CHAPTER 2 – AT-SITE FLOOD FREQUENCY ANALYSIS

George Kuczera and Stewart Franks

School of Engineering, University of Newcastle

1. INTRODUCTION

Flood frequency analysis refers to procedures that use recorded and related flood data to select and fit a probability model of flood peaks at a particular location in the catchment. The flood probability model can then be used to perform risk-based design and flood risk assessment and to provide input to regional flood estimation methods.

This chapter considers the issues and techniques useful when undertaking a flood frequency analysis using flood data at a single site. It represents an evolutionary update of Chapter 10 of Australian Rainfall and Runoff (Pilgrim and Doran, 1987) – where appropriate the original contribution by Pilgrim and Doran has been retained. The major changes include introduction of non-homogeneous probability models to account for long-term climate variability, abandonment of product log-moments with the log-Pearson III distribution, use of Bayesian methods to make better use of available flood information (such as censored flow data, rating error and regional information) and reduced prescription about the choice of flood probability model.

The primary purpose of this chapter is to present guidelines on performing flood frequency analyses. Often judgment will need to be exercised. To inform such judgments, this chapter describes the key conceptual foundations that underpin flood frequency analysis – the reader will need an understanding of elementary probability theory and statistics to get maximum benefit. In addition a number of worked examples are provided to provide deeper insight with the implied caveat that the examples are not exhaustive in their scope While it is expected that most users will use software written by others to implement the methods described in this chapter, sufficient information is provided to enable users to develop their own software applications.

2. CONCEPTUAL FRAMEWORK

2.1 Definition of Flood Probability Model

In flood frequency analysis flood peaks are considered to be random variables. Following convention the random variable denoting the flood peak is denoted by an upper-case symbol (e.g., Q) whereas a specific realization (or sample) is denoted by the lower-case symbol (e.g., q) – where there is no ambiguity lower-case symbols will be used.

It is assumed that each realization q is statistically independent of other realizations. This is the standard assumption in flood frequency analysis and is believed to be widely applicable [e.g., Stedinger et al., 1993].

In its most general form, the flood probability model can be described by its probability density function (pdf) $p(q|\theta,x,M)$ where q is the flood peak. The pdf of q is determined by the vector of parameters θ belonging to the probability distribution family M and by x defined as the vector of exogenous or external variables which affect the values of θ. The notation "|" refers to conditioning: the variables to the left of "|" depend on the values taken by variables to the right of "|".

The distribution function of Q is defined as the non-exceedance probability $P(Q \le q)$ and is related to the pdf by

$$P(Q \le q \mid \theta, x, M) = \int_0^q p(z \mid \theta, x, M)dz \qquad (1)$$

Empirically the pdf of q is the limiting form of the histogram of q as the number of samples approaches infinity. Importantly, eqn (1) shows that the area under the pdf is interpreted as probability.

Homogeneous flood probability model

The simplest form of the flood probability model arises when the parameter vector θ does not depend on an exogenous vector x. In that case each flood peak is considered to be a random realization from the same probability model $p(q|\theta,M)$. Under this assumption flood peaks form a homogeneous time series.

Non-homogeneous flood probability model

A more complicated situation arises when flood peaks do not form a homogeneous time series. This may arise for a number of reasons including the following:

- Rainfall and flood mechanisms may be changing over time. For example, long-term climate change due to global warming, land use change and river regulation may render the flood record non-homogeneous.
- Climate may experience pseudo-periodic shifts that persist over periods lasting from several years to several decades. There is growing evidence that parts of Australia are subject to such forcing and that this significantly affects flood risk [for example, Warner and Erskine, 1988; Harvey et. al, 1991; Franks and Kuczera, 2002; Kiem et al., 2003; Micevski et al., 2003]

Although this chapter will provide some guidance on non-homogeneous flood probability models it needs to be stressed that this is an area of continuing research and, therefore, users are therefore advised to keep abreast of new developments.

2.2 Flood Risk Perspectives

Flood frequency analysis deals with the probability distribution of significant flood peaks. Throughout the year, there are typically many flood peaks associated with individual storm events - this is illustrated in Figure 1 which presents a time series plot of a streamflow discharge. There are two ways of describing the probability of exceeding a significant flood magnitude:

1. The probability distribution of the largest flood peak occurring over a particular interval of time, which, in practice, is one year.
2. The probability distribution of the time between consecutive flood peaks that exceed a particular magnitude.

These two perspectives are intimately connected as the following exposition will show. Central to this connection is the annual maximum series, obtained by extracting the largest flood peak in every year of the record, and the peak-over-threshold (POT) series, obtained by extracting independent flood peaks above some threshold discharge.

Referring to Figure 1 let the random variable q be a local peak discharge defined as a discharge which has lower discharge on either side of the peak. This presents an immediate problem as any bump on the hydrograph would produce a local peak. To circumvent this problem we focus on peak flows greater than some threshold discharge defined as q_o. The threshold is selected so that the peaks above the threshold are sufficiently separated in time to be statistically independent of each other.

Suppose over a time interval of length T there are n peaks over the threshold q_o. This defines the POT time series of n independent realizations $\{q_1,\ldots,q_n\}$.

Let w be the maximum value in the POT time series; that is,

$$w = \max\ \{q_1,\ldots,q_n\} \tag{2}$$

For w to be the maximum value each observed peak must be less than or equal to w. In probability theory this condition is expressed by the joint event consisting of the intersection of the following n events

$$\{(q_1 \le w) \cap (q_2 \le w) \cap .... \cap (q_n \le w)\}$$

Because the peaks are assumed to be statistically independent the probability of the joint event is the product of the probabilities of the individual events. Therefore the probability that the random variable $W \le w$ in a POT series with n events occurring over the interval T simplifies to

$$\begin{aligned} P(W \le w \mid n,T) &= P\left[(q_1 \le w) \cap .... \cap (q_n \le w)\right] \\ &= P(q_1 \le w)P(q_2 \le w)....P(q_n \le w) \\ &= P(q \le w)^n \end{aligned} \tag{3}$$

The last term in eqn (3) comes from the assumption that all the peaks above the threshold q_o are sampled from the same distribution with the pdf $p(q|q>q_o)$.

The number of POT events n occurring over an interval T is random. Suppose that the random variable n follows a Poisson distribution with ν being the average number of POT events per unit time; that is,

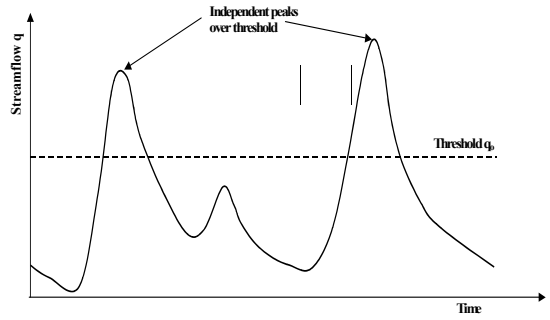

Figure 1. Peak-over-threshold series.

$$P(n \mid \nu,T) = \frac{(\nu T)^n \exp(-\nu T)}{n!}, n = 0,1,2,.... \tag{4}$$

Eqn (4) can be applied to cases where the distribution of POT occurrences varies according to the seasons encountered within time interval T. Provided the distribution of POT occurrences in each season is Poisson, the distribution of POT occurrences over interval T is Poisson with the parameter νT equal to the sum of the average number of occurrences in each season – this result is a consequence of the fact that the sum of Poisson random variables is Poisson distributed.

After some algebra, application of the total probability theorem yields the distribution of the largest flood peak magnitude over the interval with duration T

$$\begin{aligned} P(W \le w \mid T) &= \sum_{n=0}^{\infty} P(W \le w \mid n,T)\,P(n \mid \nu,T) \\ &= \exp\left[-(\nu T)\,P(Q > w)\right] \end{aligned} \tag{5}$$

where $P(W \le w|T)$ is the probability that the largest flood peak over time interval T is less than or equal to w. This result hinges on the assumption that all the peaks above the threshold q_o are sampled from the same distribution. If the pdf $p(q|q>q_o)$ exhibits significant seasonal differences eqn (3) cannot be simplified necessitating a more involved analysis.

2.2.1 Distribution of time between floods

The objective is to derive the probability distribution of the time between consecutive flood peaks with magnitude in excess of w. With regard to eqn (5), if the largest flood peak during time T is less than or equal to w, then the time to the next peak with magnitude in excess of w must be greater than T. It therefore follows from eqn (5) that the distribution of time between flood peaks with magnitude exceeding w is

$$\begin{aligned} &P(\text{Time to next peak exceeding } w \le T) \\ &\quad = 1 - \exp\left[-\nu P(Q > w)T\right] \end{aligned} \tag{6}$$

This is recognized as an exponential distribution with parameter $\nu P(Q>w)$ which is interpreted as the expected number of peaks exceeding w per unit time. An important property of the exponential distribution is that its expected value is equal to the inverse of its parameter. It therefore follows that the expected time interval between peaks that exceed w is

$$T_P(w) = \frac{1}{\text{Expected number of peaks} > w \text{ per unit time}} = \frac{1}{\nu P(Q > w)} \quad (7)$$

$T_P(w)$ is termed the average recurrence interval (ARI) for magnitude w and provides a convenient probability measure for POT series.

2.2.2 Annual maximum flood risk

With regard to eqn (5) $P(W \le w|T=1)$ represents the probability that the largest flood peak during the year is less than or equal to w. $1 - P(W \le w|T=1)$ is referred to as the annual exceedance probability AEP(w) and provides a convenient probability measure for annual maximum series. It is understood that the 1 in Y AEP flood has a probability of 1/Y of being exceeded in any one year.

An alternative measure is based on the average recurrence interval of annual maximum flood peaks exceeding magnitude w. The probability of waiting n years for the next annual maximum flood to exceed w can be shown to be

$$P(\text{Annual maximum peak exceeds w in } n^{th} \text{ year}) = P(W \le w \mid T = 1)^{n-1} P(W > w \mid T = 1) \quad (8)$$

This is known as the geometric distribution. The expected number of years between annual maximum flood peaks with magnitude in excess of w can be shown to be

$$T_A(w) = \frac{1}{1 - P(W \le w \mid T = 1)} = \frac{1}{AEP(w)} \quad (9)$$

2.2.3 Choice of series

It is important to distinguish between T_P and T_A – they have different meanings even though they are referred to as average recurrence intervals. From eqns (5), (7) and (9) it follows that the relationship between $T_P(w)$ and $T_A(w)$ is

$$T_A(w) = \frac{1}{1 - \exp\left(-\frac{1}{T_P(w)}\right)} \quad (10)$$

Figure 2 illustrates this relationship. For an ARI of 10 or more years the difference between T_A and T_P is minimal - POT and annual maximum analyses should yield similar results. However, for ARIs less than 10 years, the ARI for annual maximum series will exceed the ARI for the POT series for the same discharge. This arises because in years with large annual maximum flood the smaller floods of that year may exceed the annual maximum flood in other years.

These considerations make it essential when quoting ARIs to make clear whether they refer to annual maximum or POT series. The annual maximum and POT ARIs refer to different properties of the flood time series. This motivates the convention used in Australian Rainfall and Runoff, namely AEPs are used when referring to annual maximum series , while ARIs are used when referring to POT series.

Figure 2 leads to the following guidelines:

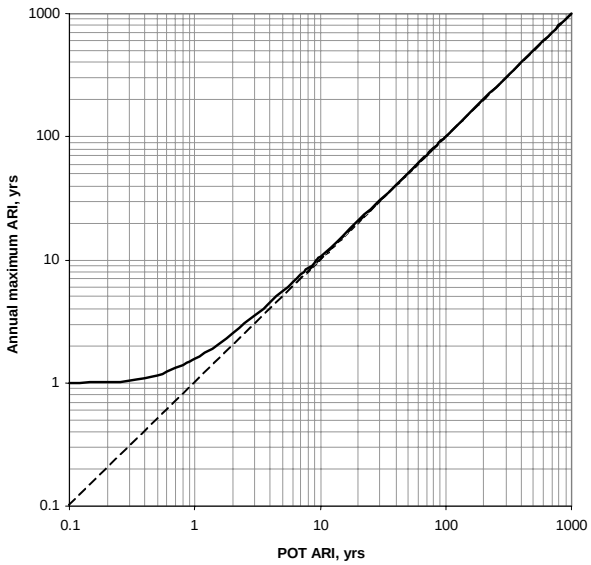

Figure 2. Relationship between POT and annual maximum average recurrence intervals.

(i) T_A > 10 years

Use of annual maximum series is generally preferred because it is straightforward to ensure statistical independence between annual maxima and allows the analysis to better focus on low AEP events. This series is generally used in design, as low AEPs in this range are generally required for estimation of a design flood for a structure or works at a particular site.

(ii) T_A < 10 years

Use of POT series is generally preferred because all floods are of interest in this range, whether they are the highest in the particular year of record or not. The annual maximum series may omit many floods of interest. The POT series is appropriate for estimating design flows of low ARI in urban stormwater contexts including water quality treatment devices and measures and for diversion works, coffer dams and other temporary structures. However, it is infrequently used because flow records are not often available at sites where minor works with a design ARI of less than 10 years are required. A significant application is in the development of regional flood estimation methods for small to medium sized catchments (e.g. Pilgrim and McDermott, 1982; Flavell, 1983; Adams and McMahon, 1985; and Adams, 1987). It should also be noted that the design rainfall data in Book II Section 1 effectively represent the results of frequency analysis of POT series rainfall data.

2.3 Advantages and Disadvantages of Flood Frequency Analysis

Users need to be aware of the advantages and disadvantages of flood frequency analysis.

Flood peaks are the product of a complex joint probability process involving the interaction of many random variables associated with the rainfall event, antecedent conditions and rainfall-runoff transformation. Peak flood records represent the integrated response of the storm event with the

catchment. They provide a direct measure of flood exceedance probabilities. As a result flood frequency analysis is less susceptible to bias, possibly large, that can affect alternative methods based on design rainfall [Kuczera et al., 2003].

Other advantages of flood frequency analysis include its comparative simplicity and capacity to quantify uncertainty arising from limited information. It remains a moot point whether flood frequency methods are less accurate than rainfall-based methods for which rigorous uncertainty analysis is yet to be developed.

Offsetting these significant advantages are several disadvantages:

- The true probability distribution family M is unknown. Unfortunately, different models can fit the bulk of the flood data with similar capability, yet can diverge in the right hand tail when extrapolated beyond the data and have difficulty dealing with the left hand tail.
- Short records may produce flood estimates with considerable uncertainty. However, the ready ability to compute confidence limits directly informs the user about the credibility of the estimate.
- It may be difficult or impossible to adjust the data if the catchment conditions under which the flood data were obtained have changed during the period of record, or are different to those applying to the future design life of a structure or works being designed.
- Considerable extrapolation of rating curves is necessary to convert recorded stage to discharge for the largest flood peaks at most Australian gauging stations. Also the probability of malfunction of recording instruments is increased during major floods. Suspect floods and the years in which they occurred may be omitted in analysis of annual maximum series, but this reduces the sample size and may introduce bias if the suspect floods are all major events. Though these problems are inherent to the calibration of all methods employing major flood peaks, flood frequency analysis is more sensitive because of its need to use major flood peaks.

2.4 Range of Application

As noted, the true flood probability family is unknown. In practice the choice of model is guided by goodness of fit to data, understanding of the hydrologic characteristics of the region, and regional hydrologic experience. Therefore, use of the fitted frequency curve for ARIs up to that of the data is regarded as an interpolation exercise deemed to be reliable in the sense that confidence limits capture the uncertainty. However, when the frequency curve is extrapolated well beyond the observed data, confidence limits which quantify the effect of sampling variability on parameter uncertainty may underestimate the true uncertainty - model bias may be significant and even dominant. Example 1 demonstrates the need to understand the processes affecting flood peaks beyond the observed record and illustrates the pitfall of blind extrapolation.

Large extrapolations of flood frequency analyses are not recommended. It is acknowledged that prescribing strict limits on maximum ARIs does not have a strong conceptual foundation. The limits to extrapolation should be guided by consideration of confidence limits, which are affected by the information content of the data and choice of flood model, and by judgments about model bias which cannot be quantified. In situations where the analyst is prepared to make the judgment that the processes operating in the range of the observed record continue to dominate for larger floods, model bias may be deemed to be manageable – of course the effects of sampling uncertainty may be so amplified for significant extrapolation to render the frequency estimate of little value.

In the absence of user analysis about the degree of model bias when extrapolating, the following guidelines are offered to promote consistency with previous practice: As discussed further in Book VI Section 1.2, the 1 in 100 AEP flood is the largest event that should be estimated by direct frequency analysis for important work. The maximum flood that should be estimated by this means under any circumstances is the 1 in 500 AEP event. Where a regional flood frequency method transfer of data from an adjacent catchment is used, the limiting AEP should be 1 in 100. Book VI Section 1 describes procedures for estimating floods beyond the probabilities noted above. These procedures interpolate flood magnitudes between the 1 in 100 AEP event and the probable maximum flood. While these procedures involve some arbitrary assumptions, they provide a consistent approach and overcome many of the problems involved in extrapolation of flood frequency analysis resulting from choice of the appropriate probability distribution.

3. SELECTION AND PREPARATION OF DATA

3.1 Requirements of Data for Valid Analysis

For valid frequency analysis, the data used should constitute a random sample of independent values, ideally from a homogeneous population. Streamflow data are collected as a continuous record, and discrete values must be extracted from this record as the events to be analyzed. The problem of assessing independence of events, and of selecting all independent events, is illustrated by the streamflow record for a 1000 km^2 catchment in Figure 3. There is little doubt that peaks A and B are not independent or that they are serially correlated, while peak D is independent of A and B. However, the independence of peak C from A and B is open to question, and there is doubt as to whether the independent peaks in the record are B and D, or B, C and D. Methods for selecting the peaks to be included in the analysis are described in the following subsections.

Lack of homogeneity of the population of floods is also a practical problem, particularly as the data sample from the past is used to derive flood estimates applicable to the design life of the structure or works in the future. Examples of changes in the collection of the data or in the nature of the catchment that lead to lack of homogeneity are:

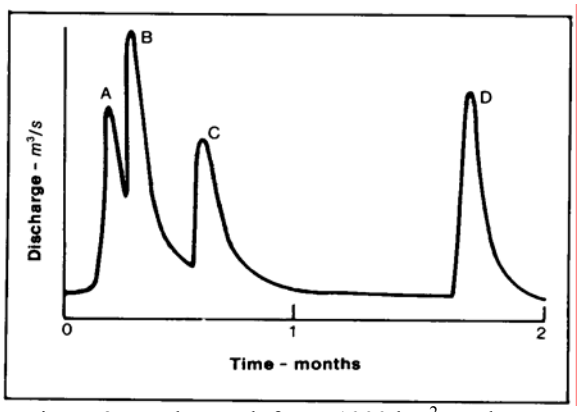

Figure 3. Hydrograph for to 1000 km^2 catchment illustrating difficulty of assessing independence of floods.

1. undetected change in station rating curve;
2. change of gauging station site;
3. construction of large storages, levees and channel improvements; and
4. changes in land use such as clearing, growth in the number of farm dams on the catchment, different farming practices, soil conservation works, reafforestation, and urbanization.

The record should be carefully examined for these and other causes of lack of homogeneity. In some cases recorded values can be adjusted by means such as routing pre-dam floods through the storage to adjust them to equivalent present values, correcting rating errors where this is possible, or making some adjustment for urbanization. Such decisions must be made largely by judgment. As with all methods of flood estimation, it is important that likely conditions during the design life be considered rather than those existing at the time of design. Some arbitrary adjustment of derived values for likely changes in the catchment may be possible, but the recorded data must generally be accepted for analysis and design. Fortunately, the available evidence indicates that unless changes to the catchment involve large proportions of the total area or large changes in the storage on the catchment, the effects on flood magnitudes are likely to be modest. Also, the effects are likely to be larger for small floods than for the large floods that are of primary interest in design.

