

The FORGEX method of rainfall growth estimation

I: Review of requirement

E. J. Stewart, D. W. Reed, D. S. Faulkner and N. S. Reynard

Institute of Hydrology, Wallingford, Oxfordshire, OX10 8BB, UK
e-mail for corresponding author: d.reed@mail.nwl.ac.uk

Abstract

A growth factor is the ratio of the T -year extreme value to an index extreme value such as the mean of annual maxima. Whereas a record length of ten or more years may suffice to estimate the index variable, it is generally necessary to blend data from several sites if estimates of exceptional extreme values are to be obtained. Methods of rainfall growth estimation are reviewed, including traditional methods which extend frequency curves to long return period by a distributional assumption, and methods which study spatial dependence in extreme rainfalls. It is desirable that estimates at neighbouring sites, and across different durations and return periods, are internally consistent. The review concludes that rather special techniques may be required if this goal of estimating extreme rainfall depths consistently is to be met. The motivation of the Focused Rainfall Growth Extension (FORGEX) method is presented.

Introduction

Conceived almost ten years ago, the Focused Rainfall Growth Estimation (FORGE) method remains distinctive: a method of rainfall frequency estimation that exploits spatial dependence in rainfall extremes to reach out to long return periods but makes no major distributional assumption. The accompanying papers present and assess the Focused Rainfall Growth Extension (FORGEX) method. This reworking retains the strongly empirical character of its predecessor, and provides rainfall growth curves which extend to return periods as long as 2000 years.

After introducing terminology, the paper examines the context in which rainfall frequency estimates are required. Thereafter, methods are reviewed in two main groups: broadly traditional methods which extend growth curves to long return period by a distributional assumption, and less conventional approaches that consider spatial dependence in rainfall extremes. In between, the paper considers briefly the *station-year method*, which falls in neither category. While theoretically flawed, its strengths and weaknesses motivate the development of approaches based on dependence modelling.

TERMINOLOGY

A *rainfall frequency estimate* is an estimate of the depth of rain falling in a given duration—such as one hour—that is expected to be exceeded only occasionally. The frequency

or rarity of the fall is usually defined by reference to the *annual maximum series*, which comprises the largest values observed in each year. The frequency of a given rainfall value is specified by its *annual exceedance probability* (AEP), i.e. the probability that the value is exceeded in any year. Equivalently, the rarity of the value can be specified in terms of *return period*. The return period—strictly the *return period on the annual maximum scale*, T_{AM} —is the reciprocal of the AEP, and is measured in years. It is the average interval between years containing one or more exceedances of the given value. Because of the possibility of more than one extreme event occurring in a given year, T_{AM} is slightly longer than the true return period, which is the average interval between all exceedances without regard to measurement years. The difference is unimportant when the return period of interest is longer than about 20 years. Hence, T_{AM} is often abbreviated to T . A depth that has a return period of T years is referred to as the T -year maximum rainfall of the given duration, or just the T -year rainfall. Theoreticians usually refer to this as a quantile estimate.

An *index variable* is a typical extreme value. The index variable most often adopted is the mean of the annual maxima. The FORGEX method chooses the median annual maximum rainfall, $RMED$. Half of annual maxima are larger than the median so that $RMED$ has an annual exceedance probability of precisely 0.5, corresponding to the 2-year rainfall. The index variable is chosen to be a

relatively common event that can be reliably estimated from the record lengths of data typically available. The index variable is used both to characterize the typical size of extreme values expected at the site, and as a basis for inter-site comparisons.

A *regional analysis* is one in which data are combined, or *pooled*, from several sites. Sometimes a region corresponds to a fixed geographic area. In other cases, the region refers only to the gauged sites being grouped. It is usual for the regional analysis to seek to estimate a *growth curve*. The growth curve represents a set of dimensionless factors which specify the sizes of rare extreme events relative to the size of a common extreme event. The *T-year growth factor* is simply the ratio of the *T-year* rainfall depth to the index rainfall depth. In the regional approach, the estimate of rainfall frequency at a particular site is obtained as the product of the rainfall growth curve (derived in a regional analysis of rainfall growth) and the index rainfall depth. The index rainfall is either estimated from gauged data at the site, or interpolated from values estimated at neighbouring sites.

CONTEXT

How rare was that event?

Rainfall frequency analysis is a useful first step when major flooding has occurred or been narrowly averted. If the flood-producing rainfall can be shown to be exceptionally rare, a sufficient response may be to review operational procedures for flood warning and flood fighting, and rebuild dikes to their former level. If the damage incurred can be assessed, answering the question *How rare was that event?* provides a key datum in the assessment of the benefit-cost efficiency of flood defence. If the economic efficiency is low, or the impact of works for riverine habitats or flood risk elsewhere is high, it may be prudent to rebuild the defences only selectively and to accept a greater risk of inundation in some areas.

While frequency assessments based on *flood data* are to be preferred, there is greater scope to assess rainfall rarity because available records are typically longer and more justifiably transferable (from site to site) for rainfall than for floods. Thus the analyst chooses to answer the easier question (i.e. rainfall rarity estimation) well, rather than the more relevant question (i.e. flood rarity estimation) poorly.

