A direct analysis of flood interval probability using approximately 100-year streamflow datasets

Ian M. Brodie[1] and Shahjahan Khan[2]

[1]*School of Civil Engineering and Surveying, International Centre for Applied Climate Sciences, University of Southern Queensland, Toowoomba, Australia. Email* brodiei@usq.edu.au
[2]*School of Agricultural, Computational and Environmental Sciences, International Centre for Applied Climate Sciences, University of Southern Queensland, Toowoomba, Australia. Email* shahjahan.khan@usq.edu.au

Received XXXX; accepted XXXX

Abstract Series of observed flood intervals, defined as the time intervals between successive flood peaks over a threshold, were extracted directly from eleven approximately 100-year streamflow datasets from Queensland, Australia. A range of discharge thresholds were analysed that correspond to approximately 3.7 months to 6.3 year return periods. Flood interval histograms at South East Queensland gauges were consistently unimodal whereas those of the North and Central Queensland sites were often multimodal. The exponential probability distribution (pd) is often used to describe interval exceedance probabilities, but fitting utilizing the Anderson Darling statistic found little evidence that it is the most suitable. The fatigue life pd dominated sub-year return periods (<1 year), often transitioning to a log Pearson 3 pd at above-year return periods. Fatigue life pd is used in analysis of the life time to structural failure when a threshold is exceeded and this paper demonstrates its relevance also to the elapsed time between above-threshold floods. At most sites, the interval medians were substantially less than the means for sub-year return periods. Statistically the median is a better measure of central tendency of skewed distributions but the mean is generally used in practice to describe the classical concept of flood return period.

Key words Flood frequency analysis; design discharges; partial series, peaks over threshold, annual series, flood interval, return period, probability distribution, nonparametric test.

1 INTRODUCTION

Streamgauge records are statistically analyzed to provide crucial information about the probability of floods, in terms of the frequency and magnitude of peak discharges associated with individual events. Inherent within this type of analysis is the flood return period, defined as the average interval of time within which the magnitude of the event will be equalled or exceeded once (Chow 1964). In this paper, the intervening time between successive exceedances of peak discharge over a threshold is referred to as the flood interval.

Flood frequency analysis is based on data pertaining to the independent flood discharge maxima during each year (the annual series) or alternatively the partial series (PS). Various procedures for PS analysis are available; a common three-step approach is described by Kuczera and Franks (2006): (a) A sample of flood peaks is obtained from the streamflow record by selecting peaks that equal or exceed a selected threshold discharge. This sample is the PS, alternatively referred to as the peaks over threshold or partial duration series. An example PS sample including discharges X_i (i=1 to n) $\geq$ a threshold X_T is shown on Figure 1, (b) A statistical distribution relating the flood discharge peaks to exceedance probability (the probability distribution, or pd) is fitted to the sample data, and (c) The probabilities are transformed into ‘annualised’ return periods, based fundamentally on Equation (1).

$$T_A(x_T) = \frac{1}{1-\mathrm{F}_A(X_T)} \approx \frac{1}{\lambda(1-\mathrm{F}_T(X_T))} \qquad (1)$$

where $T_A\ (X_T)$ is the return period (in years) of flood threshold X_T based on being equaled or exceeded in any randomly selected year (i.e. based on annual probability exceedence) , $\mathrm{F}_A\ (X_T)$ is the cumulative frequency of X_T being equaled or exceeded in any randomly selected year, λ is the average number of PS floods per time interval, taken to be a year, $\mathrm{F}_T\ (X_T)$ is the cumulative frequency of X_T being equaled or exceeded in the population of PS floods. In all cases, suffix T denotes the partial series and A denotes the annual series equivalent.

The return period is expressed as an equivalent annual probability and the equation (1) relationship to the PS frequency is accurate for flood magnitudes with return periods longer than 5 years (Bell *et al*. 1989). A more robust relationship should use the Langbein-Takeuchi relationship (Takeuchi 1984) in defining $\mathrm{F}_A\ (X_T)$ in equation (1):

$$\mathrm{F}_A(X_T) = \exp[-\lambda\{1 - \mathrm{F}_T(X_T)\}] \qquad (2)$$

As noted by Bell *et al*. (1989), the equation 1 relation enables the conversion of PS frequency to annual estimates that suit most purposes, and as this is often the case, the cumulative frequencies F_A and F_T are confused in the literature or not explicitly defined. As this paper considers return periods shorter than 1 year, the preference is to express flood frequency in terms of the PS return period, denoted herein as $T_T\ (X_T)$. As indicated in Figure 1, a PS sample of floods exceeding X_T has a corresponding set of intervals T_i (i=1 to n-1). The PS return period is conventionally estimated as an average of the intervals. (In Australia, interval is referred to as "recurrence interval" and return period is referred to as the "average recurrence interval"). Based on this interpretation, the following equation is also considered to apply:

$$T_T(X_T) \approx \bar{T} = \frac{\sum_i^{n-1} T_i}{n-1} \qquad (3)$$

Much research has been conducted on aspects relating to all three steps of the outlined PS methodology based on flood peaks. Various pds have been fitted to PS discharges including exponential (Cunnane 1973, Madsen *et al*. 1997) and log Pearson 3 (Jayasuriya and Mein 1985). The generalized Pareto (GP) pd is a popular choice with its merits demonstrated in the late 1920s (Fisher and Tippet 1928 cited in Ribatet *et al*. 2007). Stedinger *et al.* (1993) and Rao and Hamed (2000) provide discussions and reviews of the application of various pd types to flood frequency analysis.