3.2 Types of Flood Data

It is convenient to classify the flood peak data used in a flood frequency analysis as either being gauged or censored.

3.2.1 Gauged data

Gauged data consist of a time series of flood discharge estimates. Such estimates are based on observed peak (or instantaneous) stages (or water levels). A rating curve is used to transform stage observations to discharge estimates. When extrapolated, the rating curve can introduce large systematic error into discharge estimates.

It is important to check how the peak discharges were obtained from the gauged record. Peak discharges may be derived from daily readings, possibly with some intermediate readings during some floods for part of the record, and continuous readings from the remainder of the record. If the daily reading is deemed an unreliable estimate of the peak discharge during that day, the reading need not be discarded but treated as an imprecise measurement. Micevski et al. (2005) present a likelihood-based method for dealing with such cases. They also examine the consequences of ignoring the error associated with daily readings.

3.2.2 Censored data

Censored data refer to peak flows which are treated as being above or below a known threshold. They can be expressed as a time series of indicator values defined as

$$I_t(q) = \begin{cases} 1 \text{ if the } t^{th} \text{ flood peak} > \text{threshold discharge } q \\ -1 \text{ if the } t^{th} \text{ flood peak} \le \text{threshold discharge } q \end{cases}$$

They arise in a number of ways. For example, prior to gauging, water level records may have been kept only for large floods above some perception threshold. Therefore, all we may know is that there were n_a flood peaks above the threshold and n_b peaks below the threshold.

Sometimes, we may deliberately exclude zero or small gauged floods below some threshold because the overall fit is unduly influenced by small floods. In such cases, even though the gauged flows are available, they are treated as censored data.

3.3 Annual Flood Gauged Series

This is the most common method of selecting the floods to be analyzed. The series is comprised of the highest instantaneous rate of discharge in each year of record. The year may be a calendar year or a water year, the latter usually commencing at the end of the period of lowest average flow during the year. Where flows are highly seasonal, especially with a wet summer, use of the water year is preferable. The highest flow in each year is selected whether it is a major flood or not, and all other floods are neglected, even though some will be much larger than the maximum discharges selected from some other years. For N years of data, the annual flood series will consist of N values.

The annual flood series has at least two advantages:

1. As the individual annual maximum flows are likely to be separated by considerable intervals of time, it is probable that the values will be independent. Checking of dates of the annual maxima to ensure that they are likely to be independent is a simple procedure that should always be carried out. If the highest annual value occurred at the start of a year and was judged to be strongly related to the annual maximum at the end of the previous year, the lower of these two values should be discarded, and the second highest flow in that year substituted.
2. The series is easily and unambiguously extracted. Most

data collection agencies have annual maxima on computer file and/or hard copy.

3.4 Peak-Over-Threshold Gauged Series

A POT flood series consists of all floods with peak discharges above a selected base value, regardless of the number of such floods occurring each year. The POT series may also be termed the partial duration series or basic stage series. The number of floods K generally will be different to the number of years of record N, and will depend on the selected theshold discharge. The American Society of Civil Engineers (1949) recommended that the base discharge should be selected so that K is greater than N, but that there should not be more than 3 or 4 floods above the threshold in any one year. These two requirements can be incompatible. The U.S. Geological Survey (Dalrymple, 1960) recommended that K should equal 3N. If a probability distribution is to be fitted to the POT series the desirable threshold discharge and average number of floods per year selected depend on the type of distribution. These distributions are discussed further in Section 4.3. For the compound model using a Poisson distribution of occurrences and an exponential distribution of magnitudes, Tavares and da Silva (1983) and Jayasuriya and Mein (1985) found that K should equal 2N or greater, and the U.K Flood Studies Report (Natural Environment Research Council, 1975) recommended that K should equal 3N to 5N. For fitting the log Pearson III distribution, the values of the moments depend on the number of floods selected and the base discharge. McDermott and Pilgrim (1982) and Jayasuriya and Mein (1985) found that best results were obtained in this case when K equalled N. Martins and Stedinger (2001) found that the precision of flood quantiles derived from a GEV-Poisson POT model is fairly insensitive for $K \geq N$.

An important advantage of the POT series is that when the selected base value is sufficiently high, small events that are not really floods are excluded. With the annual series, non-floods in dry years may have an undue influence on shape of the distribution. This is particularly important for Australia, where both the range of flows and the non-occurrence of floods are greater than in many other countries such as the United States and the United Kingdom [Grayson et al., 1996]. For this reason it would also be expected that the desirable ratio of K to N would be lower in Australia than in these countries.

A criterion for independence of successive peaks must also be applied in selecting events. As discussed by Laurenson (1987), statistical independence requires physical independence of the causative factors of the flood, mainly rainfall and antecedent wetness. This type of independence is desirable if the POT series is used to estimate the distribution of annual floods. On the other hand, selection of POT series floods for design flood studies should consider the consequences of the flood peaks in assessing independence of events where damages or financial penalties are the most important design variables. Factors to be considered might include duration of inundation, and time required to repair flood damage. In both cases, the size or response time of the catchment will have some effect.

The decision regarding a criterion for independence therefore requires subjective judgment by the designer or analyst in each case. There is often some conflict in that some flood effects are short-lived, perhaps only as long as inundation, while others such as the destruction of an annual crop may last as long as a year. It is thus not possible to recommend a simple and clear-cut criterion for independence. The circumstances and objectives of each study, and the characteristics of the catchment and flood data, should be considered in each case before a criterion is adopted. It is inevitable that the adopted criterion will be arbitrary to some extent.

While no specific criterion can be recommended, it may be helpful to consider some criteria that have been used in past studies:

- Bulletin 17B of the Interagency Advisory Committee on Water Data (1982) states that no general criterion can be recommended and the decision should be based on the intended use in each case, as discussed above. However in Appendix 14 of that document, a study by Beard (1974) is summarised where the criterion used is that independent flood peaks should be separated by five days plus the natural logarithm of the square miles of drainage area, with the additional requirement that intermediate flows must drop to below 75% of the lower of the two separate flood peaks. This may only be suitable for catchments larger than 1000 km^2. Jayasuriya and Mein (1985) used this criterion.
- The UK Flood Studies Report (Natural Environment Research Council, 1975) used a criterion that flood peaks should be separated by three times the time to peak and that the flow should decrease between peaks to two thirds of the first peak.
- McIllwraith (1953), in developing design rainfall data for flood estimation, used the following criteria based on the rainfall causing the floods:
    - for rainfalls of short duration up to two hours, only the one highest flood within a period of 24 hours.
    - for longer rains, a period of 24 hours in which no more than 5 mm of rain could occur between rain causing separate flood events.
- In a study of small catchments, Potter and Pilgrim (1971) used a criterion of three calendar days between separate flood events but lesser events could occur in the intervening period. This was the most satisfactory of five criteria tested on data from seven small catchments located throughout eastern New South Wales. It also gave the closest approximation to the above criteria used by McIllwraith (1953).
- Pilgrim and McDermott (1982) and McDermott and Pilgrim (1983) adopted monthly maximum peak flows to give an effective criterion of independence in developing a design procedure for small to medium sized catchments. This was based primarily on the assumption that little additional damage would be

caused by floods occurring within a month, and thus closer floods would not be independent in terms of their effects. This criterion was also used by Adams and McMahon (1985) and Adams (1987).

The criteria cited above represent a wide range, and illustrate the difficult and subjective nature of the choice. It is stressed that these criteria have been described for illustrative purposes only. In each particular application the designer or analyst should choose a criterion suitable to the analysis and relevant to all of the circumstances and objectives.

3.5 Monthly and Seasonal Gauged Series

In some circumstances, series other than the annual or POT series may be used. The monthly and seasonal series are the most useful.

Maximum monthly flows are an approximation to the POT series in most parts of Australia, as the probability of two large independent floods occurring in the same month is low - tropical northern Australia, the west coast Tasmania and the south west of Western Australia may be exceptions. While the monthly series is easier to compile than a POT series (most gauging authorities have monthly maximum flows on file) consideration needs to be given to significance of multiple floods in a month causing adverse events. It should be noted that not every monthly maximum flood will be selected, but only those large enough to exceed a selected base discharge, as is the case for the POT series. Care is required to check any floods selected in successive months for independence. Where the dates are close, the lower value should be discarded. The second highest flood in that month could then be checked from the records, but this would generally not be worthwhile. An example of use of the monthly series is described by Pilgrim and McDermott (1982).

Seasonal flood frequencies are sometimes required. For these cases, the data are selected for the particular month or season as for the annual series, and the flood frequency analysis is carried out in a similar fashion to that for the annual series.

3.6 Extension of Gauged Records

It may sometimes be possible to extend the recorded data by values estimated from longer records on adjacent catchments, by use of a catchment rainfall-runoff model, or by historical data from before the commencement of records. If this can be done validly, the effective sample size of the data will be increased and the reliability of the analysis will be greater. However, care is necessary to ensure that the extended data are valid, and that real information has been added. Several procedures can be used.

3.6.1 Extension using data from an adjacent catchment

Suppose there are n_1 years of record at site 1, the study catchment, and a longer record of length n_2 years at site 2, an adjacent catchment. Using the overlapping record a relationship can be developed between q_1 and q_2, the peak flows at sites 1 and 2, and then used to extend the shorter record at site 1.

The most intuitive way to extend the site 1 record is to use the best fit relationship between q_1 and q_2 to obtain a smoothed estimate of q_1 given the observed q_2. The best fit would usually be determined using regression methods. However, this approach is not recommended - it biases downward the variance of the extended record because the scatter about the regression line is ignored.

This bias can be overcome by one of two methods:

1. A random error sampled from the regression error distribution can be added to the smoothed estimate for q_1 [Matalas and Jacobs, 1964]. While this will preserve the expected variance of the extended record it produces an arbitrary extended record.

2. A maintenance of variance extension (MOVE) method can be used to derive a relationship between q_1 and q_2 which preserves the expected value of selected statistics for q_1 such as the variance. For example, the simplest MOVE method preserves the mean and variance of q_1 using the following relationship

$$q_1 = E(q_1) + \text{sgn}(r)\sqrt{\frac{\text{Var}(q_1)}{\text{Var}(q_2)}}[q_2 - E(q_2)]$$

where sign(r) is the correlation between q_1 and q_2 and E() and Var() are the mean and variance. For more details the reader should consult Hirsch (1982) and Grygier et al. (1989) in the case of extension of multiple records.

When annual floods are used, the dates of the corresponding annual floods may be different resulting in a lack of physical basis for the relation. It is a moot point whether this constitutes grounds for not using such data. After all, the purpose of record extension is to exploit the association between the annual peaks at the two sites – the existence of a causal linkage is not a prerequisite.

Wang (2001) presents a Bayesian approach that exploits the useful information arising from the association of peak flows between sites 1 and 2. Though not a record extension method, it does augment the information beyond that in the site 1 record.

3.6.2 Use of a catchment rainfall-runoff model

Suppose a rainfall record longer than the flow record is available at the study catchment. A rainfall-runoff model (ranging from a simple rainfall-runoff regression to a continuous soil moisture accounting model) can be calibrated and used to extend the flow record. As discussed in the previous section such an approach is likely to produce smoothed flow estimates which bias downwards the variability of the flow record. The use of smoothed estimates is therefore not recommended. If a rainfall-runoff approach is to be used the variability needs to be preserved, for example, by sampling from the error distribution describing the discrepancy between observed and simulated

flows. This error distribution can be complex with the magnitude of errors growing with the magnitude of the flow. Care needs to be exercised to ensure that such error characteristics are replicated in the sampling scheme.

3.6.3 Station-year method

This is included only to warn against its shortcomings. In this procedure, records from several adjacent catchments are joined "end-to-end" to give a single record equal in length to the sum of the lengths of the constituent records. As discussed by Clarke-Hafstad (1942) for rainfall data, spatial correlation between the records of the adjacent stations invalidates the procedure.

3.7 Rating Error in Gauged Flows

Though it is widely accepted that discharge estimates for large floods can be in considerable error, there is limited published information on these errors and how they can be allowed for in a flood frequency analysis. Rating error can arise from a number of mechanisms:

1. For large floods the rating curve typically is extrapolated or fitted to indirect discharge estimates. This can introduce a systematic but unknown bias.
2. If the gauging station is located at a site with an unstable cross section the rating curve may shift causing a systematic but unknown bias.

The conceptual model of rating error presented in this section is based on Kuczera (1999) and is considered to be rudimentary and subject to refinement. It is assumed the cross section is stable with the primary source of rating error arising from extension of the rating curve to large floods.

Potter and Walker (1981, 1985) observe that flood discharge is inferred from a rating curve which is subject to discontinuous measurement error. Consider Figure 4 which depicts a rating curve with two regions having different error characteristics. The interpolation zone consists of that part of the rating curve well defined by discharge-stage measurements; typically the error coefficient of variation (CV) would be practically negligible, say 5%. In the extension zone the rating curve is extended by methods such as slope-conveyance, log-log extrapolation or fitting to indirect discharge estimates. Typically such extensions are smooth and, therefore, can induce systematic under- or over-estimation of the true discharge over a range of stage - in Figure 4 the illustrated extension systematically underestimates the true discharge. The extension zone is considerable. Hatfield and Muir (1984) report that the highest gauged flow at over 50% of NSW stations was less than 20% of the estimated highest recorded flow. The extension error CV is not well known but Potter and Walker (1981, 1985) suggest it may be as high as 30%. Brown (1983) concluded that the accuracy of high floods at most stations is probably not much better than ±25% and in many cases much worse.

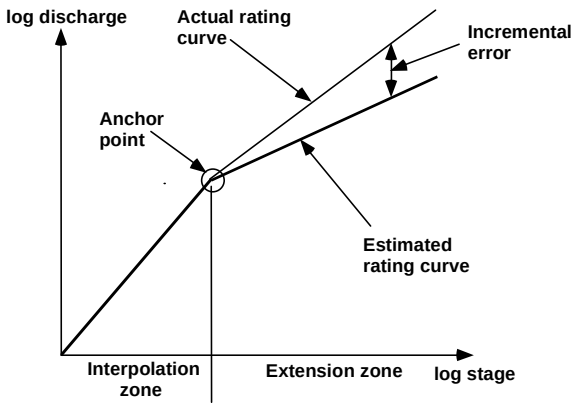

Figure 4. Rating curve extension error.

Though Figure 4 represents an idealization of actual rating curve extension two points of practical significance are noted:

1. The error is systematic in the sense that the extended rating is likely to diverge from the true rating as discharge increases. The error, therefore, is likely to be highly correlated - in fact, it is perfectly correlated in the idealization of Figure 4.
2. The interpolation zone anchors the error in the extension zone. Therefore, the error in the extension zone depends on the distance from the anchor point and not from the origin. This error is termed incremental because it originates from the anchor point rather than the origin of the rating curve.

The conceptual rating error model is incorporated into the Bayesian fitting procedure described in Section 6.3.

3.8 Historical and Paleo Flood Information

A flood may have occurred before, during or after the period of gauged record, and is known to be the largest flood, or flood of other known rank, over a period longer than that of the gauged record. Such floods can provide valuable information, and should be included in the analysis if possible.

Care is needed in assessing historical floods. Only stages are usually available, and these may be determined by flood marks recorded on buildings or structures, by old newspaper reports, or from verbal evidence. Newspaper or other photographs can provide valuable information. Verbal evidence is often untrustworthy, and structures may have been moved. A further problem is that the channel morphology, and hence the stage-discharge relation of the stream, may have changed from those applying during the period of gauged record.

It is desirable to carry out frequency analyses both by including and excluding the historical data. The analysis including the historical data should be used unless in the comparison of the two analyses, the magnitudes of the observed peaks, uncertainty regarding the accuracy of the historical peaks, or other factors, suggest that the historical peaks are not indicative of the extended period or are not accurate. All decisions made should be thoroughly documented.

Paleofloods are major floods that have occurred outside the historical record, but which are evidenced by geological, geomorphological or botanical information. Techniques of

paleohydrology have been described by Costa (1978, 1983, 1986), Kochel et al. (1982) and O'Connell et al. (2002) with a succinct summary given by Stedinger and Cohn (1986). Although high accuracy is not possible with these estimates, they may only be marginally less accurate than other estimates requiring extrapolation of rating curves, and they have the potential for greatly extending the data base and providing valuable information on the right hand tail of the underlying flood probability distribution. Hosking and Wallis (1986) and Jin and Stedinger (1989) consider procedures for assessing the value of paleoflood estimates in flood frequency analysis. Only a little work on this topic has been carried out in Australia, but its potential has been indicated by its use to identify the five largest floods in the last 700 years in the Finke River Gorge in central Australia (Baker et al., 1983; Baker, 1984), and for more frequent floods, by identification of the six largest floods that occurred since a major flood in 1897 on the Katherine River in the Northern Territory (Baker, 1984). The use of paleoflood data should be considered where this is possible in view of the potential benefits, and further development of the technique would be desirable. The two major problems in its use are that there are not many sites where it can be employed, and climate changes may have affected the homogeneity of long-term flood data.

3.9 Data Characterizing Climate Long-Term Persistence

There is growing evidence that flood peaks are not identically distributed from year to year in some parts of Australia and that flood risk is dependent on long-term climate variability. The idea of alternating flood and drought dominated regimes that exist on decadal and longer timescales was first proposed by Warner and Erskine (1988). The climate-dependence of flood risk is an important consideration when assessing flood risk. Most flood frequency applications will require assessment of long-term flood risk; that is, flood risk that is independent of a particular current climate state. If a flood record was sufficiently long to sample all climate states affecting flood risk, a traditional analysis assuming homogeneity would yield the long-term flood risk. Unfortunately many flood records are relatively short and may be dominated by one climate state. Blind use of such data can result in substantial bias in long-term flood risk estimates. For this reason it may be necessary to obtain climate variable data which characterizes long-term persistence in climate and to investigate the homogeneity of the flood distribution.

A number of known climate phenomena impact on Australia climate variability. Most well known is the inter-annual El Nino/Southern Oscillation (ENSO). The cold ENSO phase, La Nina, results in a marked increase in flood risk across Eastern Australia, whereas El Nino years are typically without major floods (Kiem et al., 2003).

There is also mounting evidence that longer-term climate processes also have a major impact on flood risk. The Interdecadal Pacific Oscillation (IPO) is a low frequency climate process related to the variable epochs of warming and cooling in the Pacific Ocean and is described by an index derived from low pass filtering of sea surface

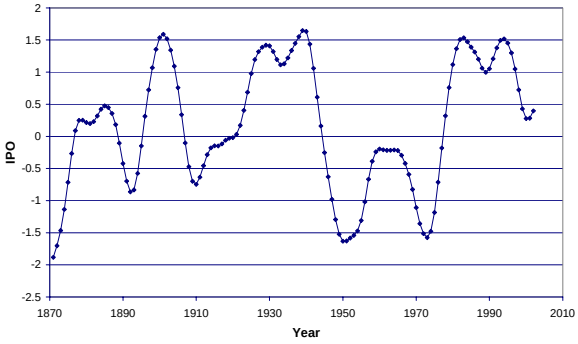

Figure 5. Annual average IPO time series.

temperature (SST) anomalies in the Pacific Ocean [Power et al., 1998, 1999; Allan, 2000]. The IPO is similar to the Pacific Decadal Oscillation (PDO) of Mantua et al. (1997), which is defined as the leading principal component of North Pacific monthly sea surface temperature variability.