Judging flood rarity from rainfall rarity is hazardous on catchments where other factors influence the flood response strongly. Thus the practice cannot be recommended for catchments known to be prone to snowmelt flooding, for very permeable catchments (where antecedent wetness is likely to be highly influential), or for sites subject to strong tidal or tributary interaction. Conversely, the inference of flood rarity from rainfall data will be more assured on naturally impermeable or heavily urbanized catchments which are routinely responsive to heavy rainfall. If long-term rainfall records are available for many

sites within the catchment, it may be helpful to carry out a frequency analysis of catchment-average maximum rainfalls. More typically, networks of long-term computerized records are sparse, and the question is answered by reference to rainfall data for a representative site within the catchment.

T-year rainfall frequency estimation as an input to flood design

Here the question is the more classical one: *What large rain-depth R , falling within duration D , can be expected to be exceeded only with a certain small probability in any year?* Rainfall frequency estimates are crucial to flood frequency estimation procedures based on the concept of a *design event*. A design event is a hypothetical rainfall event—usually in the form of a *T-year* rainfall depth of a duration thought to be most relevant to the catchment—which is assumed to occur when the catchment is in a stated wetness condition. A *design event procedure* sets criteria for choosing the hypothetical rainfall event, and the hypothetical antecedent catchment wetness, so that the desired *T-year* flood event is synthesized. Examples include the *Flood Studies Report* rainfall-runoff method (NERC, 1975), the *Australian Rainfall and Runoff* unit hydrograph and runoff routing methods (Pilgrim, 1987), and various Nordic methods based on the HBV model (e.g. Bergström *et al.*, 1992). The procedure is usually devised so that the *T-year* flood is synthesized from the *T-year* rainfall, although this is not a requirement. Because it provides a design hydrograph rather than just a flood peak, the approach is widely used in the design of storm-sewer systems, storm balancing ponds, flood storage reservoirs, and in appraisals of dam safety. Rainfall frequency estimation is also a prerequisite to flood frequency estimation by the Gradex method and its derivatives (e.g. Margoum *et al.*, 1994).

Drought rarity assessment

Rainfall frequency analysis in support of *drought* rarity assessment differs in two main respects: estimates are often required *during* an event, and the duration over which rainfall extremes are assessed is very much longer (typically, many months). Drought effects are mainly determined by cumulative shortfalls (of runoff or recharge) over many months, but can be influenced by pronounced seasonal and inter-annual effects. A resource system is not uniquely sensitive to a drought of a particular duration, or beginning in a particular month, and this interferes with attempts to infer drought severity from rainfall data. Although some of these issues may be resolvable (Reed, 1995), the difficulty remains that a given period of record includes relatively few independent events of long duration that can be considered genuinely extreme. Prolonged droughts are typically spatially extensive. In consequence, there is less new information to be gained by regional pooling of data than in the short-duration case considered here.

Design of communication systems

A different requirement arises in the design of communication systems which are subject to interference by intense rainfall. These design problems require the analysis of rainfall rates along a transect (e.g. Sims and Jones, 1974). However, the design durations and return periods are typically rather short. For example, communication systems sensitive to extreme 1-minute rainfall rates are sometimes designed so that rainfall disables communication for only 0.01% of the time. This represents about 53 minutes per year, meaning that there will be failures in most years. Rather than designing communication equipment to be fully reliable in extreme conditions, it may be cost-effective to accept occasional failures or to meet the target reliability by networked systems that are resilient to the failure of individual links.

Storm impacts

The *hail impact* of a storm is relatively direct, immediate and localized: it depends on the intensity of precipitation and its form, in particular the size of the hailstones. In contrast, the *flood impact* of a storm is intricate. It depends on many factors: the location, depth and temporal distribution of heavy rainfall in relation to the layout of drainage paths, the natural response characteristics of the catchments that these drainage paths define, and the rainfall history prior to the storm.

The link between storm rarity and flood rarity might be expected to be clearer in the very highest impact events. The *stochastic storm transposition* method (Fontaine and Potter, 1989; Wilson and Fofoula-Georgiou, 1990) provides a possible approach to the difficult problem of estimating rainfall depths of very long return period. The method focuses on storms (rather than point or catchment rainfalls), and includes extreme storms that have not occurred over the catchment of interest—but could have. In somewhat similar vein, *storm modelling* (e.g. Collier and Hardaker, 1996) perhaps provides the only satisfactory route to estimating probable maximum precipitation, if this is held to be a useful concept.

Should design rainfalls respect storm properties?

Efficient design requires the correct sizing of flood defences in relation to the full range of flood conditions experienced. Flood-producing rainfalls are catchment rainfalls *not* storms. They may represent a whole storm or half a storm or two storms, an unusually stationary storm or a passing squall, a multi-cellular thunderstorm or a small part of widespread frontal rainfall. It is this richness of scenarios, coupled with the requirement to make frequency statements, that encourages a statistical analysis of extreme gauged rainfall depths rather than a meteorological study of storm properties.

It is usual to concentrate on the frequency analysis of extreme rainfall depths in duration D , and to see the subsequent construction of a *design storm* as a separate step.

