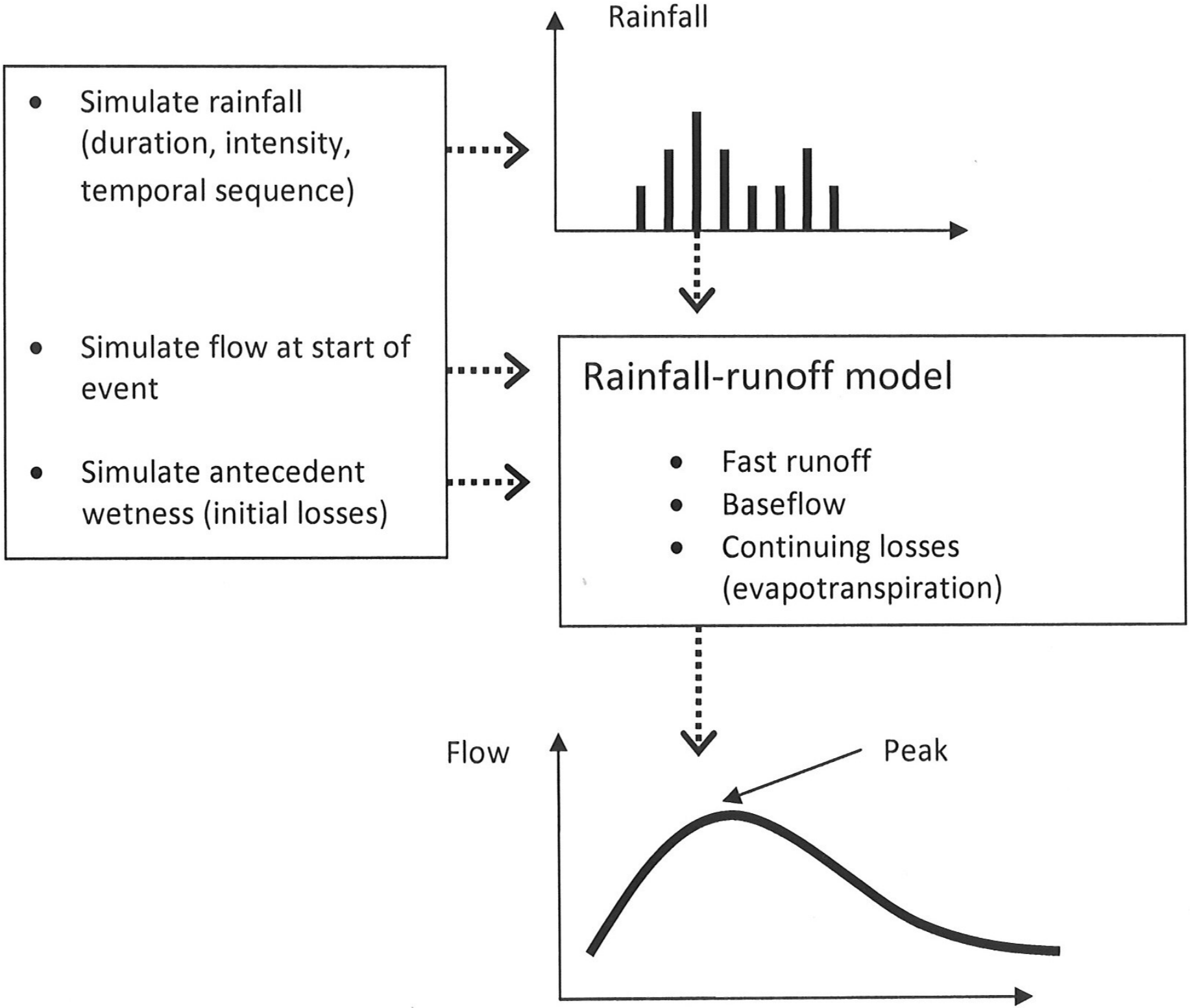
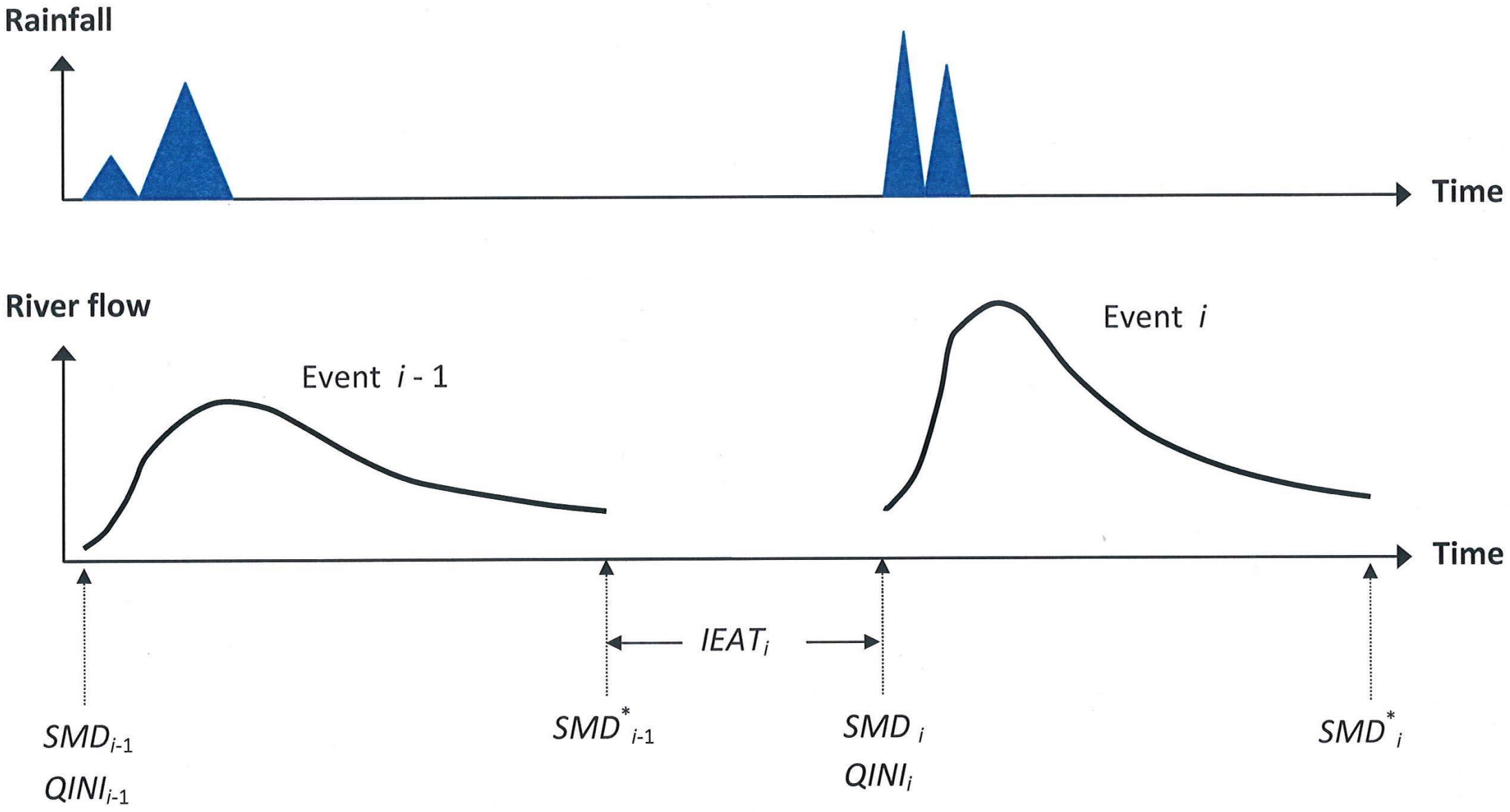
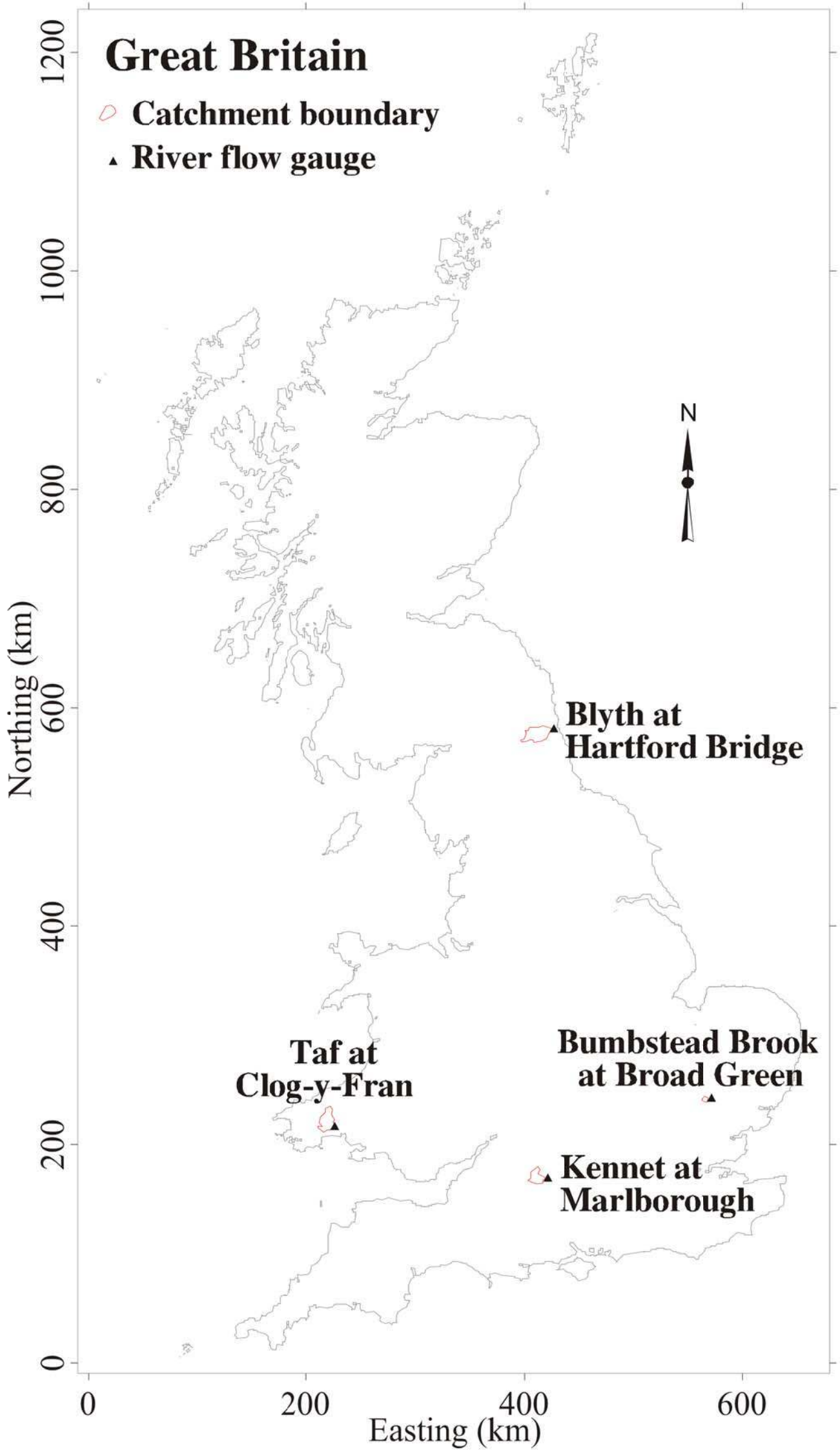
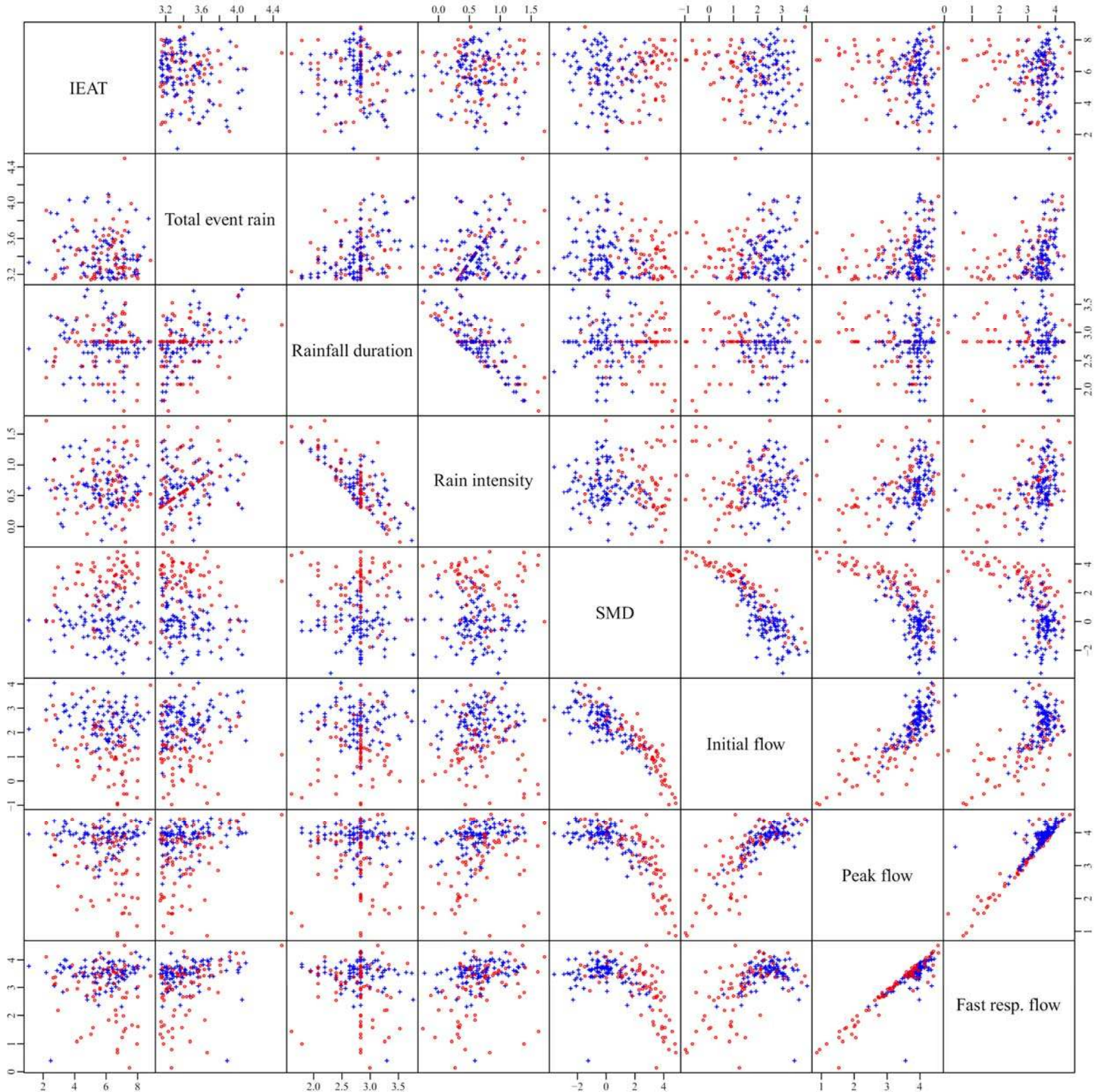
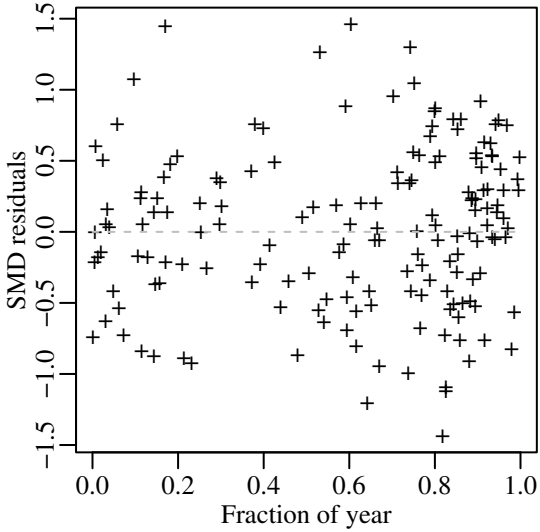

