Hydrological Extremes: Understanding, Predicting, Mitigating (Proceedings of IUGG 99 Symposium HS1, Birmingham, July 1999). IAHS Publ. no. 255, 1999.

Regional frequency analysis: a new vocabulary

DUNCAN W. REED, DÖRTE JAKOB, ALICE J. ROBSON, DUNCAN S. FAULKNER & ELIZABETH J. STEWART
Institute of Hydrology, Wallingford, Oxfordshire OX10 8BB, UK
e-mail: d.reed@mail.nwl.ac.uk

Abstract Regional frequency analysis solves the problem of estimating the magnitude of extreme events by pooling information from observations at several sites. The topic is reviewed with particular reference to evolution in the last decade. It is suggested that the terminology of regional frequency analysis has lagged behind developments, and a new vocabulary is introduced. This is illustrated for the relevant flood frequency estimation developed for the *Flood Estimation Handbook.*

INTRODUCTION

Regional frequency analysis pools information from several sites in a region. Pooling combines data, such as annual maximum floods, to provide an estimate of the average behaviour in the region. The approach allows estimates to be made at ungauged sites, and reduces the sensitivity of frequency analyses at gauged sites to the limited sample of data typically available. Deconstruction suggests that the principles of regional frequency analysis are intricate, with several ideas interwoven. This review looks first at principles, before turning to the building blocks and vocabulary of regional frequency analysis.

REVIEW OF PRINCIPLES

A fundamental idea in regional frequency analysis is *to avoid undue extrapolation of a fitted distribution function*. This means that the region should be large enough to provide sufficient data to underpin estimation of the required rare event. In this paper, the return period for which estimates are required is termed the *target return period.* A second idea is that *there is similar extremal behaviour at sites in the region*. This implies that the region must be small enough—or very carefully chosen—to justify the assumption that extremes at the different sites are identically distributed, either directly or after a standard transformation. Such a region is said to be *homogeneous*.

There is obvious tension between these two ideas. If the target return period is relatively short (e.g. 20 years), it is helpful to choose a compact region. Especially for meteorological variables, this improves the prospect that the region will be homogeneous. However, if the target return period is long (e.g. 200 years), it is advisable to pool (i.e. average) data from a greater number of sites. Without this stabilizing effect, estimates at different sites, or of different durations, may be unreasonably different; for example, it is inconsistent if the estimated 200-year 1-day rainfall depth exceeds the corresponding 2-day depth. In some cases, it may be preferable to relax concerns about regional heterogeneity. An immediate consequence of this tension is that regions need to be of *flexible* size, so that they can be tailored to the problem posed.

Use of a compact region is, in one respect, not ideal. A geographically compact region increases scope for the analysis to be susceptible to period-of-record effects. Reed & Stewart (1991) illustrate how a regional frequency estimate can be highly sensitive to the occurrence of a widespread extreme event.

A third idea is that *there is prior information about suitable regions*. A notable physiographic feature, such as a range of hills of distinctive geology, can sometimes provide a natural region. However, a comprehensive regionalization—partitioning a whole province into regions—is more difficult to achieve, and the final choice is likely to be subjective. This will be based on perceptions of important hydrological differences within the province, but may also reflect variations in gauge density, and concerns to avoid regional boundaries dividing neighbouring sites for which estimates will be required.

Most prior knowledge relates to *unsuitable* regions. For example, it is well known that highly permeable catchments produce floods relatively reluctantly and episodically, because of the great control exerted by antecedent wetness conditions. This argues against any scheme that sets highly permeable and impermeable catchments in the same flood frequency region.

A fourth idea is that *regional frequency analysis helps to estimate extremal behaviour at ungauged sites*. Such cases occur frequently in practice, and it is easy to associate regional frequency analysis solely with solving the ungauged-site problem. This preoccupation may explain why there has been no strong drive amongst users to see fixed regions discarded in favour of flexible pooling schemes.

REVIEW OF BUILDING BLOCKS

This section presents a brief historical review of the building blocks of regional frequency analysis.

Pooling

An early example of regional frequency analysis is Bilham (1936), who presents a general formula for the frequency of short-duration extreme rainfalls. Bilham pools data from 12 rain-recorders in England and Wales. Subsequently, Holland (1964) found the formula generally adequate if restricted to rainfall frequency estimation at lowland sites. This condition was necessary because the formula does not expressly allow for some sites being systematically wetter than others.

Standardization

Standardization is the process by which data from different sites are made comparable: for example, to reflect that some sites are systematically wetter than others, and some catchments are larger than others. The *index flood procedure* of Dalrymple (1960) is an early example in which flood peaks at different sites are made comparable by dividing by a reference value called the *index variable*. Often, the mean of annual maxima is chosen as the index variable. The *Flood Estimation Handbook* (IH, 1999) prefers to use the median, *QMED*. This is a more robust measure than the mean, being less sensitive to particular extreme values in the annual maximum series. It is also helpful that *QMED* has a fixed return period of two years, when rarity is measured on the annual maximum scale.

