainfall and potential evaporation

The task addressed in Part I is the development of a national procedure for single-site or catchment average rainfall and evaporation modelling for the continuous simulation of streamflows.

In the literature, a number of single-site models have been developed and validated with a range of rainfall data.The work presented in Chapter 2 identifies the most appropriate model for application in the UK. Because it is not feasible to express the likelihood function in a tractable form, these models have generally been fitted using the Generalized Method Of Moments (GMOM). (Appendix C summarises the analytical expressions of the most important statistics characteristic of rainfall.) The GMOM is however unable to ensure a good reproduction of extreme values since it is based upon the use of statistics reflecting the average behaviour of rainfall. Moreover, a number of authors have found the parameter estimates it produces to be poorly identified. The work presented here therefore addresses the issue of the uncertainty in parameter estimation. It exploits the poor identifiability characterising the GMOM method to develop a two-phase fitting method. This chooses, among parameters with non-significantly different performances in terms of the average rainfall characteristics, those which optimally reproduce the extremal behaviour of the observed rainfall.

This methodology is used to compare the performance of a number of models based upon Poisson-cluster processes at three different sites in the UK. A model is chosen on the basis of the GMOM, and the model variant is then selected on the basis of its optimal extreme-value performance. The chosen model is the Random-Parameter Bartlett-Lewis model with an exponential distribution for cell intensities. Its reproduction of extreme-values is significantly improved as a result of the two-phase fitting method, as compared with the performance recorded in the literature. Model parameter uncertainty results using various record lengths also appear to identify genuine opportunities for parameter bounding, which with further work will aid application of these models in a national setting.

Although the performance is improved with the introduction of the two-phase fitting method, some deficiencies remain in terms of reproduction of extreme statistics. We conclude that the model and optimisation technique are appropriate for use with caution, with the user undertaking careful verification of the simulated outputs. More extensive testing is required on a wider range of sites before the model can be recommended for general application across the UK. It is recommended that alternate means of improving the extreme value performance of the model (such as the use of skewness in fitting) are compared to the newly introduced fitting technique.

In Chapter 3, methods are presented for the modelling of Potential Evaporation (PE). The models developed here are heteroscedastic autoregressions. They allow for the simulation of weekly or daily sequences, and also that of realistic daily sequences consistent with a given weekly sequence. The weekly PE is modelled using joint mean-variance models. Seasonality and linear trends in observed MORECS data are represented successfully by these joint models. Using daily data from a single site, a model is developed which represents daily PE as multivariate normal distribution conditional on given weekly evaporation totals. This can either be used as a

disaggregation tool to generate daily PE values from weekly (e.g. MORECS) PE data, or together with the weekly PE generator to enable the simulation of daily PE values.

Chapter 4 presents the results from an extensive program of testing that examines the performance of the single site rainfall models for several UK catchments. Catchment average rainfall data are used, as a precursor to joint testing of the rainfall and rainfall-runoff modelling procedure developed under FD2106.

The results of fitting the single-site model to these data are broadly satisfactory, but there are some notable discrepancies. In particular, some problems in the characterisation of mean rainfall depth and the persistence of wet and dry episodes have been identified. A first analysis found that the original formulation of the model led to the appearance of an unrealistically large number of high magnitude rainfall events. This was resolved by a small modification to the simulation methodology. The modified simulation scheme was tested on three catchments, where the results were much improved. Further testing is required before conclusions regarding the model's performance in different regions in UK can be reached.

Part of the problem of reproducing rainfall characteristics may lie in the way in which the hourly catchment average data series is produced. Hourly catchment average series are obtained by first averaging single-site daily series for all the available gauges in the catchment. Using the hourly profile at a recording raingauge with data of acceptable quality (or a weighted average of such profiles if several such gauges are available), hourly catchment averages are obtained. It is emphasised that the models applied here were designed for use with single-site rainfall (not catchment average rainfall) - with no account taken of catchment statistics being different than those of a single site. It is recommended that the models applied here are further tested regarding the conditions of applicability in terms of catchment size and data availability. Analysis of observed data might allow this correction to be estimated empirically, improving applicability for larger areas.

Chapter 5 completes the investigation into the use of the selected single-site model for the representation of catchment averages. The focus is on the ability to generate realistic streamflows by inputting the generated rainfalls into the continuous rainfall-runoff model developed in FD2106. The analysis has shown that hourly flow data are generally well simulated. However, flow duration curves based on simulated data are generally underestimated, while an analysis of peak-over-threshold (POT) series has shown that flood peaks are generally overestimated, especially for return periods greater than ten years. The underestimation of the flow duration curve is probably linked to underestimation of rainfall persistence within the rainfall model. Consequently, antecedent catchment wetness, conditioned by the presence or absence of antecedent rainfall, may not be adequately simulated and estimates of flow are therefore too low. Some overestimation of flood frequency curves may result from the occassional generation of very large storms by the rainfall model.

We conclude that the models are appropriate to use with caution, but careful assessment on a large number of catchments would be needed before the single-site Random-Parameter Bartlett-Lewis rainfall generator could be used with confidence to provide data for flood risk assessment for ungauged catchments. Any user should validate the model using local data to have sufficient confidence for practical application. More work is recommended to pursue the issue of parameter constraints

and non-feasible simulations. There are also ideas for model extension that should be explored to seek improvements in performance.

14.2 Part II: Spatial-temporal modelling based on daily rainfall modelling and spatial-temporal disaggregation

Part I of this report has been concerned with the development and application of models for temporal rainfall at a single site. The models are constructed in continuous time, enabling their properties to be aggregated into arbitrary discrete time intervals for model fitting and for applications. In principle, these stochastic models can be generalised to simulate spatial-temporal fields of rainfall (such as that presented within Chapter 9). At present, however, such stationary models suffer from some major drawbacks. First they rely on radar data for calibration. Apart from the data quality issues discussed in Appendix F, radar technology is relatively new so that record lengths are limited. In view of the general consensus that climate varies on decadal or longer timescales (IPCC, 2001, Section 1.2.2), it may be unwise to rely too heavily upon such short runs of data to provide model parameterisations that will hold into the future. A second drawback of the stationary stochastic models is that they are unsuitable for reproducing such hydrologically important features as long term changes in climate or systematic orographic effects. This is addressed to some extent in the developments presented in Section 6.5.

Although radar and subdaily raingauge data are relatively scarce in the UK at present, long daily raingauge records are much more abundant. This suggests that for short to medium term application, it would be useful to develop models for daily rainfall that can be used to study nonstationarities and to simulate rainfall sequences at a daily timescale. In principle, the output from such models can be downscaled to any desired resolution using some appropriate method. Hourly data may be adequate for most flow simulation purposes.

Chapter 6 provides an overview of the models for daily rainfall sequences. Examples of their application to two contrasting data sets are provided in Chapter 7. Chapter 8 then presents and illustrates a methodology for disaggregating a multisite daily rainfall sequence to a subdaily time scale. When combined with the capability for simulating multi-site daily sequences, this enables (possibly nonstationary) multi-site subdaily sequences to be generated.

Many techniques are available for modelling and simulating daily rainfall sequences; some of these are reviewed by Wilks and Wilby (1999), for example. Here, the aim is to produce a methodology that is:

- Conceptually simple
- Suitable for routine application, in terms both of computational demands and of the time taken to fit, validate and simulate models
- Capable of representing realistically, and in reasonable detail, the observed temporal structure of daily rainfall sequences in the UK

- Suitable for the generation of multi-site rainfall sequences, at scales up to a few thousand square kilometres

The final point above ensures that the methodology is suitable for application to all but the largest UK catchments.

In Chapter 6, GLMs are introduced as an intuitive and interpretable means of modelling daily rainfall sequences, either at a single site or at multiple sites. They are able to accommodate nonstationarity, through covariates that change in space or time; and are also able to generate dependent sequences at neighbouring sites. The GLIMCLIM software package (Chandler, 2002) provides a means of fitting and simulating the models, without worrying too much about the technical details or the manipulation of large volumes of data. However, as with all modelling tools, care is required in the implementation of the methodology. Some guidelines are given in the GLIMCLIM user manual (Section 6). Users should also note the following:

1. The models are designed to be fitted to data from a network of daily rain gauges. The amounts models, in particular, assume that the coefficient of variation in rainfall amounts is constant. Our experience is that this is true in general of daily raingauge data in many regions of the world. However, it may not be true for gridded data derived by interpolation from a network of gauges, because the variability at each grid node will depend on the density of gauges in its neighbourhood (among other things). In general therefore, we do not recommend that the models are fitted to gridded data. If gridded rainfall data are required (for example, as input to a rainfall-runoff model), we recommend that GLMs are fitted directly to the available daily raingauge records and then simulated on a regular grid. If necessary, the simulations can be conditioned on all available rain gauge data (this is essentially an imputation exercise, which can be handled by GLIMCLIM). This ensures that the simulated sequences have, as far as possible, the properties that would have been observed using raingauges at each grid node.

2. The spatial dependence structures available in the GLIMCLIM software (in particular relating to rainfall occurrence) are designed primarily for use when inter-site dependence is uniformly high. This is the case at spatial scales where most sites can reasonably be assumed to be affected by the same weather systems on any day; hence the dependence structures are probably suitable for all but the largest UK catchments. At larger scales (say, above 2000 km^2), the methodology may still produce reasonable results; however, users should check carefully to ensure that the results of any simulation exercise appear reasonable.

3. The beta-binomial model for dependence in rainfall occurrence is unlikely to capture the strong dependence between sites that are too close together (e.g. tens of metres). It is therefore advisable to ensure a reasonable spacing between sites in any simulation exercise (note, however, that there is no problem with *fitting* the models to data from sites that are very close together).