The conventional three-step method is based on fitting a probability density function (pdf) to the PS peak discharges (X sample). By comparison, alternative approaches such as directly using interval statistics to determine return periods has received less attention, i.e. utilize the T sample shown on Figure 1. One such approach is to assume that the count of exceeding floods in a given time interval follows a Poisson pd (Kirby 1969) and, if this is the case, it can be demonstrated the resulting T values are exponentially distributed (Shane and Lynn 1964). Early work by Thom (1959) on rainfall intervals questioned the use of a mean recurrence interval based on simple averaging (number of years divided by the number of events) to indicate design risk and proposed an alternative approach based on the Poisson pd. Thom (1959) also recognized that the rainfall interval can be considered a random variable and has a probability distribution.

Ben-Zvi and Azmon (2010) offer a direct approach based on fitting a gamma pd to the T sample rather than the PS flood discharge peaks. The approach was tested on data from nine streamgauges located in Israel. Estimates of discharge for various return periods were based on the statistics of the *T* sample using slope methods described by Brauner (1997). It was found that the exponential pd gave a poor representation of the observed *T* distribution. Previous analysis of the same streamgauges (Ben-Zvi 1999) showed that the gamma pd fitted the *T* samples of the high peaks better than the exponential pd.

Other researchers who have adopted a more direct approach using *T* samples include Keylock (2005) who adapted Tsallis statistics to describe extreme flood return periods for the Po River, Italy. This analysis built on earlier work by Mazzarella and Rapetti (2004) who classified historical flood descriptions observed since 1780. Interestingly, Kishore *et al.* (2011) used random walk analysis to derive a theoretical interval frequency relationship of extreme events within complex networks such as transport and flood drainage systems and this expression had a similar form as described by Keylock (2005).

This broad approach of directly working with *T* samples extracted from observed streamflow data as an alternative to the more classical statistical analysis of *X* samples is taken up by this paper. The premise is simple; ensembles of *T* samples corresponding to various thresholds X_T were extracted from observed streamflows and their probability distributions were investigated. Relatively long streamflow records (approximately 100-year length) available at several sites in Queensland, Australia were used in the study.

The specific research questions considered in the study were twofold. Firstly "Do the *T* samples resulting from different flood thresholds follow a consistent pd type?" In other words, can the flood interval distribution corresponding to relatively frequent discharges inform the distribution for larger flood magnitudes with sparse observed data? As precedence, Ashkar and Rouselle (1983) suggest that if the Poissonian flood occurrence is found applicable at a certain threshold discharge, then it should also apply at higher threshold discharges.

Evidence of a consistent pd to explain flood interval frequency (using *T* sample) offers the potential to provide an alternative to more conventional methods of analysis based on flood discharges (using *X* sample). The exponential pd was assessed as a candidate for a consistent pd. Data that fits an exponential pd exhibits a standard deviation equal to the mean and a skewness of 2. These statistical properties were used in the evaluation of the exponential pd. Testing of statistical differences between flood interval samples were also carried out, together with goodness-of-fit analyses against numerous pd candidates.

The second research question was "Does the use of the median significantly influence the return period estimate?" The return period is generally expressed as the numerical average of the intervals, consistent with equation 3. However as the pds relevant to the *T* samples may be highly skewed, the median may be a superior statistic to denote the central tendency of observed flood intervals. This aspect was explored by preparing flood frequency charts at selected streamgauges and plotting quantiles based on the median and the mean for direct comparison.

2 DATA AND METHODS

2.1 Streamflow data

The Queensland Department of Natural Resources and Mines operates a large number of streamgauge sites throughout Queensland. A total of 11 streamgauge sites (Table 1) were identified that had approximately 100-year periods of record and also had limited periods of missing or poor quality data. The selected streamgauges represent a range of Köppen climate zones (Stern and de Hoedt 2000) from tropical to temperate. The catchments can be broadly divided into two groups; mostly 'coastal' that drain eastward and some 'inland' that drain towards the west and as such are defined by their location relative to the Great Dividing Range. Catchment area varies from <100 to >130,000 square kilometers in size. The locations of the streamgauges are mapped on Figure 2.

Table 1 and Figure 2 here

2.2 Methodology used in T sample analysis

Daily flow volumes and discharge maxima were obtained at each streamguage for the full record period available to March 2013. Independent flood events were identified based on using the long-term average daily volume as a flow marker to delineate floods (Brodie 2013). A flood was defined as a period within the record starting when daily flows rose above the average volume and ceasing when flows fell below the average. Daily discharge peaks and their time of occurrence within each 'above average' period were identified to form a series of candidate flood peaks for PS analysis. The mean number of these 'above average' floods per year ranged from 3.5 (130003 Fitzroy River) to 8.1 (143303 Stanley River) with the mean across the streamgauges of 5.1 floods/year.

The flood occurrence times based on daily peaks were utilized to derive intervals between floods that equal or exceed a given threshold. As a further check to limit the inclusion of dependent floods, the minimum interval was set at 3 days, as recommended by Potter and Pilgrim (1971) based on hydrological analysis of several rural catchments located in eastern New South Wales, Australia. This check was the most satisfactory of five sets of criteria that were tested.

A consistent range of thresholds were adopted in the analysis to make direct comparisons between streamgauges. With reference to equation 4, N based on the available records varied from 80 to 103 years (93.7 ± 7.1) and this was considered to be consistent enough to adopt a standard set of n values throughout the analysis. Selected n values were 15, 30, 60, 90, 120, 150, 180, 210, 240, 270 and 300. This range corresponds to return periods, on average, that varies from approximately 3.7 months to 6.3 years. In this study, the threshold discharge was iteratively adjusted until the number of floods in the PS sample matched the selected n value.

$\lambda = \frac{n}{N}$ (4)

where λ is the average number of PS floods per year, n is the number of floods in the PS sample and N is the period of streamflow record (years).

It was assumed that no floods occurred during periods of missing data. This assumption is expected to become more valid as flood threshold increases (n reduces). Edge intervals (from the start of streamgauge data to the occurrence of the first flood and from the last flood to the end of data) were excluded from the flood interval series. There is loss of information associated with edge intervals and this effect becomes more pronounced as n declines.