Standardization greatly widens the applicability of pooled analyses. It is no longer necessary to assume that the distribution of extreme values is the same at all sites: something that is restrictive for rainfall data, and implausible for flood data. The advance gave birth to regional flood frequency analysis, allowing major generalizations of flood frequency such as that developed for the *Flood Studies Report* (NERC, 1975). Instead the assumption is made that the distribution of *standardized* extreme values is essentially identical at all sites in the region.

Hydrological regions

Mosley (1981) considers the delimitation of general-purpose hydrological regions. Gottschalk (1985) suggests that the formation of regions for flood frequency analysis can gain from joint studies of the regional variation of other (non-flood) hydrological variables.

L-moment methods

Based on research published in a series of reports and papers, Hosking & Wallis (1997) present a comprehensive treatment of regional frequency analysis by L-moment methods. L-moment statistics (such as the L-CV and L-skewness) are analogous to statistics based on conventional moments, but have important theoretical and practical advantages. Hosking & Wallis (1997) rationalize, describe and illustrate tools for regional frequency analysis. Techniques are presented to test for discordant sites, heterogeneity and goodness of fit. A discordant site is one that exhibits unusual behaviour in comparison to other sites. A heterogeneity test inspects the likely validity of the assumption that growth rates (i.e. extreme values after standardization by the index variable) are identically distributed at all sites. A goodness-of-fit test examines whether a given statistical distribution (e.g. the Generalized Extreme Value) fits the data adequately.

Pooling according to catchment properties

Acreman & Sinclair (1986) and Wiltshire (1986) were among the first to advocate that flood data be pooled according to catchment type rather than geographical position. Wiltshire's (1986) scheme classifies 376 stations by reference to catchment wetness, size, and urbanization: producing five groups of similar size. The classification by Acreman & Sinclair (1986) uses six descriptors to divide 168 catchments into five groups; 118 catchments were allocated to a single group.

Focused pooling

In contrast to earlier methods based on the prior construction of a small number of fixed groups, *focused pooling* constructs a grouping that is specifically designed for (and focused on) the site of interest. Acreman & Wiltshire (1989) suggest that flood data be pooled thus, with the grouping carried out in catchment-descriptor space. Reed & Stewart (1989) present a general method of rainfall frequency estimation in which data are pooled according to distance from the site of interest. Focused pooling is also a feature of the region-of-influence approach to flood frequency analysis (Burn, 1990).

More elaborate pooling schemes

Based on a suggestion by Reed (1994), Burn (1997) applies a focused pooling scheme in which catchment similarity is judged from flood seasonality. Preferably, measures of flood seasonality are based on data abstracted in peaks-over-threshold form (Bayliss & Jones, 1993).

The *Flood Estimation Handbook* (IH, 1999) presents a method in which the size of region is tailored to the target return period. The method uses a focused pooling scheme in which site similarity is initially judged in terms of catchment size, wetness and soils.

REVIEW OF VOCABULARY

Why a change in terminology is needed

In general parlance, a *region* denotes a contiguous zone identified by geographical or political boundaries. In applied hydrology, a region is often defined by hydrometric or operational boundaries. With respect to regional frequency analysis, the term is satisfactory when data are pooled from fixed regions but less suitable for focused pooling schemes. Although clear to the researcher, the word *region* confuses the practitioner who: (a) has previous experience of procedures based on fixed regions, (b) works for a regional agency, or (c) obtains data from archives arranged by measurement region.

In pooling schemes which group catchments according to hydrological similarity rather than geographical position, the scope for confusion is greater still. Such groups of sites are often geographically dispersed (see Fig. 1), and continued use of the word *region* simply cannot be justified.

A new vocabulary

To overcome these problems, the *Flood Estimation Handbook* (IH, 1999) breaks with traditional terminology. A scheme for choosing how to group sites for frequency analysis (formerly a *regionalization*) is termed a *pooling scheme*, while a region for pooled frequency analysis is called a *pooling group*. These terms communicate the concepts of *grouping* sites and *pooling* data, but are free from geographical connotation. The new vocabulary is summarized in Table 1.

Table 1 Terminology.