This is not to deny that the empirical analysis of extreme rainfall depths can benefit from detailed meteorological understanding. Revfeim (1992) refers to situations where annual maximum rainfalls abstracted for a long duration may in reality represent much shorter-duration falls. The phenomenon of preferential storm durations—prevalent in tropical climates—complicates the derivation of depth-duration-frequency relationships that are internally consistent. From a hydrological perspective, the worry is that incorrect deductions will be drawn if the T -year D -hour rainfall estimate is applied to synthesize a design flood on a catchment purported to be sensitive to heavy rainfall of a duration that never occurs in practice! While this suggests a possible weakness in undertaking rainfall frequency analyses without reference to meteorological properties, the particular difficulty arises from the use to which the rainfall frequency estimates are put. Limitations in the *design event* approach to flood frequency estimation do not invalidate the method of rainfall frequency estimation.

PREVIOUS REVIEWS

Sevruk and Geiger (1981) present a chronology (from 1852) of the development of frequency analysis methods—for floods as well as rainfall—but make no mention of regional analysis (i.e. pooling of data). Buishand (1989) presents a wide-ranging review of the statistical issues met in extreme-value applications in climatology: advocating regional analyses and discussing problems arising from inter-site dependence.

Frequency analysis methods (principally for floods) are reviewed by Cunnane (1989). Rossi and Villani (1994) review flood frequency estimation methods and elaborate the *hierarchical approach* to regional frequency estimation. This associates magnitude, variability and skewness of the distribution of extreme values with local, district and regional effects. The approach is no less appropriate for extreme rainfalls than for floods. GREHYS (1996) reviews methods of regional flood frequency analysis, giving particular emphasis to the delineation of homogeneous regions.

Hosking and Wallis (1997) present a general approach to regional frequency analysis based on L-moments. This authoritative text amplifies and extends their earlier publications on L-moment (and probability-weighted moment) methods.

Methods

PEAK-OVER-THRESHOLD ANALYSIS

Most procedures for rainfall frequency estimation are based on the analysis of annual maximum rainfall depths of a given duration, often daily. This contrasts with theoretical guidance, which invariably recommends the analysis of extremes abstracted in *peak-over-threshold* (POT) format. Davison and Smith (1990) consider POT methods

in general, while their application to extreme rainfalls is considered by Van Montfort and Witter (1986) and Fitzgerald (1989). Amongst other virtues, POT methods importantly include all large events (rather than no more than one per year) and exclude annual maximum events that are unusually small (and might mislead an analysis of extremes).

Madsen, Rasmussen and Rosbjerg (1997) and Madsen, Pearson and Rosbjerg (1997) compare annual maximum and partial duration (i.e. POT) methods for single-site and regional analyses respectively. The papers conclude that the POT approach is generally superior. Such conclusions are unsurprising given that a POT series provides more comprehensive and extensive information than the annual maximum series: more comprehensive because all extreme events are represented, and more numerous because the threshold is typically set to yield a mean number of events per year between two and five. The greater power of analyses based on POT data is also evident when testing time-series of extreme values for non-stationarity (Robson *et al.*, 1998).

Unfortunately, the theoretical advantages of POT methods can be outweighed by the more demanding task of extracting peak-over-threshold events. An irritant is the lack of clear-cut rules by which to determine which peaks to exclude from the analysis when successive peaks represent different parts of the same event. Such decisions are difficult even when a continuous record of the variable is available, as is usual with water level recorders in rivers. But rainfall data are most often recorded as aggregate depths falling within fixed intervals, the longest and most numerous records being based on daily observations. Thus it is necessary to make rather arbitrary assumptions about temporal dependence: for example, that a large 1-day rainfall depth represents only one extreme event, and that large 1-day rainfalls on consecutive days always (or, alternatively, never) represent independent extreme events.

The treatment of brief missing periods in records is also more demanding in the POT approach. If the recorded annual maximum is large, it may be reasonable to assume that this represents the true annual maximum. However, it is difficult to be certain that a threshold exceedance (i.e. a POT event) has not occurred during the period for which records are missing. This is another instance where the problem is typically more difficult for rainfall depths than river levels. Whereas a continuous river level record yields evidence of its operation (or non-operation) at all times, a daily raingauge provides only a single datum per day, many of which take the rather uninformative value of zero. Smith and Shively (1995) discuss the treatment of missing periods of record in the POT approach.

The definition and treatment of inter-site dependence are much less straightforward for extremes defined as peaks over a threshold than for annual maxima. This is a specific discouragement using POT data in the context of

developing a method of rainfall growth estimation that takes account of spatial dependence in extreme rainfalls.

SINGLE-SITE ANALYSIS OF ANNUAL MAXIMA

The analysis of annual extremes of a variable measured at a particular site is a relatively mature branch of applied statistics, with many important advances being stimulated by workers in meteorology (e.g. Jenkinson, 1955) and hydrology (e.g. Cunnane, 1978; Hosking *et al.*, 1985). The key decisions centre on the choice of an appropriate distribution and the choice of fitting method. Single-site methods are reviewed in detail by Stedinger *et al.* (1992), who also discuss goodness-of-fit tests and the construction of confidence intervals for T -year values.