The IPO time series from 1880 is displayed in Figure 5. It reveals extended periods where the index either lies below or above zero. Power et al. (1999) have shown that the association between ENSO and Australian climate is modulated by the IPO - a strong association was found between the magnitude of ENSO impacts during negative IPO phases, whilst positive IPO phases showed a weaker, less predictable relationship. Additionally, Kiem et al. (2003) and Kiem and Franks (2004) analyzed NSW flood and drought data and demonstrated that the IPO negative state magnified the impact of La Nina events. Moreover, they demonstrated that the IPO negative phase, related to midlatitude Pacific Ocean cooling, appears to result in an increased frequency of cold La Nina events. The net effect of the dual modulation of ENSO by IPO is the occurrence of multi-decadal periods of elevated and reduced flood risk. To place this in context, Figure 6 shows regional flood index curves based on about 40 NSW sites for the different IPO states (Kiem et al., 2003) – the 1 in 100 AEP flood during years with a positive IPO index corresponds to the 1 in 6 AEP flood during years with a negative IPO index. Micevski et al. (2003) investigating a range of sites in NSW

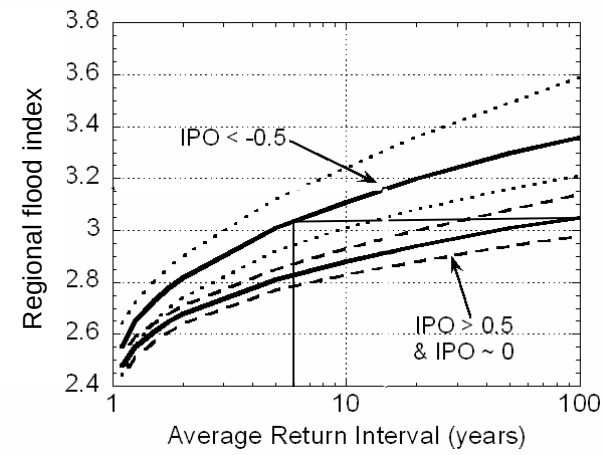

Figure 6. NSW regional flood index frequency curves for positive and negative IPO epochs (dashed lines represent 90% confidence limits) [Kiem et al., 2003].

found that that floods occurring during IPO "negative" periods were, on average, about 1.8 times bigger than floods with the same frequency during IPO "positive" periods.

The finding that flood risk in parts of Australia is modulated by low frequency climate variability is recent. Users are reminded this is an area of active research and to keep abreast of future developments. Nonetheless, the practical implication of these findings is considerable. A standard flood frequency analysis can produce significantly biased estimates of flood risk if this phenomenon is ignored. Section 4.2.6 provides guidance on how to conduct a flood frequency analysis in such circumstances.

3.10 Regional Flood Information

Whereas the primary focus of this chapter is flood frequency analysis using at-site information, the accuracy of the frequency analysis can be improved, substantially in some cases, by augmenting at-site information with regional information. Subsequent chapters in this Book describe methods for estimating flood frequency at ungauged sites. Provided such methods also provide estimates of uncertainty, the regional information can be pooled with the at-site information to yield more accurate results. Section 6.3.5 shows how regional information on flood probability model parameters can be pooled with at-site information. When pooling at-site and regional information it is important to establish that both sources of information are consistent – that is, they yield statistically consistent results.

3.11 Missing Records

Streamflow data frequently contain gaps for a variety of reasons including the malfunction of recording equipment. Rainfall records on the catchment and streamflow data from nearby catchments may indicate the likelihood of a large flood having occurred during the gap. A regression may be able to be derived to enable a missing flood to be estimated, but as discussed in Section 3.6.1, the degree of correlation is often insufficient for a quantitative estimate.

For annual series the missing record period is of no consequence and can be included in the period of record, if it can be determined that the largest flow for the year occurred outside the gap, or that no large rains occurred during the gap. However the rainfall records and streamflow on nearby catchments might indicate that a large flood could have occurred during the period of missing record. If a regression with good correlation can be derived from concurrent records, the missing flood can be estimated and used as the annual flood for the year. If the flood cannot be estimated with reasonable certainty, the whole year should be excluded from the analysis.

For POT series data, treatment of missing records is less clear. McDermott and Pilgrim (1982) tested seven methods, leading to the following recommendations based on the assumption that the periods of missing data are random occurrences and are independent of the occurrence of flood peaks.

1. Where a nearby station record exists covering the missing record period, and a good relation between the flood peaks on the two catchments can be obtained, then use this relation and the nearby station record to fill in the missing events of interest.
2. Where a nearby station record exists covering the missing record period, and the relation between the flood peaks on the two catchments is such that only the occurrence of an event can be predicted but not its magnitude, then:
   - for record lengths less than 20 years, ignore the missing data and include the missing period in the overall period of record;
   - for record lengths greater than 20 years, subtract an amount from each year with missing data proportional to the ratio of the number of peaks missed to the total number of ranked peaks in the year.
3. Where no nearby station record exists covering the missing record period, or where no relation between flood peaks on the catchment exists, then ignore the missing data and include the missing record period in the overall period of record.

4. CHOICE OF FLOOD PROBABILITY MODEL

4.1 General

As noted in Section 2, it is assumed that each flood peak in an annual or POT series is statistically independent of other flood peaks in the series. In addition the flood probability model, described by its probability density function (pdf) $p(q|\theta,x,M)$, must be specified.

There is no universally accepted flood probability model. Many types of probability distributions have been applied to flood frequency analysis. Unfortunately, it is not possible to determine the true or correct form of distribution (see, for example, Cunnane, 1985), and there is no rigorous analytical proof that any particular probability distribution for floods is the correct theoretical distribution. The appropriateness of these distributions can be tested by examining the fit of each distribution to observed flood data. Various empirical tests of different distributions have been carried out with recorded data from many catchments but conclusive evidence is not possible as a result of the wide range of samples that can be realised from a given population. Examples of sampling experiments illustrating this problem are given by Alexander (1957) and Benson in Dalrymple (1960). The choice of flood probability model is further exacerbated by recent evidence that in certain parts of Australia the flood record is not homogeneous due to variations in long-term climate controls.

Given these considerations it is considered inappropriate to be prescriptive with regard to choice of flood probability model. As a general rule the selected probability distribution family M should be consistent with available data. It is recognized that more than one probability distribution family may be consistent with the data. One approach to deal with this problem is to select the

distribution family on the basis of best overall fit to a range of catchments within a region or landscape space – one approach for assessing overall goodness of fit is based on the use of L moment diagrams [see Stedinger et al., 1993, Section 18.3.3 for more details].

4.2 Choice of Distribution Family for Annual Series

Two distribution families are suggested as reasonable initial choices, namely the generalized extreme value (GEV) and log Pearson III (LP3) families. These families have been shown to fit most flood data adequately. Nonetheless the user is reminded that there is no rigorous justification for these families, which is particularly important when extrapolating – Example 1 demonstrates the importance of understanding the mechanisms controlling flood response. The following sections describe the GEV and LP3 distributions and also some other distributions which may be more appropriate in certain circumstances.

4.2.1 Generalized Extreme Value (GEV) distribution

Table 1 lists the pdf $p(q|\theta)$, distribution function $P(Q \le q|\theta)$ and product moments for the generalized extreme value (GEV) distribution. It has three parameters: τ, the location parameter, α, the scale parameter and κ, the shape parameter. When the shape parameter κ equals 0, the GEV simplifies to the Gumbel distribution whose details are provided in Table 1. For positive values of κ there exists an upper bound, while for negative κ there exists a lower bound.

Of the widely used distribution families, the GEV distribution has the strongest theoretical appeal because it is the asymptotic distribution of extreme values for a wide range of underlying parent distributions. In the flood frequency context, suppose there are N flood peaks in a year. Provided N is large and the flood peaks are identically and independently distributed, the distribution of the largest peak flow in the year approaches the GEV under quite general conditions. Unfortunately, hydrologic reality does not always fit comfortably with these assumptions. First, the number of independent flood peaks in any year may not be sufficient to ensure asymptotic behaviour. Second, it remains unclear whether the within-year independent flood peaks are random realizations from the same probability distribution.

The GEV has gained widespread acceptance (for example, Natural Environment Research Council, 1975; Wallis and Wood, 1985; Handbook of Hydrology, Stedinger et al., 1993).

4.2.2 Log Pearson III (LP3) distribution

Table 1 lists the pdf $p(q|\theta)$ and product moments for log Pearson III (LP3) distribution. It has three parameters: τ, the scale parameter, α and β, the shape parameters. When the skew of log q is zero, the distribution simplifies to the log normal whose details are provided in Table 1.

The LP3 distribution has been widely accepted in practice because it consistently fits flood data as well if not better than other probability families. It has performed best of those that have been tested on data for Australian catchments (Conway 1970; Kopittke et al., 1976; McMahon, 1979, McMahon and Srikanthan, 1981). It is the recommended distribution for the United States in Bulletin 17B of the Interagency Advisory Committee on Water Data (1982).

The distribution, however, is not well-behaved from an inference perspective. Direct inference of the parameters α, β and τ can cause numerical problems. For example, when the skew of log q is close to zero, the shape parameter α tends to infinity. Experience shows that the parameters can be estimated reasonably using the first three moments of log(q). Note that τ is a lower bound for positive skew and an upper bound for negative skew.

A problem arises when the absolute value of the skew of log q exceeds 2; that is, when $\alpha < 1$. When $\alpha > 1$, log q has a gamma-shaped density. When $\alpha \le 1$, the density of log q changes to a J-shaped function. Indeed when $\alpha = 1$, the pdf of log q degenerates to that of an exponential distribution with scale parameter β and location parameter τ. For $\alpha \le 1$, the J-shaped density seems to be over-parameterized with three parameters. In such circumstances, it is pointless to use the LP3. It is suggested either the GEV or the generalized Pareto distributions be used as a substitute.

An analytical form of the distribution function is not available for the LP3 and log-normal distributions. To compute the quantile q_Y (that is, the discharge with a 1 in Y AEP) the following equation may be used:

$$\log q_Y = m + K_Y(g)\, s \qquad (11)$$

where m, s and g are the mean, standard deviation and skewness of the log discharge and K_Y is a frequency factor well-approximated by the Wilson-Hilferty transformation

$$K_Y(g) = \begin{cases} \dfrac{2}{g}\left[\left(\dfrac{g}{6}\left(z_Y - \dfrac{g}{6}\right)+1\right)^3 - 1\right] & \text{if } |g| > 0 \\ 0 & \text{if } g = 0 \end{cases} \qquad (12)$$

for $|g| < 2$ and AEPs ranging from 0.99 to 0.01. The term z_Y is the frequency factor for the standard normal distribution which has a mean of zero and standard deviation of 1 – z_Y is the value of the standard normal deviate with exceedance probability 1/Y. Table 2 lists z_Y for selected exceedance probabilities. Comprehensive tables of K_Y can be found in Australian Rainfall and Runoff (1987, p.208).

4.2.3 Generalized Pareto distribution

In POT modelling the generalized Pareto (GP) distribution is often found to satisfactorily fit the data. Table 1 lists the pdf $p(q|\theta)$, distribution function $P(Q \le q|\theta)$ and product moments for the GP distribution. It has three parameters: q_*, the location parameter, β, the scale parameter and κ, the shape parameter. When κ equals zero, the distribution simplifies to the exponential distribution which is summarized in Table 1.

The GP distribution has an intimate relationship with the

Table 1. Selected homogeneous probability models families for use in flood frequency analysis

Family	Distribution	Moments
Generalized extreme value (GEV)	$p(q \mid\mid \theta) = \frac{1}{\alpha}\exp\left\{-\left[1-\frac{\kappa(q-\tau)}{\alpha}\right]^{\frac{1}{\kappa}}\right\}\left[1-\frac{\kappa(q-\tau)}{\alpha}\right]^{\frac{1}{\kappa}-1}$ $P(Q \le q \mid \theta) = \exp\left\{-\left[1-\frac{\kappa(q-\tau)}{\alpha}\right]^{\frac{1}{\kappa}}\right\}$ when $\kappa>0$, $q<\tau+\frac{\alpha}{\kappa}$; when $\kappa<0$, $q>\tau+\frac{\alpha}{\kappa}$	$Mean(q) = \tau+\frac{\alpha}{\kappa}\left[1-\Gamma(1+\kappa)\right]$ $for\ \kappa > -1$ $Variance(q)=\frac{\alpha^2}{\kappa^2}\left[\Gamma(1+2\kappa)-\left[\Gamma(1+\kappa)\right]^2\right]$ $for\ \kappa > -\frac{1}{2}$ where $\Gamma()$ is the gamma function
Gumbel	$p(q \mid \theta) = \frac{1}{\alpha}\exp\left[-\frac{(q-\tau)}{\alpha}\right]\exp\left\{-\exp\left[-\frac{(q-\tau)}{\alpha}\right]\right\}$ $P(Q \le q \mid \theta) = \exp\left\{-\exp\left[-\frac{(q-\tau)}{\alpha}\right]\right\}$	$Mean(q) = \tau + 0.5772\alpha$ $Variance(q) = \frac{\pi^2\alpha^2}{6}$ $Skew(q) = 1.1396$
Log-Pearson III (LP3)	$p(q \mid \theta) = \frac{\lvert\beta\rvert}{q\Gamma(\alpha)}\left[\beta(\log_e q-\tau)\right]^{\alpha-1}\exp\left[-\beta(\log_e q-\tau)\right]$ $\alpha > 0$ when $\beta>0, \log_e q>\tau$; when $\beta<0, \log_e q<\tau$	$Mean(\log_e q) = m = \tau+\frac{\alpha}{\beta}$ $Variance(\log_e q) = s^2 = \frac{\alpha}{\beta^2}$ $Skew(\log_e q) = g = \begin{cases}\frac{2}{\sqrt{\alpha}} & \text{if } \beta>0\\ -\frac{2}{\sqrt{\alpha}} & \text{if } \beta<0\end{cases}$
Log-normal	$p(q \mid \theta) = \frac{1}{q\sqrt{2\pi s^2}}\exp\left[-\frac{1}{2s^2}(\log_e q-m)^2\right]$ $q>0, s>0$	$Mean(\log_e q) = m$ $Variance(\log_e q) = s^2$
Generalized Pareto	$p(q \mid \theta) = \frac{1}{\beta}\left(1-\frac{\kappa(q-q_*)}{\beta}\right)^{\frac{1}{\kappa}-1}$ $P(Q \le q \mid \theta) = 1-\left(1-\frac{\kappa(q-q_*)}{\beta}\right)^{\frac{1}{\kappa}}$ when $\kappa<0, q_* \le q < \infty$; when $\kappa>0, q_* \le q \le \frac{\beta}{\kappa}$	$Mean(q) = q_* + \frac{\beta}{1+\kappa}$ $Variance(q) = \frac{\beta^2}{(1+\kappa)^2(1+2\kappa)}$ $Skew(q) = \frac{2(1-\kappa)(1+2\kappa)^{1/2}}{(1+3\kappa)}.\ \kappa > -1/3$
Exponential	$p(q \mid \theta) = \frac{1}{\beta}\exp\left(-\frac{q-q_*}{\beta}\right)$ $P(Q \le q \mid \theta) = 1-\exp\left(-\frac{q-q_*}{\beta}\right), q \ge q_*$	$Mean(q) = q_* + \beta$ $Variance(q) = \beta^2$ $Skew(q) = 2$

GEV. If the GP describes the distribution of peaks over a threshold, then for Poisson arrivals of the peaks with ν being the average number of arrivals per year, it can be shown that the distribution of annual maximum peaks is GEV with shape parameter κ, scale parameter $\alpha = \beta\nu^{-\kappa}$ and location parameter $\tau = q_* + \frac{\beta}{\kappa}$ [see Stedinger et al., 1993, p18.39 for more details]. Likewise when peaks over a threshold follow an exponential distribution then for the Poisson arrival the distribution of annual maxima follows a Gumbel distribution.

4.2.4 Threshold mixture model

In certain parts of Australia annual maximum flood series

Table 2. Frequency factors for standard normal distribution.

Y in AEP of 1 in Y	z_Y
2	0.0000
5	0.8416
10	1.2816
20	1.6449
50	2.0537
100	2.3263
200	2.5758
500	2.8782
1000	3.0902

may contain one or more years of zero or negligible flow. This causes a lumpiness in the flood probability model which is best handled by generalizing the pdf to the threshold mixture form

$$p(q \mid \theta, M) = \begin{cases} P(q_*) \text{ if } q \le q_* \\ \left(1 - P(q_*)\right) p(q \mid q > q_*, \theta, M) \text{ if } q > q_* \end{cases} \quad (13)$$

where $P(q_*)$ is the probability of a flow less than or equal to q_*, the threshold value and $p(q|q>q_*,\theta,M)$ is the conditional pdf of a flood peak exceeding the threshold q_*.

The mixture model is interpreted as follows: There is a probability $P(q_*)$ of the annual maximum flow being less than or equal to the threshold q_* in any year. If the annual maximum flood exceeds the threshold its value is randomly sampled from the conditional pdf $p(q|q>q_*,\theta,M)$.

4.2.5 Multi-component or mixture models

In some areas, flooding may be affected by different types of meteorological events (for example, intense tropical cyclones and storms characterized by more usual synoptic conditions) or by changing hydraulic controls (e.g., see Example 1), causing abnormal slope changes in the frequency curve. In such cases the GEV or LP3 families may not adequately fit the data. This type of problem may be of particular importance in northern Australia, and an example is described by Ashkanasy and Weeks (1975).

Where components affecting flooding have distinctly different distributions one can separate the data by cause and analyze each set separately using simple distributions. Ashkanasy and Weeks (1975) used a combination of two log Normal distributions. Waylen and Woo (1982) considered separately summer runoff and winter snowmelt floods. They showed that separation of the flood record into separate series can result in a better estimate of the probability of extreme events because the data that describes phenomena that produce those large events is better represented in the analysis.

A two-component extreme value (TCEV) distribution has been described by Rossi et al. (1984) and Fiorentino et al. (1985), and developed and applied to U.K. data by Beran et al. (1986). It corresponds to the maximum of two independent Gumbel distributions. Generally, one of the two distributions is thought of as describing the bulk of the data, and the other as the outlier distribution. Because the TCEV has four parameters, it is very flexible (Beran et al., 1986). However, if only annual maximum series are available, regional estimation methods must be used to help estimate the parameters.

4.2.6 Non-homogeneous models

If the evidence suggests that flood risk is affected by multi-decadal climate persistence, the use of non-homogeneous probability models may need to be investigated. The concern is that ignoring the non-homogeneity of the flood record may lead to biased estimates of long-term flood risk.

If a non-homogeneous probability model with pdf $p(q|\theta,x,M)$ is identified this cannot be used to estimate long-term or marginal flood risk. This is because the flood risk is dependent on the exogenous climate-related variables x. To estimate long-term flood risk the dependence on x must be removed using total probability to give

$$P(Q \le q \mid \theta, M) = \int_X \left(\int_o^q p(z \mid \theta, x, M) dz \right) p(x) dx \quad (14)$$

where p(x) is the pdf of the exogenous variables.

If the gauged record adequately samples the distribution of x, it is not necessary to identify a non-homogeneous model. It suffices to fit a probability model to all the record to estimate the long-term flood risk.

However, if the gauged record does not adequately sample the exogenous variable x, significant bias in long-term flood risk may result. In such instances, one can:

- Use a regional frequency method based on sites with sufficiently long records to fully sample x.
- Extend the flood record using nearby sites which have more fully sampled the exogenous variable x.

Example 6 illustrates the impact on flood risk arising from multi-decadal persistence in climate state as represented by the IPO index. It demonstrates the serious bias in flood risk that can arise if short records do not adequately sample different climate states. It also demonstrates use of eqn (14).

In the future it is expected regional frequency methods explicitly dealing with non-homogeneity will be developed. This will allow combining regional and site information (see Example 4) to yield the most accurate frequency analysis.

4.3 Choice of Distribution for POT Series

In some cases, it may be desirable to analytically fit a probability distribution to POT data. Distributions that have been used to describe the flood peak above a threshold include exponential, GP and LP3.

5. CHOICE OF FLOOD ESTIMATOR

If the true value of θ were known, then the pdf $p(q|\theta,M)$ could be used to compute the flood quantile $q_Y(\theta,M)$. For annual maximum series the quantile for a 1 in Y AEP is defined as

$$P(q > q_Y \mid \theta, M) = \frac{1}{Y} = \int_{q_Y}^{\infty} p(q \mid \theta, M)\, dq \qquad (15)$$

However, in practice the true value of θ (as well as the distribution family M) is unknown. All that is known about θ, given the data D, is summarized by a probability distribution with pdf p(θ|D,M). This distribution may be the sampling distribution from a bootstrap analysis or the posterior distribution from a Bayesian analysis. Note that if an informative prior distribution is employed the Bayesian analysis will in practice produce different results to a bootstrap analysis.