Rainfall
• Simulate rainfall (duration, intensity, temporal sequence)
• Simulate flow at start of event
• Simulate antecedent wetness (initial losses)
Rainfall-runoff model
• Fast runoff
• Baseflow
• Continuing losses (evapotranspiration)
Flow
Peak

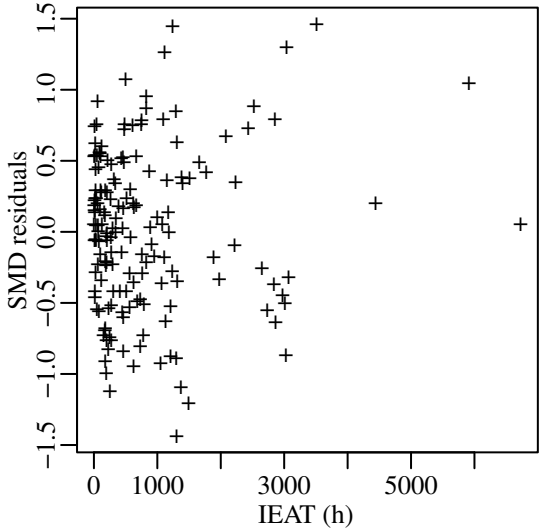

Rainfall
Time
River flow
Event i - 1
Event i
Time
IEAT_i
SMD_i-1
QINI_i-1
SMD*_i-1
SMD_i
QINI_i
SMD*_i

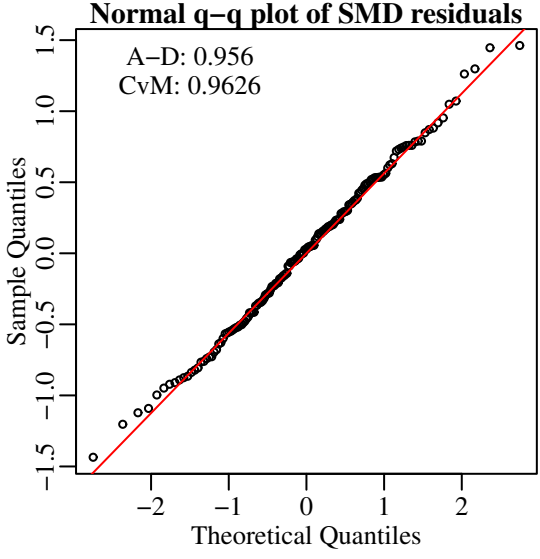

Great Britain
Catchment boundary
River flow gauge
N
Blyth at
Hartford Bridge
Taf at
Clog-y-Fran
Bumbstead Brook
at Broad Green
Kennet at
Marlborough
Northing (km)
0
200
400
600
800
1000
1200
0
200
400
600
Easting (km)

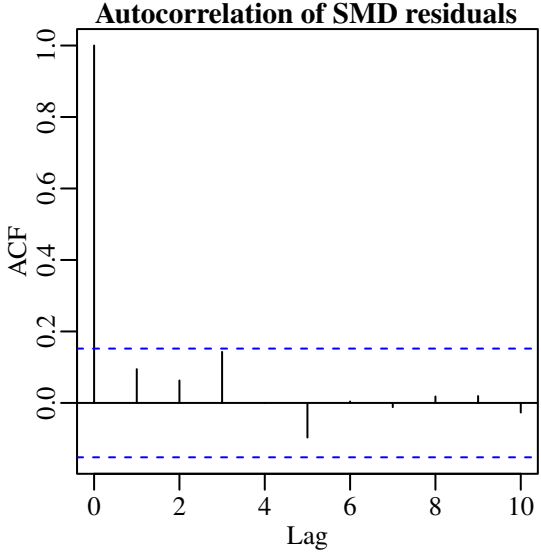

IEAT
Total event rain
Rainfall duration
Rain intensity
SMD
Initial flow
Peak flow
Fast resp. flow