Choice of distribution

Sevruk and Geiger (1981) report that the Gumbel, Generalized Extreme Value (GEV), and Log-Normal distributions are the most commonly assumed distributions for annual maximum rainfalls. Their plea for research to inform the choice of distribution from physical reasoning does not appear to have been widely heeded. One exception is the two-component analysis by Pegram and Adamson (1988) of annual maximum rainfalls in sub-tropical climates. In the two-component approach (e.g. Arnell and Gabriele, 1988), annual maxima are assumed to come from a mix of populations: one component representing *normal* extreme events, the other representing *exceptional* extreme events (or outliers). Ideally, annual maxima are classified prior to analysis, from physical reasoning and concurrent weather observations. Although Pegram and Adamson are unable to do this, they compensate by offering a physical interpretation for the spatial variation seen in the proportion of annual maximum events assigned to the outlier component. They find that the proportion is smaller for stations in the hinterland than near the coast, reflecting the less frequent intrusion of tropical cyclones.

Method of fitting

Graphical techniques, and moment methods, were formerly the most commonly used methods of fitting. More recent studies (e.g. Griffiths and Pearson, 1993; Smithers, 1996) have typically adopted L-moment methods (see Hosking, 1990 or Hosking and Wallis, 1997). In a theoretical comparison of annual maximum and POT methods, Madsen, Rasmussen and Rosbjerg (1997) also consider *maximum likelihood* methods. While providing efficient estimates for large sample sizes, maximum likelihood methods can be rather poor when applied to the small sample sizes typical of applications to environmental extremes, as these authors demonstrate for the GEV distribution.

Graphical display

Although no longer fashionable as a method of fitting, graphical methods are indispensable to a single-site

extreme value analysis. An extreme value plot is constructed by ranking the data in order of magnitude and plotting them according to a *plotting position* formula linked to a standard frequency scale. The manner of display inevitably affects interpretations: a specific choice would be appropriate if the underlying population were known. However, this is never the case for environmental extremes and it is necessary to follow some convention. Gringorten's plotting position formula and Gumbel's reduced variate scale provide a fairly neutral choice for the typical range of extreme value samples met in meteorology and hydrology. A straight-line plot indicates that the sample is consistent with a Gumbel distribution, the plotting position formula being nearly unbiased for samples drawn from a Gumbel distribution (Gringorten, 1963; NERC, 1975; Cunnane, 1978).

The extreme value plot provides an immediate signal of whether there are any unusually high or low values in the sample. It is important to inspect the extent to which the chosen fitting method has neglected or accommodated any unusual value, by superimposing the fitted distribution on the extreme value plot. Additionally, a time-series plot of the annual maxima is advisable as a check for possible non-stationarity.

REGIONAL ANALYSIS OF ANNUAL MAXIMA

Consistent estimates of climatological extremes of very long return period can only be obtained by combining information from several sites (e.g. NERC, 1975; Buishand, 1989; Cunnane, 1989).

A long-standing procedure for rainfall frequency estimation in the UK is the regional method developed by Jenkinson and presented in the *Flood Studies Report* (NERC, 1975). This adopts an *index variable* approach (see introductory section on terminology) in which the T -year D -hour rainfall depth is synthesized as the product of the 5-year value (denoted by M_5) and a growth factor (MT/M_5) which specifies the ratio of T -year and 5-year values. At first sight, the procedure has only a broad regionalization, with the UK divided into just two geographical regions for rainfall growth estimation. However, additional regionalization is introduced by pooling data (for the purpose of growth curve derivation) according to the M_5 rainfall depth. The properties of this intricate regionalization scheme are difficult to decipher. More usually, stations are pooled according to geographical or administrative region, or according to a variable, such as the long-term average annual rainfall, which acts as a proxy for the rainfall regime.

Dales and Reed (1989) adopt fixed geographical regions and derive a regional rainfall frequency model by pooled analysis of annual maxima. They fit a Generalized Extreme Value (GEV) distribution by the method of regional Probability-Weighted Moments. This is a forerunner of regional L-moment methods (see below). Schaefer (1990)

takes an iterative approach to forming regions, beginning with a division suggested in an earlier study. He uses regional PWM values to fit the very flexible 5-parameter Wakeby distribution, enabling a choice to be made between competing 3-parameter distributions. After the exploratory analysis, Schaefer adopts the GEV distribution and applies it within a regionalization scheme which takes account of the typical decrease in annual maximum variability and skewness with increasing mean annual precipitation. Buishand (1991) introduces a regional analysis based on maximizing a joint likelihood function across all sites in a region, again adopting a GEV distribution.

Hosking and Wallis (1997) present a comprehensive approach to regional frequency analysis based on L-moment methods. L-moment ratios derived from single-site analyses are combined by regional averaging, weighted according to record length. An aid to visualizing the approach is to superimpose in one diagram the regionally-fitted distribution and individual extreme-value plots for each station; each station-record is brought to a common base by dividing by an index variable such as the mean of the annual maxima. Pooling data from a larger number of stations, by selecting a larger region, can induce additional smoothing that is helpful in stabilizing estimates at very long return periods. However, this gain may be at the price of an additional bias incurred through excessive generalization. Moreover, the accompanying hypothesis—that stations share a common underlying distribution of rainfall growth—is inevitably weaker for a larger group of stations.