This uncertainty should be accounted for when making predictions, particularly when extrapolating to floods beyond the observed record. There are a number of ways the uncertainty in θ can be treated leading to different flood quantile estimators. The following sections consider two estimators, expected parameter quantiles and expected probability quantiles.

5.1 Expected Parameter Quantiles

In general the estimation of a design flood should be guided by the consequences of over- and underdesign [Slack et al., 1975]. In the presence of parameter uncertainty, an expected loss approach can be employed.

Define the loss function $L[\hat{q}(D), q(\theta)]$ as the loss incurred when the true quantile q(θ), which depends on θ, is replaced by the quantile estimate $\hat{q}(D)$, which depends on the data D. Because the data do not exactly specify θ, the loss itself is not exactly known.

One approach to find the best estimate is to seek the estimate $\hat{q}(D)_{opt}$ which minimizes the expected loss defined by

$$\hat{q}(D)_{opt} \leftarrow \min \int_{\theta} L[\hat{q}(D), q(\theta)]\, p(\theta \mid D, M)\, d\theta \qquad (16)$$

If the loss function is quadratic

$$L[\hat{q}(D), q(\theta)] = \alpha[\hat{q}(D) - q(\theta)]^2 \qquad (17)$$

where α is a constant, the consequences of under- and overdesign are assumed equal. DeGroot (1970) shows that for this loss function, termed a mean-squared-error (MSE) loss function, the optimal estimate is the expected value of q(D).

It follows that the optimal MSE quantile estimate is the expected value of the quantile defined as

$$E[q_Y \mid D, M] = \int_{\theta} q_Y(\theta, M)\, p(\theta \mid D, M)\, d\theta \qquad (18)$$

Using a first-order Taylor series approximation yields

$$E[q_Y|D,M] \approx q_Y[E(\theta|D,M)] \qquad (19)$$

where E(θ|D,M) is the expected parameter given the data D and model family M.

The term $q_Y[E(\theta|D,M)]$ is referred to as the 1 in Y AEP quantile. When this term is used it is understood that the expected parameter vector conditioned on the data D, E(θ|D,M)], is being treated as if it were the true value of θ.

A closely related estimate to E(θ|D,M) is θ_{MP} defined as the value of θ that maximizes the posterior pdf p(θ|D,M). Asymptotically θ_{MP} and E(θ|D,M) should converge to the same value. Martins and Stedinger (2000, 2001) show that when a prior distribution is used to restrict parameters to a statistically/physically reasonable range θ_{MP} (or generalized maximum likelihood estimate) yields MSE quantile estimates for GEV distributions with negative shape factors which are substantially better for than method of moment and L moment estimates.

5.2 Expected Probability Quantiles

Stedinger (1983) showed that the dependence of the flood quantile on uncertain parameters can be removed using total probability to yield the design flood distribution

$$p(q \mid D, M) = \int_{\boldsymbol{\theta}} p(q \mid \theta, M)\, p(\theta \mid D, M)\, d\theta \qquad (20)$$

This distribution only depends on the data D and the assumed probability family M.

Suppose the 1 in Y AEP quantile q_Y is evaluated. Using the design flood distribution given by eqn (20) and changing the order of integration yields the AEP for q_Y as

$$\begin{aligned} P(q > q_Y \mid D, M) &= \int_{\theta} \left(\int_{q_Y}^{\infty} p(q \mid \theta, M)\, dq \right) p(\theta \mid D, M)\, d\theta \\ &= \int_{\theta} P(q > q_Y \mid \theta, M)\, p(\theta \mid D, M)\, d\theta \end{aligned} \qquad (21)$$

Eqn (21) demonstrates that $P(q>q_Y|D,M)$ is an expected probability. For this reason, the design flood distribution is also called the expected probability flood distribution.

In general the presence of parameter uncertainty results in $P(q>q_Y|D,M)$ exceeding 1/Y. That is, the magnitude of the 1 in Y AEP quantile is expected to be exceeded more often than 1 in Y years.

The term $P(q>q_Y|D,M)$ is called the expected probability for quantile q_Y.

5.3 Selection of Flood Quantile Estimator

Ideally the flood estimator should be derived by minimizing a loss criterion such as the expected loss given by eqn (16). This is particularly important because the loss function is not necessarily quadratic – see Stedinger (1997) and Al-Futaisi and Stedinger (1999) – thereby undermining the justification for use of eqn (19).

However, in most situations, it is neither practical nor warranted to perform a formal loss analysis. In such cases the decision therefore reduces to choosing between the 1 in Y AEP quantile or the expected probability to the quantile defined by eqn (19).

The decision is somewhat subjective and a matter of policy. It depends on whether design is being carried out for many sites or a single site, and whether risk, the actual discharge or average annual damage is of primary interest in design. As a general guide, some recommendations are given below. The list is not intended to be exhaustive.

Situations where expected probability should be considered:

1. Where design is required for many sites in a region, and the objective is to attain a desired average exceedance frequency over all sites.
2. Where flood estimates are made at multiple sites with different record lengths, and a comparable basis is required for flood estimates of a given AEP.
3. Design for a single site, where risk or frequency of exceedance is of primary importance, and it is desired that the actual frequency of exceedance will equal the design frequency.
4. Flood estimation for flood plain management where the actual frequency of flooding is required to equal the design frequency, as in point 3.
5. Design for sites where underestimated frequency of surcharging carries large penalties, such as would occur with too-frequent closure of a railway or major road, or overtopping of a levee.
6. The investigation of appropriate levels of design frequency for particular sites or structures, and the development of generalized data-based design standards.
7. Split sample testing of flood estimation procedures (Beard, 1960).

Situations where neglect of expected probability should be considered:

1. Where sizing of a structure is of primary interest rather than actual frequency of flooding, such as some bridges, culverts or drainage pipes where overdesign may have large cost penalties relative to the benefits.
2. Where design is carried out to an arbitrary standard which has been judged to be satisfactory on the basis of experience from previous practice. Many of the design standards in Book III Sections 1.8 and 1.9, fall into this category.
3. Where unbiased estimates of annual flood damages are required (Doran and Irish, 1980).

6. FITTING FLOOD PROBABILITY MODELS TO ANNUAL SERIES

6.1 Overview of Methods

Fitting a flood probability model involves three steps:

1. Calibrating the model to available data D to determine the parameter values consistent with the data D.
2. Estimation of confidence limits, flood quantiles and expected probability floods.
3. Evaluation of goodness of fit and consistency of model with data.

Two calibration approaches are described. The recommended approach uses Bayesian methods on account of their ability to handle gauged and censored data, errors in data and regional information. The second approach based on L moments is included because of its intrinsic value in regionalization studies and usefulness in dealing with difficult-to-fit data. For each approach the methods are documented and illustrated with worked examples. Implementation of the methods in software requires specialist skill; therefore the typical user is advised to make use of available software.

Other methods may be used for calibration provided they make efficient use of information. The use of the method of product-moments applied to the LP3 is not recommended because more efficient approaches are available.

6.2 Probability Plots

An essential part of a flood frequency analysis is the construction of an empirical distribution function, better known as a probability plot. In such a plot an estimate of AEP is plotted against the observed discharge. This enables one to draw a smooth curve as an empirical probability distribution or to visually check the adequacy of a fitted distribution.

The following steps describe the production of a probability plot for gauged annual maximum floods:

- Rank the gauged flows in descending order (that is, from largest to smallest) yielding the series $\{q_{(1)}, q_{(2)},\ldots,q_{(n)}\}$ where $q_{(i)}$ is the rank i or the i^{th} largest flood.
- Estimate the AEP for each $q_{(i)}$ – the estimate is referred to as a plotting position.
- Using suitable scales plot the estimated AEP against $q_{(i)}$.

For analysis of the annual flood series, a general formula (Blom, 1958) for estimating the AEP of an observed flood is

$$P_{(i)} = \frac{i - \alpha}{n + 1 - 2\alpha} \qquad (22)$$

where i is the rank of the gauged flood, n is the number of years of gauged floods and α is a constant whose value is selected to preserve desirable statistical properties.

There are several choices for α:

- $\alpha = 0$ yields the Weibull plotting position which produces unbiased estimates of the AEP of $q_{(i)}$.
- $\alpha = 0.375$ yields the Blom's plotting position which produces unbiased quantile estimates for the normal distribution.
- $\alpha = 0.4$ yields the Cunnane's (1978) plotting position which produces nearly unbiased quantile estimates for a range of probability families.

While there are arguments in favour of plotting positions that yield unbiased AEPs, usage has favoured plotting positions that yield unbiased quantiles. To maintain consistency it is recommended that the Cunnane plotting position be used, namely

$$P_{(i)} = \frac{i - 0.4}{n + 0.2} \tag{23}$$

A fuller discussion on plotting positions can be found in Stedinger et al. (1993).

It is stressed that plotting positions should not be used as an estimate of the actual AEP or ARI of an observed flood discharge. Such estimates should be obtained from the fitted distribution.

Judicious choice of scale for the probability plot can assist the evaluation of goodness of fit. The basic idea is to select a scale so that the data plot as a straight line if the data are consistent with the assumed probability model.

This is best illustrated by an example. Suppose it is assumed that floods follow an exponential distribution. From Table 1 the distribution function is

$$P(Q \le q) = 1 - \exp\left(-\frac{q - q_*}{\beta}\right) \tag{24}$$

Replacing q by $q_{(i)}$ and 1-P(Q≤q) by the plotting position of $q_{(i)}$ gives

$$1 - P_{(i)} = 1 - \exp\left(-\frac{q_{(i)} - q_*}{\beta}\right) \tag{25}$$

Making $q_{(i)}$ the subject of the equation yields

$$q_{(i)} = q_* - \beta \log P_{(i)} \tag{26}$$

If $q_{(i)}$ is plotted against $\log P_{(i)}$ the data will plot approximately as a straight line if they are consistent with the exponential distribution.

Examples for other distributions include:

- For the Gumbel distribution plot $q_{(i)}$ against $-\log[-\log(1-P_{(i)})]$. Data following a GEV distribution will plot as a curved line.
- For the log normal distribution plot $\log q_{(i)}$ against the standard normal deviate with exceedance probability $P_{(i)}$. Data following a LP3 distribution will plot as a curved line.

When visually evaluating the goodness of fit care needs to exercised in judging the significance of departures from the assumed distribution. Plotting positions are not scattered about the true distribution independently of each other. As a result, plotting positions exhibit a wave-like behaviour about the true distribution. To guard against misinterpreting such behaviour it is suggested that statistical tests be used to assist goodness-of-fit assessment. Stedinger et al. (1993) discuss the use of the Kolmogorov-Smirnov test, the Filiben probability plot correlation test and L moment diagrams and ratio tests.

The estimation of plotting positions for censored and historic data is more involved and in some cases can be inaccurate. See Stedinger et al. (1993, p 18.41) for more details.

6.3 Bayesian Calibration

6.3.1 Overview

The Bayesian approach is a very general approach for calibrating and identifying models. The Handbook of Hydrology [Stedinger et al., 1993, p18.8] observes that "the Bayesian approach .. allows the explicit modeling of uncertainty in parameters and provides a theoretically consistent framework for integrating systematic flow records with regional and other hydrologic information". However, it is only with the advent of new computational methods that Bayesian methods can be routinely applied to flood frequency applications.

The core of the Bayesian approach is described below – see Lee (1989) and Gelman et al., 1995) for general expositions. Consider a set of data D hypothesized to be a random realization from the probability model M with pdf $p(D|\theta,M)$ where θ is an unknown finite-dimensioned parameter vector. The pdf $p(D|\theta,M)$ is given two labels depending on the context. When $p(D|\theta,M)$ is used to describe the probability model generating the sample data D for a given θ it is called the sampling distribution. However, when inference about the parameter θ is sought, $p(D|\theta,M)$ is called the likelihood function to emphasize that the data D is known and the parameter θ is the object of attention. The same notation for the sampling distribution and likelihood function is used to emphasize its oneness.

In Bayesian inference the parameter vector θ is considered to be a random vector whose probability distribution describes what is known about the true value of θ. Prior to analyzing the data D knowledge about θ given the probability model M is summarized by the pdf $p(\theta|M)$. This density, referred to as the prior density, can incorporate subjective belief about θ.

Bayes theorem is then used to process the information contained in the data D by updating what is known about the true value of θ as follows:

$$p(\theta \mid D, M) = \frac{p(D \mid \theta, M)\, p(\theta \mid M)}{p(D \mid M)} \tag{27}$$

The posterior density $p(\theta|D,M)$ describes what is known about the true value of θ given the data D and the model M. The denominator $p(D|M)$ is the marginal likelihood defined as

$$p(D \mid M) = \int p(D \mid \theta, M)\, p(\theta \mid M) d\theta \tag{28}$$

6.3.2 Likelihood Function

The key to a Bayesian analysis is the formulation of the likelihood function. In the context of flood frequency analysis two formulations are considered. The first assumes there is no error in the flood data. The focus is on the contribution to the likelihood function made by gauged and

censored data. The second case then generalizes the likelihood function to allow for error-in-discharge estimates.

6.3.3 Likelihood function: No-error-discharge Case

Suppose the following data are available:

1. A gauged record of n true flood peaks $\{q_1,...,q_n\}$; and
2. m censored records in which a_i annual flood peaks in $(a_i + b_i)$ ungauged years exceeded a threshold with true discharge Q_i, i=1,..,m.

Denote the flood data as D = $\{q_i, i=1,..,n; (a_i, b_i, Q_i), i=1,..,m\}$. The likelihood function is the joint pdf of D given the parameter vector θ. Noting the statistical independence of flood peaks it follows that the likelihood function for the gauged data is the joint pdf of the n gauged floods (Stedinger and Cohn, 1986); that is,

$$p(q_1,..,q_n \mid \theta, M) = \prod_{i=1}^{n} p(q_i \mid \theta, M) \tag{29}$$

The probability of observing exactly k exceedances in n years is given by the binomial distribution

$$p(k \mid n, \pi) = {}^{n}C_{k} \left(1-\pi\right)^{n-k} \pi^{k} \tag{30}$$

where π is the probability of an exceedance. Hence the likelihood of the censored data becomes

$$\begin{aligned} &p(\text{censored data} \mid \theta, M) \\ &= \prod_{i=1}^{m} {}^{a_i+b_i}C_{a_i} \left[1 - P(Q \le Q_i \mid \theta, M)\right]^{a_i} P(Q \le Q_i \mid \theta, M)^{b_i} \\ &= \prod_{i=1}^{m} P(a_i, b_i \mid Q_i, \theta, M) \end{aligned} \tag{31}$$

where $P(a_i,b_i|Q_i,\theta,M)$ is the binomial probability of observing exactly a_i exceedances above the threshold discharge Q_i in (a_i+b_i) years.

6.3.4 The Likelihood Function: Error-in-discharge Case

The incorporation of rating errors into the likelihood function complicates matters. A simple model based on Section 3.7 is presented. Figure 7 presents a rating error space diagram. In zone 1, the interpolation zone it is assumed the rating error multiplier e_1 equals 1 – that is, errors within the rated part of the rating curve are deemed negligible. As a result the estimated flow w equals the true flow q. However, in zone 2, the extension zone, the rating error multiplier e_2 is assumed to be a random variable with mean of 1. The anchor point (q_1,w_1) separates the interpolation and extension zones. The rating error model can represented mathematically as

$$w = \begin{cases} q & \text{if } q \le q_1 \\ w_1 + e_2 (q - q_1) & \text{if } q > q_1 \end{cases} \tag{32}$$

The rating error multiplier e_2 is sampled only once at the time of extending the rating curve. Therefore, all flood discharge estimates exceeding the anchor value of q_1 (which equals w_1) are corrupted by the same rating error multiplier. It must be stressed that the error e_2 is not known – at best, only its probability distribution can be estimated. For practical applications one can assume e_2 is distributed as either a log-normal or normal distribution with mean 1 and standard deviation σ_2.

Data are assigned to each of the two zones, i=1,2, in the rating error space diagram. The rating error multiplier standard deviation for the extension zone σ_2 is assigned a value with $\sigma_1 = 0$. There are n_i annual flood peak estimates w_{ji} satisfying the zone constraint $w_{i-1} \le w_{ji} < w_i$, j=1,..,n_i where w_0=0 and w_2=∞. In addition, there are m_i threshold discharge estimates W_{ji} for which there are a_{ji} exceedances in $(a_{ji}+b_{ji})$ years, j=1,..,m_i. Collectively this data is represented as

$$\begin{aligned} D &= \{D_i, i=1,2\} \\ &= \{[w_{ji}, j=1,..,n_i; W_{ji}, a_{ji}, b_{ji}, j=1,..,m_i], i=1,2\} \end{aligned}$$

Following Kuczera (1999) it can be shown for the two-zone rating error model of Figure 7 the likelihood reduces to

$$\begin{aligned} p(D_1, D_2 \mid \theta, \sigma_2, M) &= p(D_1, e_1 = 1 \mid \theta, M) \\ &\left[\int_0^{\infty} p(D_2, e_2 \mid \theta, M)\, g(e_2 \mid \sigma_2)\, de_2\right] \end{aligned} \tag{33}$$

where

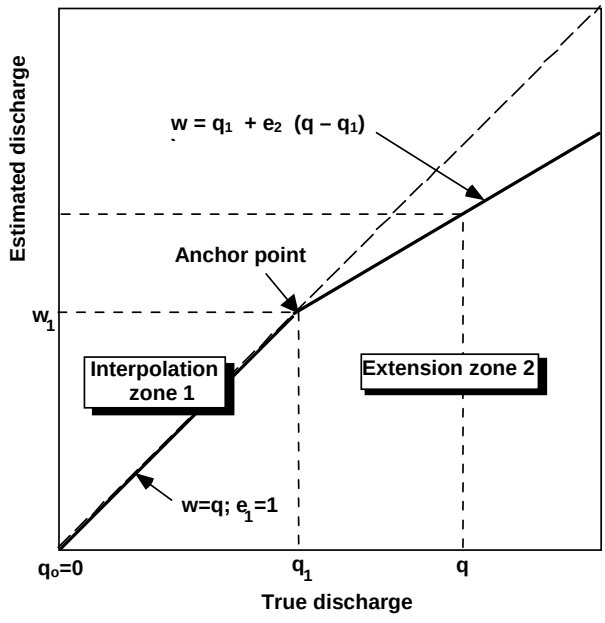

Figure 7. Rating error multiplier space diagram for rating curve shown in Figure 5.

$$p(D_i, e_i \mid \theta, M) = \prod_{j=1}^{n_i} \frac{1}{e_i} p(q_{i-1} + \frac{w_{ji} - q_{i-1}}{e_i} \mid \theta) \prod_{j=1}^{m_i} P[a_{ji}, b_{ji} \mid q_{i-1} + \frac{W_{ji} - q_{i-1}}{e_i}, \theta, M]$$

and $g(e_i|\sigma_i)$ is the rating error multiplier pdf with mean 1 and standard deviation σ_i. This is a complex expression which can only be evaluated numerically. However, it makes the fullest use of information on annual flood peaks and binomial-censored data in the presence of rating curve error. Section 3.7 offers some guidance on the choice of σ_2.

6.3.5 Prior Distributions

The prior pdf $p(\theta|M)$ reflects the worth of prior information on θ obtained preferably from a regional analysis. A convenient distribution is the multivariate normal with mean μ_p and covariance Σ_p; that is,

$$\theta \sim N(\mu_p, \Sigma_p) \tag{34}$$

In the absence of prior information, the covariance matrix can be made non-informative as illustrated in the following equation for a three-parameter model:

$$\Sigma_p \underset{M \to \infty}{\rightarrow} \begin{pmatrix} M & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & M \end{pmatrix} \tag{35}$$

Informative prior information can be obtained from a regional analysis of flood data using methods such index flood, skew maps, and generalized least squares regression. Importantly the reliability of the prior information must accompany the estimates.