Taf at Clog−y−Fran

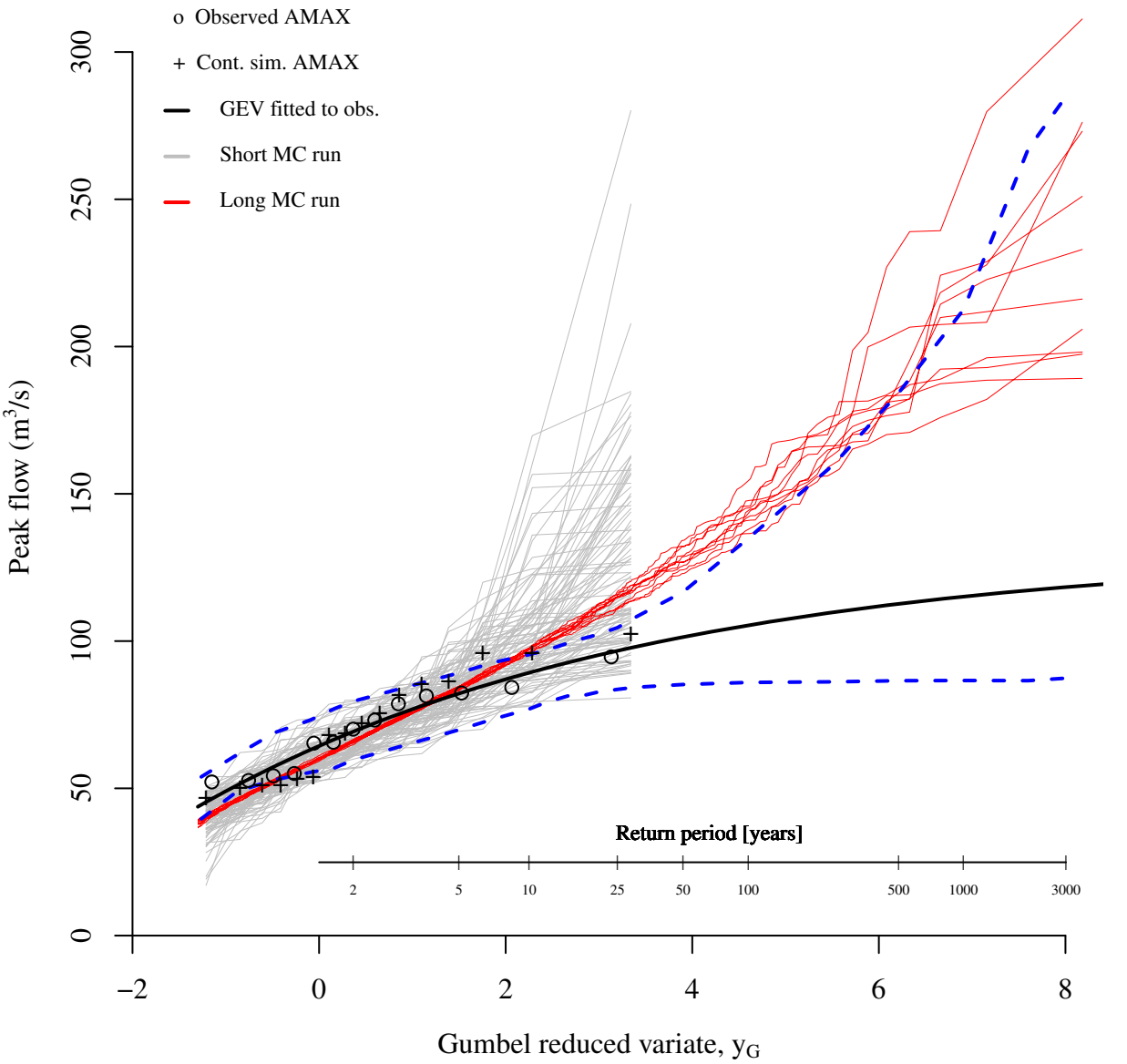


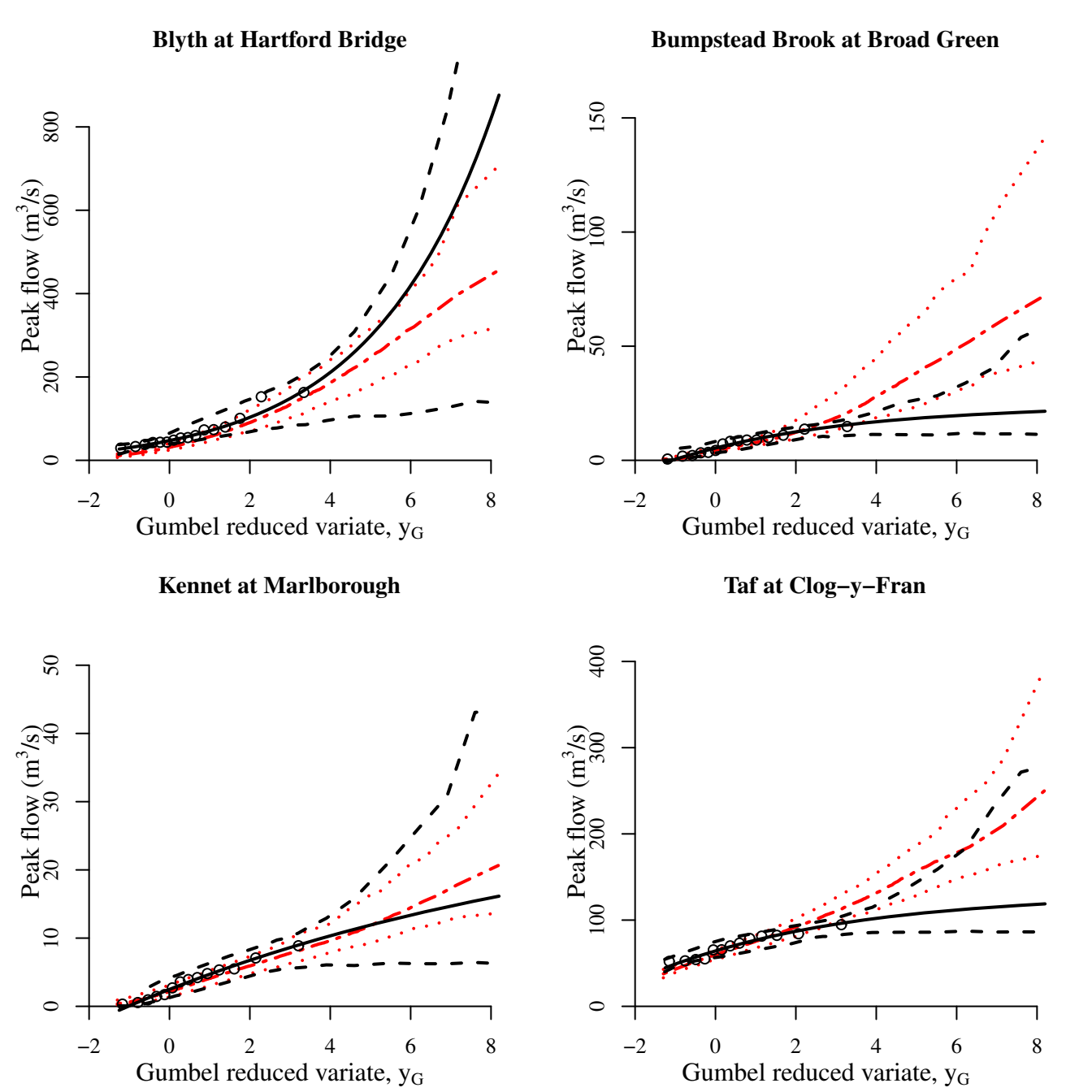


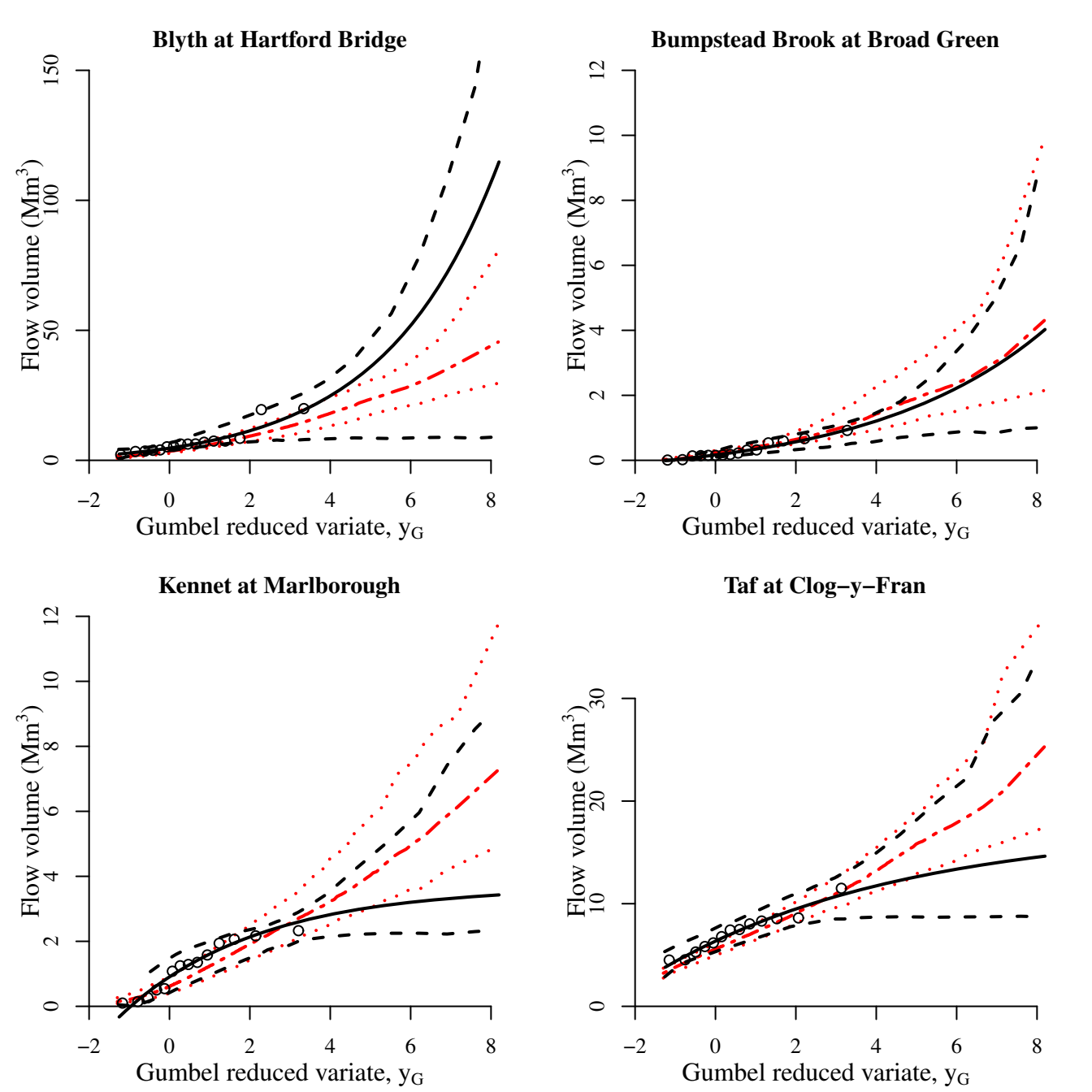




Taf at Clog–y–Fran

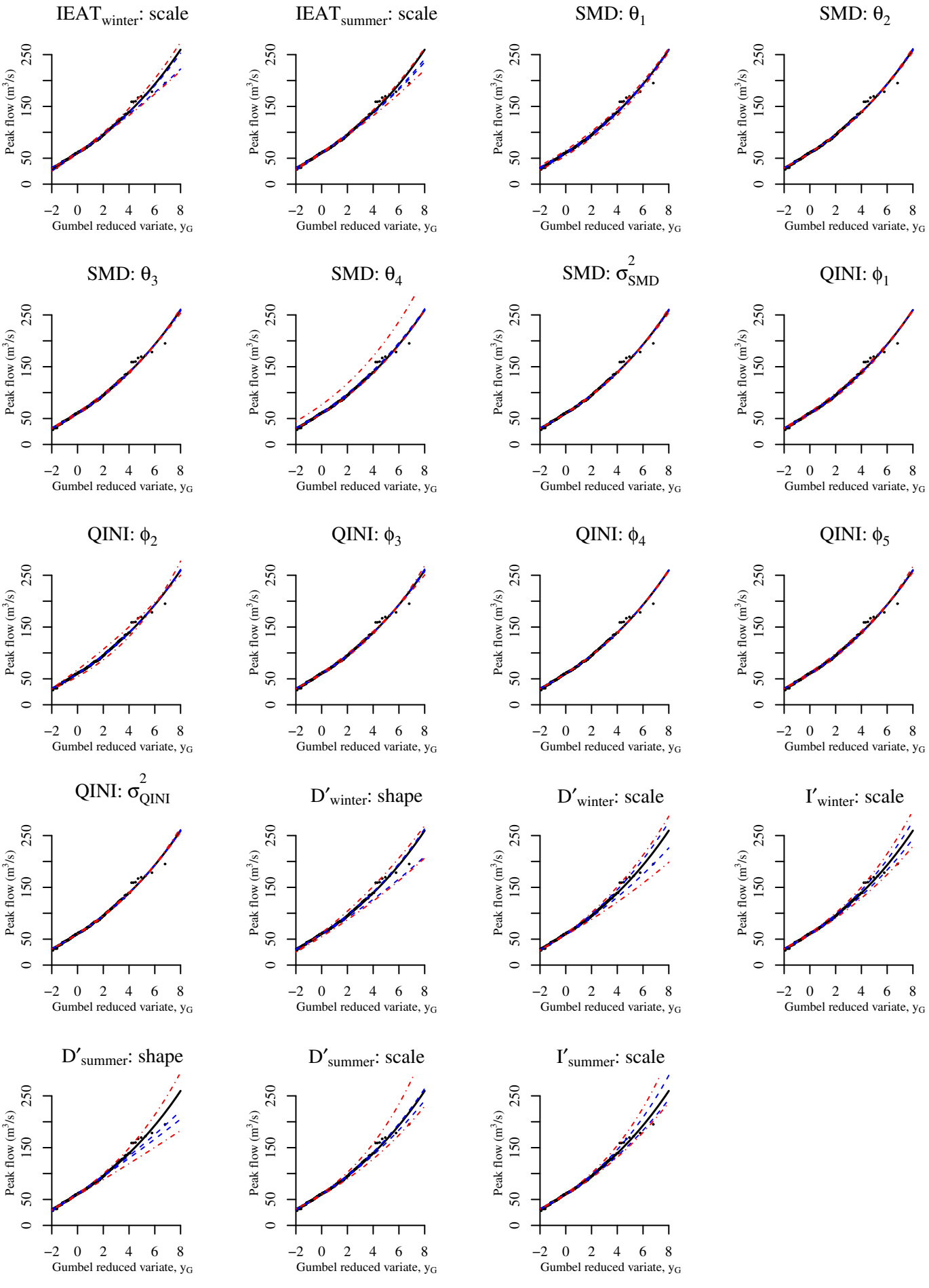



Blyth at Hartford Bridge
Peak flow (m3/s)
0
200
400
600
800
−2
0
2
4
6
8
Gumbel reduced variate, yG
Bumpstead Brook at Broad Green
Peak flow (m3/s)
0
50
100
150
−2
0
2
4
6
8
Gumbel reduced variate, yG
Kennet at Marlborough
Peak flow (m3/s)
0
10
20
30
40
50
−2
0
2
4
6
8
Gumbel reduced variate, yG
Taf at Clog−y−Fran
Peak flow (m3/s)
0
100
200
300
400
−2
0
2
4
6
8
Gumbel reduced variate, yG

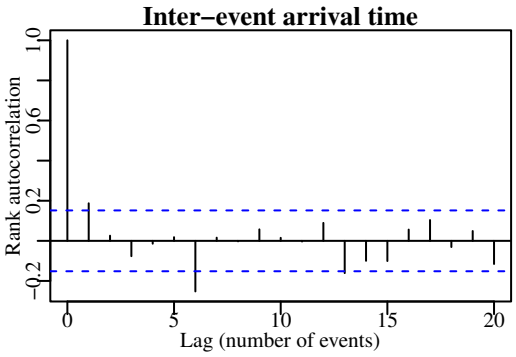

Blyth at Hartford Bridge
Flow volume (Mm3)
0
50
100
150
−2
0
2
4
6
8
Gumbel reduced variate, yG
Bumpstead Brook at Broad Green
Flow volume (Mm3)
0
2
4
6
8
10
12
−2
0
2
4
6
8
Gumbel reduced variate, yG
Kennet at Marlborough
Flow volume (Mm3)
0
2
4
6
8
10
12
−2
0
2
4
6
8
Gumbel reduced variate, yG
Taf at Clog−y−Fran
Flow volume (Mm3)
0
10
20
30
−2
0
2
4
6
8
Gumbel reduced variate, yG

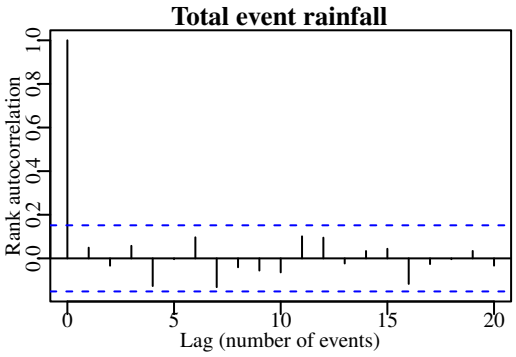
Freq. curves based on upper and lower confidence intervals around parameter estimates