4. A beta-binomial distribution for the number of wet sites may not always be compatible with the probabilities of rain predicted by the logistic regression model for rainfall occurrence. If the GLIMCLIM software package encounters

such incompatibility, it will report a warning and modify the predicted probabilities. However, in practice such incompatibility is caused by predicted probabilities that are unrealistic (usually because the probabilities at one or two sites are very different from the remainder). Warning messages from the software should therefore be taken as an invitation to improve the rainfall occurrence model.

Chapter 7 has illustrated the use of GLMs to model and simulate daily rainfall sequences, and has also highlighted some issues that are likely to be encountered when using daily raingauge data from the UK. The results from the two extensive and contrasting case studies were very similar, which gives some confidence that they can be used as a guide to what can be expected when applying the methodology at other locations in the UK.

The main points to emerge from these applications are as follows:

- Daily rainfall records in the UK should be scrutinised carefully before use. Duplicate records should be removed to avoid biasing any analyses, as should unflagged monthly totals. A further discussion of data quality issues can be found in Yang et al. (2006a).

- After excluding records that are clearly erroneous, daily rainfall records can exhibit apparent spatial inconsistencies. These affect rainfall occurrence in particular, and are in part caused by differences in the characteristics of very small rainfall amounts. The reason for these differences is unknown. Having experimented with a variety of different schemes, and compared simulation results, it seems that thresholding the data using a soft threshold of 0.5 mm can help to remove inconsistencies without affecting the overall rainfall properties to any appreciable degree. Although such thresholding does not always remove the inconsistencies completely, it does seem to reduce them. As far as rainfall occurrence is concerned, this reduction can be substantial. We therefore recommend thresholding daily rainfall in this way as a matter of course, prior to analysis.

- The GLM framework provides a natural and interpretable means of incorporating systematic structures into simulated rainfall sequences. As well as modelling the dynamics of the rainfall process itself, corrections can be made for features such as changes in recording resolution and different gauge types. Such features can, and should, be accommodated via the use of indicator variables.

- In both of the case studies considered here, the fitted models had very similar structures. The main differences were in the representation of regional variation, which is more complex for northeast Lancashire than for the Blackwater. This is to be expected, since northeast Lancashire has a more complex topography. The models for the two regions were built independently of each other. Their similarity therefore suggests that the covariates identified here may be suitable for use in other areas of the UK.

- Pearson residuals can be used to check the systematic structure of the fitted models. The GLIMCLIM software calculates the mean and standard deviation of

Pearson residuals by month, site and year; plots of these can be used to highlight any structure that is not captured by the models.

- GLM simulations are able to reproduce a wide variety of properties of observed rainfall sequences, over a range of spatial scales (scales up to around 2500 km^2 — the size of the northeast Lancashire region — have been considered here). These include extremal properties as well as means, variances, proportions of wet days and so on. Further results can be found in Yang et al. (2006b) and Wheater et al. (2006).

- The GLM for the Blackwater data has been combined with the spatial-temporal model described in Appendix F, to generate nonstationary rainfall simulations in continuous space and time.

Elsewhere in the FD2105 project, GLMs have been fitted to additional data sets. Some results can be found in Chapters 8 and 13. In general, the performance is promising. After extensive testing therefore, we conclude that the methodology is suitable for routine use in the UK.

In Chapter 8, a simple disaggregation method has been proposed, which enables multi-site rainfall sequences to be generated at a subdaily time scale. In principle, the method can be used based on an observed record of daily rainfall from a raingauge network, with one or more observed sub-daily raingauges. More generally, long sequences of daily rainfall can be simulated, and stochastic modelling used to generate a temporally-disaggregated hourly series. The proposed modelling scheme retains a relatively simple structure, and uses the GLMs of Chapter 6 to include spatial and temporal nonstationarities in rainfall properties on a daily basis. Hourly data are then obtained using the HYETOS technique (based on the Poisson cluster models of Chapter 2) to obtain a subdaily temporal profile at a master gauge; this profile is then applied to all gauges in a catchment.

The performance of the proposed disaggregation procedure has been assessed in terms of the reproduction of a set of standard statistics and with regard to extremes. Results from a network of 21 rain gauges on the River Lee are promising. Simulated rainfall properties match the empirical statistics to an acceptable level for practical purposes. Although the model overestimates the size of the extreme value properties of rainfall, the larger observed maxima fall within a 95% confidence interval of simulated extremes. Additionally, the combined scheme has the potential to transfer information from sites included in the calibration of the GLM to nearby locations which are not used to calibrate the model.

Further research is required to improve the reproduction of the proportion of dry periods, which in turn will benefit the simulation of extreme values. Further work is possible to improve the performance of the temporal disaggregation model. The first step would be to identify single-site parameters that provide a better fit to the proportion of dry periods for the months of January, June and August; and also the lag 1 autocorrelation in October and November. A second direction where development can take place involves further investigation on the impact of the adjustment procedure within HYETOS on the overestimation of the standard deviation. Improvements in the reproduction of the proportion of dry periods and the standard deviation would lead to data that are more temporally correlated.

The major limitation of this scheme is the overestimation of the spatial correlation due to the fact that a single temporal profile is applied at all gauges. Further refinements could be introduced to address this. For example, in order to decrease this correlation, the temporal profile could be shifted in time for some of the gauges (either randomly or taking into account the distance from the master gauge). More generally, the assumption of complete spatial dependence is expected to become an increasing limitation for the generated rainfall field as the spatial scale of application increases. It will be important to identify at what scale allowance for spatial heterogeneity must be made. To address this issue, evaluation of the methodology could be extended to a larger region. Various options could be pursued for more complex procedures. One possible development would be to run independent disaggregation procedures for discrete sub-areas, another would be to build in spatial correlation to a multi-site procedure. More generally, if adequate radar data were available then the continuous space-time models of Chapters 9 and 10 could be used for subdaily data, conditioned on daily GLM simulation.

However, the scheme proposed here has the considerable advantage of simplicity and can be readily implemented. Further extension would only be required if the limitations are important in the context of rainfall-runoff modelling. We note, for example, that although the spatial variability of rainfall may be expected to increase with increasing scale, the effects of hydrological damping will increase with increasing catchment area. In the context of flood management, necessary further work includes the testing of the modelled rainfall as input to rainfall-runoff models to assess the impact of the simulated rainfall on runoff generation. A first study of this aspect is presented in Chapter 12.

14.3 Part III: Development and testing of a non-stationary spatial-temporal rainfall model

We turn now, in Part III, to consider rainfall as a fully spatial-temporal process where, for maximum flexibility, both space and time vary continuously. Here, the ultimate goal is to develop a model that can be simulated to produce high-resolution rainfall sequences. These can then be aggregated as appropriate and used as input to rainfall-runoff models for flood risk assessment and design. In Chapter 9, the spatial-temporal models are presented, in Appendix F, the development of a re-calibrated long radar record is described, and results are presented in Chapter 10.

In earlier work (FD0426) a spatial-temporal rainfall model based on Poisson cluster processes was developed, and some preliminary testing was carried out with a single, very limited, radar data set (from the Wardon Hill radar). This model was stationary both in space and time, apart from allowing basic seasonality. The report on that project noted the need for further development of the model and testing on a more substantial data set. Thus, objectives of the current DEFRA-funded project (FD2105) were to carry out further model testing, and theoretical model development as necessary, and also to develop the model to incorporate temporal and spatial non-stationarities (enabling the representation of possible climate change scenarios, and topographic effects). An additional objective was to review recent developments in

fractal methods for rainfall and to compare these with the Poisson-based approach, to ensure that our spatial-temporal model development is based on the most appropriate theory.

A substantial review paper on fractal methods methods for rainfall is attached to this report as Appendix E. The content is somewhat technical and we give here only a brief summary of the main points and conclusions. The basic idea of a multifractal model is that it enables the parsimonious description of rainfall variability over a large range of scales in space and time. Most multifractal fields are based on multiplicative cascade processes, and these may develop through either a continuous or a discrete range of scales. Continuous cascades do not take exactly zero values so that an arbitrary threshold has to be imposed to enable spatial-temporal regions to be defined as 'dry'. The generating mechanisms of discrete cascades are mathematically simpler and more intuitive, but the resulting fields have a rather artificial rectangular structure.

Much of the work on multifractal models is most applicable to the purely spatial representation of total rainfall over some fixed time period. The contrast between the directionality of time and the symmetry of spatial dimensions means that space-time processes cannot sensibly be represented simply by extending spatial models to include an extra spatial dimension. Advances are being made in research on genuinely spatial-temporal fractal models, but many outstanding modelling issues remain to be resolved. We note also that the simulation of such models is highly computer-intensive.

In general, multifractal models for (usually, spatial) rainfall have been validated in terms of their ability to reproduce the scaling properties of observed rainfall, but their performance with regard to other properties that are crucial for rainfall-runoff, such as wet-dry properties and extremes, have not yet been adequately explored. In addition, there are questions to do with parameter estimation and model identification that need to be addressed. The overall conclusion is that fractal models for spatial-temporal processes are not yet at the stage where they are readily applicable for continuous simulation of rainfall, and that therefore the Poisson-based approach is the most promising for medium-term application.

Given these conclusions, we concentrate on the development of Poisson-based nonstationary continuous simulation models in this part of the programme. Thus, there has been considerable further testing of the stationary continuous simulation model, together with development of this model to allow for spatial and temporal nonstationarities. The stationary continuous simulation model is described in Chapter 9. It comprises two main sub-models. The first is a model for the spatial-temporal rainfall field interior to a rain event; details are given in Section 9.2. The second sub-model is for the advection process by which rain events move across a catchment; details can be found in Section 9.3.

Substantial effort has been expended in developing and checking the software used for fitting the model for the rain event interiors, and for the continuous simulation. The stationary version of the model has been fitted to data from the Chenies radar, and extensive model validation has been carried out. The data and their calibration was described in Appendix F, while the results of the continuous simulation is discussed in Chapter 10. The model is fitted on a month by month basis so that basic seasonality is incorporated, but it is otherwise stationary in space and time.

