

A depth–duration–frequency analysis for short-duration rainfall events in England and Wales

Ilaria Prosdocimi, Elizabeth J. Stewart and Gianni Vesuviano

ABSTRACT

This study presents a depth–duration–frequency (DDF) model, which is applied to the annual maxima of sub-hourly rainfall totals of selected stations in England and Wales. The proposed DDF model follows from the standard assumption that the block maxima are generalised extreme value (GEV) distributed. The model structure is based on empirical features of the observed data and the assumption that, for each site, the distribution of the rainfall maxima of all durations can be characterised by common lower bound and skewness parameters. Some basic relationships between the location and scale parameters of the GEV distributions are enforced to ensure that frequency estimates for different durations are consistent. The derived DDF curves give a good fit to the observed data. The rainfall depths estimated by the proposed model are then compared with the standard DDF models used in the United Kingdom. The proposed model performs well for the shorter return periods for which reliable estimates of the rainfall frequency can be obtained from the observed data, while the standard methods show more variable results. Although the standard methods used no or little sub-hourly data in their calibration, they give fairly reliable estimates for the estimated rainfall depths overall.

Key words | depth–duration–frequency, intensity–duration–frequency, short-duration rainfall, statistical modelling

Ilaria Prosdocimi (corresponding author)
Department of Mathematical Sciences,
University of Bath,
Claverton Down,
Bath BA2 7AY,
UK
E-mail: prosdocimi.ilaria@gmail.com

Elizabeth J. Stewart
Gianni Vesuviano
Centre for Ecology & Hydrology,
Maclean Building,
Crowmarsh Gifford, Wallingford,
Oxfordshire OX10 8BB,
UK

INTRODUCTION

Estimates of the magnitude of rainfall events of a given duration with an expected annual exceedance probability p , are an important component of current methods of flood frequency estimation, used in the design and assessment of flood defence schemes, bridges and reservoir spillways, as well as urban drainage systems. Rainfall frequency estimates are also a key input to mapping studies of the risk of surface water flooding. The estimates can be obtained from depth–duration–frequency (DDF) models, in which the relationship between the rainfall depth, event duration and event rarity is integrated in a unique framework. In a DDF

model, it is required that frequency curves for different durations do not cross, meaning that the rainfall depth that is exceeded with probability p should increase monotonically with increasing event duration. The probability p is typically expressed as a return period T , with $p = 1/T$, as events larger than those corresponding to the quantile that is expected to be exceeded with probability p should happen, on average, every T years.

DDF models, which are often referred to as intensity–duration–frequency models, can then serve two purposes: to estimate the rainfall depth of a hypothetical event with a given duration and rarity, and to assess the rarity of a storm event with known rainfall depth and duration. Svensson & Jones (2010) give an overview of different DDF models used in several countries, showing the large array

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

of possible approaches to rainfall frequency estimation. Many of the countries included in the review use some form of index rainfall approach combined with regional estimation of growth curves for different durations, although some countries were reported to use the linear regression approach. The idea behind the latter approach is to fit a statistical distribution separately to the single series of block maxima of different accumulation periods and then to fit regression models across the different durations or frequencies, so that increasing rainfall depths are estimated for increasing durations given a certain frequency. See [Koutsoyiannis et al. \(1998\)](#) for a discussion of the mathematical formulation of the relationship between the duration and frequency of rainfall events, and a general discussion of DDF modelling. Although the relationship between rainfall depths and frequencies has been studied for several decades, there is still much interest in identifying methods to derive DDF curves (e.g., [Overeem et al. 2008](#)) and in the actual derivation of DDF curves to be used at different sites of interest (e.g., [Jiang & Tung 2013](#)).

One interesting finding of the review in [Svensson & Jones \(2010\)](#) is that, in several countries, different models are used depending on the duration and rarity of the rainfall events of interest. The need for different models for different durations and frequencies stems from the difficulty of developing models that can provide reliable results across several rainfall durations and frequencies. One country where several DDF models are currently in use is the UK: the main models are presented below and are the main focus of this study.

In the UK, the most widely used DDF models are those presented in volume II of the *Flood Studies Report* (FSR, [Natural Environment Research Council 1975](#)) and in volume 2 of the *Flood Estimation Handbook* (FEH99, [Faulkner 1999](#)), which mostly superseded the FSR methods. Recently, a new model (FEH13, [Stewart et al. 2013](#)) has been developed, with the specific aim of overcoming the issues encountered when the original FEH99 model is used to estimate rare events. Estimates from the FEH13 model have only been available to practitioners since November 2015, and have therefore not yet been widely used in practice. Furthermore, the performance of the FEH13 model for short duration events (i.e., under 1 hour) is still being assessed, since most of the model evaluation

focused on the estimation of the frequency of long-duration events. Considering that the FEH13 model aimed to improve rainfall frequency estimates for rare events with durations longer than 1 hour, it is not yet clear how it will perform for the more frequent events of very short duration which are of interest in this study.

The FSR and FEH99 DDF models are based on an index-rainfall approach and were developed with the scope of providing nationwide rainfall frequency estimates. The FEH99 method was calibrated on a larger network of stations with longer records than the FSR method and, unlike the FSR method, incorporated a spatial model in which data from nearby stations were used for rainfall frequency estimation at a given location. On the other hand, the FEH99 method was calibrated using data with an accumulation period of at least 1 hour while, in the development of the FSR method, some data with an accumulation period of 1 minute were also used. Compared to the FSR method, the FEH99 method has been found to give much larger estimates of rainfall depth for the very long return periods required for reservoir safety assessment ([Babtie Group in association with CEH Wallingford & Rodney Bridle Ltd 2000](#); [MacDonald & Scott 2001](#)). As a result, the FSR and FEH99 methods are both still used, but for different cases that depend on the duration and rarity of the design event to be estimated ([ICE 2015](#)). As [Svensson & Jones \(2010\)](#) report, the FSR method can be used to estimate return periods of rainfall events with accumulation periods between 1 minute and 25 days and return periods longer than 1,000 years, and is recommended for the estimation of rainfall depths associated with return periods up to 10,000 years. The FEH99 method provides estimates of rainfall accumulations between 1 hour and 8 days, with return periods shorter than 1,000 years and, although rainfall frequencies up to return periods of 10,000 years can technically be estimated, their use is not recommended. The newly developed FEH13 might replace the FSR and the FEH99 as the recommended model to use to estimate the magnitude of very rare events, but the official guidelines have not yet been amended. The FEH99 method can also be extended to estimate the frequency of rainfall events with accumulation periods shorter than 1 hour, although, as no sub-hourly data were used in the calibration of the method, extrapolation to durations below 30 minutes is

strongly discouraged. The coexistence of the FEH99 and FSR methods results in uncertainty when estimates are needed for sub-hourly rainfall events. These cases go beyond the range of reliable estimates for the FSR, a relatively old model that was calibrated on fairly short records with very limited sub-hourly data, and beyond the intended use of the FEH99, a more complex and structured model that was calibrated using a dense network of stations but no sub-hourly data at all.

Small catchments (i.e., smaller than 25 km²) and plot-sized areas are expected to be particularly vulnerable to short, intense cloudbursts, due to their short response times. As Faulkner *et al.* (2012) emphasise, reliable estimates of sub-hourly design rainfalls are therefore needed to allow credible flow and hydrograph estimates for the smallest catchments using rainfall-runoff techniques. The suggestions in Faulkner *et al.* (2012) motivated the second phase of the Environment Agency's (EA) *Estimating Flood Peaks and Hydrographs for Small Catchments* project. The project aims to improve the estimation of flood frequencies in small catchments and encompasses, among other things, an assessment of the most appropriate methods to estimate the frequency of very short duration rainfall, which this study is concerned with. A novel at-site DDF modelling strategy is discussed and an application of the proposed model is presented using data series available at selected sites that give a reasonable geographical coverage of England and Wales, for which relatively long records of sub-hourly rainfall are available. The proposed model does not follow the traditional approaches and uses instead the data across all durations to fit a unique model. Rainfall frequency curves estimated with the proposed method are compared to those estimated with the FSR and FEH99 DDF models, and to empirical return level estimates.