To facilitate the comparison between streamgauges and thresholds, the T values were reported as a ratio to the mean, i.e. $T_i/\bar{T}$, with $\bar{T}$ computed using equation (3). The unit-free ratio is referred to in this paper as the 'flood interval ratio'.

The data on the ratio enables us to check if the distribution of the flood interval differs significantly at the different streamgauges. An appropriate statistical test to test the equality of the locations or centres of the distributions at different sites should be applied to determine any significant differences. A one-way analysis of variance to test equality of the means would be appropriate if the distribution of the ratio is symmetrical. Otherwise a non-parametric equivalent test should be used to test the equality of medians. As indicated later in section 3.2, the non-parametric Kruskal-Wallis test was adopted.

The $T/\bar{T}$ samples were then analyzed using EasyFit Professional (version 5.5, http://www.mathwave.com/) to identify suitable fitting pdfs. The Anderson Darling goodness of fit statistic was adopted to aid in pdf selection from a range of 65 different types of pdf. The Anderson Darling test is considered superior to alternatives such as the Kolmogorov-Smirnov and chi-square tests for pd evaluation (Ahmad *et al.* 1988). A range of pd fitting methods are used depending on the distribution. With reference to the pds found to be most relevant (as evident later in Table 2), the methods included the Method of Moments (exponential, gamma, log Pearson 3 pdfs), the Maximum Likelihood Method (Dagum, fatigue life and Pareto pdfs) and Method of L-Moments (generalized extreme value, Wakeby and generalized Pareto pdfs).

The inclusion of the fatigue life pd is noteworthy as it is rarely used in flood hydrology. Also known as the Birnbaum-Saunders distribution, this pd was developed to model structural failure due to cracking (Birnbaum and Saunders 1969). The fatigue life pd has a unimodal cumulative density function (cdf) as follows:

$$\mathrm{F}(x) = \Phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{x-\gamma}{\beta}} - \sqrt{\frac{\beta}{x-\gamma}}\right)\right) \quad (5)$$

where α is the shape parameter, γ is the location parameter, β is the scale parameter and Φ is the cdf of the standard normal distribution.

Flood frequency charts were prepared for all streamgauges in order to assess the influence of using the median or mean as the adopted measure of central tendency relating to flood return period. For comparison, the PS discharges were ranked and graphed against a plotting position estimate of return period based on the Cunnane formula (Cunnane 1978):

$$T_T(X_T) \approx T_{PP} = \frac{N+0.2}{m-0.4} \qquad (6)$$

where T_{PP} is the return period plotting position of the rank m flood and N is the number of years of record.

To summarize, the adopted methodology involved the following computational steps for each streamgauge:

(1) Daily peak and mean volumetric discharges were obtained for the streamgauge and periods of missing data were identified (refer to Table 1). The daily peaks were used in Step (3) and the daily mean volumetric discharges employed in Step (2).
(2) The average daily mean volumetric discharge over the full streamflow record was computed and this was utilised to identify sequences of flood periods in accordance to the approach described by Brodie (2013). Periods with daily flows less than the average or missing were included as non-flood periods. The historical streamgauge record was defined as a chronological sequence of flood and non-flood periods.
(3) The peak flood discharge and date of peak for each flood sequence was identified from the daily peak data.
(4) A target n was adopted. A trial discharge was then selected. Individual intervals were computed based on the intervening dates when the flood peak discharge equaled or exceeded the trial discharge. This resulted in a series of flood intervals (expressed in days between peak discharges). Checks were made to ensure the minimum interval exceeded 3 days.
(5) The flood intervals were counted and step (4) was repeated with an adjusted trial discharge until the flood interval count matched the target n
(6) Steps (4) and (5) were repeated for other n targets to cover the full range of analysis (n = 15 to 300)
(7) Each flood interval series were converted to a flood interval ratio series based on calculating $T_i/\bar{T}$ values
(8) Sample statistics of the flood interval ratio series (mean, median, standard deviation and skewness) were estimated and plotted as boxplots. Kruskal-Wallis testing was conducted to evaluate the statistical differences between streamgauges.
(9) Each flood interval ratio series was then individually analyzed using EasyFit Professional to identify and rank suitable fitting pdfs.
(10) Flood frequency charts were finally prepared for each streamgauge to compare peak discharge estimates based on plotting positions in addition to mean and median flood interval ratios.

3 RESULTS AND DISCUSSION

3.1 Sample statistics of flood interval ratio

Median, standard deviation and skewness statistics of $T/\bar{T}$ pooling all streamgauges are presented as boxplots in Figure 3. Medians are generally significantly less than the means, with the low average flood interval ratios evident for the very frequent floods (≈ 0.4 to 0.5 for n = 210 to 300) increasing to approximately 80% of the mean at n = 90 to 150. The average median drops for larger flood thresholds (n = 15 to 60). A value of n = 90 approximates the 1-year return period.

The average standard deviations of $T/\bar{T}$ falls in the 1.0 to 1.2 range and tends to decrease as n becomes smaller (increasing return period). Average skewness is generally in the 1.0 to 1.5 range, and a large variance of skewness is present. These general observations indicate that the exponential pd (Standard deviation = 1, Skewness = 2) would have limited application in describing interval frequencies, consistent with the findings of Ben-Zvi and Azmon (2010).