A nonstationary spatial-temporal model is achieved by combining the stationary model, which provides the required finescale spatial and temporal structure, with a multiplicative rescaling field that drives the spatial and temporal heterogeneities. This extension of the stationary model is described in Section 9.5. The rescaling field is obtained by fitting Generalised Linear Models (see Part II of this report) to daily rain gauge data. A benefit of using gauge data here is that simulations of the final model will be calibrated directly to ground truth, so that careful calibration of the radar data themselves is not essential. Simulation of the nonstationary model enables ensembles of future rainfall over a catchment, allowing for its topographic features, to be explored. The results from the continuous simulation of the nonstationary model are discussed in Chapter 10.

The overall conclusions of this work are, firstly, that the model for rain event interiors has been extensively validated and provides a good representation of the basic statistical properties of the fine scale spatial and temporal structure of rainfall. Second, methodology to produce a continuous stationary simulation of rainfall across a catchment has been further developed. This has been applied in conjunction with the Chenies radar data and the simulated realisations of the model have been shown to have statistical properties that are generally quite a good representation of those of the empirical data. Third, for the first time, a way has been demonstrated to extend the stationary spatial-temporal model to allow topographic effects and possible climate change scenarios to be included. In the current report, validation of the continuous simulations has focussed on wet/dry properties and first and second-order moments of the spatial-temporal process. In future work, and especially for hydrological applications, it will be important to look at the tail of the intensity distribution and properties of extremes.

The model validation has raised many complex issues, many of which are to do with the radar data themselves and their calibration to reflect ground truth. Ultimately, the results of the continuous simulations can only be as good as the data used to fit the models, and various limitations of currently available radar data are discussed in Appendix F. We have applied the UK standard calibration procedure (Moore et al., 1994) and found that, for the Chenies data, it is possible to improve the agreement between the calibrated radar values and data from rain gauges under the radar region, by using a modified procedure. As well as improving the agreement between the distributions as a whole, emphasis has been put on reproducing the larger hourly values which are especially important for rainfall-runoff applications. A particular issue (see Appendix F) is that arbitrary choices in the method of calibrating the radar data lead to calibrated data with different statistical properties (often with reduced temporal autocorrelations) and thus, potentially, to differences in the fitted model and its continuous simulation. This sensitivity of statistical properties to the calibration procedure is a source of some concern and means that, at present, models based on radar data must be used with some caution in flood risk assessment. A substantial further study is needed to gain a thorough understanding of all the issues involved.

At present, a limitation to the use of the continuous simulation model is that the within-event model can only be fitted to data from events that cover the entire radar window for a substantial period of time, and can be clearly distinguished from radar clutter and anomalies. This means that, for example, periods of very light scattered showers or rainfall events that pass over only a part of the window are not fitted. The

omission of such events from the continuous simulation may lead to discrepancies between the properties of the simulations and those of the observed data; the significance of these omissions is discussed in Chapter 10. In the comparisons presented there, a threshold coverage of 15% has been applied to all images, with all pixels in images below this coverage being given intensities of zero. Given this threshold, the models generally perform acceptably in reproducing the statistical properties of the observed data. However, further model development is needed if the model is to be able to represent light and spatially and/or temporally intermittent rainfall.

A number of other outstanding issues have been identified both in terms of the identification and sampling of events, and the modelling of the advection process. These are discussed in Chapter 10. In connection with the former, a threshold rule has been applied to the calibrated Chenies radar series for the period 1990–2002 in order to select rainfall events to which the event interiors model can be fitted. This thresholding has resulted in the rather surprising fact that there appear to be an increasing number of rain events over the period, with a jump upwards in the number of wet events around 1994, and an increasing trend thereafter. In part, these observations can be explained by changes in radar recording practice, but inspection of rain gauge data suggests that at least part of the increasing trend during the 1990s is genuine. Such a trend would have various implications for the continuous simulation of rainfall, as discussed in Section F.5. If there are genuine changes in the numbers of events and their characteristics (and to determine long-term patterns far more than 12 years' data are needed), the need for nonstationary continuous simulation, as discussed in Section 9.5, becomes apparent.

In the light of experience with the Chenies radar data, modifications have been made to some of the procedures described in our earlier work for DEFRA (FD0426). Changes have been made to the way in which models are fitted, as well as to the software implementation of both model fitting and simulation. Considerable time and ingenuity were required to identify the causes of problems and to resolve them. However, as a result, the procedures developed to fit the spatial-temporal models and to produce continuous simulations are now much more robust and firmly based.

The extension of the stationary spatial-temporal model to allow the inclusion of topographic effects and possible climate change scenarios is a significant step forward. The results are dependent on the adequacy of the stationary model, and the rescaling method employed involves the implicit assumption that climate change would leave the wet-dry structure of rainfall unchanged. Thus the model will need to be further developed to remove this rather implausible assumption. Nevertheless, the model proposed here is an important first step in demonstrating the feasibility of obtaining nonstationary simulations. For the future, it will be of particular interest to incorporate physically-based climate change scenarios into the rainfall simulations. This can be achieved by conditioning the GLMs on the output of numerical climate models. Research in this area is currently being carried out as part of DEFRA project FD2113.

14.4 Part IV: Effect of spatial structure of catchment-scale rainfall on runoff estimation

Clearly, a key issue for the application of methods to simulate spatial-temporal rainfall fields as input to rainfall-runoff models is the importance of the spatial structure of rainfall and its representation for flood runoff generation. We address this issue through a review of the international literature, in Chapter 11, and, in Chapter 12, a numerical experiment, based on the River Lee North of London, in which alternative rainfall representations are used as input to a distributed, event-based rainfall-runoff model. The aim of these chapters is to quantify the significance of spatial rainfall as a function of rainfall type, catchment type and spatial scale. Finally, in Chapter 13, we include a relevant study, undertaken by CEH Wallingford outside the FD2105 programme, that takes the GLM methodology and applies it to the daily modelling of a UK catchment.

The review of the literature on the importance of spatial rainfall for runoff generation, reported in Chapter 11, has focussed on three types of situation: (a) urban areas, (b) semi-arid areas associated with rapid runoff and localised convective rainfall and (c) temperate regions where storms are the results of frontal and convective activities, or a combination of both, and the catchment is more heterogeneous in nature. In rainfall-runoff modelling, a simulated hydrograph is generally derived from a transformation of the rainfall input via the use of parameters that are the results of a calibration procedure at a fixed spatial scale. The modelling strategy adds another level of complexity and can mask the effect of spatial rainfall, therefore its impact should ideally be quantified independently of the effect of rainfall. However, the main advantage of watershed modelling is that it provides a framework to use rainfall patterns to test hypotheses about the behaviour of the catchment, and hence it provides the main basis for the literature reviewed.

Across all catchment types, the review stresses the nonlinearity of the runoff response to rainfall input (Lopes, 1996; Shah et al., 1996b) and highlights three main interactions: (a) rainfall, (b) catchment properties and (c) scale. The effect of rainfall includes variability in space and time, storm movement and spatial organisation. The damping effect of the basin is an intrinsic property of the catchment and can be characterised by the lag time (sometimes referred to as the time of concentration, characteristic time scale, mean runoff time, catchment response time). It is a measure of the catchment response time and expresses the average travel time from all points on the catchment to the outlet (Lekkas and Onof, 2005). As the scale increases, the importance of spatial rainfall decreases and there is a transfer from spatial variability of rainfall to catchment response time as the dominant factor governing runoff generation. Hence these three interactions principally control the runoff response; however their relative contributions differ depending on the type of catchment considered.

Small urban catchments are fast-responding and sensitive to storm movement, storm variability and event magnitude. Their hydrological modelling requires fine spatial and temporal rainfall resolution, of the order of minutes and kilometers. Since the damage as a result of flooding can be important in urban areas, there is a strong incentive to use radar data or develop disaggregation techniques (Onof et al., 2005) to meet the

requirement of urban hydrological modelling. Berne et al. (2004) show that finer spatial and temporal resolutions are required at smaller scale.

In semi-arid areas dominated by convective thunderstorms, the catchment response is also very rapid and a detailed representation of the rainfall distribution is necessary to reproduce the outflow accurately. The findings from this review highlight the importance of spatial rainfall at the small scale and suggest a spatial resolution of 1 km below 50 km^2, and 2 km between 50 and 500 km^2. This is a hard recommendation to implement in practice using raingauge data as common operational networks are usually not as extensively instrumented as the catchments reviewed citepReefsgaard96.

Contrasting observations arise from the studies in sub-humid areas where more complex interactions between spatial rainfall and catchment characteristics exist. A key paper is presented by Woods and Sivapalan (Woods and Sivapalan, 1999). Their analytical framework includes the main components of the hydrological cycle and quantifies the effects of rainfall and catchment response. The distribution of rainfall-excess with flow distance to the outlet was found to be a dominant factor influencing the runoff response. This finding was also identified by other authors (Smith et al., 2004; Naden, 1992; Dodov and Foufoula-Georgiou, 2005). In the event studied, although substantial variation in rainfall distribution was observed, the distribution of rainfall-excess with flow distance was relatively uniform so that knowledge of catchment-averaged rainfall was sufficient to model the runoff response. Obled et al. (1994); Naden (1992)and Smith et al. (2004) reached similar conclusions. Woods and Sivapalan (1999) conclude that the mean runoff time is dominated by storm duration at small scale and for larger catchments the mean travel time gradually becomes more important. Similarly Smith et al. (2004) suggest that when catchments are characterised by marked spatial variability in precipitation and less of a filtering effect of the rainfall input signal, improved outlet simulations are obtained from distributed versus lumped models. The raingauge network densities proposed by Obled et al. (1994) and Andréassian et al. (2001) as adequate to model the behaviour of humid catchments were 5, 8 and 33 raingauges at scales of 71, 1120 and 10700 km^2 respectively.