The stations and datasets used in the study are introduced in the next section. Subsequently, a unified generalised extreme value (GEV) model is proposed and its performance for the stations under study is discussed. The performance of the unified GEV, FSR and FEH99 models for short-duration rainfall frequency estimation are compared in the section *Comparisons of the unified GEV results to current methods*. The final section of the paper contains the conclusions and final remarks.

DATA

From the large number of tipping bucket rain gauges managed by the EA and Natural Resources Wales (NRW) and providing sub-hourly rainfall data, a subset with sufficiently long records was identified that could allow for good spatial coverage of the area. Sub-hourly data for the rainfall stations are available as time of tip (ToT) at some sites series and as aggregated 15-minute accumulation series at other sites. In the selection of stations to be included in this scoping study, priority was given to those for which ToT data were available, to allow very short durations to be investigated. It appears that long ToT series are more readily available in some regions (the English Midlands and Wales), hence the final subset of stations included in the study is a compromise between the competing needs of having long series and maintaining a good coverage of England and Wales (E&W). In particular, the sites were chosen to be at least 35 km apart. The final selected stations are shown in Figure 1. The shortest series in the dataset is 15 years long; the longest two are each 46 years long. A total of nine ToT series and ten 15-minute series are included in the study dataset. The analysis was performed on the annual and seasonal maxima of the different accumulations, with two six-month seasons included in the study. The final dataset was compiled from the ToT and 15-minute series, following two slightly different workflows as outlined below.

- From the original ToT data, 1-minute accumulation series were composed. From these, 1-minute monthly maxima were extracted and, by cumulating successive data-points, monthly maxima for 2-, 5-, 10-, 15-, 30-, 45-, 60-, 90- and 120-minute accumulations were extracted.
- From the 15-minute accumulation data, monthly maxima for the 15-, 30-, 45-, 60-, 90- and 120-minute accumulations were extracted.

For all series, a month was considered complete if at least 75% of the data in the month were non-missing. Finally, the annual and seasonal maxima series were constructed from the monthly maxima series. A year or season was considered complete if no more than one monthly record within that year or season was incomplete. Approximately 89% of the station-seasons have at least

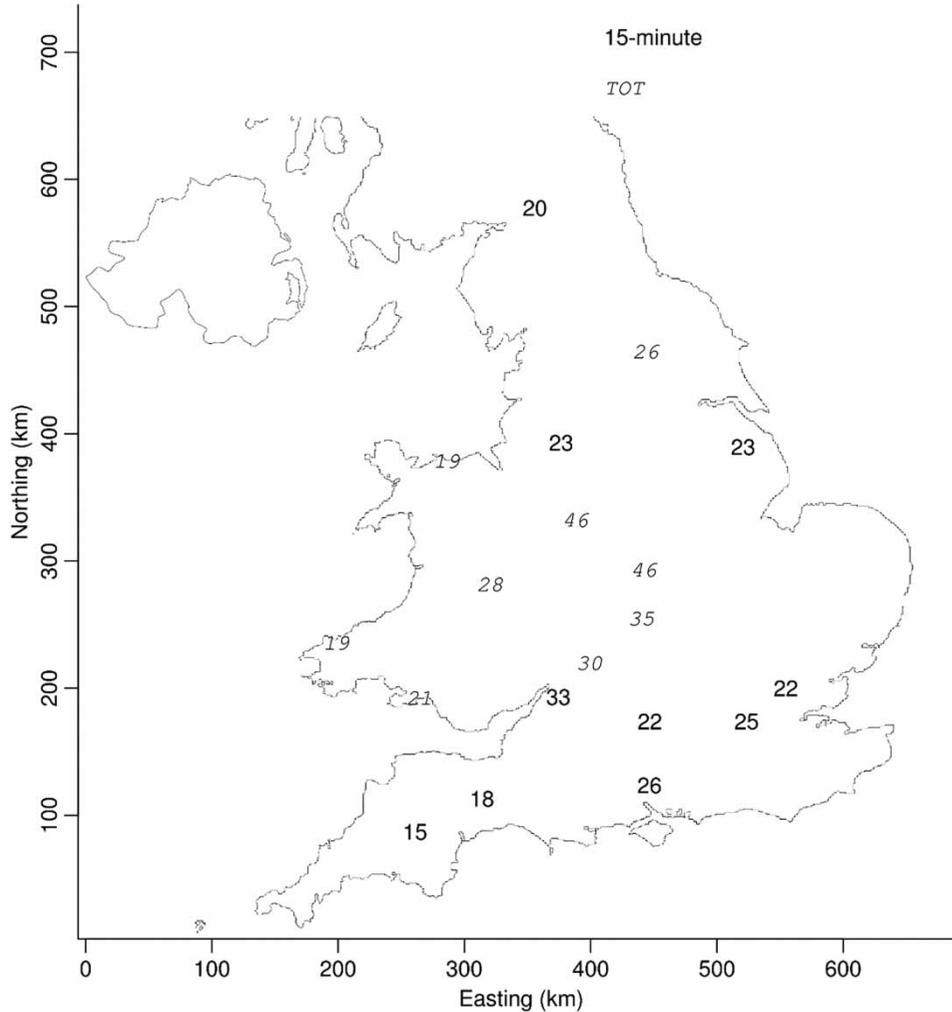


Figure 1 | Location of the 19 stations included in the study. The record length of the annual maxima series is indicated in the location of each station; numbers in italics indicate ToT stations, numbers in roman indicate 15-minute stations.

99% of valid data points, 98% of the station-seasons have at least 90% of valid data points and in just one instance is the percentage of valid data points in a season lower than 80% (summer rainfall series of 1995 at Victoria Park, which has a total of 79.3% valid data points). Overall, for all stations, for the series across all years and seasons, more than 99% of the total number of data points are recorded as valid, giving reasonable confidence in the quality of the available data and confidence that the maxima were captured. Annual maxima were extracted as the maximum single value recorded in each calendar year. Summer maxima were extracted as the maximum value recorded in the months from May to October inclusive. Winter maxima were

extracted as the maximum value recorded in the months from November to April inclusive.

The availability of the raw ToT information for the tipping bucket stations allows for the extraction of series at a 1-minute resolution and additionally at coarser or even finer resolutions. However, the level of precision that can be reached in high resolution series depends greatly on the tip volume of the instrument, a property that might change slightly in time (e.g., due to sediment collecting in the bucket) or more significantly over time (e.g., if the specific gauge used at a station is replaced by a different model). Furthermore, the tip volume might be different at different stations, thus creating inconsistencies in the precision

across different stations. The discrete nature of the tipping bucket measurements is also the underlying reason why, in a number of months, the recorded 1-minute and 2-minute maxima have the same value, and why several annual and seasonal maxima are identical across a number of years. These issues are more common in the earlier years of the record, during which time the data were measured at a coarser resolution. The issues connected to systematic errors in tipping buckets are known (Molini *et al.* 2005, and references therein). In particular, lower intensities tend to be overestimated and higher intensities tend to be underestimated. Methods to quantify the systematic error of each station are beyond the scope of this study, and the data extracted from the original series are used in all subsequent analysis without further adjustment. The issues connected with the original data series should, nevertheless, be acknowledged as they can have an impact on the estimation procedures discussed in the section *Results for the at-site analysis* and in the comparisons discussed in the section *Comparisons of the unified GEV results to current method*.

Due to differences in the underlying data collection methods, the series of maxima extracted from the ToT and the 15-minute series do not provide the same information for accumulations of 15 minutes or greater. The ToT maxima are computed using a sliding window, so the 15-minute annual maximum value (for example) corresponds to the actual largest amount of rainfall recorded in any 15-minute interval in the year. However, the maximum obtained from the 15-minute records instead corresponds to the maximum amount recorded in one predefined 15-minute interval, which is likely to be lower than the actual maximum amount of rainfall that could have been recorded in a 15-minute interval without a fixed start time. The true maximum rainfall is most likely to be under-recorded when its duration is the same as the fixed-duration recording unit, as the rainfall event is very unlikely to align neatly with the station clock. However, when longer durations are

considered, the alignment between the rainfall event and the station clock is less important, as the depths of rainfall at the tail ends of the storm, which are difficult to capture exactly, become less and less important to the storm depth as a whole.