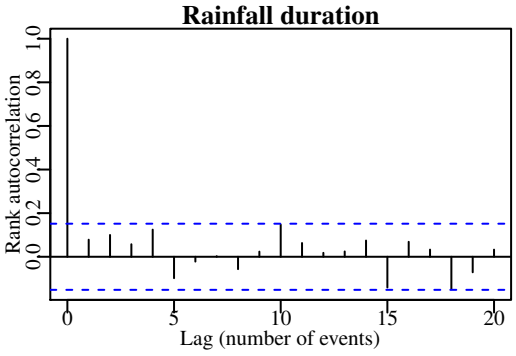
Taf at Clog–y–Fran

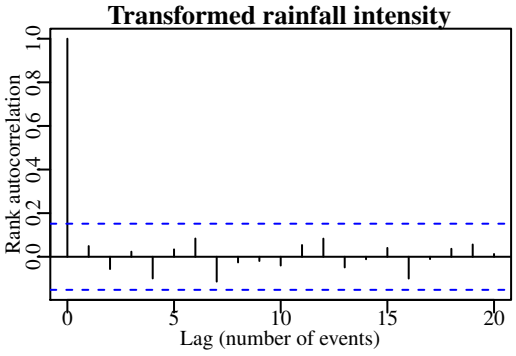


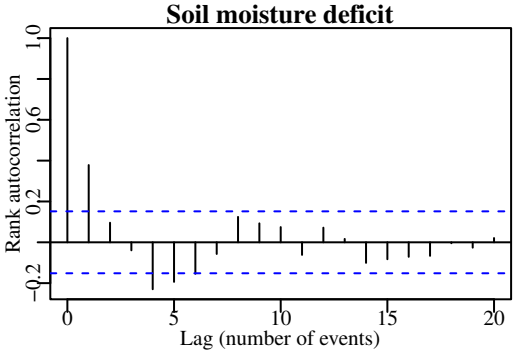
Taf at Clog−y−Fran, whole year

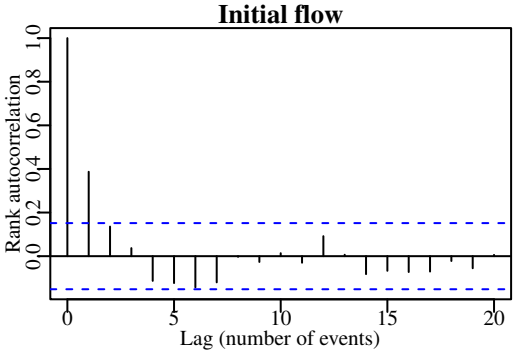


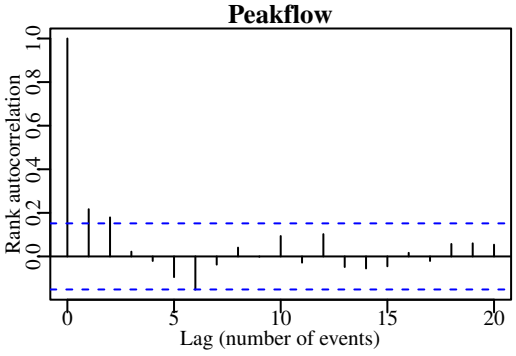


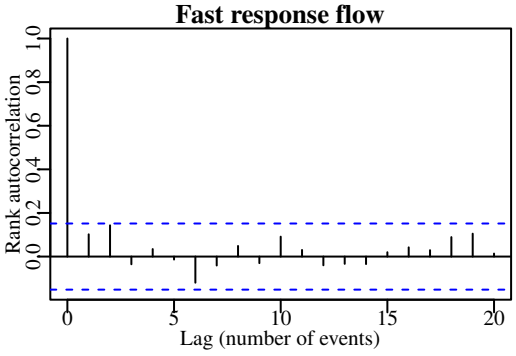


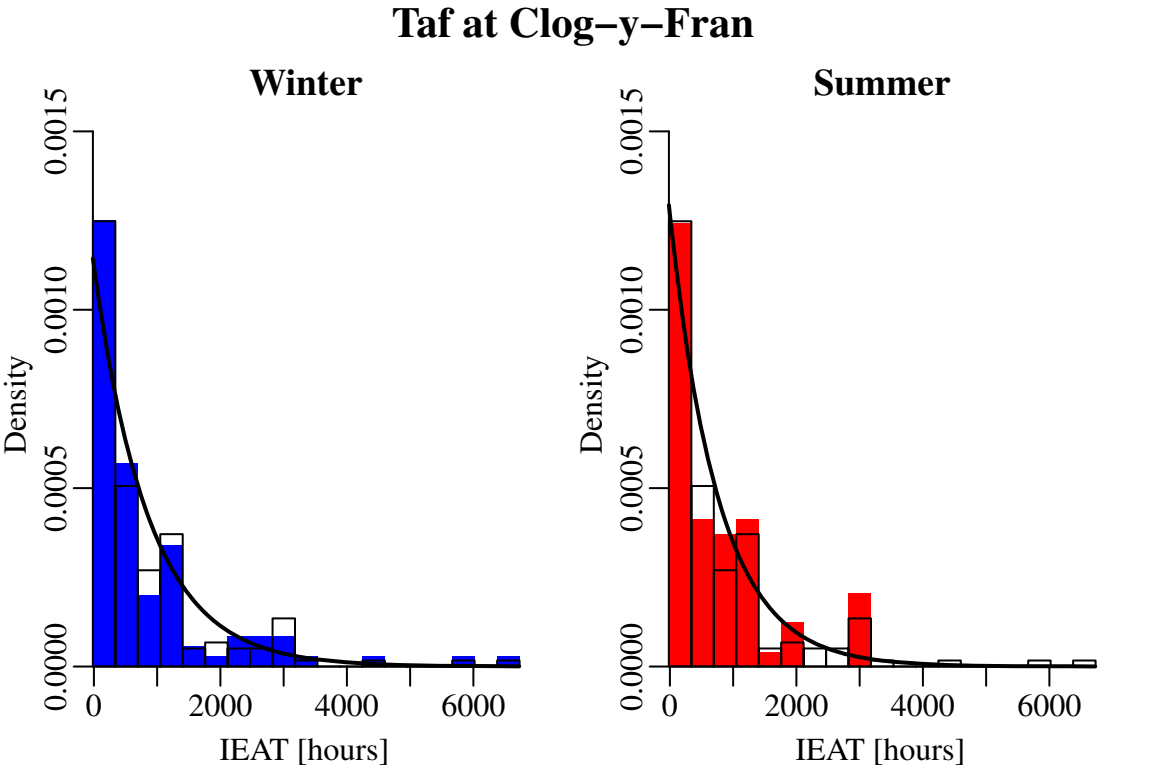


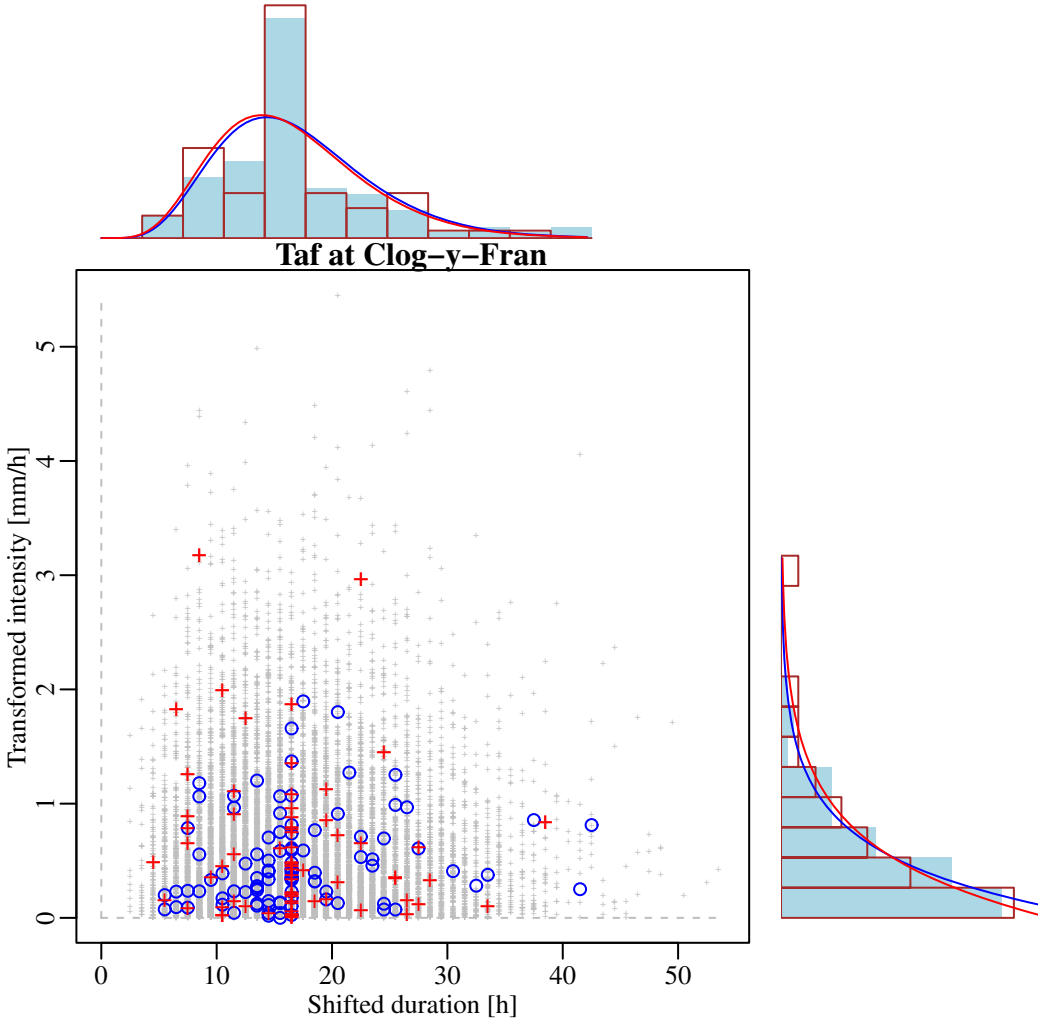


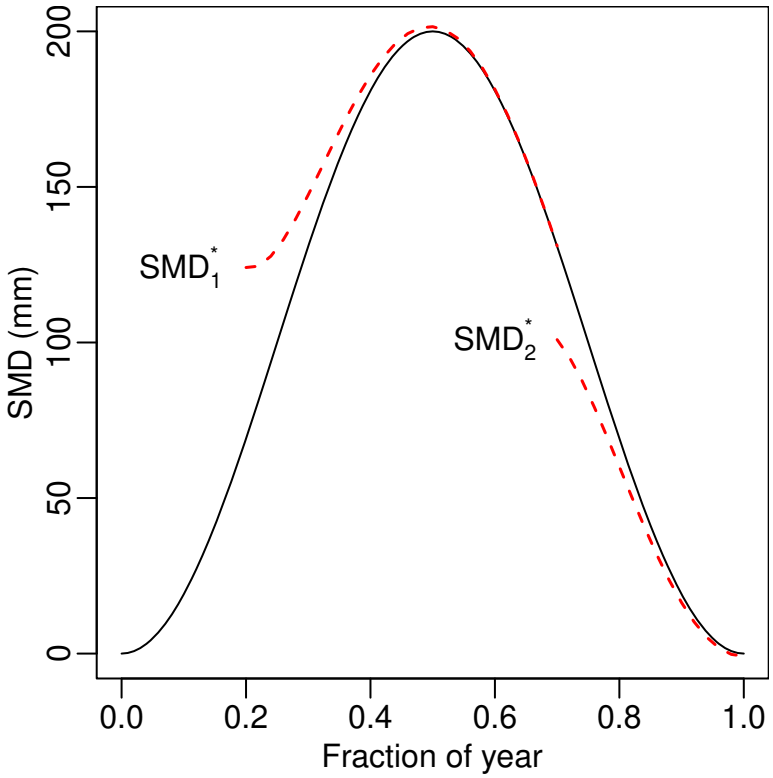


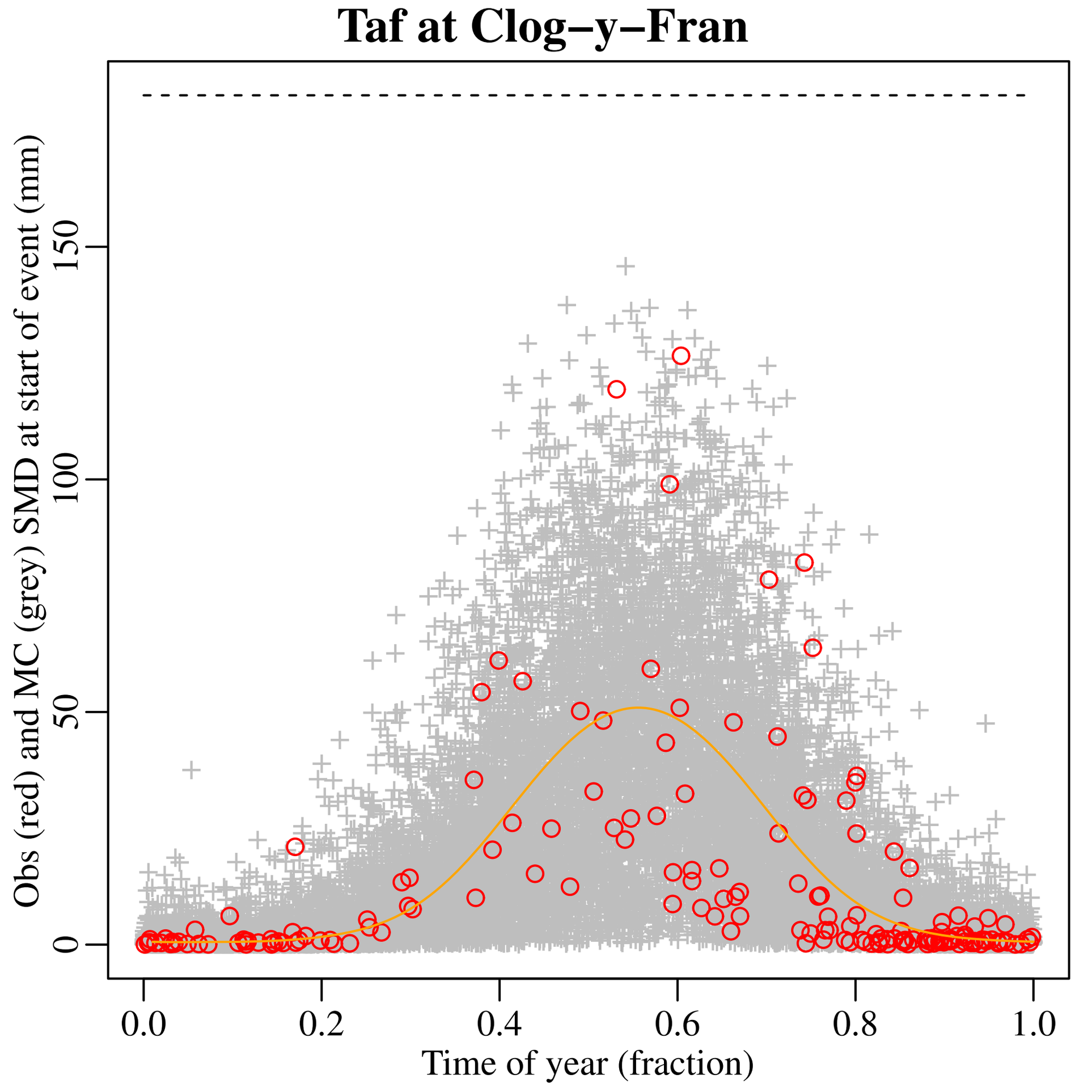





Taf at Clog−y−Fran
Winter
Summer
Density
0.0000
0.0005
0.0010
0.0015
0
2000
4000
6000
IEAT [hours]


Taf at Clog−y−Fran
Transformed intensity [mm/h]
0
1
2
3
4
5
0
10
20
30
40
50
Shifted duration [h]


SMD (mm)
0
50
100
150
200
0.0
0.2
0.4
0.6
0.8
1.0
Fraction of year
SMD*1
SMD*2


Taf at Clog–y–Fran
Obs (red) and MC (grey) SMD at start of event (mm)
0
50
100
150
0.0
0.2
0.4
0.6
0.8
1.0
Time of year (fraction)

Flood frequency estimation using a joint probability approach within a Monte Carlo framework

Prédétermination des crues en utilisant une méthode de probabilité combinée dans une structure basée sur la méthode de Monte-Carlo.