This review has identified a number of open problems in the importance of spatial rainfall for runoff modelling in temperate humid areas. It is not clear under what conditions an explicit rainfall representation is required. The findings are usually to some extent subjective and there is a need to quantify the effect of rainfall, catchment response and catchment scale independently. It is also difficult to distinguish between the effect of spatial rainfall and the impact of the model spatial discretisation.

To provide further insight for UK catchments, in Chapter 12, an investigation on the Lee catchment, in the Thames region was undertaken in the context of flood management, to explore some of the issues highlighted above. This analysis had three goals: (a) to test whether spatial rainfall is important to model medium to large scale catchments in humid areas; (b) assuming that it is important, to investigate the rainfall spatial resolution required at catchment scale, for a range of scales and catchment types; (c) to identify the modelling needs.

The requirements of a rainfall-runoff model for this analysis are as follows: (a) a realistic model incorporating nonlinear routing, (b) the capability of accepting a large

number of distributed rainfall input data to assess the importance of rainfall representation, (c) an event-based structure to analyse the impact of the type of rainfall event, (d) a small number of parameters to facilitate parameter identification of catchment response. The event-based, semi-distributed hydrological model RORB was selected to carry out the investigation. The study is based on 15 years of data from the Chenies radar, 16 raingauges and 12 flow stations from the 1000 km^2 upper catchment of the river Lee, UK.

The experimental design was constructed so that spatial rainfall was tested for different representations of spatial rainfall and different rainfall types. Selected events from the period 1987-2002 from 2km grid radar data, subdaily data from a network of 16 raingauges and subsets of the raingauges were used. To investigate the case where only daily data are available, the spatially uniform temporal disaggregation scheme developed for this project (see Chapter 8), conditioned on observed daily data, was also tested.

The Lee was divided into six subcatchments, representing the main tributaries. Each subcatchment was first modelled independently and the simulated runoff was then routed to generate the total runoff hydrograph at the watershed outlet. Hence the relationship between the spatial structure of observed rainfall and runoff production for a range of catchments could be analysed. To investigate the sensitivity to catchment type, the existing (sub)catchments were then artificially turned into fast-responding urban basins. The catchment response time was taken as an independent variable of catchment response. This modelling strategy allowed analysis of the spatial structure of catchment-scale rainfall at scales ranging from 80 to 1000 km^2.

An index of spatial rainfall variability (SDIR) was defined, based on the difference between reference rainfall (defined by the full raingauge network) and alternative rainfall representations (using 1 raingauge per subcatchment, 1 raingauge for the whole catchment, or radar), and this proved to be the main controlling factor in determining the sensitivity of catchment response to spatial rainfall description. For 'natural catchments', at subcatchment scale, spatial deviations in rainfall of more than 10% led to variations in peak flow of 20% or more and variations in runoff volume of 10% or more, with marked further deterioration in runoff performance with spatial deviations in excess of 20%. Defining a Nash-Sutcliffe Efficiency for simulated runoff with respect to reference simulated runoff, spatial deviations above 20% lead to NSE values of less than 0.8. Using radar data, results for NSE were generally robust (usually above 0.9), although results for the largest summer event considered showed high spatial deviation (30%) between radar and reference rainfall. Using 1 raingauge per subcatchment achieved good performance at the whole catchment scale, but variable performance at subcatchment scale - generally close to the radar performance, but with some larger errors, particularly for some events on the Stort. Results using a single raingauge for the catchment showed significant degradation in runoff performance.

Results were analysed for catchment properties. No clear patterns emerged as a function of catchment area or catchment lag time, but the catchments themselves showed no direct relationship between lag time and scale. However, at the whole catchment scale (1000 km^2), effects of spatial rainfall description on runoff were generally less than for individual subcatchments (80 - 280 km^2), and the clay or part-urban catchments were more sensitive to spatial rainfall than the chalk

subcatchments. When the subcatchments were artificially turned into 'urban' catchments, response times decreased, and sensitivity to rainfall variability increased for the most spatially-variable rainfall events.

Clearly these results demonstrate that spatial variability of rainfall is important, but that quantification depends in a complex manner on catchment properties, rainfall type and scale. For largely rural catchments, a network of 16 raingauges seems appropriate at 1000 km^2 scale; between 4 and 7 gauges are required at 80 - 280 km^2, but this recommendation may not be appropriate for summer convective events. As urbanisation increases, the requirements become greater.

The analysis of the approach to disaggregation of daily rainfall data was instructive. We recall that the spatial pattern of daily rainfall is preserved, but a single sub-daily temporal pattern is applied uniformly in space. For the natural catchments, a close fit (NSE close to 1) is obtained, suggesting that the damping effect of the catchments compensates for the lack of rainfall temporal variability. However, for the artificially 'urbanised' catchments, the peak flow is somewhat reduced (by an amount comparable to using the radar data) and larger errors in the timing of the peak are observed. This confirms that urban catchments require a fine spatial and temporal resolution to adequately reproduce the streamflow.

Finally in Part IV, we include a study undertaken by CEH-Wallingford in which the GLIMCLIM GLM software was used to generate daily sequences of rainfall, and applied to rainfall-runoff modelling at the daily time-scale. Although the main focus of FD2105 is on hourly simulation, the report provides some insights into the use of GLM software by a third party and the associated simulation performance.

In terms of reproduction of rainfall characteristics it was found for the model specified that the temporal pattern of the generated rainfall (specifically monthly means) is good for the upper catchments where the rainfall is heavier. Rainfall for the lower catchment is overestimated in all months. Extreme rainfalls, in terms of 1-day and 5-day catchment averages, are well simulated in the winter but underestimated in all other seasons. No check was made of rainfall extremes at single sites (grid squares). It is noted that the GLM was applied to spatially averaged data (over a grid), with irregular number and location of point data within each grid. The GLM was designed for point data, and hence the reproduction of observed statistics can in some part be attributed to the spatial averaging process.

As a note on use of the software, whilst it was found that specification of the model was easy, it was found that constructively interpreting the output from GLIMCLIM while fitting the occurrence and amounts models was less so.

Regarding the flood frequency curves generated, it was found that the flood frequency determined from flows simulated from the GLM series compares well with that simulated from observed rainfall, as flood events are predominantly caused by heavy rainfall during the winter over the upper catchments. Flows, apart from those during flood events, are too high as for four of the regions the rainfall is overestimated. If rainfall from the GLM were to be used for further study, the occurrence and amounts should be revised to try and improve, at least, the spatial distribution of the rainfall (reduce the mean site residuals).

14.5 Recommendations

This project represents an extensive evaluation of rainfall and evaporation simulation methods to support continuous simulation rainfall-runoff modelling. Here we provide guidance to users concerning the applicability of the methods and recommendations for further work.

14.5.1 Single site rainfall and evaporation modelling

After the comparative testing of alternative single site rainfall models, the six parameter Random-Parameter Bartlett-Lewis model is preferred. A two stage fitting process is recommended, in which near-equivalent parameter sets, with respect to the fitting of statistical moments of the data, are identified and used to simulate test rainfall sequences. The preferred parameter set is selected on the basis of simulated extreme value performance. A minimum of 15 years of horuly data is recommended for model fitting. Some trade-offs occur between different aspects of performance, for example the daily and hourly statistics. The user has the opportunity to weight the optimisation criteria to reflect the needs of particular applications.

Performance for individual raingauges is variable - some results are good, some less so. Extreme value performance is often difficult to assess given the limited lengths of available sub-daily rainfall data for some sites and the associated large confidence limits. However, the model tends to underestimate hourly and overestimate daily extreme values, and performance for individual months was often better than for the annual maxima. A problem arose with the simulation of very long (1000 year) sequences, in that unrealistically-large events would very occasionally be simulated. This problem has largely been rectified by imposing bounds on parameters used in simulation, but users must inspect simulations to check for infeasible extremes.

The single site rainfall models were applied to represent catchment-average rainfall and then as input to the simulation of river flows, using the model of FD2106. Some theoretical issues were identified, in that catchment-average sequences were often based on individual gauge sub-daily time-series, giving rise to statistical discrepancies between the moments estimation for daily and hourly data. As with the evaluation for individual raingauges, performances were mixed. Some catchments gave very encouraging results, others less so. One generic problem was that the flow duration curves tended to be underestimated.

We conclude that the models are appropriate to use with caution, but that further testing on a wider range of catchments is necessary before general application to ungauged catchments can be recommended. Any user should validate the model using local data to have sufficient confidence for practical application.

More work is recommended to pursue the issue of parameter constraints and non-feasible simulations. There are also ideas for model extension that should be explored to seek improvements in performance.

The project has also developed a new set of modelling approaches for potential evaporation. These include:

- modelling weekly potential evaporation time series given rainfall;
- modelling daily potential evaporation time series given rainfall; and,
- downscaling weekly potential evaporation to daily values.

The models have had limited testing, and as yet have not been used with rainfall-runoff models. Nevertheless, the methods appear robust and defensible, and in our opinion can be applied in practice with a reasonable level of confidence.

14.5.2 Modelling spatial-temporal rainfall in continuous space and time

Based on our review, we conclude that models for continuous simulation of spatial-temporal rainfall should be based on the Poisson point process approach for medium-term application; fractal models are not yet at the stage where they are readily applicable.

The modelling of spatial-temporal rainfall in continuous space and time requires radar data to define the full structure of the spatial rainfall fields. This presents a number of problems; long records of radar data are not yet available, and calibration issues are complex. Radar data need careful quality assessment before use. Calibration procedures have significant effects effect on the statistical properties of the calibrated fields, and hence on the properties used in the simulation models. A substantial further study is needed to gain a thorough understanding of the issues involved relating to radar calibration and the use of calibrated data in fitting spatial-temporal models.

The project has explored the performance of the Poisson process modelling approach for continuous space-time modelling more extensively than has been possible hitherto. Performance is encouraging. The model for event interiors can be used to represent short-term fine scale spatial and temporal rainfall variability of typical rain events. Continuous rainfall space-time fields can be generated by generating storm arrivals and sampling model parameters on a monthly basis from a library of storm parameters that has been derived from extensive analysis of a 13 year radar sequence. A new methodology has been developed to represent spatial non-stationarity and appears promising.