Although this Chapter has focused on the use of at-site data, the importance of using well-founded prior information cannot be stressed strongly enough. The use of prior information on one or more parameters confers very significant benefits as demonstrated in Example 4.

Martins and Stedinger (2000, 2001) demonstrate that the method of generalized maximum likelihood produce results superior to L moment and method-of-moment estimators for the GEV. This result is attributed to the use of a geophysical prior on the shape parameter κ. The prior, a beta distribution over the range $-0.5 < \kappa < 0.5$ with mean -0.10 and standard deviation 0.122, confines κ to a range consistent with worldwide experience on rainfall depths and flood flows. Where possible, users are encouraged to make use of well-founded prior information.

6.3.6 Monte Carlo Sampling from the Posterior Distribution

The posterior pdf $p(\theta|D,M)$ fully defines the parameter uncertainty. However, until the last decade, interpreting this distribution was problematic. Modern Monte Carlo methods for sampling from the posterior have overcome this limitation – for example, see Gelman et al. (1995). The sampling procedure described here involves a method called importance sampling. Three steps are involved:

Step 1: Find most probable posterior parameters

Any robust search method can be used to locate the value of θ which maximizes the logarithm of the posterior probability density; that is,

$$\hat{\theta} \leftarrow \max_{\theta} \log p(\theta \mid D, M) \tag{36}$$

where $\hat{\theta}$ is the most probable value of θ. The shuffled complex evolution algorithm of Duan et al. (1992) is a recommended search method.

Step 2: Multinormal approximation to posterior distribution

Almost always, the log of posterior pdf $p(\theta|D,M)$ can be approximated by a second-order Taylor series expansion about the most probable parameter to yield the multivariate normal approximation

$$\theta \mid D, M \sim N(\hat{\theta}, \Sigma) \tag{37}$$

where $\hat{\theta}$ is interpreted as the mean and the posterior covariance Σ is defined as the inverse of the Hessian

$$\Sigma = \left(- \frac{\partial^2 \log_e p(\theta \mid D, M)}{\partial^2 \theta} \right)^{-1} \tag{38}$$

An adaptive difference scheme should be used to evaluate the Hessian. Particular care needs to be exercised when selecting finite difference perturbations for the GEV and LP3 distributions when upper or lower bounds are close to the observed data.

Step 3: Importance sampling of posterior distribution

Importance sampling is a Monte Carlo technique [Gelman et al., 1995] for random sampling from a probability distribution. The basic idea is to randomly sample parameters (called particles) from a reasonable approximation to the posterior distribution (for example, the multinormal approximation obtained in step 2) and then make a correction to the probability mass assigned to the particle. The algorithm proceeds as follows:

i. Sample N particles according to $\theta_i \leftarrow p_N(\theta), i = 1,..,N$ where $p_N(\theta)$ is the pdf of the multinormal approximation in Step 2.

ii. Calculate particle probability weights according to $P(\theta_i) = \frac{p(\theta_i \mid D, M)}{p_N(\theta_i)}, \quad i = 1,..,N$

iii. Scale the particle weights so they sum to 1.

6.3.7 Quantile Confidence Limits and Expected Probability

Using the N particles produced by the importance sampling the posterior distribution of any function dependent on θ can be readily approximated.

Confidence limits[1] describe the uncertainty about quantiles arising from uncertainty in the fitted parameters. They are used in conjunction with the probability plot to evaluate goodness-of-fit. 100(1-α)% quantile confidence limits, or more correctly probability limits, can be obtained as follows: Rank in ascending order the N quantiles $\{q_Y(\theta_{(i)},M), i=1,..,N\}$. For each ranked quantile evaluate the non-exceedance probability $\sum_{j=1}^{i} P(\theta_{(j)})$ where $P(\theta_{(j)})$ is the particle weights for the j^{th} ranked quantile $q_Y(\theta_{(j)},M)$. The lower and upper confidence limits are the quantiles whose non-exceedance probabilities are nearest to $\frac{\alpha}{2}$ and $1-\frac{\alpha}{2}$ respectively.

The expected posterior parameters can be estimated

$$E[\theta \mid D,M]=\sum_{i=1}^{N} P(\theta_i)\, q_Y(\theta_i,M) \qquad (39)$$

These parameters can then be used to compute the 1 in Y AEP quantiles described in Section 5.1.

Finally, the expected exceedance probability for a flood of magnitude q_Y can be estimated

$$E\left[P(q > q_Y \mid D,M)\right]=\sum_{i=1}^{N} P(\theta_i)\, P(q > q_Y \mid \theta_i,M) \qquad (40)$$

6.3.8 Software

The Bayesian approach to calibrating flood probability models is numerically complex and is best implemented in a high level programming language such as FORTRAN or C. The reader is referred the Australian Rainfall and Runoff web site www.arr.org.au for available software.

6.3.9 Treatment of Poor Fits

The standard probability models described in Section 4 may sometimes poorly fit flood data. Typically the goodness-of-fit is assessed by comparing observed data against the fitted probability model and its confidence limits. A poor fit may be characterized by:

- Presence of outliers in the upper or lower tail of the distribution. Outliers in flood frequency analysis represent observations that are inconsistent with the trend of the remaining data and typically would lie well outside appropriate confidence limits.
- Systematic discrepancies between observed and fitted distributions. Caution is required in interpreting systematic departures on probability plots. Confidence limits can help guide the interpretation.

Poor fits to the standard probability models may arise for a variety of reasons including the following:

1. By chance, one or more observed floods may be unusually rare for the length of gauged record. Recourse to the historical flood record may be useful in resolving this issue.
2. Rating curve extensions are biased resulting in a systematic under or over-estimate of large floods (see discussion in Section 3.7).
3. A change in hydraulic control with discharge may affect the shape of the frequency curve as illustrated in Example 1.
4. Storm events responsible for significant flooding may be caused by different meteorological mechanisms which lead to a mixed population not amenable to three-parameter distributions. This may arise when the majority of flood-producing storms are generated by one meteorological mechanism and the minority by an atypical mechanism such as a tropical cyclone.
5. Nonhomogeneity of the flood record.

The potential causes of a poor fit need careful investigation.

If it is decided the poor fit is due to inadequacy of the probability model, two strategies are available to deal with the problem:

- The data responsible for the unsatisfactory fit may be given less weight; or
- A more flexible probability model be used to fit the data.

Data in the upper part of the distribution is typically of more interest and therefore a strong case needs to be made to justify reduction in weight of such data. A corollary of this is that data in the lower part of the distribution is typically of less interest and therefore more amenable to weighting. Examples 5 and 8 illustrate two approaches to devaluing low flow data which are responsible for a poor fit to the GEV distribution.

6.3.10 Worked Examples

Example 2 illustrates fitting a LP3 distribution to a 31-year gauged record. Although the fit is judged satisfactory, considerable uncertainty in the 1 in 100 AEP quantile is noted.

Example 3 is a continuation of Example 2 and illustrates the benefit of incorporating censored historic flood information. In the 118 years prior to gauging only one flood exceed the largest gauged flood. This information is shown to substantially reduce quantile uncertainty.

Example 4 is a continuation of Example 2. It illustrates the value of regional information in reducing uncertainty in parameters and quantiles.

Example 5 illustrates how three–parameter distributions such as the GEV and LP3 can be made to fit data exhibiting sigmoidal behaviour. Because interest is in fitting the higher flows the low flows are de-emphasized by treating them as

[1] In Bayesian statistics the correct terminology is credible interval rather than confidence interval. Given the widespread usage of the term and its popular mis-interpretation as a probability interval the term 'confidence limit' is retained.

censored observations.

Finally Example 6 illustrates application of a non-homogeneous flood probability model conditioned on the IPO index. It shows how the long-term flood risk may be estimated.

6.4 L Moments Approach

6.4.1 Overview

L moments were developed to provide estimators less sensitive to outliers than product moments which involve raising observations to an integer power. They are of particular interest in regionalization studies because they yield estimators of dimensionless moment ratios that are almost unbiased and normally distributed. [Stedinger et al., 1993].

For at-site flood frequency analysis the evidence suggests L moments provide reasonably efficient estimators. Wang (1999) compared the efficiency of L moments using the GEV against product-moments using LP3. He undertook a Monte Carlo study fitting LP3 with product moments and GEV with L moments to data generated from LP3 distributions fitted to 151 Australian flood records. He showed that the GEV L moment estimates of 50- and 100-year quantiles had mean-squared errors substantially lower than the LP3 product-moment estimates for most of the 151 distributions. However, Rosbjerg et al. (1992), Madsen et al. (1997) and Martins and Stedinger (2000, 2001) demonstrate the overall superiority of the method of moments over L moments when estimating GEV quantiles.

In view of this the Bayesian approach is the preferred one. This section presents L moments because of their intrinsic value in regionalization studies. Moreover, when standard three-parameter distributions do not adequately fit the data, a generalization of L moments, called LH moments, can be used to obtain a good fit to the right hand tail. This can be used to assist setting the censoring threshold as illustrated in Example 5.

6.4.2 L-moments for summarizing distributions

Hosking (1990) developed the L-moment theory based on order statistics. The first four L-moments are defined as

$$\lambda_1 = E[X_{1:1}]$$

$$\lambda_2 = \frac{1}{2}E[X_{2:2} - X_{1:2}]$$

$$\lambda_3 = \frac{1}{3}E[X_{3:3} - 2X_{2:3} + X_{1:3}] \qquad (41)$$

$$\lambda_4 = \frac{1}{4}E[X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}]$$

where $X_{j:m}$ is the j^{th} smallest variable in a sample of size m and E stands for expectation.

Wang (1996) justifies L-moments as follows: "When there is only one value in a sample, it gives a feel of the magnitude of the random variable. When there are two values in a sample, their difference gives a sense of how varied the random variable is. When there are three values in a sample, they give some indication on how asymmetric the distribution is. When there are four values in a sample, they give some clue on how peaky, roughly speaking, the distribution is. "

When many such samples are considered, the expectations λ_1 and λ_2 give measures of location and scale. Moreover, the L moment ratios

$$\tau_3 = \frac{\lambda_3}{\lambda_2}$$
$$\tau_4 = \frac{\lambda_4}{\lambda_2} \qquad (42)$$

give measures of skewness and kurtosis respectively. Hosking termed τ_3 L-skewness and τ_4 L-kurtosis. Hosking also defined the L-coefficient of variation as

$$\tau_2 = \frac{\lambda_2}{\lambda_1} \qquad (43)$$

Table 3 summarizes L moments for a range of distributions.

6.4.3 L-moments estimates for gauged data sample data

The traditional L moment estimator is based on probability weighted moments. However, Wang (1996) derived the following sample estimators directly from the definition of the first four L-moments:

$$\hat{\lambda}_1 = \frac{1}{{}^{n}C_1}\sum_{i=1}^{n} q_{(i)}$$

$$\hat{\lambda}_2 = \frac{1}{2}\frac{1}{{}^{n}C_2}\sum_{i=1}^{n}\left({}^{i-1}C_1 - {}^{n-i}C_1\right)q_{(i)}$$

$$\hat{\lambda}_3 = \frac{1}{3}\frac{1}{{}^{n}C_3}\sum_{i=1}^{n}\left({}^{i-1}C_2 - 2\,{}^{i-1}C_1\,{}^{n-i}C_1 + {}^{n-i}C_2\right)q_{(i)} \qquad (44)$$

$$\hat{\lambda}_4 = \frac{1}{4}\frac{1}{{}^{n}C_4}\sum_{i=1}^{n}\left({}^{i-1}C_3 - 3\,{}^{i-1}C_2\,{}^{n-i}C_1 + 3\,{}^{i-1}C_1\,{}^{n-i}C_2 - {}^{n-i}C_3\right)q_{(i)}$$

where $q_{(i)}$, i=1,2,..,n are gauged peak flows ranked in ascending order and

$${}^{m}C_k = \begin{cases} \dfrac{m!}{k!(m-k)!} & \text{if } k \le m \\ 0 & \text{if } k > m \end{cases}$$

is the number of combinations of any k items from m items.

6.4.4 Parameter and quantile estimation

The method of L moments involves matching theoretical and sample L moments to estimate parameters. The L moments in Table 3 are replaced by their sample estimates given by eqn (44). The resulting equations are then solved to obtain estimates of the parameters. These parameters are

Table 3. L moments for several distributions (from Stedinger et al., 1993)

Family	L Moments
Generalized extreme value (GEV)	$\lambda_1 = \tau + \frac{\alpha}{\kappa}\left[1-\Gamma(1+\kappa)\right] \quad \lambda_2 = \frac{\alpha}{\kappa}\Gamma(1+\kappa)\left[1-2^{-\kappa}\right]$ $\tau_3 = \frac{2(1-3^{-\kappa})}{1-2^{-\kappa}} - 3 \quad \tau_4 = \frac{1-5(4^{-\kappa})+10(3^{-\kappa})-6(2^{-\kappa})}{1-2^{-\kappa}}$, $\kappa \neq 0$
Gumbel	$\lambda_1 = \tau + 0.5772\alpha \quad \lambda_2 = \alpha \log_e 2 \quad \tau_3 = 0.1699 \quad \tau_4 = 0.1504$
Generalized Pareto	$\lambda_1 = q_* + \frac{\beta}{1+\kappa} \quad \lambda_2 = \frac{\beta}{(1+\kappa)(2+\kappa)} \quad \tau_3 = \frac{1-\kappa}{3+\kappa} \quad \tau_4 = \frac{(1-\kappa)(2-\kappa)}{(3+\kappa)(4+\kappa)}$

used to calculate the 1 in Y AEP quantiles.

6.4.5 LH-moments for fitting the GEV distribution

When the selected probability model does not adequately fit all the gauged data the lower flows may exert undue influence on the fit and give insufficient weight to the higher flows which are the principal object of interest. To deal with this situation Wang (1997) introduced a generalization of L moments called LH moments.

LH moments are based on linear combinations of higher order-statistics. A shift parameter η=0,1,2,3… is introduced to give more emphasis on higher ranked flows. LH moments are defined as:

$$\lambda_1^\eta = E[X_{(\eta+1):(\eta+1)}]$$

$$\lambda_2^\eta = \frac{1}{2}E[X_{(\eta+2):(\eta+2)} - X_{(\eta+1):(\eta+2)}]$$

$$\lambda_3^\eta = \frac{1}{3}E[X_{(\eta+3):(\eta+3)} - 2X_{(\eta+2):(\eta+3)} + X_{(\eta+1):(\eta+3)}] \qquad (45)$$

$$\lambda_4^\eta = \frac{1}{4}E[X_{(\eta+4):(\eta+4)} - 3X_{(\eta+3):(\eta+4)} + 3X_{(\eta+2):(\eta+4)} - X_{(\eta+1):(\eta+4)}]$$

Table 4 presents the relationship between the first four LH moments and the parameters of the GEV and Gumbel distributions.

For ease of computation Wang (1997) derived the following approximation for the shape parameter κ:

$$\kappa = a_0 + a_1[\tau_3^\eta] + a_2[\tau_3^\eta]^2 + a_3[\tau_3^\eta]^3 \qquad (46)$$

where the polynomial coefficients vary with η according to Table 5.

Wang (1997) derived the following estimators for LH moments with shift parameter η:

$$\hat{\lambda}_1^\eta = \frac{1}{{}^nC_{\eta+1}}\sum_{i=1}^{n} {}^{i-1}C_\eta x_{(i)}$$

$$\hat{\lambda}_2^\eta = \frac{1}{2}\frac{1}{{}^nC_{\eta+2}}\sum_{i=1}^{n}\left({}^{i-1}C_{\eta+1} - {}^{i-1}C_\eta \, {}^{n-i}C_1\right)q_{(i)} \qquad (47)$$

Table 4. LH moments for GEV and Gumbel distributions (from Wang, 1997)

Family	LH Moments
Generalized extreme value (GEV)	$\lambda_1^\eta = \tau + \frac{\alpha}{\kappa}\left[1-\Gamma(1+\kappa)(\eta+1)^{-\kappa}\right] \qquad \lambda_2^\eta = \frac{(\eta+2)\alpha\Gamma(1+\kappa)}{2!\kappa}\left[-(\eta+2)^{-\kappa} + (\eta+1)^{-\kappa}\right]$ $\lambda_3^\eta = \frac{(\eta+3)\alpha\Gamma(1+\kappa)}{3!\kappa}\left[-(\eta+4)(\eta+3)^{-\kappa} + 2(\eta+3)(\eta+2)^{-\kappa} - (\eta+2)(\eta+1)^{-\kappa}\right]$ $\lambda_4^\eta = \frac{(\eta+4)\alpha\Gamma(1+\kappa)}{4!\kappa}\left[-(\eta+6)(\eta+5)(\eta+4)^{-\kappa} + 3(\eta+5)(\eta+4)(\eta+3)^{-\kappa} - 3(\eta+4)(\eta+3)(\eta+2)^{-\kappa} + (\eta+3)(\eta+2)(\eta+1)^{-\kappa}\right]$ where $\kappa \neq 0$
Gumbel	$\lambda_1^\eta = \tau + \alpha\left[0.5772 + \ln(\eta+1)\right] \quad \lambda_2^\eta = \frac{(\eta+2)\alpha}{2!}\left[\ln(\eta+2) - \ln(\eta+1)\right]$ $\lambda_3^\eta = \frac{(\eta+3)\alpha}{3!}\left[(\eta+4)\ln(\eta+3) - 2(\eta+3)\ln(\eta+2) + (\eta+2)\ln(\eta+1)\right]$ $\lambda_4^\eta = \frac{(\eta+4)\alpha}{4!}\left[(\eta+6)(\eta+5)\ln(\eta+4) - 3(\eta+5)(\eta+4)\ln(\eta+3) + 3(\eta+4)(\eta+3)\ln(\eta+2) - (\eta+3)(\eta+2)\ln(\eta+1)\right]$

Table 5. Polynomial coefficients for use with eqn (46)

η	a_0	a_1	a_2	a_3
0	0.2849	-1.8213	0.8140	-0.2835
1	0.4823	-2.1494	0.7269	-0.2103
2	0.5914	-2.3351	0.6442	-0.1616
3	0.6618	-2.4548	0.5733	-0.1273
4	0.7113	-2.5383	0.5142	-0.1027

$$\hat{\lambda}_3^{\eta} = \frac{1}{3}\frac{1}{{}^{n}C_{\eta+3}}\sum_{i=1}^{n}\left({}^{i-1}C_{\eta+2} - 2\,{}^{i-1}C_{\eta+1}\,{}^{n-i}C_1 + {}^{i-1}C_{\eta}\,{}^{n-i}C_2\right)q_{(i)}$$

$$\hat{\lambda}_4^{\eta} = \frac{1}{4}\frac{1}{{}^{n}C_{\eta+4}}\sum_{i=1}^{n}\left({}^{i-1}C_{\eta+3} - 3\,{}^{i-1}C_{\eta+2}\,{}^{n-i}C_1 + 3\,{}^{i-1}C_{\eta+1}\,{}^{n-i}C_2 - {}^{i-1}C_{\eta}\,{}^{n-i}C_3\right)q_{(i)}$$

The selection of the best shift parameter requires some form of goodness-of-fit test. Wang (1998) argued that the first three LH moments are used to fit the GEV model leaving the fourth LH moment available for testing the adequacy of the fit. He proposed the following approximate test statistic:

$$z = \frac{\hat{\tau}_4^{\eta} - \tau_4^{\eta}}{\sigma(\hat{\tau}_4^{\eta} \mid \hat{\tau}_3^{\eta} = \tau_3^{\eta})} \qquad (48)$$

where $\hat{\tau}_4^{\eta}$ is the sample estimate of the LH-kurtosis, τ_4^{η} is the LH-kurtosis derived from the GEV parameters fitted to the first three LH moments, and $\sigma(\hat{\tau}_4^{\eta} \mid \hat{\tau}_3^{\eta} = \tau_3^{\eta})$ is the standard deviation of $\hat{\tau}_4^{\eta}$ assuming the sample LH-skewness equals the LH-skewness derived from the GEV parameters fitted to the first three LH moments. Under the hypothesis that the underlying distribution is GEV, the test statistic z is approximately normal distributed with mean 0 and variance 1. Wang (1998) describes a simple relationship to estimate $\sigma(\hat{\tau}_4^{\eta} \mid \hat{\tau}_3^{\eta} = \tau_3^{\eta})$.