Traditional	Flood Estimation Handbook
Region	Pooling group
Regional frequency analysis	Pooled frequency analysis
Regional growth curve	Pooled growth curve
Regionalization	Pooling scheme

COMPARISON OF POOLING SCHEMES

Pooling schemes for UK flood data

Table 2 summarizes two pooling schemes for UK flood frequency estimation. These are the schemes presented in the *Flood Studies Report* (NERC, 1975) and the *Flood Estimation Handbook* (IH, 1999). The Flood Estimation Handbook (FEH) scheme adjusts the pooling-group size to include five times as many station-years of data as the target return period T. This *5T rule* is a compromise between the use of small groups (which are more likely to be homogeneous) and the use of large groups (which support estimates without undue extrapolation).

Table 2 Pooling schemes for UK flood data.

	Flood Estimation Handbook	*Flood Studies Report*
Domain of pooling	Size–wetness–soils space	Geographic space
Basis of site selection	Flexible: similarity of catchment size, wetness and soils	Fixed: according to hydrometric region
Size of pooling group	Chosen to provide $5T$ station-years of record, where T is target return period	Fixed

Example

Figure 1 illustrates pooled flood growth curves for the Ae Water at Ae Village, a 70 km^2 ungauged catchment in southern Scotland. The growth curve indicates the factor by which the index variable (*Q*MED) is to be multiplied to estimate a flood peak of given rarity. The FEH curve uses the Flood Estimation Handbook pooling scheme, while the FSR curve is derived when flood data are pooled according to the relevant Flood Studies Report region. In both cases, the pooled growth curve is derived as a Generalized Logistic (GL) distribution fitted by a weighted L-moment method. For the FSR pooling, weighting is by record length (Hosking & Wallis, 1997). For the FEH pooling, the weighting reflects both record length and similarity ranking (Jakob *et al.*, 1999). The FEH pooling group is sized (by the $5T$ rule) to meet a target return period of 100 years.

A heterogeneity measure, H_2, is used to assess the heterogeneity of the pooling groups; this summarizes inter-site variation in L-CV and L-skewness of the annual maximum floods. [H_2 corresponds to Hosking & Wallis's H statistic evaluated on their alternative measure of dispersion, V_2.] By this criterion, the FEH pooling group is judged to be acceptably homogeneous ($H_2 = 0.79 < 1.0$) while the FSR pooling group is strongly heterogeneous ($H_2 = 4.35 >> 2.0$). Part of the reason that the FEH pooling

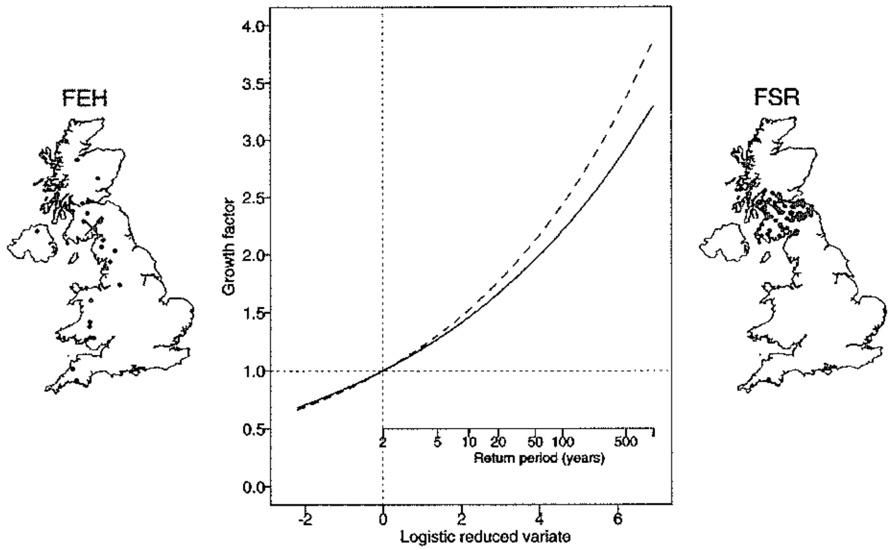

Fig. 1 Pooled flood growth curves for Ae Water at Ae Village in southern Scotland; the solid line is based on the FEH pooling group (locations inset to left), while the broken line is based on the FSR pooling group (locations inset to right).

group is better behaved is its smaller size (25 sites compared to 83 sites). If the FEH pooling group is expanded to embrace 1982 station-years of data—matching the number in the FSR pooling group—the H_2 value is 1.95. This is close to the test value of 2.0 which Hosking & Wallis (1997) suggest for practical use in distinguishing possible heterogeneity from definite heterogeneity, but is still much less heterogeneous than the FSR pooling group ($1.95 \ll 4.35$). Thus, the FEH pooling scheme is considered to outperform the FSR scheme for the Ae Water example.