Cecilia Svensson[a], Thomas R. Kjeldsen[a], David A. Jones[a]

[a]*Centre for Ecology & Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire OX10 8BB, UK. Tel.: +44 1491 692363, Fax: +44 1491 692430*

Email addresses: csve@ceh.ac.uk (C. Svensson), trkj@ceh.ac.uk (T. R. Kjeldsen), daj@ceh.ac.uk (D. A. Jones)

Abstract

Event-based methods are used in flood estimation to obtain the entire flood hydrograph. Previously, such methods adopted in the UK have relied on pre-determined values of the input variables (e.g. rainfall and antecedent conditions) to a rainfall-runoff model, which is expected to result in an output flood of a particular return period. In contrast, this paper presents a method that allows all the input variables to take on values across the full range of their individual distributions. These values are then brought together in all possible combinations as input to an event-based rainfall-runoff model in a Monte Carlo simulation approach. Further, this simulation strategy produces a long string of events (on average 10 per year), where dependencies from one event to the next, as well as between different variables within a single event, are accounted for. Frequency analysis is then applied to the annual maximum peak flows and flow volumes.

Résumé

Les méthodes basées sur des évènements sont utilisées dans la prédétermination des crues afin d'obtenir un hydrographe de crue complet. Précédemment, ces méthodes adoptées au Royaume-Uni dépendaient de valeurs pré-déterminées des variables d'entrées (ex. données de pluies et conditions antérieures) au modèle qui devrait produire en résultat une crue correspondant à une certaine période de retour. Par opposition, cet article présente une méthode permettant aux variables d'entrée de prendre l'ensemble des valeurs de leur distribution propre. Ces valeurs sont alors assemblées en décrivant toutes les combinaisons possibles de données d'entrée au modèle pluie-débit d'évènements avec une approche basée sur des simulations de Monte-Carlo. De plus, cette stratégie de simulation produit une longue séquence d'évènements (en moyenne 10 par an) où sont pris en compte aussi bien les dépendences entre un évènement et le suivant que celles entres diverses variables relative au même évènement. Une analyse de fréquence est alors appliquée aux maxima annuels de débits de pointe et de débits volumiques.

Keywords

Flood frequency estimation; joint probability; Monte Carlo simulation; hydrological event modelling; uncertainty analysis; sensitivity analysis, river flow, rainfall, soil moisture deficit, Great Britain.

Mots-clefs

Estimation des fréquences de crues, probabilité combinée, simulation de Monte-Carlo, modélisation d'évènements hydrologiques, analyse des incertitudes, analyse de sensibilité, débit fluvial, précipitation, déficit hydrique des sols, Grande Bretagne.

1 INTRODUCTION

Estimation of design floods is a key task for practicing hydrologists, and the importance of the topic is illustrated by the sheer number of methodologies that can be found in the scientific and practical literature: see for example Lamb (2005) for a comprehensive review. The methods can involve simply fitting a probability distribution to a series of annual maximum peak flows and perhaps bringing in data from neighbouring stations in a regionalisation scheme to improve the flood estimates. However, when a design volume is required, the full flow hydrograph rather than just the peak flow is needed, and event-based or continuously simulated flow modelling is used. The latter may involve sophisticated rainfall-runoff modelling systems combined with stochastic rainfall generators. The actual choice of method for a particular study is often determined by practical considerations such as type and importance of the study, existing guidance documents (e.g. Bulletin 17B in the US or the Flood Estimation Handbook in the UK), previous experience of the analyst, and the availability of relevant data on which to base the analysis.

In most event-based methods, which are the concern of the present study, a set of pre-defined input values of antecedent conditions and a design rainfall are transformed into a design flood hydrograph of a particular return period via a rainfall-runoff method. Examples of event-based methods that have found widespread use in practise include, for example, the American TR-55 method (SCS 1986), the Australian rainfall-runoff method (IOEA 2001) and the British revitalised rainfall-runoff method (CEH 2007). However, a drawback of these methods is the difficulty in selecting the appropriate magnitude of the pre-determined input values, as poor choices can lead to biased flood estimates. A key question is therefore the treatment of the joint probability of occurrence between the individual flood producing processes and the resulting flood response. In a seminal paper, Eagleson (1972) provided a foundation for establishing a flood frequency relation in the absence of streamflow records by deriving it from storm and basin variables. An advantage of using such a joint probability method is that a combination of moderately high values of the input variables, which can be well estimated within their observed range, can result in a large flood. Therefore, the output flood can be estimated

through a method using less extrapolation (and expected greater precision) than would be involved if the flood were to be estimated from flow records alone.

Since Eagleson's paper was published, further studies of joint probability methods have been undertaken. With the advent of improved computing power, the generation of multiple realisations of the input variables within a Monte Carlo simulation framework has become a useful tool. The Monte Carlo-type methods generally involve the stochastic simulation of input variables (such as rainfall, antecedent soil moisture, initial flow), followed by the use of these as input to a rainfall-runoff model that may be fully deterministic or have stochastic components (e.g. Aronica and Candela 2007, Kwon *et al.* 2007, Sivapalan *et al.* 2005, Arnaud and Lavabre 2002, Loukas 2002, Rahman *et al.* 2002). The output consists of many realisations of flood hydrographs, which can be used for frequency analysis of peak flows or volumes. The aim of the present study is to use such a Monte Carlo framework to estimate frequencies of peak flow and total event flow volume. New, relatively simple models for the input variables will be constructed that adequately represent any dependencies between them, as well as any seasonality.

Compared with event-based methods, the flood estimates from continuous simulation methods are not as sensitive to potentially biased estimates of the initial conditions at the start of the model run. Here, the rainfall-runoff model automatically provides continuous model estimates of the antecedent soil moisture and initial river flow, that influence the river's response to a rainfall event. However, for simulating long flow series, the continuous simulation method relies on a rainfall generator that is correctly formulated to generate dry spells and low, moderate and extreme rainfalls in a realistic manner. In an event-based approach, only the larger rainfalls need to be specified, which makes for a simpler rainfall model. In the past, researchers have sometimes assumed independence between rainfall duration and rainfall intensity. Others have developed elaborate methods to account for the dependence between duration on the one hand, and rainfall intensity or total on the other (e.g. Balistrocchi and Bacchi 2011, Zegpi and Fernandez 2010, Rahman *et al.* 2002, Kurothe *et al.* 1997). We propose using a simple variable transformation to obtain independence between (transformed)

rainfall duration and intensity. Frequency distributions fitted separately to observations of the two transformed variables can then be used for simulation of the rainfall input to the rainfall-runoff model.

Purely event-based methods do not necessarily incorporate a seasonal component, although, for example, Sivapalan *et al.* (2005) and Swain *et al.* (1998) use observed monthly values to derive seasonally varying distributions for one or more of the input variables. However, individual events are simulated independently of each other (apart from any serial dependence arising from the seasonal component), which may only be appropriate provided that the time between events is long enough. We propose models of the input variables that are seasonally varying, and in addition we build in serial dependence between successively simulated events. This reflects the actual dependence that occurs for observed events and should lead to improved frequency estimates because any dependence in the flows for distinct events will affect the distribution of derived quantities such as annual maxima.

The paper is organised as follows. The overall simulation strategy is outlined in section 2, the original data is described in section 3, and the preparation of event data and exploratory analysis is presented in section 4. Lastly, the results are shown and discussed in section 5, and conclusions are summarised in section 6.

2 SIMULATION STRATEGY

This joint probability approach to flood estimation is based on modelling individual hydrological events through consideration of the major flood producing mechanisms and their mutual dependences. The proposed method allows all the input variables (rainfall, soil moisture, initial flow) to take on values across the full range of their individual distributions. These values are then brought together in all possible combinations as input to the rainfall-runoff model in a Monte Carlo simulation approach (Kjeldsen *et al.* 2010), as shown in Fig. 1. In a further extension of the event-based methodology, the simulation strategy proposed here will produce a long string of events, where dependencies from one

event to the next, as well as between different variables within a single event, are accounted for (Fig. 2). The method can be described as having the following steps:

1) Calibration of the rainfall-runoff model based on continuous hourly input data (observed series of hourly rainfalls, potential evaporation and river flows).
2) Selection of representative observed events (on average 10 per year), based on total event rainfall.
3) Estimation of the distributions of the input variables (inter-event arrival time, soil moisture deficit (SMD), initial flow, duration and intensity of rainfall) based on the characteristics of the selected observed events in Step 2.
4) Simulation of values from the fitted distributions in Step 3, and using them as input when running the calibrated rainfall-runoff model from Step 1 on an event basis, many times (Monte Carlo simulation).
5) Analysis of long strings of simulated events, for example, extraction of the annual maximum peak flows and fitting of a flood frequency curve.

3 DATA

The joint probability flood estimation method presented here was developed using four different catchments in Great Britain. Continuous series of catchment average rainfall, river flow, SMD and evaporation were used. These data series are 17 or 18 calendar years long depending on catchment, and are described below.

3.1 The test catchments

The four catchments were selected for their long concurrent hourly records of rainfall and river flow, and to represent different catchment permeabilities and climates (Table 1 and Fig. 3). Catchments further north were not included because snowmelt may feature more prominently, which was deemed

to be an unwanted complication for this initial study. Britain lies in the path of the mid-latitude westerly circulation, and is dominated by rainfall of frontal origin. Orographic enhancement of the rainfall is a feature in the mountainous north and west, whereas convective thunderstorms are more prominent in the continentally influenced climate of the southeast (Mayes and Wheeler 1997).

The smallest catchment in the study, the 28.3 km^2 Bumpstead Brook at Broad Green (in the southeast), is also the least permeable and fastest responding catchment. The most slowly responding catchment is the Kennet at Marlborough, which does not have the largest area, but has a very high base flow contribution to the total runoff due to its chalk geology.

3.2 Hourly river flow data and catchment average hourly rainfalls

Hourly river flow data for the four catchments were supplied by the Environment Agency (EA). The EA also provided the gauged hourly rainfalls for the calculation of the catchment average hourly rainfalls used in the study. The latter were derived by Lamb and Gannon (1996) and Crooks *et al.* (2002): areal average rainfall calculated from gauged daily data (provided by the UK Met Office) was distributed proportionally over the day using observations at an hourly gauge. When no hourly data were available, a fixed temporal profile was used to distribute the rainfall over the day.

3.3 Potential evaporation

Catchment average potential evaporation was calculated for each of the catchments based on monthly gridded data with a 40 km resolution from the UK Met Office Rainfall and Evaporation Calculation System (MORECS). These records span the study period for each of the catchments. A fixed seasonal profile of potential evaporation was used for the Monte Carlo simulations. This was calculated as the average for each month separately, i.e. the average over all Januaries, the average over all Februaries, etc.

3.4 Soil moisture deficit

A continuous hourly series of SMDs was derived for each catchment through continuous simulation using the PDM rainfall-runoff model (Moore 2007). See further details in section 4.1. Note that although continuously modelled SMDs are necessary for the development of the methodology, it is envisaged that when generalising the method to ungauged catchments, relationships will be sought directly between the (event) SMD model parameters and catchment/climate descriptors.

4 PREPARATION OF EVENT DATA AND EXPLORATORY ANALYSIS

The model building task was preceded by extraction of event data, and an exploratory analysis of marginal and joint distributions (dependence structures) of the flood producing variables, considering both variation within and between events. Individual, large rainfall events were identified from the continuous hourly series of catchment average rainfall. The initial flow at the beginning of each rainfall event was then extracted from the observed hourly river flow series, as was the associated flood peak. SMD at the beginning of each rainfall event was estimated from continuous simulation using the PDM rainfall-runoff model.

4.1 Continuous simulation

Continuous simulation using a rainfall-runoff model at an hourly time step was undertaken to i) provide continuous hourly series of SMD, ii) in a few cases to fill in missing flood peak data, and iii) to provide calibrated parameters of the rainfall-runoff model to use for the joint probability study. A lumped conceptual rainfall-runoff model was used; the probability distributed model (PDM) as described by Moore (2007) and previously applied to a range of flood studies in the UK (e.g. Lamb 1999, Kay *et al.* 2007, Roberts *et al.* 2009). The PDM represents runoff by individual point cylinders

whose storage capacity varies across the catchment according to a Pareto probability distribution. As the cylinders fill during rainfall, excess saturation occurs and the direct runoff is split between a fast and a slow (baseflow) routing component. Between rainfall events, the cylinders are depleted from evaporation and drainage. For this study, a 7 parameter form of the model was selected, which is similar to the 5-parameter model described in Calver *et al.* (2005) except that the soil store is characterised by the parameters c_{min} and b in addition to c_{max}. Using concurrent time series of observed streamflow, catchment average rainfall and potential evaporation, the model parameters of the PDM model were optimised to provide as good an agreement as possible between observed and simulated runoff. As well as simulated streamflow, the output of the PDM includes a continuous time series of soil moisture content, which has been converted into a continuous series of SMD, allowing a catchment SMD to be associated with the start of each selected observed event.

4.2 Rainfall event selection and associating flow peaks

On average 10 events per year were selected, based on the largest observed rainfall event total. The events were defined from rainfall rather than flow to ensure a sample representing a broader coverage of the joint distribution between the major runoff producing factors; rainfall and soil moisture. The catchment average hourly rainfall series and the catchment time-to-peak were used to define individual hydrological events. The event separation criteria were chosen so that each rainfall event could easily be associated with a peak in the river flow series, and are an adaptation of the criteria used by Rahman *et al.* (2002). Rainfall events were separated by at least a time span, C_d, during which the average rainfall did not exceed a threshold C_1, and no individual hour had more than C_2 mm of rain. After some trial and error these values were set as follows:

$C_2 = 0.05\ R$

where R is the 2-year return period 1-hour duration catchment average areal rainfall from the FEH depth-duration-frequency model (CEH 2009) for the particular catchment: for the selected catchments C_2 is roughly 0.5 mm. The parameter C_1 is determined by

$C_1 = 1.2\ C_2 / C_d$.