A number of issues remain to be explored or resolved before practical application could be recommended. For example, further model development is needed to enable the representation of light and spatially and/or temporally intermittent rainfall, and in future work, and especially for hydrological applications, it will be important to look at the tail of the intensity distribution and properties of extremes, for both individual rain events and the continuous simulation. Almost certainly, further model development will be needed to enable adequate representation of particularly intense events. The rescaling method whereby nonstationary continuous simulation is achieved needs substantial further development to allow the wet-dry structure of rainfall as well as its intensity to vary according to different climate change scenarios. In addition, further work will enable physically-based climate change scenarios to be incorporated into rainfall simulations.

An interesting observation was of an apparent increasing frequency of rainfall event occurrence over the period of the available radar record. Further assessment of radar data is needed to determine the causes.

14.5.3 Spatial-temporal modelling based on daily rainfall modelling and spatial-temporal disaggregation

Realistic multisite daily rainfall sequences can be generated using Generalised Linear Models (GLMs) fitted to daily raingauge data. A simple disaggregation scheme can be applied if subdaily data are required.

Daily rainfall records should be scrutinised carefully before use in any modelling exercise. Common problems include the presence of unflagged monthly totals (which, if undetected, will seriously bias any analysis of extremes) and undocumented changes in recording resolution (which can produce spurious trends in rainfall sequences). We also recommend routine thresholding of daily rainfall observations by setting to zero any values below 0.5mm, to allow for differences in observer practice in the recording of small rainfall amounts.

The GLM framework enables systematic structures to be incorporated into simulated rainfall sequences. In addition corrections can, and should, be made for features such as changes in recording resolution and measurement device.

A number of contrasting case studies have yielded GLMs for daily rainfall with a broadly similar structure. This basic structure is therefore expected to apply quite generally across the UK.

GLMs are designed to be fitted to point raingauge data. If catchment averages or gridded data are required, models should be fitted to observed raingauge data and then simulated on the required grid. For calibration of rainfall-runoff models, historical gridded values can be obtained by conditioning on all available observations. This avoids artefacts associated with the common procedure of gridding the observations prior to analysis.

As developed, the GLMs are designed primarily for use in small to medium sized catchments. At scales above 2000km^2, the methodology may still produce reasonable results; however, users should check carefully to ensure that the results of any simulation exercise appear reasonable. In this and other respects, warning messages from the simulation software should be taken seriously. Further research is required to extend the range of applicability of the GLMs to larger scales.

Some testing has been carried out to assess the suitability of the combined GLM-disaggregation scheme for use in rainfall-runoff modelling applications. Results from a single case study indicate that the scheme has a lot of potential. At present, however, a limitation is that the disaggregation scheme overestimates spatial correlation in rainfall - this, in turn, leads to some overestimation of peak flows. Further research is required to investigate this, to develop the disaggregation methodology further and to test the performance in a wider range of catchments.

The GLM software has been extensively tested by a number of people and found to be generally easy to use. However, it would be convenient to add a graphical user

interface which would speed up the process of model fitting, testing and simulation.

All of the methods developed here require a reasonable level of user awareness. The methods have the potential to deliver much more than traditional event-based techniques, but with this comes a responsibility on the part of the user to evaluate the results carefully and critically in any application.

14.5.4 The significance of spatial rainfall for runoff estimation

The literature on the significance of spatial rainfall for runoff estimation is complex and sometimes contradictory. Effects can be expected to vary depending on the nature of the rainfall, the nature of the catchment, and the spatial scale of the catchment and rainfall. It is helpful to note that as the scale increases, the importance of spatial rainfall decreases and there is a transfer from spatial variability of rainfall to catchment response time distribution as the dominant factor governing runoff generation.

Runoff from urban areas is extremely sensitive to spatial rainfall, and the same is true for arid areas, where rapid flow response is generated from spatially-localised convective rainfall. This is problematic, since the required density of raingauges to capture the spatial variability exceeds that normally available from routine monitoring networks.

Numerical experiments for the Lee catchment (1000 km^2 to Feildes Weir) showed a complex picture. The subcatchments varied greatly in geology and runoff response. Chalk catchments were less sensitive to spatial rainfall description than clay catchments or those with significant urban development. As a result, no clear pattern emerged as a function of catchment scale, or response time, except that the effect of spatial variability was damped at the whole catchment scale of 1000 km^2. The dominant effect was the variability of the rainfall; as this increased, so did the significance of appropriate rainfall characterisation. When the catchments were numerically urbanised, the significance of spatial rainfall was enhanced. It is concluded that, for largely rural catchments, a network of 16 raingauges seems appropriate at 1000 km^2 scale; between 4 and 7 gauges are required at 80 - 280 km^2, but this recommendation may not be appropriate for summer convective events. As urbanisation increases, the requirements become greater.

A notable result was the relative success of the spatial-temporal disaggregation scheme, when used in conjunction with the rainfall-runoff model. The use of this strategy to disaggregate daily data worked impressively well for the natural catchments, and the procedure can be recommended with some confidence. However, as might be expected, some deterioration in performance occurred when the catchments were artificially urbanised.

The final Chapter of this section reported the application, at large catchment scale, of the GLM models for daily rainfall to continuous simulation rainfall-runoff modelling using a distributed hydrological model. The GLM results were based on gridded data, which may have affected performance, and showed strengths and weaknesses. They were particularly good for winter rainfall and the areas of the catchment where rainfall was highest, and as a result, overall performance in flood frequency estimation at catchment scale was impressive. However, lower flows were generally over-estimated.

14.5.5 Concluding comments

This last result, as with the results using the single site rainfall model with rainfall-runoff modelling, can be regarded as a first iteration in the joint assessment of rainfall and rainfall-runoff models. An obvious next step would be to revisit the rainfall models in the light of the joint performance assessment and it seems likely that further improvement in performance could be achieved.

The rainfall models presented here have strengths and weaknesses in aspects of their performance, and the objective of continuous simulation modelling necessitates a trade-off between different aspects of performance. It is a major challenge to seek to represent such a wide range of properties of rainfall and flow, including the extremes of flood frequency and low flow durations, using coupled modelling of the rainfall and runoff processes. Nevertheless the methods presented here have clear potential for success, and we have attempted to indicate to the user where the models are in our view ready for immediate application.

Clearly more extensive testing is needed. We encourage the user community to explore the use of these methods for application in practice, and would welcome feedback on their use and performance.

Bibliography

Abramowitz, M. and Stegun, I. A. (1965). *Handbook of mathematical functions: with formulas, graphs and mathematical tables*. Dover, New York.

Ajami, N., Gupta, H., Wagener, T., and Sorooshian, S. (2004). Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system. *Journal of Hydrology*, 298:112–135.

Andréassian, V., Perrin, C., Michel, C., Usart-Sanchez, I., and Lavabre, J. (2001). Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. *Journal of Hydrology*, 250:206–223.

Arnaud, P., Bouvier, C., Cisneros, L., and Dominguez, R. (2002). Influence of rainfall spatial variability on flood prediction. *Journal of Hydrology*, 260:216–230.

Aronica, G. and Cannarozzo, M. (2000). Studying the hydrological response of urban catchments using a semi-distributed linear non-linear model. *Journal of Hydrology*, 238:35–43.

Austin, G. (2001). Weather radar: theory and practice. In Cluckie, I. and Griffith, R., editors, *Radar hydrology for real time flood forecasting*, pages 33–38. European Commission. ISBN: 92-894-1640-8.

Austin, P. H. and Houze, R. A. (1972). Analysis of the structure of precipitation patterns in New England. *J. Appl. Met*, 11:926–935.

Barnston, A. G. and Livezey, R. E. (1987). Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Monthly Weather Review*, 115:1083–1126.

Bell, V. and Moore, R. (2000). The sensitivity of catchment runoff models to rainfall data at different spatial scales. *Hydrology and Earth System Sciences*, 4(4):653–667.

Bellone, E. (2003). Calibration of Chenies Data: Problems and changes, DEFRA project, Internal report no. 10. Technical report, Imperial College London and University College London.

Berne, A., Delrieu, G., Creutin, J.-D., and Obled, C. (2004). Temporal and spatial resolution of rainfall measurements required for urban hydrology. *Journal of Hydrology*, 299:1–14.

Beven, K. and Binley, A. (1992). The future of distributed models: model calibration and predictive uncertainty. *Hydrol. Process.*, 6:279–298.

Borga, M. (2002). Accuracy of radar rainfall estimates for streamflow simulation. *Journal of Hydrology*, 267:26–39.

Borga, M., Tonelli, F., Moore, R., and Andrieu, H. (2002). Long-term assessmen of bias adjustment in radar rainfall estimation. *Water Resources Research*, 38(11):1–10.

Bowman, A. W. and Azzalini, A. (1997). *Applied smoothing techniques for data analysis — the kernel approach with S-Plus illustrations*, volume 18 of *Oxford Statistical Science series*. Oxford University Press, Oxford.

Burne, S., Wheater, H. S., Butler, A. P., Johnston, P. M., Wadey, P., Shaw, G., and Bell, J. N. B. (1994). Radionuclide transport above a near-surface water table I. An automated lysimeter facility for near-surface contaminant transport studies. *J. Env. Quality*, 23, no. 6:1318–1329.

Calver, A., Lamb, R., Kay, A. L., and Crewet, J. (2001). The continuous simulation method for river flood frequency estimation. Technical report, Defra project FD0404 Final Report,CEH Wallingford.

Calver, A. and Wheater, H. (2002). Scoping the broad scale modelling hydrology programme: Stage 2 strategic programme. Technical report, FD2104 report to DEFRA and the Environment Agency.

Calvet, L., Fisher, A., and Mandelbrot, B. (1997). Large deviations and the distribution of price changes. Cowles Foundation Discussion Paper No. 1165.