The selection of appropriate regions is a general problem. Guttman (1993) adopts fixed regions selected after an exploratory data analysis. Fixed regions are convenient to use but induce discontinuities in growth factor estimates at boundaries. This is a serious limitation except in those cases where boundaries correspond to major topographical divides. Flexible regions are less convenient for general use—since the growth curves cannot be compactly summarized in a map or table—but have the twin merits of avoiding boundary problems and focusing the analysis on the site of interest (Reed and Stewart, 1989; Burn, 1990). Rossi and Villani (1994) adopt fixed regions but their hierarchical approach could in principle be applied to flexible regions.

Station-year approach

The station-year approach to estimating long return-period events seeks to avoid or reduce the need for extrapolation of a fitted distribution. Instead, the approach combines (by concatenation) records from individual sites, to form a single record equal in length to the sum of the individual record lengths. Like other methods of pooling, the approach assumes that the distribution of extremes is identical at each site once the annual maxima have been standardized by division by an index variable. Additionally, the station-year method assumes that the

records at the sites are mutually independent. This is only valid if the sites are very widely scattered or if their periods of record are mutually exclusive. The latter condition is never met in practice, while, if the former condition holds, the underlying assumption of identical distributions is unlikely to be tenable. Thus, in practice, the approach lacks justification.

Despite these weaknesses, the station-year method continues to find application to meet the engineering requirement to extend frequency estimates to long return periods. A typical procedure when using the station-year method is to avoid pooling records from near-neighbours, to avoid excessive dependence in the pooled extremes. However, Reed and Stewart (1989) illustrate that the station-year assumption—that the combined record has an effective record length equal to the sum of the individual record lengths—introduces less bias than appears at first glance. This is because applications of the method typically allow an individual extreme event—experienced at more than one site—to contribute more than one *station-year point* to the analysis. In other words, the expected poor performance of the method is not fully realized in practice because of compensating errors: it neglects dependence both in the extremes selected for plotting and in calculating the positions at which they are plotted. In such a situation, it is reasonable to expect estimates to suffer greater variability rather than greater bias. This heuristic argument is consistent with a much-quoted finding of Hosking and Wallis (1988): that inter-site dependence in extremes reduces the confidence with which a regional growth curve is estimated but does not lead to bias. Their result is based on extensive sampling from flood-like distributions of extreme values. However, a barrier to definitive conclusions is the difficulty of demonstrating that schemes used to generate samples of partially dependent extreme values adequately replicate the dependence patterns seen in real data.

Studies of spatial dependence in rainfall extremes

Fricke *et al.* (1983) study dependence in rainfall extremes observed in a network of N gauges, and use the concept of an effective number of independent stations suggested by (amongst others) Yevjevich (1972):

$$N_e = \frac{N}{1 + \bar{r}(N-1)} \quad (1)$$

where \bar{r} is the mean correlation coefficient between all pairs of raingauges. However, this formula is only valid for defining the effective number of independent series for the purpose of defining the mean of the variable, not for extremes. Moreover, assessments of N_e are highly sensitive to the criteria for prescribing the heavy rainfalls to be used in evaluating the mean correlation coefficient, \bar{r} .

Buishand (1984) uses the theory of bivariate extremes to explore the dependence between rainfall maxima at pairs of sites. Later, Dales and Reed (1989) interpret Buishand's dependence function, q , as an effective number of independent sites, and extend the concept to assess dependence in rainfall maxima at N sites. For the bivariate case (i.e. $N = 2$ and $N_e = q$), Buishand explores the variation of N_e with inter-site distance and return period. He deduces that, for 1-day rainfall maxima in the Netherlands, the inter-site dependence is sufficiently small at long return period to justify use of the station-year method. The one proviso is that the stations grouped should avoid near-neighbours.

Dales and Reed (1989) derive a general model for spatial dependence in UK rainfall extremes. For many experimental networks of N gauges, they study the relationship between the frequency distribution of the largest standardized values observed by the network (the so-called *network maximum curve*) and the average frequency distribution of the standardized values at individual sites (the so-called *typical curve*). As noted by Reed and Stewart (1994) for any parent distribution, $F(x)$, the maximum of N independent and identically distributed random variables lies a fixed distance, $\ln N$, to the left of the parent on a Gumbel (variate versus reduced-variate) plot. Thus the horizontal separation between the network maximum and typical curves, $\ln N_e$, defines an effective number of independent sites, N_e . Dales and Reed derive a fixed N_e value for a given network of raingauges by assuming a GEV parent distribution. Their resultant model allows N_e to be estimated from N , the area spanned by the network ($AREA$ km²), and the duration of rainfall extremes (D days). The UK-average model is:

$$\ln N_e = \ln N(0.081 + 0.085 \ln AREA - 0.051 \ln N - 0.027 \ln D) \quad (2)$$

Dales and Reed define the area spanned by the network by the heuristic rule:

$$AREA = 2.5\bar{d}^2 \quad (3)$$

where \bar{d} denotes the mean inter-gauge distance in kilometres evaluated pairwise across all combinations.