To adjust the maxima extracted from the 15-minute stations so that they are closer to the higher values that would be attained using sliding windows, correction factors were introduced. For each ToT record, fixed-period (15-minute) annual and seasonal maxima were extracted for durations of 15, 30, 45, 60, 90 and 120 minutes. These series correspond to the maxima that would be obtained if the data for the ToT stations were stored as 15-minute series (fixed window) rather than ToT series (sliding window). The average ratio between the sliding window maxima and the fixed window maxima at each duration, shown in Table 1, is used as a sliding window correction factor for that duration. In the rest of this work, the maxima extracted from the 15-minute series are multiplied by the appropriate correction factor to give estimates of the equivalent sliding window maxima. Due to the different ranges of time resolution present in the two different data sources, two separate analyses are carried out: one which uses only the series extracted from the ToT stations and covers the range of durations from 1 to 120 minutes; and one in which data from all stations are included, covering the range of durations from 15 to 120 minutes.

THE UNIFIED GEV DDF MODEL

The FSR, FEH99 and FEH13 DDF models build on a large set of available gauges and allow the estimation of frequency curves for a number of durations across the whole UK. In particular, the FEH99 and the FEH13 have complex spatial model components so that estimates for rainfall frequencies at one point are built incorporating information from nearby

Table 1 | The correction factors applied to the maxima obtained from 15-minute series, for different seasons and event durations

	15 minutes	30 minutes	45 minutes	60 minutes	90 minutes	120 minutes
Annual	1.15	1.05	1.03	1.02	1.02	1.01
Winter	1.14	1.05	1.04	1.03	1.02	1.02
Summer	1.15	1.06	1.03	1.02	1.02	1.01

gauges. Such complex spatial structures are unattainable with the subset of stations available in this study. Given the exploratory scope of this work, a simpler model is proposed: the model allows the estimation of a station’s DDF curves based solely on the data series available for that station; it does not have a component to include information from nearby stations.

The proposed model builds on extreme value theory (Coles 2001), assuming that block (e.g., annual or seasonal) maxima follow a GEV distribution: $X \sim GEV(\xi, \alpha, \kappa)$ where X indicates the random variable that describes rainfall block maxima and ξ, α and κ are the location, scale and skewness parameters of the GEV distribution, respectively. The cumulative distribution function of a GEV distributed random variable $X \sim GEV(\xi, \alpha, \kappa)$ is defined as:

$$F(x) = \exp\left\{-\left[\frac{1 - \kappa(x - \xi)}{\alpha}\right]^{1/\kappa}\right\} \tag{1}$$

The set on which the variable X is defined, e.g., the values that might be observed in a sample from a population with underlying distribution X , is governed by the skewness parameter as follows:

$$\begin{cases} -\infty < x \leq \xi + \frac{\alpha}{\kappa} & \text{if } \kappa > 0 \\ -\infty < x < \infty & \text{if } \kappa = 0 \\ \xi + \frac{\alpha}{\kappa} < x < \infty & \text{if } \kappa < 0 \end{cases} \tag{2}$$

The distribution is bounded for the case in which $\kappa \neq 0$, with the lower and upper bound being a linear combination of the distribution parameters. The skewness parameter therefore defines whether an upper or lower bound for the values of X exists.

The quantile function for the GEV distribution, which is used to build frequency curves, is derived as:

$$x(F) = \begin{cases} \xi + \frac{\alpha}{\kappa} [1 - (-\log F)^\kappa] & \text{if } \kappa \neq 0 \\ \xi - \alpha \log(-\log F) & \text{if } \kappa = 0 \end{cases} \tag{3}$$

where F is the non-exceedance probability, corresponding to $F = 1 - 1/T$ for the T -year event. The desired property of a DDF model is that the quantile functions for increasing durations of rainfall accumulation, D , do not cross. This means

that, denoting by $x(F, D)$ the rainfall depths of durations D associated with a certain non-exceedance probability F , for $d_0 < d_1$ one should have $x(F, d_0) < x(F, d_1)$. The proposed model uses the relationship between the GEV parameters shown in Equation (2) and stems from some empirical properties observed via visual explorations of the estimated parameters for the different durations at each station (see the section *Results for the at-site analysis*). The GEV distribution can be shown to be the asymptotic distribution of sample maxima (see Coles 2001) and has often been used as an underlying distribution in the development of DDF models (among others, Overeem et al. 2008; Jiang & Tung 2013). According to the goodness of fit test presented in Kjeldsen & Prosdocimi (2015), the GEV distribution was deemed acceptable for a large majority of the series analysed in the study. When estimating frequency curves, it is expected that no upper limit should be attainable by the rainfall values at any duration, so the skewness parameter is constrained in the proposed model to be negative. The model development is presented below only for the case in which $\kappa < 0$, although similar ideas would apply for $\kappa > 0$. It is assumed that the skewness parameter κ is constant across all durations, while the location and scale parameters are dependent on the duration D : $\xi(D)$ and $\alpha(D)$. Taking ℓ to be the lower bound of the distribution, and assuming this to be the same for all durations, the following relationship is obtained from the inequality in Equation (2):

$$\alpha(D) = (\ell - \xi(D))\kappa. \tag{4}$$

The quantile function shown in Equation (3) can then be updated to a quantile function $x_D(F)$, which depends on the event duration D via the location parameter $\xi(D)$:

$$\begin{aligned} x_D(F) &= \xi(D) + \frac{\alpha(D)}{\kappa} [1 - (-\log F)^\kappa] \\ &= \xi(D) + (\ell - \xi(D)) [1 - (-\log F)^\kappa] \\ &= \ell [1 - (-\log F)^\kappa] + \xi(D)(-\log F)^\kappa \end{aligned} \tag{5}$$

Provided that $\xi(D)$ is monotonically increasing, the function $x_D(F)$ is a monotonically increasing function of D , so that the estimated frequency curves give consistent results for increasing durations. For the case of the British rain gauges under study, the following relationship is proposed

to model the location as a function of the event duration, based on the observed properties of the location parameter for a GEV distribution fitted separately for each different duration across all stations (see Figure 2 in the next section):

$$\xi(D) = a + b D + c (1 - \exp\{-g D\}) \quad (6)$$

which is an increasing function of D provided that its first derivative is positive:

$$b + c g \exp\{-g D\} > 0 \quad (7)$$

The scale function is determined by a combination of the lower bound (ℓ), the skewness parameter (κ) and the location function ($\xi(D)$) according to Equation (4). The proposed unified GEV model then requires the estimation of a total of six parameters (a, b, c, g, ℓ, κ), a relatively parsimonious model which, given some constraints in the location function, allows for consistent frequency estimates for different durations. It is possible that an even simpler formulation could be used for Equation (6), but the suggested function originates from the methods discussed in Stewart et al. (2013) and seems to give reasonable results.

The proposed unified GEV model uses a different strategy to obtain consistent estimates for rainfall frequencies than many published works, which use approaches based on linear regression across estimates for the different durations. The unified GEV model presented in this paper instead seeks to fit a unique model to all series at once, so that all available information is used to estimate the DDF curves. The development of the model is inspired by some of the discussion in Stewart et al. (2013), on the development of the statistical framework used in the FEH13 model.

The basic novel idea behind the proposed model is to ensure that monotonic quantile functions are obtained by constraining some of the parameters of the rainfall distribution to have common properties across different durations. It is possible that for a different set of durations, or a new set of gauging stations that exhibit different properties, the assumptions of which common distributional properties are to be shared across durations might be different. Furthermore, the functional relationship between the location and the duration shown in Equation (6) might not be valid. Nevertheless, the building blocks of the proposed model could be adapted to

accommodate different data behaviours: the unified GEV is an addition to the possible modelling approaches used for at-site estimation of DDF curves.