Cameron, D., Beven, K., and Tawn, J. (2000). An evaluation of three stochastic models. *J. Hydrol.*, 228:130–149.

Cameron, D., Beven, K., and Tawn, J. (2001). Modelling extreme rainfalls using a modified random pulse bartlett-lewis stochastic rainfall model (with uncertainty). *Advances in Water Resources*, 24:203–211.

Carpenter, T., Konstantine, P., and Sperfslagea, J. (2001). On the parametric and nexrad-radar sensitivities of a distributed hydrologic model suitable for operational use. *Journal of Hydrology*, 253:169–193.

Chandler, R. (2003). Moment-based inference for stochastic-mechanistic models. Technical report, no. 7, DEFRA project *Improved Methods for National Spatial-Temporal Rainfall and Evaporation Modelling for BSM*.

Chandler, R. E. (1997). A spectral method for estimating parameters in rainfall models. *Bernoulli*, 3, No.3:301–322.

Chandler, R. E. (2002). GLIMCLIM: Generalized linear modelling for daily climate time series (software and user guide). Technical report, no. 227, Department of Statistical Science, University College London, London WC1E 6BT. `http://www.ucl.ac.uk/Stats/research/Resrpts/abstracts.html`.

Chandler, R. E. (2005). On the use of generalized linear models for interpreting climate variability. *Environmetrics*, 16, no. 7:699–715.

Chandler, R. E. and Wheater, H. S. (1998). Climate change detection using Generalized Linear Models for rainfall — a case study from the West of Ireland. II. Modelling of rainfall amounts on wet days. Technical report, no. 195, Department of Statistical Science, University College London. `http://www.ucl.ac.uk/Stats/research/Resrpts/abstracts.html`.

Chandler, R. E. and Wheater, H. S. (2002). Analysis of rainfall variability using Generalized Linear Models — a case study from the West of Ireland. *Water Resources Research*, 38, No.10:doi:10.1029/2001WR000906.

Chaouche, K. (2001). *Approche multifractale de la modelisation stochastique en hydrologie*. PhD thesis, Ecole Nationale du Genie Rural et de Eaux et Forets, Paris.

Chaubey, I., Hann, C., Grunwald, S., and Salisbury, J. (1999). Uncertainty in the model parameters due to spatial variability of rainfall. *Journal of Hydrology*, 220:48–61.

Chow, V., Maidment, D., and Mays, L. (1988). *Applied Hydrology*. McGraw-Hill.

Chui, C. K. (1992). *An Introduction to Wavelets*, volume 1 of *Wavelet Analysis and its Applications*. Academic Press.

Cluckie, I., Tilford, K., Griffith, R., and Lane, A. (2000). Radar hydrometeorology using a vertically pointing radar. *Hydrology and Earth System Sciences*, 4:565–580.

Coe, R. and Stern, R. D. (1982). Fitting models to daily rainfall. *J. Appl. Meteorol.*, 21:1024–1031.

Collier, C. (2000). Developments in radar and remote-sensing methods for measuring and forecasting rainfall. *Proc. R. Soc. Lond.*, A360:1345–1361.

Corral, C., Sempere-Torres, D., Revilla, M., and Berenguer, M. (2000). A semi-distributed hydrological model using rainfall estimates by radar. application to mediterranean basins. *Physics and Chemistry of the Earth*, 25(10–12):1133–1136.

Cowpertwait, P. (1998). A Poisson-cluster model of rainfall: higher-order moments and extreme values. *Proc. R. Soc. Lond.*, A454:885–898.

Cowpertwait, P. S. P. (1991). Further developments of the Neyman-Scott clustered point process for modeling rainfall. *Water Resources Research*, 27:1431–1438.

Cowpertwait, P. S. P. (1994). A generalised point process model for rainfall. *Proc. R. Soc. Lond.*, 447:23–37.

Cox, D. and Isham, V. (1980). *Point Processes*. Chapman and Hall, London.

Cox, D. R. and Hinkley, D. (1974). *Theoretical Statistics*. Chapman & Hall, London.

Cox, D. R. and Isham, V. (1988). A simple spatial-temporal model of rainfall. *Proc. R. Soc. Lond.*, A415:317–328.

Cressie, N. (1991). *Statistics for spatial data*. Wiley, New York.

Crooks, S. (2002). Catchment flood management system: extensions of eurotas hydrological modelling. Technical report, Defra FD0421 Final Report.

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

Davison, A. C. (2003). *Statistical Models*. Cambridge University Press, Cambridge.

de Lima, J. and Singh, V. (2002). The influence of the pattern of moving rainstorms on overland flow. *Advances in Water Resources*, 25:817–828.

Deidda, R. (1999). Multifractal analysis and simulation of rainfall fields in space. *Physics and Chemistry of the Earth*, 24:73–78.

Deidda, R. (2000). Rainfall downscaling in a space-time multifractal framework. *Water Resources Research*, 36:1779–1794.

Dobson, A. J. (2001). *An Introduction to Generalized Linear Models (second edition)*. Chapman and Hall, London.

Dodov, B. and Foufoula-Georgiou, E. (2005). Incorporating the spatio-temporal distribution of rainfall and basin geomorphology inot nonlinear analyses of streamflow dynamics. *Advances in Water Resources*, 28:711–728.

Duan, Q., Gupta, V., and Sorroshian, S. (1993). Shuffled complex evolution approach for effective and efficient global minimization. *J. of Optimization Theory and Applications*, 76(3):501–521.

Duan, Q., Sorooshian, S., and Gupta, H. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, 28(4):265–284.

EAE (2005). Encyclopedia of the Atmospheric Environment (Online). *http://www.ace.mmu.ac.uk/eae/english.html*.

Entekhabi, D., Rodriguez-Iturbe, I., and Eagleson, P. S. (1989). Probabilistic representation of the temporal rainfall process by a modified Neyman-Scott rectangular pulses model: parameter estimation and validation. *Water Resources Research*, 25, No.2:295–302.

Falconer, K. (1990). *Fractal Geometry*. John Wiley & Sons.

Faulkner, D. (1999). *Flood Estimation Handbook, 2: Rainfall frequency estimation*, volume 3. Institute of Hydrology, United Kingdom.

Faurès, J., Goodrish, D., Woolhiser, D., and Sorooshian, S. (1995). Impact of small-scale spatial rainfall variability on runoff modeling. *Journal of Hydrology*, 173:309–326.

Feller, W. (1996). *An introduction to probability theory and its applications*. John Wiley & Sons.

Flynn & Rothwell (1991a). Report on Hydraulic Modelling (ONDA) of The River Stort. Technical report, Thames Region, National Rivers Authority.

Flynn & Rothwell (1991b). Report on Hydrologic Modelling (RORB) of Middle Lee, draft report. Technical report, Thames Region, National Rivers Authority.

Freer, J., Beven, K., and Abroise, B. (1996). Bayesian uncertainty in runoff prediction and the value of data: an application of the GLUE approach. *Water Resources Research*, 32:2163–2173.

Gannon, B. (1995). Automating areal rainfall calculations for catchments. Technical report, Institute of Hydrology, Wallingford.

Gershenfeld, N. A. (1999). *The nature of mathematical modelling*. Cambridge University Press, Cambridge.

Gupta, V. K. and Waymire, E. (1990). Multiscaling properties of spatial rainfall and river flow distributions. *Journal of Geophysical Research*, 95:1999–2009.

Gupta, V. K. and Waymire, E. C. (1993). A statistical analysis of mesoscale rainfall as a random cascade. *Journal of Applied Meteorology*, 32:251–267.

Gupta, V. K. and Waymire, E. C. (1997). Universal multifractals do exist!: Reply. *Journal of Applied Meteorology*, 36:1304–1304.

Halsey, T. C., Jensen, M. H., Kadanoff, L. P., Procaccia, I., and Shraiman, B. (1986). Fractal measures and their singularities: The characterization of strange sets. *Phys. Rev. A*, 33:1141–1151.

Harte, D. (2001). *Multifractals: Theory and Applications*. Chapman and Hall/CRC.

Hawnt, R. (1987a). Catchment runoff control study, Report on the evaluation of the hydrological response of the catchments of the Rib and Ash. Technical report, Thames Water, Rivers Division.

Hawnt, R. (1987b). Catchment runoff control study, Report on the evaluation of the hydrological response of the Mimram and Beane catchments. Technical report, Thames Water, Rivers Division.

Holley, R. and Waymire, E. C. (1992). Multifractal dimensions and scaling exponents for strongly bounded random cascades. *Annals of Applied Probability*, 2:819–845.

Horn, R. A. and Johnson, C. R. (1985). *Matrix Analysis*. Cambridge University Press, Cambridge.

Hulme, M. and Barrow, E. M., editors (1997). *Climate of the British Isles - Past, present and future.* Routledge, London.

Hurrell, J. W. (1995). Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation. *Science*, 269:676–9.

Hurvich, C. M. and Beltrao, K. (1993). Asymptotics for the low-frequency ordinates of the periodogram of a long-memory time series. *Journal of Time Series Analysis*, 14:455–472.

Hwang, Y., Clark, M., Rajagopala, B., Gangopadhyay, S., and Hay, L. (2005). Spatial interpolation schemes for daily precipitation. *Journal of Hydrometeorology*, Submitted.

Illingworth, A., Blackman, T., and Goddard, J. (2000). Improved rainfall estimates in convective storms using polarisation diversity radar. *Hydrology and Earth System Sciences*, 4:555–563.

IPCC (2001). *Climate Change 2001 — the scientific basis*. Cambridge University Press, Cambridge. Report of the Intergovernmental Panel on Climate Change.

Jensen, N. and Pedersen, L. (2005). Spatial variability of rainfall: variations within a single radar pixel. *Atmospheric Research*, In press.

Jolley, T. J. and Wheater, H. S. (1996). A large-scale grid-based hydrological model of the Severn and Thames ctachments. *J. CIWEM*, 10:253–262.