The ratio $\ln N_e / \ln N$ is the fractional distance between the fully dependent case (when $N_e = 1$) and the fully independent case (when $N_e = N$). Thus:

$$1 - \frac{\ln N_e}{\ln N} = 0.919 - 0.085 \ln AREA + 0.051 \ln N + 0.027 \ln D$$

defines the degree of inter-site dependence (Reed and Stewart, 1994) seen in rainfall extremes in the network. The degree of dependence increases as the number of gauges increases or the duration lengthens, but decreases as the spanning area increases. Dales and Reed apply the spatial dependence model (Eqn. 2) to assess the annual

collective risk, r , of a T -year D -day rainfall being experienced at one or more of a network of N sites, given by:

$$r = 1 - \left(1 - \frac{1}{T}\right)^N. \quad (4)$$

Having undertaken a pooled analysis of annual maximum depths from a network of about 115 recording rain-gauges in Washington State, Schaefer (1990) touches on inter-site dependence effects by defining an equivalent independent record length (EIRL) for the network. This is evaluated by counting the number of high exceedances recorded by the network, retaining only the rarest event when two or more annual maxima stem from the same storm or storm system. Schaefer associates the longer EIRL for 2-hour extremes (compared to 24-hour extremes) with the typically smaller inter-site correlations in short-duration annual maxima, and suggests that both effects are symptomatic of the more limited areal extent of convective storm activity, that is often responsible for 2-hour annual maxima. He suggests that determination of an EIRL is beneficial as an indicator of the acceptable limits of extrapolation of frequency curves based on regional analysis.

In a study of methods of extreme rainfall estimation, Buishand (1991) checks a general model of rainfall frequency, derived in a pooled analysis of 15 long-record stations, by counting extreme events in a much larger national dataset. For the rarest quantiles (500 and 1000-year events) the counts are smaller than expected, though within the range that might be ascribed to chance. He finds that inter-site dependence in the rarest events leads to variability in the counts and reduces the precision of any test based on counts. He extends a method by Clarke-Hafstad (1942) to provide a measure of inter-site dependence in extreme events based on counting pairs of stations with exceedances in the same year.

Coles (1993) also studies spatial dependence in rainfall extremes. His extremal coefficient corresponds to the effective number of independent sites, N_e . Reed and Stewart (1994) study inter-site and inter-duration dependence in rainfall extremes, assuming a GEV model for rainfall growth. In addition to applications in collective risk estimation (Eqn. 4), Reed and Stewart suggest that dependence models can assist in choosing the locations and durations on which to concentrate further extraction of data from chart and/or manuscript records.

Rainfall frequency estimation accounting for inter-site dependence

Reed and Stewart (1989) present a modified station-year method in which the combined record length of the pooled data is reduced to an effective record length, taking account of inter-site dependence in extremes by invoking the Dales and Reed spatial dependence model.

Additionally, Reed and Stewart introduce the concept of focusing the growth curve analysis on the site for which rainfall frequency estimates are required. Their focused rainfall growth estimation (FORGE) method is developed further in the next paper (Reed *et al.*, 1999). A feature of the method is that the rainfall growth curve is determined by the data and the spatial dependence model (Eqn. 2) alone, without an explicit distributional assumption. The resultant flexibility in derived growth curves might alternatively be gained by fitting a 5-parameter Wakeby distribution.

Nandakumar *et al.* (1997) present a variant of the FORGE method in which the effective number of independent sites, N_e , varies with return period. Results presented by Buishand (1984) and Dales and Reed (1989) suggest that N_e may increase at longer return period. This is consistent with studies of areal reduction factors (e.g. Stewart, 1989; Siriwardena and Weinmann, 1996) which suggest that rarer storms are typically somewhat less spatially extensive.

It is unreasonable, and potentially unsafe, to assume that N_e approaches N at long return periods, since this would imply that the most extreme rainfall depths will arise from highly localized storms rather than from the genesis or intrusion of a spatially extensive extreme event not previously recorded in the region. An earlier section noted the difficulty—in rainfall frequency analysis for drought rarity assessment—that a given period of record includes relatively few independent events of very long duration. The analogue is that a given national database of rainfall measurements includes relatively few spatially extensive extreme events. While acknowledging that inter-site dependence *may* reduce somewhat at long return periods, the preference here is to develop a method for rainfall growth estimation based on the Dales and Reed model of spatial dependence. This reflects both concern that available datasets provide a relatively small sample of spatially extensive extreme events and the desire to avoid embellishment that may inadvertently lead to underestimation.

Nandakumar *et al.* test several aspects of the FORGE method. In particular, they find that the growth curve estimates are relatively insensitive to the minor distributional assumption implicit in the method, namely the use by Dales and Reed (1989) of a GEV distribution in calculating the N_e values for calibration of the spatial dependence model (Eqn. 2).

Summary

This paper has reviewed the context in which rainfall frequency estimates are needed and explored the particular challenge of estimating long return-period events. The following features are suggested as practical requirements for a general method of rainfall frequency estimation:

- use of annual maximum rainfall data;
- standardization by an index variable;

- extension of the growth curve to long return period;
- avoidance of a restrictive distributional assumption;
- avoidance of excessive generalization which suppresses local variations;
- avoidance of boundary problems associated with pooling data within fixed regions.

The FORGEX method (see accompanying papers) is designed to meet these requirements. Subsequent steps to growth curve estimation are the mapping of the index variable (Faulkner and Prudhomme, 1998) and the generalization of depth-duration-frequency relationships (e.g. Faulkner, 1999).