6.4.6 Parameter Uncertainty and Quantile Confidence Limits

The sampling distribution $p(\theta|D,M)$[2] can be approximated using the Monte Carlo method known as the parametric bootstrap:

1. Fit the model M to n years of gauged flows using L or LH moments to yield the parameter estimate $\hat{\theta}$.
2. Set i=1.
3. Randomly sample n flows from the fitted distribution; that is, $q_{ji} \leftarrow p(q \mid \hat{\theta}, M), j = 1,..,n$
4. Fit the model to the sampled flows $\{q_{ji}, j=1,..,n\}$ using L or LH moments to yield the parameter estimate θ_i.
5. Increment i. Go to step 3 if i does not exceed N.

This procedure yields N equi-weighted samples that approximate the sampling distribution $p(\theta|D,M)$. As a result, they can be used to quantify parameter uncertainty and estimate quantile confidence limits. However, because the parametric bootstrap assumes $\hat{\theta}$ is the true parameter, it underestimates the uncertainty and therefore should not be used to estimate expected probabilities.

6.4.7 Software

Implementation of L and LH moments requires extensive computation. The reader is referred the Australian Rainfall and Runoff web site www.arr.org.au for available software. An extensive library of FORTRAN L moment procedures can be found at www.research.ibm.com/people/h/hosking.

6.4.8 Worked Examples

Example 7 illustrates fitting the GEV distribution using L moments to a 47-year gauged record. Example 8 revisits Example 5 demonstrating the search procedure for finding the optimal shift in LH moments fitting.

6.5 Method of Product Moments Approach

The method of product moments used in conjunction with the LP3 distribution was the recommended method in Australian Rainfall and Runoff (1987). The method was simple to implement and was consistent with US practice. However, in view of its inferior performance to other methods, its use is no longer recommended with LP3 using at-site data.

7. Fitting Flood Probability Models to POT Series

In Section 2.2.3 it was noted that when interest is in events with $T_A < 10$ years it is generally preferable to work with a POT series rather than an annual series. This is because all floods (above some threshold) are of interest in this range, whether they are the highest in the particular year of record or not. The annual maximum series may omit many floods of interest see Examples 7 and 9. This section overviews procedures for fitting a probability model to a POT series.

7.1 Probability Plots

As with the analysis of annual maximum series it is recommended that a probability plot of the POT data series be prepared. The plot involves plotting an estimate of the observed ARI against the discharge. The ARI of a gauged flood can be estimated using

[2] Some liberties are taken with notation to provide a consistent treatment of uncertainty. In Section 6.3 $p(\theta|D,M)$ referred to a Bayesian posterior pdf. Here it refers to the sampling distribution of θ. Although the distributions are different, they do describe uncertainty about θ. That said, the Bayesian approach is more general in its ability to make full use of the information in the data and accept prior information.

$$T_{(i)} = \frac{n + 0.2}{i - 0.4} \quad (49)$$

where i is the rank of the gauged flood (in descending order) and n is the number of years of record.

7.2 Fitting POT Models

In some cases, it may be desirable to analytically fit a probability distribution to POT data. Two approaches can be employed.

The first employs the following steps:

1. For a given threshold q_0, fit a probability distribution with pdf $p(q|q>q_o,\theta)$ to the POT series.
2. Estimate ν the average number of POT events per unit.
3. The expected time interval between peaks that exceed w is given by eq (7) as

$$T_P(w) = \frac{1}{\nu P(Q > w \mid q_o, \theta)} .$$

4. Check that the fitted distribution is consistent with the POT data using the probability plot described in Section 7.1.

A number of authors have investigated the proposition that it is better to infer the annual flow series distribution from the POT series on the grounds that the POT series considers all relevant data. For the exponential-Poisson POT model with an average arrival rate $\nu > 1.65$ floods per year, Cunnane (1973) showed that the sample variance of T-year quantiles (T > 10 years) is less than that obtained for a Gumbel distribution fitted to annual maximum series. Jayasuriya and Mein (1985), Ashkar and Rousselle (1983) and Tavares and da Silva (1983) noted that the exponential-Poisson POT model was fairly successful, but some results diverged from the distribution derived directly from the annual series. Madsen et al. (1997) investigated this issue using a GP-Poisson POT model. They concluded that for negative shape parameters κ the POT model is generally preferable for at-site quantile estimation. However, this conclusion is dependent on the choice of estimator. Martins and Stedinger (2001) showed for negative κ that the generalized maximum likelihood estimator performs better than method of moments and L moments. Significantly, they found that the precision of flood quantiles derived from a POT model is insensitive to the arrival rate ν and for the 100-year quantile is similar to the precision obtained fitting a GEV model to annual maximum data.

At this stage, the POT approach cannot be recommended as a replacement for the analysis of annual series data. If possible, the threshold for the exponential-Poisson POT model should be selected so that the number of floods in the POT series K is at least 2 to 3 times the number of years of record N. For the GP-Poisson POT model it suffices that exceed N. However, it may be necessary to use a much lower value of K in regions with low rainfall where the number of recorded events that could be considered as floods is low.

The second approach uses a probability distribution as an arbitrary means of providing a consistent and objective fit to POT series data. For example, McDermott and Pilgrim (1982), Adams and McMahon (1985) and Jayasuriya and Mein 1985) used the LP3 distribution – they found that selecting a threshold discharge such that K equalled N was best.

8. REFERENCES

Adams, C.A. (1987) Design flood estimation for ungauged rural catchments in Victoria Road Construction Authority, Victoria. Draft Technical Bulletin.

Adams, C.A. and McMahon, T.A. (1985) Estimation of flood discharge for ungauged rural catchments in Victoria. Hydrol. and Water Resources Symposium 1985, Inst Engrs Aust., Natl Conf. Publ. No. 85/2, pp. 86-90.

Alexander, G.N. (1957) Flood flow estimation, probability and the return period. Jour. Inst. Engrs Aust, Vol. 29, pp. 263-278.

Al-Futaisi, A., and Stedinger, J.R. (1999) Hydrologic and Economic Uncertainties in Flood Risk Project Design, J. Water Resources Planning and Management, 125(6), 314-324.

Allan, R.J. (2000) ENSO and climatic variability in the last 150 years, in El Nino and the Southern Oscillation, Multi-scale variability, Global and Regional Impacts, edited by H.F. Diaz and V. Markgraf, Cambridge University Press, Cambridge, uk, p3-56.

American Society of Civil Engineers (1949) Hydrology Handbook.

Ashkanasy, N.M. and Weeks, W.D. (1975) Flood frequency distribution in a catchment subject to two storm rainfall producing mechanisms. Hydrol. Symposium 1975, Inst Engrs Aust, Natl Conf. Publ. No. 75/3, pp. 153-157.

Ashkar, F. and Rousselle, J. (1983) Some remarks on the truncation used in partial flood series models. Water Resources Research, Vol. 19, pp. 477-480.

Australian Rainfall and Runoff: A guide to flood estimation (1987) Pilgrim, D.H. (ed), The Institution of Engineers, Australia, Canberra.

Baker, V. (1984) Recent paleoflood hydrology studies in arid and semi-arid environments (abstract). EOS Trans. Amer. Geophys. Union, Vol. 65, p. 893.

Baker, V., Pickup, G. and Polach, H.A. (1983) Desert paleofloods in central Australia. Nature, Vol. 301, pp. 502-504.

Beard, L.R (1960) Probability estimates based in small normal-distribution samples. Jour. of Geophysical Research, Vol.65, pp. 2143-2148.

Beard, L.R (1974) Flood flow frequency techniques Univ. of Texas at Austin, Center for Research in Water Resources, Tech. Report CRWR119, October.

Beran, M., Hosking, J.RM. and Arnell, N. (1986) Comment on "Two-component extreme value distribution for flood frequency analysis." Water Resources Research Vol.22, pp. 263-266.

Blom, G. (1958) Statistical Estimates and Transformed Beta-Variables. Wiley, New York, 176 p.

Brown, J.A.H. (1983), Australia's Surface Water Resources - Water 2000: Consultants Report 1, Australian Government Publishing Service, Canberra.

Clarke-Hafstad, K. (1942) Reliability of station-year rainfall-frequency determinations. Trans. Amer. Soc. Civ. Engrs, Vol.107, pp. 633-652.

Conway, K.M. (1970) Flood frequency analysis of some N.S.W. coastal rivers. Thesis (M.Eng Sc.), Univ. of New South Wales.

Costa, J.E. (1978) Holocene stratigraphy in flood frequency analysis. Water Resources Research, Vol.14, pp. 626-632.

Costa, J.E. (1983) Palaeohydraulic reconstruction or flashflood peaks from boulder deposits in the Colorado Front Range. Geol. Soc. America Bulletin, Vol. 94, AP. 986-1004.

Costa, J.E. (1986) A history of paleoflood hydrology in the United States, 1800-1970. EOS Trans. Aoler Geophysical Union, Vol. 67, No. 17, pp. 425-430.

Cunnane, C. (1973) A particular comparison of annual maxima and partial duration series methods for flood frequency prediction, Journal of Hydrology, Vol. 18, 257-271.

Cunnane, C. (1978) Unbiased plotting positions - a review. Journal of Hydrology, Vol. 37, pp. 205-222.

Cunnane, C. (1985) Factors affecting choice of distribution for flood series. Hydrological Sciences Journal, Vol. 30, pp. 25-36.

Dalrymple, T. (1960) Flood-frequency analyses. Manual of Hydrology: Section 3. Flood-flow Techniques. U.S. Geological Survey Water Supply Paper 1543-A, 79 p.

DeGroot, M.H., Optimal statistcal decisions, McGraw-Hill, 1970.

Doran, D.G. and Irish, J.L. (1980) On the nature and extent of bias in flood damage estimation. Hydrol. and Water Resources Symposium 1980, Inst. Engrs Aust, Natl Conf. Publ. No. 80/9, pp. 135-139.

Duan, Q., Sorooshian, S. and V. Gupta, Effective and efficient global optimization for conceptual rainfall-runoff models, Water Resources Research, 28(4), 1015-1031, 1992.

Erskine, W. D., and Warner, R.F. (1988), Geomorphic effects of alternating flood and drought dominated regimes on a NSW coastal river, in Fluvial Geomorphology of Australia, edited by R.F. Warner, pp. 223-244, Academic Press, Sydney.

Fiorentino, M., Versace, P. and Rossi, F. (1985) Regional flood frequency estimation using the two-component extreme value distribution. Hydrol. Sciences Jour., Vol. 30, pp. 51-64.

Flavell, D.J. (1983) The Rational Method applied to small rural catchments in the south west of Western Australia Civ. Engg Trans., Inst. Engrs Aust., Vol. CE25, pp. 121-127.

Franks, S.W. and Kuczera, G. (2002) Flood Frequency Analysis: Evidence and Implications of Secular Climate Variability, New South Wales, Water Resources Research, 38(5), 10.1029/2001WR000232.

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (1995) Bayesian data analysis, Chapman and Hall, p.526.

Grayson, R.B., Argent, R.M., Nathan, R.J., McMahon, T.A. and Mein, R.G. (1996) Hydrologic Recipes: Estimation techniques in Australian Hydrology, CRC for Catchment Hydrology.

Grygier, J. C., Stedinger, J.R. and Yin, H–B. (1989) A generalized maintenance of variance extension procedure for extending correlated series, Water Resources Research, 25(3), 345–349.

Harvey, F., Doran, D.G., Cordery, I, and Pilgrim, D.H. (1991) Regional flood frequency characteristics in rural New South Wales, Proc. Int. Hydrology and Water Resources Symposium, Perth, Australia, IEAust NCP 91/19, ISBN 85825 540 5, pp 775-780.

Hatfield, E.R. and Muir, G.L. (1984), Surface water resources assessment in New South Wales, Proc. Workshop on Surface Water Resources Data, AWRC Conference Series No. 10, pp 263-292.

Hirsch, R. M. (1982) A Comparison of Four Record Extension Techniques, Water Resources Research, 18(4), 1081-1088.

Hosking, J.R.M. (1990) L-moments: Analysis and estimation of distributions using linear combinations of order statistics, J. Roy. Statist. Soc., Ser. B 52(2), 105-124.

Hosking, J.R.M. and Wallis, J.R (1986) Paleoflood hydrology and flood frequency analysis. Water Resources Research, Vol. 22, pp. 543-550.

Houghton, J.C. (1978) Birth of a parent: The Wakeby distribution for modeling flood flows. Water Resources Research, Vol. 14, pp. 1105-1109.

Interagency Advisory Committee on Water Data (1982) Guidelines for determining flood flow frequency. Bulletin 17B of the Hydrology Sub-committee, Office of Water Data Coordination, Geological Survey, U.S. Dept of the Interior.

Jayasuriya, M.D.A. and Mein, RG. (1985) Frequency analysis using the partial series. Hydrol. and Water Resources Symposium 1985, Inst. Engrs Aust., Natl Conf. Publ. No. 85/2, pp. 81-85.

Jin, M. and Stedinger, J.R. (1989) Flood Frequency Analysis with Regional and Historical Information, Water Resources Research, 25(5), 925-936.

Kiem, A. S., and Franks, S.W. (2004) Multidecadal variability of drought risk - eastern Australia, Hydrol. Proc., 18, doi:10.1002/hyp.1460.

Kiem, A.S., Franks, S.W. and Kuczera, G. (2003) Multi-decadal variability of flood risk, Geophysical Research Letters, 30(2), 1035, DOI:10.1029/2002GL015992.

Kochel. RC., Baker, V.R and Patton, P.C. (1982) Paleohydrology of southwestern Texas. Water Resources Research, Vol. 18, pp. 1165-1183.

Kopittke, RA., Stewart, B.J. and Tickle, K.S. (1976) Frequency analysis of flood data in Queensland. Hydrol. Symposium 1976, Inst Engrs Aust., Natl Conf. Publ. No. 76/2, pp. 20-24.

Kuczera, G. (1999) Comprehensive at-site flood frequency analysis using Monte Carlo Bayesian inference, Water Resources Research, 35(5), 1551-1558.

Kuczera, G., Lambert, M., Heneker, T., Jennings, S., Frost, A. and Coombes, P. (2003) Joint probability and design storms at the crossroads, Hydrology and Water Resources Symposium, Institution of Engineers, Australia, Wollongong.

Lee, P.M. (1989) Bayesian statistics: An introduction, Oxford University Press (NY).

Laurenson, E.M. (1987) Back to basics on flood frequency analysis. Civ. Engg Trans., Inst Engrs Aust, Vol. CE29, pp. 47-53.

Madsen, H., Pearson, C.P. Rasmussen, P.F. and Rosbjerg, D. (1997) Comparison of Annual Maximum Series and Partial Duration Series Methods for Modeling Extreme Hydrologic Events 1. At-Site Modeling, Water Resources Research, 33(4), 747-758.

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis (1997), A Pacific interdecadal climate oscillation with impacts on salmon production, Bull. Amer. Meteorol. Soc., 78(6), 1069-1079.

Martins, E. S. and Stedinger, J.R. (2000) Generalized Maximum Likelihood GEV Quantile Estimators for Hydrologic Data, Water Resources Research, 36(3), 737-744.

Martins, E. S. and Stedinger, J.R. (2001) Generalized Maximum Likelihood Pareto-Poisson Flood Risk Analysis for Partial Duration Series, Water Resources Research, 37(10), 2559-2567.

Matalas, N.C. and Jacobs, B. (1964) A correlation procedure for augmenting hydrologic data. U. S. Geological Survey Professional Paper 434-E.

McDermott, G.E. and Pilgrim, D. H. (1982) Design flood estimation for small catchments in New South Wales. Dept of National Development and Energy, Aust Water Resources Council Tech. Paper No. 73, 233 p.

McDermott, G.E. and Pilgrim, D.H. (1983) A design flood method for arid western New South Wales based on bankfull estimates. Civ. Engg Trans., Inst Engrs Aust, Vol. CE25, pp. 114-120.

McIllwraith, J.F. (1953) Rainfall intensity-frequency data for New South Wales stations. Jour. Inst Engrs Aust, Vol. 25, pp. 133-139.

McMahon, T.A. (1979) Hydrologic characteristics of Australian streams. Civ. Engg Research Reports, Monash Univ., Report No. 3/1979.

McMahon, T.A. and Srikanthan, R (1981) Log Pearson III distribution- is it applicable to flood frequency analysis of Australian streams? Jour. of Hydrology, Vol.52, pp. 139-147.

Micevski, T., Kiem, A.S., Franks, S.W. and Kuczera, G. (2003) Multidecadal Variability in New South Wales Flood Data, Hydrology and Water Resources Symposium, Institution of Engineers, Australia, Wollongong.

Micevski T., G. Kuczera, S. W. Franks (2005), Flood frequency censoring errors associated with daily-read flood observations, Water Resour. Res., 41, W07002, doi:10.1029/2004WR003881.

Natural Environment Research Council (1975) Flood Studies Report, Vol. 1, Hydrological Studies. London.

O'Connell, D.R., Ostemaa, D.A., Levish, D.R. and Klinger, R.E. (2002) Bayesian flood frequency analysis with paleohydrologic bound data, Water Resources Research, 38(5), 1058, DOI:10.1029/2000WRR000028.

Pilgrim, D.H. and Doran, D.G. (1987) Flood frequency analysis, in Australian Rainfall and Runoff: A guide to flood estimation, Pilgrim, D.H. (ed), The Institution of Engineers, Australia, Canberra.

Pilgrim, D.H. and McDermott, G.E. (1982) Design floods for small rural catchments in eastern New South Wales. Civ. Engg Trans., Inst Engrs Aust., Vol. CE24, pp. 226-234.

Potter, D.J. and Pilgrim, D.H. (1971) Flood estimation using a regional flood frequency approach. Final Report, Vol. 2, Report on Analysis Components. Aust. Water Resources Council, Research Project 68/1, Hydrology of Small Rural Catchments. Snowy Mountains Engg Corporation, April.

Potter, K.W. and Walker, J.F. (1981) A model of discontinuous measurement error and its effects on the probability distribution of flood discharge measurements, Water Resources Research, 17(5), 1505-1509.

Potter, K.W. and Walker, J.F. (1985) An empirical study of flood measurement error, Water Resources. Research, 21(3), 403-406.

Power, S., F. Tseitkin, S. Torok, B. Lavery, R. Dahni, and B. McAvaney (1998), Australian temperature, Australian rainfall and the Southern Oscillation, 1910-1992: coherent variability and recent changes, Aust. Met. Mag., 47(2), 85-101.

Power, S., T. Casey, C. Folland, A. Colman, and V. Mehta (1999), Inter-decadal modulation of the impact of ENSO on Australia, Climate Dynamics,15(5), 319-324.

Rosbjerg, D., Madsen, H. and Rasmussen, P.F. (1992) Prediction in partial duration series with generalized Pareto-distributed exceedances, Water Resources Research, 28(11), 3001-3010.

Rossi, F., Fiorentino, M. and Versace, P. (1984) Two-component extreme value distribution for flood frequency analysis. Water Resources Research, Vol.20, pp.847-856.

Slack, J.R., Wallis, J.R. and Matalas, N.C. (1975) On the value of information in flood frequency analysis, , Water Resources. Research, 11(5), 629-648.

Stedinger, J.R. (1983) Design events with specified flood risk, Water Resources Research, 19(2), 511-522.