CONCLUDING REMARKS

No changes were made in the above example to the pooling groups derived according to the basic FEH and FSR pooling schemes (Table 2). The *Flood Estimation Handbook* does, however, encourage users to examine the selected catchments carefully. It is recommended that catchments known to be strongly dissimilar to the subject catchment (e.g. in an important feature such as reservoir exploitation) are deleted from the group. However, the FEH advises that catchments that are discrepant only in terms of their L-moment ratios should be retained in the pooling group. Such catchments may appear discordant simply because they have experienced an unusual number of extreme floods during the period of gauged record.

A new vocabulary has been introduced for regional frequency analysis, and illustrated for the FEH pooling scheme. The scheme pools catchment flood data according to size–wetness–soils similarity, and adjusts the size of pooling group to reflect the return period for which estimates are required. A refinement of the procedure might further adjust the size of pooling group to reflect the degree of site similarity found.

For applications such as flood frequency estimation, it is suggested that pooling schemes based on hydrological similarity are more relevant than pooling schemes based on geographical regions. Where possible, studies should consider data from all sites or catchments that exhibit similarity, without reference to regional or national boundaries.

Acknowledgements The Flood Estimation Handbook research programme was funded by the Ministry of Agriculture Fisheries and Food, the Environment Agency, the Department of Agriculture Northern Ireland, and a consortium led by the Scottish Office. All are UK organizations. The authors are grateful to co-workers in regional frequency analysis for helpful discussions, in particular: David Jones, Jon Hosking and Don Burn. The new vocabulary, suggested by David, meets a requirement identified by Linda Aucott and Andy Pepper.

REFERENCES

Acreman, M. C. & Wiltshire, S. E. (1989) The regions are dead: long live the regions. Methods of identifying and dispensing with regions for flood frequency analysis. In: *FRIENDS in Hydrology* (ed. by L. Roald, K. Nordseth & K. A. Hassel), 175–188. IAHS Publ. no. 187.

Acreman, M. C. & Sinclair, C. D. (1986) Classification of drainage basins according to their physical characteristics, and application for flood frequency analysis in Scotland. *J. Hydrol.* **84**(3), 365–380.

Bayliss, A. C. & Jones, R. C. (1993) Peaks-over-threshold flood database: summary statistics and seasonality. *Report no. 121, Institute of Hydrology, Wallingford, UK.*

Bilham, E. G. (1936) Classification of heavy falls in short periods. In: *British Rainfall 1935*, 262–280. HMSO, London.

Burn, D. H. (1990) Evaluation of regional flood frequency analysis with a region of influence approach. *Wat. Resour. Res.* **26**, 2257–2265.

Burn, D. H. (1997) Catchment similarity for regional flood frequency analysis using seasonality measures. *J. Hydrol.* **202**, 212–230.

Dalrymple, T. (1960) Flood frequency analyses. *Wat. Supply Paper 1543-A, US Geol. Survey, Reston, Virginia.*

Gottschalk, L. (1985) Hydrological regionalization of Sweden. *Hydrol. Sci. J.* **30**(1), 65–83.

Holland, D. J. (1964) Rain intensity frequency relationships in Britain. *Hydrol. Memo. 33, Meteorological Office, Bracknell, UK.*

Hosking, J. R. M. & Wallis, J. R. (1997) *Regional Frequency Analysis: An Approach Based on L-moments.* Cambridge Univ. Press, Cambridge, UK.

IH (Institute of Hydrology) (1999) *Flood Estimation Handbook* (five volumes). Institute of Hydrology, Wallingford, UK.

Jakob, D., Robson, A. J. & Reed, D. W. (1999) Choosing a pooling-group. Chapter C6 in: *Statistical Procedures for Flood Frequency Estimation, Flood Estimation Handbook*, vol. 3. Institute of Hydrology, Wallingford, UK.

Mosley, M. P. (1981) Delimitation of New Zealand hydrologic regions. *J. Hydrol.* **49**(1), 173–192.

NERC (Natural Environment Research Council) (1975) *Flood Studies Report* (five volumes). Natural Environment Research Council, London.

Reed, D. W. (1994) Plans for the Flood Estimation Handbook. In: *Proc. MAFF Conf. of River and Coastal Engineers* (Loughborough, July 1994), 8.3.1–8.3.8. Ministry of Agriculture Fisheries and Food, London.

Reed, D. W. & Stewart, E. J. (1989) Focus on rainfall growth estimation. In: *Proc. 2nd National Hydrol. Symp.* (Sheffield, September 1989), 3.57–3.65. British Hydrological Society.

Reed, D. W. & Stewart, E. J. (1991) Discussion on dam safety: an evaluation of some procedures for design flood estimation. *Hydrol. Sci. J.* **36**(5), 499–502.

Wiltshire, S. E. (1986) Identification of homogeneous regions for flood frequency analysis. *J. Hydrol.* **84**(3), 387–402.