Acknowledgements

The authors are grateful to rainfall frequency researchers with whom they have exchanged ideas in the decade during which the FORGEX method has evolved, including Adri Buishand, Jon Hosking, Bruce Kelbe, Nanda Nandakumar, Jonathan Tawn and Erwin Weinmann. The Australian Bureau of Meteorology is thanked for supplying extensive hourly rainfall records at a critical stage in the research. Comments from the anonymous referees are gratefully acknowledged.

References

- Arnell, N.W. and Gabriele, S., 1988. The performance of the two-component extreme value distribution in regional flood frequency analysis. *Wat. Resour. Res.*, **24**, 879–887.
- Bergström, S., Harlin, J. and Lindström, G., 1992. Spillway design floods in Sweden, I: New guidelines. *Hydrol. Sci. J.*, **37**, 505–519.
- Buishand, T.A., 1984. Bivariate extreme-value data and the station-year method. *J. Hydrol.*, **69**, 77–95.
- Buishand, T.A., 1989. Statistics of extremes in climatology. *Statistica Neerlandica*, **43**, 1–30.
- Buishand, T.A., 1991. Extreme rainfall estimation by combining data from several sites. *Hydrol. Sci. J.*, **36**, 345–365.
- Burn, D.H., 1990. Evaluation of regional flood frequency analysis with a region of influence approach. *Wat. Resour. Res.*, **26**, 2257–2265.
- Clarke-Hafstad, K., 1942. Reliability of station-year rainfall-frequency determinations (with discussion). *Trans. Am. Soc. Civ. Engrs*, **107**, 633–683.
- Coles, S.G., 1993. Regional modelling of extreme storms via max-stable processes. *J. Roy. Statist. Soc.*, **B55**, 797–816.
- Collier, C.G. and Hardaker, P.J., 1996. Estimating Probable Maximum Precipitation using a storm model approach. *J. Hydrol.*, **183**, 277–306.
- Cunnane, C., 1978. Unbiased plotting positions – a review. *J. Hydrol.*, **37**, 205–222.
- Cunnane, C., 1989. *Statistical distributions for flood frequency analysis*. World Meteorological Organization Operational. Hydrol. Rep. No. 33, WMO Publ. No. 718, Geneva, 73pp.
- Dales, M.Y. and Reed, D.W., 1989. Regional flood and storm hazard assessment. Institute of Hydrology Rep. No. 102, IH, Wallingford, UK. 159pp.
- Davison, A.C. and Smith, R.L., 1990. Models for exceedances over high thresholds. *J. Roy. Statist. Soc.*, **B52**, 393–440.
- Faulkner, D.S., 1999. Building the depth-duration-frequency model. Chapter 10 in *Rainfall frequency estimation*, Volume 2 of *Flood Estimation Handbook*, Institute of Hydrology, Wallingford, UK.
- Faulkner, D.S. and Jones, D.A., 1999. The FORGEX method of rainfall growth estimation, III: Examples and confidence intervals. *Hydrol. Earth System Sci.*, **3**, 205–212.
- Faulkner, D.S. and Prudhomme, C., 1998. Mapping an index of extreme rainfall across the UK. *Hydrol. Earth System Sci.*, **2**, 183–194.
- Fitzgerald, D.L., 1989. Single station and regional analysis of daily rainfall extremes. *Stoch. Hydrol. and Hydraul.*, **3**, 281–292.
- Fontaine, T.A. and Potter, K.W., 1989. Estimating probabilities of extreme rainfalls. *J. Hydraul. Eng., Am. Soc. Civil Engineers*, **115**, 1562–1575.
- Fricke, T.J., Kennedy, M.R. and Wellington, N.B., 1983. *The use of rainfall correlation in determining design storms for waterways on a long railway line*. Proc. Hydrol. and Water Resources Symp., Hobart, November 1983, 5pp.
- GREHYS (Groupe de recherche en hydrologie statistique) 1996. Presentation and review of some methods for regional flood frequency analysis. *J. Hydrol.*, **186**, 63–84.
- Gringorten, I.I., 1963. A plotting rule for extreme probability paper. *J. Geophys. Res.*, **68**, 813–814.
- Griffiths, G.A. and Pearson, C.P., 1993. Distribution of high intensity rainfalls in metropolitan Christchurch, New Zealand. *J. Hydrol., (N.Z.)*, **31**, 5–22.
- Guttman, N.B., 1993. The use of L-moments in the determination of regional precipitation climates. *J. Climate*, **6**, 2309–2325.
- Hosking, J.R.M., 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. *J. Roy. Statist. Soc.*, **B52**, 105–124.
- Hosking, J.R.M. and Wallis, J.R., 1988. The effect of inter-site dependence on regional flood frequency analysis. *Wat. Resour. Res.*, **24**, 588–600.
- Hosking, J.R.M. and Wallis, J.R., 1997. *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press, 224pp.
- Hosking, J.R.M., Wallis, J.R. and Wood, E.F., 1985. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, **27**, 251–261.
- Jenkinson, A.F., 1955. The frequency distribution of the annual maximum (or minimum) of meteorological elements. *Quart. J. Roy. Meteorol. Soc.*, **81**, 158–171.
- Madsen, H., Pearson, C.P. and Rosbjerg, D., 1997. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events, 2: Regional modeling. *Wat. Resour. Res.*, **33**, 759–769.
- Madsen, H., Rasmussen, P.F. and Rosbjerg, D., 1997. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events, 1: Regional modeling. *Wat. Resour. Res.*, **33**, 747–757.
- Margoum, M., Oberlin, G., Lang, M. and Weingartner, R., 1994. Estimation des crues rares et extrêmes: principes du modèle Agregee. *Hydrol. continent.*, **9**, 85–100.
- Nandakumar, N., Weinmann, P.E., Mein, R.G. and Nathan, R.J., 1997. *Estimation of extreme rainfalls for Victoria using the CRC-FORGE method*. Report 97/4, Co-operative Research Centre for Catchment Hydrology, Monash University, Australia, 118pp.