Stedinger, J.R., Vogel, R.M. and Foufoula-Georgiou, E. (1993) Frequency analysis of extreme events in Handbook of Hydrology, Maidment, D.R. (ed.), McGraw-Hill, NY.

Stedinger, J.R and Cohn, T.A. (1986) Flood frequency analysis with historical and paleoflood information. Water Resources Research, Vol. 22, pp. 785-793.

Stedinger, J.R. (1998) Expected Probability and Annual Damage Estimators, J. of Water Resources Planning and Management, 123(2), 125-35. [With discussion, Leo R. Beard, J. of Water Resources Planning and Management, 124(6), 365-366, 1998.]

Tavares, L.V. and da Silva, J.E. (1983) Partial series method revisited. Jour. of Hydrology, Vol. 64, pp. 1-14.

Wallis, J.R and Wood, E.F. (1985) Relative accuracy log Pearson III procedures. Proc. Amer. Soc. Civ. Engrs. Jour. of Hydraulic Eng, Vol. lll, No.7, pp.1043-10

Wang, Q. J. (1996) Direct sample estimators of L-moments, Water Resources Research, 32(12), 3617-3619, 1996.

Wang, Q. J. (1997) LH moments for statistical analysis of extreme events, Water Resources Research, 33(12), 2841-2848.

Wang, Q. J. (1998) Approximate goodness-of-fit tests of fitted generalized extreme value distributions using LH moments, Water Resources Research, 34(12), 3497-3502.

Wang, Q.J. (1999)

Wang, Q.J. (2001) A Bayesian joint probability approach for flood record augmentation, Water Resources Research, 37(6), pp. 1707-1712.

Waylen, P. and M.-K. Woo, M.-K. (1982) Prediction of annual floods generated by mixed processes, Water Resources Research, 18(4), 1283-1286.

EXAMPLE 1: EXTRAPOLATION AND PROCESS UNDERSTANDING

The importance of process understanding when extrapolating beyond the observed record is illustrated by a simple Monte Carlo experiment. A Poisson rectangular pulse rainfall model is used to generate a long record of high resolution rainfall. This is routed through a rainfall-runoff model to generate runoff into the stream system. The storage-discharge relationship for the stream is depicted by the bilinear relationship shown in Figure E1-1. A feature of this relationship is the activation of significant flood terrace storage once a threshold discharge is exceeded.

The routing model parameters were selected so that major flood terrace storage is activated by floods with an ARI in excess of 100 years. This situation was chosen to represent a river with multiple flood terraces with the lowest terraces accommodating the majority of floods and the highest terrace only inundated by extreme floods.

Figure E1-2 presents the flood frequency curve based on 30000 simulated years – it shows a clear break in slope around the 100 year ARI corresponding to the activation of major flood terrace storage. Indeed the flood frequency curve displays downward curvature despite that the fact the rainfall frequency curve displays upward curvature in the 100 to 1000 year ARI range. In contrast the flood frequency curve based on 100 years of "data" shows no evidence of downward curvature. This is because in a 100-year record there is little chance of the major flood terrace storage being activated. Indeed without knowledge of the underlying hydraulics one would be tempted to extrapolate the 100-year flood record using a straight line extrapolation. Such an extrapolation would rapidly diverge from the "true" frequency curve.

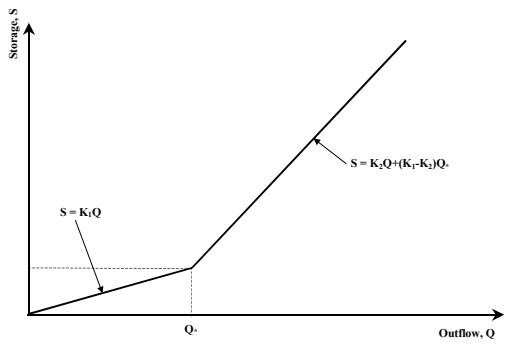

Figure E1-1. Bilinear channel storage-discharge relationship

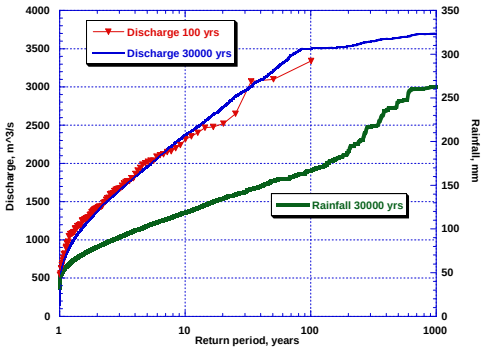

Figure E1-2. Simulated rainfall and flood frequency curves with major floodplain storage activated at a threshold discharge of 3500 m^3/s

Although the example idealizes the dominant rainfall-runoff dynamics it delivers a very strong message. Extrapolation of flood frequency curves fitted to gauged flow records requires the exercise of hydrologic judgment backed up by appropriate modelling. The problem of extrapolation is much more general. For example, in this example, if a rainfall-runoff approach were used with the rainfall-runoff model calibrated to small events the simulated flood frequency curve is likely to be compromised in a similar way.

EXAMPLE 2: FITTING A PROBABILITY MODEL TO GAUGED DATA

This example illustrates fitting a probability model to gauged annual maximum flood data. The table lists 31 years of gauged annual maximum discharges (m^3/s) for the Hunter river at Singleton:

76.19	171.7	218.0	668.2	1373.	124.0	276.0	894.8
1373.	279.9	202.4	4049.	2321.	2534.	3313.	1231.
1390.	12515	1098.	447.4	478.5	180.3	164.2	229.3
2123.	965.5	2749.	48.98	76.45	911.7	925.9	

An LP3 distribution was fitted using the Bayesian approach. The table below presents the posterior mean, standard deviation and correlation for the LP3 parameters: m, s and g which are respectively the mean, standard deviation and skewness of $\log_e q$:

LP3 parameter	Mean	Std deviation	Correlation		
m	6.426	0.236	1.000		
$\log_e$ s	0.350	0.127	0.046	1.000	
g	0.146	0.643	0.000	0.068	1.000

Figure E2-1 plots on a log normal probability plot the gauged flows, the 1 in Y AEP quantile curve (derived using the posterior mean parameters), the 90% quantile confidence limits and the expected probability curve.

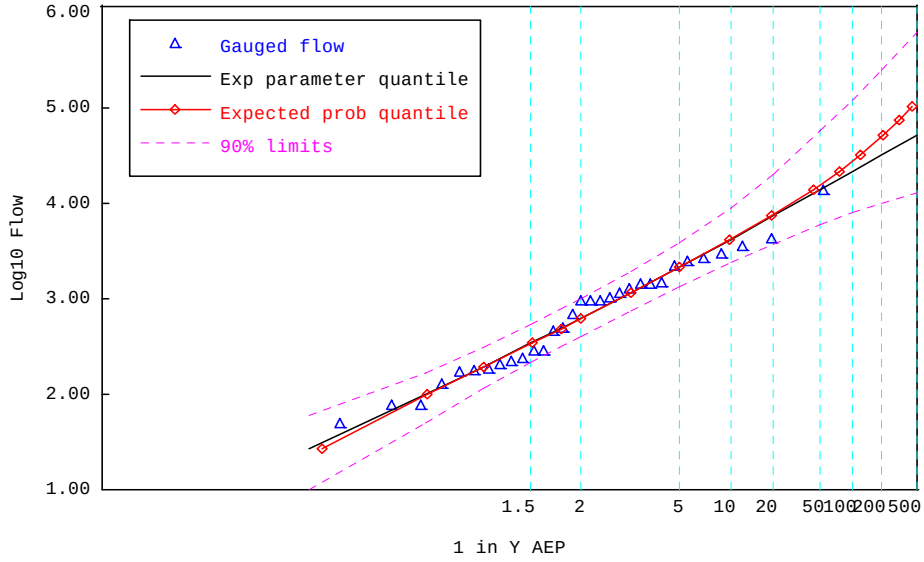

Figure E2-1. Bayesian LP3 fit to 31 years of gauged annual maximum floods.

The good fit to the gauged data and the tight confidence limits needs to be tempered by the fact that Figure E2-1 plots the logarithm of the flood peaks. The table of selected 1 in Y AEP quantiles q_Y and their 90% confidence limits presents a more sobering perspective. For example, for the 1 in 100 AEP flood the 5% and 95% confidence limits are respectively 38% and 553% of the quantile q_Y! The 1 in 500 AEP confidence limits are so wide as to render estimation meaningless. Note the expected AEP for the quantile q_Y consistently exceeds the nominal 1 in Y AEP. For example, the 1 in 100 AEP quantile of 19548 m^3/s has an expected AEP of 1 in 74.

1 in Y AEP	Quantile q_Y	Quantile confidence limits		Expected AEP for q_Y
		5% limit	95% limit	
10	3888	2222	8287	1/9.9
50	12729	5537	52157	1/43
100	19548	7372	108144	1/74
500	47366	11799	549773	1/210

EXAMPLE 3: USE OF CENSORED HISTORICAL DATA

This example is a continuation of Example 2. It illustrates the benefit of using historical flood information. The gauged record spanned 1938 to 1969. The biggest flood in that record occurred in 1955. An examination of historic records indicates that during the ungauged period 1820 to 1937 there was only one flood that exceeded the 1955 flood – this flood occurred in 1820. This represents valuable information even though the magnitude of the 1820 flood is not reliably known. This is an example of censored data. Over the ungauged period 1820 to 1937 there was one flood above and 117 floods below the threshold discharge corresponding to the 1955 flood.

An LP3 distribution was fitted to gauged and censored data using the Bayesian approach. The table below presents the posterior mean, standard deviation and correlation for the LP3 parameters: m, s and g which are respectively the mean, standard deviation and skewness of $\log_e$ q. Comparison with Example 3 reveals the censored data have reduced by almost 30% the uncertainty in the skewness g, the parameter that controls the shape of the distribution, particularly in the tail region.

LP3 parameter	Mean	Std deviation	Correlation		
m	6.359	0.226	1.000		
$\log_e$ s	0.304	0.116	-0.139	1.000	
g	0.001	0.458	-0.261	-0.487	1.000

Figure E3-1 plots on a log normal probability plot the gauged flows, the 1 in Y AEP quantile curve (derived using the posterior mean parameters), the 90% quantile confidence limits and the expected probability curve. Compared with Example 2 the tightening of the confidence limits is most noticeable.

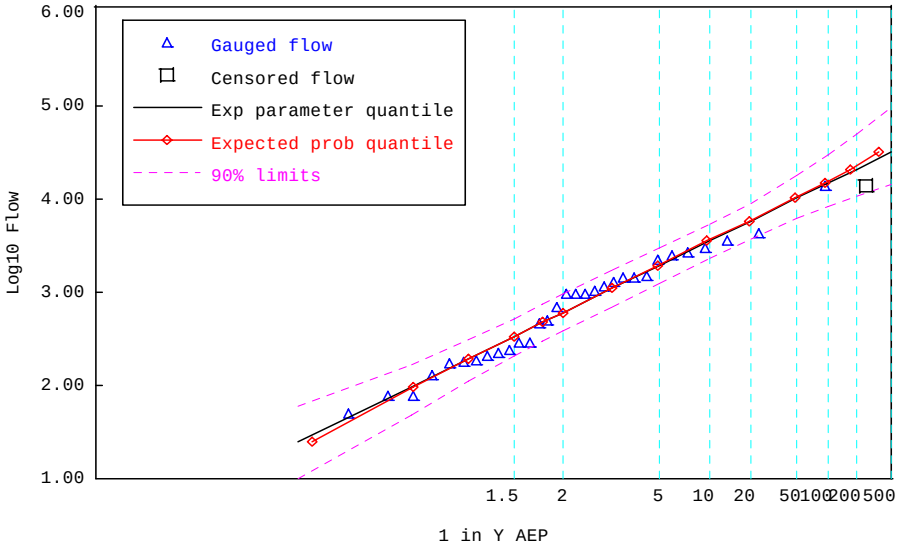

Figure E4-1. Bayesian LP3 fit to 31 years of gauged annual maximum floods.

The table of selected 1 in Y AEP quantiles q_Y and their 90% confidence limits illustrates the benefit of the information contained in the censored data. For example, for the 1 in 100 AEP flood the 5% and 95% confidence limits are respectively 57% and 200% of the quantile q_Y! This represents a major reduction in quantile uncertainty compared with Example 3 which yielded limits of 38% and 553%.

1 in Y AEP	Quantile q_Y	Quantile confidence limits		Expected AEP for q_Y
		5% limit	95% limit	
10	3281	2188	5006	1/9.6
50	9351	5786	16416	1/48
100	13535	7751	27100	1/93
500	28615	12790	87327	1/363

This example highlights the significant reductions in uncertainty that historical data can offer. However, care must be exercised ensuring the integrity of the historic information – see Section 3.8 for more details.

EXAMPLE 4: USE OF REGIONAL INFORMATION

This is a continuation of Example 2. In Example 2 the posterior mean of the skewness was estimated to be 0.146 with a posterior standard of 0.643. The uncertainty in the 1 in 100 AEP quantile was large.

Suppose hypothetically that a regional analysis of skewness was conducted. Furthermore suppose the expected regional skew is 0.00 with a standard deviation of 0.35. Following Section 6.3.5 this information can be approximated as a normally distributed prior distribution for the skewness. Figure E4-1 presents the probability plot for the LP3 model fitted to the gauged data with prior information on the skewness. Comparison with Figure E2-1 reveals substantially improved accuracy in the right hand tail.

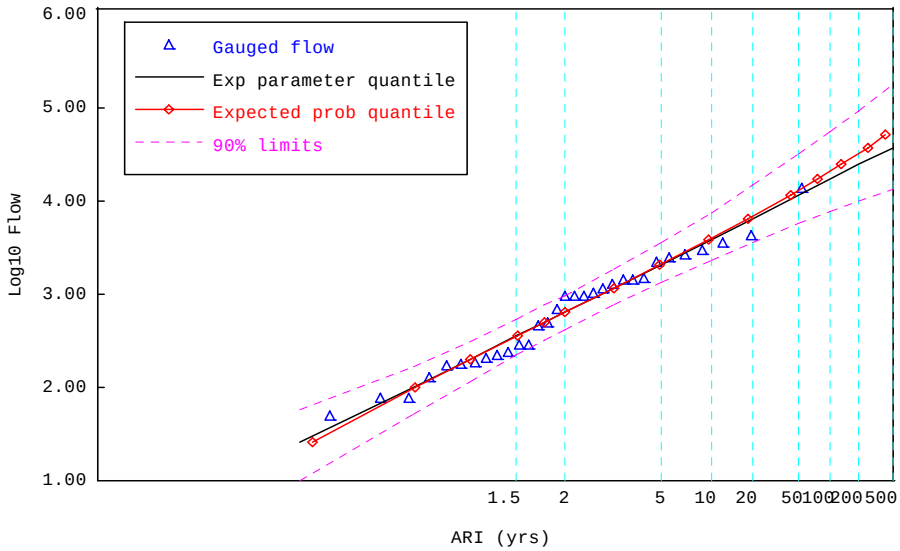

Figure E4-1. Bayesian LP3 fit to 31 years of gauged annual maximum floods with prior information on skewness.

The table compares the LP3 fitted with and without prior information on skewness. The posterior uncertainty on the skewness is about 88% of the prior standard deviation indicating the gauged data are not very informative about the shape parameter of the flood distribution.

	No prior information		With prior information	
LP3 parameter	Mean	Std deviation	Mean	Std deviation
m	6.426	0.236	6.423	0.236
$\log_e$ s	0.350	0.127	0.322	0.127
g	0.146	0.643	0.028	0.305

The table of selected 1 in Y AEP quantiles q_Y and their 90% confidence limits further illustrates the benefit of incorporating regional information. For example, for the 1 in 100 AEP flood the 5% and 95% confidence limits are respectively 38% and 553% of the quantile q_{100} when no prior information is used. These limits are reduced to 45% and 321% using prior regional information.

	No prior information			With prior information		
1 in Y AEP	Quantile q_Y	Quantile confidence limits		Quantile q_Y	Quantile confidence limits	
		5% limit	95% limit		5% limit	95% limit
10	3888	2222	8287	3619	2184	6933
50	12729	5537	52157	10677	5294	29104
100	19548	7372	108144	15670	7050	50274

EXAMPLE 5: IMPROVING POOR FITS USING CENSORING OF LOW FLOW DATA

The standard probability models such as GEV and LP3 may not adequately fit flood data for a variety of reasons. Often the poor fit is associated with a sigmoidal probability plot as illustrated in Figure E5-1. In such cases one can employ four or five-parameter distributions which have sufficient degrees of freedom to track the data in both upper and lower tails of the sigmoidal curve or adopt a calibration approach that gives less weight to smaller floods. The latter approach is illustrated in this example which considers 50 annual maximum floods (m^3/s) for the Albert River at Broomfleet presented in the following table:

13.02	15.57	15.85	16.70	22.36	36.51	72.73	78.11	87.73	88.30
95.65	99.90	113.77	116.88	124.52	131.03	156.22	156.50	190.74	210.55
220.74	249.61	249.61	271.68	285.83	302.81	305.64	362.24	384.88	461.29
466.95	676.37	752.78	761.27	761.27	860.32	863.15	865.98	1086.72	1177.28
1185.77	1214.07	1273.50	1327.27	1341.42	1364.06	1468.77	1652.72	1689.51	1765.92

Figure E5-1 displays the GEV Bayesian fit on a Gumbel probability plot. Although the observed floods are largely contained within the 90% confidence limits, the fit, nonetheless, is poor – the data exhibit a sigmoidal trend with reverse curvature developing for floods with an AEP less than 1 in 2. It appears that the confidence limits have been inflated because the GEV fit represents a poor compromise.

To deal with this poor fit, low flows can be censored to obtain a fit that favours the right hand tail of the distribution. The main idea is to force the GEV to only fit the probability of low flows falling below a threshold rather than the low flow part of the distribution – this allows the GEV to focus on the right hand tail. The censoring threshold can be obtained by trial-and-error but is better obtained by fitting LH moments as in Example 8.

Figure E5-2 illustrates one such fit. To de-emphasize the left hand tail the floods below the threshold of 250 m^3/s were censored. This means the GEV distribution was fitted to:

- a gauged record consisting of the 27 floods above 250 m^3/s; and
- a censored record consisting of 23 floods below the threshold of 250 m^3/s and 0 floods above this threshold.

The censored record provides an anchor point for the GEV distribution – it ensures that the chance of an annual maximum flood being less than 250 m^3/s is about 23/50 without forcing the GEV to fit the peaks below the 250 m^3/s threshold. The fit effectively disregards floods with an AEP greater than 1 in 2 and provides a good fit to the upper tail. Another benefit is the substantially reduced 90% confidence range.

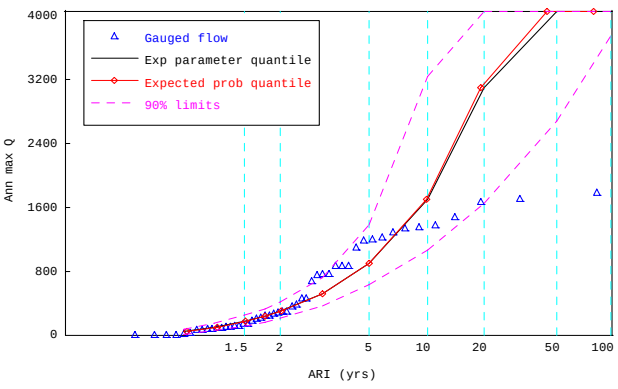

Figure E5-1. Bayesian fit to all gauged data

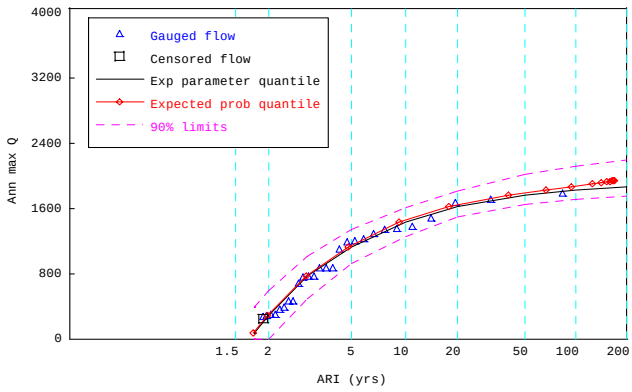

Figure E5-2. Bayesian fit with floods below 250 m^3/s threshold treated as censored observations.