Jones, B. (1983). The estimation of catchment average point rainfall profiles. Technical report, Institute of Hydrology, Wallingford.

Kakou, A. (1997). *Point process based models of rainfall*. PhD thesis, Department of Statistical Science, University College London.

Kakou, A. (1998). A point process based model for rainfall with dependent duration and intensity. Unpublished manuscript.

Kedem, B. and Chiu, L. S. (1987). Are rainrate processes self-similar? *Water Resources Research*, 23:1816–1818.

Kendall, M. and Ord, J. (1990). *Time Series (third edition)*. Edward Arnold.

Kilsby, C., Jones, P. D., Burton, A., Ford, A. C., Fowler, H. J., Harpham, J. P., Smith, A., and Wilby, R. L. (2006). A daily weather generator for use in climate change studies. *Environmental Modelling & Software*, Submitted.

Kilsby, C., Moaven-Hashemi, A., and O'Connell, P. (2004). Simulation of rainfall extremes: fitting to observed annual maxima. *First International Conference on Flood Risk*, Institute of Mathematics and its Applications, 7-8 september, University of Bath, UK.

Kitchen, M., Brown, R., and Davies, A. (1994). Real-time correction of weather radar data for the effects of bright band, range and orographic growth in widespread precipitation. *Quaterly Journal Royal Meteorological Society*, 120:1231–1254.

Koren, V., Finnerty, B., Schaake, J., Smith, M., Seo, D.-J., and Duan, Q.-Y. (1999). Scale dependencies of hydrologic models to spatial variability of precipitation. *Journal of Hydrology*, 217:285–302.

Koutsoyiannis, D. and Onof, C. (2000). HYETOS - A Computer program for stochastic disaggregation of fine-scale rainfall. *http://www.itia.ntua.gr/e/softinfo/3/*.

Koutsoyiannis, D. and Onof, C. (2001). Rainfall disaggregation using adjusting procedures on a Poisson cluster model. *Journal of Hydrology*, 246:109–122.

Krzanowski, W. (1988). *Principles of Multivariate Analysis*. Oxford University Press.

Kumar, P. and Foufoula-Georgiou, E. (1993). A multicomponent decomposition of spatial rainfall fields 1. segregation of large- and small-scale features using wavelet transforms. *Water Resources Research*, 29:2515–2532.

Kuonen, D. (1999). Saddlepoint approximations for distributions of quadratic forms in normal variables. *Biometrika*, 86(4):929–935.

Lamb, R. and Kay, A. L. (2004). Confidence intervals for a spatially generalized, continuous simulation flood frequency model for great britain. *Water Resources Research*, 40:doi:10.1029/2003WR002428.

Laurenson, E. and Mein, R. (1988). RORB version 4 Runoff Routing program, User Manual. Technical report, Monash University, Department of Civil Engineering.

Lavallée, D. (1991). *Multifractal techniques: Analysis and simulation of Turbulent fields*. PhD thesis, MCGill University.

Lavallée, D., Schertzer, D., and Lovejoy, S. (1990). On the determination of the codimension function. In Schertzer, S. and Lovejoy, S., editors, *Scaling, Fractals and Non-Linear Variablility in Geophysics*. Kluwer.

Lekkas, D. (2002). Comparison of stochastic models for single site rainfall. Technical report, no. 1, DEFRA project *Improved Methods for National Spatial-Temporal Rainfall and Evaporation Modelling for BSM*.

Lekkas, D. and Onof, C. (2003). The calibration and validation of single-site models. Technical report, no. 3, DEFRA project *Improved Methods for National Spatial-Temporal Rainfall and Evaporation Modelling for BSM*.

Lekkas, D. and Onof, C. (2005). Introducing the variation of advective time delay (adt) to transfer function models. *Environmental Modelling & Software*, In press.

Liang, K.-Y. and Zeger, S. (1995). Inference based on estimating functions in the presence of nuisance parameters. *Statistical Science*, 10:158–173.

Lopes, V. (1996). On the effect of uncertainty in spatial distribution of rainfall on catchment modelling. *CATENA*, 28:107–119.

Lotufo Conejo, J. (1979). Hydrologic simulation and flood routing models: a case study. Technical report, Imperial College.

Lovejoy, S. and Schertzer, D. (1990). Multifractals, universality classes and satellite and radar measurement of cloud and rain fields. *Journal of Geophysical Research*, 95:2021–2034.

Lovejoy, S., Schertzer, D., and Tsonis, A. A. (1987). Functional box-counting and multiple elliptical dimensions in rain. *Science*, 235:1036–1038.

Lovejoy, S. and Schertzer, S. (1991). Multifractal analysis techniques and the rain and cloud fields from 10^{-3} to 10^{6} m. In Schertzer, D. and Lovejoy, S., editors, *Non-Linear Variability in Geophysics*, pages 111–144. Kluwer Acad.

Mandelbrot, B. B. (1975). *Les objets fractals: forme, hasard et dimension.* Flammarion.

Marsan, D. and Schertzer, D. (1996). Causal space-time multifractal processes: predictability and forecasting of rain fields. *Journal of Geophysical Research*, 101:26333–26346.

Maskey, S., Guinot, V., and Price, R. (2004). Treatment of precipitation uncertainty in rainfall-runoff modelling: a fuzzy set approach. *Advances in Water Resources*, 27:889–898.

McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models (second edition)*. Chapman and Hall, London.

McIntyre, N., Lee, H., Wheater, H., Young, A., and Wagener, T. (2006). Ensemble predictions of runoff in ungaged catchments. *Water Resources Research*, In press.

Michaud, J. and Sorooshian, S. (1994). Effect of rainfall-sampling errors on simulations of desert flash floods. *Water Resources Research*, 30(10):2765–2775.

Monteith, J. L. (1965). Evaporation and the environment. *Symp. Soc. Expl. Biol.*, 19:205–234.

Montgomery, D. (1997). *Design and analysis of experiments (fourth edition)*. Wiley.

Moore, R. J. (1985). The probability-distributed principle and runoff production at point and basin scales. *Water Resources Research*, 30(2):273–297.

Moore, R. J., May, B. C., Jones, D. A., and Black, K. B. (1994). Local calibration of weather radar over london. In *Advances in Radar Hydrology*, pages 186–195. European Commission.

Morin, E., Enzel, Y., Shamir, U., and Garti, R. (2001). The characteristic time scale for basin hydrological response using radar data. *Journal of Hydrology*, 252:85–99.

Mouhous, N. (2003). *Interet des modeles de cascades multiplicatives pour la simulation de series chronologiques de pluies ponctuelles adaptees a l'hydrologie urbaine*. PhD thesis, Ecole Nationale des Ponts et Chaussees, Paris.

Naden, P. (1992). Spatial variability in flood estimation for large catchments: the exploitation of channel network structure. *Journal des Sciences Hydrologiques*, 37(1):53–71.

NERC (1975). Flood studies report. Report of the Natural Environment Research Council (5 volumes).

NERC (1975). *Flood Studies Report, Hydrological Studies*, volume 1. Natural Environment Research Council, United Kingdom.

NERC (1999). WINFAP-FEH CD-ROM version 1.0.

Ngirane-Katashaya, G. and Wheater, H. (1985). Hydrograph sensitivity to storm kinematics. *Water Resources Research*, 21(3):337–345.

Northrop, P. (2004). Estimating the parameters of rainfall models using maximum marginal likelihood. *Student*, submitted.

Northrop, P. J. (1996). *Modelling and statistical analysis of spatial-temporal rainfall fields*. PhD thesis, Department of Statistical Science, University College London.

Northrop, P. J. (1998). A clustered spatial-temporal model of rainfall. *Proc. Roy. Soc. London.*, A454:1875–1888.

Northrop, P. J. and Stone, T. M. (2005). A point process model for rainfall with truncated Gaussian rain cells. *Student*. In press.

Obled, C., Wendling, J., and Beven, K. (1994). The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data. *Journal of Hydrology*, 159:305–333.

Ogden, F. and Julien, P. (1993). Runoff sensitivity to temporal and spatial rainfall variability at runoff plane and small basin scales. *Water Resources Research*, 29(8):2589–2597.

Ogden, F. and Julien, P. (1994). Runoff model sensitivity to radar rainfall resolution. *Journal of Hydrology*, 158:1–18.

Onof, C. (1992). Stochastic modelling of British rainfall data using Poisson processes. PhD thesis, University of London.

Onof, C. (2003). Mathematical expressions for generalized moments in single-site models. Technical report, no. 8, DEFRA project *Improved Methods for National Spatial-Temporal Rainfall and Evaporation Modelling for BSM*.

Onof, C., Chandler, R. E., Kakou, A., Northrop, P., Wheater, H. S., and Isham, V. (2000). Rainfall modelling using Poisson-cluster processes: a review of developments. *Stoch. Env. Res. & Risk Ass.*, 14:384–411.

Onof, C., Townend, J., and Kee, R. (2005). Comparison of two hourly to 5-min rainfall disaggregators. *Atmospheric Research*, 77:176–187.

Onof, C. and Wheater, H. S. (1993). Modelling of British rainfall using a random parameter Bartlett-Lewis rectangular pulse model. *J. Hydrol.*, 149:67–95.

Onof, C. and Wheater, H. S. (1994). Improvements to the modelling of british rainfall using a modified random parameter bartlett-lewis rectangular pulses model. *J. Hydrol.*, 157:177–195.

Onof, C., Wheater, H. S., and Isham, V. (1994). Note on the analytical expression of the inter-event time characteristics for Bartlett-Lewis type rainfall models. *J. Hydrol.*, 157:197–210.

Over, T. (1995). *Modelling space-time rainfall at the mesoscale using random cascades*. PhD thesis, University of Colorado.

Over, T. M. and Gupta, V. K. (1996). A space-time theory of mesoscale rainfall using random cascades. *Journal of Geophysical Research*, 101:26319–26331.