RESULTS FOR THE AT-SITE ANALYSIS

For each station separately, the parameters of the unified GEV model (a, b, c, g, ℓ, κ) are estimated via maximum likelihood, which ensures some optimal properties for parameter estimates (Coles 2001). The unified GEV model is fitted to the data from all the ToT stations and to all the series with accumulations of at least 15 minutes, in two different fitting procedures. The location function, shown in Equation (6), and the relationship between the scale and other parameters, shown in Equation (4), are used in the two fitting procedures.

To illustrate the challenges relating to the model fitting procedure and to show some of the features of the fitted models, the location ($\xi(D)$) and scale ($\alpha(D)$) functions, together with the skewness (κ) and lower bound (ℓ) parameters, all as estimated by fitting the unified GEV model to the ToT annual maxima series, are shown in Figure 2. As a reference, the plot also shows estimates for the GEV parameters obtained by applying an L -moments fitting procedure (Hosking 1990) to the series of each duration separately for all stations. L -moment estimates are frequently used in hydrology due to their good performance when applied to relatively short series, such as the duration-specific rainfall series analysed here. The scatter of the duration-specific estimates inspired the use of an exponential function to describe the location of the GEV distribution as a function of the rainfall duration shown in Equation (6) and there is, indeed, a general agreement between the duration-specific estimates and the location functions estimated within the unified GEV model. Note that the GEV fitted to each duration separately would lead to non-consistent return curves across the different durations, unlike the unified GEV model: although it is desirable for the unified GEV parameter functions to resemble the estimates obtained for each duration separately, the differences in the estimates are needed to ensure the consistency of the estimated frequency curves. Moreover, the relatively large difference seen between the

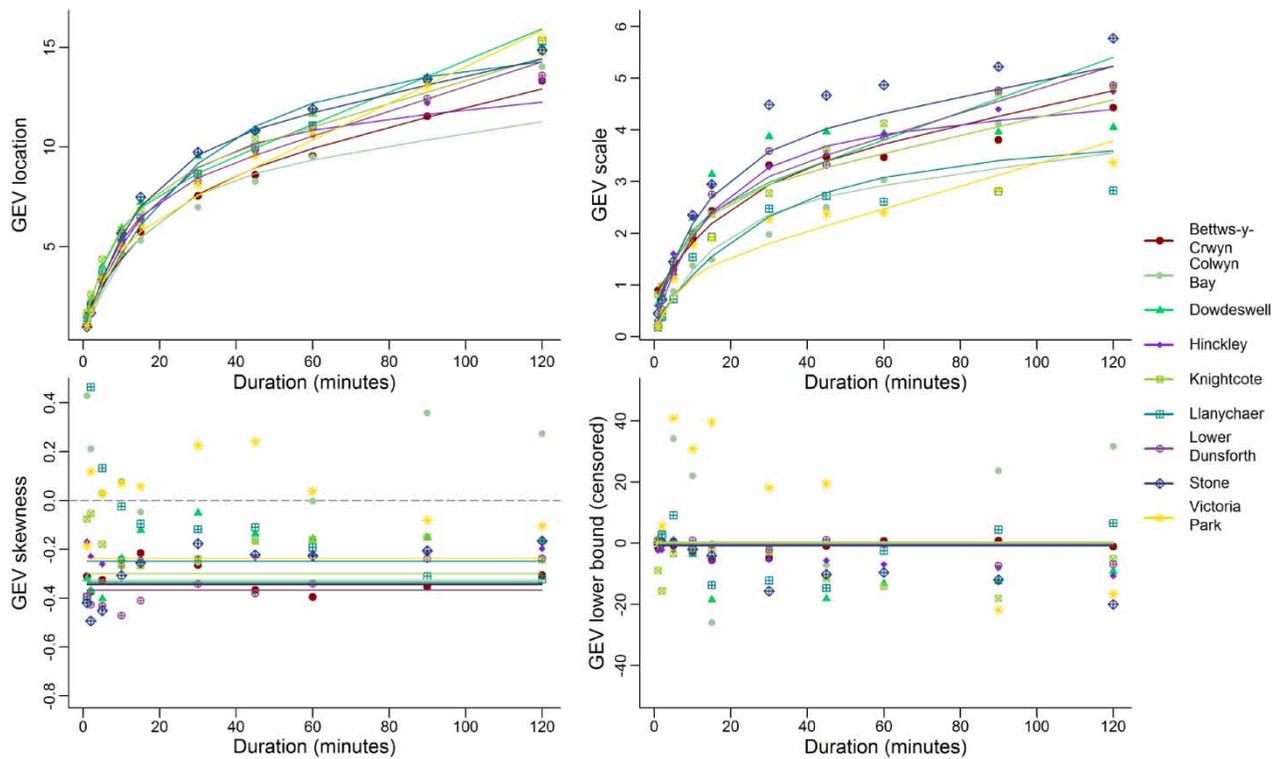


Figure 2 | Estimated parameters obtained from an L -moment estimation for each duration separately (dots) and from the proposed unified GEV distribution (lines) annual series. Colours and symbols indicate the different ToT stations. Distribution parameters in each panel are (clockwise from top left): the location, scale, lower bound and skewness. To make the figure readable, lower bound estimates below -40 are not shown.

unified GEV estimates and the separate-duration GEV parameter estimates at some stations (e.g., Victoria Park) are partially the consequence of the model structure, in which the skewness parameter, which is constrained to be negative, regulates the curvature of the scale function. For Victoria Park, for example, the raw estimate of the skewness parameter for many durations is positive or very close to zero, as shown in the lower left panel of Figure 2. The final estimated values for the unified GEV parameters maximise the overall likelihood for all durations within the constraints of the model: this could lead to large discrepancies between the estimates obtained under the unified GEV and those obtained from the GEV parameters estimated for each duration separately. The results of fitting the unified GEV to winter and summer maxima show a similar pattern.

Estimated rainfall DDF curves for the annual, winter and summer series for the ToT station at Dowdeswell are shown in Figure 3, together with the block maxima extracted from the original series plotted using Gringorten

plotting positions. The frequency curves seem to fit the data reasonably well. Due to the constraints in the model structure that ensures that the location function is monotonically increasing for increasing durations, the frequency curves computed from the formula in Equation (5) tend to fan out. A noticeable feature of the data is that the winter maxima tend to be much smaller than the summer maxima, which also appear to be the annual maxima. Results for the other ToT stations have similar properties to the ones shown in Figure 3 and are shown in Prosdocimi et al. (2016).

Figure 4 shows the estimated location and scale functions, together with the skewness and lower bound parameters of the unified GEV model, for annual data at all 19 stations, considering accumulations of 15 to 120 minutes. As in Figure 2, the original estimates for the GEV parameters obtained from an L -moment estimation procedure fitted to each duration separately are also shown. Again, the fitted location functions seem to be mostly in