Figure 3 here

Box plots of $T/\bar{T}$ statistics for individual streamgauges are compiled in Figure 4. Based on a visual assessment, the streamgauges can be placed into two groups: 1) Group I - Streamgauges that exhibit medians that have a distinctive peak close to the mean at $n \approx$ 120 to 150 then tend to reduce or oscillate at lower n (Gauges 110001, 110002, 110003, 116001 and 130003), and 2) Group II - streamgauges that tend to have a more linear trend with reducing n (the remaining gauges). Geographically, the first grouping coincides with streamgauges situated in North Queensland plus a single representation within Central Queensland (130003 Fitzroy River). The second grouping is clustered within South East Queensland. At some gauges, the differentiation between groups is not clear cut; for example, Group II Gauge 143303 Stanley River exhibits an unpronounced peak in the $n \approx$ 120 range.

Figure 4 here

3.2 Test for statistical differences between flood interval ratio samples

Plots of $T/\bar{T}$ data (not provided) showed that the distribution is not normal regardless of the size of n selected and the streamgauge in consideration. The histograms for various values of n clearly showed that the distribution of the ratio is highly skewed to the right. Thus we pursue the nonparametric method to conduct a test on the equality of medians of the flood interval ratio at different locations for various values of n.

As an example we performed the independent samples Kruskal-Wallis test for n = 15. It was found that the test is not significant at all (with p-value = 0.985). Thus we conclude that the median $T/\bar{T}$ ratio at different streamgauges are not significantly different. In other words, the geographical location of the river site has no impact on the median flood interval ratio.

The same nonparametric test was performed on the median flood interval ratios at different streamgauges for n = 300. Once again, the test provided no significant evidence against the null hypothesis of equality of median ratios at different river sites (with p-value = 0.294). The tests on the $T/\bar{T}$ ratio for other selected values of n (in between 15 and 300) lead to the same conclusion.

3.3 Fitting pds to flood interval ratio samples

The outcomes of the pd fitting using EasyFit Professional and the Anderson Darling statistic are presented in Table 2. Group I streamgauges occupy the first five rows of the table.

Some evidence of differences between Group I and II can be observed. The $T/\bar{T}$ histograms were visually inspected and many were found to be multimodal. Multimodal histograms are highlighted **bold** in Table 2. Group I streamgauges have a significantly higher incidence of attempting to fit pds to multimodal histograms, especially for n values greater than 90, corresponding to sub-year return periods (< 1 year). An example to illustrate fitting to a multimodal histogram is given in Figure 5. Partial series histograms, albeit based on discharge peaks, can exhibit multimodal shapes depending on whether different mechanisms of flood generation, such as heavy rainfall and snowmelt, are present (Adamowski *et al.* 1998). The mechanisms that apply to the predominantly tropical climate of Group I gauges is less clear, but may involve the absence or presence of cyclonic weather patterns. In contrast, the Group II histograms are almost consistently unimodal, and an example of which is also given in Figure 5. The pd fitting against the multimodal histograms is relatively poor compared to those against the unimodal histograms.

The fatigue life pd consistently dominated the Group II $n > 150$ scenarios (approx. 7.5 month return period) and also featured in the Group I $n > 240$ scenarios (approx. 4.7 month return period). At low n, the log Pearson 3 pd is preferred for 6 out of 11 gauges at $n = 15$. At intermediate $n \approx 90$ - 180, phased bi-Weibull and Dagum pds are common, but noting that for Group I gauges these are often unimodal attempts to fit multimodal histograms. For the transition from $n = 90$ towards $n = 15$, Pareto type pds and the Wakeby pd are represented within Table 2.

Table 2 here

Figure 5 here

Ben-Zvi and Azmon (2010) used a gamma pd fitting to flood intervals for 5-year return period discharges and found a good fit for 6 of the 9 Israeli arid-zone streamflow gauges, with a fair fit for the remainder. A 5-year return period corresponds to $n \approx 15$ to 30 for the Queensland stations under analysis. A log Pearson 3 or Pareto type pd yielded better fitting than the gamma pd (Table 2). However, the gamma pd does provide a superior fit at some streamgauges for some other n values. Its usefulness to represent interval probabilities cannot be discounted.

Fitting pds allows mode estimation to be performed and this data can be presented across all streamgauges as a boxplot (Figure 6). For some pds, the mode is unable to be implicitly computed. However, there is a trend of consistently low modes (< 0.1) of the flood interval ratio for $n > 120$. This corresponds to return periods shorter than 1-year. As n decreases, the scatter in the mode estimates between gauges increases significantly and the median mode approaches ≈ 0.2 for $n \leq 30$.

Figure 6 here

The pd fitting shows no evidence of the exponential pd as being preferred for high-*n* scenarios, consistent with the preliminary assessment based on Figure 3 boxplot statistics. This pd is preferred on a small number of scenarios (Gauge 143303 Stanley River, *n* = 150 - 180). The fatigue life pd dominates instead.

The identification of the fatigue life pd is an interesting outcome as its original use was in structures to describe the elapsed time until the development and growth of a dominant crack occurs that passes a threshold and causes failure (Desmond 1985). In this paper, we have identified a hydrological analogy in the form of the elapsed times between flood events that also pass a type of threshold. The fatigue life has been used in diverse applications including biological, medical and environmental studies such as air contamination (Sanhueza *et al.* 2008, Marchant *et al.* 2013) and also financial and sharemarket analyses (Ahmed *et al.* 2010) These examples demonstrate the versatility of the fatigue life pd and hence potential application to flood intervals.

A more detailed investigation of the fatigue life pd was undertaken using data at gauging station 136202 (Barambah Creek). As evident in Table 2, flood intervals at this location demonstrated the highest preference towards the fatigue life pd (9 out of the 11 $T/\bar{T}$ series). Fatigue life pds fitted to all $T/\bar{T}$ series at this streamgauge resulted in a narrow range of pdf parameters ($1.17 < \alpha < 1.53$, $0.45 < \beta < 0.63$, $-0.06 < \gamma < 0.02$). The near-zero estimate for the location parameter γ indicates that the fatigue life pdf could be further simplified to a 2-parameter version (by setting $\gamma = 0$) with little loss of accuracy. The fitted pds are presented as Figure 7 and show that the distributions have similar tails but slightly different modes – the mode peaks for the larger flood magnitudes (*n* = 15 to 30, shown as black solid curves) are less than the more frequent floods (*n* = 60 to 300, shown as grey solid curves). The exponential pdf is also presented on Figure 7 (as a black dashed curve) and shows a significant shape difference compared to the fatigue life pds.