- NERC, 1975. *Flood Studies Report* (five volumes). Natural Environment Research Council, London.
- Pegram, G. and Adamson, P., 1988. Revised risk analysis for extreme storms and floods in Natal/KwaZulu. *Die Siviele Ingenieur in Suid-Afrika*, Jan. 1988, 15–20 and 42. (Discussion: Jul. 1988, 331–336.)
- Pilgrim, D.H. (ed.), 1987. *A guide to flood estimation*. Volume 1 of *Australian Rainfall and Runoff*, Institution of Engineers, Australia.
- Reed, D.W., 1995. Rainfall assessments of drought severity and centennial events. Proc. CIWEM Centenary Conf., October 1995, Chartered Inst. Wat. Environ. Mang., London, 16.1–16.17.
- Reed, D.W., Faulkner, D.S. and Stewart, E.J., 1999. The FORGEX method of rainfall growth estimation, II: Description. *Hydrol. Earth System Sci.*, 3, 197–203.
- Reed, D.W. and Stewart, E.J., 1989. Focus on rainfall growth estimation. Proc. British Hydrol. Soc. Nation. Symp., Sheffield, September 1989, 3.57–3.65.
- Reed, D.W. and Stewart, E.J., 1991. Discussion on: "Dam safety: an evaluation of some procedures for design flood estimation". *Hydrol. Sci. J.*, 36, 487–490.
- Reed, D.W. and Stewart, E.J., 1994. Inter-site and inter-duration dependence in rainfall extremes. Chapter 8 in: *Statistics for the Environment 2: Water Related Issues* (Editors V. Barnett and K.F. Turkman), Wiley, Chichester, UK, 125–143.
- Revfeim, K.J.A., 1992. Dominant events in extreme rainfall records. *J. Hydrol.*, 134, 143–149.
- Robson, A.J., Jones, T.K., Reed, D.W. and Bayliss, A.C., 1998. A study of national trend and variation in UK floods. *Int. J. Climatol.*, 18, 165–182.
- Rossi, F. and Villani, P., 1994. Regional flood estimation methods. In *Coping with Floods* (Editors G. Rossi, N. Harmancioglu and V. Yevjevich), Kluwer, Dordrecht, 135–169.
- Schaefer, M.G., 1990. Regional analyses of precipitation annual maxima in Washington State. *Wat. Resour. Res.*, 26, 119–131.
- Sevruk, B. and Geiger, H., 1981. *Selection of distribution types for extremes of precipitation*. Operational Hydrol. Rep. No. 15, World Meteorological Organization, Geneva, WMO Publ. No. 560, 64pp.
- Sims, A.L. and Jones, D.M.A., 1974. Frequencies of short-period rainfall rates along lines. *J. App. Meteorol.*, 14, 170–174.
- Siriwardena, L. and Weinmann, P.E., 1996. *Derivation of areal reduction factors for design rainfalls in Victoria*. Report 96/4, Co-operative Research Centre for Catchment Hydrology, Monash University, Australia, 66pp.
- Smith, R.L. and Shively, T.S., 1995. Point process approach to modeling trends in tropospheric ozone based on exceedances of a high-threshold. *Atmos. Environ.*, 29, 3489–3499.
- Smithers, J.C., 1996. Short-duration rainfall frequency model selection in Southern Africa. *Water SA*, 22, 211–217.
- Stedinger, J.R., Vogel, R.M. and Foufoula-Georgiou, E., 1992. Frequency analysis of extreme events. In *Handbook of hydrology* (Ed. D.R. Maidment), McGraw-Hill, New York, 18.1–18.66.
- Stewart, E.J., 1989. Areal reduction factors for design storm construction: joint use of raingauge and radar data. In *New Directions for Surface Water Modeling*, IAHS Publ. No. 181, 31–40.
- Van Montfort, M.A.J. and Witter, J.V., 1986. The Generalized Pareto distribution applied to rainfall depths. *Hydrol. Sci. J.*, 31, 151–162.
- Wilson, L.L. and Foufoula-Georgiou, E., 1990. Regional rainfall frequency analysis via the stochastic storm transposition method. *J. Hydraul. Eng.*, 116, 859–880.
- Yevjevich, V., 1972. *Probability and statistics in hydrology*. Water Resources Publications, Fort Collins, Colorado, 276pp.