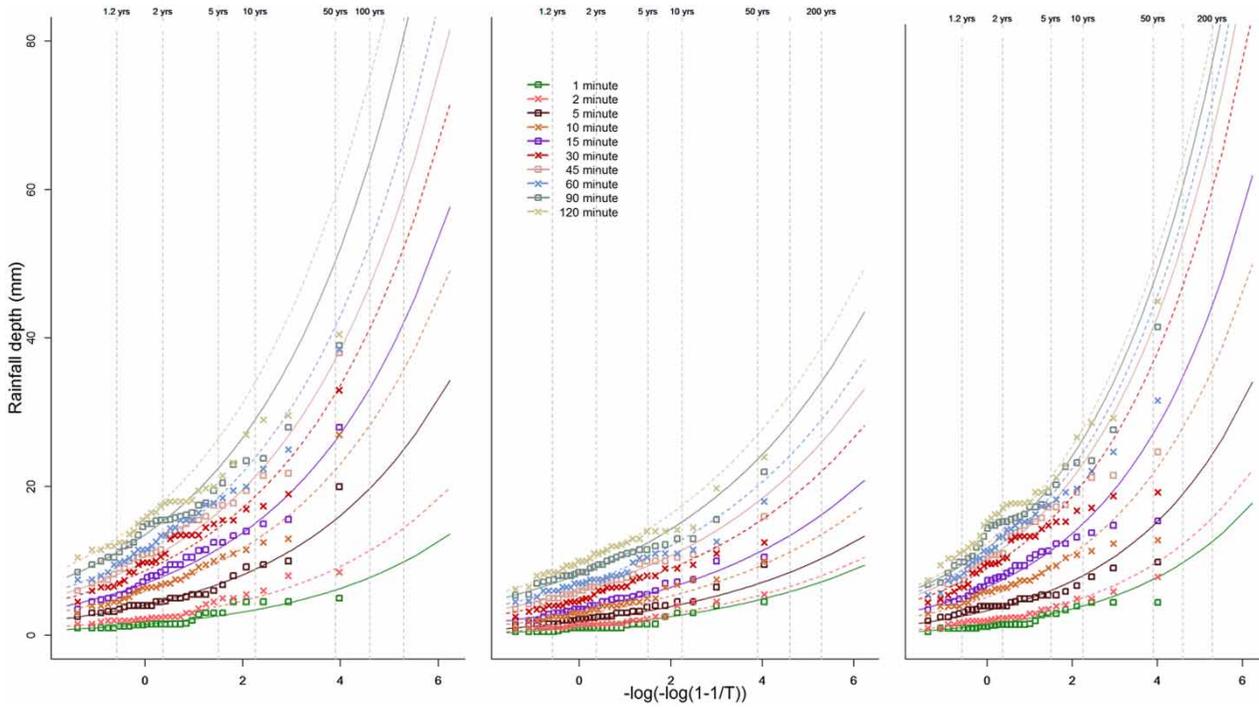


Figure 3 | Estimated frequency curves for the station at Dowdeswell for the annual (left panel), winter (central panel) and summer (right panel) series, superimposed on the Gringorten plotting positions for each duration, starting from 1 minute.

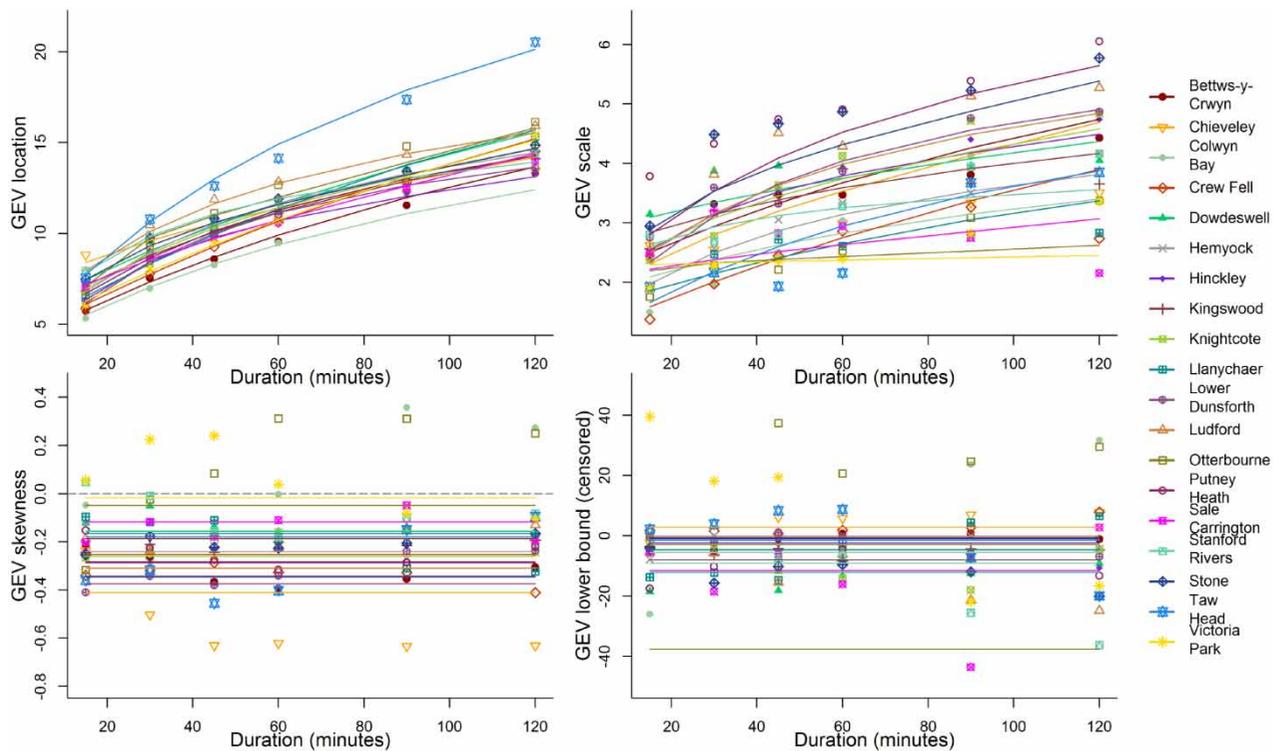


Figure 4 | Estimated parameters obtained from an *L*-moment estimation for each duration separately (dots) and from the proposed unified GEV distribution (lines) annual series. Colours and symbols indicate the different stations with series of at least 15-minute accumulations. Parameters in each panel are (clockwise from top left): the location, scale, lower bound and skewness. To make the figure readable, lower bound estimates below -40 are not shown.

agreement with the estimates obtained from the different durations, while more variability can be seen in the estimation of the scale function in the top right panel. In particular, the scale functions for Victoria Park and Otterbourne are very flat, as a result of the estimates for the skewness parameters at these stations being very close to zero. The estimated lower bounds for these two stations are also very small: -37.7 at Otterbourne and -124.6 at Victoria Park (censored in Figure 4). The fact that the skewness parameter for these stations is estimated to be very close to zero in the unified GEV model is likely to be connected to the fact that some series in these stations appear to have a finite upper bound (e.g., positive skewness) for some durations. In the unified GEV model, the skewness parameter is required to be negative and to be unique for all durations, so that the final estimate is a summary of the properties of all durations. If the behaviour of the series at a station differs across durations, the final estimates need to be a compromise between the different tendencies of each series. Nevertheless, the final fit of the estimated frequency curves compared to the annual maxima shown in Figure 5 seem to indicate that overall an acceptable fit is obtained for the series at Otterbourne. The estimated frequency

curves shown in Figure 5 have similar properties to those shown in Figure 3 – the curves have a tendency to fan out and the annual extremes appear to be mostly driven by summer rather than winter events.

Seasonal differences are not explored further in this analysis, but the estimates obtained from the different stations could be employed in the future to develop correction factors to obtain seasonal estimates from sub-hourly annual estimates, similarly to Kjeldsen et al. (2006). The unified GEV proved to be a flexible and reliable modelling approach which could give reasonable estimates across different seasons.