Figure 7 here

The specific research question posed in the paper is to identify whether a consistent pd type can be applied across the range of adopted flood thresholds. From the fitting analysis, the strongest candidate is the fatigue life pd especially for the predominately uni-modal distributions found with Group II. Flood occurrence based on Poissonian counting would suggest the exponential pd would be appropriate, but this is not supported by the fitting results.

It is notable that fatigue life pd used in the fitting is a 3-parameter pdf which potentially can be reduced to two parameters (as γ is close to zero). It is expected that more sophisticated multi-parameter pds (such as the 4-parameter generalized gamma) would yield better fits to the observed data, but this only occurred on occasions for the longer return periods (corresponding to $n \leq 120$). This effect may be simply due to increased uncertainty in identifying the 'best' pd as the sample size, represented by *n*, reduces.

This outcome is further illustrated in Table 3 which shows the Anderson Darling statistic for both the 'best' and second 'best' fitted pds for the Group II streamgauges. This table was compiled to identify whether the top-ranked pd identified by the fitting analysis is clearly separated from the next-best pd. Cases when a relatively small difference in the statistic, arbitrarily set at < 10%, are highlighted **bold** in the table. At several streamgauges, the performance of the top-fitted pd is only marginally better than the second ranked pd for sample sizes less than 90 (coinciding to return periods exceeding 1-year). At sub-year return periods, although there are some exceptions, the top-fitted pd is generally more clearly identifiable as the 'best' pd. In many cases that involved fatigue life as the top fitted pd, the log normal pd was the next 'best' distribution.

Insert Table 3 here

However, as noted by Meylan *et al.* (2012), goodness-of-fit tests such as the Anderson Darling test only makes it possible to reject or accept the null hypothesis that an individual pd provides an explanatory fit to the observed data, and is unreliable in choosing the 'best' pd amongst a range of different pd types. A diversity of pds were identified in Table 2 across the streamgauges for n = 15, but this may not be the case given that the Kruskal-Wallis tests (section 3.2) found that the median $T/\bar{T}$ ratios between streamgauges are not significantly different. This raises the possibility that a common pd, such as the fatigue life, could be applicable across the range of n values but is unable to be confirmed by the statistical tests.

3.4 Central tendency and T plotting position

Figure 4 plots highlight the wide range of T values present within each PS sample, varying by at least two orders of magnitude. Coupled with the tendency for low-mode pds, this raises the issue of whether the return period based on the arithmetic mean of the intervals is a suitable measure of central tendency. This aspect was explored further by preparing flood frequency charts at each streamgauge. Two of these plots are presented as examples in Figure 8.

Figure 8 here

Mean intervals based on the arithmetic means (equation 3) of each T series extracted for both streamgauges are also plotted and closely overlay the Cunnane plotting positions. Numerically, these two return period estimates are similar. Even at the lowest n equal to 15, the computed difference between the Cunnane and the mean estimate is less than 3%. Of interest is the median interval of each T sample, also plotted on the Figure 8 charts. Depending on the streamgauge and the flood magnitude, the flood interval is substantially different depending whether the median or the mean is adopted.

These two sites were selected for presentation with reference to individual box plots in Figure 4. They represent two extremes: 1) Gauge 110001 where there is a significant departure of the median from the mean at low n, and 2) Gauge 136202 where the mean and the median merge at low n. The flood frequency charts for the other gauges exhibited median curves relative to the mean that fall in between these extremes. The displacement of the median from the mean as plotted in Figure 4 (the gap between the closed and open circles) provides a direct guide on how the respective curves will plot on the flood frequency chart.

Apart from the major trends that characterize the Group I and Group II gauges, each streamgauge has subtle differences at low n that make it difficult to provide firm conclusions. An example is the patterns exhibited by 110001 and 110002 which are two locations on the same river (Barron River): At 110001, the median departs from the mean at low *n*, whereas the opposite trend is present at 110002.

4 CONCLUSIONS

Streamflows at Queensland gauges with approximately 100-year datasets were analyzed to determine flood interval pd characteristics. Average daily flow volume was used to identify independent floods and *T* series were determined at each of the eleven streamgauges for a range of discharge thresholds. Based on the average time period of available records ($\overline{\mathrm{N}}$ = 93.7 years), the selected range covers an average return period from 3.7 months to 6.3 years.

Several findings can be drawn from the statistical analysis and pd fitting:

(1) Limited evidence was found that the exponential pd is an appropriate distribution to use to describe the probabilities of flood intervals exceeding any of the adopted thresholds. This is consistent with the findings of Ben-Zvi and Azmon (2010) for arid catchments in Israel, even though our study encompassed temperate, subtropical and tropical climates in Australia.

(2) The streamgauges could be broadly classified into two geographical groupings: Group I situated in tropical North Queensland and a sole gauge representative in Central Queensland, and Group II situated in the temperate and subtropical South East Queensland.

(3) Group I gauges tended to exhibit multimodal interval histograms for sub-year return periods with resulting poor fitting of unimodal pds. It is postulated that the multimodal characteristic may be associated with the incidence, or not, of cyclonic activity in the region during individual wet seasons.

(4) Group II gauges had interval histograms that were almost consistently unimodal and this was conducive to satisfactory pd fitting.

(5) The statistical plots show the distribution of the interval ratios is skewed to the right and the nonparametric test on the equality of medians is not significant.