Over, Y. M. and Gupta, V. K. (1994). Statistical analysis of mesoscale rainfall: Dependence of a random cascade generator on large-scale forcing. *Journal of Applied Meteorology*, 33:1526–1542.

Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass. *Proc. Roy. Soc. London*, A193:120–145.

Percival, D. B. and Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge University Press, Cambridge, England.

Perica, S. and Foufoula-Georgiou, E. (1996). Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and sclaing descriptions. *Journal of Geophysical Research*, 101:26347–26361.

Porporato, A. and Ridolfi, L. (2001). Multivariate nonlinear prediction of river flows. *Journal of Hydrology*, 248:109–122.

Priestley, M. (1981). *Spectral Analysis and Time Series*. Academic Press.

Prudhomme, C. (2005). Joint testing of single-site rainfall simulation with continuous rainfall-runoff modelling. Technical report, no. 15, DEFRA project *Improved Methods for National Spatial-Temporal Rainfall and Evaporation Modelling for BSM*, CEH Wallingford.

Quasem, W. M. (2004). Effect of spatial variability of rainfall on runoff for the lee catchment. Master's thesis, Imperial College London.

R Development Core Team (2003). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3.

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., and Seo, D.-J. (2004). Overall distributed model intercomparison project results. *Journal of Hydrology*, 298:27–60.

Reefsgaard, J. and Knudsen, J. (1996). Operational validation and intercomparison of different types of hydrological models. *Water Resources Research*, 32(7):2189–2202.

Reynard, N., Crooks, S., and Kay, A. (2004). Impact of climate change on flood flows in river catchments. Technical report, Defra/EA project W5B-01-050 Final Report.

Rice, J. (1995). *Mathematical Statistics and Data Analysis (second edition)*. Duxbury Press.

Riedi, R. H. (1997). *An introduction to multifractals*. Rice University. `http://www.ece.rice.edu/~riedi/mfr_intro.html`.

Robson, A. and Reed, D. (1999). Statistical procedures for flood frequency estimation. flood estimation handbook vol. 3. Technical report, Institute of Hydrology, Wallingford.

Rodriguez-Iturbe, I., Cox, D. R., and Isham, V. (1987). Some models for rainfall based on stochastic point processes. *Proc. R. Soc. Lond.*, A410:269–288.

Rodriguez-Iturbe, I., Cox, D. R., and Isham, V. (1988). A point process model for rainfall: further developments. *Proc. R. Soc. Lond.*, A417:283–298.

Rodriguez-Iturbe, I. and Eagleson, P. (1987). Mathematical models of rainstorm events in space and time. *Water Resources Research*, 23(1):181–190.

Ryan, L. (1995). Comment on the article by Liang and Zeger. *Statistical Science*, 10, no.2:189–193.

Schertzer, D. (2003). Personal Communication.

Schertzer, D. and Lovejoy, S. (1987). Physicall modeling and analysis of rain and cloudsby anisotropicscaling multiplicative processes. *Journal of Geophysical Research*, 92:9693–9714.

Schertzer, D. and Lovejoy, S. (1997). Universal mulitfractals do exist!; comments on "a statistical analysis of mesoscale rainfall as a random cascade". *Journal of Applied Meteorology*, 36:1296–1303.

Shah, S., O'Connel, P., and Hosking, J. (1996a). Modelling the effects of spatial variability in rainfall on catchment response. 1. formulation and calibration of a stochastic rainfall field. *Journal of Hydrology*, 175:67–88.

Shah, S., O'Connel, P., and Hosking, J. (1996b). Modelling the effects of spatial variability in rainfall on catchment response. 2. experiments with distributed and lumped models. *Journal of Hydrology*, 175:89–111.

Shaw, E. M. (1983). *Hydrology in practice.* Van Nostrand Reinhold (UK) Co. Ltd, Wokingham.

Shuttleworth, W. J. (1993). Evaporation. In Maidment, D. R., editor, *Handbook of Hydrology*, pages 4.1–4.53. McGraw Hill.

Singh, V. (1997). Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. *Hydrological Processes*, 11:1649–1669.

Smith, M., Koren, V., Zhang, Z., Reed, S., Pan, J., and Moreda, F. (2004). Runoff response to spatial variability in precipitation: an analysis of observed data. *Journal of Hydrology*, 298:267–286.

Smith, M., Seo, D.-J., Koren, V., Reed, S., Zhang, Z., Duan, Q., Moreda, F., and Cong, S. (2003). The distributed model intercomparison project (dmip): motivation and experiment design. *Journal of Hydrology*, 298:4–26.

Stern, R. D. and Coe, R. (1984). A model fitting analysis of rainfall data (with discussion). *J. Roy. Stat. Soc.*, A147:1–34.

Sumbler, M. (1996). *British Regional Geology, London and the Thames Valley*. British Geological Survey, London, fourth edition.

Syed, K., Goodrich, D., Myers, D., and Sorooshian, S. (2003). Spatial characteristics of thunderstorm rainfall fields and their relation to runoff. *Journal of Hydrology*, 271:1–21.

Takens, F. (1981). Detecting strange attractors in turbulence. lecture notes in mathematics. In *Rand, D., Young, L.-S., (Eds.), Dynamical Systems and Turbulence, Springer Lecture Notes in Mathematics*, pages 366–381, Berlin.

Tessier, Y., Lovejoy, S., and Schertzer, D. (1993). Universal multifractals: theory and observations for rain and clouds. *Journal of Applied Meteorology*, 32:223–250.

Thompson, N., Barrie, I. A., and Ayles, M. (1981). *The Meteorological Office rainfall and evaporation calculation system: MORECS*. Hydrological memorandum, no. 45. UK Meterological Office, Bracknell.

Vanmarcke, E. (1993). *Random Fields: Analysis and Synthesis*. MIT Press, Cambridge, Mass., USA.

Velghe, T., Troch, P. A., de Troch, F. P., and Van de Velde, J. (1994). Evaluation of cluster-based rectangular pulses point process models for rainfall. *Water Resources Research*, 30, no. 10:2847–2857.

Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S (fourth edition)*. Springer-Verlag, New York.

Venugopal, V., Foufoula-Georgiou, E., and Sapozhnikov, V. (1999). A space-time downscaling model for rainfall. *Journal of Geophysical Research*, 104:19705–19721.

Verhoest, N., Troch, P. A., and de Troch, F. P. (1997). On the applicability of Bartlett-Lewis rectangular pulses models for calculating design storms at a point. *J. Hydrol.*, 202:108–120.

Wagener, T., Wheater, H., and Gupta, H. (2004). *Rainfall-runoff modelling in gauged and ungauged catchments.* Imperial College Press, London, UK.

Wallingford, C. (2005). National river catchment flood frequency method using continuous simulation. vol. 1: Main report. Technical report, Project FD2106 Final Report to Defra/Evironment Agency Flood and Coastal Management R&D Programme, CEH Wallingford.

Waymire, E. and Gupta, V. (1981). The mathematical structure of rainfall representations 1: a review of the stochastic rainfall models. *Water Resources Research*, 17, no.5:1261–1272.

Weisstein, E. (1992). Eric's Weisstein's world of mathematics. http://mathworld.wolfram.com.

Wheater, H. and Lotufo Conejo, J. (1979). River Lea flood routing study. Technical report, Imperial College.

Wheater, H. S., Chandler, R. E., Onof, C. J., Isham, V. S., Bellone, E., Yang, C., Lekkas, D., Lourmas, G., and Segond, M.-L. (2006). Spatial-temporal rainfall modelling for flood risk estimation. *Stoch. Env. Res. & Risk Ass.*, In press.

Wheater, H. S., Isham, V. S., Cox, D. R., Chandler, R. E., Kakou, A., Northrop, P. J., Oh, L., Onof, C., and Rodriguez-Iturbe, I. (2000a). Spatial-temporal rainfall fields: modelling and statistical aspects. *Hydrological and Earth Systems Science*, 4:581–601.

Wheater, H. S., Isham, V. S., Onof, C., Chandler, R. E., Northrop, P. J., Guiblin, P., Bate, S. M., Cox, D. R., and Koutsoyiannis, D. (2000b). Generation of spatially consistent rainfall data. Report to the Ministry of Agriculture, Fisheries and Food (2 volumes). Also available as Research Report no. 204, Department of Statistical

Science, University College London
(http://www.ucl.ac.uk/Stats/research/Resrpts/abstracts.html).

Wiggins, S. (1999). *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Springer, second edition.

Wilks, D. S. and Wilby, R. L. (1999). The weather generation game: a review of stochastic weather models. *Progress in Physical Geography*, 23:329–357.

Williams, D. (1982). Extra-binomial variation in logistic linear models. *Appl. Statist.*, 31, no.2:144–148.

Woods, R. and Sivapalan, M. (1999). A synthesis of space-time variability in storm response: rainfall, runoff generation and routing. *wrr*, 35(8):2469–2489.

Yan, Z., Bate, S., Chandler, R. E., Isham, V. S., and Wheater, H. S. (2002). An analysis of daily maximum windspeed in northwestern Europe using Generalized Linear Models. *J. Climate*, 15, No.15:2073–2088.

Yan, Z., Bate, S., Chandler, R. E., Isham, V. S., and Wheater, H. S. (2006). Changes in extreme wind speeds in NW Europe simulated by Generalized Linear Models. *Theoretical and Applied Climatology*, 83:121–137.

Yang, C., Chandler, R. E., Isham, V. S., and Wheater, H. S. (2006a). Quality control for daily observational rainfall series in the UK. *CIWEM*, To appear.

Yang, C., Chandler, R. E., Isham, V. S., and Wheater, H. S. (2006b). Spatial-temporal rainfall simulation using generalized linear models. *Water Resources Research*, In press.

Young, A. (2005). Stream flow simulation within uk ungauged catchments using a daily rainfall-runoff model. *Journal of Hydrology*, In press.

PB 12527/9

Nobel House
17 Smith Square
London SW1P 3JR

www.defra.gov.uk