COMPARISONS OF THE UNIFIED GEV RESULTS TO CURRENT METHODS

The estimated depths obtained with the methods currently in use (FSR and FEH99) and the proposed unified GEV, corresponding to some pre-specified frequencies, are compared to the empirical estimates obtained from the recorded data series at each station. Since reliable estimates of very rare events cannot be obtained from the relatively short records

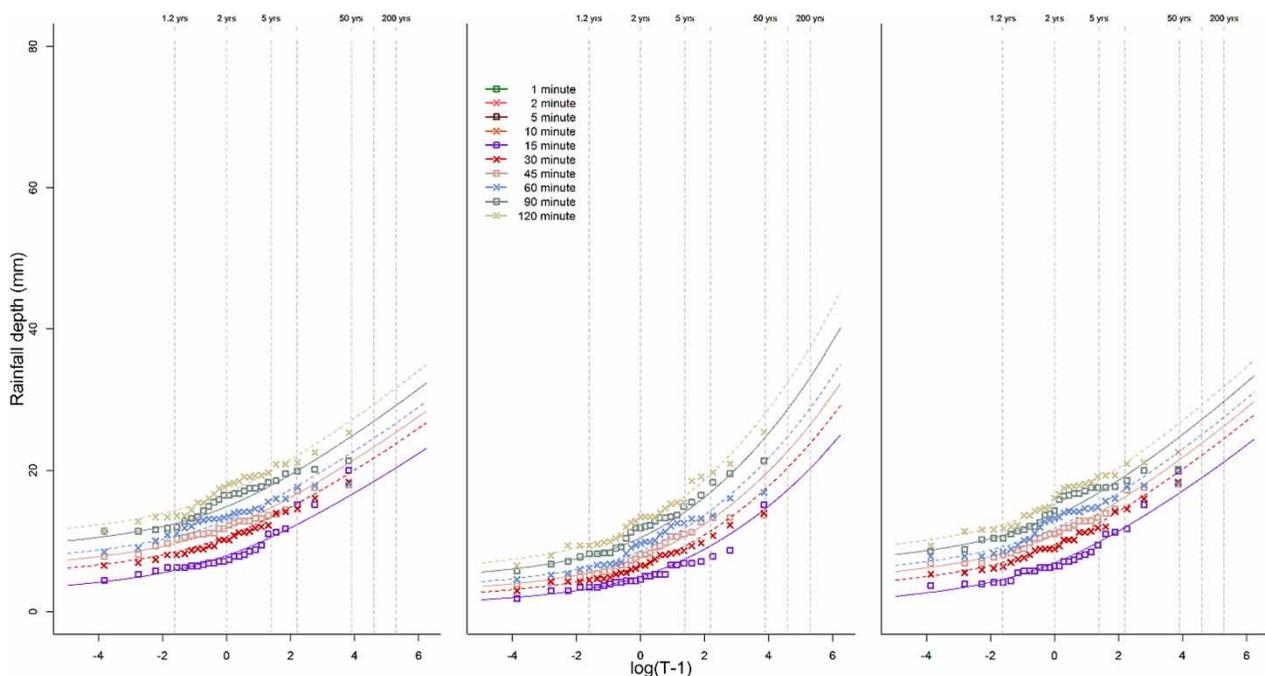


Figure 5 | Estimated frequency curves for the station at Otterbourne for the annual (left panel), winter (central panel) and summer (right panel) series, superimposed on the Gringorten plotting positions, for durations from 15 to 120 minutes.

available (the median record length for the observed series is 24 years), the comparison is limited to the 2-, 5- and 10-year return periods. The empirical estimates are obtained as the median (50th percentile), 80th percentile and 90th percentile of the recorded data series. For some series, the record length would be less than $2T$ years long when estimating the 10-year event: these empirical estimates might be less precise. The comparison is performed on every station for durations of at least 15 minutes, and the fitted unified GEV models shown in Figure 4 are used to estimate the rainfall depths.

Figures 6–8 display the relative differences between the rainfall depths, as estimated with the different methods, and the empirical quantile corresponding to the specific frequencies for the 2-, 5- and 10-year return periods, respectively. For example, the left panel of Figure 6 shows, for each station and each duration, the value $(R2_{\text{FSR}} - R2_{\text{Observed}}) / R2_{\text{Observed}}$, where $R2_{\text{FSR}}$ and $R2_{\text{Observed}}$ indicate the estimated 2-year rainfall of the given duration at a station and the empirical 2-year event estimated from the observed data, respectively. The unified GEV model is the only method directly fitted to the observed data, which explains the much better performance of that model in comparison

to the FSR and FEH99 models for the 2-year events. In particular, the FSR appears to give consistently positively biased estimates for the 2-year events (Figure 6), with lower variabilities in the error for longer durations. The FEH99 seems to perform well on average, although the results are more variable than the unified GEV. The results for the longer return periods show more variation, with the unified GEV performing slightly better in terms of the variability of the error. The unified GEV, an at-site model fitted directly to the observed data, performs quite well for most stations. Among the models currently used in the UK for rainfall frequency estimation, the FEH99 seems to give acceptable results, across all return periods, with an error variability comparable to the FSR estimate.

These comparisons are based only on empirical estimates of events with a relatively short return period, and it is not clear how the different models differ for the estimation of rare events, for which no reliable empirical estimates can be obtained from the observed series. An assessment of the accuracy of the different estimation methods for longer return periods would, in fact, require reliable information on the real frequency of short-duration rainfall events, which cannot be easily retrieved. The overall relative

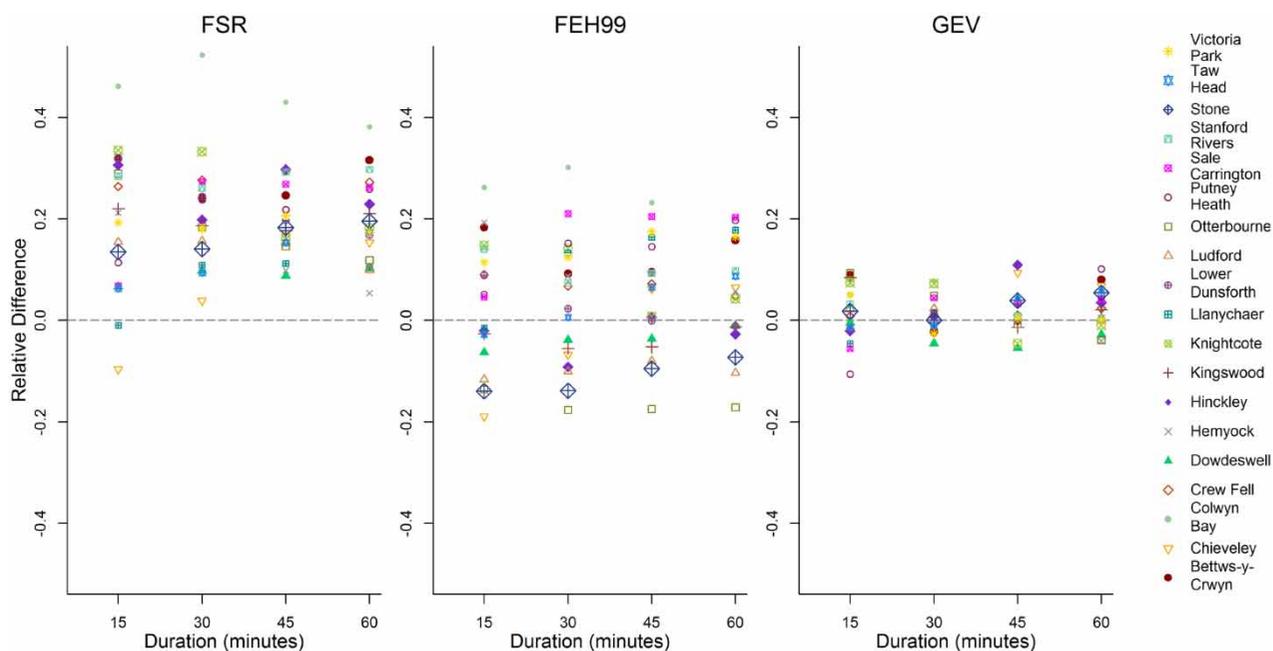


Figure 6 | Relative difference between the 2-year rainfall depths estimated by different methods and the 50th percentile of the recorded series. Larger symbols correspond to stations with longer records.