(6) Fitting based on the Anderson Darling statistic strongly indicated the fatigue life pd being appropriate to model Group II sub-year intervals of the order of less than 7.5 month return period. There is indication of a transition towards the log Pearson 3 pd as the threshold discharge increases towards 6.3 year return period, but this was not consistent for all streamgauges.

(7) Due to the similarities in their definitions, the return periods computed from the $T/\bar{T}$ series (as arithmetic mean) were consistent with Cunnane plotting position estimates for the range of event exceedances under analysis. Given the low-mode characteristic of the fitted pds, the median interval may be a better measure of central tendency. The median was often significantly less than the mean, particularly for the sub-year return periods. This observation has implications on how the return period of these relatively minor floods are best defined and statistically described.

ACKNOWLEDGEMENTS

The authors thank three anonymous reviewers and Associate Editor Dr Ioannis Nalbantis for their constructive comments.

REFERENCES

Adamowski, K., Liang, G-C. and Patry, G.G., 1998. Annual maxima and partial duration flood series analysis by parametric and non-parametric methods. *Hydrological Processes*, 12, 1685-1699.

Ahmad, M.I., Sinclair, C.D. and Spurr, B.D., 1988. Assessment of flood frequency models using empirical distribution function statistics. *Water Resources Research*, 24 (8), 1323-1328.

Ahmed, S. E., Castro-Kuriss, C., Flores, E., Leiva, V. and Sanhueza, A., 2010. A truncated version of the Birnbaum-Saunders distribution with an application in financial risk. *Pakistan Journal of Statistics*, 26 (1), 293-311.

Ashkar, F.and Rouselle, J. 1983. Some remarks on the truncation used in partial flood series models. *Water Resources Research*, 19 (2), 47-480.

Bell, F.C., Dolman, G.S. and Khu, J.F., 1989. Frequency estimation of natural hazards and extremes. *Australian Geographical Studies*, 27 (1), 67-86.

Birnbaum, Z.W. and Saunders, S.C. 1969. A new family of life distributions. *Journal of Applied Probability*, 6 (2), 319-327.

Ben-Zvi, A., 1999. Distribution of intervals between flood events. *In*: V.P Singh, I.W. Seo and J.H. Sonu eds. *Hydrologic Modeling*. Water Resources Publications, 255-261.

Ben-Zvi, A. and Azmon, B., 2010. Direct relationships of discharges to the mean and standard deviation of the intervals between their exceedences. *Hydrological Sciences Journal*. 55 (4), 565-577.

Brauner, J.S., 1997. Environmental sampling of slope: Sen’s method in environmental pollution. *In*: D. Gallagher ed. *Environmental Sampling and Monitoring Primer*, Available from:
http://www.webapps.cee.vt.edu/ewr/environmental/teach/smprimer/sen/sen.html

Brodie, I., 2013. Using volume delivery time to identify independent partial series events. *Water Resources Management*, 27, 3727-3738.

Chow, V.T., 1964. Section 8, Statistical and probability analysis of hydrologic data, Part 1, Frequency Analysis. *In*: V.T. Chow ed. *Handbook of Applied Hydrology*, New York, McGraw-Hill.

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

Cunnane, C., 1978. Unbiased plotting positions - a review. *Journal of Hydrology*, 37, 205-222.

Desmond, A., 1985. Stochastic models of failure in random environments. *Canadian Journal of Statistics*, 13, 171-183.

Fisher, R.A. and Tippett, L.H., 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Cam. Phil. Soc*. 24, 180-190.

Jayasuriya, M.D.A. and Mein, R.G., 1985. Frequency analysis using the partial series. The Institute of Engineers Australia Hydrology and Water Resources Symposium, Sydney, 14-16 May, National Conf. Publ. No. 85/2, 81-85.

Keylock, C.J., 2005. Describing the recurrence interval of extreme floods using nonextensive thermodynamics and Tsallis statistics. *Advances in Water Resources*, 28, 773-778.

Kirby, W., 1969. On the random occurrence of major floods. *Water Resources Research*, 5, 778-784.

Kishore, V., Santhanam, M.S. and Amritkar, R.R., 2011. Extreme events on complex networks. *Physical Review Letters*, 106 (18), 1-5.

Kuczera, G. and Franks, S., 2006. *Australian Rainfall and Runoff Book IV Estimation of Peak Discharge, Chapter 2 At-site Flood Frequency Analysis. Revision Draft* [online]. Engineers Australia. Available from: ww.arr.org.au.

Madsen H., Rasmussen, P.F. and Rosbjerg, D., 1997. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrological events 1. At-site modeling. *Water Resources Research*, 33 (4), 747-757.

Marchant C., Leiva, V., Cavieres, M.F. and Sanhueza, A., 2013. Air contaminant statistical distributions with application to PM10 in Santiago, Chile. *Review of Environmental Contaminant Toxicology*, 223, 1-31.

Mazzarella, A. and Rapetti, F., 2004. Scale-invariance laws in the recurrence interval of extreme floods: an application to the upper Po river valley (northern Italy). *Journal of Hydrology*, 288 (3-4), 264-271.

Meylan, P., Favre, A-C. and Musy, A., 2012. *Predictive hydrology – A frequency analysis approach*. New York, CRC Press, 103.

Potter D.J. and Pilgrim, D.H., 1971. Flood estimation using a regional flood frequency approach Vol. 2 Report on analysis components. Australian Water Resources Council, Research Project 68/1, *Hydrology of Small Rural Catchments*, Snowy Mountains Engineering Corp.

Ribatet, M., Sanquet, E., Gresillon, J-M. and Ouarda, T.B.M.J., 2007. A regional Bayesian POT model for flood frequency analysis. *Stochastic Environmental Research Risk Assessment*, 21, 327-339.