An event is accepted as valid provided it exceeds C_d times C_1 in total, and has at least one hour exceeding C_2. Using a fixed separation time for C_d (e.g. 4h or 6h) resulted in climatologically sensible rainfall durations, i.e. shorter durations in eastern Britain (where there is more convection) and longer durations in the west (which is dominated by longer, orographically enhanced, frontal events). However, particularly for the slowly responding catchments, there were often two rainfall events on the rising limb of the hydrograph, so that the flood peak was associated with two, rather than a single, rainfall event. This was overcome by relating C_d to the catchment time-to-peak, T_p, i.e. the time between a pulse of rain falling over the catchment and the flow peaking at the catchment outlet, as

$C_d = 0.75\ T_p$.

This led to longer separation times, and longer rainfall event durations, for the more slowly responding catchments.

The parameter T_p was estimated using impulse response function analysis. An auto-regressive model was fitted to the continuous hourly rainfall series, and then applied to both the rainfall and the river flow series. The lagged correlation coefficients between the residuals of the two series then represent the impulse response function, and the time lag for which the correlation is highest can be used as an estimate of T_p. Models were also fitted to the differenced series (1 time step) to remove any effects of trend or seasonality, to different seasons, and using different numbers of maximum lags. The time lag for which the correlation was largest was similar for the different variants of the analysis. Because there is a more direct relationship between rainfall and runoff in winter when SMDs are low, the final choice of T_p (Table 1) was based on the auto-regressive model with a maximum 5-hour lag, fitted directly to the original series for the winter season (November-April).

Lastly, the flow peak associated with the rainfall event was located. The search was undertaken in two stages, first looking at an interval close to the rainfall event, and then if no peak could be found, looking further ahead. This was advantageous for resolving any timing issues, particularly for the case when hourly rainfall data was unavailable, and an assigned rainfall profile had been used for

disaggregating the daily rainfall data to an hourly resolution. It also reduced the risk of linking a rainfall event to a flow peak that was really associated with a later rainfall event. The length of the first search period started at the time-of-centroid of the rainfall event and ended 1.5 T_p after the end of the rainfall event. The length of the second search period was T_p, starting at the end of the first search period. Where there were missing data in the observed hourly river flow series, the peak flow from the continuously modelled series was substituted.

4.3 Characteristics of the hydrological event

The total duration of the flow event, hereafter referred to as the hydrological event, was determined as starting at the beginning of the rainfall event and ending 3 T_p after the end of the rainfall event. The total flow volume for an event was defined as the area under the flow hydrograph for the duration of the hydrological event. A simple measure of hydrograph shape was used to characterise the length of the recession. The hydrograph shape was defined as the time elapsed between 25% and 75% of the total flow volume occurring.

4.4 Cross-correlation in the input variables

Initial scatter plots and correlation analyses of the event data revealed that there was dependence between some of the input variables, which needed to be taken into account when building the joint probability model. Examples of scatter plots for a selection of variables (in log space) are shown in Fig. 4 for the Taf at Clog-y-Fran. In these plots, the inter-event arrival time (IEAT) is the time elapsed from the end of one hydrological event to the start of the next rainfall event. The rainfall intensity is the total event rainfall divided by the rainfall event duration, and the fast response flow is the peak flow minus the initial flow at the start of the event.

The peak flow and the fast response flow variables generally show dependence with all the other variables except for the inter-event arrival time, for the year, summer and winter. However, the

relationships are rather sketchy for the rainfall duration. There is also dependence between SMD and initial flow, and between total event rainfall or intensity, on the one hand, and rainfall duration, on the other. For each variable, the samples for the winter and summer seasons generally suggest that these come from different populations, and that seasonality therefore has to be taken into account for the stochastic modelling of the input variables to the rainfall-runoff model.

A particular feature of the Taf catchment is the comparatively large number of days for which there was no recording raingauge nearby to use for distributing the daily rainfall observations over the hours of the day. Hence the standard temporal profile was used, which can be seen in the scatter plots involving the rainfall duration: there is a disproportionate number of rainfall events with a duration of 17 hours.

4.5 Serial correlation in the input variables

There is serial correlation in some of the input variables, which needs to be taken into account for the Monte Carlo simulation. Figure 5 shows the Spearman's rank autocorrelation and 95% significance level for the inter-event arrival time, SMD, and various rainfall and river flow variables for the Taf at Clog-y-Fran.

There is no lag-1 correlation in the rainfall variables, and the Taf is the only catchment with significant lag-1 correlation in the inter-event arrival time. However, this correlation is no longer significant when the analysis is done on a seasonal basis (not shown), which suggests that at least some of the dependence is due to seasonality and that this therefore has to be taken into account for the Monte Carlo simulation. The transformed rainfall intensity used for this analysis is described in section 5.2.

The SMD and initial flow variables always have significant correlations for at least one lag, which is generally carried over to the peak flow. For the fast response flow, i.e. the peak flow minus the initial flow at the start of the event, the correlations are generally reduced but remain quite large for the

Bumpstead Brook and the Kennet. They are reduced further in the seasonal analysis for these two catchments, but do not quite become insignificant.

5 STOCHASTIC MODELS FOR GENERATING INPUTS TO THE RAINFALL-RUNOFF MODEL

Stochastic models were developed for the variables controlling the simulations:

- Inter-event arrival time
- Rainfall duration, intensity and temporal profile
- SMD at the onset of the event
- Initial flow

The stochastic models were developed with a view to keeping the model structure simple, and for the parameter values for each catchment to be derived from climate or catchment characteristics. This keeps open the possibility of generalising the method at a later stage, so that it can be used for flood estimation in ungauged catchments.

5.1 Inter-event arrival time

The inter-event arrival time is the time elapsed from the end of one hydrological event (3 T_p after the end of the rainfall event) to the start of the next rainfall event. Histograms of extracted IEATs for the summer and winter seasons are shown in Fig. 6 for the Taf at Clog-y-Fran. In some cases the next observed rainfall event starts before the river flow recession is complete, and the IEAT is negative. The number of such events varies, from one event only for the Taf to nearly 18% of events for the slowly responding Kennet, with the number roughly increasing with increasing T_p of the catchment. Different statistical distributions were explored for the IEAT, but for simplicity a one-parameter exponential distribution was fitted to IEATs that had been shifted to ensure they were always positive.

The distributions were fitted separately for summer and winter, depending on the season in which the previous event occurred (Fig. 6). The shift was fixed at 2.25 T_p, as the IEAT cannot be smaller than -2.25 T_p when taking into account the criteria for rainfall event separation. When the IEAT is negative, the SMD and the initial flow for the next event are taken from the output from the event-based PDM rainfall-runoff model (for the previous event), rather than being derived from the stochastic models. On average, 10 events per year will be simulated.

5.2 Rainfall duration and intensity

The selection of rainfall events results in a lower bound for both event duration, *D*, and depth, *P*. As indicated in Fig. 4, there is evidence of positive dependence between duration and total event rainfall depth, whereas there is negative dependence between duration and rainfall intensity because of the artificial lower bound due to the event selection criteria. This artificial dependence can be removed by transforming the intensity, taking into account the lower bounds (Fig. 7). The lower bound of the rainfall depth is the threshold used for the rainfall event selection, P_{min}. For the duration, the lowest value that can occur is one hour, which is the resolution of the rainfall data. To avoid division by zero the transformed intensity, I', is defined as

$$I' = \frac{P - P_{min}}{D - 0.5} \quad (1)$$

The shifted duration, $D' = D - 0.5$, and transformed intensity, I', are uncorrelated using Spearman's rank correlation. The marginal distributions of D' and I' are modelled using a one-parameter exponential distribution for the transformed intensity and a two-parameter gamma distribution for the shifted duration. Separate models are fitted for the summer and winter half-year, and when using these for the Monte Carlo simulation the generated durations are rounded to the nearest whole hour.

Figure 7 shows histograms and fitted distributions for the shifted duration (x-axis) and transformed intensity (y-axis) for the Taf at Clog-y-Fran. There is not much difference between the seasons for this

catchment, but other catchments show longer durations and lower intensities in winter than in summer. The grey crosses in Fig. 7 represent 100 Monte Carlo simulations of 17 years, and suggest a good correspondence between the observations and the simulated values.

5.3 Rainfall temporal profile

After investigation of several temporal rainfall profiles, a double triangle was chosen. The effect of five different temporal rainfall profiles on the simulated peak flow, event flow volume and hydrograph shape was investigated. These simulated flow characteristics were compared with the corresponding observed ones using scatter plots and histograms. The profiles included a single triangle, a double triangle, two triangles separated by a dry period, a uniform rectangle and two equal-sized rectangular pulses in the first and last quarter. The three different triangular profiles showed very similar results to each other. The uniform and double rectangular pulse profiles tended to produce smaller peak flows and flow volumes than the triangular profiles, although all the profiles tended to produce somewhat smaller events than observed. There is a bias towards too long recessions (large hydrograph shapes) for all the profiles for all the catchments except for the slowly responding Kennet, but this is more pronounced for the uniform and double rectangular pulse profiles than for the triangular profiles.

The selected double triangle profile can be considered to reflect two bursts of rainfall. The profile consists of two immediately adjacent triangles (see the examples in Fig. 2). The duration of the first triangle is allowed to vary between 30 and 70% of the total event duration, and is randomly generated from a uniform distribution. The total rainfall depth in the first triangle is generated in the same way to vary between 30 and 70% of the total event rainfall depth. The second triangle makes up the balance. This simplified temporal profile does not capture the full variability of observed rainfalls, and, as discussed above, therefore does not produce quite as large flow events as observed.

5.4 Soil moisture deficit

A model for generating an appropriate value of SMD at the onset of an event has been defined as a typical value of SMD at a particular time of year, combined with a deviation from this typical value depending on both the time elapsed since the end of the previous event, and the SMD at the end of the previous event (denoted here by SMD*). This concept is illustrated in Fig. 8. Based on exploratory analysis, the study finally settled on the following model structure:

$$\underbrace{\ln\left[\frac{SMD_i^{1/3}}{S_{\max}^{1/3}-SMD_i^{1/3}}\right]}_{\text{Measure of SMD}} = \underbrace{\mu(f)}_{\substack{\text{Typical SMD}\\ \text{at the time of}\\ \text{year when}\\ \text{event } i \text{ starts}}} + \underbrace{\left(\ln\left[\frac{\left(SMD_{i-1}^*\right)^{1/3}}{S_{\max}^{1/3}-\left(SMD_{i-1}^*\right)^{1/3}}\right]-\mu(f^*)\right)}_{\substack{\text{SMD anomaly at the}\\ \text{end of the previous event}}}\underbrace{\exp(-\theta_4\, IEAT_i)}_{\substack{\text{Decay with}\\ \text{time since}\\ \text{previous event}}} + \underbrace{\varepsilon_i}_{\substack{\text{Model}\\ \text{error}}} \quad (2)$$

where $S_{\max}$ is the total available soil moisture storage obtained through calibration of the PDM model. The term f is the time of year (as a fraction of a year) of the onset of event i, and f^* is the corresponding time at the end of the hydrological event i-1. The typical SMD value at any time of the year, $\mu(f)$, can be thought of as the value that the variable $\ln\left[\frac{SMD_i^{1/3}}{S_{\max}^{1/3}-SMD_i^{1/3}}\right]$ strives to return to, and it is modelled as

$$\mu(f) = \theta_1 + \theta_2 \sin(2\pi f) + \theta_3 \cos(2\pi f)\,, \quad (3)$$

where $\boldsymbol{\theta}$ is a vector of model parameters.

Because very low values of the SMD can cause problems in the later model for the initial flow, a final calculation is carried out when simulating from this model:

$$SMD_i^{\text{final}} = \max(SMD_i, SMD_{\min})\,, \quad (4)$$

where $SMD_{\min}$ was set to an arbitrary low value of 0.05 mm.

Figure 9 shows "observed" SMD (i.e. output from the continuous PDM simulation, see section 4.1) as red circles, the predicted typical $\mu(f)$ curve (in orange) and values of SMD generated for 100 Monte Carlo simulations of 17 water years (1 October – 30 September) each (grey crosses) for the Taf at Clog-y-Fran.

The model error ε_i is independent and identically distributed, with a mean value of zero and constant variance σ^2_{SMD} (Fig. 10). However, it is also assumed that it should follow a normal distribution, which it does not do well for several of the catchments according to the Cramér-von Mises test, but particularly badly for the Bumpstead Brook. The fractions of the variance explained by the SMD models for the four catchments vary from 66% (the Taf) to 91% (the Kennet).

The particular formulation of the quantity $\left[\frac{SMD_i^{1/3}}{S_{\max}^{1/3} - SMD_i^{1/3}}\right]$ has the following reasons. The logarithm of the ratio is similar to a logit transformation, which allows the transformation of values within a bound (in this case $SMD^{1/3}$, which varies between 0 and $S_{\max}{}^{1/3}$) to an unbounded variable (the ratio) to which an error can be added at any scale. Compared with using the simpler expression $\frac{SMD_i}{S_{\max} - SMD_i}$, the cube root transformed quantity has an effect particularly on the top end of the SMDs, drawing them closer to the typical curve. It makes the residuals more identically distributed over the year and also makes them follow a normal distribution somewhat better. The explained variance of this model is slightly higher (between 1 and 6 percentage units) than for the corresponding model without the cube root transformation, but there is not much difference to the final flood frequency curves.