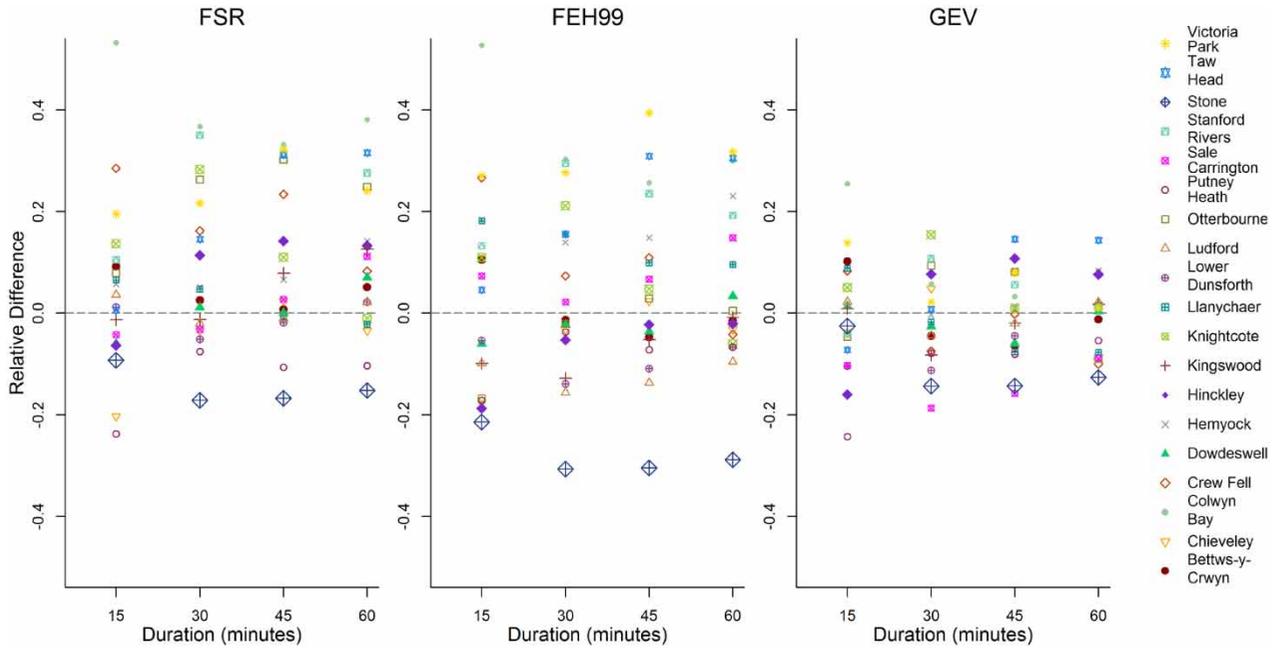


Figure 7 | Relative difference between the 5-year rainfall depths estimated by different methods and the 80th percentile of the recorded series. Larger symbols correspond to stations with longer records.

difference between the FEH99 and FSR, which were developed with the purpose of allowing DDF estimation for the whole UK, and the estimates obtained from the unified

GEV model, estimated using only at-site data is investigated in Figure 9. The figure shows, for a large range of return periods, the relative difference between the design events

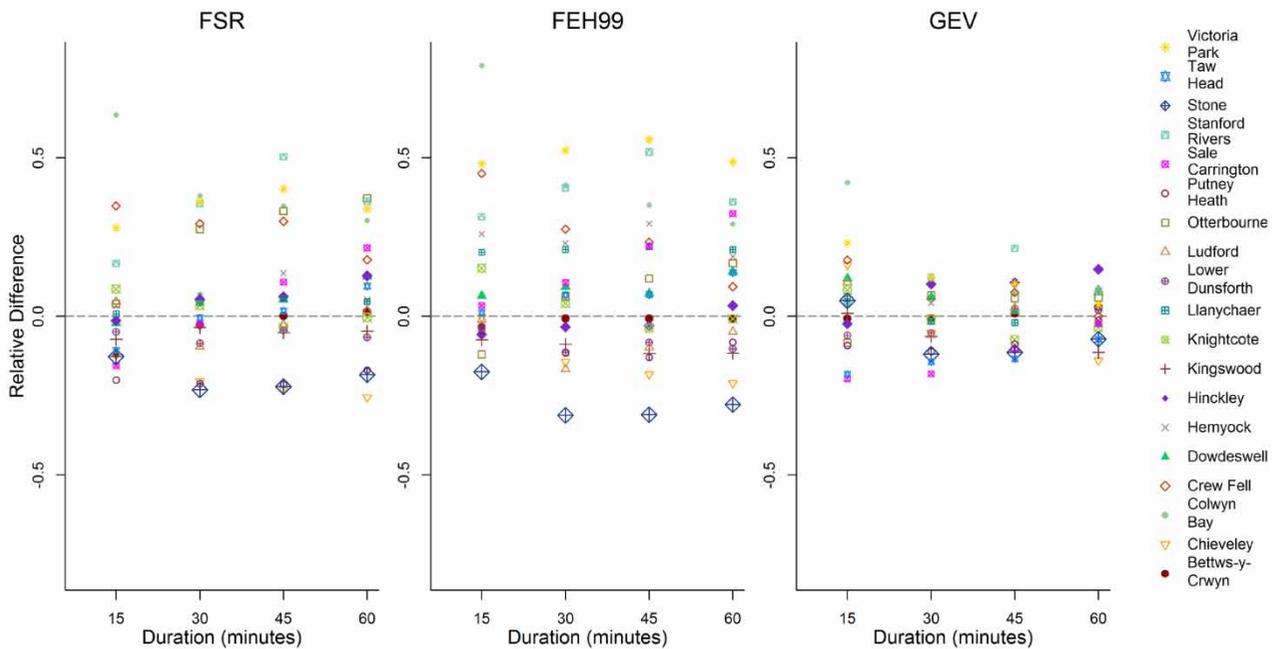


Figure 8 | Relative difference between the 10-year rainfall depths estimated by different methods and the 90th percentile of the recorded series. Larger symbols correspond to stations with longer records.

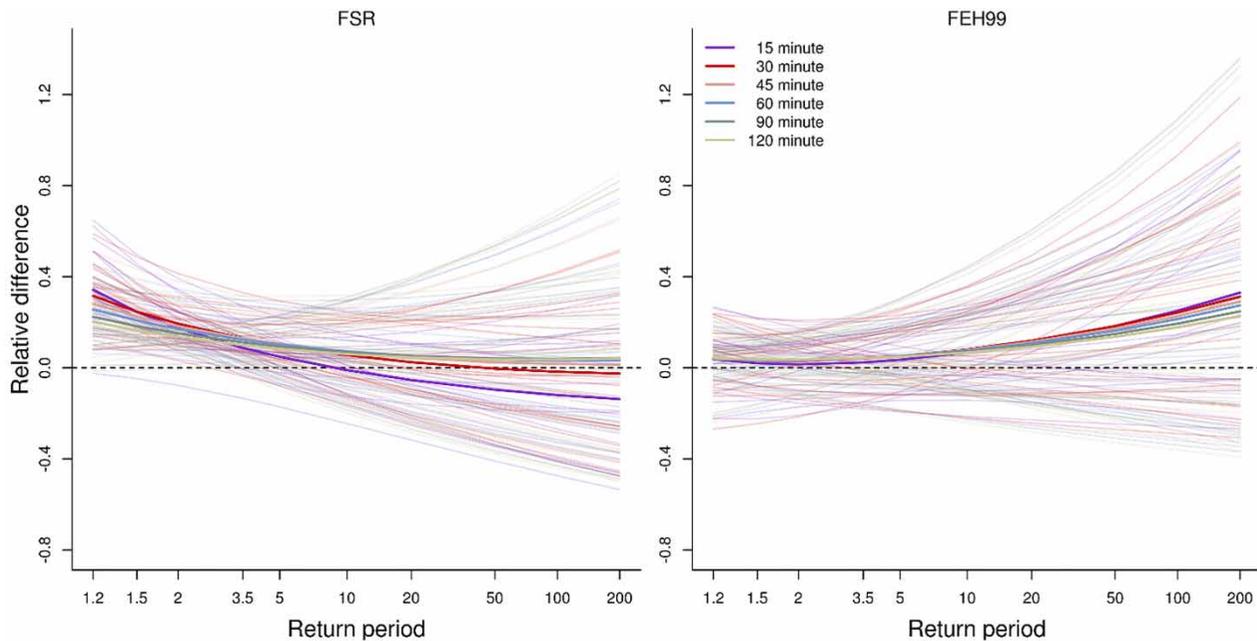


Figure 9 | Relative differences between the FSR (left panel) and FEH99 (right panel) estimates and the unified GEV estimates for all stations and all durations. Thick lines indicate the average differences for all durations across all stations.