EXAMPLE 6: EXAMPLE OF A NON-HOMOGENEOUS FLOOD PROBABILITY MODEL

The work of Micevski et al. (2003) illustrates an example of a non-homogeneous model. An indicator time series based on the IPO time series (Figure 6) was used to create the exogenous vector x

$x = \{I_t, t = 1,..,n\}$ where the indicator $I_t = \begin{cases} 1 \text{ if } IPO_t \geq IPO_{thresh} \\ 0 \text{ if } IPO_t < IPO_{thresh} \end{cases}$, IPO_t is the IPO index for year t and IPO_{thresh} is a threshold value equal to –0.125.

At each of the 33 NSW sites considered by Micevski et al. the annual maximum peak flows were stratified according to the indicator I_t. A 2-parameter log-normal distribution was fitted to the gauged flows with indicator equal to 1 – this is the IPO+ distribution. Likewise, a 2-parameter log-normal distribution was fitted to the gauged flows with indicator equal to 0 – this is the IPO- distribution. Figure E6-1 presents the histogram for the ratio of the IPO- and IPO+ floods for selected 1 in Y AEPs. If the the IPO+ and IPO- distributions were homogeneous then about half of the sites should have a flood ratio < 1 – Figure E6-1 shows otherwise.

Figures E6-2 and E6-3 present log normal fits to the IPO+ and IPO- annual maximum flood data for the Clarence river at Lilydale respectively. Though the adequacy of the log normal model to fit high floods may be questioned, in the AEP range 1 in 2 to 1 in 10 years the IPO- floods are about 2.6 times the IPO+ floods with the same AEP.

To avoid bias in estimating long-term flood risk it is essential that the gauged record adequately span both IPO+ and IPO- years. In this example, the IPO+ record is 43 years and the IPO- record is 33 years in length. With reference to Figure 6 this length of record appears to adequately sample both IPO epochs. This suggests that fitting to all the data will yield a largely unbiased estimate of the long-term flood risk. Figure E6-4 illustrates a log normal fit to all the data.

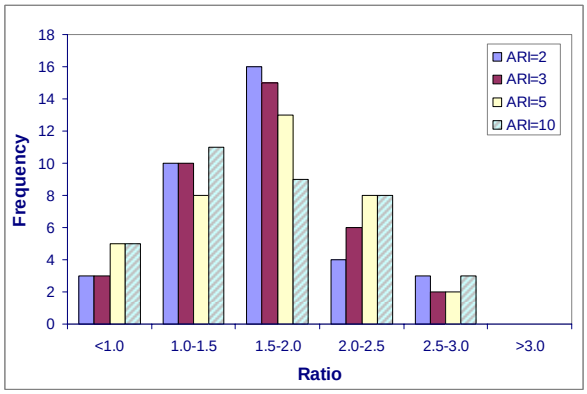

Figure E6-1. Histogram of IPO- and IPO+ flood ratios.

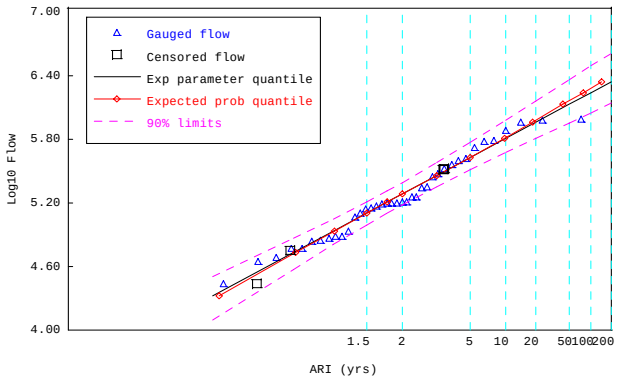

Figure E6-2. Log-normal fit to 43 years of IPO+ data for the Clarence river at Lilydale (units ML/day).

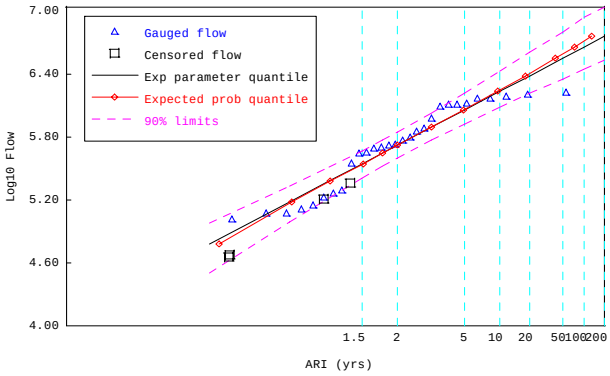

Figure E6-3. Log-normal fit to 33 years of IPO- data for the Clarence river at Lilydale (units ML/day).

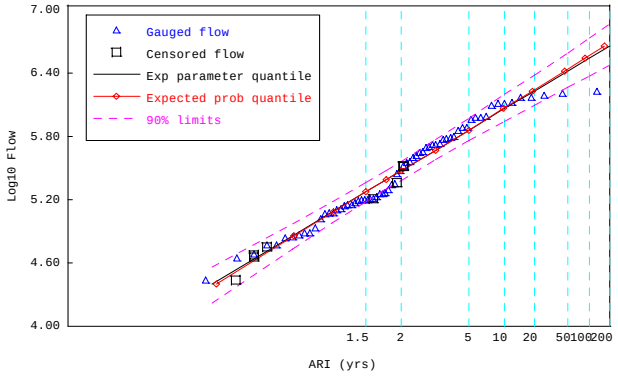

Figure E6-4. Log-normal fit to 76 years of data for the Clarence river at Lilydale (units ML/day).

EXAMPLE 6 (continued): EXAMPLE OF A NON-HOMOGENEOUS FLOOD PROBABILITY MODEL

A better appreciation of the differences in flood risk can be gleaned by considering Figure E6-5 which presents the fitted log normal distributions to the IPO+, IPO- and total data. During an IPO+ period a flood with peak discharge of 100 m^3/s has a 1 in 20 AEP while during an IPO- period it has a 1 in 4 AEP. Likewise a flood with peak discharge of 200 m^3 /s has 1 in 100 and 1 in 10 AEPs for IPO+ and IPO- periods respectively. The differences in flood risk are considerable. If a short gauged record falling largely in the IPO+ period was used, a standard flood frequency analysis could seriously underestimate the long-term or marginal flood risk.

The marginal flood risk can be derived by combining the IPO+ and IPO- distribution using eqn (14) to give

$$P(Q \le q \mid \theta) = P(x = 0)\int_0^q p(z \mid \theta, x = 0)\,dz + P(x = 1)\int_0^q p(z \mid \theta, x = 1)\,dz$$

The exogenous variable x can take two values, 0 or 1, depending on the IPO epoch. P(x=0), the probability of being in an IPO- epoch, is assigned the value 33/76 based on the observation that 33 of the 76 years of record were in the IPO- epoch. Likewise P(x=1), the probability of being in an IPO+ epoch, is assigned the value 43/76.

The derived marginal distribution is plotted in Figure E6-5. It almost exactly matches the log normal distribution fitted to all the data.

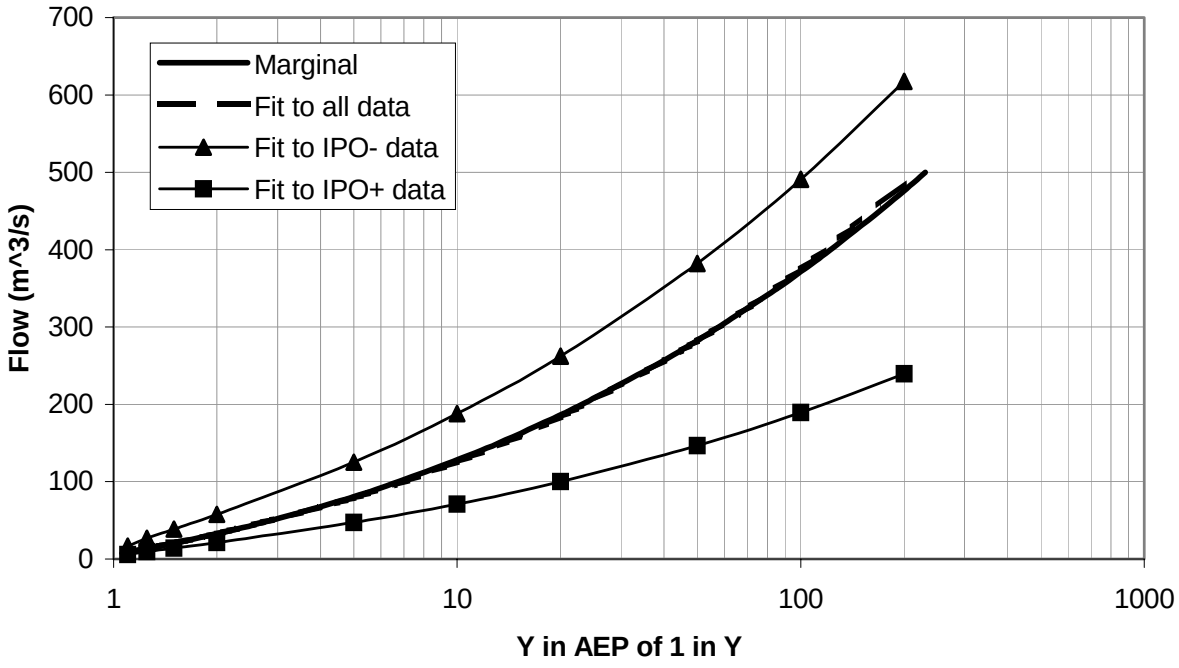

Figure E6-5. Marginal, IPO+ and IPO+ log-normal distributions for the Clarence river at Lilydale.

EXAMPLE 7: L MOMENTS FIT TO GAUGED DATA

This example illustrates fitting a GEV distribution to gauged data using L moments. The table lists 47 ranked annual maximum flood for the Styx River at Jeogla.:

878	541	521	513	436	411	405	315
309	300	294	258	255	235	221	220
206	196	194	190	186	177	164	126
117	111	108	105	92.2	88.6	79.9	74
71.9	62.6	61.2	60.3	58	53.5	39.1	26.7
26.1	23.8	22.4	22.1	18.6	13	8.18	

The following table reports the results. The GEV parameters are estimated by substituting the L moment estimates in eqn (45) to obtain κ and then using the equations in Table 3 to estimate τ and α. The standard deviation and correlation were derived from 5000 bootstrapped samples following the procedure described in Section 6.4.6.

L moment estimates		GEV parameter	Parameter estimate	Std deviation	Correlation		
$\hat{\lambda}_1$	189.238	τ	100.660	17.657	1.000		
$\hat{\lambda}_2$	92.476	α	104.157	15.554	0.597	1.000	
$\hat{\lambda}_3$	29.264	κ	-0.219	0.130	0.358	0.268	1.000

Figure E7-1 presents a probability plot of the gauged data along with GEV 1 in Y AEP quantiles and their 90% confidence limits. The confidence limits widen appreciably for the bigger floods – this is largely due to there being insufficient information to accurately infer the shape parameter κ.

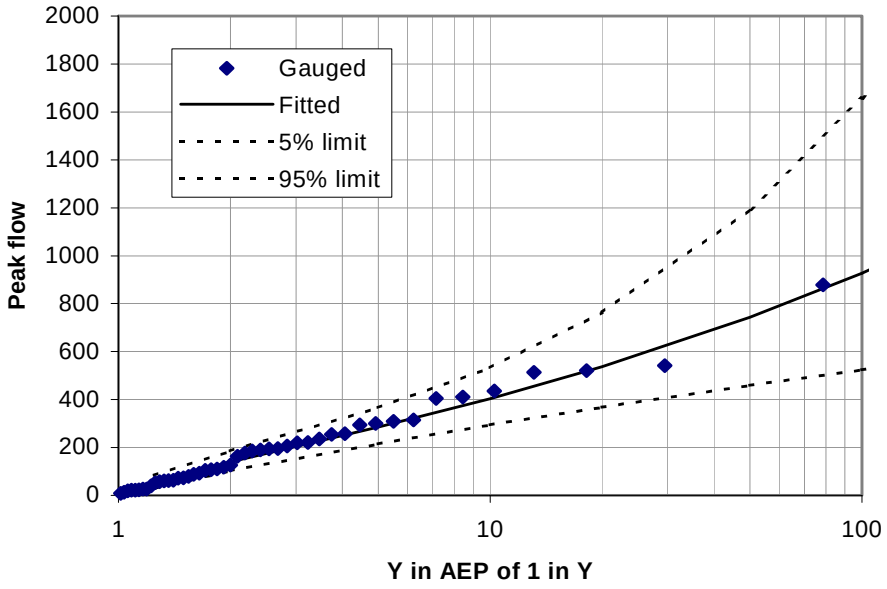

Figure E7-1. Probability plot for Styx river at Jeogla.

EXAMPLE 8: IMPROVING POOR FITS USING LH MOMENTS

This example is a continuation of Example 5. The GEV distribution was fitted to the gauged data for the Albert River at Broomfleet using LH moments.

Figure E8-1 displays the GEV L moment fit on a Gumbel probability plot. Although the observed floods are largely contained within the 90% confidence limits, the fit, nonetheless, is poor with systematic departures from the data which exhibit reverse curvature.

To deal with this poor fit, a LH moment search was conducted to find the optimal shift parameter η using the procedure described in Section 6.4.5. The optimal shift was found to be 4. Figure E8-2 presents the LH moment fit with shift η equal to 4. The fit effectively disregards floods with an AEP in excess of 1 in 2 and provides a very good fit to upper tail.

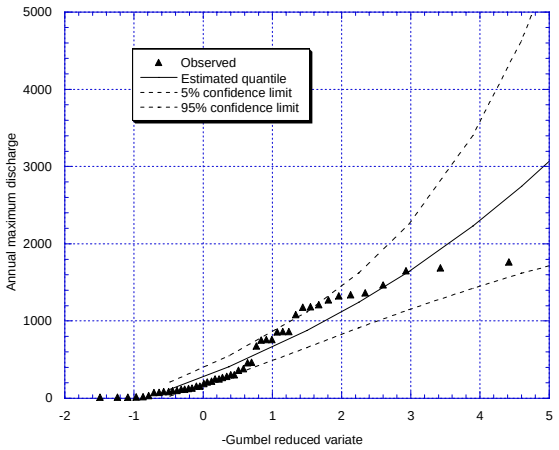

Figure E9-1. L moment fit

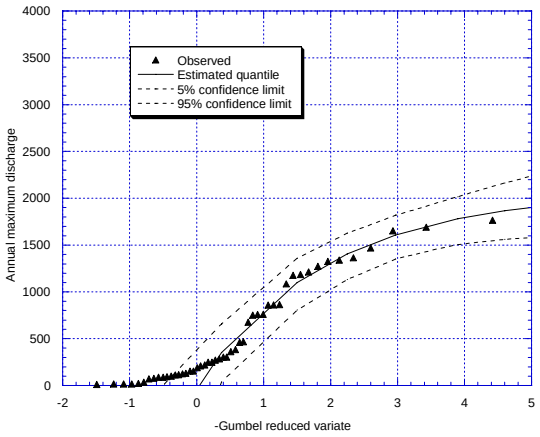

Figure E9-2. LH moment fit with shift $\eta = 4$

The very significant reduction in the width of the quantile confidence intervals is largely due to the shape parameter κ changing from –0.17 to 0.50. The L moment fit in Figure E8-1 was a compromise with the bulk of the small and medium-sized floods suggesting an upward curvature in the probability plot – as a result the GEV shape parameter κ had to be negative to enable upward curvature. In contrast, the LH moment fit favoured the large-sized floods which exhibit a downward curvature resulting in a positive shape parameter. For positive κ the GEV has an upper bound. In this case the upper bound is about 2070 m^3/s which is only 17% greater than the largest observed flood.

EXAMPLE 9: FITTING A PROBABILITY MODEL TO POT DATA

This example illustrates fitting an exponential distribution to POT data using data for the Styx River at Jeogla. The table lists all the independent peak flows recorded over a 47 year period that exceeded a threshold of 74 m^3/s – the total number of peaks was 47. Comparison with the annual maximum flood peaks in Example 7 reveals that in 15 of the 47 years of record the annual maximum peak was below the threshold of 74 m^3/s.

878	541	521	513	436	411	405	315
309	301	300	294	283	258	255	255
238	235	221	220	206	196	194	190
186	164	150	149	134	129	129	126
119	118	117	117	111	108	105	98
92.2	92.2	91.7	88.6	85.2	79.9	74	

A GP distribution with distribution function

$$P(Q \le q \mid \theta) = 1 - \left(1 - \frac{\kappa(q - q_*)}{\beta}\right)^{\frac{1}{\kappa}}$$

was fitted to the POT data yielding the following estimates: $\beta = 148.7$, κ=-0.024 and q_*=73.999. Using eqn (7) the average recurrence interval between two peaks exceeding a magnitude of w is

$$T_P(w) = \frac{1}{\nu P(Q > w)} = \frac{1}{\nu\left[1 - \frac{\kappa(w - q_*)}{\beta}\right]^{1/\kappa}}$$

where ν is the average number of flood peaks above the threshold q_* per year. Given that 47 peaks above the threshold occurred in 47 years, the best estimate of ν equals 1.0. Figure E9-1 displays a plot of the fitted POT exponential model against the observed POT series obtained using eqn (49).

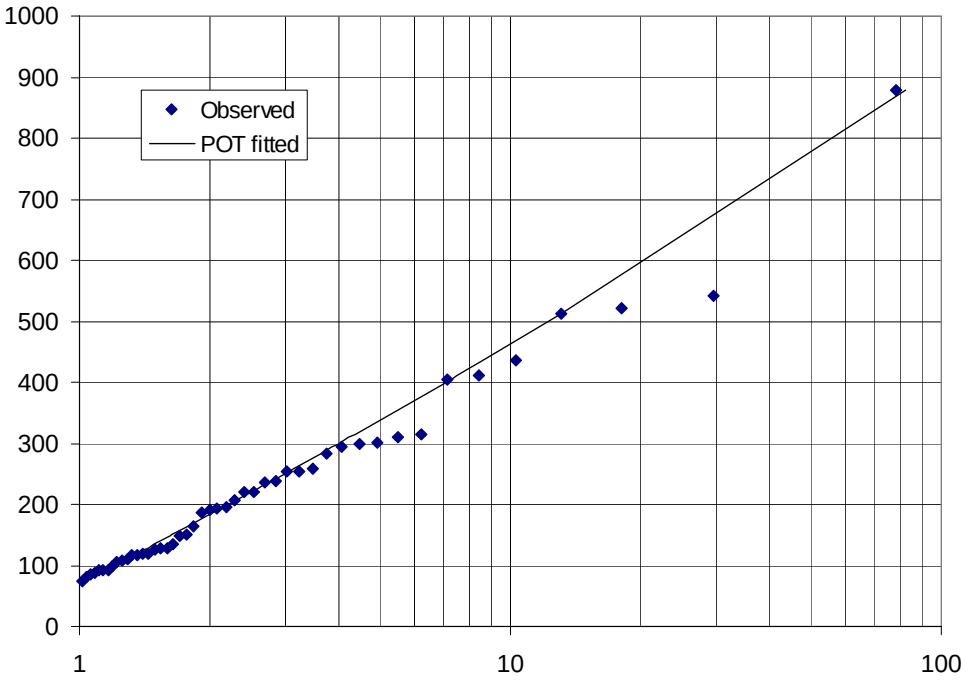

Figure E9-1. POT probability plot for Styx river at Jeogla.