5.5 Initial flow

The initial flow, $QINI_i$, at the onset of the event i is modelled as a function of the SMD at the onset of the event and the time of year at the onset of the event, and also depends on the error, ε_i, from the SMD model.

$$\underbrace{\ln(QINI_i)}_{\substack{\text{Measure of}\\ \text{initial flow}}} = \underbrace{\phi_1 + \phi_4 \sin(2\pi f) + \phi_5 \cos(2\pi f)}_{\substack{\text{Typical initial flow}\\ \text{at the time of year}\\ \text{when event } i \text{ starts}}} + \underbrace{\phi_2 \ln\left[\frac{SMD_i^{1/3}}{S_{\max}^{1/3} - SMD_i^{1/3}}\right]}_{\substack{\text{Measure of SMD at}\\ \text{the start of event } i}} + \underbrace{\phi_3 \varepsilon_i}_{\substack{\text{Error}\\ \text{from SMD}\\ \text{model for}\\ \text{event } i}} + \underbrace{\eta_i}_{\substack{\text{Model}\\ \text{error}}} \quad (5)$$

where ϕ is a vector of model parameters, and η is a set of errors that are assumed to be independent and identically distributed, and to follow a normal distribution with zero mean and constant variance σ^2_{QINI}. The inclusion of the term ε_i allows the dependence between $QINI_i$ and SMD_{i-1} (that is not explained by SMD_i) to be represented in the model. The QINI model residuals for the Bumpstead Brook again show the worst match to a normal distribution, as they did for the SMD model, and there is serial correlation for four lags events for the Kennet.

Several formulations of the relationship between *QINI* and *SMD* were explored, but ln(*QINI*) versus $\ln\left[\frac{SMD_i^{1/3}}{S_{\max}^{1/3} - SMD_i^{1/3}}\right]$ shows good linearity for all the catchments. The fractions of the variance explained by the initial flow models vary between 71% (the Bumpstead Brook) and 84% (the Taf). The addition of the term for the error ε_i from the SMD model was mainly needed for the very permeable Kennet and the reasonably permeable Taf catchments. However, it is still significant at the 5% level for the Bumpstead Brook and at the 10% level for the Blyth, although the parameter values for these catchments are negative rather than positive as for the two permeable catchments. Because the error ε_i is used as a predictor, it is convenient that both the SMD and the initial flow models use

the same SMD transformation, $\ln\left[\frac{SMD_i^{1/3}}{S_{\max}^{1/3} - SMD_i^{1/3}}\right]$, as the relationship between the error and ln(*QINI*) might otherwise have been less linear.

6 RESULTS AND DISCUSSION

6.1 Frequency analysis of flood peaks

The results from the joint probability model in the Monte Carlo framework will be shown together with the corresponding results based on the observed flow series, and in some cases also with the results based on the continuous hourly PDM simulation of river flows, hereafter referred to as the "continuous simulation".

Figure 11 shows flood frequency results for the Taf at Clog-y-Fran. The grey lines are the annual (water year) maximum flood peaks from 100 16-year long Monte Carlo simulations, which is the number of water years in the observed hourly series. Each simulation starts from a different random seed, and the figure therefore shows the effect of the sampling variability on the produced flow peak series. The red lines show 10 much longer simulations (2000 years long), and it can be seen that the lower parts of the frequency curves stabilise with increased simulation period, whereas the upper parts still diverge. These curves fit slightly better to the annual maxima from the continuously simulated flow series (crosses) than to those from the observed flow series (circles), which seems reasonable since the annual maxima from both the event-based Monte Carlo simulation and from the continuous simulation are outputs from the rainfall-runoff model (and therefore will share the same bias from its calibration). The observed rainfall events, on which the rainfall model is based, are also subject to the same standard rainfall profile as the continuous rainfall record used for the continuous simulation, for days when there was no recording gauge available to distribute the daily rainfall total over the hours of the day. Note that there are fewer annual maxima in the observed than in the continuously simulated series, because of missing data in the observed hourly flow series. At least 75% of observations in a year had to be present for an annual maximum to be derived.

A Generalised Extreme Value (GEV) distribution is fitted using L-moments to the observed maxima (thick black line), with a 95% confidence interval (dashed blue line) around the GEV frequency curve calculated based on a bootstrap method. The observed annual maxima series was resampled 999 times, and for each resample a GEV was fitted. For a selection of percentiles, the 999 quantile estimates for each percentile were ranked, and the 25th and 974th values were used to describe the 2.5 and 97.5 points of the distribution. The upper and lower confidence limits were plotted by connecting all the 2.5, and the 97.5, points, respectively.

The 95% confidence interval encompasses the bulk of the 16-year long Monte Carlo simulations for the return periods below about 5 years, but simulated outliers disperse the curves (grey lines) at higher return periods. The 2000 year long simulations (red lines) are less dispersed, and give a reasonable indication of the flow frequency curve corresponding to the joint probability model that has been fitted. The joint probability model thus suggests that the catchment could be expected to experience more high flow events in a typical 16-year period than observed during these particular 16 years of the data record.

6.2 Uncertainty analysis of predicted flood peaks

The uncertainty of the flood peak magnitudes, as predicted by the joint probability model, can be described by 95% confidence intervals estimated using a bootstrap method (Fig. 12). The original event data were resampled in blocks of one year (to retain seasonality), and the parameters for the stochastic models described in Section 5 were estimated for each of these new data series, 199 in total. One 2000-year long series was simulated for each set of parameters, and the resulting annual maxima were ranked separately for each of the 199 series. For each rank, the 5th and 195th values were plotted using the Gringorten plotting position and used to denote the 2.5 and 97.5 points of the distribution, i.e. the 95% confidence interval (dotted red lines in Fig. 12). Similarly, the 100th value for each rank was used to illustrate a “median” flood frequency curve (solid red line).

These model runs all start using the same random seed, so that the spread outlines the variability in the parameter estimates rather than the randomised sampling from the distributions. Figure 12 also shows the GEV distributions (solid black line) fitted to the observed annual maxima (circles) for each catchment, and the 95% confidence intervals around the GEV (as in Fig. 11). The confidence intervals around the fitted GEVs are larger and generally overlap the confidence interval for the Monte Carlo simulations at low return periods. However, they can diverge quite considerably for higher return periods. Particularly, the joint probability model predicts peak flow distributions that are not bounded above for any of the catchments, and does not reflect the bounded distributions suggested by the GEV fitted to the observed data for some of the catchments. The observed series are short, and as discussed earlier the low annual maxima may be a result of sampling variability.

6.3 Uncertainty analysis of predicted flood volumes

The method used to estimate the uncertainty of the predicted flood volumes is the same as that for the predicted flood peaks in the previous section. The results are shown in Fig. 13. It can be seen that the confidence intervals for the predicted volumes largely overlap the confidence interval of the GEV fitted to the observed annual maximum flood volumes (from the dataset comprising on average 10 events per year). Here, the poorest fit is for the Kennet at Marlborough, where the predicted curve turns upwards whereas the frequency curve fitted to the observations is bounded above. Similarly to the frequency curves for the flood peaks, the predicted frequency curves for the flood volumes are not bounded above for any of the catchments.

6.4 Sensitivity analysis for predicted flood peaks and volumes

The sensitivity of the flood frequencies to the sampling variability of the different stochastic model parameters described in Section 5 can be investigated using a bootstrap-based approach. Once a confidence interval for each estimated parameter has been calculated, the upper and lower bounds can

be used to generate the inputs to the rainfall-runoff model. This will show how sensitive the peak flow (or volume) estimates are to sampling variability in each parameter. A comparison of the relative sensitivities of the parameters is shown in Fig. 14 for the Taf at Clog-y-Fran, using the resulting frequency curves for peak flows. In each diagram, all parameters except the target parameter have been estimated using the original data set. The target parameter has been given the value of the upper and lower confidence interval, and hence the results are shown in pairs of frequency curves. Two different colours denote two different sensitivity measures: the pairs of red dash-dotted lines are derived from the unconditional variance for a given parameter and the pairs of blue dashed lines are from the conditional variance given all the other parameters (see Appendix). The central black line on each diagram is the frequency curve estimated using only parameters based on the original data set. Each frequency curve is a GEV fitted to a single 500 year long simulation. That is, in total five model runs were made to produce the five curves for each diagram. Each model run started from the same random seed so that any variation in the frequency curves will result from the difference in the parameter values, rather than from the effects of sampling from the parameter distributions.

The unconditional variance from the bootstrap results represents how well-determined a given parameter is by the data-fitting, given that a poor choice of value for that parameter could be compensated for by changing the values of other parameters.

The conditional variance that is derived from the covariance matrix obtained from the bootstrap results represents an approximation of how well-determined a given parameter is by the data-fitting, given that the other parameters are left at fixed values. This is just a convenient approach used here to avoid the more computationally expensive task of undertaking multiple bootstrapping experiments, with each of them fitting just one parameter while retaining fixed values for the others. The approximation assumes that the statistical relationships (arising from fitting the parameters to a finite amount of data) among the fitted parameters are approximately linear with constant unexplained variance.

The pair of frequency curves based on the unconditional variance generally envelopes the pair for the conditional one and the central estimate, as might be expected. This sometimes doesn't happen, presumably reflecting a nonlinearity in the behaviour of the modelled flow-frequencies in response to changes in the parameters.

Figure 14 shows that the peak flow frequency estimates are sensitive to sampling variability mainly in the IEAT and rainfall parameters. This is true for all the catchments. For some catchments, the duration parameters have comparatively tight ranges for the conditional bands compared to those for the unconditional ones, suggesting that the limited amount of data makes it difficult to estimate both the parameters of the gamma distribution being used here. However, there seems no reason to select any particular value for the shape parameter which might be a possible course of action: fixing the value to correspond to an exponential distribution was judged unsuitable. In addition, in the two more permeable catchments, the Kennet and the Taf, θ_4 and ϕ_2, the parameters controlling the main non-seasonal components in the SMD and initial flow models, are poorly determined by the available datasets. This also applies to the non-seasonal component of the SMD model for the fast-responding Bumpstead Brook catchment. The intercepts, seasonal components and model errors are generally well determined apart from, for the Kennet, the intercepts, θ_1 and ϕ_1, and the variance of the initial flow model error, σ^2_{QINI}.

The results for the flow volumes are very similar to those for the flow peaks (not shown).

6.5 Extrapolation to high return periods and future research

The uncertainty and sensitivity analyses have highlighted the limitations of the method that has been developed when applying it to estimate flood frequencies in the range of return periods exceeding, say, 50 to 100 years. The uncertainties mainly arise because of sensitivity to the rainfall model parameters, both for the duration and the intensity, and possibly also sensitivity to the rainfall model structure. Future studies may therefore include an exploration of a link between the

structure/parameters of the rainfall model used in this study, and the rainfall estimation models derived from regional studies using longer rainfall records pooled from a number of observation sites. It may also be useful to explore how the parameter values vary with rainfall/flood event magnitude. This could be done by using a smaller number of events per year than currently used, say, on average 2-3 events per year rather than 10. However, because the events are selected based on the rainfall, extra care would need to be taken to ensure that the full ranges of the distributions of the other variables are represented in the set of selected events. In particular, the sensitivity analysis showed that, for some of the catchments, the parameters controlling the main non-seasonal components in the SMD and initial flow models are poorly determined by the available datasets. Therefore, further limiting the dataset to fewer events would not be helpful, although longer series may mitigate the reduction in the number of events per year.

7 SUMMARY AND CONCLUSIONS

A joint probability method for flood estimation has been developed using an event-based rainfall-runoff model within a Monte Carlo simulation framework. This allows the input variables (rainfall, soil moisture, initial flow) to take on values across the full range of their distributions, thus avoiding the potential bias that can occur when a single set of pre-determined input values are used. In contrast to the continuous simulation method, a full weather generator is not needed, but a simpler model for extreme rainfall only can be used.

The overall approach is to model a string of individual hydrological events through consideration of the major flood producing mechanisms and their mutual dependences. In particular, it uses a procedure which improves on previous work on event-based simulation, by modelling the dependence between the conditions at the end of an event and those at the start of the next.

When deriving the input variables to the rainfall-runoff model, seasonality needs to be taken into account. However, this alone would not account for all the variability in the SMD and initial flow. The SMD depends also on the time elapsed since the previous event, and the initial flow depends

strongly on the SMD at the start of the event. For each event, the initial flow is modelled in a way which takes account of dependence on SMD both at the end of the previous event and at the beginning of the current event, which was found to be particularly important for the more permeable catchments. This effect was achieved by incorporating the residual from the SMD model as a component of the initial flow model.

Dependence between rainfall intensity and duration (for the same event) was removed by using transformed variables. The transformed duration was modelled using a gamma distribution and the transformed intensity using an exponential distribution.

The flood frequency curves derived using the new method do not replicate the upper bound suggested by the GEV fitted to the observed annual maxima. However, the observed series are short, and may not reflect the true distribution.

Some of the input variables are more sensitive to sampling variability than others. A bootstrap method showed that, for most catchments, the intercepts, seasonal components and model errors are generally well estimated. However, for all catchments the uncertainties in the parameters of the distributions of the time between events (IEAT), and of the rainfall descriptors, have the largest effects on the final results for the flood frequency curve. The parameters controlling the main non-seasonal components in the SMD and initial flow models are sensitive to sampling variability mainly for the two more permeable catchments.