Rao, A.R. and Hamed, K.H., 2000. *Flood Frequency Analysis*, CRC Press, Boca Raton, Florida.

Sanhueza, A., Leiva, V., and Balakrishnan, N., 2008. The generalized Birnbaum-Saunders distribution and its theory, methodology and application. *Communications in Statistics: Theory and Methods*, 37, 645-670.

Shane, R.M. and Lynn, W.R. 1964. Mathematical model of flood risk evaluation. *J. Hydraulics Div. Am. Soc. Civ. Eng*., 90 (HY6), 1-20.

Stedinger, J.R., Vogel, R.M. and Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events. *In*: Maidment, D. ed. *Handbook of Applied Hydrology*. McGraw Hill, New York.

Stern, H. and de Hoedt, G., 2000. Objective classifications of Australian climates. *Australian Meteorological Magazine*, 49, 87-96.

Takeuchi, K., 1984. Annual maximum series and partial-duration series – Evaluation of Langbein's formula and Chow's discussion, *Journal of Hydrology*, 68 (1-4), 275-284.

Thom, H.C.S., 1959. A time interval distribution for excessive rainfall. *Journal of the Hydraulics Division*, Proc. Amer. Soc. Civil Engineers, 85 (HY7), 83-91.

Tables

Table 1: Selected Queensland streamgauges

Gauge	Location	Started	Missing Data (%)	Catchment Area (km^2)	Climate, Catchment Type
110001	Barron River at Myola	1915	1.2	1940	Tr,C
110002	Barron River at Mareeba	1915	2.4	840	Tr,C
110003	Barron River at Picnic Crossing	1925	< 0.1	220	Tr,C
116001	Herbert River at Ingham	1915	2.6	8805	Tr,C
130003	Fitzroy River at Riverslea	1922	< 0.1	132090	ST,C
130302	Dawson River at Taroom	1911	< 0.1	15846	ST,C
136202	Barambah Creek at Litzowa	1920	< 0.1	640	ST, C
138001	Mary River at Miva	1910	< 0.1	4830	ST, C
143303	Stanley River at Peachester	1927	< 0.1	104	ST, C
422306	Swan Creek at Swanfels	1919	< 0.1	83	Te, I
422310	Condamine River at Warwick	1933	0.5	1360	Te, I

Note: Climate types: Tr = Tropical, ST = Subtropical, Te =Temperate. Catchment types: I=Inland, C= Coastal.

Table 2 Group I and II top fitted pds based on Anderson Darling statistic

Gauge	$n = 300$	$n = 270$	$n = 240$	$n = 210$	$n = 180$	$n = 150$	$n = 120$	$n = 90$	$n = 60$	$n = 30$	$n = 15$
110001	**Fatigue Life**	**Fatigue Life**	**Gen. Gamma**	**Dagum**	**Dagum**	**Phased Bi-Weib**	**Dagum**	Phased Bi-Weib	Gen. Logistic	Gen. Pareto	Log-Pearson 3
110002	**Fatigue Life**	**Dagum**	**Dagum**	**Gen. Gamma**	**Gen. Gamma**	Dagum	Phased Bi-Weib	Dagum	GEV	Log Normal	Gen. Pareto
110003	Fatigue Life	Fatigue Life	**Gen. Gamma**	**Phased Bi-Exp**	**Phased Bi-Weib**	**Phased Bi-Weib**	Dagum	Cauchy	Wakeby	**Gen. Pareto**	Log-Pearson 3
116001	**Fatigue Life**	**Fatigue Life**	**Fatigue Life**	**Phased Bi-Weib**	**Phased Bi-Weib**	**Phased Bi-Weib**	**Laplace**	Cauchy	Wakeby	Wakeby	Error
130003	Fatigue Life	Fatigue Life	**Fatigue Life**	**Fatigue Life**	**Fatigue Life**	**Phased Bi-Weib**	Dagum	Phased Bi-Weib	Wakeby	Burr	Log-Pearson 3
130302	Fatigue Life	Fatigue Life	**Fatigue Life**	Fatigue Life	Fatigue Life	Wakeby	Wakeby	Dagum	Burr	Weibull	Fatigue Life
136202	Fatigue Life	Fatigue Life	Fatigue Life	Fatigue Life	Fatigue Life	Fatigue Life	Fatigue Life	Fatigue Life	Fatigue Life	Log-Pearson 3	Log-Pearson 3
138001	Fatigue Life	Fatigue Life	**Fatigue Life**	Fatigue Life	Fatigue Life	Fatigue Life	Fatigue Life	**Dagum**	Pareto 2	Inv. Gaussian	Log-Pearson 3
143303	Fatigue Life	Fatigue Life	Fatigue Life	Fatigue Life	Expon.	Expon.	Gen. Gamma	Gamma	Pareto 2	Gen. Pareto	Log-Pearson 3
422306	Fatigue Life	Fatigue Life	Fatigue Life	Fatigue Life	Fatigue Life	Wakeby	Wakeby	**Wakeby**	Gamma	Dagum	Frechet
422310	Fatigue Life	Fatigue Life	Inv. Gaussian	Fatigue Life	Fatigue Life	Fatigue Life	Burr	**Fatigue Life**	Pareto 2	Dagum	Wakeby

Expon : exponential, Phased Bi –Exp : phased bi-exponential, Phased Bi-Weib : phased bi-Weibull, GEV : general extreme value, LP3 : log Pearson 3, Inv Gauss : inverse Gaussian, Gen Gamma : generalized gamma. **Bold** pdfs indicate fitting to multimodal histograms exhibiting at least two peaks.