estimated by FSR and the unified GEV (left panel) and the difference between the design events estimated by FEH99 and the unified GEV (right panel), for all stations and all durations. The average relative differences across all stations for each duration are also shown. It should be noted that a large difference between the standard methods and the unified GEV estimates does not necessarily indicate poor performance of the standard methods: the unified GEV models are fitted to the recorded data series, which are, at most, 46 years long. It is therefore very likely that unified GEV estimates would be more accurate for shorter rather than longer return periods. Nevertheless, what is visible in the plots is that the variability is much larger for the longer return periods for all durations. Furthermore, the FSR seems to give consistently larger results than the unified GEV for short return periods, but the difference between the two estimates become smaller for return periods longer than 10 years. For the 15-minute events the difference is more marked and the FSR seems to give much smaller estimates than the unified GEV for longer return periods. The difference between the FEH99 and the unified GEV results instead appears to increase for longer return periods, although for shorter return periods (up to 5 years) the difference in the two estimates is on average very small. At very

long return periods, it appears that the average difference between the unified GEV and the FEH99 estimates is smaller for events of long duration.

DISCUSSION AND CONCLUSIONS

An exploration of rainfall frequency estimation for short-duration events is presented. A new general at-site model, the unified GEV, is proposed. The model is successfully used to estimate consistent annual and seasonal rainfall frequency curves for a number of stations in England and Wales for which sub-hourly rainfall records exist. The proposed model builds on the standard assumption that block maxima follow a GEV distribution: the properties of the GEV distribution are exploited to construct a unified model which is fitted to the data of different duration simultaneously. The structure of the proposed model is indeed quite innovative and different from most of the DDF models currently used in practice. The consistency of the frequency curves is ensured by assuming that the lower bound and the skewness parameter are the same across all durations and by enforcing some basic relationships between the location and scale parameter and the

event duration. The effect of the assumptions, enforced to ensure the consistency of the frequency curves, is that curves for different durations diverge more and more as return period increases. The model might therefore give extremely large rainfall depth estimates for very long return periods. The model is designed to be fitted to block maximum series of sub-hourly data at single stations and does not have a procedure to integrate information from nearby stations in the rainfall frequency estimation. The estimation of the model parameters was carried out by maximum likelihood estimation, a procedure that attains some optimal properties when applied to large samples. The final model parameter estimates are influenced by properties of the observed data series and issues might arise when the actual properties of the observed series do not match well with the properties that are assumed during model building. Nevertheless, the proposed model gives overall satisfactory results and fits the empirical data quite well, using a relatively small number of parameters. The new estimated frequency curves are compared to those obtained using the FSR and FEH99 methods currently employed in the UK. Although no sub-hourly data were used in the model calibration, the FEH99 method seems to give acceptable results for all of the sub-hourly durations under study, at least for the return periods for which reliable empirical rainfall frequencies can be estimated. The FSR estimates seem to overestimate the rainfall depths for short return periods, although the bias is less marked for longer return periods. In addition, the difference between the FEH99 and the FSR estimates becomes larger for rarer events. However, the comparisons could only be carried out on a small set of stations, and a more in-depth analysis would be needed to give a robust indication of the behaviour of the different models. Potentially, it could be useful to develop a full DDF model for short duration rainfall events at a national scale, in which information from different stations could be used in a unique framework. The relative scarcity of long and precise records of sub-hourly data would be the major obstacle to overcome in the potential development of a DDF model for the whole UK. Most of the available ToT records are fairly short and most are located in only a few areas of the UK. Due to the nature of tipping buckets, the measurement of very short duration events is likely to be

biased, especially in less recent years, which would undermine the quality of any estimation procedure.

ACKNOWLEDGEMENTS

This work was funded by the Environment Agency Project SC090031 – Estimating flood peaks and hydrographs for small catchments (Phase 2). The original data series were provided by the Environment Agency. The authors thank Steven Cole at CEH Wallingford for providing code to process the raw time-of-tip series and David Jones for the fruitful discussions in the initial stage of the study. Part of the work presented in this study was carried out while the first author was employed at CEH Wallingford, the support of which is gratefully acknowledged.

REFERENCES

- Babtie Group in association with CEH Wallingford and Rodney Bridle Ltd 2000 *Reservoir Safety – Floods and Reservoir Safety: Clarification on the use of FEH and FSR Design Rainfalls*. Final report to the Department of the Environment, Transport and the Regions (DETR). Babtie Group, Glasgow, UK, 36 pp.
- Coles, S. G. 2001 *An Introduction to Statistical Modeling of Extreme Values*. Springer, London, UK.
- Faulkner, D. 1999 *Rainfall Frequency Estimation. Volume 2 of the Flood Estimation Handbook*. Institute of Hydrology, Wallingford, UK.
- Faulkner, D., Kjeldsen, T. R., Packman, J. C. & Stewart, E. J. 2012 *Estimating Flood Peaks and Hydrographs for Small Catchments: Phase 1*. Environment Agency, Bristol, UK.
- Hosking, J. R. M. 1990 L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society B* **52**, 105–124.
- ICE 2015 *Floods and Reservoir Safety*, 4th edn. Thomas Telford, London, UK.
- Jiang, P. & Tung, Y. 2013 [Establishing rainfall depth-duration-frequency relationships at daily raingauge stations in Hong Kong](#). *Journal of Hydrology* **504**, 80–93. doi:10.1016/j.jhydrol.2013.09.037.
- Kjeldsen, T. R. & Prosdocimi, I. 2015 [A bivariate extension of the Hosking and Wallis goodness-of-fit measure for regional distributions](#). *Water Resources Research* **51** (2), 896–907. doi: 10.1002/2014WR015912.
- Kjeldsen, T. R., Prudhomme, C., Svensson, C. & Stewart, E. J. 2006 [A shortcut to seasonal design rainfall estimates in the UK](#). *Water and Environment Journal* **20**, 282–286. doi:10.1111/j.1747-6593.2006.00028.x.

- Koutsoyiannis, D., Kozonis, D. & Manetas, A. 1998 [A mathematical framework for studying rainfall intensity-duration-frequency relationships](#). *Journal of Hydrology* **206** (1), 118–135. doi:10.1016/S0022-1694(98)00097-3.
- MacDonald, D. E. & Scott, C. W. 2001 FEH vs FSR rainfall estimates: an explanation for the discrepancies identified for very rare events. *Dams Reservoirs* **11**, 280–285.
- Molini, A., Lanza, L. & La Barbera, P. 2005 [Improving the accuracy of tipping-bucket rain records using disaggregation techniques](#). *Atmospheric Research* **77**, 203–217. doi:10.1016/j.atmosres.2004.12.013.
- Natural Environment Research Council 1975 *Flood Studies Report*. Natural Environment Research Council, Swindon, UK.
- Overeem, A., Buishand, A. & Holleman, I. 2008 [Rainfall depth-duration-frequency curves and their uncertainties](#). *Journal of Hydrology* **348**, 124–134. doi:jhydrol.2007.09.044.
- Prosdocimi, I., Stewart, L., Svensson, C. & Vesuviano, G. 2016 *Depth-Duration-Frequency Analysis for Short-Duration Rainfall Events*. Environment Agency, Bristol, UK.
- Stewart, E. J., Jones, D. A., Svensson, C., Morris, D. G., Dempsey, P., Dent, J. E., Collier, C. G. & Anderson, C. A. 2013 *Reservoir Safety – Long Return Period Rainfall. Project FD2613 WS194/2/39 Final Report (two volumes)*. London, UK.
- Svensson, C. & Jones, D. A. 2010 [Review of rainfall frequency estimation methods](#). *Journal of Flood Risk Management* **3**, 296–313. doi:10.1111/j.1753-318X.2010.01079.x.

First received 9 May 2016; accepted in revised form 11 September 2016. Available online 6 May 2017