The methodology presented here could be generalised to ungauged basins, provided that the parameters for the models of the input variables (rainfall, soil moisture, initial flow) can be derived from catchment and climate descriptors. An aim has therefore been to keep the number of parameters for these models as low as possible.

Acknowledgements The work undertaken for this project was funded by the UK Natural Environment Research Council through the thematic programme Flood Risk from Extreme Events,

project number NE/F0010037/1. Hourly rainfall and river flow data were provided by the Environment Agency. Daily rainfall data and monthly MORECS potential evaporation data were provided by the UK Met Office. Catchment average hourly rainfalls were derived from the hourly and daily rainfall data by Lamb and Gannon (1996) and Crooks *et al.* (2002). The use of these data is gratefully acknowledged. The authors would like to thank Erwin Weinmann and an anonymous referee for their constructive comments on an earlier version of this paper.

APPENDIX

Bootstrapping, and conditional and unconditional variances

For the sensitivity analysis, a total of 999 new data series were derived by resampling the original data set in blocks of one year, and the 19 model parameters were estimated for these new series. A block bootstrap was necessary because of the seasonality of the series. Confidence intervals were then calculated in three ways. Firstly, the 95% confidence interval for each parameter was estimated by taking the 2.5% and 97.5% points of the empirical distribution comprising the 999 parameter estimates from the bootstrap resamples. Secondly, confidence intervals were estimated using the conditional and unconditional variance, as described below. It is the sensitivities based these latter two estimates that are presented in Fig. 14.

The 19 by 19 covariance matrix was estimated from the 999 sets of 19 parameters. The unconditional (or marginal) variance for the parameter i is the [i, i] element of the covariance matrix. Denote this [i, i] element a, and let $\mathbf{B}$ be the ith column of the covariance matrix, but without element [i, i]. Let $\mathbf{C}$ be the covariance matrix without the ith column and the ith row. Then the conditional variance for the parameter i is: $a - \mathbf{B}'\mathbf{C}^{-1}\mathbf{B}$, where $\mathbf{B}'$ is the transpose of $\mathbf{B}$.

The 95% confidence interval, CI, for each parameter, i, can be calculated using the conditional or unconditional variance by

$$CI = \mu \pm 1.96\sqrt{\mathrm{Var}}$$

where μ is the central estimate of the parameter based on the original observed data set, Var is either the conditional or unconditional variance, and 1.96 is the quantile of the standard normal distribution corresponding to the 97.5% point.

The unconditional variance is similar in size to the variance obtained by estimating the confidence interval directly from the 2.5% and 97.5% points of the empirical distribution of parameter values from the block bootstraps, as described at the beginning of this Appendix. However, this is only true when the empirical distribution is symmetrical. When it is not, then both the conditional and the unconditional variances are only indicative. The 19 parameters generally have reasonably symmetrical distributions, apart from θ_4, whose distribution is skewed for the Kennet and the Taf.

REFERENCES

Arnaud, P. and Lavabre, J., 2002. Coupled rainfall model and discharge model for flood frequency estimation. *Water Resources Research*, 38 (6), 11-1 – 11-10.

Aronica, G.T. and Candela, A., 2007. Derivation of flood frequency curves in poorly gauged Mediterranean catchments using a simple stochastic hydrological rainfall-runoff model. *Journal of Hydrology*, 347 (1-2), 132-142.

Balistrocchi, M. and Bacchi, B., 2011. Modelling the statistical dependence of rainfall event variables through copula functions. *Hydrology and Earch System Sciences*, 15, 1959-1977.

Calver, A., Crooks, S., Jones, D., Kay, A., Kjeldsen, T. and Reynard, N., 2005. National river catchment flood frequency method using continuous simulation, Volume 1. Final report to Defra/EA for project FD2106. Centre for Ecology & Hydrology, Wallingford, UK, March 2005.

CEH, 2007. The revitalised FSR/FEH rainfall-runoff method – a user handbook. Flood Estimation Handbook Supplementary Report No. 1, Centre for Ecology & Hydrology, Wallingford, UK. (www.ceh.ac.uk/refh)

CEH, 2009. FEH CD-ROM 3, Centre for Ecology & Hydrology, Wallingford, UK.

Crooks, S., Kay, A.L. and Calver, A., 2002. Model testing in the light of extended data series. Milestone 2 report to Defra/EA, project FD2106. CEH Wallingford, UK, November 2002.

Eagleson, P.S., 1972. Dynamics of flood frequency. *Water Resources Research*, 8 (4), 878-898.

IOEA, 2001. Australian Rainfall and Runoff: A Guide to Flood Estimation , Vol. 1, Editor-in-chief D.H. Pilgrim. Institution of Engineers, Australia, Barton, ACT, Australia.

Kay, A.L., Jones, D.A., Crooks, S.M., Kjeldsen, T.R. and Fung, C.F., 2007. An investigation of site-similarity approaches to generalisation of a rainfall-runoff model. *Hydrology and Earth System Sciences*, 11 (1), 500-515.

Kjeldsen, T.R., Svensson, C. and Jones, D.A., 2010. A joint probability approach to flood frequency estimation using Monte Carlo simulation. In Proc. BHS Third International Symposium, Managing Consequences of a Changing Global Environment, Newcastle 19-23 July 2010, 5 pp.

Kurothe, R.S., Goel, N.K. and Mathur, B.S., 1997. Derived flood frequency distribution for negatively correlated rainfall intensity and duration. *Water Resources Research*, 33 (9), 2103-2107.

Kwon, H.-H., Moon, Y.-I. and Khalil, A.F., 2007. Nonparametric Monte Carlo simulation for flood frequency curve derivation: an application to a Korean watershed. *Journal of the American Water Resources Association*, 43 (5), 1316-1328.

Lamb, R., 1999. Calibration of a conceptual rainfall-runoff model for flood frequency estimation by continuous simulation. *Water Resources Research*, 35 (10), 3103-3114, doi:10.1029/1999WR900119.

Lamb, R., 2005. Rainfall-runoff modelling for flood frequency estimation, in: Anderson, M.G. and McDonnell, J.J. (Eds), Encyclopedia of Hydrological Sciences, Vol. 3. Wiley, Chichester, UK, pp. 1923-1953.

Lamb, R. and Gannon, B., 1996. The establishment of a database of hourly catchment average rainfall and flow for flood frequency estimation by continuous simulation. Milestone report to MAFF, project FD0404. Institute of Hydrology, Wallingford, UK, December 1996.

Loukas, A., 2002. Flood frequency estimation by a derived distribution procedure. *Journal of Hydrology*, 255 (1-4), 69-89.

Marsh, T.J. and Hannaford, J. (Eds.), 2008. UK Hydrometric Register. Hydrological data UK series. Centre for Ecology & Hydrology, Wallingford, UK.

Mayes, J. and Wheeler, D., 1997. The anatomy of regional climates in the British Isles, in: Wheeler, D. and Mayes, J. (Eds.), Regional Climates of the British Isles. Routledge, London, UK, pp. 9-44.

Moore, R.J., 2007. The PDM rainfall-runoff model. *Hydrology and Earth System Sciences* 11 (1), 483-499, doi:10.5194/hess-11-483-2007.

Rahman, A., Weinmann, P.E., Hoang, T.M.T. and Laurenson, E.M., 2002. Monte Carlo simulation of flood frequency curves from rainfall. *Journal of Hydrology*, 256 (3-4), 196-210.

Roberts, N.M., Cole, S.J., Forbes, R.M., Moore, R.J. and Boswell, D., 2009. Use of high-resolution NWP rainfall and river flow forecasts for advance warning of the Carlisle flood, north-west England. *Meteorological Applications*, 16 (1), doi: 23-44. 10.1002/met.94.

SCS, 1986. Urban Hydrology for Small Watersheds. Tech. Release 55, 2nd Edition. Soil Conservation Service, Washington, D.C., USA.

Sivapalan, M., Blöschl, G., Merz, R. and Gutknecht, D., 2005. Linking flood frequency to long-term water balance: incorporating effects of seasonality. *Water Resources Research*, 41, W06012, doi:10.1029/2004WR003439.

Swain, R.E., Schaefer, M.G. and Barker, B.L., 1998. Stochastic modelling of extreme floods. In Proc. of the 18th annual USCOLD lecture series on Managing the risks of dam project development, safety and operation. 10-14 August 1998, Buffalo, NY, 559-574.

Zegpi, M. and Fernández, B., 2010. Hydrological model for urban catchments – analytical development using copulas and numerical solution. *Hydrological Sciences Journal*, 55 (7), 1123-1136.

Table 1 River catchment characteristics and periods of data record.

Catchment name	Catchment number[1]	Area[1] (km^2)	Base flow index[1] (-)	Standard average annual rainfall 1961-1990[2] (mm)	Annual average potential evaporation[3] (mm)	Time to peak[3] (h)	2-year return period 1-h areal rainfall[4] (mm/h)	Period of record
Blyth at Hartford Bridge	22006	269	0.35	696	558	13	8.3	1985-2001
Bumpstead Brook at Broad Green	36010	28.3	0.23	589	629	5	11.9	1984-2001
Kennet at Marlborough	39037	142	0.94	772	616	24	10.1	1985-2001
Taf at Clog-y-Fran	60003	217	0.56	1420	607	7	10.1	1985-2001

Notes:

[1]From Marsh and Hannaford (2008). The base flow index is a measure of the proportion of the river runoff that derives from stored sources. The more permeable the rock, drift and soil material of a catchment the higher the base flow index.

[2] From the National River Flow Archive, Centre for Ecology and Hydrology, Wallingford, UK. Data source: Met Office.

[3]Calculated from the project data for the period of record.

[4]From the Flood Estimation Handbook rainfall depth-duration-frequency model (Faulkner, 1999) as implemented on the FEH CD-ROM 3 (CEH, 2009).

Figure captions

Fig. 1 Schematic diagram showing the input variables to the rainfall-runoff model and the output event flow hydrograph.

Fig. 2 Simulated string of events. The variables used for the flow simulation include rainfall, soil moisture deficit at the onset of the event (*SMD*) and at the end of the previous event (*SMD**), the initial flow (*QINI*), and the inter-event arrival time (*IEAT*).

Fig. 3 Location of the four study catchments.

Fig. 4 Scatter plots in log space of the event-based input variables, and other diagnostic variables, for the Taf at Clog-y-Fran. Events occurring in summer are plotted in red and events occurring in winter are plotted in blue.

Fig. 5 Rank auto-correlations for the event-based input variables, and other diagnostic variables, for the Taf at Clog-y-Fran. The transformed rainfall intensity is calculated as in equation (1).

Fig. 6 Observed inter-event arrival time (IEAT), and fitted exponential distributions (thick black line), for winter and summer for the Taf at Clog-y-Fran. The histogram bars outlined by thin black lines show the observed annual IEAT.

Fig. 7 Scatter plot, histograms and fitted marginal distributions for shifted rainfall duration, D' , and transformed intensity, I' , for summer (red) and winter (blue) observed rainfall events, for the Taf at Clog-y-Fran. Monte Carlo simulated events are shown as grey crosses.

Fig. 8 Concept figure of the typical seasonal variation in SMD (black line), and the modelled SMDs influenced by the previous event (dashed red line). SMD_1^* is the

SMD at the end of a PDM simulated event occurring in spring. The figure shows how the SMD modelled according to equation (2) then gradually tails off and becomes more similar to the typical seasonal curve (equation (3)) as time progresses. There is an abrupt change as a second event is simulated by the PDM in the autumn, ending with SMD_2^*. The modelled SMD then slowly approaches the typical curve again.

Fig. 9 "Observed" SMD (red circles) at start of the event, predicted typical $\mu(f)$ curve (orange) and values of SMD generated for 100 Monte Carlo simulations of 17 years each (grey crosses) versus the time of year, for the Taf at Clog-y-Fran.

Fig. 10 Plots of residuals from model fitted to the SMD at the start of the event (equation (2)) for the Taf at Clog-y-Fran. Residuals versus time of year (top left), residuals versus IEAT (top right), q-q plot for Normal distribution (bottom left) and autocorrelation function (bottom right).

Fig. 11 Uncertainty due to limited record length for the Taf at Clog-y-Fran: 100 Monte Carlo simulations each 16 (water) years long (grey lines), and 10 simulations each 2000 years long (red lines), all starting from different random seeds and plotted using Gringorten's plotting position. Observed (circles) and continuously simulated (crosses) annual maximum peak flows are also shown, together with a GEV distribution (thick black line) fitted to the observed annual maxima and the 95% confidence intervals around the fitted GEV frequency curve (blue dashed lines).

Fig. 12 Uncertainty in the flood frequency curve due to sampling variability (all four catchments). The full black line is the GEV distribution fitted to observed annual maxima (black circles), with the 95% confidence interval around the fitted GEV in dashed blue lines. The full red line is the "median" flood frequency curve from 199 2000-year long model runs (solid red line), with 95% confidence intervals based on a block bootstrap method (dotted red lines).

Fig. 13 Uncertainty in the frequency curve for flood volumes due to sampling variability (all four catchments). Notations as for Fig. 12.

Fig. 14 Relative sensitivity of the flood frequency curve to sampling variability in the individual parameters. The black line is the GEV fitted to annual maxima (black dots) from a model run based on the parameters estimated for the observed data set. This is surrounded by pairs of frequency curves where, for each diagram, the target parameter has been replaced by the value corresponding to the upper or lower limit of the 95% confidence interval of the parameter. The confidence interval is estimated using the conditional (dashed blue lines), or unconditional (dash-dotted red lines), variance.