	LEGEND
	Gamma family – gamma, generalized gamma
	Extreme value theory family - generalized extreme value, Frechet, Gumbel Max, Weibull and phased bi-Weibull
	Skewed normal family – inverse Gaussian, log normal, fatigue life and log Pearson 3
	Wakeby
	Logistic family - general logistic and log-logistic
	Pareto family – generalized Pareto, Pareto 2, Burr, Dagum
	Exponential family – exponential, phased bi-exponential, Laplace and error

Table 3 Anderson Darling statistics for Group II top-2 fitted pds

	130302	136202	138001	143303	422306	422310
n = 300	Fatigue Life 2.296 Lognormal 3.079	**Fatigue Life 0.956 Inv.Gauss. 0.992**	Fatigue Life 1.729 Lognormal 2.421	Fatigue Life 0.972 Lognormal 1.306	**Fatigue Life 0.659 Lognormal 0.689**	Fatigue Life 0.613 Inv.Gauss. 0.793
n = 270	Fatigue Life 2.172 Lognormal 3.102	Fatigue Life 0.757 Lognormal 1.316	Fatigue Life 1.465 Lognormal 2.180	Fatigue Life 0.998 Lognormal 1.504	Fatigue Life 0.612 Gen.Gamma 1.535	Fatigue Life 0.644 Inv.Gauss.0.734
n = 240	Fatigue Life 1.933 Lognormal 2.926	Fatigue Life 0.526 Lognormal 1.010	Fatigue Life 1.130 Lognormal 1.825	Fatigue Life 1.488 Lognormal 2.047	Fatigue Life 0.680 Lognormal 1.515	**Inv.Gauss. 0.845 Fatigue Life 0.876**
n = 210	Fatigue Life 4.004 Wakeby 4.628	Fatigue Life 0.709 Lognormal 1.408	Fatigue Life 1.383 Gen.Gamma 2.033	Fatigue Life 1.701 Weibull 2.034	Fatigue Life 0.682 Gen.Gamma 1.107	Fatigue Life 0.585 Inv. Gauss. 0.966
n = 180	Fatigue Life 1.941 Gamma 2.722	Fatigue Life 0.684 Lognormal 1.386	Fatigue Life 0.797 Lognormal 1.469	**Expon. 2.026 Fatigue Life 2.049**	Fatigue Life 0.952 Gen.Gamma 1.624	Fatigue Life 0.582 Inv.Gauss. 1.025
n = 150	Wakeby 1.196 Weibull 1.372	Fatigue Life 0.427 Lognormal 0.923	Fatigue Life 0.713 Gen.Gamma 0.944	**Expon. 1.631 Pareto 2 1.633**	Wakeby 1.098 Gamma 1.335	**Fatigue Life 0.711 Weibull 0.728**
n = 120	Wakeby 0.519 Burr 0.731	Fatigue Life 0.406 Gen.Gamma 0.629	**Fatigue Life 0.737 Weibull 0.751**	Gen.Gamma 0.752 Pareto 2 0.873	Wakeby 0.832 Gen.Gamma 0.917	Burr 0.612 Fatigue Life 0.751
n = 90	**Dagum 0.332 Wakeby 0.358**	Fatigue Life 0.514 Burr 0.735	Dagum 0.702 LP3 0.863	**Gamma 0.624 Pareto 2 0.631**	**Wakeby 0.488 Gamma 0.524**	Fatigue Life 0.294 LP3 0.458
n = 60	**Burr 0.191 Wakeby 0.191**	Fatigue Life 0.316 LP3 0.455	**Pareto 2 0.554 Burr 0.561**	**Pareto 2 0.316 Gen Pareto 0.317**	**Gamma 0.524 Wakeby 0.535**	**Pareto 2 0.336 Burr 0.337**
n = 30	**Weibull 0.382 LP3 0.383**	**LP3 0.177 Wakeby 0.181**	**Inv.Gauss. 0.301 Frechet 0.303**	**Gen Pareto 0.241 Burr 0.243**	**Dagum 0.301 Gamma 0.322**	**Dagum 0.262 Wakeby 0.279**
n = 15	**Fatigue Life 0.299 LP3 0.319**	**LP3 0.261 Dagum 0.271**	LP3 0.195 Ph. Bi-Weib 0.236	**LP3 0.300 Fatigue Life 0.311**	**Frechet 0.252 GEV 0.262**	**Wakeby 0.326 Gen.Pareto 0.326**

Series with fatigue life as the top fitting pd are shaded. Series with < 10% difference in Anderson Darling statistic between the first and second best pd are shown **bold**.

Figure 1. PS timeseries plot showing threshold discharge X_T, PS flood peak discharges $\geq X_T$ (X_i) and the corresponding intervals T_i. In this case, number of PS floods n = 15. The PS has been extracted from N years of streamflow data.

Figure 2 Streamgauge locations

Figure 3 Box plots of flood interval ratio medians, standard deviations and values of skewness against n for all streamgauges pooled. Horizontal lines represent the maximum and minimum. Vertical bars represent the 25th and 75th percentiles. Open circles are means of the medians and the closed circles are medians (50th percentiles) of the medians.

Figure 4 Box plots of flood interval ratio (vertical axis) against n (horizontal axis) for each streamgauge. Refer to Figure 3 for description of plot symbols.

Figure 5 Selected flood interval ratio histograms (a) Gauge 110001, n = 210 showing multimodal histogram, and (b) Gauge 130302, n=210 showing unimodal histogram. x = $T/\bar{T}$

Figure 6 Box plots of flood interval ratio modes against n for all streamgauges pooled. Refer to Figure 3 for description of plot symbols.

Figure 7 Fatigue life pdfs fitted to 136202 Barambah Creek T series. Exponential pdf (n = 180 Exp) is also shown. x = $T/\bar{T}$

Figure 8. Flood frequency charts showing peak discharge (m^3/s) against return period (years) for streamgauges 110001 and 136202. Grey crosses are PS discharges plotted using Cunnane plotting position, open circles are average intervals based on the observed T series and black diamonds are median intervals based on the observed T series.